MATHEMATICS OF COMPUTATION
Volume 81, Number 277, January 2012, Pages 57–79
S 0025-5718(2011)02494-4
Article electronically published on May 23, 2011

MULTIGRID ANALYSIS FOR THE TIME DEPENDENT STOKES PROBLEM

MAXIM A. OLSHANSKII

ABSTRACT. Certain implicit time stepping procedures for the incompressible Stokes or Navier-Stokes equations lead to a singular-perturbed Stokes type problem at each time step. The paper presents a convergence analysis of a geometric multigrid solver for the system of linear algebraic equations resulting from the disretization of the problem using a finite element method. Several smoothing iterative methods are considered: a smoother based on distributive iterations, the Braess-Sarazin and inexact Uzawa smoother. Convergence analysis is based on smoothing and approximation properties in special norms. A robust (independent of time step and mesh parameter) estimate is proved for the two-grid and multigrid W-cycle convergence factors.

1. INTRODUCTION

Let $\Omega \subset \mathbb{R}^d$ with $d = 2$ or $d = 3$, be a bounded polygonal domain. Consider the Stokes type problem given by:

$$
\begin{aligned}
-\Delta \mathbf{u} + \alpha \mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega, \\
\operatorname{div} \mathbf{u} &= g \quad \text{in } \Omega, \\
\mathbf{u} &= 0 \quad \text{on } \partial\Omega.
\end{aligned}
\tag{1.1}
$$

The mean value condition $\int_\Omega g \, \mathrm{d}\mathbf{x} = \int_\Omega p \, \mathrm{d}\mathbf{x} = 0$ should be imposed to make the problem well-posed for all $\alpha \geq 0$. The system (1.1) often appears as the auxiliary one for certain implicit time stepping procedures for the incompressible Stokes or Navier-Stokes equations; see e.g. [31]. The parameter α is typically proportional to the inverse of the time step scaled with viscosity parameter. This results in large values of α making the problem singular-perturbed. On the other hand, for slow flows the value of α can be modest or small. Discretization of (1.1) with finite element method or other conventional methods leads to a system of linear algebraic equations of saddle-point type with symmetric indefinite matrix. Hence one is interested in solvers for such a system which are robust with respect to the variation of α.

Among various existing solvers for discrete saddle-point systems, resulting from discretizations of PDEs, one may distinguish between iterative methods with block preconditioners and direct multigrid methods; see e.g. [3]. This paper deals with direct (coupled) multigrid methods for (1.1). The well-known and efficient multigrid

Received by the editor April 29, 2009 and, in revised form, October 20, 2010.
2010 *Mathematics Subject Classification.* Primary 65N55, 65N30, 65N15, 65F10.
The author was partially supported through the RFBR Grant 11-01-00767 and 09-01-00115.

techniques include the one based on distributive smoothing iterations [26, 33], coupled saddle-point smoothers [6, 37] and block Gauss-Seidel type smoothers (Vanka multigrid) [32]; see also the overview in [34]. While the analysis of robust block preconditioners for the time-dependent Stokes problem can be found in [7, 18, 22, 24], we are not aware of any studies proving the efficiency of multigrid methods for (1.1) in the range of $\alpha \in [0, \infty)$. The analysis of various multigrid methods for the Stokes problem ($\alpha = 0$) can be found in several papers; see [5, 8, 21, 26, 29, 33]. The smoothing analysis from [6, 27, 37] can also be merged with the approximation property from [19, 33] to establish the convergence of the two-grid method for the case of $\alpha = 0$.

The major obstacle for extending existing analyses for the case of $\alpha > 0$ is the lack of an appropriate approximation property. Such an approximation property is established in this paper. To handle the case of $\alpha \geq 0$ we introduce special parameter-dependent norms. Using these norms involves some equivalence results and representations from [18, 22, 24] for the discrete and continuous pressure Schur complement operators. In appropriate norms we consider smoothing properties of distributive iterations and coupled iterations similar to the methods of Braess and Sarazin [6] and Bank et al. [2]. From these results the convergence of the *two-grid* method follows immediately. To establish the *multigrid* convergence we additionally prove the stability of the prolongation operator and smoothing iterations in suitable norms.

The mesh-dependent norms introduced here to prove approximation property (Theorem 5.1) seem to be a natural extension for $\alpha > 0$ of the norms used in [33]. However, to prove some specific norm equivalence results (Lemma 2.2) we need the assumption on pressure finite element space to be a subspace of $H^1(\Omega)$. Not all stable discretizations of (1.1) satisfy this assumption, but many popular discretizations do satisfy, e.g., the family of Taylor-Hood elements or MINI element. All other assumptions which are used in proving approximation and smoothing properties are quite standard and collected in the next section. From approximation and smoothing properties the uniform convergence of the two-grid method follows. No extra assumptions are needed to pass from two-grid to multigrid convergence result.

The remainder of the paper is organized as follows. Section 2 introduces necessary spaces and norms. An important technical result is given by Lemma 2.2. In Section 3 we prove *a priori* estimates and error bounds for the solution of (1.1) and its finite element counterpart. Section 4 provides an algebraic framework for multigrid analysis. Based on results of Section 3, the approximation property is proved in Section 5. In Section 6 we deduce smoothing and stability properties for the distributive, Braess-Sarazin and inexact Uzawa smoothing iterations. Section 7 contains multigrid convergence estimates. Finally, in Section 8 we include results of a few numerical experiments.

2. Preliminaries

Throughout the paper we use the notation $(\cdot,\cdot)$ and $\|\cdot\|$ for the scalar product and norm in $L^2(\Omega)$ and $L^2(\Omega)^d$. Define the following spaces:

$$\mathbf{V} := \{\mathbf{v} \in H^1(\Omega)^d \mid \mathbf{v} = 0 \text{ on } \partial\Omega\},$$

$$\mathbb{Q} := \{q \in L^2(\Omega) \mid \int_\Omega q\, d\mathbf{x} = 0\}.$$

On $\mathbf{V}$ and $\mathbb{Q}$ we introduce the norms:

$$\|\mathbf{v}\|_{\mathbf{V}} := \left(\|\nabla \mathbf{v}\|^2 + \alpha\|\mathbf{v}\|^2\right)^{\frac{1}{2}}, \qquad \|q\|_Q := \sup_{\mathbf{v}\in\mathbf{V}} \frac{(\operatorname{div}\mathbf{v}, q)}{\|\mathbf{v}\|_{\mathbf{V}}}. \tag{2.1}$$

By $\mathbf{V}^*$ we define the dual space to $\mathbf{V}$. Consider the operator $S := \operatorname{div}(\Delta - \alpha I)^{-1}\nabla$, where $-(\Delta - \alpha I)^{-1}$ is the solution operator to the following elliptic problem:

$$\begin{aligned} -\Delta\mathbf{u} + \alpha\mathbf{u} &= \mathbf{f} \quad \text{in } \Omega, \\ \mathbf{u} &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Operator S is self-adjoint positive definite on $\mathbb{Q}$ and

$$(1+\alpha)^{-1}\|q\|^2 \lesssim (S\,q, q) = \|q\|_Q^2 \quad \text{for } q \in \mathbb{Q}. \tag{2.2}$$

Indeed, the identities

$$\begin{aligned} (S\,p, q) &= (\operatorname{div}(\Delta - \alpha I)^{-1}\nabla p, q) = -\langle(\Delta - \alpha I)^{-1}\nabla p, \nabla q\rangle_{\mathbf{V}\times\mathbf{V}^*} \\ &= -\langle \nabla p, (\Delta - \alpha I)^{-1}\nabla q\rangle_{\mathbf{V}^*\times\mathbf{V}} = (p, \operatorname{div}(\Delta - \alpha I)^{-1}\nabla q) = (p, S\,q) \quad \forall\, p, q \in \mathbb{Q} \end{aligned}$$

show that S is self-adjoint on $\mathbb{Q}$; the equality $(S\,q, q) = \|q\|_Q^2$ is shown, e.g., in [18], eq. (3.2); and the lower bound in (2.2) follows with the help of the Friedrichs and the Nečas inequalities: $\|\mathbf{v}\|_{\mathbf{V}} \lesssim (1+\alpha)^{\frac{1}{2}}\|\nabla\mathbf{v}\|$ for $\mathbf{v} \in \mathbf{V}$ and $\|q\| \lesssim \sup_{\mathbf{v}\in\mathbf{V}} \frac{(\operatorname{div}\mathbf{v}, q)}{\|\nabla\mathbf{v}\|}$ for $q \in \mathbb{Q}$.

In order to avoid the repeated use of generic but unspecified constants, here and further by $x \lesssim y$ we mean that there is a constant c such that $x \le c\,y$, and c does not depend of the parameters of which x and y may depend on, e.g., α and mesh size. Obviously, $x \gtrsim y$ is defined as $y \lesssim x$, and $x \simeq y$ when both $x \lesssim y$ and $y \lesssim x$.

By $\|\cdot\|_{Q^*}$ we denote the dual norm to $\|\cdot\|_Q$ with respect to the L^2-duality. On the product space $\mathbf{V}\times\mathbb{Q}$ we define the product norm and the bilinear form:

$$\begin{aligned} |||[\mathbf{v}, q]||| &= \left(\|\mathbf{v}\|_{\mathbf{V}}^2 + \|q\|_Q^2\right)^{\frac{1}{2}}, \\ a(\mathbf{u}, p; \mathbf{v}, q) &= (\nabla\mathbf{u}, \nabla\mathbf{v}) + \alpha(\mathbf{u}, \mathbf{v}) - (p, \operatorname{div}\mathbf{v}) + (q, \operatorname{div}\mathbf{u}). \end{aligned}$$

The weak formulation of the Stokes type problem (1.1) reads: Given $\mathbf{f} \in \mathbf{V}^*$ and $g \in \mathbb{Q}$ find $\mathbf{u} \in \mathbf{V}$ and $p \in \mathbb{Q}$ such that

$$a(\mathbf{u}, p; \mathbf{v}, q) = (\mathbf{f}, \mathbf{v}) + (g, q) \quad \forall\, \mathbf{v} \in \mathbf{V},\ q \in \mathbb{Q}. \tag{2.3}$$

Remark 1. Note that the norms in (2.1) are based on the "velocity part" of the Stokes problem (1.1) and its pressure Schur complement operator S (cf. (2.2)). In this way the α-dependence is taken into account in the norms and the uniform stability and continuity results for the bilinear form $a(\cdot;\cdot)$ easily follow; see Lemma 2.1. For the completeness of presentation we shall give a proof. At the same time, the use of the Schur complement norms in analysis needs some care, especially when both differential and finite element problems are involved. The second lemma in this section (Lemma 2.2) provides useful results for the further usage of such norms.

Lemma 2.1. *Bilinear form $a(\cdot,\cdot;\cdot,\cdot)$ satisfies the uniform stability and continuity estimates:*

$$|||[\mathbf{u}, p]||| \lesssim \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{a(\mathbf{u}, p; \mathbf{v}, q)}{|||[\mathbf{v}, q]|||} \quad \forall\, \{\mathbf{u}, p\} \in \mathbf{V}\times\mathbb{Q}, \tag{2.4}$$

$$a(\mathbf{u}, p; \mathbf{v}, q) \lesssim |||[\mathbf{u}, p]|||\, |||[\mathbf{v}, q]||| \quad \forall\, \{\mathbf{u}, p\},\ \{\mathbf{v}, q\} \in \mathbf{V}\times\mathbb{Q}. \tag{2.5}$$

Proof. Take an arbitrary pair $\{\mathbf{u}, p\} \in \mathbf{V} \times \mathbb{Q}$. Define $\mathbf{w} \in \mathbf{V}$ as a solution to the problem

$$(\nabla \mathbf{w}, \nabla \mathbf{z}) + \alpha(\mathbf{w}, \mathbf{z}) = (p, \operatorname{div} \mathbf{z}) \qquad \forall\, \mathbf{z} \in \mathbf{V}. \tag{2.6}$$

Then it holds that

$$(p, \operatorname{div} \mathbf{w}) = \|\mathbf{w}\|^2_{\mathbf{V}} = \|p\|^2_Q. \tag{2.7}$$

Indeed, the first equality in (2.7) follows by taking $\mathbf{z} = \mathbf{w}$ in (2.6). The second equality follows from dividing both sides of (2.6) by $\|\mathbf{z}\|_{\mathbf{V}}$ and taking the supremum over all $0 \neq \mathbf{z} \in \mathbf{V}$. The definition of $a(\cdot;\cdot)$ and identities (2.7) give

$$\begin{aligned} a(\mathbf{u}, p; \mathbf{u}, p) &= \|\mathbf{u}\|^2_{\mathbf{V}}, \\ a(\mathbf{u}, p; -\mathbf{w}, 0) &= -(\nabla \mathbf{u}, \nabla \mathbf{w}) - \alpha(\mathbf{u}, \mathbf{w}) + \|p\|^2_Q \\ &\geq -\|\mathbf{u}\|_{\mathbf{V}}\|\mathbf{w}\|_{\mathbf{V}} + \|p\|^2_Q \geq \frac{1}{2}(\|p\|^2_Q - \|\mathbf{u}\|^2_{\mathbf{V}}). \end{aligned}$$

If we add these two relations and take $\mathbf{v} = \mathbf{u} - \mathbf{w}$ and $q = p$ we obtain

$$a(\mathbf{u}, p; \mathbf{v}, q) \geq \frac{1}{2}(\|\mathbf{u}\|^2_{\mathbf{V}} + \|p\|^2_Q) = \frac{1}{2}\,|\!|\!|[\mathbf{u}, p]|\!|\!|^2. \tag{2.8}$$

Now the result in (2.4) follows from (2.8) and the following estimate:

$$|\!|\!|[\mathbf{v}, q]|\!|\!|^2 \leq 2\|\mathbf{u}\|^2_{\mathbf{V}} + 2\|\mathbf{w}\|^2_{\mathbf{V}} + \|p\|^2_Q = 2\|\mathbf{u}\|^2_{\mathbf{V}} + 3\|p\|^2_Q \leq 3\,|\!|\!|[\mathbf{u}, p]|\!|\!|^2.$$

With the help of the Cauchy inequality the continuity estimate (2.5) follows directly from the definition of $a(\cdot;\cdot)$ and the definition of norms. □

We will also assume the following *H^2-regularity condition*: The domain Ω is such that the Stokes problem (1.1) with $\alpha = 0$ and $g = 0$ is H^2-regular; i.e., there is a constant c_R such that for any $\mathbf{f} \in L^2(\Omega)^d$ the solution $\{\mathbf{u}, p\}$ is an element of $H^2(\Omega)^d \times H^1(\Omega)$ and satisfies

$$\|\mathbf{u}\|_{H^2(\Omega)} + \|\nabla p\| \leq c_R\|\mathbf{f}\|. \tag{2.9}$$

The condition is satisfied for convex domains [12].

For the discretization of (1.1), we introduce a quasi-uniform family of nested triangulations of Ω (triangles in 2D, tetrahedra in 3D) based on *global regular refinement*. We use conforming finite elements with piecewise polynomial functions. This results in a hierarchy of nested finite element spaces for velocity and pressure:

$$\begin{aligned} \mathbf{V}_0 \subset \mathbf{V}_1 \subset \cdots \subset \mathbf{V}_k \subset \cdots \subset \mathbf{V}, \\ \mathbb{Q}_0 \subset \mathbb{Q}_1 \subset \cdots \subset \mathbb{Q}_k \subset \cdots \subset \mathbb{Q}. \end{aligned}$$

The corresponding mesh size parameter is denoted by h_k and satisfies $h_k/h_0 \simeq 2^{-k}$. We assume the discrete LBB condition to be valid:

$$\sup_{\mathbf{u}_k \in \mathbf{V}_k} \frac{(\operatorname{div} \mathbf{u}_k, p_k)}{\|\nabla \mathbf{u}_k\|} \gtrsim \|p_k\| \qquad \forall\, p_k \in \mathbb{Q}_k. \tag{2.10}$$

We will also refer to the following inequality known as a weak inf-sup condition for the case of $\mathbb{Q}_k \subset H^1(\Omega)$:

$$\sup_{\mathbf{u}_k \in \mathbf{V}_k} \frac{(\operatorname{div} \mathbf{u}_k, p_k)}{\|\mathbf{u}_k\|} \gtrsim \|\nabla p_k\| \qquad \forall\, p_k \in \mathbb{Q}_k. \tag{2.11}$$

The proof of the inequality (2.11) for the Taylor-Hood and isoP2-P1 elements can be found in [4, 24], another example is the MINI element proposed in [1].

Assume the following standard approximation properties of the finite element spaces ($H^0 := L^2(\Omega)^d$):

$$\inf_{\mathbf{v}_k\in\mathbf{V}_k}\|\mathbf{v}-\mathbf{v}_k\|_{H^\ell}\lesssim h_k\|\mathbf{v}\|_{H^{\ell+1}}\quad\text{for } \mathbf{v}\in H^{\ell+1}(\Omega)^d\cap\mathbf{V},\ \ell=0,1, \tag{2.12}$$

$$\inf_{q_k\in\mathbb{Q}_k}\|q-q_k\|\lesssim h_k\|q\|_{H^1}\quad\text{for } q\in H^1(\Omega)\cap\mathbb{Q}, \tag{2.13}$$

and for the case $\mathbb{Q}_k\subset H^1(\Omega)$:

$$\inf_{q_k\in\mathbb{Q}_k}\|q-q_k\|_{H^1}\lesssim h_k\|q\|_{H^2}\quad\text{for } q\in H^2(\Omega)\cap\mathbb{Q}. \tag{2.14}$$

The discrete problem on grid level k is given by: Find $\mathbf{u}_k\in\mathbf{V}_k$, $p_k\in\mathbb{Q}_k$ such that

$$a(\mathbf{u}_k,p_k;\mathbf{v}_k,q_k)=(\mathbf{f},\mathbf{v}_k)+(g,q_k)\quad\forall\ \mathbf{v}_k\in\mathbf{V}_k,\ q_k\in\mathbb{Q}_k. \tag{2.15}$$

Due to (2.10) there exists a unique solution to (2.15).

Besides the product norm $|||[\cdot,\cdot]|||$ defined above we endow every finite element subspace pair $\mathbf{V}_k\times\mathbb{Q}_k$ with the level-dependent product norm:

$$|||[\mathbf{v}_k,q_k]|||_k=\left(\|\mathbf{v}_k\|^2_{\mathbf{V}}+\|q_k\|^2_{Q_k}\right)^{\frac12},\qquad\text{with } \|q\|_{Q_k}:=\sup_{\mathbf{v}_k\in\mathbf{V}_k}\frac{(\operatorname{div}\mathbf{v}_k,p_k)}{\|\mathbf{v}_k\|_{\mathbf{V}}}.$$

Note that the latter relation defines a norm on $\mathbb{Q}_k$ due to the LBB condition (2.10). Again, $\|\cdot\|_{Q_k^*}$ denotes the dual norm to $\|\cdot\|_{Q_k}$ with respect to the L^2-duality. The choice of the norm yields the stability estimate on $\mathbf{V}_k\times\mathbb{Q}_k$ similar to (2.4):

$$|||[\mathbf{u}_k,p_k]|||_k\lesssim\sup_{\mathbf{v}_k,q_k\in\mathbf{V}_k\times\mathbb{Q}_k}\frac{a(\mathbf{u}_k,p_k;\mathbf{v}_k,q_k)}{|||[\mathbf{v}_k,q_k]|||_k}\quad\forall\ \{\mathbf{u}_k,p_k\}\in\mathbf{V}_k\times\mathbb{Q}_k. \tag{2.16}$$

The proof of (2.16) repeats the proof of Lemma 3.1 with $\mathbf{V}$, $\mathbb{Q}$ replaced by $\mathbf{V}_k$, $\mathbb{Q}_k$.

In the following lemma we prove an important technical result.

Lemma 2.2. *Assume $\mathbb{Q}_k\subset H^1(\Omega)$ and (2.11). Then it holds that*

$$\|p_k\|_Q\lesssim\|p_k\|_{Q_k}\lesssim\|p_k\|_Q\qquad\forall\ p_k\in\mathbb{Q}_k \tag{2.17}$$

and

$$\|p_k\|_{Q^*}\lesssim\|p_k\|_{Q_k^*}\lesssim\|p_k\|_{Q^*}\qquad\forall\ p_k\in\mathbb{Q}_k. \tag{2.18}$$

Proof. The upper bound in (2.17) immediately follows from the definition of the norms and the embedding $\mathbf{V}_h\subset\mathbf{V}$.

To prove the lower bound we use the following two inequalities [22, 24]:

$$\|p\|^2_Q\lesssim\inf_{q\in H^1(\Omega)}(\|p-q\|^2+\alpha^{-1}\|\nabla q\|^2)\qquad\forall\ p\in\mathbb{Q}, \tag{2.19}$$

$$\|p_k\|^2_{Q_k}\gtrsim\inf_{q_k\in\mathbb{Q}_k}(\|p_k-q_k\|^2+\alpha^{-1}\sup_{\mathbf{v}_k\in\mathbf{V}_k}\frac{(q_k,\operatorname{div}\mathbf{v}_k)^2}{\|\mathbf{v}_k\|^2})\qquad\forall\ p_k\in\mathbb{Q}_k. \tag{2.20}$$

In particular, (2.19)–(2.20) follows from relations (2.25)–(2.26) in [24] and further application of the upper bound in Theorem 3.2 from [24] (to show (2.19)) and the lower bound in Theorem 4.1 from [24] (to show (2.20)).

Now the lower bound in (2.17) follows from (2.11), (2.19), (2.20), and the embedding $\mathbb{Q}_k\subset H^1(\Omega)$. To prove (2.18) we use the following results [18, 24]:

$$\|p\|^2_{Q^*}\simeq\|p\|^2+\alpha((-\Delta)^{-1}p,p)\ \forall\ p\in\mathbb{Q}, \tag{2.21}$$

$$\|p_k\|^2_{Q_k^*}\simeq\|p_k\|^2+\alpha((-\Delta)_k^{-1}p_k,p_k)\qquad\forall\ p_k\in\mathbb{Q}_k, \tag{2.22}$$

where Δ^{-1} and Δ_k^{-1} are the solution operators for the Poisson-Neumann problem and the finite element Poisson-Neumann problem, respectively. We remark that relation (2.21) follows from Theorems 2.1 and 3.1 in [18] and (2.22) follows from Theorem 4.1 and the analysis of §4.1 in [24]. We note that the existing proofs of (2.20)–(2.22) use the H^2-regularity assumption. For any $p_k \in \mathbb{Q}_k$ it holds that

$$-(\Delta^{-1}p_k, p_k) = \sup_{q\in H^1(\Omega)} \frac{(p_k,q)^2}{\|\nabla q\|^2} \quad \text{and} \quad -(\Delta_k^{-1}p_k, p_k) = \sup_{q_k\in \mathbb{Q}_k} \frac{(p_k,q_k)^2}{\|\nabla q_k\|^2}. \tag{2.23}$$

Therefore, the upper bound in (2.18) immediately follows from (2.21), (2.22), (2.23) and the embedding $\mathbb{Q}_k \subset H^1(\Omega)$.
Furthermore, denote by $P_k q \in \mathbb{Q}_k$ the L^2-projection of $q \in H^1(\Omega)\cap\mathbb{Q}$ on $\mathbb{Q}_k$. Given our assumptions on the triangulation one has $\|\nabla P_k q\| \lesssim \|\nabla q\|$; cf. [9]. Therefore,

$$\begin{aligned} -(\Delta^{-1}p_k, p_k) &= \sup_{q\in H^1(\Omega)} \frac{(p_k,q)^2}{\|\nabla q\|^2} \lesssim \sup_{q\in H^1(\Omega)} \frac{(p_k,P_k q)^2}{\|\nabla P_k q\|^2} \\ &= \sup_{q_k\in \mathbb{Q}_k} \frac{(p_k,q_k)^2}{\|\nabla q_k\|^2} = -(\Delta_k^{-1}p_k, p_k). \end{aligned}$$

This estimate together with (2.21) and (2.22) yields the lower bound in (2.20). □

Remark 2. Inequalities (2.18) do not follow directly from (2.17), since the inverse of the L^2-projection of the operator S on $\mathbb{Q}_k$ is not necessarily equal to the L^2-projection of S^{-1} on $\mathbb{Q}_k$.

3. A priori and error estimates

First we prove two useful a priori estimates for the solution of (1.1).

Lemma 3.1. *Let $\mathbf{f} \in L^2(\Omega)^d$. The following estimate holds for the solution of (1.1):*

$$\alpha\|\mathbf{u}\|^2 + \|\nabla\mathbf{u}\|^2 + \|p\|_Q^2 \lesssim (1+\alpha)^{-1}\|\mathbf{f}\|^2 + \|g\|_{Q^*}^2. \tag{3.1}$$

Furthermore, if the H^2-regularity condition holds and $g = 0$, then $\mathbf{u} \in H^2(\Omega)^d$, $p \in H^1(\Omega)$, and

$$\|\mathbf{u}\|_{H^2} + \|p\|_{H^1} \lesssim \|\mathbf{f}\|. \tag{3.2}$$

Proof. The stability estimate (2.4), identity (2.3), and the Friedrichs inequality $(1+\alpha)\|\mathbf{v}\| \lesssim \|\mathbf{v}\|_{\mathbf{V}}$ on $\mathbf{V}$ imply:

$$\begin{aligned} |||[\mathbf{u},p]||| &\lesssim \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{a(\mathbf{u},p;\mathbf{v},q)}{|||[\mathbf{v},q]|||} = \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{(\mathbf{f},\mathbf{v})+(g,q)}{|||[\mathbf{v},q]|||} \\ &\le \left(\|\mathbf{f}\|_{\mathbf{V}^*}^2 + \|g\|_{Q^*}^2\right)^{\frac12} \lesssim \left((1+\alpha)^{-1}\|\mathbf{f}\|^2 + \|g\|_{Q^*}^2\right)^{\frac12}. \end{aligned}$$

Thus we prove (3.1).

Assume now $g = 0$ and consider $\tilde{\mathbf{f}} = (\mathbf{f} - \alpha\mathbf{u})$, then $\mathbf{u}, p$ solves the Stokes problem

$$\begin{aligned} -\Delta\mathbf{u} + \nabla p = \tilde{\mathbf{f}}, \quad \operatorname{div}\mathbf{u} &= 0 \qquad \text{in } \Omega, \\ \mathbf{u} &= 0 \qquad \text{on } \partial\Omega. \end{aligned}$$

Since $\tilde{\mathbf{f}} \in L^2(\Omega)^d$ and thanks to (3.1) it holds that $\|\tilde{\mathbf{f}}\| \le \|\mathbf{f}\| + \alpha\|\mathbf{u}\| \lesssim \|\mathbf{f}\|$. Now applying the standard regularity result (2.9) for the Stokes problem proves (3.2). □

Further in this section we prove several finite element convergence results for the generalized Stokes problem.

Lemma 3.2. *Assume* $\mathbb{Q}_k \subset H^1(\Omega)$ *and* (2.11). *Let* $\{\mathbf{u}, p\}$ *be the solution to* (2.3) *and* $\{\mathbf{u}_k, p_k\}$ *solve* (2.15), *then it holds that*

$$\|\![\mathbf{u}-\mathbf{u}_k, p-p_k]\|\! \lesssim \inf_{\mathbf{v}_k\in\mathbf{V}_k}\inf_{p_k\in\mathbb{Q}_k} \|\![\mathbf{u}-\mathbf{v}_k, p-q_k]\|\!\, . \tag{3.3}$$

Proof. Let $\mathbf{u}_I$ be the best possible approximation to $\mathbf{u}$ in $\mathbf{V}_k$ with respect to the $\|\cdot\|_{\mathbf{V}}$ norm and let p_I be the best possible approximation to p in $\mathbb{Q}_k$ with respect to the $\|\cdot\|_Q$ norm. The norm equivalence (2.17), stability (2.16), continuity (2.5) estimates and the orthogonality property of finite element error function give:

$$\begin{aligned}
\|\![\mathbf{u}_I-\mathbf{u}_k, p_I-p_k]\|\! &\lesssim \|\![\mathbf{u}_I-\mathbf{u}_k, p_I-p_k]\|\!_k \\
&\lesssim \sup_{\mathbf{v}_k,q_k\in\mathbf{V}_k\times\mathbb{Q}_k} \frac{a(\mathbf{u}_I-\mathbf{u}_k, p_I-p_k;\mathbf{v}_k,q_k)}{\|\![\mathbf{v}_k,q_k]\|\!_k} \\
&\lesssim \sup_{\mathbf{v}_k,q_k\in\mathbf{V}_k\times\mathbb{Q}_k} \frac{a(\mathbf{u}_I-\mathbf{u}_k, p_I-p_k;\mathbf{v}_k,q_k)}{\|\![\mathbf{v}_k,q_k]\|\!} \\
&= \sup_{\mathbf{v}_k,q_k\in\mathbf{V}_k\times\mathbb{Q}_k} \frac{a(\mathbf{u}_I-\mathbf{u}, p_I-p;\mathbf{v}_k,q_k)}{\|\![\mathbf{v}_k,q_k]\|\!} \lesssim \|\![\mathbf{u}_I-\mathbf{u}, p_I-p]\|\!\, .
\end{aligned}$$

With the help of this estimate and the triangle inequality we get

$$\|\![\mathbf{u}-\mathbf{u}_k, p-p_k]\|\! \lesssim \|\![\mathbf{u}_I-\mathbf{u}, p_I-p]\|\! = \inf_{\mathbf{v}_k\in\mathbf{V}_k}\inf_{p_k\in\mathbb{Q}_k} \|\![\mathbf{u}-\mathbf{v}_k, p-q_k]\|\!\, . \qquad \square$$

Taking $\mathbf{v}_k = 0$ and $q_k = 0$ on the right-hand side of (3.3) leads to

$$\|\![\mathbf{u}-\mathbf{u}_k, p-p_k]\|\! \lesssim \|\![\mathbf{u}, p]\|\!\, . \tag{3.4}$$

With the help of a standard duality argument we prove the lemma below.

Lemma 3.3. *Let* $\mathbf{u}, p$ *be the solution to* (2.3) *and* $\mathbf{u}, p$ *solves* (2.15), *then*

$$\|\mathbf{u}-\mathbf{u}_k\| \lesssim \min\{h_k, \alpha^{-\frac{1}{2}}\}\, \|\![\mathbf{u}-\mathbf{u}_k, p-p_k]\|\!\, . \tag{3.5}$$

Proof. Denote $\mathbf{e}_k = \mathbf{u}-\mathbf{u}_k$, $r_k = p - p_k$. Consider $\mathbf{w} \in H^2(\Omega)^d$, $q \in H^1(\Omega)\cap\mathbb{Q}$ solving the Stokes type problem

$$\begin{aligned}
-\Delta\mathbf{w} + \alpha\mathbf{w} - \nabla q = \mathbf{e}_k, \quad \operatorname{div}\mathbf{w} &= 0 \qquad \text{in } \Omega, \\
\mathbf{w} &= 0 \qquad \text{on } \partial\Omega.
\end{aligned}$$

Using the weak form of the problem and the orthogonality property for $\mathbf{e}_k, r_k$, we get

$$\|\mathbf{e}_k\|^2 = a(\mathbf{w}-\mathbf{w}_k, q-q_k; \mathbf{e}_k, r_k)$$

with arbitrary $\mathbf{w}_k \in \mathbf{V}_k$, $q_k \in \mathbb{Q}_k$. Thanks to (2.5), interpolation properties (2.12)–(2.13), and a priori estimates from Lemma 3.1, we get

$$\begin{aligned}
\|\mathbf{e}_k\|^2 &\lesssim \|\![\mathbf{w}-\mathbf{w}_k, q-q_k]\|\!\, \|\![\mathbf{e}_k, r_k]\|\! \lesssim h_k \left(\|\mathbf{w}\|_{H^2}^2 + \alpha\|\nabla\mathbf{w}\|^2 + \|\nabla q\|^2\right)^{\frac{1}{2}} \|\![\mathbf{e}_k, r_k]\|\! \\
&\lesssim h_k\|\mathbf{e}_k\|\, \|\![\mathbf{e}_k, r_k]\|\!
\end{aligned}$$

and

$$\begin{aligned}
\|\mathbf{e}_k\|^2 &\lesssim \|\![\mathbf{w}-\mathbf{w}_k, q-q_k]\|\!\, \|\![\mathbf{e}_k, r_k]\|\! \lesssim \left(\|\nabla\mathbf{w}\|^2 + \alpha\|\mathbf{w}\|^2 + \|q\|_Q^2\right)^{\frac{1}{2}} \|\![\mathbf{e}_k, r_k]\|\! \\
&\lesssim \alpha^{-\frac{1}{2}}\|\mathbf{e}_k\|\, \|\![\mathbf{e}_k, r_k]\|\!\, . \qquad \square
\end{aligned}$$

Now we are in position to prove the main result of this section.

Theorem 3.4. *Let* $\mathbf{f} \in L^2(\Omega)^d$*. Assume* $\mathbb{Q}_k \subset H^1(\Omega)$ *and* (2.11)*. Let* $\{\mathbf{u}, p\}$ *be the solution to* (2.3) *and* $\{\mathbf{u}_k, p_k\}$ *solve* (2.15)*, then the following error estimate holds:*

$$\begin{aligned}(3.6)\qquad &\|\mathbf{u}-\mathbf{u}_k\| + \min\{h_k, \alpha^{-\frac{1}{2}}\}\|p-p_k\|_Q \\ &\quad \lesssim \min\{h_k^2, \alpha^{-1}\}\left(\|\mathbf{f}\| + \max\{h_k^{-1}, \alpha^{\frac{1}{2}}\}\|g\|_{Q^*}\right).\end{aligned}$$

Proof. For arbitrary $\mathbf{v} \in \mathbf{V}$, $q \in \mathbb{Q}$ we denote by $\tilde{\mathbf{v}}_k$ and $\tilde{q}_k$ a unique projection on $\mathbf{V}_k$, $\mathbb{Q}_k$ such that

$$a(\mathbf{v}-\tilde{\mathbf{v}}_k, q-\tilde{q}_k; \mathbf{w}_k, r_k) = 0 \qquad \forall\ \mathbf{w}_k \in \mathbf{V}_k,\ r_k \in \mathbb{Q}_k.$$

Below we consequently use (2.4), orthogonality properties, estimates (3.4) and (3.5) for the differences $\mathbf{v}-\tilde{\mathbf{v}}_k$ and $q-\tilde{q}_k$:

$$\begin{aligned}|||[\mathbf{u}-\mathbf{u}_k, p-p_k]||| &\lesssim \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{a(\mathbf{u}-\mathbf{u}_k, p-p_k; \mathbf{v}, q)}{|||[\mathbf{v},q]|||} \\ &= \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{a(\mathbf{u}-\mathbf{u}_k, p-p_k; \mathbf{v}-\tilde{\mathbf{v}}_k, q-\tilde{q}_k)}{|||[\mathbf{v},q]|||} \\ &= \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{a(\mathbf{u}, p; \mathbf{v}-\tilde{\mathbf{v}}_k, q-\tilde{q}_k)}{|||[\mathbf{v},q]|||} \\ &= \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{(\mathbf{f}, \mathbf{v}-\tilde{\mathbf{v}}_k) + (g, q-\tilde{q}_k)}{|||[\mathbf{v},q]|||} \\ &\le \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{\|\mathbf{f}\|\,\|\mathbf{v}-\tilde{\mathbf{v}}_k\| + \|g\|_{Q^*}\|q-\tilde{q}_k\|_Q}{|||[\mathbf{v},q]|||} \\ &\lesssim \sup_{\mathbf{v},q\in\mathbf{V}\times\mathbb{Q}} \frac{\|\mathbf{f}\|\,\min\{h_k,\alpha^{-\frac{1}{2}}\}\,|||[\mathbf{v},q]||| + \|g\|_{Q^*}\,|||[\mathbf{v},q]|||}{|||[\mathbf{v},q]|||} \\ (3.7)\qquad &= \min\{h_k, \alpha^{-\frac{1}{2}}\}\,\|\mathbf{f}\| + \|g\|_{Q^*}.\end{aligned}$$

We proceed using (3.5) and (3.7):

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_k\| + \min\{h_k, \alpha^{-\frac{1}{2}}\}\|p-p_k\|_Q &\lesssim \min\{h_k, \alpha^{-\frac{1}{2}}\}\,|||[\mathbf{u}-\mathbf{u}_k, p-p_k]||| \\ &\lesssim \min\{h_k^2, \alpha^{-1}\}\left(\|\mathbf{f}\| + \max\{h_k^{-1}, \alpha^{\frac{1}{2}}\}\|g\|_{Q^*}\right).\end{aligned}$$ □

A few remarks are in order.

Remark 3. In the proof of the theorem the extra assumption $\mathbb{Q}_k \subset H^1(\Omega)$ was involved only through the usage of the estimate (3.4). We conjecture, however, that (3.4) still holds for the more general case of LBB stable elements.

Remark 4. For the illustrative purpose we discuss the implication of the analysis of Sections 2 and 3 for the limit cases of $\alpha = 0$ and $\alpha \to \infty$. The case $\alpha = 0$ corresponds to the Stokes problem. Substituting in (1.1) $\mathbf{u} \to \alpha^{-1}\mathbf{u}$, and $g \to \alpha^{-1}g$ we may consider as the limit case $\alpha \to \infty$ the Darcy equations (see also the next remark):

$$(3.8)\qquad \begin{aligned}\mathbf{u} + \nabla p = \mathbf{f}, \quad \operatorname{div}\mathbf{u} &= g \quad \text{in } \Omega, \\ \mathbf{u}\cdot\mathbf{n} &= 0 \quad \text{on } \partial\Omega.\end{aligned}$$

Note that the lack of the second derivatives for $\mathbf{u}$ results in boundary conditions only for the normal component of $\mathbf{u}$. For $\mathbf{f} = 0$ the system (3.8) can be observed as the mixed formulation of the Poisson problem $-\Delta p = g$ with Neumann boundary conditions [10]. Thus a proper setting for the limit problem would be either in $\mathbf{H}_0(\mathrm{div}) \times L_0^2$ or in $\mathbf{L}^2 \times \left(H^1 \cap L_0^2\right)$ spaces; see also the remark in [20] on p. 1608. As we shall show below, our analysis is consistent with the latter choice. To see this we first note the following relations:

$$\alpha^{-\frac{1}{2}}\|\mathbf{v}\|_{\mathbf{V}} \to \|\mathbf{v}\|, \quad \alpha^{\frac{1}{2}}\|q\|_Q \to \|\nabla q\|, \quad \text{for } \alpha \to \infty,\ \forall\ \mathbf{v} \in \mathbf{V},\ q \in H^1 \cap L_0^2.$$

Thus, in the limit case of $\alpha = \infty$ the result of Lemma 2.1 is the infsup stability and continuity of the bilinear form

$$a(\mathbf{u}, p; \mathbf{v}, q) := (\mathbf{u}, \mathbf{v}) + (\mathbf{u}, \nabla q) - (\mathbf{v}, \nabla p) \quad \text{on } H^1 \cap L_0^2$$
$$\text{with respect to } \|[\mathbf{v}, q]\| := (\|\mathbf{v}\|^2 + \|\nabla q\|^2)^{\frac{1}{2}}.$$

The discrete inf-sup compatibility conditions (2.10) and (2.11) are the limit cases of (2.17) from Lemma 2.2.

Likewise, the uniform estimate in Lemma 3.2 for $\alpha = 0$ becomes a standard error estimate for the Stokes problem. As for $\alpha \to \infty$, the optimal error estimate for the finite element solution to the Darcy problem (3.8) in $(\|\mathbf{v}\|^2 + \|\nabla q\|^2)^{\frac{1}{2}}$ norm is recovered from (3.6). The error estimate of Theorem 3.4 makes sense in the case of $\alpha = 0$ only if $g = 0$, recovering the $O(h^2)$ convergence for velocity and $O(h)$ for pressure in the L^2-norm provided $\mathbf{f} \in \mathbf{L}^2$: $\|\mathbf{u} - \mathbf{u}_h\| + h_k\|p - p_h\| \lesssim h_k^2\|\mathbf{f}\|$. If $g \neq 0$, the second term on the right-hand side of (3.6) blows up for $h_k \to 0$. The latter can be seen as a consequence of the lack of H^2 regularity for the Stokes problem for a generic g in a convex domain [12], which implies that an extra convergence order for the L^2-norm of the velocity error is not recovered by the standard duality arguments for $g \neq 0$. If $\alpha \to \infty$, the estimate of Theorem 3.4 results in the energy type a priori estimate for the Darcy problem solutions: $\|\mathbf{u} - \mathbf{u}_h\| + \|\nabla(p - p_h)\| \lesssim \|\mathbf{f}\| + \|g\|_{H^{-1}}$. Thus, Theorem 3.4 interpolates between this two results and provides an approximation property sufficient for the multigrid analysis below.

Remark 5. The time dependent Stokes problem (1.1) can also be considered as a model for porous media flow coupled with viscous fluid flow in a single form of equation. With such physical background it appears in the literature as the Darcy-Stokes or Brinkman equations, e.g., [13, 20, 35]. While the analysis of a multigrid solver convergence is not tied to any particular modeling content, the common notion of uniformly stable elements for the Darcy-Stokes-Brinkman equations differs from the estimate given by Lemma 3.2: For the Darcy-Stokes-Brinkman equations the uniform stability is typically sought in the $\left(\alpha^{-1}\mathbf{H}_0^1 \cap \mathbf{H}_0(\mathrm{div})\right) \times L_0^2$ space [20], while (3.3) shows uniform stability in $\left(\alpha^{-1}\mathbf{H}_0^1 \cap \mathbf{L}^2\right) \times \left(H^1 \cap L_0^2\right)$.

4. Multigrid method and algebraic framework

For the approximate solution of the discrete problem (2.15) we apply a multigrid method. The method and its convergence analysis will be presented in a matrix-vector form as in Hackbush [16]. To this end consider a space $\mathbb{Q}_k^+ := \mathbb{Q}_k \oplus \mathrm{span}\{1\}$, i.e., a pressure finite element space without orthogonality condition. Denote by $\{\phi_i\}_{1 \le i \le n_k}$ and $\{\psi_i\}_{1 \le i \le m_k}$ the standard nodal bases in $\mathbf{V}_k$ and $\mathbb{Q}_k^+$. Consider the isomorphisms:

$$P_k:\ \mathbf{X}_k := \mathbb{R}^{n_k} \to \mathbf{V}_k, \quad P_k\mathrm{u} = \sum_{i=1}^{n_k} u^i\phi_i,$$
$$R_k:\ \mathbb{Y}_k := \mathbb{R}^{m_k} \to \mathbb{Q}_k^+, \quad R_k\mathrm{p} = \sum_{i=1}^{m_k} p^i\psi_i.$$

Both on $\mathbf{X}_k$ and $\mathbb{Y}_k$ we use a Euclidean scalar product scaled with h_k^d, e.g., on $\mathbf{X}_k$ we use $\langle \mathrm{u},\mathrm{v}\rangle = h_k^d\sum_{i=1}^{n_k} u^i v^i$ and corresponding norm denoted by $\|\cdot\|_k$. The following norm equivalences hold on $\mathbf{X}_k$ and $\mathbb{Y}_k$:

$$\|\mathrm{u}\|_k \lesssim \|P_k\mathrm{u}\| \lesssim \|\mathrm{u}\|_k \quad \forall\, \mathrm{u}\in\mathbf{X}_k, \tag{4.1}$$
$$\|\mathrm{p}\|_k \lesssim \|R_k\mathrm{p}\| \lesssim \|\mathrm{p}\|_k \quad \forall\, \mathrm{p}\in\mathbb{Y}_k. \tag{4.2}$$

Let the matrices $\mathrm{A}_k \in \mathbb{R}^{n_k\times n_k}$, $\mathrm{B}_k \in \mathbb{R}^{m_k\times n_k}$ and the velocity and pressure mass matrices $\mathrm{M}_u \in \mathbb{R}^{n_k\times n_k}$ and $\mathrm{M}_p \in \mathbb{R}^{m_k\times m_k}$ be given by

$$\begin{aligned}
\langle \mathrm{A}_k\mathrm{u},\mathrm{v}\rangle &= (\nabla\mathbf{u}_k,\nabla\mathbf{v}_k) + \alpha(\mathbf{u}_k,\mathbf{v}_k) \quad \forall\, \mathrm{u},\mathrm{v}\in\mathbb{R}^{n_k},\ \mathbf{u}_k = P_k\mathrm{u},\ \mathrm{v}_k = P_k\mathrm{v},\\
\langle \mathrm{B}_k\mathrm{u},\mathrm{p}\rangle &= (\mathrm{div}\,\mathbf{u}_k, p_k) \quad \forall\, \mathrm{u}\in\mathbb{R}^{n_k},\ \mathrm{p}\in\mathbb{R}^{m_k},\ \mathbf{u}_k = P_k\mathrm{u},\ p_k = R_k\mathrm{p},\\
\langle \mathrm{M}_u\mathrm{u},\mathrm{v}\rangle &= (\mathbf{u}_k,\mathbf{v}_k) \quad \forall\, \mathrm{u},\mathrm{v}\in\mathbb{R}^{n_k},\ \mathbf{u}_k = P_k\mathrm{u},\ \mathrm{v}_k = P_k\mathrm{v},\\
\langle \mathrm{M}_p\mathrm{q},\mathrm{p}\rangle &= (q_k,p_k) \quad \forall\, \mathrm{q},\mathrm{p}\in\mathbb{R}^{m_k},\ q_k = R_k\mathrm{q},\ p_k = R_k\mathrm{p}.
\end{aligned} \tag{4.3}$$

Finite element formulation (2.15) can be written as a linear system of the form

$$\begin{pmatrix} \mathrm{A}_k & \mathrm{B}_k^T \\ \mathrm{B}_k & 0 \end{pmatrix}\begin{pmatrix} \mathrm{u} \\ \mathrm{p} \end{pmatrix} = \begin{pmatrix} \mathrm{f} \\ \mathrm{g} \end{pmatrix} \tag{4.4}$$

with f and g such that $\langle \mathrm{f},\mathrm{v}\rangle = (\mathbf{f}, P_k\mathrm{v})$ for all $\mathrm{v}\in\mathbb{R}^{n_k}$ and $\langle \mathrm{g},\mathrm{q}\rangle = (g, R_k\mathrm{q})$ for all $\mathrm{q}\in\mathbb{R}^{m_k}$. By $\mathcal{A}_k$ and S_k we denote the stiffness and pressure Schur complement matrices of the finite element problem (2.15) on level k:

$$\mathcal{A}_k := \begin{pmatrix} \mathrm{A}_k & \mathrm{B}_k^T \\ \mathrm{B}_k & 0 \end{pmatrix}, \qquad \mathrm{S}_k := \mathrm{B}_k\mathrm{A}_k^{-1}\mathrm{B}_k^T. \tag{4.5}$$

Remark 6. Note that both S_k and $\mathcal{A}_k$ are singular matrices and have a one-dimensional kernel. Define the constant vector $\mathrm{e} := R_k^{-1}1 = (1,\dots,1)^T \in \mathbb{R}^m$. Then we have $\ker(\mathrm{S}) = \mathrm{span}\{\mathrm{e}\}$. Note that

$$(R_k\mathrm{p},1) = 0 \ \Leftrightarrow\ (R_k\mathrm{p},R_k\mathrm{e}) = 0 \ \Leftrightarrow \langle \mathrm{M}_p\mathrm{p},\mathrm{e}\rangle = 0 \ \Leftrightarrow\ \langle \mathrm{p},\mathrm{M}_p\mathrm{e}\rangle = 0.$$

Thus the orthogonality condition in $\mathbb{Q}_k$ corresponds to the orthogonality to the vector $\mathrm{M}_p\mathrm{e}$ in $\mathbb{R}^{m_k}$. This can be written as $\mathbb{Q}_k = \{\, R_k\mathrm{p} \mid \mathrm{p}\in(\mathrm{M}_p\mathrm{e})^\perp \,\}$. Denote $\tilde{\mathbb{Y}}_k := (\mathrm{M}_p\mathrm{e})^\perp$. Below we always consider S_k on the subspace $\tilde{\mathbb{Y}}_k$ and $\mathcal{A}_k$ on the subspace $\mathbf{X}_k\times\tilde{\mathbb{Y}}_k$. Moreover, from the definition of the $\|\cdot\|_{Q_k}$ norm and S_k it follows that

$$\langle \mathrm{S}_k\mathrm{p},\mathrm{p}\rangle = \|p_k\|_{Q_k}^2 \quad \forall\, \mathrm{p}\in\mathbb{R}^{m_k},\ p_k = R_k\mathrm{p}. \tag{4.6}$$

Furthermore, we define two product norms on $\mathbf{X}_k\times\tilde{\mathbb{Y}}_k$:

$$\begin{aligned}
\|\mathrm{u},\mathrm{p}\|_{S_k} &:= \left(\|\mathrm{u}\|_k^2 + \min\{h_k^2,\alpha^{-1}\}\langle \mathrm{S}_k\mathrm{p},\mathrm{p}\rangle\right)^{\frac12},\\
\|\mathrm{u},\mathrm{p}\|_{S_k^{-1}} &:= \left(\|\mathrm{u}\|_k^2 + \max\{h_k^{-2},\alpha\}\langle \mathrm{S}_k^{-1}\mathrm{p},\mathrm{p}\rangle\right)^{\frac12}.
\end{aligned} \tag{4.7}$$

Denote $\mathrm{D}_k = \mathrm{diag}(\mathrm{A}_k)$. Due to regular mesh refinement and (2.10) the following relations hold (cf. [23], [24]):

$$(1+\alpha)\mathrm{I}_k \;\lesssim\; \mathrm{A}_k \;\lesssim\; \mathrm{D}_k, \tag{4.8}$$

$$\mathrm{D}_k \;\simeq\; (h_k^{-2}+\alpha)\mathrm{I}_k, \tag{4.9}$$

$$\mathrm{S}_k^{-1} \;\simeq\; \mathrm{I}_k + \alpha(\mathrm{B}_k\mathrm{M}_u^{-1}\mathrm{B}_k^T)^{-1}. \tag{4.10}$$

Here and further I_k is the identity matrix for a corresponding vector space and $A \le B$ for two square matrices if $B-A$ is non-negative definite.

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

$$\begin{aligned} &\mathbf{p}_k : \mathbf{X}_{k-1}\times\tilde{\mathbb{Y}}_{k-1} \to \mathbf{X}_k\times\tilde{\mathbb{Y}}_k, \quad \mathbf{p}_k = P_k^{-1}P_{k-1}\times R_k^{-1}R_{k-1} \\ &\mathbf{r}_k : \mathbf{X}_k\times\tilde{\mathbb{Y}}_k \to \mathbf{X}_{k-1}\times\tilde{\mathbb{Y}}_{k-1}, \quad \mathbf{r}_k = P_{k-1}^*(P_k^*)^{-1}\times R_{k-1}^*(R_k^*)^{-1}. \end{aligned} \tag{4.11}$$

Note that both $\mathbf{p}_k$ and $\mathbf{r}_k$ keep the pressure vector in the right subspace.

In Section 6 we consider several linear smoothing iterations of the form

$$\{\mathrm{u}^{\mathrm{new}},\mathrm{p}^{\mathrm{new}}\} = \{\mathrm{u}^{\mathrm{old}},\mathrm{p}^{\mathrm{old}}\} - \mathcal{W}_k^{-1}(\mathcal{A}_k\{\mathrm{u}^{\mathrm{old}},\mathrm{p}^{\mathrm{old}}\}-b) \quad \text{for } \{\mathrm{u}^{\mathrm{old}},\mathrm{p}^{\mathrm{old}}\}, b\in\mathbf{X}_k\times\tilde{\mathbb{Y}}_k$$

with the corresponding iteration matrix denoted by

$$\mathcal{L}_k = \mathcal{I}_k - \mathcal{W}_k^{-1}\mathcal{A}_k. \tag{4.12}$$

With the components defined above a standard multigrid algorithm with ν pre-smoothing iterations can be formulated (cf. [16]) with an iteration matrix that satisfies the recursion

$$\mathcal{M}_0 = 0, \quad \mathcal{M}_k = \left(\mathcal{I}_k - \mathbf{p}_k(\mathcal{I}_k - \mathcal{M}_{k-1}^{\gamma})\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k\right)\mathcal{L}_k^{\nu}, \quad k=1,2,\ldots.$$

The choices $\gamma=1$ and $\gamma=2$ correspond to the V- and W-cycle, respectively. For the analysis of this multigrid method we use the framework of [16] based on the approximation and smoothing property. Below we derive these properties for the generalized Stokes problem.

5. Approximation property

The theorem below states the necessary approximation property.

Theorem 5.1 (Approximation property)**.** *Let $\mathcal{A}_k$ be the stiffness matrix from* (4.5) *and $\mathbf{p}_k,\mathbf{r}_k$ the prolongation and restriction as in* (4.11)*. Then under the assumptions of Theorem* 3.4 *the following approximation property holds:*

$$\|\mathcal{A}_k^{-1} - \mathbf{p}_k\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\|_{S_k^{-1}\to S_k} \lesssim \min\left\{h_k^2,\alpha^{-1}\right\}.$$

Proof. Take $\{\mathrm{f}_k,\mathrm{g}_k\}\in\mathbf{X}_k\times\tilde{\mathbb{Y}}_k$. Let $\{\mathbf{u},p\}\in\mathbf{V}\times\mathbb{Q}$, $\{\mathbf{u}_k,p_k\}\in\mathbf{V}_k\times\mathbb{Q}_k$, and $\{\mathbf{u}_{k-1},p_{k-1}\}\in\mathbf{V}_{k-1}\times\mathbb{Q}_{k-1}$ be such that

$$a(\mathbf{u},p;\mathbf{v},q) = ((P_k^*)^{-1}\mathrm{f}_k,\mathbf{v}) + ((R_k^*)^{-1}\mathrm{g}_k,r) \quad \forall\,\{\mathbf{v},q\}\in\mathbf{V}\times\mathbb{Q},$$

$$a(\mathbf{u}_k,p_k;\mathbf{v},q) = ((P_k^*)^{-1}\mathrm{f}_k,\mathbf{v}) + ((R_k^*)^{-1}\mathrm{g}_k,r) \quad \forall\,\{\mathbf{v},q\}\in\mathbf{V}_k\times\mathbb{Q}_k,$$

$$a(\mathbf{u}_{k-1},p_{k-1};\mathbf{v},q) = ((P_k^*)^{-1}\mathrm{f}_k,\mathbf{v}) + ((R_k^*)^{-1}\mathrm{g}_k,r) \quad \forall\,\{\mathbf{v},q\}\in\mathbf{V}_{k-1}\times\mathbb{Q}_{k-1}.$$

Putting $\mathbf{f}=(P_k^*)^{-1}\mathrm{f}_k$ and $g=(R_k^*)^{-1}\mathrm{g}_k$ in Theorem 3.4, we obtain

$$\begin{aligned} &\|\mathbf{u}-\mathbf{u}_l\| + \min\{h_l,\alpha^{-\frac12}\}\|p-p_l\|_Q \\ &\qquad\qquad \lesssim \min\{h_l^2,\alpha^{-1}\}\left(\|(P_k^*)^{-1}\mathrm{f}_k\| + \max\{h_l^{-1},\alpha^{\frac12}\}\|(R_k^*)^{-1}\mathrm{g}_k\|_{Q^*}\right) \end{aligned}$$

with $l=k$ and $l=k-1$. Due to $h_{k-1}\lesssim h_k$ this yields

$$\|\mathbf{u}_k-\mathbf{u}_{k-1}\|+\min\{h_k,\alpha^{-\frac12}\}\|p_k-p_{k-1}\|_Q \lesssim \min\{h_k^2,\alpha^{-1}\}\left(\|(P_k^*)^{-1}\mathrm{f}_k\|+\max\{h_k^{-1},\alpha^{\frac12}\}\|(R_k^*)^{-1}\mathrm{g}_k\|_{Q^*}\right).$$

Now we use the result of Lemma 2.2 to obtain

$$\|\mathbf{u}_k-\mathbf{u}_{k-1}\|+\min\{h_k,\alpha^{-\frac12}\}\|p_k-p_{k-1}\|_{Q_k} \lesssim \min\{h_k^2,\alpha^{-1}\}\left(\|(P_k^*)^{-1}\mathrm{f}_k\|+\max\{h_k^{-1},\alpha^{\frac12}\}\|(R_k^*)^{-1}\mathrm{g}_k\|_{Q_k^*}\right).$$

From the definition of $\mathcal{A}_k$ and (4.11) it follows that $\{\mathbf{u}_k,p_k\}=(P_k,R_k)^T\mathcal{A}_k^{-1}\{\mathrm{f}_k,\mathrm{g}_k\}$ and $\{\mathbf{u}_{k-1},p_{k-1}\}=(P_{k-1},R_{k-1})^T\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\{\mathrm{f}_k,\mathrm{g}_k\}$. Thus, using (4.1) and (4.6), we get

$$\begin{aligned}\|(\mathcal{A}_k^{-1}-&\mathbf{p}_k\mathcal{A}_{k-1}^{-1}\mathbf{r}_k)\{\mathrm{f}_k,\mathrm{g}_k\}\|_{S_k}\\ &\simeq \|\mathbf{u}_k-\mathbf{u}_{k-1}\|+\min\{h,\alpha^{-\frac12}\}\|p_k-p_{k-1}\|_{Q_k}\\ &\lesssim \min\{h_k^2,\alpha^{-1}\}\left(\|(P_k^*)^{-1}\mathrm{f}_k\|+\max\{h^{-1},\alpha^{\frac12}\}\|(R_k^*)^{-1}\mathrm{g}_k\|_{Q_k^*}\right)\\ &\simeq \min\{h_k^2,\alpha^{-1}\}\|\mathrm{f}_k,\mathrm{g}_k\|_{S_k^{-1}},\end{aligned}$$

which proves the theorem. □

Based on the "inexact" Schur complement $\widehat{\mathrm{S}}_k=\mathrm{B}_k\mathrm{D}_k^{-1}\mathrm{B}_k^T$, we define two more product norms on $\mathbf{X}_k\times\tilde{\mathbb{Y}}_k$:

$$\begin{aligned}\|\mathrm{u},\mathrm{p}\|_{\widehat{S}_k}&:=\left(\|\mathrm{u}\|_k^2+\min\{h_k^2,\alpha^{-1}\}\langle\widehat{\mathrm{S}}_k\mathrm{p},\mathrm{p}\rangle\right)^{\frac12},\\ \|\mathrm{u},\mathrm{p}\|_{\widehat{S}_k^{-1}}&:=\left(\|\mathrm{u}\|_k^2+\max\{h_k^{-2},\alpha\}\langle\widehat{\mathrm{S}}_k^{-1}\mathrm{p},\mathrm{p}\rangle\right)^{\frac12}.\end{aligned}\tag{5.1}$$

Thanks to (4.8) it holds that $\widehat{\mathrm{S}}_k\lesssim\mathrm{S}_k$. Therefore, we get from Theorem 5.1:

Corollary 5.2. *Under the assumptions of Theorem* 5.1 *the following approximation property holds:*

$$\|\mathcal{A}_k^{-1}-\mathbf{p}_k\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\|_{\widehat{S}_k^{-1}\to\widehat{S}_k}\lesssim\min\left\{h_k^2,\alpha^{-1}\right\}.$$

6. Smoothing property

In this section we prove a smoothing property for several iterative methods (smoothers) known from the literature. This smoothing property will complement the approximation property from the previous section, resulting in the uniform estimate of the two-grid convergence. We also analyze stability of smoothing iterations, since this property is used for proving multigrid W-cycle convergence.

We will need the following result; cf. e.g. [17].

Lemma 6.1. *Assume $A,\widehat{A},\widehat{S}$ are symmetric positive definite and $S=BA^{-1}B^T$. Assume also that the inequalities*

$$\rho_1\widehat{A}\ \le\ A\ \le\ \rho_2\widehat{A},\tag{6.1}$$

$$\mu_1\widehat{S}\ \le\ S\ \le\ \mu_2\widehat{S}\tag{6.2}$$

hold with positive constants $\rho_1, \rho_2, \mu_1, \mu_2$. Then all eigenvalues of the problem

$$\begin{array}{rcl} A\mathrm{u} + B^T\mathrm{p} &=& \lambda \widehat{A}\mathrm{u}, \\ B\mathrm{u} &=& \lambda \widehat{S}\mathrm{p}, \end{array} \tag{6.3}$$

belong to

$$[\rho_1, \rho_2] \bigcup \left[\frac{\rho_1 + \sqrt{\rho_1^2 + 4\rho_1\mu_1}}{2}, \frac{\rho_2 + \sqrt{\rho_2^2 + 4\rho_2\mu_2}}{2} \right] \bigcup \left[\frac{\rho_2 - \sqrt{\rho_2^2 + 4\rho_2\mu_2}}{2}, \frac{\rho_1 - \sqrt{\rho_1^2 + 4\rho_1\mu_1}}{2} \right]. \tag{6.4}$$

Remark 7. Similar eigenvalue bounds for (6.3) can be found in other papers, e.g., [28]. We will use the result of the lemma for the case when $\widehat{S}$ and S are symmetric positive definite on the subspace $\tilde{\mathbb{Y}}_k$ and the problem (6.3) has a zero eigenvalue corresponding to the eigenvector $\{0, \mathrm{e}\}$.

6.1. **Distributive iterations.** Writing the system (4.4) in the general form $\mathcal{A}\,\mathrm{x}_k = \mathrm{b}$ the idea behind the distributive smoothing iterations can be expressed as follows. One chooses matrices $\mathcal{B}$ and $\mathcal{C}$ and consider, smoothing iterations of the form:

$$\mathrm{y}^{\nu+1} = \mathrm{y}^{\nu} - \mathcal{C}^{-1}(\mathcal{A}\,\mathcal{B}\,\mathrm{y}^{\nu} - b), \quad \mathrm{x}_k = \mathcal{B}\,\mathrm{y}. \tag{6.5}$$

One possibility is to set $\mathcal{B} = \mathcal{C}^{-1}\mathcal{A}$. If $\mathcal{C} = \mathcal{C}^T > 0$, $\mathcal{C}^{-1}\mathcal{A}\mathcal{B}$ is a positive definite matrix and self-adjoint in a proper scalar product. We consider block Jacobi type iterations, i.e., $\mathcal{C}$ is a block diagonal matrix defined below. Let N_k be a matrix of the preconditioner for the discrete pressure Neumman problem, such that

$$\mathrm{N}_k \;\simeq\; \mathrm{B}_k \mathrm{M}_u^{-1} \mathrm{B}_k^T. \tag{6.6}$$

Define a block diagonal matrix $\mathcal{D}_k$ as

$$\mathcal{D}_k^{-1} \;=\; \begin{pmatrix} \mathrm{D}_k^{-1} & 0 \\ 0 & \mathrm{I}_k + \alpha \mathrm{N}_k^{-1} \end{pmatrix} \tag{6.7}$$

and set $\mathcal{C} = \omega\mathcal{D}_k$ with a parameter $\omega > 0$. The iteration matrix $\mathcal{L}_k$ in this case can be written in the form (4.12) with $\mathcal{W}_k^{-1} = \omega^2 \mathcal{D}_k^{-1} \mathcal{A}_k \mathcal{D}_k^{-1}$.

Theorem 6.2 (Smoothing property)**.** *Assume $\omega > 0$ is small enough, but independent of α and k. It holds that*

$$\|\mathcal{A}_k \mathcal{L}_k^{\nu}\|_{S_k \to S_k^{-1}} \lesssim (h_k^{-2} + \alpha) \frac{1}{\sqrt{2\nu + 1}}. \tag{6.8}$$

Proof. With the auxiliary matrix

$$\mathcal{D}_s = \begin{pmatrix} \mathrm{I}_k & 0 \\ 0 & \min\{h_k^2, \alpha^{-1}\}\mathrm{S}_k \end{pmatrix}, \tag{6.9}$$

it holds that

$$\begin{aligned} \|\mathcal{A}_k \mathcal{L}_k^{\nu}\|_{S_k \to S_k^{-1}} &= \|\mathcal{A}_k (\mathcal{I}_k - \omega^2 \mathcal{D}_k^{-1} \mathcal{A}_k \mathcal{D}_k^{-1} \mathcal{A}_k)^{\nu}\|_{S_k \to S_k^{-1}} \\ &= \|\mathcal{D}_s^{-\frac{1}{2}} \mathcal{A}_k (\mathcal{I}_k - \omega^2 \mathcal{D}_k^{-1} \mathcal{A}_k \mathcal{D}_k^{-1} \mathcal{A}_k)^{\nu} \mathcal{D}_s^{-\frac{1}{2}}\|. \end{aligned}$$

Here and further $\|\mathcal{C}\|$ denotes the spectral norm of a matric $\mathcal{C}$.

Denote $\bar{\mathcal{A}} = \omega\mathcal{D}_k^{-\frac{1}{2}}\mathcal{A}_k\mathcal{D}_k^{-\frac{1}{2}}$ and observe the equality

$$\|\mathcal{D}_s^{-\frac{1}{2}}\mathcal{A}_k(\mathcal{I}_k - \omega^2\mathcal{D}_k^{-1}\mathcal{A}_k\mathcal{D}_k^{-1}\mathcal{A}_k)^\nu\mathcal{D}_s^{-\frac{1}{2}}\| = \|\omega^{-1}\mathcal{D}_s^{-\frac{1}{2}}\mathcal{D}_k^{\frac{1}{2}}\bar{\mathcal{A}}(\mathcal{I}_k - \bar{\mathcal{A}}^2)^\nu\mathcal{D}_k^{\frac{1}{2}}\mathcal{D}_s^{-\frac{1}{2}}\|.$$

We get

$$\|\mathcal{A}_k\mathcal{L}_k^\nu\|_{S_k\to S_k^{-1}} \le \omega^{-1}\|\mathcal{D}_k\mathcal{D}_s^{-1}\|\|\bar{\mathcal{A}}(\mathcal{I}_k - \bar{\mathcal{A}}^2)^\nu\|.$$

Thanks to the eigenvalue estimate of Lemma 6.1 and bounds in (4.8) and (6.6) one can choose such $\omega \gtrsim 1$ that $\mathrm{sp}(\bar{\mathcal{A}}) \in [-1, 1]$. Hence

$$\|\bar{\mathcal{A}}(\mathcal{I}_k - \bar{\mathcal{A}}^2)^\nu\| \le \max_{x\in[-1,1]} |x(1-x^2)^\nu| \le \frac{1}{\sqrt{2\nu+1}}.$$

Finally, we use (4.9), (4.10) and (6.6) to verify that

$$\|\mathcal{D}_k\mathcal{D}_s^{-1}\| \lesssim (h_k^{-2} + \alpha). \qquad \square$$

Theorem 6.3 (Stability of smoother)**.** *With the same choice of ω as in Theorem 6.2 it holds that*

$$\|\mathcal{L}_k^\nu\|_{S_k\to S_k} \lesssim 1. \tag{6.10}$$

Proof. From (4.9), (4.10) and (6.6) we get $\max\{h_k^{-2}, \alpha\}\mathcal{D}_s \simeq \mathcal{D}_k$ for the matrices $\mathcal{D}_s$ and $\mathcal{D}_k$ defined in (6.7) and (6.9). Therefore,

$$\|\mathcal{L}_k^\nu\|_{S_k\to S_k} = \|\mathcal{D}_s^{\frac{1}{2}}\mathcal{L}_k^\nu\mathcal{D}_s^{-\frac{1}{2}}\| \simeq \|\mathcal{D}_k^{\frac{1}{2}}\mathcal{L}_k^\nu\mathcal{D}_k^{-\frac{1}{2}}\| = \|(\mathcal{I}_k - \bar{\mathcal{A}}^2)^\nu\| \tag{6.11}$$

with $\bar{\mathcal{A}} = \omega\mathcal{D}_k^{-\frac{1}{2}}\mathcal{A}_k\mathcal{D}_k^{-\frac{1}{2}}$. In the proof of Theorem 6.2 we have shown that $\mathrm{sp}(\bar{\mathcal{A}}) \in [-1, 1]$. Hence it holds

$$\|(\mathcal{I}_k - \bar{\mathcal{A}}^2)^\nu\| \le 1. \tag{6.12}$$

Inequalities (6.11)–(6.12) yield (6.10). $\square$

For the Stokes problem ($\alpha = 0$) similar smoothing iterations were considered first in [33] and [26]. The smoother from [33] and [26] can be written in the form of (6.5) with

$$\mathcal{C} = \omega\begin{pmatrix} h_k^{-2}\mathrm{I}_k & 0 \\ 0 & \mathrm{I}_k \end{pmatrix}.$$

Clearly, its analysis fits in the framework given in this paper.

6.2. **Braess-Sarazin and inexact Uzawa smoothers.** In this section D_k is an arbitrary symmetric matrix satisfying (4.8) and (4.9). One may still think of D_k as $\mathrm{D}_k = \mathrm{diag}(\mathrm{A}_k)$. Other reasonable choices are $\mathrm{D}_k = (h^{-2} + \alpha)\mathrm{I}_k$ or $\mathrm{D}_k = (h^{-2} + \alpha)\mathrm{M}_u$, where M_u is the velocity mass matrix or its diagonal approximation. Let ω be some given positive parameter. Consider iterations of the form:

$$\begin{pmatrix} \mathrm{u}^{\mathrm{new}} \\ \mathrm{p}^{\mathrm{new}} \end{pmatrix} = \begin{pmatrix} \mathrm{u}^{\mathrm{old}} \\ \mathrm{p}^{\mathrm{old}} \end{pmatrix} - \begin{pmatrix} \omega\mathrm{D}_k & \mathrm{B}_k^T \\ \mathrm{B}_k & 0 \end{pmatrix}^{-1} \left\{ \begin{pmatrix} \mathrm{A}_k & \mathrm{B}_k^T \\ \mathrm{B}_k & 0 \end{pmatrix} \begin{pmatrix} \mathrm{u}^{\mathrm{old}} \\ \mathrm{p}^{\mathrm{old}} \end{pmatrix} - \begin{pmatrix} \mathrm{f} \\ \mathrm{g} \end{pmatrix} \right\}. \tag{6.13}$$

At each iteration (6.13) one has to solve the auxiliary system:

$$\begin{pmatrix} \omega\mathrm{D}_k & \mathrm{B}_k^T \\ \mathrm{B}_k & 0 \end{pmatrix} \begin{pmatrix} \mathrm{v} \\ \mathrm{q} \end{pmatrix} = \begin{pmatrix} \mathrm{r}^{\mathrm{old}} \\ \mathrm{B}_k\,\mathrm{u}^{\mathrm{old}} - \mathrm{g} \end{pmatrix}. \tag{6.14}$$

To solve (6.14) one can eliminate v from the system (6.14) and obtain a problem for the q variable (we recall the notation $\widehat{\mathrm{S}}_k = \mathrm{B}_k\,\mathrm{D}_k^{-1}\mathrm{B}_k^T$):

$$\widehat{\mathrm{S}}_k\mathrm{q} = \mathrm{B}_k\,\mathrm{D}_k^{-1}\mathrm{r}^{\mathrm{old}} - \omega(\mathrm{B}_k\mathrm{u}^{\mathrm{old}} - \mathrm{g}). \tag{6.15}$$

The upper bound in (4.8) yields $\lambda_{\max}(\mathrm{D}_k^{-1}\mathrm{A}_k) \lesssim 1$. Thus one can choose ω satisfying

$$\omega > \lambda_{\max}(\mathrm{D}_k^{-1}\mathrm{A}_k) \quad \text{and} \quad \omega \simeq 1. \tag{6.16}$$

Remark 8. Smoothing iterations (6.13) were first proposed in [6] with $\mathrm{D}_k = \mathrm{I}_k$ for the case $\mathrm{g} = 0$; see also [5]. A more general choice of D_k was analyzed in [37] and [19]. Considering a general $\mathrm{g} \in \tilde{\mathbb{Y}}_k$ causes no additional difficulties.

The method requires an *exact* solution of the problem (6.15) which can be interpreted as a discrete pressure Poisson problem. Note that the distributive smoother from Section 6.1 requires an *approximate* solution of a similar problem; cf. (6.6). Below we also consider a smoother closely related to (6.13), which avoids the exact solution of (6.15). Hence consider the block iterative method from [2], which can be seen as a variant of inexact Uzawa method. Let G_k be a preconditioner for $\widehat{\mathrm{S}}_k$ such that

$$\mathrm{G}_k < \omega^{-1}\widehat{\mathrm{S}}_k \leq (1+\beta)\mathrm{G}_k, \qquad \beta > 0. \tag{6.17}$$

One step of the method can be divided in the following three substeps:

$$\omega \mathrm{D}_k(\mathrm{u}^{\mathrm{aux}} - \mathrm{u}^{\mathrm{old}}) = \mathrm{f} - \mathrm{A}_k\mathrm{u}^{\mathrm{old}} - \mathrm{B}_k^T\mathrm{p}^{\mathrm{old}}, \tag{6.18}$$

$$\mathrm{G}_k(\mathrm{p}^{\mathrm{new}} - \mathrm{p}^{\mathrm{old}}) = \mathrm{B}_k\mathrm{u}^{\mathrm{aux}} - \mathrm{g}, \tag{6.19}$$

$$\omega \mathrm{D}_k(\mathrm{u}^{\mathrm{new}} - \mathrm{u}^{\mathrm{aux}}) = -\mathrm{B}_k^T(\mathrm{p}^{\mathrm{new}} - \mathrm{p}^{\mathrm{old}}). \tag{6.20}$$

The iteration matrix of the method (6.18)–(6.20) is written in the form (4.12) with

$$\mathcal{W}_k = \begin{pmatrix} \omega \mathrm{D}_k & \mathrm{B}_k^T \\ \mathrm{B}_k & \omega^{-1}\widehat{\mathrm{S}}_k - \mathrm{G}_k \end{pmatrix}.$$

Thus iterations (6.13) can be interpret as (6.18)–(6.20) with exact preconditioner for the "inexact" Schur complement $\widehat{S}_k$ (for the sake of analysis we need a strict lower bound in (6.17), however). The smoothing property of (6.18)–(6.20) is based on the following lemma from [37].

Lemma 6.4. *Assume* (6.16) *and* (6.17)*. Denote* $\widetilde{\mathcal{D}}_s = \begin{pmatrix} \omega\mathrm{D}_k - \mathrm{A}_k & 0 \\ 0 & \omega^{-1}\widehat{\mathrm{S}}_k - \mathrm{G}_k \end{pmatrix}$, *then the matrix* $\bar{\mathcal{L}}_k = \widetilde{\mathcal{D}}_s^{\frac{1}{2}}\mathcal{L}_k\widetilde{\mathcal{D}}_s^{-\frac{1}{2}}$ *is symmetric and*

$$\mathrm{sp}\,(\bar{\mathcal{L}}_k) \in [-\beta - \sqrt{\beta^2+\beta}, 1].$$

Moreover, the identity $\mathcal{A}_k\mathcal{L}_k^{\nu} = \widetilde{\mathcal{D}}_s^{\frac{1}{2}}(\mathcal{I}_k - \bar{\mathcal{L}}_k)\bar{\mathcal{L}}_k^{\nu-1}\widetilde{\mathcal{D}}_s^{-\frac{1}{2}}$ *holds.*

Now Lemma 6.4 leads us to the smoothing property for (6.13) and (6.18)–(6.20) which complements the approximation property from Corollary 5.2:

Theorem 6.5 (Smoothing property)**.** *Let* $\mathcal{L}_k$ *be the iteration matrix of* (6.18)–(6.20)*. Assume* (6.16) *and* (6.17) *with* $\beta < \frac{1}{3}$*, then*

$$\|\mathcal{A}_k\mathcal{L}_k^{\nu}\|_{\widehat{S}_k \to \widehat{S}_k^{-1}} \lesssim (h_k^{-2} + \alpha)\frac{1}{\nu - 1}, \qquad \nu > 1. \tag{6.21}$$

Proof. Define the auxiliary matrix $\widehat{\mathcal{D}}_s = \begin{pmatrix} \mathrm{I}_k & 0 \\ 0 & \min\{h_k^2, \alpha^{-1}\}\widehat{\mathrm{S}}_k \end{pmatrix}$, then $\|\cdot\|_{\widehat{S}_k} = \langle \widehat{\mathcal{D}}_s \cdot, \cdot \rangle^{\frac{1}{2}}$. Thanks to (4.9), (6.16) and (6.17) we obtain $\|\widehat{\mathcal{D}}_s^{-1}\widetilde{\mathcal{D}}_s\| \lesssim h_k^{-2} + \alpha$. Therefore, Lemma 6.4 and assumption $\beta < \frac{1}{3}$ yield

$$\begin{aligned}
\|\mathcal{A}_k\mathcal{L}_k^\nu\|_{\widehat{S}_k \to \widehat{S}_k^{-1}} &= \|\widehat{\mathcal{D}}_s^{-\frac{1}{2}}\mathcal{A}_k\mathcal{L}_k^\nu\widehat{\mathcal{D}}_s^{-\frac{1}{2}}\| = \|\widehat{\mathcal{D}}_s^{-\frac{1}{2}}\widetilde{\mathcal{D}}_s^{\frac{1}{2}}(\mathcal{I}_k - \bar{\mathcal{L}}_k)\bar{\mathcal{L}}_k^{\nu-1}\widetilde{\mathcal{D}}_s^{-\frac{1}{2}}\widehat{\mathcal{D}}_s^{-\frac{1}{2}}\| \\
&\le \|\widehat{\mathcal{D}}_s^{-1}\widetilde{\mathcal{D}}_s\|\|(\mathcal{I}_k - \bar{\mathcal{L}}_k)\bar{\mathcal{L}}_k^{\nu-1}\| \\
&\lesssim (h_k^{-2} + \alpha) \max_{x \in [-\beta - \sqrt{\beta^2+\beta}, 1]} |(1-x)x^{\nu-1}| \\
&\lesssim \frac{h_k^{-2} + \alpha}{\nu - 1}.
\end{aligned}$$

□

Theorem 6.5 together with Theorem 5.1 guarantee the uniform convergence estimates for the two-grid method with Braess-Sarazin or inexact Uzawa smoothings. To analyze multigrid convergence we need stability property from the theorem below.

Theorem 6.6 (Stability of smoother)**.** *Assume* (6.16) *and* (6.17) *with* $\beta \le \frac{1}{3}$, *then*

$$\|\mathcal{L}_k^\nu\|_{\widehat{S}_k \to \widehat{S}_k} \lesssim 1. \tag{6.22}$$

Proof. Define the following product norms on $\mathbf{X}_k \times \tilde{\mathbb{Y}}_k$:

$$|||\mathrm{u}, \mathrm{p}||| := \left(\omega\langle \mathrm{D}_k\mathrm{u}, \mathrm{u}\rangle + \omega^{-1}\langle \widehat{\mathrm{S}}_k\mathrm{p}, \mathrm{p}\rangle\right)^{\frac{1}{2}}.$$

Due to $\omega \simeq 1$ and (6.16) we have $\min\{h_k^2, \alpha^{-1}\}|||\mathrm{u}, \mathrm{p}||| \simeq \|\mathrm{u}, \mathrm{p}\|_{\widehat{S}_k}$. This implies

$$\|\mathcal{L}_k^\nu\|_{\widehat{S}_k \to \widehat{S}_k} \simeq |||\mathcal{L}_k^\nu|||. \tag{6.23}$$

The assumption $\omega > \lambda_{\max}(\mathrm{D}_k^{-1}\mathrm{A}_k)$ implies the eigenvalue bound

$$|\lambda(\mathrm{I}_k - \omega^{-1}\mathrm{D}_k^{-1}\mathrm{A}_k)| < 1.$$

Now we apply Theorem 2.1 from [30] with $\widehat{A} = \omega\mathrm{D}_k$, $\widehat{G} = \mathrm{B}_k\widehat{A}^{-1}\mathrm{B}_k^T$, $\beta = \frac{1}{3}$, $c = 1$ to conclude that the right-hand side of (6.23) is less than 1. □

Remark 9. Both distributive iterations and inexact Uzawa are not parameter-free smoothers. For distributive iterations parameter ω should be sufficiently *small*, but still of $O(1)$ order and independent of h and α. Let $\lambda_{\max} = \max\lambda(\mathrm{D}_k^{-1}\mathrm{A}_k)$ and $\mu_{\max} = \max\lambda((\mathrm{I}_k + \mathrm{N}_k^{-1})\mathrm{S}_k)$. Both $\lambda_{\max}$ and $\mu_{\max}$ can be bounded by $O(1)$ constants independent of h and α. Lemma 6.1 and the proof of Theorem 6.2 yield the sufficient bound on the relaxation parameter of the distributive iterations: $\omega < 2\left(\lambda_{\max} + \sqrt{\lambda_{\max}^2 + 4\mu_{\max}\lambda_{\max}}\right)^{-1}$. For the inexact Uzawa iterations parameter ω should be sufficiently *large*, but also of $O(1)$ order and independent of h and α. In this case the restriction is $\omega > \lambda_{\max}$. Numerical experiments show that optimal ω is largely insensitive to the mesh level k. Hence running a few coarse-grid tests may provide an appropriate choice of ω.

We also note that for a fixed mesh level k the eigenvalue bounds given by Lemma 6.1 with $A = \mathrm{A}_k$, $\widehat{A} = \mathrm{D}_k$, $S = \mathrm{S}_k$, and $\widehat{S} = (I_k + \mathrm{N}_k^{-1})^{-1}$ are robust with respect to α; thus the same arguments as in the proofs of Theorems 6.3 and 6.6 show that both smoothing iterations converge robust with respect to α (but not h).

7. Multigrid convergence

In this section we prove the convergence result for the multigrid method. The result is based on the approximation, smoothing and stability properties from the previous sections. First, however, we prove the following technical lemma.

Lemma 7.1. *Let* $\mathbf{p}_k$ *and* $\mathbf{r}_k$ *be the prolongation and restriction operators defined in* (4.11). *For all* $\{\mathrm{u},\mathrm{p}\} \in \mathbf{X}_{k-1}\times\tilde{\mathbb{Y}}_k$ *it holds that*

$$\|\mathbf{p}_k\{\mathrm{u},\mathrm{p}\}\|_{S_k} \simeq \|\mathrm{u},\mathrm{p}\|_{S_{k-1}} \quad \textit{and} \quad \|\mathbf{p}_k\{\mathrm{u},\mathrm{p}\}\|_{\widehat{S}_k} \simeq \|\mathrm{u},\mathrm{p}\|_{\widehat{S}_{k-1}}. \tag{7.1}$$

For the case of distributive smoothings it holds that

$$\|\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k\mathcal{L}_k^{\nu}\|_{S_k\to S_{k-1}} \lesssim 1, \tag{7.2}$$

and for smoothings (6.13) *or* (6.18)–(6.20)

$$\|\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k\mathcal{L}_k^{\nu}\|_{\widehat{S}_k\to \widehat{S}_{k-1}} \lesssim 1. \tag{7.3}$$

Proof. For arbitrary $\mathrm{u} \in \mathbf{X}_{k-1}$, $\mathrm{p} \in \tilde{\mathbb{Y}}_{k-1}$ consider $\mathbf{u}_{k-1} = P_{k-1}\mathrm{u} \in \mathbf{V}_{k-1}$ and $p_{k-1} = R_{k-1}\mathrm{p} \in \mathbb{Q}_{k-1}$, then from the definition of $\mathbf{p}_k$ and thanks to (4.1), (4.6) we conclude that

$$\begin{aligned} \|\mathrm{u},\mathrm{p}\|_{S_{k-1}} &\simeq \|\mathbf{u}_{k-1}\| + \min\{h_{k-1},\alpha^{-\frac{1}{2}}\}\|p_{k-1}\|_{Q_{k-1}}, \\ \|\mathbf{p}_k\{\mathrm{u},\mathrm{p}\}\|_{S_k} &\simeq \|\mathbf{u}_{k-1}\| + \min\{h_k,\alpha^{-\frac{1}{2}}\}\|p_{k-1}\|_{Q_k}. \end{aligned} \tag{7.4}$$

With the help of (2.17) we obtain

$$\|p_{k-1}\|_{Q_{k-1}} \simeq \|p_{k-1}\|_{Q_k}. \tag{7.5}$$

Since $h_k \simeq h_{k-1}$ relations (7.4) and (7.5) prove the first relation in (7.1).

Now consider the relations

$$\lim_{\alpha\to\infty}\alpha^{-1}\mathrm{S}_k = \mathrm{B}_k\mathrm{M}_u^{-1}\mathrm{B}_k^T \simeq \min\{h_k^2,\alpha^{-1}\}\widehat{\mathrm{S}}_k.$$

Thus we prove the second equivalence in (7.1) passing to the limit ($\alpha\to\infty$) in the first relation from (7.1) and applying the scaling argument.

Now we are going to prove (7.2). Thanks to (7.1) we have with distributive smoother:

$$\|\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k\mathcal{L}_k^{\nu}\|_{S_k\to S_{k-1}} \lesssim \|\mathbf{p}_k\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k\mathcal{L}_k^{\nu}\|_{S_k\to S_k}.$$

Now observe the following identity:

$$\mathbf{p}_k\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k\mathcal{L}_k^{\nu} = (\mathcal{A}_k^{-1} - \mathbf{p}_k\mathcal{A}_{k-1}^{-1}\mathbf{r}_k)(\mathcal{A}_k\mathcal{L}_k^{\nu}) - \mathcal{L}_k^{\nu}.$$

Hence using the approximation, smoothing and stability properties we get

$$\begin{aligned} &\|\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k\mathcal{L}_k^{\nu}\|_{S_k\to S_{k-1}} \\ &\qquad \le \|(\mathcal{A}_k^{-1} - \mathbf{p}_k\mathcal{A}_{k-1}^{-1}\mathbf{r}_k)\|_{S_k^{-1}\to S_k}\|(\mathcal{A}_k\mathcal{L}_k^{\nu})\|_{S_k\to S_k^{-1}} + \|\mathcal{L}_k^{\nu}\|_{S_k\to S_k} \lesssim 1. \end{aligned}$$

The estimate (7.3) is proved similarly. □

The iteration matrix of the multigrid W-cycle with ν pre-smoothings satisfies the recursion

$$\mathcal{M}_0 := 0,\ \mathcal{M}_k = \mathcal{T}_k + \mathbf{p}_k(\mathcal{M}_{k-1})^2\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k\mathcal{L}_k^{\nu}\ , \tag{7.6}$$

where $\mathcal{T}_k = (\mathcal{I}_k - \mathbf{p}_k\mathcal{A}_{k-1}^{-1}\mathbf{r}_k\mathcal{A}_k)\mathcal{L}_k^{\nu}$ is the iteration matrix of the two-grid method. Approximation and smoothing properties yield the estimates

$$\|\mathcal{T}_k\|_{S_k\to S_k} \lesssim \frac{1}{\sqrt{2\nu+1}}, \tag{7.7}$$

if distributive smoothings are used and

$$\|\mathcal{T}_k\|_{\widehat{S}_k\to \widehat{S}_k} \lesssim \frac{1}{\nu-1}, \qquad \nu>1, \tag{7.8}$$

if smoothings (6.13) or (6.18)–(6.20) are used.

Theorem 7.2. *Assume that the number of smoothing steps on every grid level is sufficiently large, but independent of all relevant parameters. Then for the contraction number of the multigrid* W*-cycle with distributive smoothings the inequality*

$$\|\mathcal{M}_k\|_{S_k\to S_k} \le \xi^* \ , \quad k>0 \ ,$$

holds with a constant $\xi^*<1$ independent of k and α. For the contraction number of the multigrid W*-cycle with smoothings* (6.13) *or* (6.18)*–*(6.20) *the inequality*

$$\|\mathcal{M}_k\|_{\widehat{S}_k\to \widehat{S}_k} \le \xi^* \ , \quad k>0$$

holds with a constant $\xi^*<1$ independent of k and α.

Proof. Consider the W-cycle with distributive smoothings. Define $\xi_k := \|\mathcal{M}_k\|_{S_k\to S_k}$. Using the recursion relation (7.6) for $\mathcal{M}_k$ and (7.1), (7.2) it follows that

$$\begin{aligned}\xi_k &\le \|\mathcal{T}_k\|_{S_k\to S_k} + \|\mathbf{p}_k\|_{S_k\to S_k}\|\mathcal{A}_k\mathbf{r}_k\mathcal{A}_{k-1}^{-1}\mathcal{L}^{\nu}\|_{S_k\to S_k}\|\mathcal{M}_{k-1}^{\mathrm{mgm}}\|^2_{S_k\to S_k}\\ &\le \|\mathcal{T}_k\|_{S_k\to S_k} + C\xi_{k-1}^2\end{aligned}$$

with a positive constant $C>0$ independent of all parameters. Now use the two-grid bound given in (7.7) with sufficiently large ν and a fixed-point argument. It is clear that the proof of the theorem for the case of smoothings (6.13) or (6.18)–(6.20) is literally the same with the only difference being that instead of (7.2) and (7.7) one should use (7.3) and (7.8). □

Remark 10. We briefly remark on the implications of Theorem 7.2 in two limit cases: $\alpha=0$ and $\alpha\to\infty$. For $\alpha=0$ (4.10) yields that both S_k and $\widehat{\mathrm{S}}_k$ are spectrally equivalent to the identity matrix; therefore, we recover the standard h-independent convergence bound for the multigrid W-cycle in the weighted ℓ^2-norm similar to the one in [33]. In the other limit case, $\alpha=\infty$, the result of the theorem can be interpret as the spectral equivalence of $\mathrm{B}_k\mathrm{M}_u^{-1}\mathrm{B}_k^T$ (the mixed approximation of the pressure Poisson problem) and the conforming approximation of the Laplacian operator for finite elements satisfying weak infsup condition (2.11).

8. Numerical example

Numerical results demonstrating the efficiency and robustness with respect to h and α of the multigrid method can be found in [6] for smoothing iterations (6.13) and [19] for an inexact variant of these smoothings. Below we include some numerical results which illustrate the robustness of the multigrid method with distributive and inexact Uzawa smoothings. It is not the intention of this paper to compare systematically the performance of different solvers for (4.4) including multigrids and preconditioned Krylov subspace methods. Such comparative studies can be found, e.g., in [14, 19].

TABLE 8.1. The number of iterations for V-cycle/W-cycle with distributive and inexact Uzawa smoother

mesh size h	parameter α					
	0	10^2	10^4	10^6	10^8	10^{10}
	distributive smoother					
1/32	46 / 43	42 / 41	43 / 35	77 / 72	79 / 74	79 / 74
1/64	47 / 42	45 / 42	29 / 27	67 / 63	74 / 68	74 / 68
1/128	47 / 42	45 / 43	34 / 34	60 / 54	73 / 68	73 / 68
	inexact Uzawa smoother					
1/32	19 / 19	19 / 19	13 / 13	13 / 13	13 / 12	13 / 12
1/64	18 / 18	18 / 18	16 / 16	13 / 12	13 / 12	13 / 12
1/128	17 / 17	17 / 17	16 / 16	12 / 12	12 / 11	12 / 11

We consider the generalized Stokes problem as in (1.1) on the domain $\Omega = (0,1)^2$. The right-hand side $\mathbf{f}$ and g are such that the continuous solution is

$$\mathbf{u} = \begin{pmatrix} 4(2y-1)(1-x)x \\ -4(2x-1)(1-y)y \end{pmatrix}, \quad p = (x^3+y^3)+C \tag{8.1}$$

with a constant C such that $\int_\Omega p\, d\mathbf{x} = 0$. For the discretization we used isoP_2-P_1 finite elements on a uniform west-north triangulation.

To define the distributive smoother we set in (6.7) $D_k = 2\mathrm{diag}(A_k)$ and N_k^{-1} is defined through the one V(2,2)-cycle of the inner MG method with damped Jacobi smoothings applied to the conforming P_1 discretization of the pressure Poisson equation: for a given $r_k \in \mathbb{Q}_h$ find $p_k \in \mathbb{Q}_k$ from

$$(\nabla p_k, \nabla q_k) = (r_k, q_k) \qquad \forall\ q_k \in \mathbb{Q}_k. \tag{8.2}$$

Note that for any $p_k \in \mathbb{Q}_k$ and corresponding coefficients vector $\mathrm{p} \in \mathbb{R}^m$ it holds that

(8.3)

$$\langle \mathrm{B}_k \mathrm{M}_u^{-1} \mathrm{B}_k^T \mathrm{p}, \mathrm{p}\rangle = \sup_{\mathbf{u}_k \in \mathbf{V}_k} \frac{(\mathrm{div}\, \mathbf{u}_k, p_k)^2}{\|\mathbf{u}_k\|^2} \quad \text{and} \quad \langle N_k \mathrm{p}, \mathrm{p}\rangle \lesssim \|\nabla p_k\|^2 \le \langle N_k \mathrm{p}, \mathrm{p}\rangle.$$

Therefore, the necessary condition (6.6) follows from the weak infsup inequality (2.11) and (8.3). The damping parameter is set as $\omega^2 = 0.8$.

For the inexact Uzawa smoothings (6.18)–(6.20) we set $D_k = \mathrm{diag}(A_k)$ and $\omega = 1.25$. This choice of ω is recommended in [19] as close to an optimal one. Let N_k be the same matrix as defined above through the one multigrid V(2,2)-cycle for solving (8.2). We set G_k from (6.17) to be $G_k = \rho N_k$ with $\rho = 0.8\omega^{-1}\|\mathrm{M}_u\|\|D_k\|^{-1}$. For the uniform triangulation it holds that

$$\begin{aligned}\|D_k\|\langle \mathrm{B}_k \mathrm{D}_k^{-1} \mathrm{B}_k^T \mathrm{p}, \mathrm{p}\rangle &= \langle \mathrm{B}_k \mathrm{I}_k \mathrm{B}_k^T \mathrm{p}, \mathrm{p}\rangle \le \|\mathrm{M}_u\|\langle \mathrm{B}_k \mathrm{M}_u^{-1} \mathrm{B}_k^T \mathrm{p}, \mathrm{p}\rangle \\ &= \|\mathrm{M}_u\| \sup_{\mathbf{u}_k \in \mathbf{V}_k} \frac{(\mathrm{div}\, \mathbf{u}_k, p_k)^2}{\|\mathbf{u}_k\|^2} \le \|\mathrm{M}_u\|\|\nabla p_k\|^2 \le \|\mathrm{M}_u\|\langle N_k \mathrm{p}, \mathrm{p}\rangle.\end{aligned}$$

We get $\omega^{-1}\mathrm{B}_k \mathrm{D}_k^{-1} \mathrm{B}_k^T \le \frac{5}{4} G_k$. Thus the upper estimate in (6.17) holds with $\beta = \frac{1}{4}$, which is the admissible value for the smoothing property from theorem 6.5. Numerical experiments show that violating the upper bound of $\frac{1}{3}$ for β in (6.17) leads to the divergence of the multigrid method, while satisfying the lower bound in (6.17)

TABLE 8.2. Dependence of the number of iterations for V(n,n)-cycle on the number of pre- and post-smoothings for distributive / inexact Uzawa smoothings

α	number of pre- and post-smoothings				
	1	2	4	8	16
0	181 / div	91 / 18	47 / 8	25 / 6	17 / 5
10^2	173 / div	87 / 18	45 / 8	24 / 6	16 / 4
10^4	169 / div	62 / 15	29 / 7	15 / 4	8 / 2

TABLE 8.3. The number of iterations for BiCGstab method with multigrid V-cycle as a preconditioner: distributive / inexact Uzawa smoothings

mesh size h	parameter α					
	0	10^2	10^4	10^6	10^8	10^{10}
1/32	10 / 7	10 / 7	7 / 4	11 / 4	11 / 4	11 / 4
1/64	11 / 7	10 / 7	8 / 5	11 / 4	13 / 4	13 / 4
1/128	11 / 7	11 / 7	9 / 6	10 / 4	12 / 4	12 / 4

is less crucial. This is consistent with observations in [37], where the phenomena is explained heuristically (cf. [37], Remark 5). However, the strong violation of the lower bound may require more smoothing steps to make the iterations converge. Note that computational complexity of both distributive and inexact Uzawa smoothers defined above scales linearly with the number of unknowns.

In all numerical tests we stop the iteration once the ℓ_2-norm of the initial residual has been reduced by at least nine orders of magnitude, and we always use a vector with equally distributed on $[0, 1]$ random entries as the initial guess. In Table 8.1 we show the number of iterations for various values of mesh size and α. Here we use a multigrid V(4,4) and W(4,4) cycles in the case of distributive smoothing iterations and V(2,2) and W(2,2) cycles in the case of inexact Uzawa smoothing iterations. Both methods are robust with respect to the variation of parameters. The method with inexact Uzawa smoothings shows significantly better results in terms of iteration numbers (each iteration is also computationally cheaper in this case). Although our analysis proves the convergence only for the W-cycle, numerical experiments show similar convergence results for V-cycles.

The dependence of iteration number on the number of pre- and post-smoothing steps is shown in Table 8.2. Similar to [6] we note that making only 1 post- and 1 pre-smoothing steps is not enough for convergence. Finally, Table 8.3 demonstrates that using the multigrid methods as preconditioners in a Krylov subspace method may reduce the number of iterations significantly compared to the case when multigrid methods are used as stand-alone solvers. To produce these results we use one V(4,4)-cycle with distributive smoothings or one V(2,2)-cycle with inexact Uzawa smoothings to define a preconditioner for the BiCGstab method to solve (4.4). One iteration of the BiCGstab method from Table 8.3 is approximately

TABLE 8.4. The number of iterations for V-cycle with distributive and inexact Uzawa smoother for isoP2-P0 elements

mesh	parameter α						parameter α					
size h	0	10^2	10^4	10^6	10^8	10^{10}	0	10^2	10^4	10^6	10^8	10^{10}
	distributive smoother						inexact Uzawa smoother					
1/32	82	67	13	28	29	29	11	10	5	7	7	7
1/64	85	83	29	28	29	29	11	10	6	7	7	7
1/128	86	92	71	21	29	29	11	11	8	6	7	7
1/256	87	94	163	34	28	29	13	14	10	5	7	7

twice as expensive as one iteration of the multigrid method as a stand-alone solver as in Table 8.1.

Finally, we check numerically if the uniform convergence of the multigrid methods is still observed if the assumptions $Q_h \subset H^1(\Omega)$ and (2.11) are violated. To this end, we repeat the same experiments with isoP2-P0 finite elements. We note that this stable Stokes element is not stable in the Darcy limit ($\alpha \to \infty$) with respect to the $H^1_0(\mathrm{div}) \times L^2_0$-norm [20]. This poses the challenge of building a preconditioner for the Schur complement matrices of the limit problem: $\mathrm{S}_k = \mathrm{B}_k \mathrm{M}_u^{-1} \mathrm{B}_k^T$ or $\widehat{\mathrm{S}}_k = \mathrm{B}_k \mathrm{D}_k^{-1} \mathrm{B}_k^T$; see (6.6) and (6.17). A standard multigrid approach may fail to provide an h-scalable preconditioner and more elaborated methods should be used; see [25] and [36] for such developments. In our experiments we used the preconditioner built as one W(4,4) cycle applied to solve a system with $\widehat{\mathrm{S}}_k$, which is a sparse matrix. The multigrid uses SOR iteration as smoother with a tuned relaxation parameter ($\kappa = 1.8$) and the Galerkin coarse grid matrices: $\mathrm{S}_{k-1} := \mathbf{r}_k \mathrm{S}_k \mathbf{p}_k$. In general, the latter definition leads to the increase of the fill-in pattern on coarser grid levels. However, the construction of S_{k-1} directly from the coarse grid finite element matrices B_{k-1} and D_{k-1} resulted in our experiments in a non-convergent multigrid method. Such an expensive preconditioner can be seen as a price paid for the lack of the element stability in the Darcy limit. In Table 8.4 we show the number of iterations for various values of mesh size and α. Here we use a multigrid V(4,4) cycle both in the case of distributive smoothing iterations and inexact Uzawa smoothing iterations (less smoothing iterations with inexact Uzawa was not enough for multigrid convergence). Results with W-cycle showed the same trends with respect to h and α. The parameter ω in both cases was set equal to 1. The results indicate that in spite of h-independent convergence results in the limit cases of $\alpha = 0$ and $\alpha = \infty$ the multigrid method with distributive smoother seems not to ensure robust convergence for the whole range of parameters. This and the higher cost of the method itself may result from the fact that isoP2-P0 is not a stable Stokes-Darcy element.

REFERENCES

[1] D.N. Arnold, F. Brezzi, and M. Fortin, *A stable finite element method for the Stokes equations*, Calcolo, 21 (1984), pp. 337-344. MR799997 (86m:65136)

[2] R. Bank, B.D. Welfert, and H. Yserentant, *A class of iterative methods for solving saddle point problems*, Numer. Math., 56 (1990), pp. 645-666 MR1031439 (91b:65035)

[3] M. Benzi, G.H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005), pp. 1–137. MR2168342 (2006m:65059)

[4] M. Bercovier and O. Pironneau, *Error estimates for finite element method solution of the Stokes problem in primitive variables* Numer. Math., 33 (1979), pp. 211–224. MR549450 (81g:65145)

[5] D. Braess and W. Dahmen, *A cascadic multigrid algorithm for the Stokes equations*, Numer. Math., 82 (1999), pp. 179–191. MR1685458 (2000c:65114)

[6] D. Braess and R. Sarazin, *An efficient smoother for the Stokes problem*, Applied Numerical Mathematics, 23 (1997), pp. 3–19. MR1438078 (97k:65277)

[7] J.H. Bramble and J.E. Pasciak, *Iterative techniques for time dependent Stokes problems,* Comput. Math. Appl., 33 (1997), pp. 13–30. MR1442058 (98e:65091)

[8] S. Brenner, *A nonconforming multigrid method for the stationary Stokes problem,* Math. Comp., 55 (1990), pp. 411–473. MR1035927 (91d:65167)

[9] J.H. Bramble, J.E. Pasciak and O. Steinbach, *On the stability of the L^2 projection in $H^1(\Omega)$*, Math. Comp., 71 (2001), pp. 147–156. MR1862992 (2002h:65175)

[10] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, New York, Springer-Verlag, 1991. MR1115205 (92d:65187)

[11] P.G. Ciarlet, *Basic Error Estimates for Elliptic Problems.* Handbook of Numerical Analysis, Vol. II, P.G. Ciarlet and J.L. Lions, eds., Elsevier Science, 1991, pp. 17–351. MR1115237

[12] M. Dauge, *Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. Part I. Linearized Equations*, SIAM J. Math. Anal., 20 (1989), pp. 74–97 MR977489 (90b:35191)

[13] L. Durlofsky and J. F. Brady, *Analysis of the Brinkman equation as a model for flow in porous media*, Phys. Fluids, 30 (1987), pp. 3329–3342.

[14] H.C. Elman, *Multigrid and Krylov subspace methods for the discrete Stokes equations*, Int. J. Numer. Methods Fluids, 22 (1998), pp. 755 – 770.

[15] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations,* Springer, Berlin, 1986. MR851383 (88b:65129)

[16] W. Hackbusch, *Multi-grid Methods and Applications*, Berlin, Heidelberg: Springer, 1985.

[17] Yu. Iliash, R. Tuomo, and J. Toivanen, *Two iterative methods for solving the Stokes problem*, Laboratory of Scientific Computing, University of Jyvaskyla, Report 2, 1993.

[18] G.M. Kobelkov and M.A. Olshanskii, *Effective Preconditioning of Uzawa Type Schemes for Generalized Stokes Problem*, Numer. Math., 86 (2000), pp. 443–470. MR1785417 (2001j:65168)

[19] M. Larin and A. Reusken, *A comparative study of efficient iterative solvers for generalized Stokes problem*, Numer. Linear Algebra Appl., 15 (2008), pp. 13–34. MR2388412 (2009a:76094)

[20] K.-A. Mardal, X.-C. Tai, and R. Winther, *A robust finite element method for Darcy-Stokes flow*, SIAM J. Num. Anal. 40 (2002), 1605-1631. MR1950614 (2003m:76110)

[21] S. Manservisi, *Numerical Analysis of Vanka-Type Solvers for Steady Stokes and Navier-Stokes Flows*, SIAM J. Numer. Anal. 44 (2006), pp. 2025–2056. MR2263039 (2007k:76036)

[22] K.-A., Mardal and R. Winther, *Uniform preconditioners for the time dependent Stokes problem*, Numer. Math. 98 (2004), pp. 305–327. MR2092744 (2005g:65063)

[23] M. A. Olshanskii and A. Reusken, *On the convergence of a multigrid method for linear reaction-diffusion problems*, Computing, 65 (2000), pp. 193–202 MR1810324 (2001m:65186)

[24] M. A. Olshanskii, J. Peters, and A. Reusken, *Uniform preconditioners for a parameter dependent saddle point problem with application to generalized Stokes interface equations*, Numer. Math., 105 (2006), pp. 159–191. MR2257388 (2007i:76026)

[25] T. Rusten, P.S. Vassilevski, and R. Winther, *Interior penalty preconditioners for mixed finite element approximations of elliptic problems*, Math. Comp., 65 (1996), pp. 447–466. MR1333325 (96j:65127)

[26] V.V. Shaidurov, *Multigrid methods for finite elements,* Moscow: Nauka, 1989. (in Russian) MR1007306 (91h:65193)

[27] J. Schoeberl and W. Zulehner, *On Schwarz-type smoothers for saddle point problems*, Numer. Math., 95 (2002), pp. 377–399. MR2001083 (2004g:65149)

[28] D. Silvester and A. Wathen, *Fast iterative solution of stabilised stokes systems. Part II: Using general block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367. MR1293519 (95g:65132)

[29] R. Stevenson, *Nonconforming finite elements and the cascadic multi-grid method*, Numer. Math., 91 (2002), pp. 351–387. MR1900924 (2003c:65139)

[30] Z. Tong and Z. Sameh, *On an iterative method for saddle point problems*, Numer. Math., 79 (1998), pp. 643-646. MR1635705 (99f:65054)
[31] S. Turek, *Efficient Solvers for Incompressible Flow Problems. An Algorithmic and Computational Approach*, Lecture Notes in Computational Science and Engineering, Vol. 6, Springer-Verlag, Berlin/Heidelberg/New York, 1999. MR1691839 (2000g:76071)
[32] S.P. Vanka, *Block-implicit multigrid solution of Navier-Stokes equations in primitive variables*, J. Comput. Phys., 65 (1986), pp. 138-158. MR848451 (87g:76009)
[33] R. Verfürth, *A multilevel algorithm for mixed problems*, SIAM J. Num. Anal., 21 (1984), pp. 264–271. MR736330 (85f:65112)
[34] P. Wesseling and C.W. Oosterlee, *Geometric multigrid with applications to computational fluid dynamics*, J. Comput. Appl. Math., 128 (2001) pp. 311–334. MR1820879 (2002b:65187)
[35] X. Xie, J. Xu, G. Xue, *Uniformly-stable finite element methods for Darcy-Stokes-Brinkman models*, J. Comp. Math., 26 (2008), pp. 437–455. MR2421892 (2009g:76087)
[36] J. Xu, *The auxiliary space method and optimal multigrid preconditioning techniques for unstructure grids*, Computing, 56 (1996), pp. 215–235. MR1393008 (97d:65062)
[37] W. Zulehner, *A class of smoothers for saddle point problems*, Computing, 65 (2000), pp. 227–246. MR1810326 (2001j:65160)

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