A comparative study of efficient iterative solvers for generalized Stokes equations

Maxim Larin[*,†] and Arnold Reusken

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

SUMMARY

We consider a generalized Stokes equation with problem parameters $\xi \geq 0$ (size of the reaction term) and $\nu > 0$ (size of the diffusion term). We apply a standard finite element method for discretization. The main topic of the paper is a study of efficient iterative solvers for the resulting discrete saddle point problem. We investigate a coupled multigrid method with Braess-Sarazin and Vanka type smoothers, a preconditioned MINRES method and an inexact Uzawa method. We present a comparative study of these methods. An important issue is the dependence of the rate of convergence of these methods on the mesh size parameter and on the problem parameters ξ and ν. We give an overview of the main theoretical convergence results known for these methods. For a three dimensional problem, discretized by the Hood-Taylor $\mathcal{P}_2 - \mathcal{P}_1$ pair, we give results of numerical experiments.

KEY WORDS generalized Stokes problem, preconditioned MINRES, inexact Uzawa method, multigrid methods, Vanka and Braess-Sarazin smoothers.

1. Introduction

Let $\Omega \subset \mathbb{R}^3$ be a bounded polygonal domain with a Lipschitz boundary $\Gamma = \partial\Omega$. We consider the following generalized Stokes problem:

Given $\vec{f}$, find a velocity $\vec{u}$ and a pressure p such that

$$\begin{aligned} \xi\vec{u} - \nu\Delta\vec{u} + \nabla p &= \vec{f} \quad \text{in} \quad \Omega, \\ \nabla \cdot \vec{u} &= 0 \quad \text{in} \quad \Omega, \\ \vec{u} &= 0 \quad \text{on} \quad \Gamma. \end{aligned} \tag{1}$$

*Correspondence to: Maxim Larin, Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, 52056 Aachen, Germany.
†E-mail: larin@igpm.rwth-aachen.de .

Contract/grant sponsor: SFB; contract/grant number:

The parameters $\nu > 0$ (viscosity) and $\xi \geq 0$ are given. Often the latter is proportional to the inverse of the time step in an implicit time integration method applied to a nonstationary Stokes problem. Note that this general setting includes the classical (stationary) Stokes problem ($\xi = 0$). This problem is discretized on a tetrahedral grid with a pair of conforming finite element spaces that is inf-sup stable. In our experiments we use the $\mathcal{P}_2 - \mathcal{P}_1$ Hood-Taylor pair. The resulting discrete problem is of saddle-point type with a symmetric indefinite matrix. In this paper we study efficient iterative solvers for this linear system. In particular the efficiency (robustness) of the solvers with respect to variation in the mesh size parameter and in the problem parameters ξ and ν is studied. We consider a multigrid method with Vanka and Braess-Sarazin smoothers, a preconditioned MINRES method and an inexact Uzawa method. In the latter two methods multigrid preconditioners are used for the scalar problems for each velocity component. A comparative study of the preconditioned MINRES and inexact UZAWA method with other preconditioned Krylov subspace methods for this problem class is given in [18]. Numerical studies on the performance of coupled multigrid problems for (generalized) Stokes and Navier-Stokes can be found in, for example, [9, 10, 11, 12, 21]. We do not know of any literature in which a systematic comparison between coupled multigrid, preconditioned MINRES and inexact Uzawa type of methods for this class of generalized Stokes equations is given. In this paper we present such a comparative study. We give an overview of the main theoretical results that are available for these methods. From this it follows that concerning theoretical convergence results the state of affairs is much better for the preconditioned MINRES and the inexact Uzawa method than for the coupled multigrid method. We pay special attention to the case $\xi = 0$, $\nu > 0$ variable. In this case variation of the parameter ν corresponds to a rescaling of the velocity unknowns. We show that for all methods considered here the rate of convergence is essentially independent of this rescaling.

We also investigate the efficiency of the different methods by means of numerical experiments. It turns out that all methods show good robustness properties with respect to variation in the mesh size and in the parameters ξ and ν.

The paper is organized as follows. In section 2 the weak formulation and the finite element discretization are given. In section 3 the coupled multigrid method with Braess-Sarazin and Vanka smoothers is desribed. For the method with Braess-Sarazin smoother a convergence analysis known from the literature is presented in a different form. In section 4 we discuss the preconditioned MINRES and inexact Uzawa methods. We recall known convergence results for these methods. A comparison of all these methods from a theoritical point of view is given in section 5. In the final section 6 a numerical study of these methods is presented and conclusions are drawn.

2. Weak formulation and finite element discretization

The weak formulation of (1) is as follows:

$$\begin{aligned} &\textit{Given } \vec{f} \in L^2(\Omega)^3, \textit{ we seek } \vec{u} \in H_0^1(\Omega)^3 \textit{ and } p \in L_0^2(\Omega) \textit{ such that} \\ &\qquad \xi(\vec{u},\vec{v}) + \nu(\nabla\vec{u},\nabla\vec{v}) - (\text{div } \vec{v}, p) = (\vec{f},\vec{v}) \qquad \text{for all } \vec{v} \in H_0^1(\Omega)^3, \\ &\qquad\qquad (\text{div } \vec{u}, q) = 0 \qquad \text{for all } q \in L_0^2(\Omega). \end{aligned} \tag{2}$$

Here $(\cdot,\cdot)$ denotes the L^2 scalar product.

For descretization of (2) we use a standard finite element approach. Based on a quasi-uniform family of nested tetrahedral grids $\mathcal{T}_0 \subset \mathcal{T}_1 \subset \ldots$ we use a sequence of nested finite element spaces

$$(\mathbf{V}_{l-1}, Q_{l-1}) \subset (\mathbf{V}_l, Q_l), \quad l = 1, 2, \ldots .$$

The pair of spaces $(\mathbf{V}_l, Q_l)$, $l \geq 0$, is assumed to be stable. By h_l we denote the mesh size parameter corresponding to $\mathcal{T}_l$. We assume that h_{l-1}/h_l is uniformly bounded in l. For the theoretical analysis we assume that the pair has the following approximation property:

$$\inf_{\vec{v} \in \mathbf{V}_l} \|\vec{u} - \vec{v}\|_1 + \inf_{q \in Q_l} \|p - q\|_{L^2} \leq c\, h_l \big(\|\vec{u}\|_2 + \|p\|_1\big), \;\; \forall\, \vec{u} \in (H^2(\Omega) \cap H_0^1(\Omega))^3,\; p \in H^1(\Omega) \cap L_0^2(\Omega).$$

We use the notation $\|\cdot\|_k$, $k = 1, 2$, for the norms in $H^k(\Omega)$. In our numerical experiments we use the Hood-Taylor $\mathcal{P}_2 - \mathcal{P}_1$ pair. The discrete problem is given by the Galerkin discretization of (2) with the pair $(\mathbf{V}_l, Q_l)$. We are interested in the solution of this discrete problem on a given finest discretization level $l = L$. To solve this discrete problem we introduce the standard nodal bases in these finite element spaces. The representation of the discrete problem on level l in these bases results in a linear saddle point problem of the form:

$$\mathcal{A}_l \mathbf{x}_l = \mathbf{b}_l \quad \text{with} \quad \mathcal{A}_l = \begin{pmatrix} A_l & B_l^T \\ B_l & 0 \end{pmatrix}, \quad \mathbf{x}_l = \begin{pmatrix} \mathbf{u}_l \\ \mathbf{p}_l \end{pmatrix}. \tag{3}$$

The dimensions of the spaces $\mathbf{V}_l$ and Q_l are denoted by n_l and m_l, respectively. The matrix $A_l \in \mathbb{R}^{n_l \times n_l}$ is the discrete representation of the differential operator $\xi I - \nu\Delta$ and is symmetric positive definite. Note that A_l depends on the parameters ξ and ν. The matrix $\mathcal{A}_l$ depends on these parameters, too, and is symmetric and strongly indefinite. In the remainder of the paper we consider iterative solvers for the system (3) on the finest level L.

Remark 1. The matrix $\mathcal{A}_l$ is singular. Below we always consider $\mathcal{A}_l$ on the subspace $\mathbb{R}^{n_l} \times 1^{\perp}$, where $1^{\perp}$ is the subspace of $\mathbb{R}^{m_l}$ of vectors for which the corresponding finite element functions $q_l \in Q_l$ satisfy $\int_\Omega q_l\, dx = 0$. The mapping $\mathcal{A}_l : \mathbb{R}^{n_l} \times 1^{\perp} \to \mathbb{R}^{n_l} \times 1^{\perp}$ is invertible.

3. A coupled multigrid method

We consider a multigrid method for the coupled system in (3). Below we discuss the components of this multigrid solver.

The grid transfer operations. For the prolongation and restriction of vectors (or corresponding finite element functions) between different level we use the canonical operators. The prolongation between level $l-1$ and l is given by

$$P_l = \begin{pmatrix} P_{\mathbf{V}} & 0 \\ 0 & P_Q \end{pmatrix},$$

where the matrices $P_{\mathbf{V}} : \mathbb{R}^{n_{l-1}} \to \mathbb{R}^{n_l}$ and $P_Q : \mathbb{R}^{m_{l-1}} \to \mathbb{R}^{m_l}$ are matrix representations of the embeddings $\mathbf{V}_{l-1} \subset \mathbf{V}_l$ (quadratic interpolation for $\mathcal{P}_2$) and $Q_{l-1} \subset Q_l$ (linear interpolation for $\mathcal{P}_1$), respectively. For the restriction operator R_l between the levels l and $l-1$ we take the adjoint of P_l (w.r.t. a scaled Euclidean scalar product).

Coarse grid operators. In the multigrid solver for the problem on the finest level L we need operators on the coarser levels $L-1,\ldots,0$. We use the matrices $\mathcal{A}_l$ in (3). These result from the finite element discretization method applied on level l. For these matrices the Galerkin relation $\mathcal{A}_{l-1} = R_l \mathcal{A}_l P_l$ holds.

The smoothers. In this paper we consider two popular smoothers for Stokes type of problems, namely the Braess-Sarazin smoother and a Vanka-type smoother. These smoothers and their properties are discussed in the following two subsections.

3.1. Braess-Sarazin smoother

This smoother is introduced in [4]. With $D_l = \mathrm{diag}(A_l)$ and a given $\alpha > 0$ the smoothing iteration has the form

$$\begin{pmatrix} \mathbf{u}_l^{(j+1)} \\ \mathbf{p}_l^{(j+1)} \end{pmatrix} = \begin{pmatrix} \mathbf{u}_l^{(j)} \\ \mathbf{p}_l^{(j)} \end{pmatrix} - \begin{pmatrix} \alpha D_l & B_l^T \\ B_l & 0 \end{pmatrix}^{-1} \left\{ \begin{pmatrix} A_l & B_l^T \\ B_l & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}_l^{(j)} \\ \mathbf{p}_l^{(j)} \end{pmatrix} - \begin{pmatrix} \mathbf{f}_l \\ 0 \end{pmatrix} \right\}. \tag{4}$$

Each iteration (4) requires the solution of the auxiliary problem

$$\begin{pmatrix} \alpha D_l & B_l^T \\ B_l & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{u}}_l \\ \hat{\mathbf{p}}_l \end{pmatrix} = \begin{pmatrix} \mathbf{r}_l^{(j)} \\ B_l \mathbf{u}_l^{(j)} \end{pmatrix} \tag{5}$$

with $\mathbf{r}_l^{(j)} = A_l \mathbf{u}_l^{(j)} + B_l^T \mathbf{p}_l^{(j)} - \mathbf{f}_l$. From (5) one obtains

$$B_l \hat{\mathbf{u}}_l = B_l \mathbf{u}_l^{(j)},$$

and hence,

$$B_l \mathbf{u}_l^{(j+1)} = B_l(\mathbf{u}_l^{(j)} - \hat{\mathbf{u}}_l) = 0 \quad \text{for all} \quad j \geq 0. \tag{6}$$

Therefore, the Braess-Sarazin method can be considered as a smoother on the subspace of vectors that satisfy the constraint equation $B_l \mathbf{u}_l = 0$.

The problem (5) can be reduced to a problem for the auxiliary pressure unknown $\hat{\mathbf{p}}_l$:

$$Z_l \hat{\mathbf{p}}_l = B_l D_l^{-1} \mathbf{r}_l^{(j)} - \alpha B \mathbf{u}_l^{(j)}, \tag{7}$$

where $Z_l = B_l D_l^{-1} B_l^T$.

Remark 2. The matrix Z_l is a sort of discrete Laplace operator on the pressure space. In practice the system (7) is solved approximately using an iterative solver.

Once $\hat{\mathbf{p}}_l$ is known (approximately), an approximation for $\hat{\mathbf{u}}_l$ can easily be determined from $\alpha D_l \hat{\mathbf{u}}_l = \mathbf{r}_l^{(j)} - B_l^T \hat{\mathbf{p}}_l$. The iteration matrix of the smoother (4) is denoted by $\mathcal{S}_l$. For a two-grid method, with ν_1 pre- and ν_2 post-smoothing steps, applied to (3) the iteration matrix is given by

$$\mathcal{M}_l = \mathcal{S}_l^{\nu_2} (I - P_l \mathcal{A}_{l-1}^{-1} R_l \mathcal{A}_l) \mathcal{S}_l^{\nu_1} \tag{8}$$

We derive a convergence result for this multigrid method with the Braess-Sarazin smoother. A multigrid convergence analysis is given in [4], however, only for the case where D_l is replaced by the identity matrix. In that paper the reduction of the velocity error $\mathbf{u}_l^{(j)} - \mathbf{u}_l$ in the subspace of vectors that satisfy the constraint equation $B_l \mathbf{u}_l = 0$ is analyzed. The pressure error does not

play a role. In the present paper, due to the parameters contained in the left-upper A_l-block we are interested in the dependence of the multigrid convergence behaviour on the scaling of the A_l-block. This dependence can be analysed very easily (Lemma 3 below) if we present the analysis from [4] in a different form in which both the errors in the velocity and pressure components are taken into account. In this modified analysis we consider the smoother as in (4) and not the one (as in [4]) in which D_l is replaced by the identity. The analysis is formulated in terms of a smoothing- and approximation property. We use the following norms. By $\|\cdot\|$ we denote the Euclidean norm on $\mathbb{R}^k$. On $\mathbb{R}^{n_l+m_l}$ we also use the following norm:

$$\left\| \begin{pmatrix} \mathbf{u}_l \\ \mathbf{p}_l \end{pmatrix} \right\|_h^2 := \|\mathbf{u}_l\|^2 + h_l^2 \|\mathbf{p}_l\|^2 = \|\Lambda_l \begin{pmatrix} \mathbf{u}_l \\ \mathbf{p}_l \end{pmatrix}\|^2 \quad \text{with} \quad \Lambda_l := \begin{pmatrix} I_{n_l} & 0 \\ 0 & h_l I_{m_l} \end{pmatrix}.$$

Corresponding matrix norms are denoted by $\|\cdot\|$, $\|\cdot\|_h$, too. For the Braess-Sarazin smoother we have the following result.

Lemma 1. *For the method* (4)*, with iteration matrix* $\mathcal{S}_l$*, the following holds:*

$$\mathcal{A}_l \mathcal{S}_l^{\nu_1} = \begin{pmatrix} I_{n_l} & 0 \\ 0 & 0 \end{pmatrix} \mathcal{A}_l \mathcal{S}_l^{\nu_1} \tag{9}$$

$$\|\mathcal{A}_l \mathcal{S}_l^{\nu_1}\|_h = \|\mathcal{A}_l \mathcal{S}_l^{\nu_1}\| \leq \frac{\alpha}{e(\nu_1 - 2) + 1} \|D_l\| \quad \textit{if} \;\; \alpha \geq \lambda_{\max}(D_l^{-1} A_l),\, \nu_1 \geq 2. \tag{10}$$

Proof: The result in (9) follows from $B_l \mathbf{u}_l = 0$ and (6). Introduce $\tilde{A}_l = D_l^{-1/2} A_l D_l^{-1/2}$, $\tilde{B}_l = B_l D_l^{-1/2}$. A simple computation shows that for the iteration matrix we have

$$\mathcal{S}_l = \begin{pmatrix} D_l^{-\frac{1}{2}} & 0 \\ 0 & I_{m_l} \end{pmatrix} \begin{pmatrix} (I_{n_l} - \tilde{B}_l^T (\tilde{B}_l \tilde{B}_l^T)^{-1} \tilde{B}_l)(I_{n_l} - \alpha^{-1} \tilde{A}_l) & 0 \\ (\tilde{B}_l \tilde{B}_l^T)^{-1} \tilde{B}_l (\alpha I_{n_l} - \tilde{A}_l) & 0 \end{pmatrix} \begin{pmatrix} D_l^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix}.$$

The operator

$$T_l := I_{n_l} - \tilde{B}_l^T (\tilde{B}_l \tilde{B}_l^T)^{-1} \tilde{B}_l$$

is an orthogonal projector on $\mathrm{Kern}(\tilde{B}_l)$, thus $\tilde{B}_l T_l = 0$. With $M_l := T_l(I_{n_l} - \alpha^{-1}\tilde{A}_l)T_l$, we obtain, for $\nu_1 \geq 2$,

$$\begin{aligned} \mathcal{S}_l^{\nu_1} &= \mathcal{S}_l^{\nu_1 - 1} \mathcal{S}_l \\ &= \begin{pmatrix} D_l^{-\frac{1}{2}} & 0 \\ 0 & I_{m_l} \end{pmatrix} \begin{pmatrix} M_l^{\nu_1 - 1} & 0 \\ \alpha (\tilde{B}_l \tilde{B}_l^T)^{-1} \tilde{B}_l (I_{n_l} - \alpha^{-1} \tilde{A}_l) M_l^{\nu_1 - 2} & 0 \end{pmatrix} \begin{pmatrix} T_l (I_{n_l} - \alpha^{-1} \tilde{A}_l) D_l^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned} \tag{11}$$

Note that

$$\begin{pmatrix} D_l^{-\frac{1}{2}} & 0 \\ 0 & I_{m_l} \end{pmatrix} \mathcal{A}_l \begin{pmatrix} D_l^{-\frac{1}{2}} & 0 \\ 0 & I_{m_l} \end{pmatrix} = \begin{pmatrix} \tilde{A}_l & \tilde{B}_l^T \\ \tilde{B}_l & 0 \end{pmatrix}.$$

Combined with (11) this yields

$$\begin{aligned} &\begin{pmatrix} D_l^{-\frac{1}{2}} & 0 \\ 0 & I_{m_l} \end{pmatrix} \mathcal{A}_l \mathcal{S}_l^{\nu_1} \\ &\quad = \begin{pmatrix} \tilde{A}_l M_l^{\nu_1 - 1} + \alpha (I_{n_l} - T_l)(I_{n_l} - \alpha^{-1} \tilde{A}_l) T_l M_l^{\nu_1 - 2} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} T_l (I_{n_l} - \alpha^{-1} \tilde{A}_l) D_l^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned} \tag{12}$$

¿From this and (9) it follows that $\Lambda_l \mathcal{A}_l \mathcal{S}_l^{\nu_1} \Lambda_l^{-1} = \mathcal{A}_l \mathcal{S}_l^{\nu_1}$ and thus the identity in (10) holds. ¿¿From (12) we obtain for $\alpha \geq \lambda_{\max}(D_l^{-1} A_l) = \lambda_{\max}(\tilde{A}_l)$ and $\nu_1 \geq 2$,

$$\begin{aligned}
\|\mathcal{A}_l \mathcal{S}_l^{\nu_1}\| &\leq \|D_l^{\frac{1}{2}}\| \left\| \begin{pmatrix} D_l^{-\frac{1}{2}} & 0 \\ 0 & I_{m_l} \end{pmatrix} \mathcal{A}_l \mathcal{S}_l^{\nu_1} \right\| \\
&\leq \|D_l^{\frac{1}{2}}\| \|\tilde{A}_l M_l^{\nu_1 - 1} + \alpha (I_{n_l} - T_l)(I_{n_l} - \alpha^{-1} \tilde{A}_l) T_l M_l^{\nu_1 - 2}\| \|T_l\| \|I_{n_l} - \alpha^{-1} \tilde{A}_l\| \|D_l^{\frac{1}{2}}\| \\
&\leq \|\tilde{A}_l M_l^{\nu_1 - 1} + \alpha (I_{n_l} - T_l)(I_{n_l} - \alpha^{-1} \tilde{A}_l) T_l M_l^{\nu_1 - 2}\| \|D_l\|. \qquad (13)
\end{aligned}$$

Note that M_l is symmetric positive definite with $\sigma(M_l) \subset [0, 1]$. Using this and $T_l \tilde{A}_l T_l = \alpha(T_l - M_l)$ we obtain

$$\begin{aligned}
\|\tilde{A}_l M_l^{\nu_1 - 1} + \alpha (I_{n_l} - T_l)(I_{n_l} - \alpha^{-1} \tilde{A}_l) T_l M_l^{\nu_1 - 2}\| &= \left\| \left(\tilde{A}_l M_l - \tilde{A}_l T_l + \alpha (T_l - M_l) \right) M_l^{\nu_1 - 2} \right\| \\
&= \|(\tilde{A}_l - \alpha I_{n_l})(M_l - T_l) M_l^{\nu_1 - 2}\| \\
&\leq \|\tilde{A}_l - \alpha I_{n_l}\| \|(M_l - T_l) M_l^{\nu_1 - 2}\| \\
&= \|\tilde{A}_l - \alpha I_{n_l}\| \|(M_l - I_{n_l}) M_l^{\nu_1 - 2}\| \\
&\leq \frac{\alpha}{e(\nu_1 - 2) + 1}.
\end{aligned}$$

Using this in (13) we get the inequality in (10). ∎

We now consider the approximation property.

Lemma 2. *Take $\xi = 0$ and $\nu = 1$ in (2). Assume that Ω is such that the problem (2) is H^2-regular. Then there exists a constant C_A independent of l such that*

$$\left\| \left(\mathcal{A}_l^{-1} - P_l \mathcal{A}_{l-1}^{-1} R_l \right) \begin{pmatrix} I_{n_l} & 0 \\ 0 & 0 \end{pmatrix} \right\|_h \leq C_A \|D_l\|^{-1} \qquad (14)$$

holds.

Proof: Let $\{\phi_i\}_{1 \leq i \leq n_l}$, $\{\tilde{\phi}_i\}_{1 \leq i \leq m_l}$ be the nodal bases in $\mathbf{V}_l$ and Q_l and

$$F_l \mathbf{u} := \sum_{i=1}^{n_l} u_i \phi_i \qquad \text{and} \qquad \tilde{F}_l \mathbf{p} := \sum_{i=1}^{m_l} p_i \tilde{\phi}_i$$

the finite element isomorphisms $\mathbb{R}^{n_l} \to \mathbf{V}_l$ and $\mathbb{R}^{m_l} \to Q_l$, respectively. On $\mathbb{R}^{n_l}$ we use a scaled Euclidean inner product $\langle \mathbf{v}, \mathbf{u} \rangle_l = h_l^3 \sum_{i=1}^{n_l} v_i u_i$, and similarly on Q_l. The norms $\|\mathbf{u}\|_l$ ($\|\mathbf{p}\|_l$) and $\|F_l \mathbf{u}\|_{L^2}$ ($\|\tilde{F}_l \mathbf{p}\|_{L^2}$) are uniformly (w.r.t. $\mathbf{u}$, $\mathbf{p}$ and l) equivalent. Let $\mathcal{A}_l$ (and $\mathcal{A}_{l-1}$) be scaled such that

$$\langle A_l \mathbf{u}, \mathbf{v} \rangle_l = (\nabla F_l \mathbf{u}, \nabla F_l \mathbf{v}), \ \langle B_l \mathbf{u}, \mathbf{p} \rangle_l = (\mathrm{div} F_l \mathbf{u}, \tilde{F}_l \mathbf{p}) \quad \text{for all } \ \mathbf{u}, \mathbf{v} \in \mathbb{R}^{n_l}, \, \mathbf{p} \in \mathbb{R}^{m_l} \qquad (15)$$

For $f_l \in \mathbf{V}_l$ let $\vec{u} \in H_0^1(\Omega)^3$, $p \in L_0^2(\Omega)$ be the solution of

$$\begin{aligned}
(\nabla \vec{u}, \nabla \vec{v}) - (\mathrm{div}\ \vec{v}, p) &= (f_l, \vec{v}) \qquad \text{for all } \vec{v} \in H_0^1(\Omega)^3, \\
(\mathrm{div}\ \vec{u}, q) &= 0 \qquad \text{for all } q \in L_0^2(\Omega).
\end{aligned} \qquad (16)$$

Let $(\vec{u}_l, p_l)$ be the Galerkin solution of this problem in the pair of spaces $(\mathbf{V}_l, Q_l)$. The matrix $\mathcal{A}_l$ is the matrix representation of the finite element discretization of the problem (16), cf.

(15). Using this, the approximation property of the spaces $(\mathbf{V}_l, Q_l)$ and standard finite element techniques (duality argument) we obtain

$$\left\| \left(\mathcal{A}_l^{-1} - P_l \mathcal{A}_{l-1}^{-1} R_l\right) \begin{pmatrix} I_{n_l} & 0 \\ 0 & 0 \end{pmatrix} \right\|_h \leq c \sup_{f_l \in \mathbf{V}_l} \frac{\|\vec{u} - \vec{u}_l\|_{L^2} + h_l \|p - p_l\|_{L^2}}{\|f_l\|_{L^2}} \leq \tilde{c} h_l^2.$$

Using the scaling of A_l as in (15) and standard properties of the finite element nodal basis we obtain $\|D_l\| \geq c h_l^{-2}$ with a constant $c > 0$. Thus we obtain the bound in (14). ∎

In the approximation property it is important to have the projection factor $\begin{pmatrix} I_{n_l} & 0 \\ 0 & 0 \end{pmatrix}$ in (14). Without this factor one has to consider the Stokes problem in (16) where in the second equation the right handside 0 is replaced by (g_l, q) with a $g_l \in Q_l$. The regularity properties of such a problem are in general less favorable as for the case with a 0 right hanside. In particular, for H^2-regularity one has to assume certain compatibility conditions on g_l that are not satisfied for all $g_l \in Q_l$, cf. [6].

Combination of the smoothing- and approximation property yields a two-grid convergence result.

Theorem 1. *Take $\xi = 0$ and $\nu = 1$ in (2). Assume that Ω is such that the problem (2) is H^2-regular. For the iteration matrix of the two-grid method with $\nu_2 = 0$ the following holds:*

$$\|\mathcal{M}_l\|_h \leq \frac{C_A}{e(\nu_1 - 2) + 1} \quad \text{for} \quad \nu_1 \geq 2,$$

with a constant C_A independent of l.

Proof: Due to (9) we have

$$\mathcal{M}_l = \left(\mathcal{A}_l^{-1} - P_l \mathcal{A}_{l-1}^{-1} R_l\right) \begin{pmatrix} I_{n_l} & 0 \\ 0 & 0 \end{pmatrix} \mathcal{A}_l \mathcal{S}_l^{\nu_1}.$$

The desired result follows from (10) and Lemma 2. ∎

In [4] such a result is proved for a two-grid method in which an additional projection step is used in the coarse-grid correction. In [3], however, it is noted that this projection step is superfluous.

Using this two-grid contraction number bound one can derive a multigrid W-cycle convergence result using techniques from [7, 8].

The numerical experiments in section 6 clearly show that the multigrid W-cycle method with Braess-Sarazin smoother is robust with respect to variation in the problem parameters ν and ξ. We do not know of any analysis in which such a robustness property is proved. An elementary scaling argument can be used to derive a robustness result for the (less interesting) case $\xi = 0$, $\nu > 0$ arbitrary. For this scaling argument we introduce the notation

$$I_{l,\delta} = \begin{pmatrix} I_{n_l} & 0 \\ 0 & \delta I_{m_l} \end{pmatrix}, \qquad \tilde{I}_{l,\delta} = \begin{pmatrix} \delta I_{n_l} & 0 \\ 0 & I_{m_l} \end{pmatrix} \tag{17}$$

with I_k the identity matrix in $\mathbb{R}^k$. For $\delta \neq 0$ let $\mathcal{M}_{l,\delta}$ be the iteration matrix of the two-grid method with ν_1 pre- and ν_2 post-smoothing iterations (4) applied to the matrix

$$\mathcal{A}_{l,\delta} := \begin{pmatrix} \delta A_l & B_l^T \\ B_l & 0 \end{pmatrix}. \tag{18}$$

Note that $\mathcal{M}_{l,1} = \mathcal{M}_l$ as in (8). The effect of the scaling on the two-grid iteration matrix is given in the following lemma.

Lemma 3. *For $\delta \neq 0$ the relation*

$$\mathcal{M}_{l,\delta} = I_{l,\delta}\, \mathcal{M}_{l,1}\, I_{l,\delta}^{-1}$$

holds.

Proof: Let $\mathcal{S}_{l,\delta}$ be the iteration matrix of the Braess-Sarazin smoother applied to $\mathcal{A}_{l,\delta}$. The following relations hold

$$\mathcal{A}_{l,\delta} = \tilde{I}_{l,\delta}\, \mathcal{A}_{l,1}\, I_{l,\delta}^{-1}, \quad \mathcal{S}_{l,\delta} = I_{l,\delta}\, \mathcal{S}_{l,1}\, I_{l,\delta}^{-1}, \quad I_{l,\delta}^{-1}\, P_l\, I_{l-1,\delta} = P_l, \quad \tilde{I}_{l-1,\delta}^{-1}\, R_l\, \tilde{I}_{l,\delta} = R_l. \tag{19}$$

Using these we get

$$\mathcal{M}_{l,\delta} = \mathcal{S}_{l,\delta}^{\nu_2}(I_l - P_l \mathcal{A}_{l-1,\delta}^{-1} R_l \mathcal{A}_{l,\delta})\mathcal{S}_{l,\delta}^{\nu_1} = I_{l,\delta}\, \mathcal{M}_{l,1}\, I_{l,\delta}^{-1}$$

and thus the result is proved. ∎

Corollary 1. Introduce the norm

$$\| \begin{pmatrix} \mathbf{u}_l \\ \mathbf{p}_l \end{pmatrix} \|_{h,\delta}^2 := \|\Lambda_l I_{l,\delta}^{-1} \begin{pmatrix} \mathbf{u}_l \\ \mathbf{p}_l \end{pmatrix} \|^2 = \|\mathbf{u}_l\|^2 + \frac{h_l^2}{\delta^2}\|\mathbf{p}_l\|^2$$

with a corresponding matrix norm denoted by $\|\cdot\|_{h,\delta}$. Then we have

$$\|\mathcal{M}_{l,\delta}\|_{h,\delta} = \|\mathcal{M}_{l,1}\|_h = \|\mathcal{M}_l\|_h,$$

and thus the convergence result in Theorem 1 immediately yields an analogous result for the multigrid method applied to the scaled system.

3.2. Vanka smoother

The Vanka-type smoothers, originally proposed by Vanka [22] for finite difference schemes, are block Gauß-Seidel type of methods. If one uses such a method in a finite element setting then a block of unknowns consists of all degrees of freedom that correspond with one element. Numerical tests given in [12] show that the use of this element-wise Vanka smoother can be problematic for continuous pressure approximations. In [12] the pressure-oriented Vanka smoother for continuous pressure approximations has been suggested as a good alternative. In this method a local problem corresponds to the block of unknowns consisting of one pressure unknown and all velocity degrees of freedom that are connected with this pressure unknown. In the present paper we consider this type of Vanka smoother. We first give a more precise description of this method.

We consider a fixed level l in the discretization. To simplify the presentation we drop the level index l from the notation, i.e. we write, for example, $\begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} \in \mathbb{R}^{n+m}$ instead of $\begin{pmatrix} \mathbf{u}_l \\ \mathbf{p}_l \end{pmatrix} \in \mathbb{R}^{n_l+m_l}$. Let $r_P^{(j)} : \mathbb{R}^m \to \mathbb{R}$ be the pressure projection (injection)

$$r_P^{(j)}\mathbf{p} = p_j, \qquad j = 1, \ldots, m.$$

For each j $(1 \leq j \leq m)$ let the set of velocity indices that are "connected" to j be given by

$$\mathcal{V}_j = \{1 \leq i \leq n \mid (r_P^{(j)} B)_i \neq 0\}.$$

Define $d_j := |\mathcal{V}_j|$ and write $\mathcal{V}_j = \{i_1 < i_2 < \ldots < i_{d_j}\}$. A corresponding velocity projection operator $r_V^{(j)} : \mathbb{R}^n \to \mathbb{R}^{d_j}$ is given by

$$r_V^{(j)}\mathbf{u} = (u_{i_1}, u_{i_2}, \ldots, u_{i_{d_j}})^T.$$

The combined pressure and velocity projection is given by

$$r^{(j)} = \begin{pmatrix} r_V^{(j)} & 0 \\ 0 & r_P^{(j)} \end{pmatrix} \in \mathbb{R}^{(d_j+1)\times(n+m)}.$$

Furthermore, define $p^{(j)} = \left(r^{(j)}\right)^T$. Using these operators we can formulate a standard multiplicative Schwarz method. Define

$$\mathcal{A}^{(j)} := r^{(j)}\mathcal{A}p^{(j)} =: \begin{pmatrix} A^{(j)} & {B^{(j)}}^T \\ B^{(j)} & 0 \end{pmatrix} \in \mathbb{R}^{(d_j+1)\times(d_j+1)}.$$

Note that $B^{(j)}$ is a row vector of length d_j. In addition, we define

$$\mathcal{D}^{(j)} = \begin{pmatrix} \operatorname{diag}(A^{(j)}) & {B^{(j)}}^T \\ B^{(j)} & 0 \end{pmatrix} = \begin{pmatrix} \ddots & 0 & \vdots \\ 0 & \ddots & \vdots \\ \cdots & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{(d_j+1)\times(d_j+1)}.$$

The *full* Vanka smoother is a multiplicative Schwarz method with iteration matrix

$$\mathcal{S}_{\text{full}} = \prod_{j=1}^{m} \left(I - p^{(j)}(\mathcal{A}^{(j)})^{-1} r^{(j)}\mathcal{A}\right). \tag{20}$$

The *diagonal* Vanka smoother is similar, but with $\mathcal{D}^{(j)}$ instead of $\mathcal{A}^{(j)}$:

$$\mathcal{S}_{\text{diag}} = \prod_{j=1}^{m} \left(I - p^{(j)}(\mathcal{D}^{(j)})^{-1} r^{(j)}\mathcal{A}\right). \tag{21}$$

Thus, a smoothing step with a Vanka-type smoother consists of a loop over all pressure degrees of freedom ($j = 1, \ldots, m$), where for each j a linear system of equations with the matrix $\mathcal{A}^{(j)}$ (or $\mathcal{D}^{(j)}$) has to be solved. The degrees of freedom are updated in a Gauss-Seidel manner. These two methods are well-defined if all matrices $\mathcal{A}^{(j)}$ and $\mathcal{D}^{(j)}$ are nonsingular.

The linear systems with the diagonal Vanka smoother can be solved very efficiently using the special structure of the matrix $\mathcal{D}^{(j)}$ whereas for the systems with the full Vanka smoother a direct solver for the systems with the matrices $\mathcal{A}^{(j)}$ is required. The computational costs for solving a local (i.e. for each block) linear system of equations is $\sim d_j$ for the diagonal Vanka smoother and $\sim d_j^3$ for the full Vanka smoother. Typical values for d_j are given in Table 2.

As far as we know *there is no convergence analysis of a multigrid method with a Vanka smoother* applied to two- or three-dimensional Stokes problems. In [14] a convergence analysis, based on Fourier transformations, is given for a model one- or two-dimensional Poisson problem in the mixed finite element formulation. Note that in this case the finite element spaces are different from the ones used for Stokes problems. Recently, in [13] convergence for certain Vanka-type iterative methods applied to Stokes and Navier-Stokes problems has been proved.

We discuss the effect of a rescaling of the system as in (18) on the behaviour of the two-grid method with a full- or diagonal Vanka smoother. Let $\mathcal{S}_{V,\delta}$ be the iteration matrix of a Vanka smoother as in (20) or (21) with $\mathcal{A} = \mathcal{A}_l$ replaced by $\mathcal{A}_\delta = \mathcal{A}_{l,\delta}$ as in (18). Let $\mathcal{M}_{V,\delta}$ be the iteration matrix of the corresponding two-grid method (on level l):

$$\mathcal{M}_{V,\delta} = \mathcal{S}_{V,\delta}^{\nu_2}(I - P_l \mathcal{A}_{l-1,\delta}^{-1} R_l \mathcal{A}_{l,\delta})\mathcal{S}_{V,\delta}^{\nu_1}.$$

For this method the same scaling result as for the two-grid method with Braess-Sarazin smoother in Lemma 3 holds.

Lemma 4. *For $\delta \neq 0$ the relation*

$$\mathcal{M}_{V,\delta} = I_{l,\delta}\, \mathcal{M}_{V,1}\, I_{l,\delta}^{-1} \tag{22}$$

holds.

Proof: We consider the full Vanka smoother. The same analysis, with obvious modifications, applies to the diagonal Vanka smoother, too. We write $\mathcal{A}_\delta$ instead of $\mathcal{A}_{l,\delta}$. Define $\mathcal{A}_\delta^{(j)} := r^{(j)}\, \mathcal{A}_\delta\, p^{(j)}$. Using the relations

$$p^{(j)} I_{l,\delta} = I_{l,\delta} p^{(j)}, \quad p^{(j)} I_{l,\delta}^{-1} = I_{l,\delta}^{-1} p^{(j)}, \quad r^{(j)} \tilde{I}_{l,\delta} = \tilde{I}_{l,\delta} r^{(j)}, \quad r^{(j)} \tilde{I}_{l,\delta}^{-1} = \tilde{I}_{l,\delta}^{-1} r^{(j)},$$

we obtain

$$p^{(j)} (\mathcal{A}_\delta^{(j)})^{-1}\, r^{(j)} = I_{l,\delta}\, p^{(j)}\, (\mathcal{A}_1^{(j)})^{-1}\, r^{(j)}\, \tilde{I}_{l,\delta}^{-1}.$$

Thus we get

$$I - p^{(j)} (\mathcal{A}_\delta^{(j)})^{-1}\, r^{(j)}\, \mathcal{A}_\delta = I_{l,\delta} \left(I - p^{(j)} (\mathcal{A}_1^{(j)})^{-1}\, r^{(j)}\, \mathcal{A}_1 \right) I_{l,\delta}^{-1}.$$

This yields $\mathcal{S}_{V,\delta} = I_{l,\delta}\, \mathcal{S}_{V,1}\, I_{l,\delta}^{-1}$. In combination with the properties in (19) we obtain the result in (22). ∎

4. Other coupled iterative methods

In this section we consider a preconditioned minimal residual (PMINRES) method and an inexact Uzawa method for solving the discretized Stokes problem. We drop the level index l in the notation, i.e., the system matrix is denoted by

$$\mathcal{A} = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}.$$

Both methods require preconditioners for the matrix $A \in \mathbb{R}^{n \times n}$ and for the Schur complement $S = BA^{-1}B^T \in \mathbb{R}^{m \times m}$. Note that both A and S are symmetric positive definite (S: on $1^\perp$). Let Q_A and Q_S be symmetric positive definite preconditioners of A and S, respectively. Let $\gamma_A > 0$, $\gamma_S > 0$, Γ_A and Γ_S be spectral bounds such that

$$\gamma_A Q_A \leq A \leq \Gamma_A Q_A \tag{23}$$

and

$$\gamma_S Q_S \leq S \leq \Gamma_S Q_S. \tag{24}$$

Below we first specify the choice of Q_A and of Q_S for the discrete generalized Stokes problem. Then we discuss the PMINRES and inexact Uzawa method.

4.1. Preconditioners for Q_A and Q_S

The matrix A has block diagonal form with identical blocks. Such a block corresponds to the finite element discretization of a scalar reaction-diffusion problem of the form $-\nu\Delta u+\xi u = f$. For Q_A we use one iteration of a symmetric V-cycle multigrid method (for each of the blocks in A). In [15] it is shown that for this preconditioner the inequalities

$$\gamma_A Q_A \leq A \leq Q_A \tag{25}$$

hold with a constant $\gamma_A > 0$ *independent of* l, ν *and* ξ. Note that the upper spectral constant is $\Gamma_A = 1$. For typical multigrid methods the spectral constant γ_A is close to one (typically $\gamma_A \geq 0.85$).

We now discuss the choice of Q_S. For this we introduce an auxiliary Neumann problem in the pressure space, with a given $g \in L^2(\Omega)$:

$$\begin{aligned} &\textit{Find } w \in H^1(\Omega) \cap L^2_0(\Omega) \textit{ such that} \\ &\qquad (\nabla w, \nabla \phi) = (g, \phi) \quad \textit{for all} \quad \phi \in H^1(\Omega) \cap L^2_0(\Omega). \end{aligned}$$

Let $N = N_l$ be the stiffness matrix resulting from a finite element discretization of this problem in the pressure finite element space Q_l. Let $M = M_l$ be the mass matrix for the pressure space. Let Q_N a preconditioner of N induced by one symmetric V-cycle multigrid iteration applied to the discrete problem with stiffness matrix N. The (Cahouet-Chabard) Schur complement preconditioner Q_S is given by

$$Q_S^{-1} := \tau M^{-1} + \xi Q_N^{-1}, \quad \tau = \max\{\nu, \xi h_l^2\}. \tag{26}$$

In [5, 16] it is shown that under certain regularity assumptions for the Stokes problem this preconditioner has corresponding spectral bounds $\gamma_S > 0$, Γ_S in (24) that are *independent of the parameters* l, ξ and ν.

4.2. The preconditioned minimal residual method

In the preconditioned minimal residual (PMINRES) method used for solving a linear system with matrix $\mathcal{A}$ we use a block diagonal preconditioner defined by

$$\mathcal{M} = \begin{pmatrix} Q_A & 0 \\ 0 & Q_S \end{pmatrix}. \tag{27}$$

For a discussion and an efficient implementation of the PMINRES method we refer to the literature [2, 17]. In an efficient implementation one needs per PMINRES iteration one evaluation of Q_A^{-1}, one evaluation of Q_S^{-1} and one matrix-vector product with $\mathcal{A}$ (and a few other inexpensive operations).

Let $\mathbf{r}^{(k)}$ be the residual in the k-th iteration of this method. The convergence of the PMINRES can be analyzed based on the well-known residual bound

$$\frac{\|\mathbf{r}^{(k)}\|_{\mathcal{M}^{-1}}}{\|\mathbf{r}^{(0)}\|_{\mathcal{M}^{-1}}} \leq \min_{p_k \in \Pi_k, p_k(0)=1} \max_{\lambda \in \sigma(\mathcal{M}^{-1}\mathcal{A})} |p_k(\lambda)|. \tag{28}$$

For the spectrum of the preconditioned matrix $\mathcal{M}^{-1}\mathcal{A}$ the following result is given in [19, 20]:

$$\begin{aligned} \sigma(\mathcal{M}^{-1}\mathcal{A}) \quad &\subset \quad \left[\tfrac{1}{2}\left(\gamma_A - \sqrt{\gamma_A^2 + 4\Gamma_S\Gamma_A}\right), \tfrac{1}{2}\left(\gamma_A - \sqrt{\gamma_A^2 + 4\gamma_S\gamma_A}\right)\right] \\ &\cup \quad \left[\gamma_A, \tfrac{1}{2}\left(\Gamma_A + \sqrt{\gamma_A^2 + 4\Gamma_S\Gamma_A}\right)\right]. \end{aligned} \tag{29}$$

This general result implies that the rate of convergence of the PMINRES method is robust with respect to variation of parameters (in our case: l, ν and ξ) if the spectral constants in (23) and (24) do not depend on these parameters. For our choice of the preconditioners this is indeed the case and thus the PMINRES method with Q_A and Q_S as explained in section 4.1 *has a rate of convergence that is robust with respect to variation in the parameters l, ν and ξ.*

Remark 3. Consider the special case of only a rescaling of the A_l-block in the matrix $\mathcal{A}_l$ as in (18) (parameter: $\xi = \nu = \delta$). The multigrid preconditioner and the Cahouet-Chabard Schur complement preconditioner automatically take this scaling into account. Let $\mathcal{M}_\delta$ be the block preconditioner as in (27) for $\mathcal{A}_\delta = \mathcal{A}_{l,\delta}$. Then $\mathcal{M}_\delta^{-1}\mathcal{A}_\delta = I_\delta \mathcal{M}_1^{-1}\mathcal{A}_1 I_\delta^{-1}$ (with $I_\delta = I_{l,\delta}$ as in (17)) and thus $\sigma(\mathcal{M}_\delta^{-1}\mathcal{A}_\delta) = \sigma(\mathcal{M}_1^{-1}\mathcal{A}_1)$ for all $\delta > 0$.

4.3. The inexact Uzawa method

For the derivation of the inexact Uzawa method we consider the exact block factorization of the matrix $\mathcal{A}$

$$\mathcal{A} = \begin{pmatrix} A & 0 \\ B & -I \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & S \end{pmatrix}. \tag{30}$$

An *approximate* Schur complement is given by

$$\hat{S} = BQ_A^{-1}B^T. \tag{31}$$

Using the block factorization (30) and substituting $A^{-1} \approx Q_A^{-1}$ and $S^{-1} \approx \hat{S}^{-1}$ we obtain the approximate inverse of $\mathcal{A}$

$$\mathcal{A}^{-1} \approx \begin{pmatrix} I & -Q_A^{-1}B\hat{S}^{-1} \\ 0 & \hat{S}^{-1} \end{pmatrix} \begin{pmatrix} Q_A^{-1} & 0 \\ -BQ_A^{-1} & -I \end{pmatrix}. \tag{32}$$

In general the application of Q_A is feasible, but $\hat{S}^{-1}\mathbf{w}$ can not be determined with acceptable computational costs. Therefore we use $\hat{S}^{-1}\mathbf{w} \approx \Psi(\mathbf{w})$ with $\Psi(\mathbf{w})$ the result of a PCG method with zero starting vector and preconditioner Q_S applied to $\hat{S}\mathbf{z} = \mathbf{w}$ such that

$$\|\Psi(\mathbf{w}) - \mathbf{z}\|_{\hat{S}} \leq \theta \|\mathbf{z}\|_{\hat{S}} \tag{33}$$

holds for some prescribed tolerance $\theta < 1$.

Based on this, the inexact Uzawa method is as follows. Let $\mathbf{x}^{(0)} = (\mathbf{u}^{(0)}, \mathbf{p}^{(0)})^T$ be an initial approximation, and $\mathbf{r}^{(0)} = \mathbf{f} - \mathcal{A}\mathbf{x}^{(0)} = (\mathbf{r}_u^{(0)}, \mathbf{r}_p^{(0)})$ the initial residual. For $k = 0,\ 1,\ \ldots$:

- Compute $\mathbf{v} = \mathbf{u}^{(k)} + Q_A^{-1}\mathbf{r}_u^{(k)}$.
- Solve $\hat{S}\mathbf{z} = B\mathbf{v}$ approximately: $\mathbf{z} = \Psi(B\mathbf{v})$ as in (33).
- Update the approximation for the velocity $\mathbf{u}^{(k+1)} = \mathbf{v} - Q_A^{-1}B^T\mathbf{z}$.
- Update the approximation for the pressure $\mathbf{p}^{(k+1)} = \mathbf{p}^{(k)} + \mathbf{z}$.
- Compute the residual $\mathbf{r}^{(k+1)} = \mathbf{f} - \mathcal{A}\mathbf{x}^{(k+1)}$ with $\mathbf{x}^{(k+1)} = (\mathbf{u}^{(k+1)}, \mathbf{p}^{(k+1)})$.
- Check stopping criterion.

Remark 4. We briefly comment on the computational costs per iteration of this method. These costs depend on the number of PCG iterations needed to satisfy the tolerance requirement in (33). Assume that this number is q. In [18] it is shown that per iteration we then need $q+1$ evaluations of Q_A^{-1}, q evaluations of Q_S^{-1}, $q+1$ matrix-vector multiplications

with B, q matrix-vector multiplications with B^T and one matrix-vector multiplication with A. Moreover, in [18] it is also argued that in general a value for the tolerance parameter $\theta \approx \frac{1}{2}$ in (33) is close to optimal (w.r.t. efficiency) resulting in very low values for q (typically $q = 1$ or $q = 2$).

We give a convergence result for this inexact Uzawa method that is proved in [18].

Theorem 2. *Assume that* (23), (24) *hold with* $\Gamma_A = 1$. *Define*

$$\mu_A = 1 - \gamma_A, \quad g(\mu_A, \theta) = 2\mu_A + \theta(1 + \mu_A).$$

Consider the inexact Uzawa method with Ψ *such that* (33) *holds. For the error we have*

$$\max\left\{\|\mathbf{e}_u^{(k+1)}\|_{Q_A}, \|\mathbf{e}_p^{(k+1)}\|_{\hat{S}}\right\} \leq g(\mu_A, \theta) \max\left\{\|\mathbf{e}_u^{(k)}\|_{Q_A}, \|\mathbf{e}_p^{(k)}\|_{\hat{S}}\right\}$$

and

$$\|\mathbf{e}_u^{(k)}\|_{Q_A} + \|\mathbf{e}_p^{(k)}\|_{\hat{S}} \leq \frac{7}{2}\left(\frac{g(\mu_A, \theta) + \sqrt{g(\mu_A, \theta)^2 - 4\mu_A\theta}}{2}\right)^k (\|\mathbf{e}_u^{(0)}\|_{Q_A} + \|\mathbf{e}_p^{(0)}\|_{\hat{S}}).$$

Note that the assumption $\Gamma_A = 1$ is satisfied for the multigrid preconditioner Q_A that we use, cf. (25). For $\mu_A \to 0$ we obtain the contraction factor of the exact Uzawa method: $g(0, \theta) = \theta$. We also have $g(\mu_A, \theta) \geq \frac{1}{2}\big(g(\mu_A, \theta) + \sqrt{g(\mu_A, \theta)^2 - 4\mu_A\theta}\big)$ and

$$g(\mu_A, \theta) < 1 \quad \text{iff} \quad 0 \leq \theta < \frac{1 - 2\mu_A}{1 + \mu_A}\ , \tag{34}$$

$$\frac{1}{2}\big(g(\mu_A, \theta) + \sqrt{g(\mu_A, \theta)^2 - 4\mu_A\theta}\big) < 1 \quad \text{iff} \quad 0 \leq \theta < 1 - 2\mu_A\ . \tag{35}$$

Hence, for $\mu_A < \frac{1}{2}$ and θ sufficiently small (as quantified in (34), (35)) we have a convergent method. Moreover, these bounds for the contration number are independent of parameters (in our case: l, ν and ξ) if μ_A and θ are independent of these parameters. For the multigrid preconditioner we have $\mu_A \approx 0.15$, independent of l, ν and ξ, cf. section 4.1. We take $\theta \approx \frac{1}{2}$, cf. remark 4. Due to the fact that the Schur complement preconditioner Q_S discussed in section 4.1 has spectral constant γ_S and Γ_S independent of l, ν and ξ, an approximate solution that satisfies (33) can be computed with a low and uniformly (w.r.t. parameters) bounded number of PCG iterations. This implies that this method is optimal in the sense that the amount of work per iteration is proportional to that of a few matrix-vector multiplications and the rate of convergence is robust w.r.t. variation in the parameters l, ν and ξ.

Remark 5. Consider the special case of only a rescaling of the A_l-block in the matrix $\mathcal{A}_l$ as in remark 3. The multigrid preconditioner and the Cahouet-Chabard Schur complement preconditioner automatically take this scaling into account. One may check that the result of the inexact Uzawa method are not influenced by this scaling.

5. Comparison of methods

Before we turn to numerical experiments in the next section, we compare the methods discussed in the sections 3 and 4 from a theoretical point of view. Three issues are relevant here, namely

the arithmetic work per iteration, the rate of convergence and the dependence of this rate of convergence on parameters. We consider the following four methods: multigrid with Braess-Sarazin smoother (denoted by BS-MGM), multigrid with diagonal Vanka smoother (denoted by V-MGM), preconditioned MINRES with preconditioner as explained in section 4.1 (denoted by PMINRES) and the inexact Uzawa method given in section 4.3 with preconditioners as in section 4.1 (denoted by MGUZAWA).

Arithmetic work per iteration. For the methods V-MGM, PMINRES and MGUZAWA the arithmetic work per iteration is bounded by $c(m_l + n_l)$ with a constant c independent of l. For the BS-MGM such a bound only holds if the linear system in (7) is solved approximately, cf. remark 2. In that case, however, the convergence analysis does not apply anymore.

Convergence analysis. We consider *fixed* values for the problem parameters ξ and ν. For the BS-MGM there is a convergence result as in theorem 1, provided the linear system in (7) is solved exactly. We do not know any literature in which for a V-MGM two- or multigrid method a convergence rate independent of the mesh size is proved. For PMINRES a rate of convergence independent of l follows from (28) and (29). For the MGUZAWA method such a convergence result follows from theorem 2.

Robustness w.r.t. variation in parameters. First consider the special case of a rescaling of the A_l block of the matrix with a parameter δ as in (18). For all four methods the rate of convergence can be shown to be essentially independent of δ (cf. corollary 1, lemma 4 and the remarks 3 and 5). For the general case of variable problem parameters $\xi \geq 0$ and $\nu > 0$ there are no robustness results on the convergence of BS-MGM or V-MGM. The rate of convergence of the PMINRES and MGUZAWA methods is robust w.r.t. variation in these parameters. This follows from the results presented in section 4.

Summarizing, we conclude that concerning theoretical convergence results the state of affairs is much better for the PMINRES and MGUZWA methods than for the coupled multigrid methods.

Parameters in the methods. The V-MGM and the PMINRES methods are parameter free. In the BS-MGM method one has to choose a value for the parameter α in the smoother, cf. (4). In the MGUZAWA method the accuracy parameter θ in (33) occurs.

6. Numerical experiments

We consider the generalized Stokes equation as in (1) on the domain $\Omega = [0,1]^3$. The right-hand side $\mathbf{f}$ is taken such that the continuous solution is

$$\begin{aligned} \mathbf{u}(x,y,z) &= \frac{1}{3}\begin{pmatrix} \sin(\pi x)\sin(\pi y)\sin(\pi z) \\ -\cos(\pi x)\cos(\pi y)\sin(\pi z) \\ 2\cdot\cos(\pi x)\sin(\pi y)\cos(\pi z) \end{pmatrix}, \\ p(x,y,z) &= \cos(\pi x)\sin(\pi y)\sin(\pi z) + C \end{aligned}$$

with a constant C such that $\int_\Omega p\,dx = 0$. For the discretization we start with a uniform tetrahedral grid with $h_0 = \frac{1}{2}$ and we apply regular refinements to this starting discretization.

For the finite element discretization we use the Hood-Taylor $\mathcal{P}_2$-$\mathcal{P}_1$ pair. In Table 1 the dimension of the system to be solved on each level and the corresponding step size are given.

	$h_0 = 2^{-1}$	$h_1 = 2^{-2}$	$h_2 = 2^{-3}$	$h_3 = 2^{-4}$	$h_4 = 2^{-5}$
n_l	81	1029	10125	89373	750141
m_l	27	125	729	4913	35937

Table 1. Dimensions: n_l = number of velocity unknowns, m_l = number of pressure unknowns.

In all tests below the iterations were repeated until the condition

$$\frac{\|\mathbf{r}^{(k)}\|}{\|\mathbf{r}^{(0)}\|} < 10^{-10},$$

with $\mathbf{r}^{(k)} = \mathbf{b} - \mathcal{A}\mathbf{x}^{(k)}$, was satisfied. The methods are implemented in the DROPS package [1]. All calculations were performed on AMD Athlon XP 1700 (1465 MHz) in double precision.

We first consider an experiment to show that for this problem class the multigrid method with full Vanka smoother is very time consuming. In Table 2 we show the maximal and mean values of d_j on the level l. These numbers indicate the dimensions of the local systems that have to be solved in the Vanka smoother, cf. section 3.2.

	$h_0 = 2^{-1}$	$h_1 = 2^{-2}$	$h_2 = 2^{-3}$	$h_3 = 2^{-4}$	$h_4 = 2^{-5}$
$\text{mean}(d_j)/\max_j d_j$	21.8 / 82	51.7 / 157	88.8 / 157	119.1 / 165	138.1 / 166

Table 2. The maximal and mean values of d_j on different grids.

We use a multigrid W-cycle with 2 pre- and 2 post-smoothing iterations. In Table 3 we show the computing time (in seconds) and the number of iterations needed both for the full Vanka $\mathcal{S}_{\text{full}}$ and the diagonal Vanka $\mathcal{S}_{\text{diag}}$ smoother.

$\xi = 0$	$\mathcal{S}_{\text{full}}$, $h_3 = 2^{-4}$	$\mathcal{S}_{\text{diag}}$,$h_3 = 2^{-4}$	$\mathcal{S}_{\text{full}}$, $h_4 = 2^{-5}$	$\mathcal{S}_{\text{diag}}$,$h_4 = 2^{-5}$
$\nu = 1$	287 (4)	19 (10)	3504 (5)	224 (13)
$\nu = 10^{-1}$	283 (4)	19 (10)	3449 (5)	238 (13)
$\nu = 10^{-2}$	284 (4)	19 (10)	3463 (5)	238 (13)
$\nu = 10^{-3}$	356 (5)	20 (11)	3502 (5)	238 (13)

Table 3. CPU time and number of iterations for multigrid with the full and the diagonal Vanka smoother.

As can be seen from these results, the rather high dimensions of the local systems lead to high computing times for the multigrid method with the full Vanka smoother compared to the method with the diagonal Vanka smoother. In numerical experiments we observed that the multigrid W-cycle with only *one* pre- and post-smoothing iteration with the diagonal Vanka method sometimes diverges. Further tests indicate that often for the method with diagonal Vanka smoothing the choice $\nu_1 = \nu_2 = 4$ is (slightly) better (w.r.t. CPU time) than $\nu_1 = \nu_2 = 2$.

Based on numerical experiments, in the multigrid W-cycle with Braess-Sarazin smoother we use $\nu_1 = \nu_2 = 2$ and $\alpha = 1.25$. For other values $\alpha \in [1.1, 1.5]$ the efficiency is very similar. The linear system in (7) is solved approximately using a conjugate gradient method with a fixed relative tolerance $\varepsilon_{CG} = 10^{-2}$.

We compare the following four methods:

PMINRES – preconditioned minimal residual methods described in section 4.2 with preconditioners as discussed in section 4.1.
MGUZAWA – the inexact Uzawa method described in section 4.3 with preconditioners as discussed in section 4.1. The accuracy parameter is set to $\theta = 0.1$.
BS-MGM – the multigrid method with the Braess-Sarazin smoother described in section 3.1 with parameter $\alpha = 1.25$. We take $\nu_1 = \nu_2 = 2$.
V-MGM – the multigrid method with the diagonal Vanka smoother described in section 3.2. We take $\nu_1 = \nu_2 = 4$.

Results for these methods are given in Table 4 and Table 5.

$\xi = 0$	$h_3 = 2^{-4}$			
ν	MGUZAWA	V-MGM	BS-MGM	PMINRES
$\nu = 1$	39 (13)	19 (5)	20 (11)	49 (74)
$\nu = 10^{-1}$	38 (13)	19 (5)	20 (11)	52 (79)
$\nu = 10^{-3}$	43 (14)	19 (5)	17 (8)	53 (80)
$\xi = 10$	$h_3 = 2^{-4}$			
ν	MGUZAWA	V-MGM	BS-MGM	PMINRES
$\nu = 1$	47 (15)	19 (5)	20 (11)	57 (79)
$\nu = 10^{-1}$	34 (11)	17 (4)	20 (11)	61 (89)
$\nu = 10^{-3}$	34 (13)	15 (3)	21 (7)	51 (74)
$\xi = 100$	$h_3 = 2^{-4}$			
ν	MGUZAWA	V-MGM	BS-MGM	PMINRES
$\nu = 1$	36 (12)	17 (4)	20 (11)	51 (73)
$\nu = 10^{-1}$	29 (10)	15 (3)	19 (7)	49 (69)
$\nu = 10^{-3}$	38 (15)	15 (3)	19 (6)	58 (85)

Table 4. CPU time and the number of iterations for MGUZAWA, PMINRES, BS- and V-MGM methods.

From these results we conclude the following:

(1) Not only the PMINRES and MGUZAWA methods are robust with respect to variation in the parameters ξ and ν (as predicted by theory), but also both multigrid methods are.
(2) For $\xi = 0$ variation of ν is the same as rescaling of the A_l block as in (18). The results show that for all four methods the rate of convergence is essentially independent of this scaling parameter ν, as predicted by theory.
(3) In most cases V-MGM is the most efficient method and PMINRES has the lowest efficiency.
(4) The parameter free PMINRES method shows a close to constant behaviour (only relatively small changes in CPU time and number of iterations) if the parameters ν and ξ are varied. In this sense this method is the most robust one.

$\xi = 0$	$h_4 = 2^{-5}$			
ν	MGUZAWA	V-MGM	BS-MGM	PMINRES
$\nu = 1$	361 (14)	198 (5)	274 (14)	445 (75)
$\nu = 10^{-1}$	315 (12)	199 (5)	276 (14)	476 (81)
$\nu = 10^{-3}$	319 (12)	198 (5)	241 (11)	441 (74)
$\xi = 10$	$h_3 = 2^{-5}$			
ν	MGUZAWA	V-MGM	BS-MGM	PMINRES
$\nu = 1$	419 (15)	190 (5)	244 (13)	538 (82)
$\nu = 10^{-1}$	321 (12)	189 (5)	224 (10)	548 (87)
$\nu = 10^{-3}$	265 (11)	145 (3)	238 (7)	540 (87)
$\xi = 100$	$h_3 = 2^{-5}$			
ν	MGUZAWA	V-MGM	BS-MGM	PMINRES
$\nu = 1$	305 (11)	190 (5)	241 (13)	484 (75)
$\nu = 10^{-1}$	261 (10)	167 (4)	243 (13)	488 (77)
$\nu = 10^{-3}$	329 (14)	122 (2)	282 (9)	441 (68)

Table 5. CPU time and the number of iterations for MGUZAWA, PMINRES, BS- and V-MGM methods.

REFERENCES

1. HTTP://WWW.IGPM.RWTH-AACHEN.DE/DROPS/
2. M. Benzi, G. H. Golub, J. Liesen, *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005), pp. 1–137.
3. D. Braess, W. Dahmen, C. Wieners, *A multigrid algorithm for the mortar finite element method*, SIAM J. Numer. Anal., 37 (1999), pp. 48–69.
4. D. Braess, R. Sarazin, *An efficient smoother for the Stokes problem*, Applied Numerical Mathematics, 23 (1997), pp. 3–19.
5. J.H. Bramble, J.E. Pasciak, *Iterative techniques for time dependent Stokes problems*, Comput. Math. Appl., 33 (1997), pp. 13–30.
6. M. Dauge, *Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. Part I: linearized equations*, SIAM J. Math. Anal., 22 (1989), pp. 74–97.
7. W. Hackbusch, *Multigrid methods and applications*, Springer, Berlin - Heidelberg - New York, 1985.
8. W. Hackbusch, *Iterative solution of large sparse systems of equations*, Springer, New York, 1994.
9. V. John, *A comparison of parallel solvers for the incompressible Navier-Stokes equations*, Comput. Visual. Sci. 1 (1999), pp. 193–200.
10. V. John, G. Matthies, T.I. Mitkova, L. Tobiska, P.S. Vassilevski, *A comparison of three solvers for the incompressible Navier-Stokes equations*, in: Large-Scle Scientific Computations of Engineering and Environmental Problems II, ed. by M. Griebel et al., Notes on Numerical Fluid Mechanics, 73 (2000), pp. 215–222.
11. V. John, L. Tobiska, *A coupled multigrid method for nonconforming finite element discretizations of the 2d-Stokes equation*, Computing, 64 (2000), pp. 307–321.
12. V. John, G. Matthies, *Higher order finite element discretizations in a benchmark problem for incompressible flows*, Int. J. Num. Meth. Fluids, 37 (2001), pp. 885–903.
13. S. Manservisi, *Numerical analysis of Vanka-type solvers for steady Stokes and Navier-Stokes flows*, SIAM J. Numer. Anal., 44 (2006), pp. 2025–2056.
14. J. Molenaar, *A two-grid analysis of the combination of mixed finite elements and Vanka-type relaxation*, In: Multigrid methods, III (W. Hackbusch and U. Trottenberg eds.), International Series of Numerical Mathematics, 98 (1991), Birkhäuser Verlag, Basel, pp. 313–323.
15. M. Olshanskii, A. Reusken, *On the convergence of a multigrid method for linear reaction-diffusion problems*, Computing, 65 (2000), pp. 193–202.
16. J. Peters, M. A. Olshanskii, 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.

17. C. Paige, M. Sounders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Num. Anal., 12 (1975), pp. 617–629.
18. J. Peters, V. Reichelt, A. Reusken, *Fast iterative solvers for discrete Stokes equations*, SIAM J. Sci. Comput., 27 (2005), pp. 646–666.
19. T. Rusten, R. Winther, *A preconditioned iterative method for saddlepoint problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.
20. D. Silvester, A. Wathen, *Fast iterative solution of stabilised Stokes systems. Part II: Using general block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367.
21. S. Turek, *Solvers for incompressible flow problems*, Springer, Berlin, 1999.
22. S. Vanka, *Block-implicit multigrid calculation of two-dimensional recirculating flows*, Computer Methods in Applied Mechanics and Engineering, 59 (1986), pp. 29–48.