FAST ITERATIVE SOLVERS FOR DISCRETE STOKES EQUATIONS

JÖRG PETERS, VOLKER REICHELT AND ARNOLD REUSKEN*

Abstract. We consider saddle point problems that result from the finite element discretization of stationary and instationary Stokes equations. Three efficient iterative solvers for these problems are treated, namely the preconditioned CG method introduced by Bramble and Pasciak, the preconditioned MINRES method and a method due to Bank *et al.* We give a detailed overview of algorithmic aspects and theoretical convergence results. For the method of Bank *et al* a new convergence analysis is presented. A comparative study of the three methods for a 3D Stokes problem discretized by the Hood-Taylor $P_2 - P_1$ finite element pair is given.

AMS subject classifications. 65N30, 65F10

Key words. Stokes equations, inexact Uzawa methods, preconditioned MINRES, multigrid

1. Introduction. We consider a class of Stokes equations on a bounded connected polyhedral Lipschitz domain Ω in d-dimensional Euclidean space. We use the notation $\mathbf{V} := H_0^1(\Omega)^d$ for the velocity space and $M = L_0^2(\Omega) := \{\, p \in L^2(\Omega) \mid \int_\Omega p(x)\,dx = 0 \,\}$ for the pressure space. The variational problem is as follows: given $\mathbf{f} \in L^2(\Omega)^d$ find $\{\mathbf{u}, p\} \in \mathbf{V} \times M$ such that

$$\begin{cases} (\nabla \mathbf{u}, \nabla \mathbf{v}) + \xi(\mathbf{u}, \mathbf{v}) - (\operatorname{div} \mathbf{v}, p) &= (\mathbf{f}, \mathbf{v}) \quad \text{for } \mathbf{v} \in \mathbf{V}, \\ (\operatorname{div} \mathbf{u}, q) &= 0 \quad \text{for } q \in M\ . \end{cases} \tag{1.1}$$

with a constant $\xi \geq 0$. Here and in the remainder the L^2 scalar product and associated norm are denoted by $(\cdot,\cdot)$, $\|\cdot\|$, respectively. The zero order term $\xi(\mathbf{u}, \mathbf{v})$ is included in view of implicit time integration methods applied to instationary Stokes equations. For the discretization of this problem we use a pair of conforming LBB stable finite element spaces. This results in a saddle point problem of the form

$$\begin{pmatrix} \mathrm{A} & \mathrm{B}^T \\ \mathrm{B} & 0 \end{pmatrix} \begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} = \begin{pmatrix} \mathrm{f} \\ 0 \end{pmatrix} \tag{1.2}$$

Many different iterative methods for solving this discrete problem are known. One can apply multigrid techniques to the whole coupled system in (1.2), cf. [15, 27, 28]. Most other approaches are based on the prominent classical Uzawa method. This Uzawa method requires that $\mathrm{A}^{-1}\mathrm{x}$ can be computed exactly ([1]). In many variants of this method, which are often called *inexact* Uzawa methods, one tries to avoid the exact inversion by using an inner iterative method, for example, a Jacobi-like iteration (Arrow-Hurwicz algorithm, [1]) or a multigrid method, cf. [24]. We also mention (variants of) the preconditioned MINRES method [21, 22, 23, 26] and the method from [2]. A different approach is presented in [3]. There the indefinite problem (1.2) is reformulated as a symmetric positive definite problem. For most of these methods theoretical convergence analyses are known, cf. [2, 3, 5, 12, 21, 23, 24, 26, 29]. In [11] the performance of a few of these methods is compared by means of systematic numerical experiments for a stationary 2D Stokes problem.

In this paper we consider three representative methods from the large class of inexact Uzawa methods, namely the preconditioned CG method from [3] (denoted by

*Institut für Geometrie und Praktische Mathematik, RWTH-Aachen, D-52056 Aachen, Germany; email: reusken@igpm.rwth-aachen.de. This work was supported by the German Research Foundation through SFB 540.

BPCG), the preconditioned MINRES method from [21, 23, 26] (denoted by PMINRES) and the method from [2] (denoted by MGUZAWA). The topics treated in this paper are the following:

- For these three methods we discuss costs per iteration, known theoretical convergence results and some implementation issues. This makes it possible to make a fair comparison of these methods.
- For the MGUZAWA method we present a convergence analysis. This analysis is much simpler than the analyses presented in [2, 29]. The result that we obtain is different from the ones in [2, 29] and gives a better explanation of the observation that if one uses a very good preconditioner for A (like multigrid) then even with a very low accuracy in the inner iteration the MGUZAWA method converges (cf. remark 7).
- We present a comparative study of the performance of the three methods. For this we consider a Stokes problem as in (1.1) in 3D. We treat both the stationary ($\xi = 0$) and instationary ($\xi > 0$) case. For the discretization we apply the popular Hood-Taylor $P_2 - P_1$ finite element pair. As a preconditioner for A a standard multigrid method is used. For the Schur complement preconditioner we use the mass matrix for the stationary case and a more sophisticated preconditioner analyzed in [4] for the instationary case.

The main results of this paper are a detailed comparative study of the three fast iterative solvers for a 3D Stokes problem and a new convergence analysis for the MGUZAWA method.

The paper is organized as follows. In section 2 we formulate the standard Galerkin discretization of the Stokes problem. In section 3.1 we consider the BPCG method. An efficient implementation of this method is discussed and a main convergence result known from the literature is formulated. In section 3.2 the PMINRES method is treated. Also for this method a known convergence result is given. In section 3.3 we discuss the MGUZAWA method and give an effcient implementation of this method. In section 4 a convergence analysis of this method is given. Theorem 4.3 contains the main theoretical result of this paper. In section 5 we discuss the preconditioners for the A-block and for the Schur complement. In section 6 results of numerical experiments are presented that illustrate the performance of the three methods. Finally, in section 7 we give some conclusions.

2. The discrete Stokes equations. For the discretization of the Stokes problem (1.1) we assume a family of triangulations $\{\mathcal{T}_h\}$ in the sense of [9, 10] and a pair of finite element spaces $\mathbf{V}_h \subset \mathbf{V}$ and $M_h \subset M$ that is LBB stable with a constant $\hat{\beta}$ independent of h:

$$\inf_{q_h \in M_h} \sup_{\mathbf{v}_h \in \mathbf{V}_h} \frac{(\operatorname{div} \mathbf{v}_h, q_h)}{\|\nabla \mathbf{v}_h\| \|q_h\|} \geq \hat{\beta} > 0 \tag{2.1}$$

The Galerkin discretization is as follows: Find $\{\mathbf{u}_h, p_h\} \in \mathbf{V}_h \times M_h$ such that

$$\begin{cases} (\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) + \xi(\mathbf{u}_h, \mathbf{v}_h) - (\operatorname{div} \mathbf{v}_h, p_h) = (\mathbf{f}, \mathbf{v}_h) & \text{for } \mathbf{v}_h \in \mathbf{V}_h, \\ (\operatorname{div} \mathbf{u}_h, q_h) = 0 & \text{for } q_h \in M_h \, . \end{cases} \tag{2.2}$$

In practice the discrete space M_h for the pressure is constructed by taking a standard finite element space, which we denote by M_h^+ (for example, continuous piecewise linear functions) and then adding an orthogonality condition:

$$M_h = \{\, p_h \in M_h^+ \mid (p_h, 1) = 0 \,\}$$

Note that $\dim(M_h) = \dim(M_h^+) - 1$. Let $n := \dim(\mathbf{V}_h),\ m := \dim(M_h^+)$. We assume standard (nodal) bases in $\mathbf{V}_h$ and M_h^+ and corresponding isomorphisms

$$J_V : \mathbb{R}^n \to V_h,\ ,\quad J_M : \mathbb{R}^m \to M_h^+ \ .$$

Let the stiffness matrices $\mathrm{D} \in \mathbb{R}^{n\times n},\ \mathrm{B} \in \mathbb{R}^{m\times n}$ and the mass matrix $\mathrm{M} \in \mathbb{R}^{m\times m}$ be given by

$$\begin{aligned} \langle \mathrm{Dx}, \mathrm{y}\rangle &= (\nabla J_V \mathrm{x}, \nabla J_V \mathrm{y}) \quad \text{for all } \ \mathrm{x},\mathrm{y} \in \mathbb{R}^n \ , \\ \langle \mathrm{Bx}, \mathrm{y}\rangle &= (\operatorname{div} J_V \mathrm{x}, J_M \mathrm{y}) \quad \text{for all } \ \mathrm{x} \in \mathbb{R}^n, \mathrm{y} \in \mathbb{R}^m \ , \\ \langle \mathrm{Mx}, \mathrm{y}\rangle &= (J_M \mathrm{x}, J_M \mathrm{y}) \quad \text{for all } \ \mathrm{x},\mathrm{y} \in \mathbb{R}^m \ . \end{aligned} \tag{2.3}$$

Here $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean scalar product. The discrete problem has a matrix-vector representation of the form

$$\begin{pmatrix} \mathrm{A} & \mathrm{B}^T \\ \mathrm{B} & 0 \end{pmatrix} \begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} = \begin{pmatrix} \mathrm{f} \\ 0 \end{pmatrix} \ , \quad \mathrm{A} := \mathrm{D} + \xi \mathrm{M} \ , \tag{2.4}$$

with f such that $\langle \mathrm{f}, \mathrm{y}\rangle = (\mathbf{f}, J_V \mathrm{y})$ for all $\mathrm{y} \in \mathbb{R}^n$. We introduce the notation

$$\mathrm{K} = \begin{pmatrix} \mathrm{A} & \mathrm{B}^T \\ \mathrm{B} & 0 \end{pmatrix} \ , \quad \mathrm{S} := \mathrm{B}\mathrm{A}^{-1}\mathrm{B}^T \ .$$

Define the constant vector $\mathrm{e} := J_M^{-1} 1 = (1, \ldots, 1)^T \in \mathbb{R}^m$. Note that $\mathrm{B}^T \mathrm{e} = 0$ and that both the Schur complement S and the matrix K are singular and have a one-dimensional kernel. For the Schur complement this kernel is $\ker(\mathrm{S}) = \operatorname{span}\{\mathrm{e}\}$. Note that

$$(J_M \mathrm{y}, 1) = 0 \ \Leftrightarrow\ (J_M \mathrm{y}, J_M \mathrm{e}) = 0 \ \Leftrightarrow \langle \mathrm{My}, \mathrm{e}\rangle = 0 \ . \tag{2.5}$$

Hence, with $\mathrm{e}^{\perp_M} := \{\, \mathrm{y} \in \mathbb{R}^m \mid \langle \mathrm{y}, \mathrm{Me}\rangle = 0 \,\}$ we have $M_h = \{\, J_M \mathrm{y} \mid \mathrm{y} \in \mathrm{e}^{\perp_M} \,\}$ and we get the following matrix-vector representation of the discrete problem (2.2):

$$\text{Find } \ \mathrm{x} \in \mathbb{R}^n,\ \mathrm{y} \in \mathrm{e}^{\perp_M} \ \text{ such that (2.4) holds.} \tag{2.6}$$

3. Iterative solvers. In this section we describe three iterative solvers for the problem (2.6). The first two methods, a preconditioned conjugate gradient method (BPCG) and a preconditioned minimal residual method (PMINRES) are known from the literature and often used in practice. The third method, which may be less known, is a simple variant of the Uzawa method as presented in [2]. In section 4 a convergence analysis of this method is given.

In all three solvers we need symmetric positive definite preconditioners Q_A of A and Q_S of S. The quality of these preconditioners is described by the following spectral equivalences. Let $\gamma_A > 0,\ \Gamma_A,\ \gamma_S > 0,\ \Gamma_S$ be such that

$$\gamma_A \mathrm{Q}_A \le \mathrm{A} \le \Gamma_A \mathrm{Q}_A \ , \tag{3.1}$$

$$\gamma_S \mathrm{Q}_S \le \mathrm{S} \le \Gamma_S \mathrm{Q}_S \quad \text{on } \ \mathrm{e}^{\perp_M} \ . \tag{3.2}$$

In section 5 we will discuss concrete choices for the preconditioners Q_S and Q_A.

3.1. Preconditioned conjugate gradient method. We describe the preconditioned conjugate gradient (BPCG) method introduced in [3]. For this method we need the following assumption on the preconditioner Q_A

$$0 < \mathrm{Q}_A < \mathrm{A} \ , \tag{3.3}$$

i.e., $\gamma_A > 1$ in (3.1). Clearly, using some suitable scaling one can always satisfy the condition (3.3). Premultiplication of (2.4) by the matrix

$$\mathrm{G} = \begin{pmatrix} \mathrm{Q}_A^{-1} & 0 \\ \mathrm{B}\mathrm{Q}_A^{-1} & -\mathrm{I} \end{pmatrix}$$

yields an equivalent system

$$\hat{\mathrm{K}} \begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} = \mathrm{G} \begin{pmatrix} \mathrm{f} \\ 0 \end{pmatrix} \ , \quad \hat{\mathrm{K}} := \mathrm{G}\mathrm{K} = \begin{pmatrix} \mathrm{Q}_A^{-1}\mathrm{A} & \mathrm{Q}_A^{-1}\mathrm{B}^T \\ \mathrm{B}\mathrm{Q}_A^{-1}\mathrm{A} - \mathrm{B} & \mathrm{B}\mathrm{Q}_A^{-1}\mathrm{B}^T \end{pmatrix} \ . \tag{3.4}$$

Due to the assumption (3.3) the bilinear form

$$\left[\begin{pmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \end{pmatrix}, \begin{pmatrix} \mathrm{y}_1 \\ \mathrm{y}_2 \end{pmatrix} \right] = \langle (\mathrm{A} - \mathrm{Q}_A)\mathrm{x}_1, \mathrm{y}_1 \rangle + \langle \mathrm{x}_2, \mathrm{y}_2 \rangle \tag{3.5}$$

defines an inner product. In [3] it is shown that the matrix $\hat{\mathrm{K}}$ is *symmetric and positive definite* on $\mathbb{R}^n \times \mathrm{e}^{\perp_M}$ with respect to this inner product. The matrix

$$\tilde{\mathrm{K}} := \begin{pmatrix} \mathrm{I} & 0 \\ 0 & \mathrm{Q}_S \end{pmatrix}$$

is also symmetric positive definite with respect to this scalar product. Hence we can apply the preconditioned CG method in the scalar product $\left[\cdot, \cdot\right]$ and with preconditioner $\tilde{\mathrm{K}}$ to the linear system in (3.4). The algorithm is as follows:

$$\begin{cases} \mathrm{v}^0 = \begin{pmatrix} \mathrm{x}^0 \\ \mathrm{y}^0 \end{pmatrix} \quad \text{a given starting vector} \ ; \ \bar{\mathrm{r}}^0 := \begin{pmatrix} \mathrm{f} \\ 0 \end{pmatrix} - \mathrm{K}\mathrm{v}^0 \ , \quad \mathrm{r}^0 := \mathrm{G}\bar{\mathrm{r}}^0 \\ \text{for } k \geq 0 \quad (\text{if } \mathrm{r}^k \neq 0) : \\ \text{Solve } \mathrm{z}^k \text{ from } \tilde{\mathrm{K}}\mathrm{z}^k = \mathrm{r}^k \ . \\ \beta_k^{(n)} := \left[\mathrm{z}^k, \mathrm{r}^k\right] , \quad \beta_k := \beta_k^{(n)} / \beta_{k-1}^{(n)} \quad (\beta_0 := 0) \\ \mathrm{p}^k := \mathrm{z}^k + \beta_k \, \mathrm{p}^{k-1} \quad (\mathrm{p}^0 := \mathrm{z}^0) \\ \alpha_k^{(d)} := \left[\hat{\mathrm{K}}\mathrm{p}^k, \mathrm{p}^k\right] , \quad \alpha_k := \beta_k^{(n)} / \alpha_k^{(d)} \\ \mathrm{v}^{k+1} := \mathrm{v}^k + \alpha_k \, \mathrm{p}^k \\ \mathrm{r}^{k+1} := \mathrm{r}^k - \alpha_k \, \hat{\mathrm{K}}\mathrm{p}^k \end{cases} \tag{3.6}$$

This is the standard PCG algorithm. We now discuss an efficient implementation of this algorithm based on ideas from [11]. In particular we want to avoid the evaluation of Q_A. We use indizes 1 and 2 to denote the first and second vector components, respectively (as in (3.5)). First note that due to the structure of $\tilde{\mathrm{K}}$ we have $\mathrm{z}_1^k = \mathrm{r}_1^k$. We introduce $\bar{\mathrm{r}}_1^k := \mathrm{Q}_A \mathrm{r}_1^k$ and $\mathrm{d} := \mathrm{A}\mathrm{r}_1^k$. We then have

$$\begin{aligned} \left[\mathrm{z}^k, \mathrm{r}^k\right] &= \langle (\mathrm{A} - \mathrm{Q}_A)\mathrm{z}_1^k, \mathrm{r}_1^k \rangle + \langle \mathrm{z}_2^k, \mathrm{r}_2^k \rangle \\ &= \langle (\mathrm{A} - \mathrm{Q}_A)\mathrm{r}_1^k, \mathrm{r}_1^k \rangle + \langle \mathrm{z}_2^k, \mathrm{r}_2^k \rangle \\ &= \langle \mathrm{d}, \mathrm{r}_1^k \rangle - \langle \bar{\mathrm{r}}_1^k, \mathrm{r}_1^k \rangle + \langle \mathrm{z}_2^k, \mathrm{r}_2^k \rangle \ . \end{aligned} \tag{3.7}$$

Also define $\mathrm{s}^k := \mathrm{A}\mathrm{p}_1^k$. From $\mathrm{p}^k = \mathrm{z}^k + \beta_k \mathrm{p}^{k-1}$ it follows that for s^k we have the relation

$$\mathrm{s}^k = \mathrm{A}\mathrm{r}_1^k + \beta_k \, \mathrm{s}^{k-1} = \mathrm{d} + \beta_k \, s^{k-1} \ .$$

With

$$\begin{pmatrix} \mathrm{t} \\ \mathrm{u} \end{pmatrix} := \mathrm{K}\mathrm{p}^k = \begin{pmatrix} \mathrm{s}^k + \mathrm{B}^T \mathrm{p}_2^k \\ \mathrm{B}\mathrm{p}_1^k \end{pmatrix} \ , \quad \begin{pmatrix} \mathrm{y} \\ \mathrm{z} \end{pmatrix} := \hat{\mathrm{K}}\mathrm{p}^k = \begin{pmatrix} \mathrm{Q}_A^{-1}\mathrm{t} \\ \mathrm{B}\mathrm{Q}_A^{-1}\mathrm{t} - \mathrm{u} \end{pmatrix}$$

we obtain

$$\begin{aligned} \left[\hat{\mathrm{K}}\mathrm{p}^k, \mathrm{p}^k\right] &= \langle (\mathrm{A} - \mathrm{Q}_A)\mathrm{y}, \mathrm{p}_1^k \rangle + \langle \mathrm{z}, \mathrm{p}_2^k \rangle \\ &= \langle \mathrm{y}, \mathrm{A}\mathrm{p}_1^k \rangle - \langle \mathrm{t}, \mathrm{p}_1^k \rangle + \langle \mathrm{z}, \mathrm{p}_2^k \rangle = \langle \mathrm{y}, \mathrm{s}^k \rangle - \langle \mathrm{t}, \mathrm{p}_1^k \rangle + \langle \mathrm{z}, \mathrm{p}_2^k \rangle \ . \end{aligned} \tag{3.8}$$

Finally note that for $\bar{\mathrm{r}}_1^k$ we have the recursion

$$\bar{\mathrm{r}}_1^{k+1} := \bar{\mathrm{r}}_1^k - \alpha_k \, (\mathrm{K}\mathrm{p}^k)_1 = \bar{\mathrm{r}}_1^k - \alpha_k \mathrm{t} \ .$$

Summarizing we then obtain the following efficient implementation of the BPCG algorithm:

$$\left\{ \begin{array}{l} \mathrm{v}^0 = \begin{pmatrix} \mathrm{x}^0 \\ \mathrm{y}^0 \end{pmatrix} \quad \text{a given starting vector} \ ; \ \bar{\mathrm{r}}^0 := \begin{pmatrix} \mathrm{f} \\ 0 \end{pmatrix} - \mathrm{K}\mathrm{v}^0 \ , \quad \mathrm{r}^0 := \mathrm{G}\bar{\mathrm{r}}^0 \\ \text{for } k \geq 0 \quad (\text{if } \mathrm{r}^k \neq 0) : \\ \text{Solve } \mathrm{z}_2^k \text{ from } \mathrm{Q}_S \mathrm{z}_2^k = \mathrm{r}_2^k \ , \ \ \mathrm{z}_1^k := \mathrm{r}_1^k \ . \\ \mathrm{d} := \mathrm{A}\mathrm{r}_1^k \\ \beta_k^{(n)} := \langle \mathrm{d}, \mathrm{r}_1^k \rangle - \langle \bar{\mathrm{r}}_1^k, \mathrm{r}_1^k \rangle + \langle \mathrm{z}_2^k, \mathrm{r}_2^k \rangle, \ \ \beta_k := \beta_k^{(n)} / \beta_{k-1}^{(n)} \quad (\beta_0 := 0) \\ \mathrm{p}^k := \mathrm{z}^k + \beta_k \, \mathrm{p}^{k-1} \quad (\mathrm{p}^0 := \mathrm{z}^0) \ , \quad \mathrm{s}^k = \mathrm{d} + \beta_k \, s^{k-1} \quad (\mathrm{s}^0 := \mathrm{d}) \\ \begin{pmatrix} \mathrm{t} \\ \mathrm{u} \end{pmatrix} := \begin{pmatrix} \mathrm{s}^k + \mathrm{B}^T \mathrm{p}_2^k \\ \mathrm{B}\mathrm{p}_1^k \end{pmatrix} \ , \quad \begin{pmatrix} \mathrm{y} \\ \mathrm{z} \end{pmatrix} := \begin{pmatrix} \mathrm{Q}_A^{-1}\mathrm{t} \\ \mathrm{B}\mathrm{Q}_A^{-1}\mathrm{t} - \mathrm{u} \end{pmatrix} \\ \alpha_k^{(d)} := \langle \mathrm{y}, \mathrm{s}^k \rangle - \langle \mathrm{t}, \mathrm{p}_1^k \rangle + \langle \mathrm{z}, \mathrm{p}_2^k \rangle, \ \ \alpha_k := \beta_k^{(n)} / \alpha_k^{(d)} \\ \mathrm{v}^{k+1} := \mathrm{v}^k + \alpha_k \, \mathrm{p}^k \\ \mathrm{r}^{k+1} := \mathrm{r}^k - \alpha_k \begin{pmatrix} \mathrm{y} \\ \mathrm{z} \end{pmatrix} \ , \ \ \bar{\mathrm{r}}_1^{k+1} := \bar{\mathrm{r}}_1^k - \alpha_k \, \mathrm{t} \end{array} \right. \tag{3.9}$$

In this algorithm we do *not* need the evaluation of Q_A and only *one* evaluation of Q_A^{-1} per iteration. Furthermore, per iteration we need one evaluation of Q_S^{-1}, two matrix-vector products with B, one with B^T and one with A.

We want to use this algorithm to solve the problem (2.6). Then it is desirable that the iterands $\mathrm{v}^k = (\mathrm{x}^k, \mathrm{y}^k)$ are such that $\mathrm{y}^k \in \mathrm{e}^{\perp_M}$ holds. Therefore we assume that the preconditioner Q_S satisfies the consistency condition

$$\mathrm{Q}_S \mathrm{e} = \alpha \, \mathrm{M}\mathrm{e} \ , \quad \alpha \in \mathbb{R} \ . \tag{3.10}$$

From this assumption it follows that $\mathrm{y}^0 \in \mathrm{e}^{\perp_M}$ implies $\mathrm{y}^k \in \mathrm{e}^{\perp_M}$ for all $k \geq 1$. Moreover, if (3.10) holds then $\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}}$ is bijective on $\mathbb{R}^n \times \mathrm{e}^{\perp_M}$. We now formulate a

main convergence result for the BPCG algorithm due to [3, 29]. Let $\kappa_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}})$ be the spectral condition number of the matrix $\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}}$ on the subspace $\mathbb{R}^n \times \mathrm{e}^{\perp_M}$. This quantity plays a key role in the convergence analysis of the PCG method.

THEOREM 3.1. *Assume that* (3.1), (3.2), (3.3) *and* (3.10) *hold. For* $x \in \mathbb{R}$ *define*

$$g(x) = \frac{1}{2}(1+x) + \frac{1}{2}\sqrt{(1+x)^2 - 4x\Gamma_A^{-1}} \ .$$

For the BPCG method we have

$$\kappa_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}}) \leq \frac{\Gamma_S}{\gamma_S}\frac{\Gamma_A g(\Gamma_S) g(\gamma_S)}{\Gamma_S} \ . \tag{3.11}$$

Proof. We can apply theorem 4.1 in [29] in the subspace $\mathbb{R}^n \times \mathrm{e}^{\perp_M}$. This yields $\kappa_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}}) \leq (1-\rho_1)/(1-\rho_2)$, with

$$\rho_1 := \frac{2-(1+\Gamma_S)\Gamma_A}{2} - \sqrt{\frac{(2-(1+\Gamma_S)\Gamma_A)^2}{4} + \Gamma_A - 1} \ ,$$
$$\rho_2 := \frac{2-(1+\gamma_S)\Gamma_A}{2} + \sqrt{\frac{(2-(1+\gamma_S)\Gamma_A)^2}{4} + \Gamma_A - 1} \ .$$

Elementary manipulations show that $(1-\rho_1)/(1-\rho_2) = \frac{\Gamma_S}{\gamma_S}\frac{\Gamma_A g(\Gamma_S) g(\gamma_S)}{\Gamma_S}$ holds. ∎

REMARK 1. Note that $\Gamma_A \geq 1$ and thus $\frac{1}{2}(1+x) + \frac{1}{2}|1-x| \leq g(x) \leq \frac{1}{2}(1+x) + \frac{1}{2}|1+x|$. Hence, $1 \leq g(x) \leq 1+x$ if $x \in [0,1]$ and $x \leq g(x) \leq 1+x$ if $x \geq 1$. From this it follows that for the factor $C(\gamma_S, \Gamma_S) := \frac{g(\Gamma_S)g(\gamma_S)}{\Gamma_S}$ in (3.11) we have

$$\frac{1}{\Gamma_S} \leq C(\gamma_S,\Gamma_S) \leq \frac{4}{\Gamma_S} \quad \text{if } \ \Gamma_S \leq 1 \ ,$$
$$1 \leq C(\gamma_S,\Gamma_S) \leq 2 + \frac{2}{\Gamma_S} \leq 4 \quad \text{if } \ \gamma_S \leq 1 \leq \Gamma_S \ ,$$
$$\gamma_S \leq C(\gamma_S,\Gamma_S) \leq 3 + \gamma_S \leq 4\gamma_S \quad \text{if } \ 1 \leq \gamma_S \ .$$

From this we see that although the condition number of the preconditioned Schur complement $\kappa_*(\mathrm{Q}_S^{-1}\mathrm{S}) \leq \Gamma_S/\gamma_s$ is *independent* of the scaling of the preconditioner Q_S, the bound for the condition number $\kappa_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}})$ in (3.11) does depend on the scaling of Q_S. A poor scaling (i.e., $\Gamma_S \ll 1$ or $\gamma_S \gg 1$) leads to a large upper bound. We consider an example, which shows that this bound is sharp. Take $\mathrm{Q}_A = \Gamma_A^{-1}\mathrm{A}$ and $\mathrm{Q}_S = \Gamma^{-1}\mathrm{S}$ (on $\mathrm{e}^{\perp}$) with $\Gamma_A > 1$ and $\Gamma > 0$. Then (3.2) is fulfilled with $\gamma_S = \Gamma_S = \Gamma$. A simple computation shows that $\sigma_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}}) = \{\Gamma/g(\Gamma), \Gamma_A, \Gamma_A g(\Gamma)\}$ and thus $\kappa_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}}) = \Gamma_A\frac{g(\Gamma)^2}{\Gamma}$, which agrees with the bound in (3.11). Hence $\kappa_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}}) \geq \frac{\Gamma_A}{\Gamma}$ if $\Gamma \leq 1$ and $\kappa_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}}) \geq \Gamma_A\Gamma$ if $\Gamma \geq 1$. Note, however, that although $\kappa_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}})$ is large if $\Gamma \ll 1$ or $\Gamma \gg 1$ the BPCG method will converge in three iterations because $\sigma_*(\tilde{\mathrm{K}}^{-1}\hat{\mathrm{K}})$ contains only three elements.

3.2. Preconditioned minimal residual method. In this section we consider a preconditioned minimal residual (PMINRES) method for solving the discretized

Stokes problem. This class of methods has been developed in [20, 21, 23]. We consider a symmetric positive definite preconditioner

$$\tilde{\mathrm{K}} = \begin{pmatrix} \mathrm{Q}_A & 0 \\ 0 & \mathrm{Q}_S \end{pmatrix} \tag{3.12}$$

for K. Define the norm $\|\mathrm{v}\|_{\tilde{K}} := \langle \tilde{\mathrm{K}}\mathrm{v}, \mathrm{v}\rangle^{\frac{1}{2}}$ for $\mathrm{v} \in \mathbb{R}^{n+m}$. Given a starting vector v^0 with corresponding error $\mathrm{e}^0 := \mathrm{v}^* - \mathrm{v}^0$, then in the preconditioned MINRES method one computes the vector $\mathrm{v}^k \in \mathrm{v}^0 + \mathrm{span}\{\tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{e}^0, \ldots, (\tilde{\mathrm{K}}^{-1}\mathrm{K})^k\mathrm{e}^0\}$ which minimizes the preconditioned residual $\|\tilde{\mathrm{K}}^{-1}\mathrm{K}(\mathrm{v}^* - \mathrm{v})\|_{\tilde{\mathrm{K}}}$ over this subspace.
An efficient implementation of this method can be derived using the Lanczos algorithm and Givens rotations. For such an implementation we refer to the literature, e.g. [19, 14]. In an efficient implementation of this method one needs per iteration one evaluation of Q_A^{-1}, one evaluation of Q_S^{-1} and one matrix-vector product with K. We assume that the preconditioner Q_S satisfies the consistency condition (3.10). Then it follows that the approximations y^k of the discrete pressure are in the subspace $\mathrm{e}^{\perp_M}$ if the starting vector y^0 is in this subspace.

Note that for this algorithm one does *not* need the scaling condition (3.3) for Q_A. We give two main theoretical convergence results on the PMINRES algorithm from the literature (cf. [14, 21, 23]).

THEOREM 3.2. *Let* $\mathrm{v}^0 = (\mathrm{x}^0, \mathrm{y}^0)$ *be a given starting vector with* $\mathrm{y}^0 \in \mathrm{e}^{\perp_M}$. *For* v^k, $k \geq 0$, *computed in the PMINRES algorithm we define* $\tilde{\mathrm{r}}^k = \tilde{\mathrm{K}}^{-1}\big((\mathrm{f}, 0) - \mathrm{K}\mathrm{v}^k\big)$. *The following holds:*

$$\begin{aligned} \|\tilde{\mathrm{r}}^k\|_{\tilde{K}} &= \min_{p_k \in \mathcal{P}_k; p_k(0)=1} \|p_k(\tilde{\mathrm{K}}^{-1}\mathrm{K})\tilde{\mathrm{r}}^0\|_{\tilde{K}} \\ &\leq \min_{p_k \in \mathcal{P}_k; p_k(0)=1} \max_{\lambda \in \sigma_*(\tilde{\mathrm{K}}^{-1}\mathrm{K})} |p_k(\lambda)| \, \|\tilde{\mathrm{r}}^0\|_{\tilde{K}} \ , \end{aligned} \tag{3.13}$$

where $\sigma_*(\cdot)$ *denotes the spectrum on the subspace* $\mathbb{R}^n \times \mathrm{e}^{\perp_M}$.

THEOREM 3.3. *Assume that* (3.1), (3.2) *and* (3.10) *hold. For the spectrum of the preconditioned matrix* $\tilde{\mathrm{K}}^{-1}\mathrm{K}$ *we have:*

$$\begin{aligned} \sigma_*(\tilde{\mathrm{K}}^{-1}\mathrm{K}) \subset & \Big[\frac{1}{2}\big(\gamma_A - \sqrt{\gamma_A^2 + 4\Gamma_S\Gamma_A}\big) \, , \, \frac{1}{2}\big(\gamma_A - \sqrt{\gamma_A^2 + 4\gamma_S\gamma_A}\big) \Big] \\ & \bigcup \Big[\gamma_A \, , \, \frac{1}{2}\big(\Gamma_A + \sqrt{\Gamma_A^2 + 4\Gamma_S\Gamma_A}\big) \Big] \end{aligned} \tag{3.14}$$

From the results in these theorems it follows that the rate of convergence of the PMINRES method is robust with respect to variation in the parameters h and ξ if the spectral constants $\gamma_S, \Gamma_S, \gamma_A, \Gamma_A$ do not depend on these parameters. Moreover we can expect faster convergence if we have better preconditioners.

REMARK 2. Related to the scaling we consider the same example as in remark 1: $\mathrm{Q}_A = \Gamma_A^{-1}\mathrm{A}$, $\mathrm{Q}_S = \Gamma^{-1}\mathrm{S}$ (on $\mathrm{e}^{\perp}$). We then have $\gamma_A = \Gamma_A$, $\gamma_S = \Gamma_S = \Gamma$. A simple computation yields $\sigma_*(\tilde{\mathrm{K}}^{-1}\mathrm{K}) = \{ \frac{1}{2}\big(\Gamma_A - \sqrt{\Gamma_A^2 + 4\Gamma\Gamma_A}\big), \Gamma_A, \frac{1}{2}\big(\Gamma_A + \sqrt{\Gamma_A^2 + 4\Gamma\Gamma_A}\big) \}$. This shows that for this example the result in (3.14) is sharp. Furthermore, the scaling of Q_S has a significant influence on the endpoints of the intervals in (3.14). However, for this example, the PMINRES method will converge in three iterations since $\sigma_*(\tilde{\mathrm{K}}^{-1}\mathrm{K})$ contains only three elements.

3.3. An inexact Uzawa method. A basic method for saddle point problems is the Uzawa method. This method is closely related to the block factorization

$$\mathrm{K} = \begin{pmatrix} \mathrm{A} & 0 \\ \mathrm{B} & -\mathrm{I} \end{pmatrix} \begin{pmatrix} \mathrm{I} & \mathrm{A}^{-1}\mathrm{B}^T \\ 0 & \mathrm{S} \end{pmatrix}$$

Solving the problem $\mathrm{K}\begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} = \begin{pmatrix} \mathrm{f} \\ 0 \end{pmatrix}$ by block forward-backward substitution yields the equivalent problem:

$$\text{1. Solve } \ \mathrm{Az} = \mathrm{f} \ . \tag{3.15}$$

$$\text{2. Solve } \ \mathrm{Sy} = \mathrm{Bz} \ , \quad \mathrm{y} \in \mathrm{e}^{\perp_M} \ . \tag{3.16}$$

$$\text{3. Solve } \ \mathrm{Ax} = \mathrm{z} - \mathrm{B}^T\mathrm{y} \ . \tag{3.17}$$

In the Uzawa method one applies an iterative solver (e.g., CG) to the Schur complement system in step 2. The A-systems that occur in each iteration of this method and in the steps 1 and 3 are solved sufficiently accurate using some fast Poisson solver.
Here we reconsider a simple variant of this method which was introduced in [2]. Let Q_A be a preconditioner of A as in (3.1). We use this preconditioner in the steps 1 and 3 and also for the approximation of the Schur complement is step 2. For this we introduce the notation

$$\hat{\mathrm{S}} := \mathrm{B}\mathrm{Q}_A^{-1}\mathrm{B}^T \tag{3.18}$$

Note that $\hat{\mathrm{S}} : \mathrm{e}^{\perp_M} \to \mathrm{e}^{\perp}$ is bijective. We use a (nonlinear) approximate inverse of $\hat{\mathrm{S}}$ denoted by $\Psi : \mathrm{e}^{\perp} \to \mathrm{e}^{\perp_M}$. For each $\mathrm{w} \in \mathrm{e}^{\perp}$, $\Psi(\mathrm{w})$ is an approximation to the solution $\mathrm{z}^* \in \mathrm{e}^{\perp_M}$ of $\hat{\mathrm{S}}\mathrm{z} = \mathrm{w}$. We assume that

$$\|\Psi(\mathrm{w}) - \mathrm{z}^*\|_{\hat{\mathrm{S}}} \leq \delta \|\mathrm{z}^*\|_{\hat{\mathrm{S}}} \quad \text{for all } \ \mathrm{w} \in \mathrm{e}^{\perp} \tag{3.19}$$

holds with some $\delta < 1$. In our experiments Ψ will be the PCG method.
Let $(\mathrm{x}^k, \mathrm{y}^k)$ be a given approximation to the solution (x, y). Note that using the block factorization of K we get

$$\begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} = \begin{pmatrix} \mathrm{x}^k \\ \mathrm{y}^k \end{pmatrix} + \begin{pmatrix} \mathrm{I} & -\mathrm{A}^{-1}\mathrm{B}^T\mathrm{S}^{-1} \\ 0 & \mathrm{S}^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{A}^{-1} & 0 \\ \mathrm{B}\mathrm{A}^{-1} & -\mathrm{I} \end{pmatrix} \Big(\begin{pmatrix} \mathrm{f} \\ 0 \end{pmatrix} - \mathrm{K} \begin{pmatrix} \mathrm{x}^k \\ \mathrm{y}^k \end{pmatrix} \Big) \tag{3.20}$$

With $\mathrm{A}^{-1} \approx \mathrm{Q}_A^{-1}$, $\mathrm{S}^{-1}\mathrm{w} \approx \hat{\mathrm{S}}^{-1}\mathrm{w} \approx \Psi(\mathrm{w})$ and $\mathrm{r}_1^k := \mathrm{f} - \mathrm{A}\mathrm{x}^k - \mathrm{B}^T\mathrm{y}^k$ we obtain the (nonlinear) iterative method

$$\begin{aligned} \mathrm{x}^{k+1} &= \mathrm{x}^k + \mathrm{Q}_A^{-1}\mathrm{r}_1^k - \mathrm{Q}_A^{-1}\mathrm{B}^T\Psi\big(\mathrm{B}(\mathrm{Q}_A^{-1}\mathrm{r}_1^k + \mathrm{x}^k)\big) \\ \mathrm{y}^{k+1} &= \mathrm{y}^k + \Psi\big(\mathrm{B}(\mathrm{Q}_A^{-1}\mathrm{r}_1^k + \mathrm{x}^k)\big) \end{aligned} \tag{3.21}$$

We obtain the following algorithmic structure:

$$
\begin{cases}
\mathrm{v}^0 = \begin{pmatrix} \mathrm{x}^0 \\ \mathrm{y}^0 \end{pmatrix} \quad \text{a given starting vector,} \quad \mathrm{y}^0 \in \mathrm{e}^{\perp_M};\ \mathrm{r}_1^0 := \mathrm{f} - \mathrm{A}\mathrm{x}^0 - \mathrm{B}^T\mathrm{y}^0 \\
\text{for } k \geq 0 \ : \\
\qquad \mathrm{w} := \mathrm{x}^k + \mathrm{Q}_A^{-1}\mathrm{r}_1^k & (3.22) \\
\qquad \mathrm{z} := \Psi(\mathrm{Bw}) \in \mathrm{e}^{\perp_M} & (3.23) \\
\qquad \mathrm{x}^{k+1} := \mathrm{w} - \mathrm{Q}_A^{-1}\mathrm{B}^T\mathrm{z} & (3.24) \\
\qquad \mathrm{y}^{k+1} := \mathrm{y}^k + \mathrm{z} & (3.25) \\
\qquad \mathrm{r}_1^{k+1} := \mathrm{r}_1^k - \mathrm{A}(\mathrm{x}^{k+1} - \mathrm{x}^k) - \mathrm{B}^T\mathrm{z} & (3.26)
\end{cases}
$$

This is the same method as the one presented in section 4 in [2]. An obvious choice for $\Psi(\mathrm{Bw})$, which is also used in our experiments, is the following:

$$
\Psi(\mathrm{Bw}) = \begin{cases} \text{Result of } \ell \text{ PCG iterations with startvector 0 and} \\ \text{preconditioner } \mathrm{Q}_S \text{ applied to } \hat{\mathrm{S}}\mathrm{y} = \mathrm{Bw}. \end{cases} \tag{3.27}
$$

REMARK 3. We comment on some implementation issues and on the arithmetic costs per iteration of the algorithm (3.22)-(3.27). In each iteration of the PCG algorithm in (3.27) one needs an evaluation of Q_S^{-1} and of $\hat{\mathrm{S}} = \mathrm{B}\mathrm{Q}_A^{-1}\mathrm{B}^T$. Thus per PCG iteration we need one evaluation of Q_S^{-1}, one of Q_A^{-1} and a matrix-vector multiplication with B and with B^T. In the PCG method one computes approximations $\mathrm{z}^0 = 0, \mathrm{z}^1, \ldots, \mathrm{z}^\ell =: \mathrm{z}$. For these approximations we have a recursion of the form $\mathrm{z}^{i+1} = \mathrm{z}^i + \alpha_i \mathrm{p}^i$ (cf. algorithm (3.6)) and the vectors $\hat{\mathrm{S}}\mathrm{p}^i = \mathrm{B}(\mathrm{Q}_A^{-1}(\mathrm{B}^T\mathrm{p}^i))$ are computed. Hence the vectors $\bar{\mathrm{z}}^i := \mathrm{B}^T\mathrm{z}^i$, $\hat{\mathrm{z}}^i := \mathrm{Q}_A^{-1}\bar{\mathrm{z}}^i$ can be computed by cheap recursions (AXPY operations). Using this approach in the PCG algorithm we can obtain $\bar{\mathrm{z}}^\ell = \mathrm{B}^T\mathrm{z}^\ell = \mathrm{B}^T\mathrm{z}$ and $\hat{\mathrm{z}}^\ell = \mathrm{Q}_A^{-1}\bar{\mathrm{z}}^\ell = \mathrm{Q}_A^{-1}\mathrm{B}^T\mathrm{z}$ with negligible computational costs. Hence for the computation of $\mathrm{B}^T\mathrm{z}$ and $\mathrm{Q}_A^{-1}\mathrm{B}^T\mathrm{z}$ in (3.24), (3.26) we do not need significant arithmetic work. Summarizing, per outer iteration of the algorithm (3.22)-(3.27) we need $\ell + 1$ evaluations of Q_A^{-1}, ℓ evaluations of Q_S^{-1}, $\ell + 1$ matrix-vector multiplications with B, ℓ matrix-vector multiplications with B^T and one matrix-vector multiplication with A.

Clearly an important issue in the algorithm is the stopping criterion for the inner iteration. From the analysis presented in section 4 it follows that for an efficient method this stopping criterion should depend on the accuracy of the preconditioner Q_A. A very efficient Poisson solver is the multigrid method. Thus for the preconditioner Q_A we propose the following:

$$
\mathrm{Q}_A^{-1}\mathrm{w} = \begin{cases} \text{One symmetric multigrid } V\text{-cycle iteration, using one} \\ \text{symmetric Gauss-Seidel iteration for pre- and post-} \\ \text{smoothing, with startvector 0 applied to } \mathrm{Ax} = \mathrm{w}. \end{cases} \tag{3.28}
$$

The convergence factor per iteration of this method typically is between 0.1 and 0.2. From the analysis (cf. remark 4) we deduce the following two possibilities for the stopping criterion of the PCG method:

- Let r^i be the residual in the PCG algorithm. We stop the PCG iteration if

$$
\|\mathrm{r}^\ell\|_2 \leq \sigma_{\mathrm{acc}}\|\mathrm{r}^0\|_2 \quad \text{is satisfied, with a given } \sigma_{\mathrm{acc}} \in [0.2, 0.6]. \tag{3.29}
$$

- We use a fixed low number of iterations: $\ell \in [1, 4]$.

4. Analysis of the inexact Uzawa algorithm. The inexact Uzawa method has been analyzed in [2, 29]. Here we present an alternative analysis. This analysis is much simpler as those presented in [2] and [29] and the result is different. In remark 7 we discuss the main differences with the results from [2, 29].

For simplicity we assume that for the preconditioner Q_A we have

$$\mathrm{A} \leq \mathrm{Q}_A, \quad \text{i.e. } \gamma_A \mathrm{Q}_A \leq \mathrm{A} \leq \mathrm{Q}_A \tag{4.1}$$

We make the scaling assumption $\mathrm{A} \leq \mathrm{Q}_A$ because, opposite to the scaling condition in (3.3), it is fulfilled for the multigrid preconditioner in (3.28). Moreover, it improves the presentation of the analysis. However, this *scaling assumption is not essential neither for the algorithm nor for the convergence analysis.*
In the analysis we use the following natural norms:

$$\|\mathrm{x}\|_Q := \|\mathrm{Q}_A^{\frac{1}{2}}\mathrm{x}\| = \langle \mathrm{Q}_A \mathrm{x}, \mathrm{x}\rangle^{\frac{1}{2}} \quad \text{for} \quad \mathrm{x} \in \mathbb{R}^n, \quad \|\mathrm{y}\|_{\hat{\mathrm{S}}} := \|\hat{\mathrm{S}}^{\frac{1}{2}}\mathrm{y}\| = \langle \hat{\mathrm{S}}\mathrm{y}, \mathrm{y}\rangle^{\frac{1}{2}} \quad \text{for} \quad \mathrm{y} \in \mathrm{e}^{\perp_M}$$

For the error in algorithm (3.22)-(3.26) we use the notation

$$\mathrm{e}^k = \begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} - \begin{pmatrix} \mathrm{x}^k \\ \mathrm{y}^k \end{pmatrix} =: \begin{pmatrix} \mathrm{e}_1^k \\ \mathrm{e}_2^k \end{pmatrix} .$$

Note that $\mathrm{e}_2^k \in \mathrm{e}^{\perp_M}$ holds.

LEMMA 4.1. *For* w *as in* (3.22) *we have the bound*

$$\langle \hat{\mathrm{S}}^{-1}\mathrm{Bw}, \mathrm{Bw}\rangle^{\frac{1}{2}} \leq (1-\gamma_A)\|\mathrm{e}_1^k\|_Q + \|\mathrm{e}_2^k\|_{\hat{\mathrm{S}}} .$$

Proof. For $\mathrm{y} \in \mathrm{range}(\hat{\mathrm{S}}) = \mathrm{range}(\mathrm{B})$ define $\|\mathrm{y}\|_{\hat{\mathrm{S}}^{-1}} = \langle \hat{\mathrm{S}}^{-1}\mathrm{y}, \mathrm{y}\rangle^{\frac{1}{2}}$. Note that for the exact discrete solution x we have $\mathrm{Bx} = 0$. Using this and the definition in (3.22) we get

$$\mathrm{Bw} = \mathrm{Bx}^k - \mathrm{Bx} + \mathrm{BQ}_A^{-1}(\mathrm{Ae}_1^k + \mathrm{B}^T\mathrm{e}_2^k) = -\mathrm{B}(\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{e}_1^k + \hat{\mathrm{S}}\mathrm{e}_2^k .$$

Hence,

$$\begin{aligned} \langle \hat{\mathrm{S}}^{-1}\mathrm{Bw}, \mathrm{Bw}\rangle^{\frac{1}{2}} = \|\mathrm{Bw}\|_{\hat{\mathrm{S}}^{-1}} &\leq \|\mathrm{B}(\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{e}_1^k\|_{\hat{\mathrm{S}}^{-1}} + \|\hat{\mathrm{S}}\mathrm{e}_2^k\|_{\hat{\mathrm{S}}^{-1}} \\ &\leq \|\hat{\mathrm{S}}^{-\frac{1}{2}}\mathrm{B}(\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{Q}_A^{-\frac{1}{2}}\|\|\mathrm{Q}_A^{\frac{1}{2}}\mathrm{e}_1^k\| + \|\hat{\mathrm{S}}^{\frac{1}{2}}\mathrm{e}_2^k\| . \end{aligned}$$

Now note that

$$\|\hat{\mathrm{S}}^{-\frac{1}{2}}\mathrm{B}(\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{Q}_A^{-\frac{1}{2}}\| \leq \|\hat{\mathrm{S}}^{-\frac{1}{2}}\mathrm{BQ}_A^{-\frac{1}{2}}\|\|\mathrm{Q}_A^{\frac{1}{2}}(\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{Q}_A^{-\frac{1}{2}}\|$$

and

$$\begin{aligned} \|\hat{\mathrm{S}}^{-\frac{1}{2}}\mathrm{BQ}_A^{-\frac{1}{2}}\|^2 &= \rho(\hat{\mathrm{S}}^{-\frac{1}{2}}\mathrm{BQ}_A^{-1}\mathrm{B}^T\hat{\mathrm{S}}^{-\frac{1}{2}}) = \rho(\mathrm{I}) = 1 , \\ \|\mathrm{Q}_A^{\frac{1}{2}}(\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{Q}_A^{-\frac{1}{2}}\| &= \rho(\mathrm{I} - \mathrm{Q}_A^{-\frac{1}{2}}\mathrm{AQ}_A^{-\frac{1}{2}}) \leq 1 - \gamma_A . \end{aligned}$$

This completes the proof. ∎

We now formulate the main convergence result.

THEOREM 4.2. *Consider the inexact Uzawa method* (3.22)-(3.27) *with* Ψ *such that* (3.19) *holds. For the error* $\mathrm{e}^k = (\mathrm{e}_1^k, \mathrm{e}_2^k)$ *we have the bounds*

$$\|\mathrm{e}_1^{k+1}\|_Q \leq (1-\gamma_A)\|\mathrm{e}_1^k\|_Q + \|\mathrm{e}_2^{k+1}\|_{\hat{\mathrm{S}}} \ , \tag{4.2}$$

$$\|\mathrm{e}_2^{k+1}\|_{\hat{\mathrm{S}}} \leq (1-\gamma_A)(1+\delta)\|\mathrm{e}_1^k\|_Q + \delta\|\mathrm{e}_2^k\|_{\hat{\mathrm{S}}} \tag{4.3}$$

with γ_A *from* (4.1) *and* δ *from* (3.19).

Proof. For the error component e_1^{k+1} we have the relations

$$\begin{aligned}
\mathrm{e}_1^{k+1} &= \mathrm{x} - \mathrm{x}^{k+1} = \mathrm{x} - \mathrm{w} + \mathrm{Q}_A^{-1}\mathrm{B}^T\mathrm{z} \\
&= \mathrm{x} - \mathrm{x}^k - \mathrm{Q}_A^{-1}(\mathrm{A}\mathrm{e}_1^k + \mathrm{B}^T\mathrm{e}_2^k - \mathrm{B}^T\mathrm{z}) \\
&= (\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{e}_1^k - \mathrm{Q}_A^{-1}\mathrm{B}^T(\mathrm{e}_2^k - \mathrm{z}) \\
&= (\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{e}_1^k - \mathrm{Q}_A^{-1}\mathrm{B}^T\mathrm{e}_2^{k+1} \ .
\end{aligned}$$

Hence,

$$\|\mathrm{e}_1^{k+1}\|_Q \leq \|\mathrm{I} - \mathrm{Q}_A^{-\frac{1}{2}}\mathrm{A}\mathrm{Q}_A^{-\frac{1}{2}}\|\|\mathrm{e}_1^k\|_Q + \|\mathrm{Q}_A^{-\frac{1}{2}}\mathrm{B}^T\hat{\mathrm{S}}^{-\frac{1}{2}}\|\|\mathrm{e}_2^{k+1}\|_{\hat{\mathrm{S}}}$$

In combination with

$$\|\mathrm{I} - \mathrm{Q}_A^{-\frac{1}{2}}\mathrm{A}\mathrm{Q}_A^{-\frac{1}{2}}\| \leq 1 - \gamma_A \ , \quad \|\mathrm{Q}_A^{-\frac{1}{2}}\mathrm{B}^T\hat{\mathrm{S}}^{-\frac{1}{2}}\| = 1 \ .$$

this proves the inequality in (4.2). For the error component e_2^{k+1} we obtain

$$\mathrm{e}_2^{k+1} = \mathrm{y} - \mathrm{y}^{k+1} = \mathrm{e}_2^k - \mathrm{z} = (\mathrm{e}_2^k - \hat{\mathrm{S}}^{-1}\mathrm{B}\mathrm{w}) + (\hat{\mathrm{S}}^{-1}\mathrm{B}\mathrm{w} - \mathrm{z}) \ . \tag{4.4}$$

Using $\mathrm{B}\mathrm{x} = 0$ we get

$$\begin{aligned}
\|\mathrm{e}_2^k - \hat{\mathrm{S}}^{-1}\mathrm{B}\mathrm{w}\|_{\hat{\mathrm{S}}} &= \|\mathrm{e}_2^k - \hat{\mathrm{S}}^{-1}\mathrm{B}(\mathrm{x}^k + \mathrm{Q}_A^{-1}\mathrm{A}\mathrm{e}_1^k + \mathrm{Q}_A^{-1}\mathrm{B}^T\mathrm{e}_2^k)\|_{\hat{\mathrm{S}}} \\
&= \|\mathrm{e}_2^k + \hat{\mathrm{S}}^{-1}\mathrm{B}(\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{e}_1^k - \hat{\mathrm{S}}^{-1}\mathrm{B}\mathrm{Q}_A^{-1}\mathrm{B}^T\mathrm{e}_2^k\|_{\hat{\mathrm{S}}} \\
&= \|\hat{\mathrm{S}}^{-1}\mathrm{B}(\mathrm{I} - \mathrm{Q}_A^{-1}\mathrm{A})\mathrm{e}_1^k\|_{\hat{\mathrm{S}}} \\
&\leq \|\hat{\mathrm{S}}^{-\frac{1}{2}}\mathrm{B}\mathrm{Q}_A^{-\frac{1}{2}}\|\|\mathrm{I} - \mathrm{Q}_A^{-\frac{1}{2}}\mathrm{A}\mathrm{Q}_A^{-\frac{1}{2}}\|\|\mathrm{e}_1^k\|_Q \\
&\leq (1-\gamma_A)\|\mathrm{e}_1^k\|_Q \ .
\end{aligned} \tag{4.5}$$

Furthermore, using (3.19) and lemma 4.1 we also have

$$\begin{aligned}
\|\hat{\mathrm{S}}^{-1}\mathrm{B}\mathrm{w} - \mathrm{z}\|_{\hat{\mathrm{S}}} &= \|\hat{\mathrm{S}}^{-1}\mathrm{B}\mathrm{w} - \Psi(\mathrm{B}\mathrm{w})\|_{\hat{\mathrm{S}}} \leq \delta\,\|\hat{\mathrm{S}}^{-1}\mathrm{B}\mathrm{w}\|_{\hat{\mathrm{S}}} \\
&= \delta\,\langle\hat{\mathrm{S}}^{-1}\mathrm{B}\mathrm{w}, \mathrm{B}\mathrm{w}\rangle^{\frac{1}{2}} \leq \delta\big((1-\gamma_A)\|\mathrm{e}_1^k\|_Q + \|\mathrm{e}_2^k\|_{\hat{\mathrm{S}}}\big) \ .
\end{aligned} \tag{4.6}$$

Combination of the results in (4.4), (4.5), (4.6) proves the inequality in (4.3). ∎

As a simple consequence of this theorem we obtain the following convergence result:

THEOREM 4.3. *Define*

$$\mu_A := 1 - \gamma_A, \quad g(\mu_A, \delta) := 2\mu_A + \delta(1+\mu_A) \ . \tag{4.7}$$

Consider the inexact Uzawa method (3.22)-(3.27) *with* Ψ *such that* (3.19) *holds. For the error* $\mathrm{e}^k = (\mathrm{e}_1^k, \mathrm{e}_2^k)$ *we have the bounds*

$$\max\left\{\|\mathrm{e}_1^{k+1}\|_Q, \|\mathrm{e}_2^{k+1}\|_{\hat{\mathrm{S}}}\right\} \leq g(\mu_A,\delta)\, \max\left\{\|\mathrm{e}_1^{k}\|_Q, \|\mathrm{e}_2^{k}\|_{\hat{\mathrm{S}}}\right\}\,, \tag{4.8}$$

$$\|\mathrm{e}_1^k\|_Q + \|\mathrm{e}_2^k\|_{\hat{\mathrm{S}}} \leq 3\frac{1}{2}\Big(\frac{g(\mu_A,\delta)+\sqrt{g(\mu_A,\delta)^2-4\mu_A\delta}}{2}\Big)^k (\|\mathrm{e}_1^0\|_Q + \|\mathrm{e}_2^0\|_{\hat{\mathrm{S}}})\,. \tag{4.9}$$

Proof. Define the matrix

$$\mathrm{C} = \begin{pmatrix} \mu_A(2+\delta) & \delta \\ \mu_A(1+\delta) & \delta \end{pmatrix}\,.$$

Due to theorem 4.2 we obtain

$$\begin{pmatrix} \|\mathrm{e}_1^{k+1}\|_Q \\ \|\mathrm{e}_2^{k+1}\|_{\hat{\mathrm{S}}} \end{pmatrix} \leq \mathrm{C} \begin{pmatrix} \|\mathrm{e}_1^{k}\|_Q \\ \|\mathrm{e}_2^{k}\|_{\hat{\mathrm{S}}} \end{pmatrix}\,.$$

From $\|\mathrm{C}\|_\infty = g(\mu_A,\delta)$ we obtain the result in (4.8). A simple (MAPLE) computation yields the eigenvector decomposition $\mathrm{C} = \mathrm{V}\mathrm{D}\mathrm{V}^{-1}$ with a diagonal matrix D. From this we get

$$\rho(\mathrm{C}) = \frac{1}{2}\Big(g(\mu_A,\delta)+\sqrt{g(\mu_A,\delta)^2-4\mu_A\delta}\Big)$$
$$\max_{0<\mu_A,\delta<1} \|\mathrm{V}\|_1\|\mathrm{V}^{-1}\|_1 \leq 3\frac{1}{2}$$

From this the result in (4.9) follows. ∎

Corollary 4.4. Clearly, the bound for the contraction factor in (4.8) and the bound for the asymptotic convergence factor in (4.9) depend only on μ_A and δ and the bounds are monotonic functions of these parameters. Note that for $\mu_A \to 0$ we obtain the contraction factor of the exact Uzawa method: $g(0,\delta) = \delta$. We also have $g(\mu_A,\delta) \geq \frac{1}{2}\big(g(\mu_A,\delta)+\sqrt{g(\mu_A,\delta)^2-4\mu_A\delta}\big)$ and

$$g(\mu_A,\delta) < 1 \quad \text{iff} \quad 0 \leq \delta < \frac{1-2\mu_A}{1+\mu_A}\,, \tag{4.10}$$

$$\frac{1}{2}\big(g(\mu_A,\delta)+\sqrt{g(\mu_A,\delta)^2-4\mu_A\delta}\big) < 1 \quad \text{iff} \quad 0 \leq \delta < 1-2\mu_A\,. \tag{4.11}$$

Hence, for $\mu_A < \frac{1}{2}$ and δ sufficiently small (as quantified in (4.10), (4.11)) we have a convergent method.

Remark 4. We comment on the important special case where we take for Q_A^{-1} a symmetric multigrid V-cycle as in (3.28). Then μ_A is the contraction number (w.r.t. the norm $\|\cdot\|_A$) of this multigrid method. It is known from numerical experiments that typically we have $0.1 \leq \mu_A \leq 0.2$. The result in (4.11) shows that for $\mu_A \approx 0.15$ we have a convergent method if for the accuracy of the inner method we take $\delta \lesssim 0.7$. Clearly it is not efficient to use a very small value for δ. We comment on the choice of this parameter δ. We consider the case $\mu_A = 0.1$. For the contraction number in (4.8) we then have $g(0.1,\delta) = 0.2 + 1.1\delta$. For a given $\mathrm{eps} \ll 1$ let k be such that

$$g(0.1,\delta)^k \approx \mathrm{eps}\,.$$

Then

$$k \approx \frac{\ln(\text{eps})}{\ln(0.2 + 1.1\delta)}$$

holds. We use the ansatz (cf. (3.27)) $\delta = \beta^\ell$ with some $\beta < 1$. From remark 3 it follows that the arithmetic costs are dominated by $k(\ell+1)$ evaluations of Q_A^{-1}. Hence, one wants to minimize

$$k(\ell+1) \approx \frac{\ln(\text{eps})}{\ln \beta} \frac{\ln \delta + \ln \beta}{\ln(0.2 + 1.1\delta)} =: k_{\text{Uzawa}} K(\beta, \delta)$$

where $k_{\text{Uzawa}} := \ln(\text{eps})/\ln \beta$ is the number of iterations of the exact Uzawa method ($\mu_A = 0$). For different β values the function $\delta \to K(\beta, \delta)$ is given in figure 4.1.

We see that an almost optimal value for δ is obtained in a very broad range. For

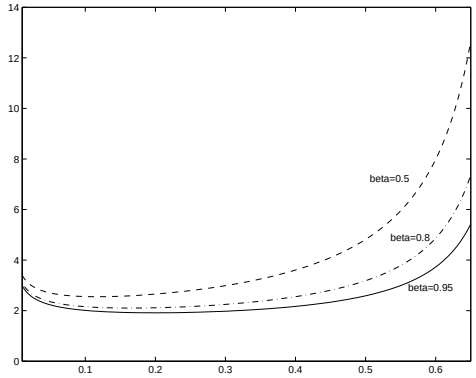

FIG. 4.1. *Function $\delta \to K(\beta, \delta)$ for $\beta = 0.5, 0.8, 0.95$*

other μ_A values (< 0.2) one observes a similar behaviour.

In the PCG method (3.27) we use a preconditioner Q_S of $\hat{\mathrm{S}}$. We assume that (3.2) and (3.10) hold. Using (4.1) it follows that

$$\gamma_A \gamma_S \mathrm{Q}_S \leq \hat{\mathrm{S}} \leq \Gamma_S \mathrm{Q}_S \tag{4.12}$$

holds.

REMARK 5. Note that opposite to the BPCG method treated in section 3.1 and the PMINRES method treated in section 3.2 the scaling of the preconditioner Q_S does *not* influence the rate of convergence of the inexact Uzawa method (cf. the remarks 1 and 2).

In the convergence analysis we use the norm

$$\sqrt{\|\mathrm{x}\|_Q^2 + \|\mathrm{y}\|_{\hat{\mathrm{S}}}^2} =: |||(\mathrm{x}, \mathrm{y})||| \quad \text{on} \quad \mathbb{R}^n \times \mathrm{e}^{\perp_M} \tag{4.13}$$

REMARK 6. If we consider a stationary Stokes problem then the norm in (4.13) is very natural. In this setting one usually takes $\mathrm{Q}_S = \mathrm{M}$ (pressure mass matrix as in (2.3)). Then the constants $\gamma_A, \gamma_S, \Gamma_S$ can be shown to be independent of h and from this one obtains

$$c_1 \big(\|\nabla J_V \mathrm{x}\|^2 + \|J_M \mathrm{y}\|^2 \big)^{\frac{1}{2}} \leq |||(\mathrm{x}, \mathrm{y})||| \leq c_2 \big(\|\nabla J_V \mathrm{x}\|^2 + \|J_M \mathrm{y}\|^2 \big)^{\frac{1}{2}} \quad \forall\, (\mathrm{x}, \mathrm{y}) \in \mathbb{R}^n \times \mathrm{e}^{\perp_M}$$

with constants $c_1 > 0$ and c_2 independent of h. Now note that the norm $(\mathbf{u}_h, p_h) \to \big(\|\nabla \mathbf{u}_h\|^2 + \|p_h\|^2\big)^{\frac{1}{2}}$ is the canonical one on $\mathbf{V}_h \times M_h$.

We indicate how the norm of the error, $|||\mathrm{e}^k||| = (\|\mathrm{e}_1^k\|_Q + \|\mathrm{e}_2^k\|_{\hat{\mathrm{S}}}^2)^{\frac{1}{2}}$, can be estimated from quantities that are available in the inexact Uzawa algorithm (3.22)-(3.26). Let $\tilde{\mathrm{K}}$ be the symmetric positive definite block diagonal matrix as in (3.12) with Q_A such that (4.1) holds. From (4.12) we get

$$\gamma_A \gamma_S \|\mathrm{v}\|_{\tilde{\mathrm{K}}}^2 \leq |||\mathrm{v}|||^2 \leq \Gamma_S \|\mathrm{v}\|_{\tilde{\mathrm{K}}}^2 \quad \text{for all} \;\; \mathrm{v} \in \mathbb{R}^n \times \mathrm{e}^{\perp_M} \tag{4.14}$$

Note that $\tilde{\mathrm{K}} : \; \mathbb{R}^n \times \mathrm{e}^{\perp_M} \to \mathbb{R}^n \times \mathrm{e}^{\perp}$ is bijective. From this and from theorem 3.3 with $\Gamma_A = 1$ we obtain

$$\begin{aligned}
\|\tilde{\mathrm{K}}^{-\frac{1}{2}} \mathrm{K} \tilde{\mathrm{K}}^{-\frac{1}{2}} \mathrm{w}\| &\leq \frac{1}{2}\big(1 + \sqrt{1 + 4\Gamma_S}\big)\|\mathrm{w}\| =: c_1(\Gamma_S)\|\mathrm{w}\| \quad \text{for all} \;\; \mathrm{w} \in \mathbb{R}^n \times \mathrm{e}^{\perp} \\
\|\tilde{\mathrm{K}}^{-\frac{1}{2}} \mathrm{K} \tilde{\mathrm{K}}^{-\frac{1}{2}} \mathrm{w}\| &\geq \min\big\{\gamma_A, \frac{1}{2}\big(\sqrt{\gamma_A^2 + 4\gamma_S\gamma_A} - \gamma_A\big)\big\}\|\mathrm{w}\| \\
&=: c_2(\gamma_S, \gamma_A)\|\mathrm{w}\| \quad \text{for all} \;\; \mathrm{w} \in \mathbb{R}^n \times \mathrm{e}^{\perp}
\end{aligned}$$

This then yields

$$\begin{aligned}
|||\mathrm{v}|||_{\tilde{\mathrm{K}}}^2 &= \frac{\langle \tilde{\mathrm{K}}\mathrm{v}, \mathrm{v}\rangle}{\langle \tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{v}, \mathrm{K}\mathrm{v}\rangle} \langle \tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{v}, \mathrm{K}\mathrm{v}\rangle \leq c_2(\gamma_S, \gamma_A)^{-2} \langle \tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{v}, \mathrm{K}\mathrm{v}\rangle \;\; \forall \;\; \mathrm{v} \in \mathbb{R}^n \times \mathrm{e}^{\perp_M} \\
|||\mathrm{v}|||_{\tilde{\mathrm{K}}}^2 &= \frac{\langle \tilde{\mathrm{K}}\mathrm{v}, \mathrm{v}\rangle}{\langle \tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{v}, \mathrm{K}\mathrm{v}\rangle} \langle \tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{v}, \mathrm{K}\mathrm{v}\rangle \geq c_1(\Gamma_S)^{-1} \langle \tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{v}, \mathrm{K}\mathrm{v}\rangle \;\; \forall \;\; \mathrm{v} \in \mathbb{R}^n \times \mathrm{e}^{\perp_M}
\end{aligned}$$

In combination with (4.14) we then obtain

$$c_2(\gamma_S, \gamma_A)^2 \Gamma_S^{-1} |||\mathrm{v}||| \leq \langle \tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{v}, \mathrm{K}\mathrm{v}\rangle \leq c_1(\Gamma_S)^2 \gamma_A^{-1} \gamma_S^{-1} |||\mathrm{v}||| \;\; \forall \;\; \mathrm{v} \in \mathbb{R}^n \times \mathrm{e}^{\perp_M}$$

Hence we can use $\langle \tilde{\mathrm{K}}^{-1}\mathrm{K}\mathrm{e}^k, \mathrm{K}\mathrm{e}^k\rangle$ as a measure for the error $|||\mathrm{e}^k|||^2$. Let $(\mathrm{r}_1^k, \mathrm{r}_2^k) = \mathrm{r}^k = \mathrm{K}\mathrm{e}^k$ be the residual. Then we have

$$|||\mathrm{e}^k|||^2 \sim \langle \tilde{\mathrm{K}}^{-1}\mathrm{r}^k, \mathrm{r}^k\rangle = \langle \mathrm{Q}_A^{-1}\mathrm{r}_1^k, \mathrm{r}_1^k\rangle + \langle \mathrm{Q}_S^{-1}\mathrm{r}_2^k, \mathrm{r}_2^k\rangle \tag{4.15}$$

Note that $\mathrm{Q}_A^{-1}\mathrm{r}_1^k$ is available from (3.22) and $\mathrm{Q}_S^{-1}\mathrm{r}_2^k$ can be computed (or approximated accurately) with relatively low costs.

REMARK 7. We discuss the main differences of our convergence result and those in [2, 29]. In [2], theorem 10, and [29], theorem 5.5, an inequality of the form

$$|||\mathrm{e}^{k+1}|||_* \leq g_B(\mu_A, \delta)|||\mathrm{e}^k|||_*$$

is proved. The norms $|||\cdot|||_*$ used in [2] and [29] are different and also differ from the norm $|||\cdot|||$ in (4.13) that is used in our analysis. A not so nice property of the norms $|||\cdot|||_*$ is that they depend on the (nonlinear) function Ψ. In case of the PCG method as in (3.27) the norm then depends on the iteration index ℓ. In [2] a bound of the form

$$g_B(\mu_A, \delta) = \max\big\{\mu_A, \frac{2\delta}{1-\delta}\big\} \tag{4.16}$$

is proved. In [29] a more general setting is considered. However, in the special case of the inexact Uzawa method (3.22)-(3.26) the bound has a similar form as in (4.16). If we compare $g_B(\mu_A, \delta)$ with our bound $g(\mu_A, \delta) = 2\mu_A + \delta(1 + \mu_A)$ then it is clear that the bounds are quite different. For $g_B(\mu_A, \delta) < 1$ one needs $\mu_A < 1$ and $\delta < \frac{1}{3}$, whereas for $g(\mu_A, \delta) < 1$ we need $\mu_A < \frac{1}{2}$ and $\delta < \frac{1-2\mu_A}{1+\mu_A}$. In the extreme case $\delta = 0$ we have $g_B(\mu_A, 0) = \frac{1}{2} g(\mu_A, 0)$, whereas if $\mu_A = 0$ we get $g(0, \delta) = \frac{1}{2}(1 - \delta) g_B(0, \delta)$. We are particularly interested in the case where for Q_A we take a multigrid method. Therefore, as an example we consider $\mu_A = 0.1$. Then we have

$$\begin{aligned} g(0.1, \delta) < 1 \quad &\text{if} \quad \delta < 0.72 \\ g_B(0.1, \delta) < 1 \quad &\text{if} \quad \delta < \frac{1}{3} \\ g(0.1, \delta) < g_B(0.1, \delta) \quad &\text{if} \quad \delta > 0.16 \end{aligned}$$

The bound $g_B(\mu_A, \delta)$ is better for "small" δ-values, whereas the bound $g(\mu_A, \delta)$ is better for "larger" δ-values. In our opinion, the bound $g(\mu_A, \delta)$ gives a better explanation of the observation (that has been made in many numerical experiments) that if one uses a very good preconditioner Q_A (such as multigrid) then even with (very) low accuracy in the inner iteration the inexact Uzawa method converges.

5. Preconditioners. In this section we discuss the choice of the preconditioners Q_A and Q_S. In particular we want that for the preconditioners the constants γ_S, Γ_S, γ_A, Γ_A in the spectral equivalences (3.2) and (3.1) are independent of the parameters h and ξ.
It is known that if for Q_A we take a symmetric multigrid V-cycle, $\mathrm{Q}_A = \mathrm{Q}_{MG}$, then we have (cf. [6, 15, 16])

$$(1 - \mu_{MG})\mathrm{Q}_{MG} \leq \mathrm{A} \leq \mathrm{Q}_{MG}, \tag{5.1}$$

with a constant $\mu_{MG} < 1$ independent of h. In [18] it is shown that for H^2-regular problems μ_{MG} is also independent of ξ. Note if we use a scaling $\mathrm{Q}_{MG} \to \alpha \mathrm{Q}_{MG}$ with $\alpha < 1 - \mu_{MG}$ then we obtain a preconditioner that satisfies (3.3). In the numerical experiments in section 6 we will use a (scaled) symmetric multigrid V-cycle as in (3.28).

It is well-known that for the case $\xi = 0$ the pressure mass matrix M is an appropriate preconditioner for S. From the discrete LBB property (2.1) and the fact that for arbitrary $\mathrm{y} \in \mathbb{R}^m$, we have (if $\xi = 0$)

$$\langle \mathrm{Sy}, \mathrm{y} \rangle = \sup_{\mathbf{u}_h \in \mathbf{V}_h} \frac{(\operatorname{div} \mathbf{u}_h, J_M \mathrm{y})^2}{\|\nabla \mathbf{u}_h\|^2} \ , \tag{5.2}$$

it follows that M is uniformly (w.r.t. h) spectrally equivalent to S on the subspace $\mathrm{e}^{\perp_M}$:

$$\hat{\beta}^2 \mathrm{M} \leq \mathrm{S} \leq d\,\mathrm{M} \quad \text{on} \quad \mathrm{e}^{\perp_M} \tag{5.3}$$

with $\hat{\beta}$ the LBB constant from (2.1). The *lumped* mass matrix $\bar{\mathrm{M}}$ is the diagonal matrix with diagonal entries $\bar{\mathrm{M}}_{ii} = \sum_{j=1}^{m} \mathrm{M}_{ij}$. In [25] it is proved that this matrix is uniformly in h spectrally equivalent to M and thus also (for $\xi = 0$) to S on the subspace $\mathrm{e}^{\perp_M}$. A further elementary observation is

$$\bar{\mathrm{M}}\mathrm{e} = \mathrm{M}\mathrm{e} \tag{5.4}$$

Hence, both for $\mathrm{Q}_S = \mathrm{M}$ and $\mathrm{Q}_S = \bar{\mathrm{M}}$ the consistency condition (3.10) is fulfilled.

For the case $\xi \gg 1$ the (lumped) mass matrix is not an appropriate preconditioner for the Schur complement. In [4, 8, 17] robust (w.r.t. h and ξ) preconditioning strategies are introduced. We briefly explain the method from [4].

For $g \in L^2(\Omega)$ consider the Neumann problem: find $w \in H^1(\Omega) \cap M$ such that

$$(\nabla w, \nabla \phi) = (g, \phi) \quad \text{for all} \ \ \phi \in H^1(\Omega) \cap M \tag{5.5}$$

Let T_h be the stiffness matrix of the Galerkin discretization of this problem in $M_h^+ \subset H^1(\Omega)$:

$$\langle \mathrm{T}_h \mathrm{x}, \mathrm{y} \rangle = (\nabla J_M \mathrm{x}, \nabla J_M \mathrm{y}) \quad \text{for all} \ \ \mathrm{x}, \mathrm{y} \in \mathbb{R}^m$$

Note that $\ker(\mathrm{T}_h) = \mathrm{span}(\mathrm{e})$ and that $\mathrm{T}_h : \ \mathrm{e}^{\perp_M} \to \mathrm{e}^{\perp}$ is bijective. We define $\tilde{\mathrm{Q}}_S^{-1} : \mathrm{e}^{\perp} \to \mathrm{e}^{\perp_M}$ by:

$$\tilde{\mathrm{Q}}_S^{-1} = \begin{cases} \mathrm{M}^{-1} + \xi \mathrm{T}_h^{-1}, & \text{if} \quad \xi \leq h^{-2}, \\ \xi h^2 \mathrm{M}^{-1} + \xi \mathrm{T}_h^{-1}, & \text{if} \quad h^{-2} \leq \xi \end{cases} \tag{5.6}$$

In the analysis it is convenient to extend $\tilde{\mathrm{Q}}_S^{-1}$ on $\mathbb{R}^m$ by defining $\tilde{\mathrm{Q}}_S^{-1}\mathrm{Me} := \mathrm{e}$. In practice, however, this extension is not needed. Note, that for $\xi = 0$ we get $\tilde{\mathrm{Q}}_S = \mathrm{M}$.

In [4] it is shown that $\tilde{\mathrm{Q}}_S$ is uniformly in h and ξ spectrally equivalent to S. We rephrase the main theorem 4.1 of [4] (for the special case $S_h = W_h$ in [4]):

THEOREM 5.1. *Assume that for $\xi = 0$ and for all $\mathbf{f} \in L_2(\Omega)^d$ the solution $(\mathbf{u}, p)^T$ of (1.1) satisfies the regularity condition $\|\mathbf{u}\|_2 + \|p\|_1 \leq C\|\mathbf{f}\|$ with a constant C independent of $\mathbf{f}$. Assume that the problem (5.5) is H^2-regular. Furthermore, for the pair of finite element spaces $(\mathbf{V}_h, M_h)$ we assume that the LBB condition (2.1) is satisfied, that $M_h^+ \subset H^1(\Omega)$, $1 \in M_h^+$ and that the following approximation and inverse properties hold with constants C independent of h:*

$$\begin{aligned} \inf_{\mathbf{w} \in \mathbf{V}_h} \|\mathbf{v} - \mathbf{w}\|_1 &\leq Ch\|\mathbf{v}\|_2 && \text{for all} \quad \mathbf{v} \in (H^2(\Omega) \cap H_0^1(\Omega))^d, \\ \inf_{q \in M_h^+} \|p - q\| &\leq Ch\|p\|_1 && \text{for all} \quad p \in H^1(\Omega), \\ \|\mathbf{v}\|_s &\leq Ch^{-1}\|\mathbf{v}\|_{s-1} && \text{for all} \quad \mathbf{v} \in \mathbf{V}_h, \ s \in \{0, 1\}, \\ \|p\| &\leq Ch^{-1}\|p\|_{-1} && \text{for all} \quad p \in M_h \end{aligned}$$

Then there are constants $C_0 > 0$, C_1 independent of h and ξ, such that

$$C_0 \langle y, \tilde{\mathrm{Q}}_S^{-1} y \rangle \leq \langle \mathrm{S}\tilde{\mathrm{Q}}_S^{-1} y, \tilde{\mathrm{Q}}_S^{-1} y \rangle \leq C_1 \langle y, \tilde{\mathrm{Q}}_S^{-1} y \rangle \quad \text{for all} \ \ y \in \mathrm{e}^{\perp} \tag{5.7}$$

Note that many pairs of finite element spaces satisfy the assumptions in this theorem. In our experiments we use such a pair, namely the Hood-Taylor $P_2 - P_1$ pair.
Although the Laplace problem for the pressure that has to be solved in each application of $\tilde{\mathrm{Q}}_S^{-1}$ is much smaller than the problems involving A or Q_A, often it is inefficient to solve it exactly. Let Q_T be a preconditioner for T_h that is spectrally equivalent to T_h, uniformly in h. Instead of $\tilde{\mathrm{Q}}_S^{-1}$ in (5.6) one can then use

$$\mathrm{Q}_S^{-1} := \begin{cases} \mathrm{M}^{-1} + \xi \mathrm{Q}_T^{-1}, & \text{if} \quad \xi \leq h^{-2}, \\ \xi h^2 \mathrm{M}^{-1} + \xi \mathrm{Q}_T^{-1}, & \text{if} \quad h^{-2} \leq \xi \end{cases} \tag{5.8}$$

In [4] (remark 4.1) it is shown that for this preconditioner a result as in (5.7) holds with constants $C_0 > 0$ and C_1 independent of h and ξ. Hence, the preconditioner Q_S as in (5.8) is spectrally equivalent to S (on $\mathrm{e}^{\perp_M}$) uniformly in h and ξ.
In the numerical experiments in section 6 for Q_T we use a symmetric multigrid V-cycle method.

6. Numerical experiments. In this section we present a comparison of the three iterative solvers discussed above. These solvers are denoted by BPCG (algorithm 3.9), PMINRES (section 3.2) and MGUZAWA (algorithm (3.22)-(3.27)). We consider a problem as in (1.1) on the unit cube $\Omega = (0,1)^3$. For the discretization we start with a uniform tetrahedral grid with $h = \frac{1}{2}$ and we apply regular refinements to this starting triangulation. For the finite element discretization we used the LBB stable pair of Hood-Taylor $P_2 - P_1$, i.e. continuous piecewise quadratics for the velocity and continuous piecewise linears for the pressure. We performed computations for the cases $h = 1/16$, $h = 1/32$. Note that for $h = 1/32$ we have approximately $7.5{\cdot}10^5$ velocity unknowns and $3.3{\cdot}10^4$ pressure unknowns ($n \approx 7.5{\cdot}10^5,\ m \approx 3.3{\cdot}10^4$).
We consider the linear system as in (2.4) with solution $(\mathrm{x}, \mathrm{y}) = 0$. We take a fixed arbitrary starting vector $(\mathrm{x}^0, \mathrm{y}^0)$, with $\mathrm{y}^0 \in \mathrm{e}^{\perp_M}$.
The iteration is stopped after a reduction of the Euclidean norm of the starting residual by a factor 10^6, i.e., $\|\mathrm{r}^k\| \leq 10^{-6}\|\mathrm{r}^0\|$, with $\mathrm{r}^k = (\mathrm{f}, 0) - \mathrm{K}(\mathrm{x}^k, \mathrm{y}^k)$.

Experiments for the case $\xi = 0$. We apply the BPCG, PMINRES, MGUZAWA methods described above. For the preconditioner Q_A we use one multigrid V-cycle iteration as in (3.28). For the preconditioner Q_S for the Schur complement we use one multigrid V-cycle iteration as in (3.28), but now applied to the mass matrix M.

For the BPCG method we need a suitable scaling of Q_A. To determine the scaling parameter, we compute $\tilde{\lambda} \approx \lambda_{\max}(I - \mathrm{Q}_{\mathrm{MG}}^{-1}\mathrm{A})$ by a few power-iterations and use $\mathrm{Q}_A := (1 - \alpha\,\tilde{\lambda})\mathrm{Q}_{MG}$, $\alpha \in [1, \tilde{\lambda}^{-1})$, as scaled preconditioner. The approximation $\tilde{\lambda}$ satisfies $\tilde{\lambda} \leq \lambda_{\max}(I - \mathrm{Q}_{\mathrm{MG}}^{-1}\mathrm{A})$, hence we take $\alpha \geq 1$. In our experiments we use $\alpha = 1.1$. Although not negligible, the arithmetic costs for determining this scaling parameter are not taken into account in the tables presented below. The PMINRES method is parameter free. In the MGUZAWA method we have to chose a stopping criterion for the inner iteration. We use the one in (3.29) with $\sigma_{\mathrm{acc}} = 0.5$. For all three methods the costs are dominated by the evaluations of Q_A^{-1}. Therefore to measure the arithmetic costs we count the number of Q_A^{-1} evaluations, which is denoted by $\#\mathrm{Q}_{MG}$. Results are presented in table 6.1.

	$h = 1/16$	$h = 1/32$
BPCG	29	29
PMINRES	49	49
MGUZAWA	33	30

TABLE 6.1
$\#\mathrm{Q}_{MG}$ for the three methods.

In figure 6.1 we give the convergence behaviour of $\|\mathrm{r}^k\|$ for the case $h = 1/32$. Note that in the MGUZAWA method in each iteration we have an inner iteration. In table 6.2 we give the number of inner iterations per outer iteration.

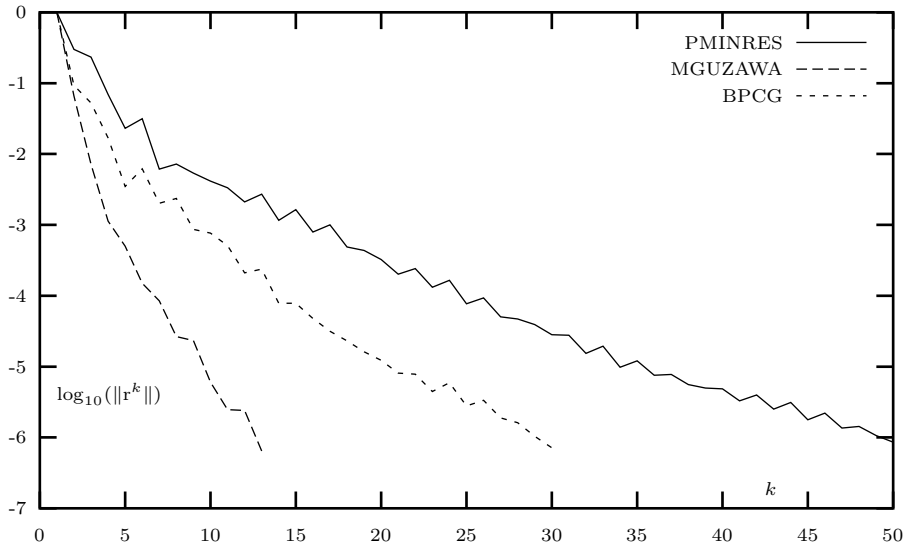

FIG. 6.1. *Convergence behaviour of* $\|r^k\|$ *for the case* $h = 1/32$.

k	1	2	3	4	5	6	7	8	9	10	11	12
# inner it.	1	1	1	2	1	2	1	2	1	1	3	2

TABLE 6.2
Number of inner iterations in MGUZAWA, $h = 1/32$.

The behaviour of the BPCG method depends on the scaling parameter α in $(1 - \alpha\,\tilde{\lambda})Q_{MG}$. In table 6.3 we show the dependence of $\#Q_{MG}$ on this scaling parameter for $\tilde{\lambda} = \lambda_{\max}(I - Q_{MG}^{-1}A)$. The convergence of the MGUZAWA method depends on σ_{acc}. In table 6.4 we show the dependence of $\#Q_{MG}$ on this parameter.

α	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.5	2.0
$\#Q_{MG}$	div	div	28	28	29	30	31	32	36

TABLE 6.3
Dependence of BPCG method on scaling parameter; $h = 1/32$.

σ_{acc}	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\#Q_{MG}$	38	33	34	30	30	29	30	30	52

TABLE 6.4
Dependence of MGUZAWA method on σ_{acc}; $h = 1/32$.

According to remarks 1, 2, 5 the bounds for the convergence rates of BPCG and PMINRES depend on the scaling parameter ρ if we use ρQ_S ($\rho > 0$) as a preconditioner for the Schur complement. The convergence rate of MGUZAWA does not depend on the scaling of Q_S. Table 6.5 shows $\#Q_{MG}$ for different values of ρ.

ρ	10^{-4}	10^{-2}	1.0	10^2	10^4
BPCG	110	58	29	44	45
PMINRES	135	91	49	30	40
MGUZAWA	30	30	30	30	30

TABLE 6.5
Dependence of $\#Q_{MG}$ on the scaling of Q_S; $h = 1/32$.

Experiments for the case $\xi > 0$. We again apply the three methods described above. For the preconditioner Q_A we use the same multigrid V-cycle iteration as for the case $\xi = 0$. For the preconditioner for the Schur complement we use the one in (5.8). For Q_T we apply one multigrid V-cycle iteration with one symmetric Gauss-Seidel pre- and post-smoothing iteration. For the BPCG method we take the same scaling of Q_A as we used for the case $\xi = 0$. Note that this may not be the optimal scaling if $\xi > 0$. In the MGUZAWA method we take $\sigma_{\text{acc}} = 0.6$. In the tables 6.6 and 6.7 we present results for $\xi = h^{-1}$ and $\xi = h^{-2}$.

	$h = 1/16$	$h = 1/32$
BPCG	29	28
PMINRES	48	48
MGUZAWA	26	29

TABLE 6.6
$\#Q_{MG}$ for the three methods; $\xi = h^{-1}$.

	$h = 1/16$	$h = 1/32$
BPCG	26	24
PMINRES	44	41
MGUZAWA	27	25

TABLE 6.7
$\#Q_{MG}$ for the three methods; $\xi = h^{-2}$

7. Concluding remarks. In this paper we considered three fast iterative methods for discretized Stokes equations. The BPCG method is a preconditioned CG method applied to a transformed system and with a special scalar product. The PMINRES method is a standard preconditioned minimal residual method. The MGUZAWA method results from an approximation of an exact block factorization of the matrix K (cf. (3.20),(3.21)). The MGUZAWA algorithm has the structure of an outer-inner iteration. Thus for this method a stopping criterion (tolerance parameter) for the inner iteration is needed. The PMINRES method is parameter free. In the BPCG method a scaling parameter for the preconditioner Q_A has to be determined. If for Q_A we use a multigrid solver then for all three methods the costs per iteration are dominated by the evaluations of Q_A^{-1}. In the BPCG and the PMINRES method

one needs one Q_A^{-1} evaluation per iteration. In the MGUZAWA method $\ell + 1$ Q_A^{-1} evaluations per iteration are needed (ℓ: # inner iterations). For all three methods convergence analyses are available in which bounds for the rate of convergence are derived that depend only on the constants in the spectral equivalences $\gamma_A Q_A \leq A \leq \Gamma_A Q_A$, $\gamma_S Q_S \leq S \leq \Gamma_S Q_S$. For the MGUZAWA method we presented a convergence analysis that shows that if one uses a multigrid preconditioner for A then for the inner iteration a very low accuracy is already sufficient to have a convergent and efficient method. In this convergence analysis we use a natural norm.

In the numerical experiments for the stationary case, all three methods have a rate of convergence independent of h. In our experiments the BPCG method and MGUZWA have almost the same efficiency. The PMINRES method is less efficient. The latter method, however, is the most robust one. It is parameter free and converges even without preconditioning. From table 6.3 we conclude that the performance of the BPCG-method depends only mildly on α (provided $\alpha \geq 1$). Thus, we need only a rough approximation $\tilde{\lambda} \approx \lambda_{\max}(I - Q_A^{-1}A)$ to obtain an efficient solver. The results in table 6.4 show that the MGUZAWA method is efficient for a broad range of σ_{acc}-values. Note that the method converges even for large σ_{acc} values. This confirms the theory in section 4.
The rate of convergence of both BPCG and PMINRES depend on the scaling of Q_S, whereas for the MGUZAWA method this is not the case. For particular choices of this scaling it may happen that the PMINRES method is more efficient than the BPCG method (cf. table 6.5).

The experiments for the instationary case demonstrate the effectiveness of the preconditioner in (5.8). The rate of convergence is independent of h for $\xi = h^{-1}$ and $\xi = h^{-2}$. If instead of the preconditioner in (5.8) one uses $Q_S = \bar{M}$, for $\xi = h^{-2}$ all three solvers would need several hundred to thousand iterations.

REFERENCES

[1] Arrow, K., Hurwicz, L., and Uzawa, H., *Studies in nonlinear programming*, Stanford University Press, Stanford, CA, 1958.

[2] Bank, R.E., Welfert, B.D., and Yserentant, H.: A class of iterative methods for solving saddle point problems, Numer. Math., **56**, 1990, 645–666.

[3] Bramble, J.H., Pasciak, J.E.: A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. Math. Comp., **50**, 1988, 1–17.

[4] Bramble, J.H., and Pasciak, J.E. Iterative techniques for time dependent Stokes problems, Comput. Math. Appl. **33**, 1997, 13–30

[5] Bramble, J.H., Pasciak, J.E., and Vassilev, A.T.: Analysis of the inexact Uzawa algorithm for saddle point problems. SIAM J. Numer. Anal. , **34**, 1997, 1072–1092.

[6] Bramble, J.H., *Multigrid Methods*, Pitman Research Notes in Mathematics, Vol. 294, John Wiley, New York, 1993.

[7] Brezzi, F. and Fortin, M., *Mixed and Hybrid Finite Element Methods*, Springer, New York, 1991.

[8] Cahouet, J., Chabard, J.P.: Some fast 3D finite element solvers for the generalized Stokes problem. Int. J. Numer. Methods Fluids, **8**, 1988, 869–895.

[9] Ciarlet, P.G., *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

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

[11] Elman, H.C.: Multigrid and Krylov subspace methods for the discrete Stokes equations. Int. J. Numer. Meth. Fluids, **227**, 1996, 755–770.

[12] Elman, H.C. and Golub, G.H.: Inexact and preconditioned Uzawa algorithms for saddle point problems, SIAM J. Numer. Anal., **31**, 1994, 1645–1661.

[13] Girault, V., and Raviart, P.-A., *Finite Element Methods for Navier-Stokes Equations,* Springer, Berlin, 1986.

[14] Greenbaum, A., *Iterative Methods for Solving Linear Systems*, Frontiers in Applied Mathematics, Vol. 17, SIAM, Philadelphia, 1997.
[15] Hackbush, W., *Multi-grid Methods and Applications,* Springer, Berlin, Heidelberg, 1985.
[16] Hackbush W. *Iterative solution of large sparse systems of equations,* Springer, Berlin, Heidelberg, 1994.
[17] Kobelkov, G.M. and Olshanskii, M.A. Effective preconditioning of Uzawa type schemes for a generalized Stokes problem, Numer. Math., **86**, 2000, 443–470
[18] Olshanskii, M.A., Reusken, A.: On the convergence of a multigrid algorithm for linear reaction-diffusion problems, Computing, **65**, 2000, 193–202.
[19] Paige, C.C., Saunders, M.A.: Solution of sparse indefinite systems of linear equations. SIAM J. Numer. Anal., **11**, 1974, 197–209.
[20] Powell, C., Silvester, D.: Black-box preconditioning for mixed formulation of self-adjoint elliptic PDEs. In: *Challenges in Scientific Computing - CISC 2002*, ed. E. Bänsch. Lecture Notes in Computational Science and Engineering, Vol. 35, 268–285, 2003.
[21] Rusten, T. and Winther, R.: A preconditioned iterative method for saddlepoint problems. SIAM J. Matrix Anal. Appl., **13**, 1992, 887–904.
[22] Silvester, D., Elman, H., Kay, D. and Wathen, A.: Efficient preconditioning of the linearized Navier-Stokes equations for incompressible flow. J. Comp. Appl. Math., **128**, 2001, 261–279.
[23] Silvester, D. Wathen, A.: Fast iterative solution of stabilised stokes systems. Part II: using general block preconditioners. SIAM J. Numer. Anal., **31**, 1994, 1352–1367.
[24] Verfürth, R.: A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem, IMA Journal Numer. Anal., **4**, 1984, 441–455.
[25] Wathen, A.: Realistic eigenvalues bounds for the Galerkin mass matrix. IMA J. Numer. Anal., **7**, 1987, 449–457.
[26] Wathen, A., Silvester, D.: Fast iterative solution of stabilised stokes systems. Part I: using simple diagonal preconditioners. SIAM J. Numer. Anal., **30**, 1993, 630–649.
[27] Wesseling, P., *An Introduction to Multigrid Methods*, Wiley, Chichester, 1992.
[28] , Wittum, G.: Multigrid methods for Stokes and Navier-Stokes equations. Transforming smoothers: algorithms and numerical results. Numer. Math., **54**, 1989, 543–563.
[29] Zulehner, W.: Analysis of iterative methods for saddle point problems: A unified approach, Math. Comp., **71**, 2001, 479–505.