ON THE ROLE OF COMMUTATOR ARGUMENTS IN THE DEVELOPMENT OF PARAMETER-ROBUST PRECONDITIONERS FOR STOKES CONTROL PROBLEMS

JOHN W. PEARSON*

Abstract. The development of preconditioners for PDE-constrained optimization problems is a field of numerical analysis which has recently generated much interest. One class of problems which has been investigated in particular is that of Stokes control problems, that is the problem of minimizing a functional with the Stokes (or Navier-Stokes) equations as constraints. In this manuscript, we present an approach for preconditioning Stokes control problems using preconditioners for the Poisson control problem and, crucially, the application of a commutator argument. This methodology leads to two block diagonal preconditioners for the problem, one of which was previously derived by W. Zulehner in 2011 (*SIAM. J. Matrix Anal. & Appl., v.32*) using a nonstandard norm argument for this saddle point problem, and the other of which we believe to be new. We also derive two related block triangular preconditioners using the same methodology, and present numerical results to demonstrate the performance of the four preconditioners in practice.

Key words. PDE-constrained optimization, Stokes control, saddle point system, preconditioning, Schur complement, commutator.

AMS subject classifications. 65F08, 65F10, 65F50, 76D07, 76D55, 93C20

1. Introduction. Decades ago, a significant area of research in numerical analysis was the numerical solution of the Stokes and Navier-Stokes equations, two partial differential equations (PDEs) that are crucial to the field of fluid dynamics. Preconditioned Krylov subspace methods for the solutions of the saddle point systems relating to each of these problems are given in [23] and [10] respectively, for instance. Since then, a further area of numerical analysis has become prominent: that of *PDE-constrained optimization*, which involves minimizing a functional with one or more PDEs as constraints. Consequently, the development of solvers for Stokes control problems, one of the most fundamental such problems, has itself become a well researched area.

There has been much success in this area: iterative solvers for a class of these problems that are independent of the mesh-size h have been devised for the time-independent problem in [18] and the time-dependent problem in [22]. Further, a multigrid solver constructed in [8] is shown to be itself independent of h. However, generating Krylov subspace solvers that are robust with respect to regularization parameter as well as mesh-size has proved to be a more difficult task – one notable exception is the preconditioned MINRES approach derived in [25] using a nonstandard norm argument, which does exhibit this independence.

In this manuscript, we consider the time-independent Stokes control problem where the velocity and the control variable are regularized but the pressure is not. We consider these problems using fundamental saddle point theory, and explain how it is possible to use this to construct preconditioners for the Stokes control problem, using a Poisson control preconditioner along with a commutator argument, the concept of which we shall describe.

There are many reasons why we believe such an investigation is of considerable interest. Firstly, it enables us to re-derive the preconditioner of Zulehner [25] within

*Numerical Analysis Group, Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford, OX1 3LB, UK (john.pearson@worc.ox.ac.uk)

a pure saddle point framework. We are also able to derive a new block diagonal preconditioner for this problem that is robust with respect to mesh-size and control regularization coefficient, as well as two block triangular preconditioners which appear to have the same property. Finally, and perhaps most intriguingly, we believe that the theory outlined in this paper can be applied to the much harder Navier-Stokes control problem, which we will address in a future manuscript [15]. We are also able to use the methodology presented here to explain why the choice of whether or not to regularize the pressure is crucial from a preconditioning point of view.

This manuscript is structured as follows. In Section 2, we detail two areas of background which we will make use of: those of saddle point theory, and preconditioners for Poisson control problems. In Section 3, we combine these areas with the theory of commutator arguments to derive the 4 aforementioned preconditioners for Stokes control problems (2 block diagonal and 2 block triangular). We also state the dominant computational operations required to apply our preconditioners, and discuss the importance of the inclusion or omission of a pressure regularization term. In Section 4, we provide numerical results to demonstrate how the preconditioners perform in practice, and in Section 5, we make some concluding remarks.

2. Background.

In this section, we introduce two fundamental subject areas which we will utilize in the remainder of this manuscript. Firstly, in Section 2.1, we outline some basic properties of saddle point systems which we will make use of. Secondly, in Section 2.2, we detail theory of solving Poisson control problems, which we will also exploit.

2.1. Saddle Point Systems.

The matrix systems that we will consider the iterative solution of in the remainder of this manuscript are all of *saddle point* form, that is of the form

$$\underbrace{\left[\begin{array}{cc} \Phi & \Psi^T \\ \Psi & -\Theta \end{array}\right]}_{\Lambda}\left[\begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_2 \end{array}\right]=\left[\begin{array}{c} \mathbf{b}_1 \\ \mathbf{b}_2 \end{array}\right], \tag{2.1}$$

where $\Phi \in \mathbb{R}^{m\times m}$, $\Psi \in \mathbb{R}^{p\times m}$, $p \leq m$, has full row rank, and $\Theta \in \mathbb{R}^{p\times p}$. In the problems that we study, Φ and Θ are also symmetric. A comprehensive review of systems of saddle point type is given in [1].

We will be seeking preconditioners for equations of the form (2.1). Therefore we utilize the observations that if we precondition Λ with $\bar{\mathcal{P}}_1$, $\bar{\mathcal{P}}_2$ or $\bar{\mathcal{P}}_3$, where

$$\bar{\mathcal{P}}_1=\left[\begin{array}{cc} \Phi & 0 \\ 0 & \Theta+\Psi\Phi^{-1}\Psi^T \end{array}\right],\quad \bar{\mathcal{P}}_2=\left[\begin{array}{cc} \Phi & 0 \\ \Psi & \Theta+\Psi\Phi^{-1}\Psi^T \end{array}\right],$$
$$\bar{\mathcal{P}}_3=\left[\begin{array}{cc} \Phi & 0 \\ \Psi & -\Theta-\Psi\Phi^{-1}\Psi^T \end{array}\right],$$

then the non-zero eigenvalues of $\bar{\mathcal{P}}_2^{-1}\Lambda$ and $\bar{\mathcal{P}}_3^{-1}\Lambda$ are given by

$$\lambda(\bar{\mathcal{P}}_2^{-1}\Lambda)=\{\pm 1\},\quad \lambda(\bar{\mathcal{P}}_3^{-1}\Lambda)=\{1\},$$

for any choice of Θ, and the non-zero eigenvalues of $\bar{\mathcal{P}}_1^{-1}\Lambda$ are given by

$$\lambda(\bar{\mathcal{P}}_1^{-1}\Lambda)=\left\{1,\frac{1}{2}(1\pm\sqrt{5})\right\},$$

provided $\Theta = 0$. The above results were given in [12, 13] in the case $\Theta = 0$; the eigenvalue results for $\bar{\mathcal{P}}_2^{-1}\Lambda$ and $\bar{\mathcal{P}}_3^{-1}\Lambda$ in the case $\Theta \neq 0$ were shown in [9].

Now, the matrices $\bar{\mathcal{P}}_1^{-1}\Lambda$ and $\bar{\mathcal{P}}_2^{-1}\Lambda$ are diagonalizable but $\bar{\mathcal{P}}_3^{-1}\Lambda$ is not, so consequently a Krylov subspace method for the solution of (2.1) with 'ideal' preconditioners $\bar{\mathcal{P}}_1$, $\bar{\mathcal{P}}_2$ and $\bar{\mathcal{P}}_3$ should converge in 3, 2 and 2 iterations respectively [13] in the relevant cases provided that $\mathcal{A}$ is non-singular.[1] Naturally, we will not be explicitly constructing $\bar{\mathcal{P}}_1$, $\bar{\mathcal{P}}_2$ and $\bar{\mathcal{P}}_3$ as this would be computationally wasteful, but will instead construct approximations $\widehat{\mathcal{P}}_1$, $\widehat{\mathcal{P}}_2$ and $\widehat{\mathcal{P}}_3$ such that the actions of the inverses of our preconditioners may be applied efficiently. Having developed these preconditioners, one may consider the MINRES algorithm [14] with preconditioners of the form $\widehat{\mathcal{P}}_1$, and preconditioners of the form $\widehat{\mathcal{P}}_2$ and $\widehat{\mathcal{P}}_3$ used in conjunction with solvers such as GMRES [19] and the Bramble-Pasciak Conjugate Gradient method [3].

We note that the quantity $S := \Theta + \Psi\Phi^{-1}\Psi^T$ is important in all three of the above preconditioners; this term is commonly known as the (negative) *Schur complement* of Λ, and much emphasis will be placed on approximating this quantity of the matrix systems we consider.

2.2. Optimal Control of Poisson's Equation. In literature including [17, 20, 25], the iterative solution of the matrix system resulting from the distributed Poisson control problem

$$\begin{aligned} \min_{y,u} \ & \frac{1}{2}\left\|y-\widehat{y}\right\|^2_{L_2(\Omega)} + \frac{\beta}{2}\left\|u\right\|^2_{L_2(\Omega)} \\ \text{s.t.} \ & -\nabla^2 y = u, \quad \text{in } \Omega, \\ & \quad y = g, \quad \text{on } \partial\Omega, \end{aligned}$$

is considered. Here, the domain on which we are working is denoted as $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, with boundary $\partial\Omega$. Here, y denotes the *state variable* (with $\widehat{y}$ some *desired state*), u the *control variable*, and β a *regularization parameter* (or *Tikhonov parameter*).

Discretizing the above problem using equal-order finite element basis functions for y, u and p leads to the 2×2 matrix system [17]

$$\begin{bmatrix} M & K \\ K & -\frac{1}{\beta}M \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} M\widehat{\mathbf{y}} + \mathbf{c} \\ \mathbf{d} \end{bmatrix}, \tag{2.2}$$

where $\mathbf{y}$ and $\widehat{\mathbf{y}}$ are the discretized versions of y and $\widehat{y}$ respectively, $\mathbf{c}$ and $\mathbf{d}$ are vectors corresponding to the boundary conditions, and $\mathbf{p}$ is the discretized version of the *adjoint variable* (or *Lagrange multiplier*) p, which is related to u by $\beta u - p = 0$. Here M denotes a finite element mass matrix, and K a finite element stiffness matrix, two frequently used types of matrices, both of which are positive definite.

Two preconditioners which are robust for all values of mesh-size h and regularization parameter β, which we denote $\mathcal{P}_1^P$ and $\mathcal{P}_2^P$ here, have been derived and tested

[1] In the problem we will consider, the matrix is singular, however the preconditioners described are often effective for such systems also.

for the matrix system (2.2) in [25] and [17] respectively:

$$\mathcal{P}_1^P = \begin{bmatrix} M+\sqrt{\beta}K & 0 \\ 0 & \frac{1}{\beta}(M+\sqrt{\beta}K) \end{bmatrix},$$

$$\mathcal{P}_2^P = \begin{bmatrix} M & 0 \\ 0 & \left(K+\frac{1}{\sqrt{\beta}}M\right)M^{-1}\left(K+\frac{1}{\sqrt{\beta}}M\right) \end{bmatrix}.$$

These two preconditioners have been derived in very different ways: $\mathcal{P}_1^P$ was derived using a nonstandard norm argument in [25], and $\mathcal{P}_2^P$ using the saddle point theory described in Section 2.1.[2] Each of these preconditioners for the Poisson control problem may be extended to an effective preconditioner for Stokes control, as we will demonstrate in Section 3.

3. Optimal Control of Stokes Equations. The problem that we consider for the majority of this section is the following distributed Stokes control problem:

$$\begin{aligned} \min_{\underline{\mathbf{v}},\underline{\mathbf{u}}} \ & \frac{1}{2}\left\|\underline{\mathbf{v}}-\widehat{\underline{\mathbf{v}}}\right\|_{L_2(\Omega)}^2+\frac{\beta}{2}\left\|\underline{\mathbf{u}}\right\|_{L_2(\Omega)}^2 \\ \text{s.t.} \ & -\nabla^2\underline{\mathbf{v}}+\nabla p=\underline{\mathbf{u}}, \quad \text{in } \Omega, \\ & -\nabla\cdot\underline{\mathbf{v}}=0, \quad \text{in } \Omega, \\ & \underline{\mathbf{v}}=\underline{\mathbf{g}}, \quad \text{on } \partial\Omega. \end{aligned}$$

Again we work on a domain $\Omega\subset\mathbb{R}^d$, $d\in\{2,3\}$, with boundary $\partial\Omega$, and with regularization parameter β. Here, $\underline{\mathbf{v}}$ denotes the velocity in d dimensions and p the pressure term, both of which are state variables in this problem. $\underline{\mathbf{u}}$ here is the control variable in d dimensions. We also introduce at this point the adjoint variables $\underline{\boldsymbol{\lambda}}$ (in d dimensions) and μ.

Discretizing this problem results in the matrix system [25]

$$\begin{bmatrix} M & 0 & K & B^T \\ 0 & 0 & B & 0 \\ K & B^T & -\frac{1}{\beta}M & 0 \\ B & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \mathbf{v} \\ \mathbf{p} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix}=\begin{bmatrix} M\widehat{\mathbf{v}}+\mathbf{c} \\ \mathbf{0} \\ \mathbf{d} \\ \mathbf{f} \end{bmatrix}, \tag{3.1}$$

where M and K here denote $d\times d$ block matrices with mass and stiffness matrices on the velocity space on the block diagonals,[3] and B represents the negative of the divergence operator on the finite element space in matrix form. The vector $\widehat{\mathbf{v}}$ corresponds to the target function $\widehat{\underline{\mathbf{v}}}$, $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ relate to the adjoint variables $\underline{\boldsymbol{\lambda}}$ and μ, and the vectors $\mathbf{c}$, $\mathbf{d}$ and $\mathbf{f}$ take account of boundary conditions. We note at this point that this matrix is in general singular, as it is well known that the vector of ones is a member of the nullspace of B^T (see [6, Chapter 5] for instance) – the matrix in (3.1) therefore has 2 zero eigenvalues (one corresponding to each appearance of B^T).[4] However, this

[2]The crucial step in constructing $\mathcal{P}_2^P$ is that of approximating the Schur complement of (2.1), $KM^{-1}K+\frac{1}{\beta}M$. In [17], it is shown that if we approximate this by $\left(K+\frac{1}{\sqrt{\beta}}M\right)M^{-1}\left(K+\frac{1}{\sqrt{\beta}}M\right)$, then the eigenvalues of the preconditioned Schur complement are all contained within the interval $\left[\frac{1}{2},1\right]$.

[3]Note that this definition of M and K is slightly different from the definition used in Section 2.2, where these terms simply denoted a single mass or stiffness matrix.

[4]On the continuous level, the zero eigenvalues arise from the fact that an arbitrary constant may be added to the solution of the pressure p or the adjoint variable μ and yield another solution.

may be avoided by restricting the pressure space to the orthogonal complement of the nullspace, as in this case the matrix B^T will clearly no longer have a nullspace.

We also note that, in the construction of the functional being minimized in this optimal control problem, we have not regularized the pressure term – the problem where pressure is regularized was considered in [18, 22] for instance. This is extremely important from a preconditioning point of view, and in Section 3.4 we explain why this makes a major difference.

We consider discretizing this problem using the well-studied (inf-sup stable) *Taylor-Hood* finite element basis functions, that is discretizing the velocity $\underline{\mathbf{v}}$ using **Q2**-basis functions, and the pressure p using **Q1**-basis functions. We discretize the control $\underline{\mathbf{u}}$ and adjoint variable $\underline{\boldsymbol{\lambda}}$ using **Q2**-functions, and the adjoint variable μ using **Q1**-functions.

It is not immediately obvious how the preconditioners derived for the Poisson control problem in the previous section can be applied to the more difficult Stokes control problem. In this section, we explain how this may be achieved.

3.1. Derivation of Block Diagonal Preconditioners. To commence our derivation, we reorder the matrix system (3.1) so that we are dealing with the system

$$\underbrace{\left[\begin{array}{cccc} M & K & B^T & 0 \\ K & -\frac{1}{\beta}M & 0 & B^T \\ B & 0 & 0 & 0 \\ 0 & B & 0 & 0 \end{array}\right]}_{\mathcal{A}}\left[\begin{array}{c} \mathbf{v} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \\ \mathbf{p} \end{array}\right]=\left[\begin{array}{c} M\widehat{\mathbf{v}}+\mathbf{c} \\ \mathbf{d} \\ \mathbf{f} \\ \mathbf{0} \end{array}\right].$$

This is a saddle point system of the form (2.1), with

$$\Phi=\left[\begin{array}{cc} M & K \\ K & -\frac{1}{\beta}M \end{array}\right],\quad \Psi=\left[\begin{array}{cc} B & 0 \\ 0 & B \end{array}\right],\quad \Theta=\left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right].$$

Note that the $(1,1)$-block Φ is of the form of the matrix system (2.2) relating to the Poisson control problem. We will use this to motivate two block diagonal preconditioners, related to two preconditioners for Poisson control detailed in Section 2.2. These preconditioners will be of the form

$$\mathcal{P}=\left[\begin{array}{cc} \widehat{\Phi} & 0 \\ 0 & (\Psi\Phi^{-1}\Psi^T)_{\text{approx}} \end{array}\right]. \tag{3.2}$$

Such a strategy also leads to block triangular preconditioners of the form

$$\mathcal{P}=\left[\begin{array}{cc} \widehat{\Phi} & 0 \\ \Psi & (\Psi\Phi^{-1}\Psi^T)_{\text{approx}} \end{array}\right] \quad\text{or}\quad \left[\begin{array}{cc} \widehat{\Phi} & 0 \\ \Psi & -(\Psi\Phi^{-1}\Psi^T)_{\text{approx}} \end{array}\right]. \tag{3.3}$$

We will derive two such block triangular preconditioners in Section 3.2.

3.1.1. First Preconditioner. We motivate our first preconditioner for the Stokes control system (3.1) using the preconditioner $\mathcal{P}_1^P$ for the Poisson control problem of Section 2.2. We first note that the $(1,1)$-block of the Stokes control problem (3.1) is of the form of the matrix involved in the Poisson control problem, so we write, in the notation of (3.2),

$$\Phi=\left[\begin{array}{cc} M & K \\ K & -\frac{1}{\beta}M \end{array}\right]\approx\left[\begin{array}{cc} M+\sqrt{\beta}K & 0 \\ 0 & \frac{1}{\beta}(M+\sqrt{\beta}K) \end{array}\right]=:\widehat{\Phi}.$$

Here, the notation $\Phi \approx \widehat{\Phi}$ denotes that $\widehat{\Phi}$ has been constructed with the aim that the singular values of $\widehat{\Phi}^{-1}\Phi$ are bounded within a fixed (small) interval.

The next step is to find a good approximation to the Schur complement $\Psi\Phi^{-1}\Psi^T$ of the matrix system (3.2); we justify a potential approximation by writing

$$\begin{aligned}\Psi\Phi^{-1}\Psi^T &= \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M & K \\ K & -\frac{1}{\beta}M \end{bmatrix}^{-1} \begin{bmatrix} B^T & 0 \\ 0 & B^T \end{bmatrix} \\ &\approx \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M+\sqrt{\beta}K & 0 \\ 0 & \frac{1}{\beta}(M+\sqrt{\beta}K) \end{bmatrix}^{-1} \begin{bmatrix} B^T & 0 \\ 0 & B^T \end{bmatrix} =: \Psi\widehat{\Phi}^{-1}\Psi^T \\ &= \begin{bmatrix} B(M+\sqrt{\beta}K)^{-1}B^T & 0 \\ 0 & \beta B(M+\sqrt{\beta}K)^{-1}B^T \end{bmatrix}.\end{aligned}$$

We highlight that, in general, the fact that $\widehat{\Phi} \approx \Phi$ does not necessarily tell us that $\Psi\widehat{\Phi}^{-1}\Psi^T \approx \Psi\Phi^{-1}\Psi^T$ (unless Ψ is a square and invertible matrix, which is not the case here). However, this seems a reasonable motivation for an approximation which is potentially effective, and we indeed find that this strategy does lead to a good approximation of $\Psi\Phi^{-1}\Psi^T$ for this problem. Furthermore, eigenvalue analysis carried out in [11] for this preconditioner verifies its potency for the Stokes control problem.

At this point, as in [25], we may approximate $B(M+\sqrt{\beta}K)^{-1}B^T$ by $(\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1}$ in the above expression,[5] where M_p and K_p denote finite element mass and stiffness matrices, respectively, on the pressure space. Hence, we may write that

$$\Psi\Phi^{-1}\Psi^T \approx \begin{bmatrix} (\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1} & 0 \\ 0 & \beta(\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1} \end{bmatrix} =: (\Psi\Phi^{-1}\Psi^T)_{\text{approx}},$$

Therefore, putting all of the above working together, we postulate that

$$\mathcal{P}_1 = \begin{bmatrix} M+\sqrt{\beta}K & 0 & 0 & 0 \\ 0 & \frac{1}{\beta}(M+\sqrt{\beta}K) & 0 & 0 \\ 0 & 0 & (\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1} & 0 \\ 0 & 0 & 0 & \beta(\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1} \end{bmatrix}$$

is an effective preconditioner for $\mathcal{A}$. This is exactly the preconditioner proposed by Zulehner in [25] using a nonstandard norm argument. We will demonstrate the effectiveness of this preconditioner by displaying numerical results in Section 4.

3.1.2. Second Preconditioner. We are also able to derive a new block diagonal preconditioner for the Stokes control system (3.1) using the preconditioner $\mathcal{P}_2^P$ for the Poisson control problem. We treat the $(1,1)$-block of the Stokes control system by using the preconditioner for Poisson control, writing (in the notation of (3.2))

$$\begin{aligned}\Phi = \begin{bmatrix} M & K \\ K & -\frac{1}{\beta}M \end{bmatrix} &\approx \begin{bmatrix} M & 0 \\ 0 & KM^{-1}K + \frac{1}{\beta}M \end{bmatrix} \\ &\approx \begin{bmatrix} M & 0 \\ 0 & \left(K + \frac{1}{\sqrt{\beta}}M\right) M^{-1} \left(K + \frac{1}{\sqrt{\beta}}M\right) \end{bmatrix} =: \widehat{\Phi}.\end{aligned}$$

[5]This may be done by applying the commutator argument of Section 3.1.2 with $\mathcal{L} = -\sqrt{\beta}\Delta + I := -\sqrt{\beta}\nabla^2 + I$. This is carried out in a very similar fashion in [22] to matrices of this form for time-dependent Stokes control problems.

We now once again search for a good approximation to the Schur complement – we proceed as follows:

$$\begin{aligned}\Psi\Phi^{-1}\Psi^T &\approx \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & KM^{-1}K + \frac{1}{\beta}M \end{bmatrix}^{-1} \begin{bmatrix} B^T & 0 \\ 0 & B^T \end{bmatrix} \\ &= \begin{bmatrix} BM^{-1}B^T & 0 \\ 0 & B\left(KM^{-1}K + \frac{1}{\beta}M\right)^{-1} B^T \end{bmatrix}.\end{aligned}$$

Once again, we have assumed in the above working that $\widehat{\Phi}$ being a good approximation to Φ leads to $\Psi\widehat{\Phi}^{-1}\Psi^T$ approximating $\Psi\Phi^{-1}\Psi^T$ well; for this problem we find that this heuristic does indeed lead to an effective approximation.

We do not as yet have a feasible preconditioner, as the matrices $BM^{-1}B^T$ and $B\left(KM^{-1}K + \frac{1}{\beta}M\right)^{-1} B^T$ cannot be inverted without computing the inverses of M or $KM^{-1}K + \frac{1}{\beta}M$. However, it is well known that $BM^{-1}B^T$ may be well approximated by K_p (see [6, Chapter 8]),[6] so we use this for the first block of our Schur complement approximation.

We therefore now seek an idea for approximating $\Sigma := B\left(KM^{-1}K + \frac{1}{\beta}M\right)^{-1} B^T$, so that we obtain a cheap and invertible approximation to the Schur complement. We do this using a commutator argument, a type of which is described in [6] for the Navier-Stokes equations for instance. We examine the commutator

$$\mathcal{E} = (\mathcal{L})\nabla - \nabla(\mathcal{L})_p,$$

where $\mathcal{L} = \Delta^2 + \frac{1}{\beta}I := \nabla^4 + \frac{1}{\beta}I$. This is an operator carefully chosen to give us a matrix that we can use to approximate Σ.

Now, discretizing this commutator using finite elements gives

$$\mathcal{E}_h = (M^{-1}L)M^{-1}B^T - M^{-1}B^T(M_p^{-1}L_p),$$

where $L = KM^{-1}K + \frac{1}{\beta}M$. Pre-multiplying by $BL^{-1}M$ and post-multiplying by $L_p^{-1}M_p$ where $L_p = K_pM_p^{-1}K_p + \frac{1}{\beta}M_p$ then gives

$$BM^{-1}B^TL_p^{-1}M_p \approx BL^{-1}B^T,$$

where, crucially, we assume that the commutator $\mathcal{E}_h$ is small.

We may now use the fact that $BM^{-1}B^T \approx K_p$ and substitute in the expression for L to give that[7]

$$\Sigma = B\left(KM^{-1}K + \frac{1}{\beta}M\right)^{-1} B^T \approx K_pL_p^{-1}M_p,$$

and therefore that

$$\Sigma^{-1} \approx M_p^{-1}L_pK_p^{-1} = M_p^{-1}\left(K_pM_p^{-1}K_p + \frac{1}{\beta}M_p\right)K_p^{-1} = M_p^{-1}K_pM_p^{-1} + \frac{1}{\beta}K_p^{-1}.$$

[6] The approximation $BM^{-1}B^T \approx K_p$ may be justified by the observations that $-\nabla\cdot\nabla = -\nabla^2$ on the continuous level, and that the matrices K_p, B, M and B^T relate to the continuous operators $-\nabla^2$, $-\nabla\cdot$, I and ∇ respectively.

[7] An approximation of the form $BL^{-1}B^T \approx K_pL_p^{-1}M_p$ was first introduced by Cahouet and Chabard in [4] for the forward Stokes problem. Such arguments have since been used to develop iterative solvers for a variety of fluid dynamics problems.

We note that such an argument has been used a number of times before – we give a brief summary of some applications in Section 5.

So a second possible preconditioner for $\mathcal{A}$ is

$$\mathcal{P}_2 = \begin{bmatrix} M & 0 & 0 & 0 \\ 0 & \left(K + \frac{1}{\sqrt{\beta}}M\right) M^{-1} \left(K + \frac{1}{\sqrt{\beta}}M\right) & 0 & 0 \\ 0 & 0 & K_p & 0 \\ 0 & 0 & 0 & \left(M_p^{-1} K_p M_p^{-1} + \frac{1}{\beta} K_p^{-1}\right)^{-1} \end{bmatrix},$$

which we postulate is an effective preconditioner. We again verify that this is the case with numerical results presented in Section 4.

We note at this point that this preconditioner is a more "flexible one", as we find that a preconditioner of this form may be applied to the more difficult and general Navier-Stokes control problem [15]. In more detail, when an Oseen-type iteration is applied to this problem, we may re-arrange the matrix system obtained so that we have as the $(1,1)$-block a matrix corresponding to the convection-diffusion control problem, as opposed to a Poisson control problem here. Using a preconditioner derived for the convection-diffusion control problem in [16], we may apply a similar commutator argument to approximate the Schur complement of the matrix systems for Navier-Stokes control – for this problem, we find that we need to approximate $BM^{-1}B^T$ and $B\left(FM^{-1}F^T + \frac{1}{\beta}M\right)^{-1} B^T$, where F is the differential operator relating to the Navier-Stokes equations. By doing this we arrive at iterative solvers for the Navier-Stokes control problem.

3.2. Block Triangular Preconditioners. A useful aspect of our approach is that we may consider developing robust preconditioners for the Stokes control problem that are not of the block diagonal form of $\mathcal{P}_1$ and $\mathcal{P}_2$. We do this by considering various block triangular preconditioners of the Poisson control matrix system.

Firstly, we may consider a preconditioner of the form $\begin{bmatrix} \widehat{\Phi} & 0 \\ \Psi & (\Psi\Phi^{-1}\Psi^T)_{\text{approx}} \end{bmatrix}$ stated in (3.3) that is in some sense analogous to $\mathcal{P}_1$ as derived in Section 3.1.1. We could in fact consider the same approximations $\widehat{\Phi}$ and $(\Psi\Phi^{-1}\Psi^T)_{\text{approx}}$ as we did to construct $\mathcal{P}_1$,

$$\widehat{\Phi} = \begin{bmatrix} M + \sqrt{\beta}K & 0 \\ 0 & \frac{1}{\beta}(M + \sqrt{\beta}K) \end{bmatrix},$$

$$(\Psi\Phi^{-1}\Psi^T)_{\text{approx}} = \begin{bmatrix} (\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1} & 0 \\ 0 & \beta(\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1} \end{bmatrix},$$

to develop the following block triangular preconditioner for $\mathcal{A}$:

$$\mathcal{P}_3 = \begin{bmatrix} M + \sqrt{\beta}K & 0 & 0 & 0 \\ 0 & \frac{1}{\beta}(M + \sqrt{\beta}K) & 0 & 0 \\ B & 0 & (\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1} & 0 \\ 0 & B & 0 & \beta(\sqrt{\beta}M_p^{-1} + K_p^{-1})^{-1} \end{bmatrix},$$

which may be applied within the GMRES algorithm.

In addition to this preconditioner, we may form a block lower triangular preconditioner for the Stokes control problem that is based on the following block triangular preconditioner $\mathcal{P}_3^P$ for the Poisson control problem:

$$\mathcal{P}_3^P = \begin{bmatrix} M & 0 \\ K & -\left(K + \frac{1}{\sqrt{\beta}}M\right) M^{-1} \left(K + \frac{1}{\sqrt{\beta}}M\right) \end{bmatrix},$$

which was shown to be effective for that problem in [17]. We may, once again, use this as an approximation to the $(1,1)$-block of the Stokes control matrix $\mathcal{A}$.

Let us consider how we may precondition the Schur complement of $\mathcal{A}$ while using this approximation of the $(1,1)$-block. We write, in the notation of (3.3),

$$\begin{aligned} \Psi\Phi^{-1}\Psi^T &= \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M & K \\ K & -\frac{1}{\beta}M \end{bmatrix}^{-1} \begin{bmatrix} B^T & 0 \\ 0 & B^T \end{bmatrix} \\ &\approx \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M & 0 \\ K & -\widehat{S}_P \end{bmatrix}^{-1} \begin{bmatrix} B^T & 0 \\ 0 & B^T \end{bmatrix} \\ &= \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M^{-1} & 0 \\ \widehat{S}_P^{-1}KM^{-1} & -\widehat{S}_P^{-1} \end{bmatrix} \begin{bmatrix} B^T & 0 \\ 0 & B^T \end{bmatrix} =: \Psi\widehat{\Phi}^{-1}\Psi^T \\ &= \begin{bmatrix} BM^{-1}B^T & 0 \\ B\widehat{S}_P^{-1}KM^{-1}B^T & -B\widehat{S}_P^{-1}B^T \end{bmatrix} \\ &\approx \begin{bmatrix} K_p & 0 \\ B\widehat{S}_P^{-1}KM^{-1}B^T & -\left(M_p^{-1}K_pM_p^{-1} + \frac{1}{\beta}K_p^{-1} + \frac{2}{\sqrt{\beta}}M_p^{-1}\right)^{-1} \end{bmatrix} \\ &=: (\Psi\Phi^{-1}\Psi^T)_{\text{approx}}, \end{aligned}$$

where

$$\widehat{S}_P = \left(K + \frac{1}{\sqrt{\beta}}M\right) M^{-1} \left(K + \frac{1}{\sqrt{\beta}}M\right).$$

In the working above, we have again used the approximation $BM^{-1}B^T \approx K_p$. To approximate the matrix $B\widehat{S}_P^{-1}B^T$, we have used the same commutator argument as in Section 3.1.2 except with $L = \widehat{S}_P = KM^{-1}K + \frac{1}{\beta}M + \frac{2}{\sqrt{\beta}}K$ and $L_p = K_pM_p^{-1}K_p + \frac{1}{\beta}M_p + \frac{2}{\sqrt{\beta}}K_p$.

Therefore, applying the (block triangular) saddle point theory of Section 2.1, we have arrived at a block triangular preconditioner for $\mathcal{A}$, namely:

$$\mathcal{P}_4 = \begin{bmatrix} M & 0 & 0 & 0 \\ K & -\widehat{S}_P & 0 & 0 \\ B & 0 & K_p & 0 \\ 0 & B & B\widehat{S}_P^{-1}KM^{-1}B^T & -\left(M_p^{-1}K_pM_p^{-1} + \frac{1}{\beta}K_p^{-1} + \frac{2}{\sqrt{\beta}}M_p^{-1}\right)^{-1} \end{bmatrix}.$$

Of course, we would not be able to apply the MINRES algorithm with the preconditioners $\mathcal{P}_3$ or $\mathcal{P}_4$; instead we would use the GMRES algorithm of [19].

However, numerical tests indicate that $\mathcal{P}_3$ and $\mathcal{P}_4$ are effective preconditioners for $\mathcal{A}$ nevertheless – we refer to Section 4 for a demonstration of this assertion.

3.3. Further Comments. We now wish to make some further observations about the preconditioners which we have proposed. Firstly, it is natural to consider the effectiveness of the new commutator arguments we have introduced, as such arguments are heuristic in nature. We therefore carry out numerical tests on our approximations; in particular we look for the maximum and minimum (non-zero) eigenvalues of

$$\left[M_p^{-1}K_pM_p^{-1}+\frac{1}{\beta}K_p^{-1}\right]B\left(KM^{-1}K+\frac{1}{\beta}M\right)^{-1}B^T, \tag{3.4}$$

$$\left[M_p^{-1}K_pM_p^{-1}+\frac{1}{\beta}K_p^{-1}+\frac{2}{\sqrt{\beta}}M_p^{-1}\right]B\left(KM^{-1}K+\frac{1}{\beta}M+\frac{2}{\sqrt{\beta}}K\right)^{-1}B^T, \tag{3.5}$$

which relate to the two new commutator arguments introduced in this paper, and which are utilized in the preconditioners $\mathcal{P}_2$ and $\mathcal{P}_4$ respectively. In Table 3.1, we provide eigenvalues of the matrix (3.4) for a range of mesh-sizes and values of β, and in Table 3.2, we present the same results for (3.5). For the results in both tables, an evenly spaced grid with Taylor-Hood elements was used, with the values of h stated corresponding to the distance between **Q2**-nodes. We can see that the approximations are effective ones for a range of parameter values, especially for smaller values of β. We note a benign dependence of the effectiveness of the approximations on h, but the results obtained are still very reasonable.

		β							
		10		10^{-2}		10^{-5}		10^{-8}	
		λ_2	$\lambda_{\max}$	λ_2	$\lambda_{\max}$	λ_2	$\lambda_{\max}$	λ_2	$\lambda_{\max}$
h	2^{-2}	0.0584	1.3315	0.1271	1.2617	0.4537	0.9776	0.4975	1.0096
	2^{-3}	0.0400	1.3495	0.0843	1.3245	0.2988	0.9591	0.5000	1.0090
	2^{-4}	0.0295	1.3730	0.0560	1.3560	0.1721	1.1442	0.4876	0.9994
	2^{-5}	0.0227	1.3645	0.0396	1.3624	0.1065	1.2964	0.3872	0.9968

TABLE 3.1
Maximum and minimum (non-zero) eigenvalues for commutator approximation (3.4) *used in block diagonal preconditioner, for different values of h and β.*

		β							
		10		10^{-2}		10^{-5}		10^{-8}	
		λ_2	$\lambda_{\max}$	λ_2	$\lambda_{\max}$	λ_2	$\lambda_{\max}$	λ_2	$\lambda_{\max}$
h	2^{-2}	0.0653	1.3211	0.1541	1.1475	0.3922	0.9171	0.4924	1.0026
	2^{-3}	0.0446	1.3443	0.1048	1.2563	0.2881	0.9550	0.4812	0.9839
	2^{-4}	0.0326	1.3694	0.0699	1.3167	0.1951	1.0707	0.4355	0.9876
	2^{-5}	0.0249	1.3633	0.0487	1.3466	0.1294	1.2051	0.3418	0.9968

TABLE 3.2
Maximum and minimum (non-zero) eigenvalues for commutator approximation (3.5) *used in block triangular preconditioner, for different values of h and β.*

Another pertinent question is how cheap it is to apply our proposed preconditioners. We therefore detail the main computational operations required to approximate $\mathcal{P}_1^{-1}$, $\mathcal{P}_2^{-1}$, $\mathcal{P}_3^{-1}$ and $\mathcal{P}_4^{-1}$ (excluding matrix multiplications, which are comparatively cheap). For the purposes of these descriptions, we view the preconditioners as 4×4 block matrices, and refer to each block in this way.

- **Operations needed to apply $\mathcal{P}_1^{-1}$:**
  - $(\mathbf{1},\mathbf{1})$: 1 multigrid operation for $M+\sqrt{\beta}K$
  - $(\mathbf{2},\mathbf{2})$: 1 multigrid operation for $M+\sqrt{\beta}K$
  - $(\mathbf{3},\mathbf{3})$: 1 Chebyshev semi-iteration for M_p, and 1 multigrid operation for K_p
  - $(\mathbf{4},\mathbf{4})$: 1 Chebyshev semi-iteration for M_p, and 1 multigrid operation for K_p
  - **Total**: 2 Chebyshev semi-iterations and 4 multigrids.
- **Operations needed to apply $\mathcal{P}_2^{-1}$:**
  - $(\mathbf{1},\mathbf{1})$: 1 Chebyshev semi-iteration for M
  - $(\mathbf{2},\mathbf{2})$: 2 multigrid operations for $K+\frac{1}{\sqrt{\beta}}M$
  - $(\mathbf{3},\mathbf{3})$: 1 multigrid operation for K_p
  - $(\mathbf{4},\mathbf{4})$: 2 Chebyshev semi-iterations for M_p, and 1 multigrid operation for K_p
  - **Total**: 3 Chebyshev semi-iterations and 4 multigrids.
- **Operations needed to apply $\mathcal{P}_3^{-1}$:** These are the same as for $\widehat{\mathcal{P}}_1^{-1}$, and hence total:
  - **Total**: 2 Chebyshev semi-iterations and 4 multigrids.
- **Operations needed to apply $\mathcal{P}_4^{-1}$:**
  - $(\mathbf{1},\mathbf{1})$: 1 Chebyshev semi-iteration for M
  - $(\mathbf{2},\mathbf{2})$: 2 multigrid operations for $K+\frac{1}{\sqrt{\beta}}M$
  - $(\mathbf{3},\mathbf{3})$: 1 multigrid operation for K_p
  - $(\mathbf{4},\mathbf{3})$: 1 Chebyshev semi-iteration for M, and 2 multigrid operations for $K+\frac{1}{\sqrt{\beta}}M$
  - $(\mathbf{4},\mathbf{4})$: 2 Chebyshev semi-iterations for M_p, and 1 multigrid operation for K_p
  - **Total**: 4 Chebyshev semi-iterations and 6 multigrids.

We can see from this list of operations that the application of each preconditioner (especially $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$) is fairly cheap, and therefore that our methods should be computationally effective ones.

3.4. Regularization of Pressure. In this section we briefly consider the Stokes control problem

$$\begin{aligned}
\min_{\underline{\mathbf{v}},p,\underline{\mathbf{u}}} \quad & \frac{1}{2}\left\|\underline{\mathbf{v}}-\widehat{\underline{\mathbf{v}}}\right\|_{L_2(\Omega)}^2+\frac{\alpha}{2}\left\|p-\widehat{p}\right\|_{L_2(\Omega)}^2+\frac{\beta}{2}\left\|\underline{\mathbf{u}}\right\|_{L_2(\Omega)}^2 \\
\text{s.t.} \quad & -\nabla^2\underline{\mathbf{v}}+\nabla p=\underline{\mathbf{u}}, \quad \text{in } \Omega, \\
& -\nabla\cdot\underline{\mathbf{v}}=0, \quad \text{in } \Omega, \\
& \underline{\mathbf{v}}=\underline{\mathbf{g}}, \quad \text{on } \partial\Omega,
\end{aligned}$$

which is identical to the problem we have studied in the previous sections, except we impose an additional regularization term on the pressure (with α being the corresponding regularization parameter).

It is useful to consider preconditioning the resulting matrix system [18]

$$\underbrace{\begin{bmatrix} M & K & B^T & 0 \\ K & -\frac{1}{\beta}M & 0 & B^T \\ B & 0 & 0 & 0 \\ 0 & B & 0 & \alpha M_p \end{bmatrix}}_{\mathcal{B}} \begin{bmatrix} \mathbf{v} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} M\widehat{\mathbf{v}} + \mathbf{c} \\ \mathbf{d} \\ \mathbf{f} \\ M_p\widehat{\mathbf{p}} \end{bmatrix},$$

in light of the framework discussed in this paper – in particular, whether it is possible to precondition the problem arising from the Stokes control problem *with* pressure regularization in the same way as it is to precondition the system arising *without* pressure regularization.

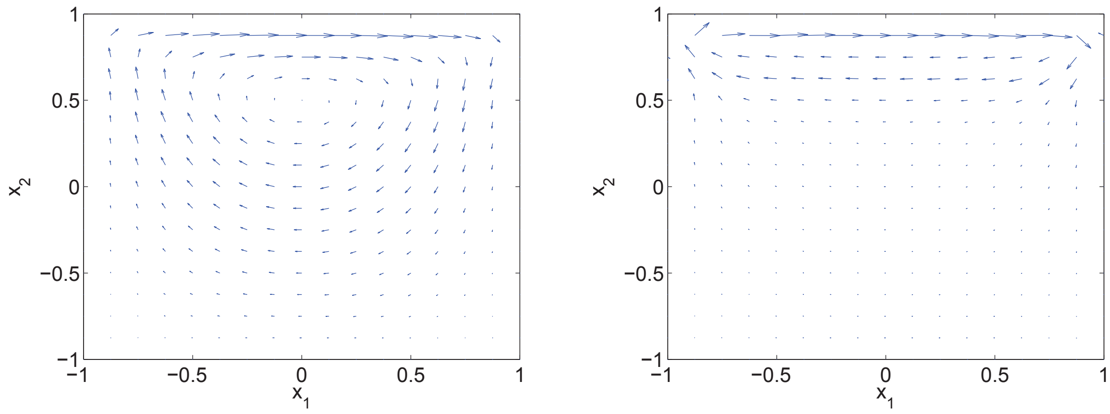
For brevity, we will simply consider developing a preconditioner of the form $\mathcal{P}_1$ to the matrix system (3.6) (we find that the same issues arise when trying to construct preconditioners of the form $\mathcal{P}_2$, $\mathcal{P}_3$ and $\mathcal{P}_4$). We may construct an approximation of the $(1,1)$-block of $\mathcal{B}$ exactly as we did for the matrix system $\mathcal{A}$ in Section 3.1.1 (as the $(1,1)$-blocks of $\mathcal{A}$ and $\mathcal{B}$ are the same). When we attempt to construct an approximation of the Schur complement of $\mathcal{B}$ in a similar way as we did for $\mathcal{A}$ in the derivation of $\mathcal{P}_1$, we obtain the following:

$$\begin{aligned} S_{\mathcal{B}} &= -\begin{bmatrix} 0 & 0 \\ 0 & \alpha M_p \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M & K \\ K & -\frac{1}{\beta}M \end{bmatrix}^{-1} \begin{bmatrix} B^T & 0 \\ 0 & B^T \end{bmatrix} \\ &\approx -\begin{bmatrix} 0 & 0 \\ 0 & \alpha M_p \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M+\sqrt{\beta}K & 0 \\ 0 & \frac{1}{\beta}(M+\sqrt{\beta}K) \end{bmatrix}^{-1} \begin{bmatrix} B^T & 0 \\ 0 & B^T \end{bmatrix} \\ &= \begin{bmatrix} B(M+\sqrt{\beta}K)^{-1}B^T & 0 \\ 0 & -\alpha M_p + \beta B(M+\sqrt{\beta}K)^{-1}B^T \end{bmatrix}. \end{aligned}$$

At this point we face a problem – the $(2,2)$-block of our proposed Schur complement approximation could be positive definite, negative definite, or indefinite, depending on the values of α, β and h used, thus creating major issues when attempting to construct a positive definite preconditioner (which we require for use with Minres). Even if the values of α, β and h were such that $-\alpha M_p + \beta B(M+\sqrt{\beta}K)^{-1}B^T$ is positive definite, it is far from clear how we may efficiently approximate this matrix in a similar way as $\beta B(M+\sqrt{\beta}K)^{-1}B^T$ was approximated in Section 3.1.1. We are therefore unable to derive a parameter-robust preconditioner using our approach.

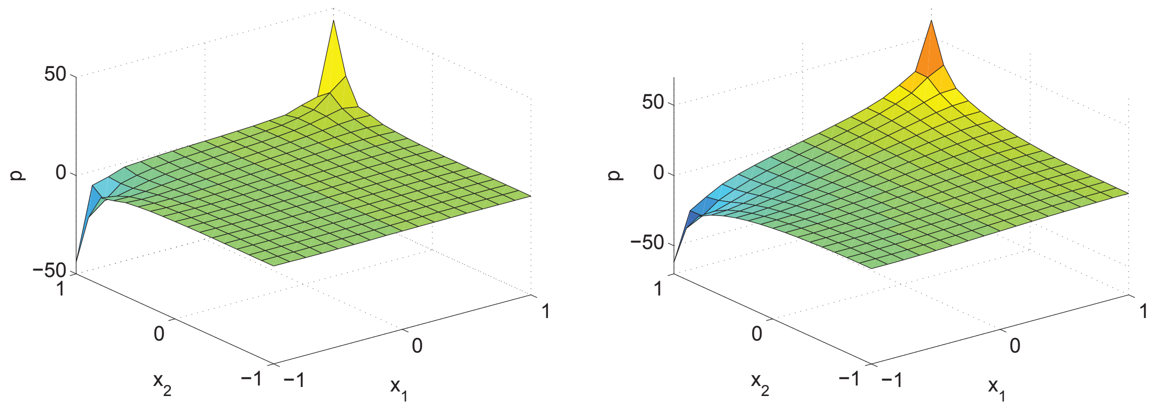
We therefore conclude that the Stokes control problem involving pressure regularization is a harder problem to solve robustly than the problem without, at least if the methodology discussed in this manuscript is considered. We point the reader to [18] for a solver for the time-independent problem with pressure regularization that is independent of the mesh-size h (but not the regularization parameter β), and to [22] for an extension of this solver to the time-dependent case.

4. Numerical Experiments. Having motivated the theoretical potential of our approach, we now seek to demonstrate how our preconditioners perform in practice. To do this, we consider two test problems. The first problem we look at is an optimal



(a) $\beta = 1$ (b) $\beta = 10^{-4}$

FIG. 4.1. *Plots of computed velocity solution to the first test problem for different* β.



(a) $\beta = 10^{-2}$ (b) $\beta = 10^{-6}$

FIG. 4.2. *Plots of computed pressure solution to the first test problem for different* β.

control analogue of the *lid-driven cavity* problem on the domain $\Omega = [-1, 1]^2$:

$$\begin{aligned} \min_{\underline{\mathbf{v}},\underline{\mathbf{u}}} \quad & \frac{1}{2} \|\underline{\mathbf{v}}\|^2_{L_2(\Omega)} + \frac{\beta}{2} \|\underline{\mathbf{u}}\|^2_{L_2(\Omega)} \\ \text{s.t.} \quad -\nabla^2 \underline{\mathbf{v}} + \nabla p = \underline{\mathbf{u}}, & \quad \text{in } \Omega, \\ -\nabla \cdot \underline{\mathbf{v}} = 0, & \quad \text{in } \Omega, \\ \underline{\mathbf{v}} = & \begin{cases} [1, 0]^T & \text{on } [-1, 1] \times \{1\}, \\ [0, 0]^T & \text{on } \partial\Omega \backslash \left([-1, 1] \times \{1\}\right). \end{cases} \end{aligned}$$

We wish to observe how well the 4 preconditioners presented in this paper perform when solving the matrix system relating to this problem. In Table 4.1, we show the number of MINRES iterations and CPU times[8] for solving this problem with preconditioner $\mathcal{P}_1$ to a tolerance of 10^{-6}, for a variety of h and β. In Table 4.2, we show the number of iterations and CPU times for solving the same problem, using MINRES with preconditioner $\mathcal{P}_2$, to the same tolerance. Finally in Tables 4.3 and

[8]The CPU times include the time taken to construct the matrices M_p and K_p involved in the preconditioner. We construct these matrices in the same way as in the Incompressible Flow & Iterative Solver Software (IFISS) package [5, 21]. Where appropriate we follow the recipe detailed in [6, Chapter 8] of imposing a Dirichlet boundary condition in the matrix K_p at the node on the velocity space corresponding to the *inflow boundary condition*.

4.4, we show the iteration count and CPU times for solving the problem to the same tolerance, with the GMRES algorithm used in the Incompressible Flow & Iterative Solver Software (IFISS) package [5, 21], preconditioned with the matrices $\mathcal{P}_3$ and $\mathcal{P}_4$.[9] In Figures 4.1 and 4.2, we display solutions to the test problem for velocity and pressure, for different values of β. In each of the tables and figures, the value of h indicated corresponds to the spacing between **Q2**-nodes.

		Size	β: 10^2	1	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
h	2^{-3}	1,318	80 0.281	80 0.283	60 0.216	44 0.189	36 0.156	(32)* (0.290)	(26)* (0.232)
	2^{-4}	4,934	84 0.755	85 0.766	66 0.601	52 0.488	37 0.475	(32)* (1.59)	(26)* (1.38)
	2^{-5}	19,078	88 3.03	90 3.08	70 2.41	58 2.04	44 2.05	32 1.54	(28)* (7.42)
	2^{-6}	75,014	86 12.6	90 13.6	74 11.0	62 10.3	50 7.96	33 8.39	(28)* (40.1)
	2^{-7}	297,478	86 62.0	88 58.8	76 53.5	66 46.2	54 39.3	40 29.2	26 28.1

TABLE 4.1
Number of iterations and CPU times (in seconds) when applying MINRES *to the first test problem with preconditioner $\mathcal{P}_1$, for a variety of h and β.*

		Size	β: 10^2	1	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
h	2^{-3}	1,318	112 0.502	107 0.480	85 0.388	59 0.317	42 0.222	(30)* (0.316)	(25)* (0.249)
	2^{-4}	4,934	125 1.51	123 1.49	97 1.18	68 0.847	48 0.768	(33)* (1.87)	(25)* (1.52)
	2^{-5}	19,078	142 6.68	137 6.42	102 4.79	75 3.59	60 3.54	39 2.37	(27)* (7.33)
	2^{-6}	75,014	156 32.5	148 30.9	106 22.3	80 17.2	67 14.1	48 17.0	(29)* (46.5)
	2^{-7}	297,478	165 141	160 138	106 91.2	84 80.5	72 61.9	54 49.4	34 89.6

TABLE 4.2
Number of iterations and CPU times (in seconds) when applying MINRES *to the first test problem with preconditioner $\mathcal{P}_2$, for a variety of h and β.*

When generating these results, we use 20 steps of Chebyshev semi-iteration to approximate the inverse of mass matrices (see [24] for more details). To approximate the inverses of K_p, $M + \sqrt{\beta}K$ and $K + \frac{1}{\sqrt{\beta}}M$ in our preconditioners (note that

[9] All results in Tables 4.1–4.4 are obtained using a tri-core 2.5 GHz workstation.

		Size	β						
			10^2	1	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
h	2^{-3}	1,318	64	62	53	44	39	$(33)^*$	$(28)^*$
			0.238	0.247	0.195	0.188	0.184	(0.287)	(0.263)
	2^{-4}	4,934	65	63	56	50	41	$(38)^*$	$(31)^*$
			0.671	0.674	0.573	0.516	0.565	(1.98)	(1.65)
	2^{-5}	19,078	63	63	56	53	48	38	$(35)^*$
			2.54	2.53	2.26	2.11	2.81	2.27	(9.42)
	2^{-6}	75,014	63	61	57	54	51	41	$(37)^*$
			13.7	12.5	13.1	11.5	10.9	13.8	(58.1)
	2^{-7}	297,478	63	62	56	52	52	48	39
			60.8	62.8	55.3	45.1	51.8	43.8	45.5

Table 4.3

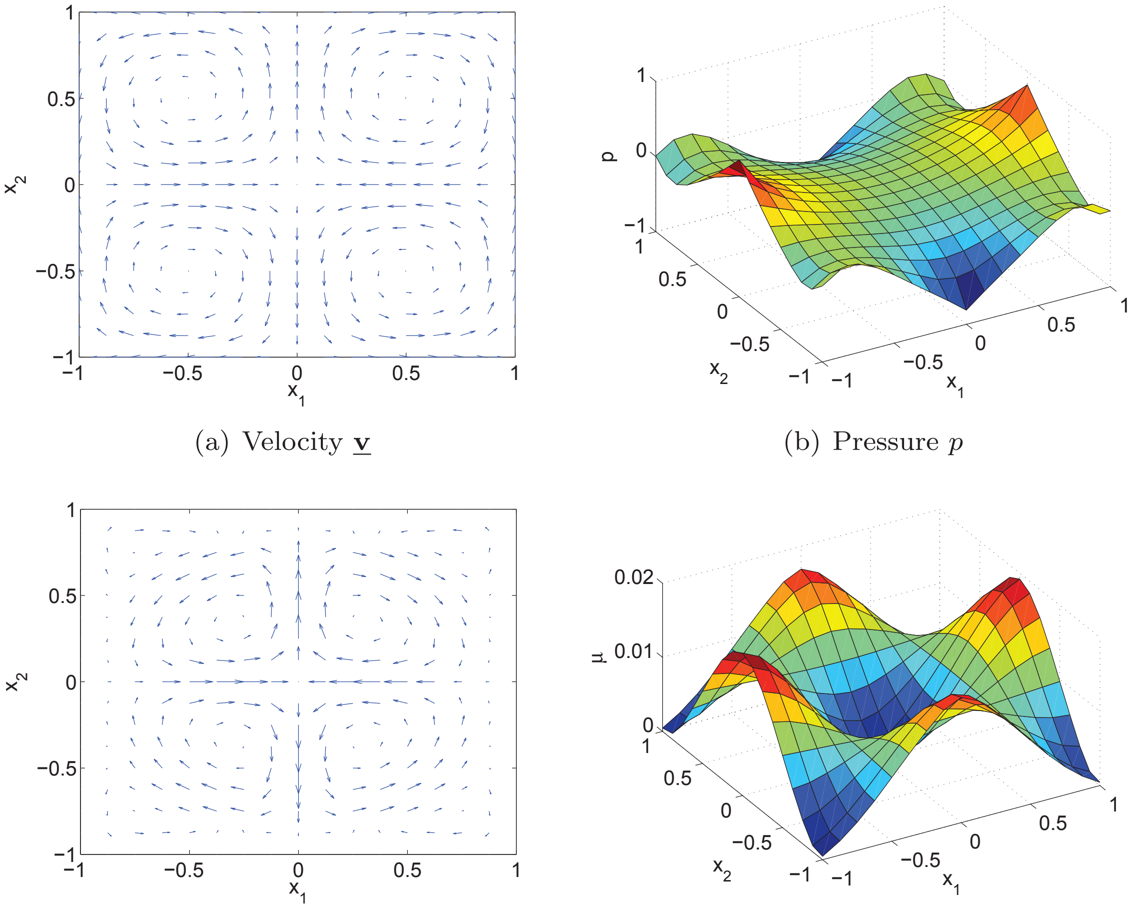
Number of iterations and CPU times (in seconds) when applying Gmres *to the first test problem with preconditioner* $\mathcal{P}_3$*, for a variety of* h *and* β*.*

		Size	β						
			10^2	1	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
h	2^{-3}	1,318	91	85	67	46	26	$(19)^*$	$(14)^*$
			0.755	0.675	0.529	0.415	0.238	(0.376)	(0.261)
	2^{-4}	4,934	107	101	79	59	34	$(24)^*$	$(15)^*$
			2.55	2.38	1.84	1.38	1.02	(2.47)	(1.63)
	2^{-5}	19,078	123	114	88	73	47	29	$(21)^*$
			11.7	10.8	8.20	6.75	5.42	3.42	(11.7)
	2^{-6}	75,014	138	131	99	81	62	37	$(25)^*$
			63.5	58.5	43.7	37.0	27.6	24.1	(74.8)
	2^{-7}	297,478	156	150	109	89	73	48	30
			327	287	224	161	130	92.3	148

Table 4.4

Number of iterations and CPU times (in seconds) when applying Gmres *to the first test problem with preconditioner* $\mathcal{P}_4$*, for a variety of* h *and* β*.*

the last two matrices are the same up to a multiplicative factor), we employ the algebraic multigrid (AMG) routine `HSL_MI20` from the Harwell Subroutine Library (HSL) [2], using 2 V-cycles with 2 pre- and post- (relaxed Jacobi) smoothing steps to approximate each matrix inverse. In all tables in this section, the symbol * denotes that the coarsening of the AMG routine failed when applied to $M+\sqrt{\beta}K$ or $K+\frac{1}{\sqrt{\beta}}M$ – this occurs in the specific and impractical parameter regime where h is large and β is small, and is caused by the presence of positive off-diagonal entries. In these cases, we present results obtained using direct solves rather than AMG.



(a) Velocity $\underline{\mathbf{v}}$ (b) Pressure p

(c) Adjoint velocity $\underline{\boldsymbol{\lambda}}$ (d) Adjoint pressure μ

FIG. 4.3. *Plots of computed solution to the second test problem with* $\beta = 10^{-4}$.

To test our methods further, we also consider the following second test problem:

$$\begin{aligned} \min_{\underline{\mathbf{v}},\underline{\mathbf{u}}} \ & \frac{1}{2}\left\|\underline{\mathbf{v}} - \widehat{\underline{\mathbf{v}}}\right\|^2_{L_2(\Omega)} + \frac{\beta}{2}\left\|\underline{\mathbf{u}}\right\|^2_{L_2(\Omega)} \\ \text{s.t.} \ & -\nabla^2\underline{\mathbf{v}} + \nabla p = \underline{\mathbf{u}}, \quad \text{in } \Omega, \\ & -\nabla\cdot\underline{\mathbf{v}} = 0, \quad \text{in } \Omega, \\ & \underline{\mathbf{v}} = \widehat{\underline{\mathbf{v}}}, \quad \text{on } \partial\Omega, \end{aligned}$$

where

$$\widehat{\underline{\mathbf{v}}} = \begin{cases} \left[-\left(\frac{1}{2} - x_2\right)x_1(1+x_1), \left(\frac{1}{2} + x_1\right)x_2(1-x_2)\right]^T & \text{in } [-1,0]\times[0,1], \\ \left[-\left(\frac{1}{2} - x_2\right)x_1(1-x_1), \left(\frac{1}{2} - x_1\right)x_2(1-x_2)\right]^T & \text{in } [0,1]\times[0,1], \\ \left[-\left(\frac{1}{2} + x_2\right)x_1(1+x_1), \left(\frac{1}{2} + x_1\right)x_2(1+x_2)\right]^T & \text{in } [-1,0]\times[-1,0], \\ \left[-\left(\frac{1}{2} + x_2\right)x_1(1-x_1), \left(\frac{1}{2} - x_1\right)x_2(1+x_2)\right]^T & \text{in } [0,1]\times[-1,0], \end{cases}$$

and $\mathbf{x} = [x_1, x_2]^T$ denotes the spatial coordinates. The target state $\widehat{\underline{\mathbf{v}}}$ within this problem set-up corresponds to a recirculating flow, with symmetry built into the problem. In Figure 4.3, we display solution plots for this problem, and in Tables 4.5 and 4.6 we present numerical results for solving this problem using MINRES preconditioned with $\mathcal{P}_1$ and $\mathcal{P}_2$. Although we do not present results for our GMRES-based solvers for

			β						
		SIZE	10^2	1	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
h	2^{-3}	1,318	50 0.181	58 0.222	58 0.221	46 0.204	38 0.201	(32)* (0.337)	(26)* (0.240)
	2^{-4}	4,934	54 0.518	62 0.594	62 0.590	50 0.491	39 0.532	(32)* (1.72)	(26)* (1.61)
	2^{-5}	19,078	56 2.05	64 2.36	64 2.31	54 1.97	42 2.09	32 1.71	(24)* (7.15)
	2^{-6}	75,014	54 8.96	68 11.3	68 10.4	58 9.69	44 7.07	30 8.63	(22)* (39.3)
	2^{-7}	297,478	52 39.4	68 51.8	70 51.8	60 42.2	46 32.4	28 22.1	21 28.4

TABLE 4.5

Number of iterations and CPU times (in seconds) when applying MINRES *to the second test problem with preconditioner* $\mathcal{P}_1$*, for a variety of* h *and* β*.*

			β						
		SIZE	10^2	1	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
h	2^{-3}	1,318	89 0.423	90 0.440	79 0.376	58 0.315	42 0.235	(31)* (0.341)	(25)* (0.276)
	2^{-4}	4,934	100 1.26	100 1.26	85 1.07	65 0.835	49 0.813	(31)* (1.87)	(25)* (1.54)
	2^{-5}	19,078	106 5.14	107 5.16	86 4.17	70 3.43	56 3.42	34 2.17	(23)* (7.36)
	2^{-6}	75,014	116 25.0	116 25.2	89 19.4	74 16.1	58 12.7	35 11.7	(21)* (39.3)
	2^{-7}	297,478	125 120	125 124	95 94.1	78 76.9	59 58.6	33 35.6	23 57.3

TABLE 4.6

Number of iterations and CPU times (in seconds) when applying MINRES *to the second test problem with preconditioner* $\mathcal{P}_2$*, for a variety of* h *and* β*.*

this problem, we note that the numerical features of these solvers are similar to those when tested on the first test problem.

The results shown in Tables 4.1–4.6 indicate that the 4 preconditioners discussed in this manuscript are robust with respect to mesh-size and regularization parameter.[10] The iteration count is low for all 4 solvers considering the complexity of the problems considered. In many practical problems, the value of β is within the range $[10^{-6}, 10^{-1}]$; all methods perform well in this regime. We note that the block diagonal preconditioner $\mathcal{P}_1$ (introduced in [25]), and the block triangular preconditioner $\mathcal{P}_3$ based on it, solve the problem in the shortest time in all cases considered, and the

[10] The only parameter regime where we do not observe complete robustness is that of very small β, when we observe some degradation in the performance of the AMG routine used.

lowest iteration count in most cases. However, the strategy involved in constructing these preconditioners is highly specific to this problem. We believe that the flexibility in the methodology used to construct $\mathcal{P}_2$ and $\mathcal{P}_4$ would enable us to consider the more general and much harder Navier-Stokes control problem, and therefore it is important to note that these preconditioners also seem to achieve robustness, albeit with larger iteration counts and CPU times than $\mathcal{P}_1$ and $\mathcal{P}_3$.

Of the two preconditioners $\mathcal{P}_2$ and $\mathcal{P}_4$, we note that the preconditioner $\mathcal{P}_4$ solves the problem in fewer iterations than $\mathcal{P}_2$, but greater CPU time due to the added complexity of the Gmres algorithm (though this could partially be offset by using restarts within the Gmres method). We find that in the Navier-Stokes control case, using preconditioners of the form $\mathcal{P}_2$ and $\mathcal{P}_4$ would result in convergence to the solution of the matrix systems involved in similar CPU times [15], because a non-symmetric solver such as Gmres has to be used in both cases, as both equivalent preconditioners would be non-symmetric in the Navier-Stokes case. We also note that in the parameter regime of small β, the iteration count when the preconditioner $\mathcal{P}_4$ is used is even smaller than that when $\mathcal{P}_1$ (or indeed $\mathcal{P}_2$) is applied. We believe that to extend this methodology to obtain an effective solver for the analogous Navier-Stokes control problem, a preconditioner of the form of either $\mathcal{P}_2$ or $\mathcal{P}_4$ can therefore be considered.

5. Concluding Remarks. The use of commutator arguments has been an extremely valuable tool when developing iterative methods for problems in fluid dynamics. In [10] for instance, such an argument was applied in order to develop a solver for the Navier-Stokes equations which performed well for a wide range of values of mesh-size and viscosity. Since then, such arguments have also been applied to good effect when deriving iterative schemes for PDE-constrained optimization problems, for example in [22] to obtain mesh-independent solvers for time-dependent Stokes control, and in [25] to arrive at a mesh- and regularization-robust solver for a class of Stokes control problems. Also, in [7], commutator arguments for the Navier-Stokes equations are analyzed for a range of boundary conditions. In this manuscript, we have used new commutator arguments to derive further mesh- and regularization-robust solvers for these problems: block diagonal and block triangular.

We have also explained the role of saddle point theory, and that of preconditioners for the Poisson control problem, in generating solvers for the more difficult Stokes control problem. We provided numerical results to justify the potency of this approach, and explained the importance of the pressure regularization term (or lack of it) from an iterative solver point of view. We believe that the arguments we have introduced in this manuscript may be extended to generate robust solvers for a class of the harder Navier-Stokes control problems – we will discuss this in a future paper [15]. In addition, future research in this area could include the application of these techniques to problems with state or control constraints, boundary control problems, or time-dependent Stokes-type problems, as well as tackling optimal control problems derived specifically from real-world data.

Acknowledgements. The author would like to thank Andy Wathen for his invaluable help and advice while this manuscript was being prepared, as well as Martin Stoll and Walter Zulehner for useful discussions about this work. He is also very grateful to two anonymous referees for their careful reading of the manuscript and helpful comments. The author was supported for this work by the Engineering and Physical Sciences Research Council (UK), Grant EP/P505216/1.

REFERENCES

[1] M. Benzi, G. H. Golub, and J. Liesen, *Numerical Solution of Saddle Point Problems*, Acta Numerica, 14, pp.1–137, 2005.

[2] J. Boyle J, M. D. Mihajlovic, and J. A. Scott, *HSL_MI20: An Efficient AMG Preconditioner for Finite Element Problems in 3D*, International Journal for Numerical Methods in Engineering, 82, pp.64–98, 2010.

[3] J. H. Bramble, and J. E. Pasciak, *A Preconditioning Technique for Indefinite Systems Resulting from Mixed Approximations of Elliptic Problems*, Mathematics of Computation, 50(181), pp.1–17, 1988.

[4] J. Cahouet, and J.-P. Chabard, *Some Fast 3D Finite Element Solvers for the Generalized Stokes Problem*, International Journal for Numerical Methods in Fluids, 8, pp.869–895, 1988.

[5] H. C. Elman, A. Ramage, and D. Silvester, *Algorithm 866: IFISS, a Matlab Toolbox for Modelling Incompressible Flow*, ACM Transactions on Mathematical Software, 33, pp.2–14, 2007.

[6] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Numerical Mathematics and Scientific Computation Series. Oxford University Press: New York, 2005.

[7] H. C. Elman, and R. S. Tuminaro, *Boundary Conditions in Approximate Commutator Preconditioners for the Navier-Stokes Equations*, Electronic Transactions on Numerical Analysis, 35, pp. 257-280, 2009.

[8] M. Hinze, M. Köster, and S. Turek, *A Hierarchical Space-Time Solver for Distributed Control of the Stokes Equation*, Priority Programme 1253, Preprint Number SPP1253-16-01, 2008.

[9] I. C. F. Ipsen, *A Note on Preconditioning Non-Symmetric Matrices*, SIAM Journal on Scientific Computing, 23(3), pp.1050–1051, 2001.

[10] D. Kay, D. Loghin, and A. Wathen, *A Preconditioner for the Steady-State Navier-Stokes Equations*, SIAM Journal on Scientific Computing, 24(1), pp.237–256, 2002.

[11] W. Krendl, V. Simoncini, and W. Zulehner, *Stability Estimates and Structural Spectral Properties of Saddle Point Problems*, Technical Report 2012-03, Institute of Computational Mathematics, Johannes Kepler University Linz, Austria, to appear in Numerische Mathematik.

[12] Y. A. Kuznetsov, *Efficient Iterative Solvers for Elliptic Finite Element Problems on Nonmatching Grids*, Russian Journal of Numerical Analysis and Mathematical Modelling, 10(3), pp.187–211, 1995.

[13] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A Note on Preconditioning for Indefinite Linear Systems*, SIAM Journal on Scientific Computing, 21(6), pp.1969–1972, 2000.

[14] C. C. Paige, and M. A. Saunders, *Solutions of Sparse Indefinite Systems of Linear Equations*, SIAM Journal on Numerical Analysis, 12(4), pp.617–629, 1975.

[15] J. W. Pearson, *Preconditioned Iterative Methods for Navier-Stokes Control Problems*, submitted to SIAM Journal on Scientific Computing, 2013.

[16] J. W. Pearson, and A. J. Wathen, *Fast Iterative Solvers for Convection-Diffusion Control Problems*, Technical Report 1410, Mathematical Institute, University of Oxford, to appear in Electronic Transactions on Numerical Analysis, 2011.

[17] J. W. Pearson, and A. J. Wathen, *A New Approximation of the Schur Complement in Preconditioners for PDE-Constrained Optimization*, Numerical Linear Algebra with Applications, 19(5), pp.816–829, 2012.

[18] T. Rees, and A. J. Wathen, *Preconditioning Iterative Methods for the Optimal Control of the Stokes Equations*, SIAM Journal on Scientific Computing, 33, pp.2903–2926, 2011.

[19] Y. Saad, and M. H. Schultz, *GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems*, SIAM Journal on Scientific Computing, 7(3), pp.856–869, 1986.

[20] J. Schöberl, and W. Zulehner, *Symmetric Indefinite Preconditioners for Saddle Point Problems with Applications to PDE-Constrained Optimization Problems*, SIAM Journal on Matrix Analysis and Applications, 29(3), pp.752–773, 2007.

[21] D. Silvester, H. Elman, and A. Ramage, *Incompressible Flow and Iterative Solver Software (IFISS)*, version 3.1, `http://www.manchester.ac.uk/ifiss/`, 2011.

[22] M. Stoll, and A. Wathen, *All-at-Once Solution of Time-Dependent Stokes Control*, Journal of Computational Physics, 232, pp.498–515, 2013.

[23] A. Wathen, and D. Silvester, *Fast Iterative Solution of Stabilised Stokes Systems Part I: Using Simple Diagonal Preconditioners*, SIAM Journal on Numerical Analysis, 30, pp.630–649, 1993.

[24] A. J. Wathen, and T. Rees, *Chebyshev Semi-Iteration in Preconditioning for Problems Including the Mass Matrix*, Electronic Transactions on Numerical Analysis, 34, pp.125–135, 2009.

[25] W. Zulehner, *Nonstandard Norms and Robust Estimates for Saddle Point Problems*, SIAM Journal on Matrix Analysis and Applications, 32(2), pp.536–560, 2011.