FAST SOLVERS FOR STOKES EQUATIONS

LONG CHEN

CONTENTS

1. INTROUDCTION

A good numerical method consists of two parts: one is the stability and accuracy of the approximation; another is an efficient way to compute this accurate approximation. In this notes, we consider efficient iterative methods for solving the discrete Stokes equation

$$\begin{pmatrix} A & B^\intercal \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}. \tag{1}$$

The system could be obtained by any stable pair using finite element or finite difference methods; see

- Finite Element Methods for Stokes Equations.
- Finite Difference Method for Stokes Equations: MAC Scheme.

The Stokes system can be factored as

$$\begin{pmatrix} A & B^\intercal \\ B & 0 \end{pmatrix} = \begin{pmatrix} A & 0 \\ B & I \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & -S \end{pmatrix} \begin{pmatrix} A & B^\intercal \\ 0 & I \end{pmatrix},$$

where $S = BA^{-1}B^\intercal$ is the Schur complement of A. Therefore by Sylverster's law of inertia, the Stokes system is a saddle point system which is much harder to solve comparing with symmetric and positive definite systems.

Date: Create: Oct. 2012; Last update: April 12, 2018.

The variable pressure p can be treat as a Lagrange multiplier to impose the divergence free constraint $\operatorname{div} u = 0$. The saddle point system (1) is the first order system of the following constraint optimization problem:

$$\min_{u,\operatorname{div} u=g} \frac{1}{2}(Au,u) - (f,u). \tag{2}$$

Methods for constraint optimization problems can then be applied. The key of multilevel methods in this approach is to construct a multilevel decomposition of the constraint. See, for example, [17, 2, 3].

We can also eliminate u to get the equation for p

$$Sp = BA^{-1}f - g. \tag{3}$$

Iterative methods can be construct by solving (3). It can be shown S is a well conditioned matrix/operator and thus (3) can be solved by well established iterative methods. The difficulty of this approach is that the Schur complement is expensive to form explicitly.

2. A Basic Iterative Method

We present a basic iterative method by solving u and p alternatively. Starting from an initial guess u^0, p^0, one iteration going from (u^k, p^k) to (u^{k+1}, p^{k+1}) is

(1) Fix p^k, solve for u^{k+1};
(2) Fix u^{k+1}, solve for p^{k+1}.

When p is fixed, there are two equations for u. We can solve the momentum equation to get u (the term $B^\intercal p$ is moved to the right hand side and treat as known), i.e., $Au = f - B^\intercal p$.

The trouble is on the update of pressure. Various methods differ in the way of updating pressure. The fundamental question is: what is the equation for pressure? The mass equation does not contain p. The momentum equation involves p but it is an over-determined system for p:

$$B^\intercal p = f - Au. \tag{4}$$

A natural way to solve an over-determined system is the least square method. Let M be an SPD operator on the velocity space. We solve equation (4) in the least square sense by minimizing the residual in the norm defined by M^{-1}. Namely we compute the solution to the minimization problem

$$\min_p \|f - Au - B^\intercal p\|_{M^{-1}} \tag{5}$$

whose solution is given by solving

$$BM^{-1}B^\intercal p = BM^{-1}(f - Au). \tag{6}$$

Thanks to the inf-sup condition, $B^\intercal$ is into and thus (6) is well defined.

Exercise 2.1. Prove that the solution of the least square problem (5) is given by (6). □

For Stokes equations, two special choices of M is of particular interest. The first one is $M = I$. If we further assume there exists an operator A_p such that $BA = A_pB$ (such A_p will be discussed in detail later), then $BAu = A_pBu = A_pg$ and equation (6) can be simplified to

$$BB^\intercal p = Bf - A_pg. \tag{7}$$

The operator $BB^\intercal = -\operatorname{div}\operatorname{grad}$ is a discretization of $-\Delta_p$ for pressure with a homogenous Neumann boundary condition (recall that in the original Stokes equations, the velocity has a Dirichlet boundary condition).

The existence of A_p such that commutator $E = BA - A_pB$ is vanished is not obvious. Indeed for the Dirichlet boundary condition of u, the pressure p will not satisfy the equation (7) due to the mis-match of the boundary condition. Nevertheless solving (7) can give a good approximation of the pressure. And (7) is an important component of the projection method which is still the dominated solver for time-dependent Stokes equations. See Section 3 for detailed discussion.

The second interesting choice is $M = A$. It is equivalent to eliminating u from the momentum equation and substituting into the mass equation. The resulting equation is the so-called Schur complement equation for p:

$$BA^{-1}B^{\intercal}p = BA^{-1}f - g. \tag{8}$$

Uzawa method can be interpret as an iterative method for solving equation (8). See Section 2 for detailed formulation and convergence analysis. Note that we do not want to form and store A^{-1} which requires $\mathcal{O}(N^2)$ algorithms with N being the size of matrix A.

We summarize our basic iterative method with exact solver at each step as follows. Given current approximation (u^k, p^k),

(1) Solve the momentum equation:
$$u^{k+1/2} = A^{-1}(f - B^{\intercal}p^k);$$
(2) Project $u^{k+1/2}$ to satisfy the constraint $Bu^{k+1} = g$:
$$u^{k+1} = P_B u^{k+1/2};$$
(3) Update pressure by solving $B^{\intercal}p^{k+1} = f - Au^{k+1}$ in the least square sense:
$$p^{k+1} = (BM^{-1}B^{\intercal})^{-1}BM^{-1}(f - Au^{k+1}).$$

The convergence analysis of this iterative method is not obvious. Even in the convergent case, we do not recommend the current form since each step an elliptic equation should be solved (for $M = A$, it is even not practical since forming and storing A^{-1} is expensive). Instead, we use this basic iterative method as a prototype to design efficient smoothers or preconditioners.

In the second step, we project u to satisfy the constraint using the projection operator P_B. If we include Step (2), it is in the type of projection methods. As an iterative solver of Stokes equations, the projection step is not necessary and without Step (2), it is in the type of Uzawa methods.

3. Uzawa-type Methods

3.1. **Uzawa method.** We shall interpret Uzawa method [19] as a Richardson method for solving the Schur complement equation

$$S\,p = \tilde{g}, \tag{9}$$

where $S = BA^{-1}B^{\intercal}$ and $\tilde{g} = BA^{-1}f - g$. Equation (9) is obtained by eliminating u from the first equation of (1). One iteration of Richardson method for solving (9) reads as

$$p^{k+1} = p^k + \omega(\tilde{g} - Sp^k). \tag{10}$$

Why not use more efficient iterative methods such as Jacobi or Gauss-Seidel methods to solve (10)? The reason is that we do not have the explicit form of S which involves A^{-1} in the definition. Although A is sparse, in general A^{-1} is not. Form A^{-1} requires $\mathcal{O}(N^2)$ complexity and memory.

For the same reason, the evaluation of the residual $\tilde{g} - Sp^k$ should be computed iteratively. By the definition of S and $\tilde{g}$, one easily verify that

$$\tilde{g} - Sp^k = BA^{-1}f - BA^{-1}B^\intercal p^k - g = Bu^{k+1} - g.$$

Here $u^{k+1} = A^{-1}(f - B^\intercal p^k)$. We summarize the Uzawa method below.

Algorithm (UM) $[u^{k+1}, p^{k+1}] \leftarrow \texttt{UzawaMethod}(u^k, p^k, f, g)$

(1) Update velocity by solving the momentum equation

$$u^{k+1} = A^{-1}(f - B^\intercal p^k),$$

(2) Update pressure

$$p^{k+1} = p^k + \omega(Bu^{k+1} - g).$$

Remark 3.1. When implement the Uzawa method for the matrix equation, operator ωI should be changed to matrix ωM_p^{-1}, where M_p is the mass matrix representing the L^2-inner product for the pressure space. See Remark 3.6. □

Recall that the error equation for Richardson method is

$$p - p^{k+1} = (I - \omega S)(p - p^k),$$

and the method converges if and only if the spectral radius $\rho(I - \omega S) < 1$. Since S is SPD, we know when $0 < \omega < 2/\lambda_{\max}(S)$, the method converges. The optimal choice of ω is

$$\omega_{\text{opt}} = \frac{2}{\lambda_{\min}(S) + \lambda_{\max}(S)},$$

and corresponding convergence rate is

$$\|p - p^k\| \leq \left(\frac{\kappa(S) - 1}{\kappa(S) + 1}\right)^k \|p - p^0\|.$$

Therefore the key is to estimate the spectrum of S and the condition number $\kappa(S)$. We begin with the following characterization of the norm $\|p\|_S = (Sp, p)^{1/2}$.

Lemma 3.2. *For $S = BA^{-1}B^\intercal$, we have*

$$\|p\|_S = \sup_{v \in \mathbb{V}} \frac{(Bv, p)}{\|v\|_A}.$$

Proof. Let $u = A^{-1}B^\intercal p$. Then

$$\|p\|_S^2 = (Sp, p) = (BA^{-1}B^\intercal p, p) = (u, B^\intercal p) = (u, A^{-1}B^\intercal p)_A = \|u\|_A^2.$$

Now we can identify $u \in \mathbb{V}'$ by the Riesz map, i.e., $(u, \cdot)_A$ defines a functional on $\mathbb{V}$. Then

$$\|u\|_A = \sup_{v \in \mathbb{V}} \frac{(u, v)_A}{\|v\|_A} = \sup_{v \in \mathbb{V}} \frac{(B^\intercal p, v)}{\|v\|_A} = \sup_{v \in \mathbb{V}} \frac{(Bv, p)}{\|v\|_A}.$$

□

Lemma 3.3. *Assuming the inf-sup condition*

$$\inf_{q \in \mathbb{P}} \sup_{v \in \mathbb{V}} \frac{(Bv, p)}{\|v\|_A \|q\|} = \beta > 0,$$

then

$$\lambda_{\min}(S) = \beta^2.$$

Proof. By the definition of $\lambda_{\min}(S)$ and Lemma 3.2, we have

$$\lambda_{\min}(S) = \inf_{p\in\mathbb{P}} \frac{(Sp,p)}{(p,p)} = \left[\inf_{q\in\mathbb{P}}\sup_{v\in\mathbb{V}} \frac{(Bv,p)}{\|v\|_A\|p\|}\right]^2. \tag{11}$$

□

From this result, the inf-sup condition for the original mixed system is equivalent to the Poincaré type inequality $\beta\|p\| \le \|p\|_S$.

Lemma 3.4. *For Stokes equations, we have* $\lambda_{\max}(S) = 1$.

Proof. This is an easy consequence of the inequality

$$\|\operatorname{div} u\| \le \|\nabla u\|, \tag{12}$$

and the fact the equality holds when $\operatorname{curl} u = 0$. □

The optimal choice of the parameter ω is then $2/(1+\beta)$. But since β is unknown, a practical choice is $\omega = 1$ and the corresponding rate is $1-\beta^2$. Note that the error operator $I-\omega S$ is also symmetric in the S-norm. The convergence analysis remains the same when the error is measured in S-norm. We summarize in the following theorem.

Theorem 3.5. *The Uzawa method converges for* $\omega = 1$ *and the convergent rate is* $1-\beta^2$ *for both the* L^2*-norm and* S*-norm, i.e.,*

$$\|p-p^k\| \le (1-\beta^2)^k\|p-p^0\|,$$
$$\|p-p^k\|_S \le (1-\beta^2)^k\|p-p^0\|_S.$$

How about the convergence of the velocity? The pair (u^{k+1},p^k) will satisfy the momentum equation and thus the error relation is

$$A(u-u^{k+1}) + B(p-p^k) = 0,$$

which implies $\|u-u^{k+1}\|_A = \|p-p^k\|_S$. As a consequence, we also have the contraction of the velocity

$$\|u-u^{k+1}\|_A \le (1-\beta^2)^k\|u-u^1\|_A.$$

Remark 3.6. In the matrix realization, the inequality

$$\beta^2(q,q) \le (Sq,q) \le (q,q)$$

is translated to

$$\beta^2 \boldsymbol{q}^\intercal M_p \boldsymbol{q} \le \boldsymbol{q}^\intercal S \boldsymbol{q} \le \boldsymbol{q}^\intercal M_p \boldsymbol{q},$$

where M_p is the mass matrix of the pressure. That is matrix S is spectral equivalent to matrix M_p. When implement the Uzawa method for the matrix equation, the operator ωI should be changed to ωM_p^{-1}.

Since M_p is well conditioned, we can use PCG to efficiently invert M_p or simply use the diagonal of M_p (with an appropriately modified step size ω). □

The inf-sup constant β depends only on the domain and thus the contraction rate $1-\beta^2$ is uniform to the size of the problem. However β is unknown or hard to estimate. And when it is close to zero, a slow convergence of Uzawa method can be observed.

3.2. **Augmented Lagrangian method.** Augmented Lagrangian method (ALM) [9] is a way to speed up the convergence of the Uzawa method. Given a parameter $r > 0$, multiplying $rB^\intercal$ to the second equation of (1) and adding it to the first one, we obtain an equivalent system

$$\begin{pmatrix} A + rB^\intercal B & B^\intercal \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f + rB^\intercal g \\ g \end{pmatrix}. \tag{13}$$

The parameter r is used to shift the spectrum of the Schur complement.

Lemma 3.7. *Let $S = BA^{-1}B^\intercal$ and $S_r = B(A + rB^\intercal B)^{-1}B^\intercal$. Then*

$$S_r = (rI + S^{-1})^{-1}, \tag{14}$$

and thus

$$\sigma(S_r) = \left\{ \frac{\lambda}{1 + r\lambda}, \lambda \in \sigma(S) \right\}.$$

Proof. The identity is based on the Sherman-Morrison-Woodbury formula [13, 14]

$$(A + rB^\intercal B)^{-1} = A^{-1} - rA^{-1}B^\intercal(I + rB^\intercal A^{-1}B)^{-1}BA^{-1}.$$

more from eigenvalue analysis. □

Theorem 3.8. *For $0 < \omega < 2r$, Uzawa method for augmented system (13) will converge and for $\omega = r$*

$$\|p - p^k\| \leq \left(\frac{1}{1 + \beta^2 r} \right)^k \|p - p^0\|.$$

Comparing with the standard Uzawa method, the rate is improved to $1/r$ asymptotically if we choose $r \gg 1$. Then the augmented Lagrangian Uzawa method could converge in few (say 2 or 3) steps. The trade-off is computing A_r^{-1} becomes extremely difficult. When r is big, the dominated operator $rB^\intercal B$ contains a large kernel. Special multigrid methods should be designed to take care of this large null space.

Remark 3.9. In the implementation, the operator form $B^\intercal B$ is better change to $B^\intercal M_p^{-1} B$. Recall that $\langle Bu, p\rangle = -(\operatorname{div} u, p)$, i.e., in matrix form $-\operatorname{div} = M_p^{-1}B$ and thus the inner product $(\operatorname{div} u, \operatorname{div} v)$ is $B^\intercal M_p^{-1} B$. □

3.3. **Inexact Uzawa method.** In each Uzawa iteration, an elliptic solver for computing A^{-1} is needed. Usually iterative methods, such as multigrid, are used. The iteration for computing A^{-1} will be called the inner iteration and the Uzawa iteration will be called outer iteration.

We could use an inexact solver when computing A^{-1} to reduce the computational cost at each Uzawa iteration. This inexact solve will increase the outer iteration steps.

Algorithm (IUM) $[u^{k+1}, p^{k+1}] \leftarrow$ `InexactUzawaMethod`(u^k, p^k, f, g)

(1) Update velocity by solving the momentum equation approximately

$$u^{k+1} = u^k + C(f - Au^k - B^\intercal p^k).$$

(2) Update pressure

$$p^{k+1} = p^k - D^{-1}(g - Bu_{k+1}).$$

We now sketch a linear algebraic proof of the convergence of IUM following [5]. In the Uzawa method, the two variables u, p can be decoupled and the analysis is reduced to p only. In IUM, these two variables are coupled and the analysis is much more involved.

- Change the error measure for u from $\|e_u\|_A$ to $\|r_u\|_{A^{-1}}$.
- Symmetrize the error operator.
- Find a right scaled norm of pressure component.

We first write out the error equations. Let $e_u^k = u - u^k$ and $e_p^k = p - p^k$. Then

$$\begin{pmatrix} e_u^{k+1} \\ e_p^{k+1} \end{pmatrix} = \begin{pmatrix} I - CA & -CB^\intercal \\ D^{-1}B & I \end{pmatrix} \begin{pmatrix} e_u^k \\ e_p^k \end{pmatrix}, \tag{15}$$

and will be denoted by $x_{k+1} = \mathcal{E} x_k$ in short. The error matrix $\mathcal{E}$ in (15) is non-symmetric and full. In UM, $I - CA = 0$ and then variables (u, p) are decoupled.

We try to transform the error equation to get a symmetric error matrix. Consider a generic error equation

$$x_{k+1} = E x_k.$$

When E is symmetric, the spectral radius $\rho(E)$ is related to the l^2-norm of x_k. We can consider the convergence in another (more appropriate) norm, e.g. $\|x\|_A = \|A^{1/2}x\|$ with an SPD operator A. So the error equation can be written as

$$A^{1/2} x_{k+1} = (A^{1/2} E A^{-1/2}) A^{1/2} x_k.$$

The convergence analysis is then reduced to the estimate of $\rho(A^{1/2} E A^{-1/2})$.

For saddle point system (1), an appropriate norm is $\|u\|_A$ and $\|p\|_S$. We can use $\mathcal{D} = \mathrm{diag}(A, S)$ and consider

$$\mathcal{D}^{1/2} \mathcal{E} \mathcal{D}^{-1/2} = \begin{pmatrix} I - A^{1/2} C A^{1/2} & -A^{1/2} C B^\intercal S^{-1/2} \\ S^{1/2} D^{-1} B A^{-1/2} & I \end{pmatrix}.$$

Now the diagonal is symmetric. But the off-diagonal is still non-symmetric.

Return to the SPD case $Ae = r$ and consider $\|r\|_{A^{-1}}^2 = (A^{-1}r, r) = (Ae, e) = \|e\|_A^2$. For saddle point problems $\mathcal{L}e = r$, the negative norm of residual may not the energy norm of the error

$$\|r\|_{\mathcal{D}^{-1}}^2 = (\mathcal{D}^{-1} r, r) = (\mathcal{D}^{-1}\mathcal{L}e, \mathcal{L}e) = (\mathcal{L}^\intercal \mathcal{D}^{-1} \mathcal{L} e, e) \neq (\mathcal{D}e, e).$$

So we shall try to use a negative norm of the residual.

Let us try this for u component first. Note that

$$r_u^k = A e_u^k + B^\intercal e_p^k \tag{16}$$

by definition. The error equation for u can be written as

$$e_u^{k+1} = e_u^k - C r_u^k = (A^{-1} - C) r_u^k + B^\intercal e_p^k.$$

Here in the second step, we solve e_u^k using (16). Substitude the above relation into the error equation for p, the error equation becomes

$$\begin{aligned} \begin{pmatrix} e_u^{k+1} \\ e_p^{k+1} \end{pmatrix} &= \begin{pmatrix} A^{-1} - C & -A^{-1} B^\intercal \\ D^{-1} B (A^{-1} - C) & I - D^{-1} S \end{pmatrix} \begin{pmatrix} r_u^k \\ e_p^k \end{pmatrix} \\ &= \begin{pmatrix} I & -A^{-1} B^\intercal \\ D^{-1} B & I - D^{-1} S \end{pmatrix} \begin{pmatrix} (A^{-1} - C) r_u^k \\ e_p^k \end{pmatrix}. \end{aligned}$$

To change e_u^{k+1} to r_u^{k+1}, we multiply a matrix to the left

$$\begin{pmatrix} A & B^\intercal \\ O & I \end{pmatrix} \begin{pmatrix} e_u^{k+1} \\ e_p^{k+1} \end{pmatrix} = \begin{pmatrix} A & B^\intercal \\ O & I \end{pmatrix} \begin{pmatrix} I & -A^{-1} B^\intercal \\ D^{-1} B & I - D^{-1} S \end{pmatrix} \begin{pmatrix} (A^{-1} - C) r_u^k \\ e_p^k \end{pmatrix},$$

and change e_p^k to $-Se_p^k$ to get

$$\begin{pmatrix} r_u^{k+1} \\ e_p^{k+1} \end{pmatrix} = \begin{pmatrix} A + B^\intercal D^{-1} B & B^\intercal D^{-1} \\ D^{-1} B & D^{-1} - S^{-1} \end{pmatrix} \begin{pmatrix} (A^{-1} - C) r_u^k \\ -S e_p^k \end{pmatrix}. \tag{17}$$

Now the matrix in (19) is symmetric but e_p^{k+1} and Se_p^k are inconsistent.

To apply the spectral analysis, we want to write $B^\intercal D^{-1} B = \bar{B}^\intercal \bar{B}$ which suggests $\bar{B} = D^{-1/2} B$ and $\bar{B}^\intercal = B^T D^{-1/2}$. The corresponding Schur complement of $\bar{S}$ is

$$\bar{S} = \bar{B} A^{-1} \bar{B}^\intercal = D^{-1/2} S D^{-1/2}, \quad S = D^{1/2} \bar{S} D^{1/2}, \quad \bar{S} D^{1/2} = D^{-1/2} S.$$

The matrix $\mathcal{M}$ becomes

$$\begin{pmatrix} A + \bar{B}^\intercal \bar{B} & \bar{B}^\intercal D^{-1/2} \\ D^{-1/2} \bar{B} & D^{-1/2} (I - \bar{S}^{-1}) D^{-1/2} \end{pmatrix}.$$

This motives a scaling of $D^{1/2}$ of e^p and $D^{-1/2} S e_p^k = \bar{S} D^{1/2} e_p^k$. The error equation is

$$\begin{pmatrix} r_u^{k+1} \\ D^{1/2} e_p^{k+1} \end{pmatrix} = \begin{pmatrix} A + \bar{B}^\intercal \bar{B} & \bar{B}^\intercal \\ \bar{B} & I - \bar{S}^{-1} \end{pmatrix} \begin{pmatrix} (A^{-1} - C) r_u^k \\ -\bar{S} D^{1/2} e_p^k \end{pmatrix}. \tag{18}$$

In the exact case, the error operator for e^p is $I - \bar{S}^{-1}$ which is expansive. We need an non expansive like $I - \bar{S}$. So we further introduce a scaling $\bar{S}^{1/2}$ for the pressure component to write the error equation as

$$\begin{pmatrix} r_u^{k+1} \\ \bar{S}^{1/2} D^{1/2} e_p^{k+1} \end{pmatrix} = \begin{pmatrix} A + \bar{B}^\intercal \bar{B} & \bar{B}^\intercal \bar{S}^{1/2} \\ \bar{S}^{1/2} \bar{B} & \bar{S} - I \end{pmatrix} \begin{pmatrix} (A^{-1} - C) r_u^k \\ -\bar{S}^{1/2} D^{1/2} e_p^k \end{pmatrix}. \tag{19}$$

The norm for p component is

$$\|\bar{S}^{1/2} D^{1/2} e_p^k\|^2 = (\bar{S}^{1/2} D^{1/2} e_p^k, \bar{S}^{1/2} D^{1/2} e_p^k) = (D^{1/2} \bar{S} D^{1/2} e_p^k, e_p^k) = (S e_p^k, e_p^k).$$

The quantity for u component is

$$\|(A^{-1} - C) r_u^k\|_A \leq \delta \|r_u^k\|_{A^{-1}}.$$

Here we basically assume the inexact solve of A^{-1} is accurate within δ. To get a symmetric error operator, we can use $\delta^{1/2}$ for u component and arrives the form

$$\begin{pmatrix} \delta^{1/2} r_u^{k+1} \\ \bar{S}^{1/2} D^{1/2} e_p^{k+1} \end{pmatrix} = \begin{pmatrix} \delta(A + \bar{B}^\intercal \bar{B}) & \delta^{1/2} \bar{B}^\intercal \bar{S}^{1/2} \\ \delta^{1/2} \bar{S}^{1/2} \bar{B} & \bar{S} - I \end{pmatrix} \begin{pmatrix} \delta^{-1/2} (A^{-1} - C) r_u^k \\ -\bar{S}^{1/2} D^{1/2} e_p^k \end{pmatrix}. \tag{20}$$

Now we use spectral analysis of the saddle point system to compute $\lambda = \lambda(\delta)$ of the following matrix

$$\begin{pmatrix} \delta(1+\mu) & \delta^{1/2} \mu \\ \delta^{1/2} \mu & \mu - 1 \end{pmatrix},$$

where $\mu \in \sigma(D^{-1} S)$. The formulae is

$$\lambda(\delta, \mu) = \frac{1}{2} \left[(\delta - 1) + (1 + \delta)\mu \pm \sqrt{[\delta - 1 + (1+\delta)\mu]^2 + 4\delta} \right].$$

We conclude in the following theorem.

Theorem 3.10. *Assume* $\|(A^{-1} - C)\phi\|_A \leq \delta \|\phi\|_{A^{-1}}$ *with* $\delta < 1/3$ *and* $\sigma(D^{-1} S) \subseteq [0, 1)$. *Then IUM converges.*

4. Iterative Methods based on Transformation

4.1. **Projection method.** Denote by

$$\mathcal{L} = \begin{pmatrix} A & B^\intercal \\ B & 0 \end{pmatrix}, \text{ and } \mathcal{M} = \begin{pmatrix} I & B^\intercal \\ 0 & -BB^\intercal \end{pmatrix}$$

then

$$\mathcal{L}\mathcal{M} = \begin{pmatrix} A & AB^\intercal - B^\intercal BB^\intercal \\ B & BB^\intercal \end{pmatrix} \approx \begin{pmatrix} A & 0 \\ B & BB^\intercal \end{pmatrix} := \widetilde{\mathcal{L}\mathcal{M}}.$$

By " $\approx$ " here we mean that the commutator $E = AB^\intercal - B^\intercal BB^\intercal$ is small or is of low rank. Thus it can be omitted in order to design relaxation methods.

Let us justify that the commutator E is kind of 'small'. In the continuous level, with certain assumptions on the smoothness and boundary conditions, the $(1,2)$ block of $\mathcal{L}\mathcal{M}$ can be manipulate as

$$(-\Delta + \text{grad div})\text{grad} = \text{curl}\,\text{curl}\,\text{grad} = 0.$$

Here we use the identity

$$-\Delta = -\text{grad div} + \text{curl}\,\text{curl}\,, \tag{21}$$

which holds in H^{-1} topology, and the fact

$$\text{curl}\,\text{grad} = 0. \tag{22}$$

In short it can be summarized as

$$\Delta_u \text{grad}_p = \text{grad}_u \Delta_p. \tag{23}$$

Here we use subscript to indicate different operators associated to velocity and pressure. Unfortunately (23) does not hold for Stokes equations due to the boundary condition of velocity.

Exercise 4.1. Verify for MAC scheme, $E(i,j) = 0$ for interior indices $1 < i, j < n$. Thus the rank of E is at most $n^2 - (n-1)^2$.

The matrix $\mathcal{M}(\widetilde{\mathcal{L}\mathcal{M}})^{-1}$ will be an good approximation of $\mathcal{L}^{-1} = \mathcal{M}(\mathcal{L}\mathcal{M})^{-1}$. It defines an iterative method for the original system

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \mathcal{M}(\widetilde{\mathcal{L}\mathcal{M}})^{-1}(\boldsymbol{b} - \mathcal{L}\boldsymbol{x}^k), \tag{24}$$

where to simply the notation $\boldsymbol{x}^k := (u^k, p^k), \boldsymbol{b} := (f, g)$.

Let us write out component-wise formulation of this iterative method using the transformation $\boldsymbol{x} = \mathcal{M}\boldsymbol{y}$ with the transformed variable $\boldsymbol{y} = (w, q)$. Let $r_u = f - Au^k - B^\intercal p^k$ and $r_p = g - Bu^k$ be the residual at the k-th step. Denote by $A_p = BB^\intercal$.

We solve the residual equation $\mathcal{L}\mathcal{M}\delta\boldsymbol{y} = \boldsymbol{r}$ by an approximate one $\widetilde{\mathcal{L}\mathcal{M}}\,\delta\boldsymbol{y} = \boldsymbol{r}$, i.e.,

$$\begin{pmatrix} A & 0 \\ B & A_p \end{pmatrix} \begin{pmatrix} \delta w \\ \delta q \end{pmatrix} = \begin{pmatrix} r_u \\ r_p \end{pmatrix}$$

and update $\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \mathcal{M}\delta\boldsymbol{y}$, i.e.,

$$\begin{aligned} u^{k+1} &= u^k + \delta w + B^\intercal \delta q; \\ p^{k+1} &= p^k - A_p \delta q. \end{aligned}$$

Introducing an intermediate update $u^{k+1/2} = u^k + \delta w$ and noting $r_p - B\delta w = g - Bu^k - B\delta w = g - Bu^{k+1/2}$, we can write the iteration as follows.

Algorithm (PM) $[u^{k+1}, p^{k+1}] \leftarrow \texttt{ProjectionMethod}(u^k, p^k, f, g)$

(1) Update velocity by solving the momentum equation

$$u^{k+\frac{1}{2}} = A^{-1}(f - B^\intercal p^k),$$

(2) Solve the pressure equation

$$\delta q = A_p^{-1}(g - Bu^{k+\frac{1}{2}}).$$

(3) Update velocity

$$u^{k+1} = u^{k+\frac{1}{2}} + B^\intercal \delta q.$$

(4) Update pressure

$$p^{k+1} = p^k + Bu^{k+1/2} - g.$$

Step (3) is a projection. One can easily verify that $Bu^{k+1} = g$. Namely $u^{k+1/2}$ usually does not satisfy the constraint and the correction $B^\intercal \delta q$ is used to project it to the affine space. When $g = 0$, that is to project the velocity $u^{k+1/2}$ to the divergence free space.

If we skip the projection step, i.e. Steps (2)-(3), the method can be interpret as the Uzawa method for $\omega = 1$. From the convergence analysis of the Uzawa method, we immediately get the

$$\|p - p^{k+1}\| \leq \|I - S\|\|p - p^k\| \leq (1 - \beta^2)\|p - p^k\|,$$

where β is the inf-sup constant for B operator. The projection method can be thus analyzed using the convergence theory of the Uzawa method.

4.2. **Projection method for time-dependent Stokes equations.** The projection method is originally designed for the time-dependent Stokes equations [6] which is still the dominated solvers in computational fluid dynamics (CFD). One can easily adapt the derivation to time-dependent Stokes equation by replacing A by $A + M/\,\mathrm{d}t$ and obtain the so-called the rotational pressure-correction projection methods proposed in [18]. For a nice survey of projection methods, we refers to [11].

Consider the time-dependent Stokes equations

$$\begin{cases} u_t - \Delta u + \operatorname{grad} p = f & \text{in } \Omega, \\ -\operatorname{div} u = 0 & \text{in } \Omega, \\ u = u_D & \text{on } \partial\Omega. \end{cases} \tag{25}$$

and the semi-discretization in time $u_t \approx D_t(u^{k+1}, u^k, u^{k-1})$ and understand $D_t u^{k+1} = D_t(u^{k+1}, 0, 0)$ and $D_t(u^k, u^{k-1}) = D_t(0, u^k, u^{k-1})$. Let $\tilde{f} = f - D_t(u^k, u^{k-1})$. Then it can be written as a generalized Stokes system

$$\begin{cases} (D_t + A)u^{k+1} + B^\intercal p^{k+1} = \tilde{f}, \\ Bu^{k+1} = 0. \end{cases} \tag{26}$$

If we solve this coupled Stokes system, we get an implicit method. The issue is the generalized Stokes system is not easy to solve, requiring special iterative methods. Projection methods are ways to decouple the system.

We update our operators as

$$\mathcal{L} = \begin{pmatrix} D_t + A & B^\intercal \\ B & 0 \end{pmatrix}, \quad \mathcal{M} = \begin{pmatrix} I & B^\intercal \\ 0 & -D_t - BB^\intercal \end{pmatrix}.$$

Multiplication will yield

$$\mathcal{LM} = \begin{pmatrix} D_t + A & W \\ B & BB^\intercal \end{pmatrix} \approx \begin{pmatrix} D_t + A & 0 \\ B & BB^\intercal \end{pmatrix}.$$

Here again

$$W = D_t B^\intercal - B^\intercal D_t + AB^\intercal - B^\intercal BB^\intercal = AB^\intercal - B^\intercal A_p. \tag{27}$$

If we use the identity $-\Delta = -\text{grad div} + \text{curl curl}$, and $\text{curl grad} = 0$, we have

$$AB^\intercal - B^\intercal BB^\intercal = (-\Delta + \text{grad div})\text{grad} = \text{curl curl grad} = 0.$$

provided $B^\intercal = \text{grad}$ in the strong sense which is very subtle.

Algorithm (PMT) $(u^{k+1}, u^{k+1/2}, p^{k+1}) \leftarrow$ `ProjectionMethod`$(u^k, p^k, \tilde{f}, g)$

(1) Solve momentum equations
$$u^{k+1/2} = (D_t + A)^{-1}(\tilde{f} - B^\intercal p^k),$$
(2) Solve the pressure equation
$$\delta q = (BB^\intercal)^{-1}(g - Bu^{k+1/2}).$$
(3) Update velocity
$$u^{k+1} = u^{k+\frac{1}{2}} + B^\intercal \delta q.$$
(4) Update pressure
$$p^{k+1} = p^k - D_t \delta q + Bu^{k+1/2} - g.$$

We can treat the projection method as an iterative method for solving (26), that is during the iteration ($k \to k+1$), the source $\tilde{f}$ remains the same. If we keep updating the source by $\tilde{f} = f - D_t(u^k, u^{k-1})$, then $k \to k+1$ represents the forward in time. As we mentioned before, it is impossible to decouple the Stokes system (u and p are coupled together through a boundary integral equation). For time dependent problem, however, the projection method produces a first order method in time $O(\delta t)$. See [10] for detailed analysis.

4.3. **Distributive Gauss-Seidel Relaxation.** The standard relaxations, e.g., the Gauss-Seidel relaxation, are not applicable to the system (1), since the matrix is not diagonally dominant, and especially the zero block in the diagonal hampers the relaxation. The idea of the distributive relaxation is to transform the principle operators to the main diagonal and apply the equation-wise decoupled relaxation.

The so-called DGS smoother introduced by Brandt and Dinar [1] comes from consecutively Gauss-Seidel relaxation applied to the operator $\widetilde{\mathcal{LM}}$. That is we are solving the transformed residual equation

$$\begin{pmatrix} A & 0 \\ B & A_p \end{pmatrix} \begin{pmatrix} \delta w \\ \delta q \end{pmatrix} = \begin{pmatrix} r_u \\ r_p \end{pmatrix}$$

approximately and update it through the distributive matrix $\mathcal{M}$. Let $\hat{A}, \hat{A}_p$ be approximation of A and $A_p = BB^\intercal$, respectively.

Algorithm (DGS) $[u^{k+1}, p^{k+1}] \leftarrow \texttt{DGS}(u^k, p^k, f, g)$

(1) Relax the momentum equation

$$u^{k+\frac{1}{2}} = u^k + \hat{A}^{-1}(f - Au^k - B^\intercal p^k),$$

(2) Relax the transformed mass equation

$$\delta q = \hat{A}_p^{-1}(g - Bu^{k+\frac{1}{2}}).$$

(3) Distribute the correction to the original variables

$$u^{k+1} = u^{k+\frac{1}{2}} + B^\intercal \delta q,$$
$$p^{k+1} = p^k - B(B^\intercal \delta q).$$

The update formulae of pressure can be interpret as a Uzawa method using $A_p\hat{A}_p^{-1}$ as an approximation for the inverse of the Schur complement. From this point of view, the step (3) can be simplified as $p^{k+1} = p^k + Bu^{k+1/2} - g$.

The name "distributive relaxation" is due to the fact that the correction δq is distributed over u and p through the distributive matrix $\mathcal{M}$.

4.4. **LSC-DGS smoother.** The success of the DGS smoother depends on the existence of a Laplacian operator A_p for the discrete pressure space such that the commutator $AB^\intercal = B^\intercal A_p$ is small. As we mentioned earlier, this is almost true in the continuous level. In the discrete level, the differential operator is approximated by matrices and some properties may not hold in the matrix level.

For example, even DGS has been constructed since late 1970s', this type of smoother, however, is only known for the Marker and Cell (MAC) scheme discretization and mini-finite element on rectangular grids [21, 7]. Generalization to other stable pairs of finite element methods seems difficult; see [12] (p. 248).

One remedy is to discretize the Poisson equation for pressure with Neumann boundary condition for the discrete pressure space which might be discontinuous. We introduce an algebraic construction to minimize the commutator following our recent work [20]. We shall use the following distributive matrix

$$\mathcal{M} = \begin{pmatrix} I & B^\intercal \\ 0 & -(BB^\intercal)^{-1}BAB^\intercal \end{pmatrix}, \tag{28}$$

which gives rise to the transformed system

$$\mathcal{L}\mathcal{M} = \begin{pmatrix} A & PAB^\intercal \\ B & BB^\intercal \end{pmatrix}, \quad \text{with } P = I - B^\intercal(BB^\intercal)^{-1}B.$$

Note that $P : \mathbb{V}_h \mapsto \ker(B)$ is the orthogonal projection operator in L^2 inner product to $\ker(B)$ and hence $PB^\intercal = 0$. Consequently, the commutator reads as

$$E := PAB^\intercal = P(AB^\intercal - B^\intercal A_p).$$

Now the only requirement is the existence of A_p such that (23) holds, and no explicit construction is needed.

For a general discrete Stokes system, it is not clear whether such A_p exists or not. However, the projection matrix makes the commutator as small as possible in the least squares sense, i.e.

$$\|PAB^\intercal\|_F \leq \min_{X:\mathbb{Q}\to\mathbb{Q}} \|AB^\intercal - B^\intercal X\|_F, \tag{29}$$

where $\|\cdot\|_F$ denotes the Frobenius norm (F-norm) of matrices. Indeed solving the least-squares problem on the right of (29) will give the solution $A_p^* = (BB^\intercal)^{-1}BAB^\intercal$ and therefore it will be called Least-Squares Commutator DGS (LSC-DGS) smoother.

The LSC was firstly developed by Elman [8], and used to construct a least-squares approximation to the Schur complement of the linearized Navier-Stokes system, yielding the so-called BFBt preconditioner. Here we use A_p^* to devise an effective DGS smoother.

We present the LSC-DGS relaxation algorithm below.

Algorithm (LSC-DGS) $[u^{k+1}, p^{k+1}] \leftarrow \texttt{LSC-DGS}(u^k, p^k, f, g)$

(1) Relax momentum equations

$$u^{k+\frac{1}{2}} = u^k + \hat{A}^{-1}(f - Au^k - B^\intercal p^k),$$

(2) Relax transformed continuity equations

$$\delta q = \hat{A}_p^{-1}(g - Bu^{k+\frac{1}{2}}).$$

(3) Transform the correction back to the original variables

$$u^{k+1} = u^{k+\frac{1}{2}} + B^\intercal \delta q, \tag{3.1}$$

$$p^{k+1} = p^k - \widetilde{A}_p^{-1} BAB^\intercal \delta q. \tag{3.2}$$

We now discuss the choices of the three approximations $\hat{A}, \hat{A}_p$ and $\widetilde{A}_p$ used in LSC-DGS. The smoother $\hat{A}^{-1}$ for the velocity can be the standard Gauss-Seidel relaxation which has been shown to be effective for the Laplacian operator. When applicable, red-black or general multi-coloring ordering is further applied to improve the smoothing effect.

In LSC-DGS, one needs to transfer the correction back to the original pressure variables by applying $(BB^\intercal)^{-1}BAB^\intercal$. In step (3.2), the matrix inversion $(BB^\intercal)^{-1}$ is replaced by a cheaper relaxation $\widetilde{A}_p^{-1}$ which is in general different with the smoothers $\hat{A}_p$ used in step (2). Recall that the commutator, i.e., (1,2) block of the transformed system $\mathcal{LM}$, will be $W = (I - B^\intercal \widetilde{A}_p^{-1} B)AB^\intercal$. The closer $\widetilde{A}_p^{-1}$ is to $(BB^\intercal)^{-1}$, the smaller $\|E\|_F$ is. As a guideline, from our empirical tests, $\widetilde{A}_p$ can be taken to be one symmetric Gauss-Seidel relaxation for structured grids and one $\mathcal{V}$-cycle iteration for unstructured grids.

The smoother $\hat{A}_p$ will affect the projection step (3.1) of $u^{k+\frac{1}{2}}$. Choosing $\hat{A}_p$ closer to $(BB^\intercal)^{-1}$ will make $u^{k+\frac{1}{2}}$ more divergence free in each level and consequently may help in accelerating the convergence of the whole multigrid procedure. Usually, the smoother $\hat{A}_p$ in step (2) can be just one Gauss-Seidel iteration. For discontinuous pressure finite element approximations, it can be chosen as an element-wise block Gauss-Seidel smoother.

Compared with the standard DGS, step (3.2) of LSC-DGS requires one more relaxation and one more matrix-vector multiplication. On the other hand, LSC-DGS is more robust, efficient, and parameter free. This is a typical trade off between robustness and operation count.

5. Block Preconditioners

5.1. Functional analysis approach.

Add Winther's approach based on Riesz map [15].

5.2. Linear algebra approach.

For Stokes system $\mathcal{L}$, the preconditioner $\mathcal{P}$ is not necessary a good approximation of $\mathcal{L}^{-1}$. A sufficient condition for a good preconditioner is that the preconditioned matrix $\mathcal{T} = \mathcal{P}^{-1}\mathcal{L}$ has a low-degree minimum polynomial or $\mathcal{T}$ having

only a few distinct eigenvalues. We follow [16] to present a block preconditioner using Schur complement for the Stokes system.

Theorem 5.1. *If*

$$\mathcal{L} = \begin{pmatrix} A & B^\intercal \\ B & 0 \end{pmatrix},$$

is preconditioned by

$$\mathcal{D} = \begin{pmatrix} A & 0 \\ 0 & BA^{-1}B^\intercal \end{pmatrix},$$

then the preconditioned matrix

$$\mathcal{T} = \mathcal{D}^{-1}\mathcal{L} = \begin{pmatrix} I & A^{-1}B^\intercal \\ (BA^{-1}B^\intercal)^{-1}B & 0 \end{pmatrix}$$

satisfies

$$(\mathcal{T} - I)(\mathcal{T}^2 - \mathcal{T} - I) = 0. \tag{30}$$

Proof. It can be proved by a straight forward computation. A key observation is the product $A^{-1}B^\intercal(BA^{-1}B^\intercal)^{-1}B$ is a projection operator (in the A-inner product). We also provide a calculation of the eigenvalue. Recall that, the Schur complement $S = BA^{-1}B^\intercal$.

As $\mathcal{D}$ is symmetric and positive definite, $\sigma(\mathcal{D}^{-1}\mathcal{L}) = \sigma(\mathcal{D}^{-1/2}\mathcal{L}\mathcal{D}^{-1/2})$ which contains only real eigen-values and eigen-vectors. Consider the generalized eigenvalue problem for the saddle point system $\mathcal{L} := (A, B^\intercal; B, O)$ in the inner product defined by (A, M) of $\mathbb{X} \times \mathbb{Y}$:

$$A\,x + B^\intercal y = \lambda A\,x, \tag{31}$$

$$B\,x \qquad\quad = \lambda S\,y. \tag{32}$$

The space $\mathbb{X}$ can be decomposed into $K = \ker(B)$ and its A-orthogonal complement $K^{\perp_A}$, i.e., $\mathbb{X} = K \oplus K^{\perp_A}$. There exists a unique decomposition $x = x_0 \oplus x_1, x_0 \in K, x_1 \in K^{\perp_A}$, i.e. $(x_0, x_1)_A = (Ax_0, x_1) = (x_0, Ax_1) = 0$.

We test (31) with $x_0 \in K$ to get

$$(A(x_0 + x_1), x_0) + (B^\intercal y, x_0) = \lambda(A(x_0 + x_1), x_0).$$

Then use $(B^\intercal y, x_0) = (y, Bx_0) = 0$ and $(Ax_1, x_0) = 0$ to get

$$(\lambda - 1)\|x_0\|_A^2 = 0 \quad \longrightarrow \quad \lambda = 1 \text{ or } x_0 = 0.$$

Case 1. $\lambda = 1$. Then Equation (31) becomes $B^\intercal y = 0$. As B is surjective, $B^\intercal$ is injective and thus $y = 0$. Substitute into (32) yields $x = x_0$. Therefore the corresponding eigenvector is in the form $(x_0, 0) \in K$. In the finite dimensional case, the multiplicity of eigenvalue 1 is $\dim K = \dim \mathbb{X} - \dim \mathbb{Y}$.

Case 2. $\lambda \neq 1$ and $x_0 = 0$. We can then solve $x = [(\lambda - 1)A]^{-1} B^\intercal y$ from (31) and substitute into (32) to get the equation

$$(\lambda - 1)^{-1} Sy = \lambda Sy. \tag{33}$$

Then we get the equation for λ

$$\lambda^2 - \lambda - 1 = 0$$

which implies (30) and the corresponding eigenvalues are

$$\lambda_\pm = \frac{1 \pm \sqrt{5}}{2}. \tag{34}$$

□

As a consequence, the GMRES applied to the preconditioned system will terminate in at most 3 steps.

A block triangular preconditioner can be construct from the block factorization

$$\begin{pmatrix} A & B^\intercal \\ B & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -S \end{pmatrix} \begin{pmatrix} I & A^{-1}B^\intercal \\ 0 & I \end{pmatrix} = \mathcal{U}^\intercal \mathcal{H} \mathcal{U}. \tag{35}$$

One can use

$$\mathcal{U}^\intercal \mathcal{H} = \begin{pmatrix} A & 0 \\ B & -S \end{pmatrix} \quad \text{or} \quad \mathcal{H}\mathcal{U} = \begin{pmatrix} A & B^\intercal \\ 0 & -S \end{pmatrix}$$

as a preconditioner. The minimal polynomial of the preconditioned system is quadratic and thus the GMRES for the preconditioned system will converge at most 2 steps!

For practical use, the Schur complement for Stokes $S = BA^{-1}B^\intercal$ should be approximated by M_p the mass matrix of the pressure; see the discussion in Uzawa method section. The eigenvalues of preconditioned system will be in three distinct regions and GMRES will converge uniformly, i.e., independent of the size of the system.

The A^{-1} can be approximated by $\hat{A}^{-1}$, e.g. several $\mathcal{V}$-cycle iterations. The M_p^{-1} can be solved efficiently by PCG with diagonal preconditioner to the tolerance 10^{-1} (usually one or two PCG iterations is enough.)

Exercise 5.2. Present a spectral analysis when the block-diagonal preconditioner is changed to $\mathcal{P} = \mathrm{diag}(A, M_p)$ and $\mathcal{P} = \mathrm{diag}(\hat{A}, M_p)$.

A symmetric and indefinite preconditioner can be also construct using an approximated block factorization. Let $\hat{A}, \hat{S}$ be an approximation of A, S, respectively. Based on (35), we introduce

$$\hat{\mathcal{H}} = \begin{pmatrix} \hat{A} & 0 \\ 0 & -\hat{S} \end{pmatrix}, \quad \hat{\mathcal{U}} = \begin{pmatrix} I & \hat{A}^{-1}B^\intercal \\ 0 & I \end{pmatrix}, \quad \hat{\mathcal{L}} = \hat{\mathcal{U}}^\intercal \hat{\mathcal{H}} \hat{\mathcal{U}}.$$

and use

$$\hat{\mathcal{L}}^{-1} = \hat{\mathcal{U}}^{-1} \left(\hat{\mathcal{U}}^\intercal \hat{\mathcal{H}}\right)^{-1} = \begin{pmatrix} I & -\hat{A}^{-1}B^\intercal \\ 0 & I \end{pmatrix} \begin{pmatrix} \hat{A} & 0 \\ B & -\hat{S} \end{pmatrix}^{-1}$$

The action $\left(\hat{\mathcal{U}}^\intercal \hat{\mathcal{H}}\right)^{-1}$ can be understood as a block Gauss-Seidel iteration and $\hat{\mathcal{U}}^{-1}$ is a distribution or change of variables. Convergence analysis of such preconditioner can be found in [4].

REFERENCES

[1] A. Brandt and N. Dinar. Multi-grid solutions to elliptic flow problems. Technical Report 79, Institute for Computer Applications in Science and Engineering. NASA Langley Research Center, Hampton, Virginia, New York, 1979. 11

[2] L. Chen. Multigrid methods for constrained minimization problems and application to saddle point problems. *Submitted for publication*, 2015. 2

[3] L. Chen. Multigrid methods for saddle point systems using constrained smoothers. *Computers & Mathematics with Applications*, 70(2):2854–2866, 2015. 2

[4] L. Chen and Y. Wu. Convergence analysis for a class of iterative methods for solving saddle point systems . *ArXiv*, 2018. 15

[5] X. Cheng. On the Nonlinear Inexact Uzawa Algorithm for Saddle-Point Problems. *SIAM J. Numer. Anal.*, 37:1930, 2000. 6

[6] A. J. Chorin. Numerical Solution of the Navier-Stokes Equations. *Mathematics of Computation*, 22(104):745, oct 1968. 10

[7] H. C. Elman. Multigrid and Krylov Subspace Methods for the Discrete Stokes Equations. *International Journal for Numerical Methods in Fluids*, 22(June 1995):755–770, 1996. 12

[8] H. C. Elman. Preconditioning for the Steady-State Navier–Stokes Equations with Low Viscosity. *SIAM Journal on Scientific Computing*, 20(4):1299–1316, jan 1999. 13

[9] M. Fortin and R. Glowinski. *Augmented Lagrangian Methods*. Applications to the numerical solution of boundary value problems, North-Holland Publishing Co., Amsterdam, 1983. 6

[10] J. Guermond and J. Shen. On the error estimates for the rotational pressure-correction projection methods. *Mathematics of Computation*, 73(248):1719–1737, 2004. 11

[11] J. L. Guermond, P. Minev, and J. Shen. An overview of projection methods for incompressible flows. *Comput. Methods Appl. Mech. Engrg.*, 195(44-47):6011–6045, 2006. 10

[12] W. Hackbusch. *Multigrid Methods and Applications*, volume 4 of *Computational Mathematics*. Springer–Verlag, Berlin, 1985. 12

[13] W. W. Hager. Updating the Inverse of a Matrix. *SIAM Rev.*, 31(2):221–239, 1989. 6

[14] H. V. Henderson and S. R. Searle. On Deriving the Inverse of a Sum of Matrices. *SIAM Rev.*, 23(1):53–60, 1981. 6

[15] K.-A. Mardal and R. Winther. Preconditioning discretizations of systems of partial differential equations. *Numerical Linear Algebra with Applications*, 18(1):1–40, jan 2011. 13

[16] M. F. Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Sci. Comput.*, 21(6):1969–1972, 2000. 14

[17] X.-C. Tai. Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities. *Numer. Math.*, 93(4):755–786, 2003. 2

[18] L. J. P. Timmermans, P. D. Minev, and F. N. Van De Vosse. An Approximate Projection Scheme for Incompressible Flow Using Spectral Elements. *International Journal for Numerical Methods in Fluids*, 22(7):673–688, apr 1996. 10

[19] H. Uzawa. Iterative methods for concave programming. *Studies in linear and nonlinear programming*, 6, 1958. 3

[20] M. Wang and L. Chen. Multigrid Methods for the Stokes Equations using Distributive Gauss-Seidel Relaxations based on the Least Squares Commutator. *Journal of Scientific Computing*, 56(2):409–431, feb 2013. 12

[21] G. Wittum. On the Convergence of Multi-Grid Methods with Transforming Smoothers. *Numerische Mathematik*, 57:15–38, 1990. 12