Stabilized Schemes for Mixed Finite Element Methods with Applications to Elasticity and Compressible Flow Problems

Daoqi Yang*

March 20, 1997

Abstract

Stabilized iterative schemes for mixed finite element methods are proposed and analyzed in two abstract formulations. The first one has applications to elliptic equations and incompressible fluid flow problems, while the second has applications to linear elasticity and compressible Stokes problems. Convergence theorems are demonstrated in abstract formulations; applications to individual physical problems are included. In contrast to standard mixed finite element methods, these stabilized schemes lead to positive definite linear systems of algebraic equations that have smaller condition numbers than penalty methods. Theoretical analysis and computational experiments both show that the stabilized schemes have very fast convergence; a few iterations are usually enough to reduce the iterative error to a prescribed precision. Numerical examples with continuous and discontinuous coefficients are presented.

Keywords: Stabilized scheme, finite element method, elliptic equations, linear elasticity problems, compressible Stokes problems.

AMS Subject Classifications: 65N12, 65N30

1 Introduction

Mixed finite element methods have been proved effective for a number of engineering problems; they provide good approximations to the solution gradient and the solution itself and have the capability of dealing with rough coefficients. However, the major shortcoming of the methods is that they result in large and ill-conditioned non-positive definite linear systems involving the solution and its gradient unknowns. On the computational side, this requires a large computer memory and fast CPU speed to handle the run-time stack and allocate space on the heap for storing and manipulating the matrix coefficients. On the algorithmic side, this causes the design of numerical methods for the resulting linear systems extremely difficult.

*Institute for Mathematics and its Applications, University of Minnesota, 514 Vincent Hall, 206 SE Church Street, Minneapolis, MN 55455-0436. Email: yang@ima.umn.edu, dyang@na-net.ornl.gov

Penalty methods [8] have been proposed to stabilize the problem by adding a penalty term, but the resulting linear system is still ill-conditioned. In particular, in order for penalty methods to have good accuracy, the penalty parameter must be chosen to be very small. This would cause a larger condition number of the system. Augmented Lagrangian methods [2, 7] have been proposed to remedy the problem by choosing a relatively larger penalty parameter and iterate the system to achieve accuracy. However, augmented Lagrangian methods have been implemented and analyzed for only a limited number of problems, for example incompressible Stokes and Navier-Stokes equations, and a general analysis of the method has not been made.

As an example, let us consider the incompressible Stokes equations: find the velocity and pressure pair $\{u,p\} \in V \times Q \equiv (H_0^1(\Omega))^d \times \{q : q \in L^2(\Omega), \int_\Omega q dx = 0\}$ such that

$$-\Delta u + \nabla p = f, \quad x \in \Omega, \tag{1}$$

$$\text{div } u = 0, \quad x \in \Omega, \tag{2}$$

$$u = 0, \quad x \in \partial\Omega, \tag{3}$$

where $\Omega \subset \mathcal{R}^d (d = 1, 2, \text{ or } 3)$ is a Lipschitz domain with boundary $\partial\Omega$, and f is a given vector function. The mixed formulation of the problem (1)-(3) reads: find $\{u,p\} \in V \times Q$ satisfying

$$(\nabla u, \nabla v) - (\text{div } v, p) = (f, v), \quad \forall v \in V, \tag{4}$$

$$(\text{div } u, q) = 0, \quad \forall q \in Q, \tag{5}$$

where $(\phi, \psi) = \int_\Omega \phi \cdot \psi dx$. Its mixed finite element approximations consist of constructing finite dimensional spaces [3] and solve (4)-(5) with $V \times Q$ replaced by these spaces. There are a few stabilized methods for this problem [3, 2, 4, 8, 7, 11], below we just mention two which are closely related to our scheme in this paper.

Penalty methods [3, 8] simply add one term to stabilize the problem: find $\{u_\epsilon, p_\epsilon\} \in V \times Q$ satisfying

$$(\nabla u_\epsilon, \nabla v) - (\text{div } v, p_\epsilon) = (f, v), \quad \forall v \in V, \tag{6}$$

$$(\text{div } u_\epsilon, q) + \epsilon(p_\epsilon, q) = 0, \quad \forall q \in Q, \tag{7}$$

where ϵ is a small positive real number. The convergence for this method is $\|u - u_\epsilon\|_V + \|p - p_\epsilon\|_Q = O(\epsilon)$. Augmented Lagrangian methods [2, 7, 11] iterate the penalty scheme: with initial guess $p_0 \in Q$, find $\{u_m, p_m\} \in V \times Q$ satisfying, for $m = 0, 1, 2, \cdots$,

$$(\nabla u_m, \nabla v) + \frac{1}{\epsilon}(\text{div } u_m, \Pi(\text{div } v)) = (f, v) + (\text{div } v, p_{m-1}), \quad \forall v \in V, \tag{8}$$

$$p_m = p_{m-1} - \frac{1}{\epsilon}\Pi(\text{div } u_m), \tag{9}$$

where Π is the projection operator onto Q. The convergence for this method is $\|u - u_m\|_V + \|p - p_m\|_Q = O(\epsilon^m)$, for $m = 0, 1, 2, \cdots$. Thus for augmented Lagrangian methods, the parameter ϵ does not have to be chosen very small so that the conditioning of the system can be improved and the accuracy is achieved through iterating the process.

At matrix level, the finite dimensional approximation of (4)-(5) has the form

$$AU + B^t P = F, \tag{10}$$

$$BU = 0, \tag{11}$$

where A and B are two coefficient matrices. Penalty methods and augmented Lagrangian methods have the form

$$AU_\epsilon + B^t P_\epsilon = F, \tag{12}$$

$$BU_\epsilon - \epsilon D P_\epsilon = 0, \tag{13}$$

and

$$AU_m + \frac{1}{\epsilon} B^t D^{-1} BU_m = F - \epsilon B^t P_{m-1}, \tag{14}$$

$$P_m = P_{m-1} + \frac{1}{\epsilon} D^{-1} BU_m, \tag{15}$$

respectively. Both penalty and augmented Lagrangian methods can decouple the velocity variable U from the pressure variable P and result in symmetric and positive definite linear systems.

Augmented Lagrangian methods are usually posed in the context of optimization, and are not widely available for general mixed finite element schemes. For example, the analysis in [2] does not seem to apply for compressible fluid problems and for second order elliptic problems with a general zero-order term. In this paper, we generalize the idea of the augmented Lagrangian methods and apply it to general framework of mixed finite element methods. Note that the computation of the projection Π in (8)-(9) for each basis function may consume a lot of CPU time. Our generalized schemes will not involve such projections and reduce to an equivalent form of (8)-(9) for incompressible Stokes equations.

2 Stabilized Schemes in abstract formulations

We first consider an abstract framework which is applicable to elliptic equations and incompressible fluid flow problems. Let V and Q be two Hilbert spaces with inner products denoted by $(\cdot,\cdot)_V$ and $(\cdot,\cdot)_Q$, norms denoted by $\|\cdot\|_V$ and $\|\cdot\|_Q$, and dual spaces denoted by V' and Q', respectively. Let $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ and $c(\cdot,\cdot)$ be bounded bilinear forms on $V\times V$, $V\times Q$ and $Q\times Q$ respectively, and f and g be two bounded linear functionals in V' and Q' respectively, with $\langle\cdot,\cdot\rangle$ denoting the duality. Find $\{u,p\}\in V\times Q$ such that

$$a(u,v) + b(v,p) = \langle f, v\rangle_{V'\times V}, \quad \forall v\in V, \tag{16}$$

$$b(u,q) - c(p,q) = \langle g, q\rangle_{Q'\times Q}, \quad \forall q\in Q, \tag{17}$$

Assume that $a(\cdot,\cdot)$ is positive semidefinite, $c(\cdot,\cdot)$ is symmetric and positive semidefinite, and $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ satisfy the conditions

$$\inf_{\substack{w\in \mathrm{Ker}(B)\\ w\neq 0}} \sup_{\substack{v\in \mathrm{Ker}(B)\\ v\neq 0}} \frac{a(w,v)}{\|w\|_V\|v\|_V} \geq \alpha_1, \tag{18}$$

$$\inf_{\substack{v\in \mathrm{Ker}(B)\\ v\neq 0}} \sup_{\substack{w\in \mathrm{Ker}(B)\\ w\neq 0}} \frac{a(w,v)}{\|w\|_V\|v\|_V} \geq \alpha_1, \tag{19}$$

$$\sup_{\substack{v \in V \\ v \neq 0}} \frac{b(v,q)}{\|v\|_V} \geq \alpha_2 \|q\|_{Q/\mathrm{Ker}(B^t)}, \quad \forall q \in Q, \tag{20}$$

where α_1 and α_2 are positive constants and $\mathrm{Ker}(B)$ denotes the kernel of the operator $B : V \to Q'$, which, together with its transpose $B^t : Q \to V'$, is defined by

$$\langle Bv, q\rangle_{Q' \times Q} = \langle B^t q, v\rangle_{V' \times V} = b(v,q), \quad \forall v \in V,\ \forall q \in Q. \tag{21}$$

Note (20) is equivalent to the fact that the image $\mathrm{Im}(B^t)$ is closed in V', that is, the image $\mathrm{Im}(B)$ of operator B is closed in Q', which is further equivalent to

$$\sup_{\substack{q \in Q \\ q \neq 0}} \frac{b(v,q)}{\|q\|_Q} \geq \alpha_2 \|v\|_{V/\mathrm{Ker}(B)}, \quad \forall v \in V. \tag{22}$$

We further assume that the bilinear form $c(\cdot,\cdot)$ is either coercive on $\mathrm{Ker}(B^t)$, that is, there is a constant $\gamma > 0$ such that

$$\gamma \|q\|_Q^2 \leq c(q,q), \quad \forall q \in \mathrm{Ker}(B^t), \tag{23}$$

or identically zero; in the latter case we also assume that $g \in \mathrm{Im}(B)$.

Let $\mathrm{Ker}(C)$ denote the kernel of C, the operator associated to the bilinear form $c(\cdot,\cdot)$. Under the assumptions above, the problem (16)-(17) has a unique solution $\{u,p\} \in V \times Q/(\mathrm{Ker}(B^t) \cap \mathrm{Ker}(C))$. See [3] for a proof. The linear system of algebraic equations from (16)-(17) is usually ill-posed and thus can not be solved easily. We propose the following iterative scheme. Taking ϵ to be a small positive number and $p_0 \in Q$ an initial guess, find $\{u_m, p_m\} \in V \times Q$ such that, for $m = 0, 1, 2, \cdots$,

$$a(u_m, v) + b(v, p_m) = \langle f, v\rangle, \quad \forall v \in V, \tag{24}$$

$$b(u_m, q) - c(p_m, q) - (\epsilon p_m, q)_Q = \langle g, q\rangle - (\epsilon p_{m-1}, q)_Q, \quad \forall q \in Q. \tag{25}$$

It is easy to see that (24)-(25) has a unique solution in $V \times Q$. We now state and prove the following convergence results.

Theorem 2.1 *Let $\{u,p\}$ be the solution to the system (16)-(17) and $\{u_m, p_m\}$ the solution of (24)-(25). Assume that $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ $c(\cdot,\cdot)$ are bounded with norms $\|a\|, \|b\|, \|c\|$, respectively, $a(\cdot,\cdot)$ is positive semidefinite, and $c(\cdot,\cdot)$ is symmetric and positive semidefinite. Assume further that (18), (19), (20), and (23) are satisfied. Then we have, for $\epsilon < 1/K$ and $p_0 \in Q$,*

$$\|p - p_m\|_Q \leq \left(\frac{K\epsilon}{1 - K\epsilon}\right)^m \|p - p_0\|_Q, \quad m = 1, 2, \cdots, \tag{26}$$

$$\|u - u_m\|_V \leq \frac{\bar{K}}{K} \left(\frac{K\epsilon}{1 - K\epsilon}\right)^m \|p - p_0\|_Q, \quad m = 1, 2, \cdots, \tag{27}$$

where

$$K = K_1 + 0.5K_2\left(K_2 + \sqrt{4K_1 + K_2^2}\right), \tag{28}$$

$$\bar{K} = \frac{\|a\| + \alpha_1}{\alpha_1\alpha_2}\left[1 + 0.5\|c\|^{1/2}\left(K_2 + \sqrt{4K_1 + K_2^2}\right)\right], \tag{29}$$

$$K_1 = \frac{\|a\|}{\gamma\alpha_1\alpha_2^2}(\|c\| + \gamma)(\|a\| + \alpha_1) + \frac{1}{\gamma}, \tag{30}$$

$$K_2 = \frac{\|a\|\|c\|^{1/2}}{\gamma\alpha_1\alpha_2^2}(\|c\| + \gamma)(\|a\| + \alpha_1). \tag{31}$$

Proof: From (16)-(17) and (24)-(25), the iterative errors $e_m = u - u_m \in V$ and $r_m = p - p_m \in Q$ satisfy

$$a(e_m, v) + b(v, r_m) = 0, \quad \forall v \in V, \tag{32}$$

$$b(e_m, q) - c(r_m, q) = -\epsilon(p_m - p_{m-1}, q)_Q, \quad \forall q \in Q. \tag{33}$$

We write e_m and r_m into

$$e_m = \hat{e}_m + \bar{e}_m, \quad \hat{e}_m \in \mathrm{Ker}(B), \quad \bar{e}_m \in \mathrm{Ker}(B)^\perp, \tag{34}$$

$$r_m = \hat{r}_m + \bar{r}_m, \quad \hat{r}_m \in \mathrm{Ker}(B^t), \quad \bar{r}_m \in \mathrm{Ker}(B^t)^\perp, \tag{35}$$

where $\mathrm{Ker}(B)^\perp$ and $\mathrm{Ker}(B^t)^\perp$ denote the orthogonal complements of $\mathrm{Ker}(B)$ and $\mathrm{Ker}(B^t)$ in the spaces V and Q, respectively. Note that the symmetry and positive semidefiniteness of $c(\cdot,\cdot)$ imply that

$$c(p, q) \le \sqrt{c(p,p)}\sqrt{c(q,q)}, \quad \forall p, q \in Q. \tag{36}$$

By (22), (33), and (36) we see that

$$\begin{aligned} \alpha_2\|\bar{e}_m\|_V &\le \sup_{\substack{q \in Q \\ q \ne 0}} \frac{b(\bar{e}_m, q)}{\|q\|_Q} \\ &= \sup_{\substack{q \in Q \\ q \ne 0}} \frac{c(r_m, q) - \epsilon(p_m - p_{m-1}, q)_Q}{\|q\|_Q} \\ &\le \epsilon\|p_m - p_{m-1}\|_Q + \|c\|^{1/2}\sqrt{c(r_m, r_m)}. \end{aligned} \tag{37}$$

By (18) and (32) we see that

$$\begin{aligned} \alpha_1\|\hat{e}_m\|_V &\le \sup_{\substack{v \in \mathrm{Ker}(B) \\ v \ne 0}} \frac{a(\hat{e}_m, v)}{\|v\|_V} = \sup_{\substack{v \in \mathrm{Ker}(B) \\ v \ne 0}} \frac{a(e_m, v) - a(\bar{e}_m, v)}{\|v\|_V} \\ &= \sup_{\substack{v \in \mathrm{Ker}(B) \\ v \ne 0}} \frac{-b(v, r_m) - a(\bar{e}_m, v)}{\|v\|_V} \le \|a\|\|\bar{e}_m\|_V. \end{aligned} \tag{38}$$

Combining (37) and (38) we have

$$
\begin{aligned}
\|e_m\|_V &\le \|\hat{e}_m\|_V + \|\bar{e}_m\|_V \le \left(\frac{\|a\|}{\alpha_1} + 1\right) \|\bar{e}_m\|_V \\
&\le \frac{\|a\| + \alpha_1}{\alpha_1 \alpha_2} \left(\epsilon \|p_m - p_{m-1}\|_Q + \|c\|^{1/2} \sqrt{c(r_m, r_m)}\right).
\end{aligned} \tag{39}
$$

In view of (20) and (32) we obtain

$$
\alpha_2 \|\bar{r}_m\|_Q \le \sup_{\substack{v \in V \\ v \neq 0}} \frac{b(v, \bar{r}_m)}{\|v\|_V} = \sup_{\substack{v \in V \\ v \neq 0}} \frac{b(v, r_m)}{\|v\|_V} = \sup_{\substack{v \in V \\ v \neq 0}} \frac{-a(e_m, v)}{\|v\|_V} \le \|a\| \|e_m\|_V. \tag{40}
$$

Note that (33) implies that

$$
c(\hat{r}_m, q) = -c(\bar{r}_m, q) + \epsilon (p_m - p_{m-1}, q)_Q, \quad \forall q \in \mathrm{Ker}(B^t). \tag{41}
$$

By the coerciveness condition (23) we have

$$
\gamma \|\hat{r}_m\|_Q \le \|c\| \|\bar{r}_m\| + \epsilon \|p_m - p_{m-1}\|. \tag{42}
$$

Combining (40), (42), and (39) we have a bound on $\|r_m\|_Q$:

$$
\begin{aligned}
\|r_m\|_Q &\le \|\hat{r}_m\|_Q + \|\bar{r}_m\|_Q \\
&\le \left(\frac{\|c\|}{\gamma} + 1\right) \|\bar{r}_m\|_Q + \frac{\epsilon}{\gamma} \|p_m - p_{m-1}\|_Q \\
&\le \frac{\|c\| + \gamma}{\gamma \alpha_2} \|a\| \|e_m\|_V + \frac{\epsilon}{\gamma} \|p_m - p_{m-1}\|_Q \\
&\le \epsilon K_1 \|p_m - p_{m-1}\|_Q + K_2 \sqrt{c(r_m, r_m)},
\end{aligned} \tag{43}
$$

where K_1 and K_2 are defined in (30)-(31).

It remains to bound $c(r_m, r_m)$. We combine (32) with $r = e_m$ and (33) with $q = r_m$ to get

$$
a(e_m, e_m) + c(r_m, r_m) = \epsilon (p_m - p_{m-1}, r_m)_Q, \tag{44}
$$

which, together with (43), implies that

$$
\begin{aligned}
c(r_m, r_m) &\le \epsilon \|p_m - p_{m-1}\|_Q \|r_m\|_Q \\
&\le K_1 \epsilon^2 \|p_m - p_{m-1}\|_Q^2 + \epsilon K_2 \|p_m - p_{m-1}\|_Q \sqrt{c(r_m, r_m)}.
\end{aligned} \tag{45}
$$

Note that $x \le y + z\sqrt{x}$ with $x \ge 0,\ y \ge 0,\ z \ge 0$ implies that $\sqrt{x} \le z/2 + \sqrt{y + z^2/4}$. Thus (45) gives

$$
\sqrt{c(r_m, r_m)} \le \epsilon \left(\frac{K_2}{2} + \sqrt{K_1 + \frac{K_2^2}{4}}\right) \|p_m - p_{m-1}\|_Q. \tag{46}
$$

In view of (46), the inequality (43) becomes

$$\|r_m\|_Q \leq \epsilon K \|p_m - p_{m-1}\|_Q, \tag{47}$$

where K is defined by (28). The formula (26) follows immediately from (47) and the triangle inequality. By (39) and (46) we obtain

$$\|e_m\|_V \leq \epsilon \bar{K} \|p_m - p_{m-1}\|_Q \leq \epsilon \bar{K}(\|r_m\|_Q + \|r_{m-1}\|_Q), \tag{48}$$

where $\bar{K}$ is given by (29). Now (27) follows from (26) and (48). This concludes the proof. ∎

Theorem 2.2 *Let $\{u, p\}$ be the solution to the system (16)-(17) and $\{u_m, p_m\}$ the solution of (24)-(25) with the bilinear form $c(\cdot,\cdot) \equiv 0$ and the right hand side functional $g \in Im(B)$. Under the assumptions of Theorem 2.1 except (23), then for $\epsilon < 1/K$ and $p_0 \in Q$,*

$$\|p - p_m\|_{Q/Ker(B^t)} \leq \left(\frac{K\epsilon}{1 - K\epsilon}\right)^m \|p - p_0\|_{Q/Ker(B^t)}, \quad m = 1, 2, \cdots. \tag{49}$$

$$\|u - u_m\|_V \leq \frac{\|a\| + \alpha_1}{\alpha_1 \alpha_2 K} \left(\frac{K\epsilon}{1 - K\epsilon}\right)^m \|p - p_0\|_{Q/Ker(B^t)}, \quad m = 1, 2, \cdots, \tag{50}$$

where

$$K = \frac{\|a\|(\|a\| + \alpha_1)}{\alpha_1 \alpha_2^2}. \tag{51}$$

Proof: From (16)-(17) and (24)-(25), the iterative errors $e_m = u - u_m \in V$ and $r_m = p - p_m \in Q$ now satisfy

$$a(e_m, v) + b(v, r_m) = 0, \quad \forall v \in V, \tag{52}$$

$$b(e_m, q) = -\epsilon(p_m - p_{m-1}, q)_Q, \quad \forall q \in Q. \tag{53}$$

Repeat the same procedure as in the proof of Theorem 2.1. But in this case, we have, instead of (39) and (43),

$$\|e_m\|_V \leq \frac{\|a\| + \alpha_1}{\alpha_1 \alpha_2} \epsilon \|p_m - p_{m-1}\|_Q, \tag{54}$$

$$\|r_m\|_{Q/\mathrm{Ker}(B^t)} = \|\bar{r}_m\|_Q \leq \frac{\|a\|}{\alpha_2} \|e_m\|_V \leq K\epsilon \|p_m - p_{m-1}\|_Q, \tag{55}$$

where K is given by (51). Since the solution for pressure is not unique in Q, we then have

$$\|e_m\|_V \leq \frac{\|a\| + \alpha_1}{\alpha_1 \alpha_2} \epsilon \left(\|r_m\|_{Q/\mathrm{Ker}(B^t)} + \|r_{m-1}\|_{Q/\mathrm{Ker}(B^t)}\right), \tag{56}$$

$$\|r_m\|_{Q/\mathrm{Ker}(B^t)} \leq K\epsilon \left(\|r_m\|_{Q/\mathrm{Ker}(B^t)} + \|r_{m-1}\|_{Q/\mathrm{Ker}(B^t)}\right). \tag{57}$$

These two inequalities lead directly to (49) and (50). ∎

Next, we consider another abstract framework which is applicable to linear elasticity problems and compressible fluid flow problems. Let V and Q be two Hilbert spaces and M be another Hilbert space such that $Q \subset M$. Assume that $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ and $c(\cdot,\cdot)$ are bounded bilinear forms on $V \times V$, $V \times M$, and $Q \times M$ respectively, such that

$$a(w,v) \le \|a\| \|w\|_V \|v\|_V, \quad \forall w \in V,\ v \in V, \tag{58}$$
$$b(w,q) \le \|b\| \|w\|_V \|q\|_M, \quad \forall w \in V,\ q \in M, \tag{59}$$
$$c(p,q) \le \|c\| \|p\|_Q \|q\|_M, \quad \forall p \in Q,\ q \in M. \tag{60}$$

Let f and g be two bounded linear functionals in V' and Q' respectively, with $\langle \cdot,\cdot \rangle$ denoting the duality. As in the previous case, we assume $b(\cdot,\cdot)$ satisfies the inf-sup condition

$$\inf_{\substack{q \in Q \\ q \ne 0}} \sup_{\substack{v \in V \\ v \ne 0}} \frac{b(v,q)}{\|v\|_V \|q\|_M} \ge \beta, \tag{61}$$

where β is a positive constant. As opposed to the previous case, we assume that $a(\cdot,\cdot)$ is coercive on V, that is, there is a constant $\sigma > 0$ such that

$$a(v,v) \ge \sigma \|v\|_V^2, \quad \forall v \in V, \tag{62}$$

and that $c(\cdot,\cdot)$ satisfies

$$c(q,q) \ge -\tau \|q\|_M^2, \quad \forall q \in Q, \tag{63}$$

where τ is a negative or small nonnegative constant. The problem can be formulated as follows. Find $\{u,p\} \in V \times Q$ such that

$$a(u,v) + b(v,p) = \langle f, v \rangle_{V' \times V}, \quad \forall v \in V, \tag{64}$$
$$b(u,q) - c(p,q) = \langle g, q \rangle_{Q' \times Q}, \quad \forall q \in Q. \tag{65}$$

This formulation is similar to the one in [10] and to Proposition 2.12 in Chapter 2 of [3]. Its stabilized iterative scheme takes the same form as in the first formulation. Letting ϵ be a small positive number and $p_0 \in Q$ an initial guess, find $\{u_m, p_m\} \in V \times Q$ such that, for $m = 0, 1, \cdots,$

$$a(u_m,v) + b(v,p_m) = \langle f, v \rangle, \quad \forall v \in V, \tag{66}$$
$$b(u_m,q) - c(p_m,q) - (\epsilon p_m, q)_M = \langle g, q \rangle - (\epsilon p_{m-1}, q)_M, \quad \forall q \in Q. \tag{67}$$

Note that the inner product in M is used in (67) for the stabilizing terms, in contrast to the inner product in Q in (25). Under the assumptions above, we can prove that there exists a unique solution in $V \times Q$ for (64)-(65) and for (66)-(67), when $\tau < \sigma \|a\|^{-2} \beta^2$.

Theorem 2.3 *Let $\{u,p\}$ be the solution to the system (64)-(65), and $\{u_m, p_m\}$ the solution of (66)-(67). Assume that (58), (59), (61), (62), and (63) hold with $\tau < \sigma \|a\|^{-2} \beta^2$, then for any $\epsilon < \sigma \|a\|^{-2} \beta^2 - \tau$ and $p_0 \in Q$,*

$$\|p - p_m\|_M \le \left(\frac{\epsilon}{\sigma \|a\|^{-2} \beta^2 - \tau - \epsilon} \right)^m \|p - p_0\|_M, \quad m = 1, 2, \cdots, \tag{68}$$

$$\|u - u_m\|_V \le \frac{\|b\|}{\sigma} \left(\frac{\epsilon}{\sigma \|a\|^{-2} \beta^2 - \tau - \epsilon} \right)^m \|p - p_0\|_M, \quad m = 1, 2, \cdots. \tag{69}$$

Proof: Subtracting the equations in (66)-(67) from the equations in (64)-(65) we have the error equations

$$a(u-u_m,v)+b(v,p-p_m)=0, \quad \forall v\in V, \tag{70}$$

$$b(u-u_m,q)-c(p-p_m,q)=-\epsilon(p_m-p_{m-1},q)_M, \quad \forall q\in Q. \tag{71}$$

We define the operators $A:V\to V'$, $B:V\to Q'$, $B^t:Q\to V'$, $C:Q\to Q'$, and $E:Q\to M'\subset Q'$ by

$$\langle Aw,v\rangle_{V'\times V}=a(w,v), \quad \forall w,\ v\in V, \tag{72}$$

$$\langle Bv,q\rangle_{Q'\times Q}=\langle B^tq,v\rangle_{V'\times V}=b(v,q), \quad \forall v\in V,\ \forall q\in Q, \tag{73}$$

$$\langle Cp,q\rangle_{Q'\times Q}=c(p,q), \quad \forall p,\ q\in Q, \tag{74}$$

$$\langle Ep,q\rangle_{M'\times M}=(\epsilon p,q)_M, \quad \forall p,\ q\in Q. \tag{75}$$

Then equations (70)-(71) become

$$A(u-u_m)+B^t(p-p_m)=0, \quad \text{in } V', \tag{76}$$

$$B(u-u_m)-C(p-p_m)=-E(p_m-p_{m-1}), \quad \text{in } Q'. \tag{77}$$

Solving for $u-u_m$ in (76) and substituting it into (77) we see that

$$-BA^{-1}B^t(p-p_m)-C(p-p_m)=-E(p_m-p_{m-1}), \quad \text{in } Q', \tag{78}$$

which leads to

$$\begin{aligned}&\langle BA^{-1}B^t(p-p_m),p-p_m\rangle_{Q'\times Q}+\langle C(p-p_m),p-p_m\rangle_{Q'\times Q}\\ &=\langle E(p_m-p_{m-1}),p-p_m\rangle_{Q'\times Q}.\end{aligned} \tag{79}$$

By the definitions of the operators B, C, and E, we have

$$b(A^{-1}B^t(p-p_m),p-p_m)+c(p-p_m,p-p_m)=(\epsilon(p_m-p_{m-1}),p-p_m)_M. \tag{80}$$

To estimate the first term in (80), we let $w=A^{-1}B^t(p-p_m)$. Then $Aw=B^t(p-p_m)$, and $a(w,v)=b(v,p-p_m)$, $\forall v\in V$. By (58), (62) (72), and (73) we have

$$\begin{aligned}b(A^{-1}B^t(p-p_m),p-p_m)&=b(w,p-p_m)\\ &=a(w,w)\\ &\geq\sigma\|w\|_V^2\\ &=\sigma\|A^{-1}B^t(p-p_m)\|_V^2\\ &\geq\sigma\left[\|A\|^{-1}\|B^t(p-p_m)\|_V\right]^2\\ &\geq\sigma\|a\|^{-2}\|B^t(p-p_m)\|_V^2.\end{aligned} \tag{81}$$

From the inf-sup condition (61), we see that

$$\begin{aligned}\|B^t(p-p_m)\|_V&=\sup_{\substack{v\in V\\ v\neq 0}}\frac{\langle B^t(p-p_m),v\rangle_{V'\times V}}{\|v\|_V}\\ &=\sup_{\substack{v\in V\\ v\neq 0}}\frac{b(v,p-p_m)}{\|v\|_V}\\ &\geq\beta\|p-p_m\|_M.\end{aligned} \tag{82}$$

For the second term of (80), we use (63) to get

$$c(p - p_m, p - p_m) \geq -\tau \|p - p_m\|_M^2. \tag{83}$$

Applying (81), (82) and (83), equation (80) gives rise to

$$(\sigma \|a\|^{-2} \beta^2 - \tau)\|p - p_m\|_M^2 \leq \epsilon \|p_m - p_{m-1}\|_M \|p - p_m\|_M.$$

By the triangle inequality we have

$$(\sigma \|a\|^{-2} \beta^2 - \tau)\|p - p_m\|_M \leq \epsilon \|p - p_m\|_M + \epsilon \|p - p_{m-1}\|_M. \tag{84}$$

The formula (68) follows from (84) directly.

Letting now $v = u - u_m$ in (70) we obtain

$$a(u - u_m, u - u_m) + b(u - u_m, p - p_m) = 0.$$

In view of the coerciveness of $a(\cdot,\cdot)$ and the boundedness of $b(\cdot,\cdot)$ we have

$$\sigma \|u - u_m\|_V^2 \leq \|b\| \|u - u_m\|_V \|p - p_m\|_M.$$

That is,

$$\|u - u_m\|_V \quad \leq \sigma^{-1} \|b\| \|p - p_m\|_M. \tag{85}$$

The inequalities (85) and (68) then imply (69). This completes the proof. ∎

Note that in the iterative schemes (24)-(25) and (66)-(67), a more general stabilizing or regularizing term like $\epsilon d(\cdot,\cdot)$, instead of $\epsilon(\cdot,\cdot)_Q$ or $\epsilon(\cdot,\cdot)_M$, could be used, where $d(\cdot,\cdot)$ is a coercive and bounded bilinear form on $Q \times Q$ for the first formulation and on $M \times M$ for the second.

The iterative schemes in the forms of (24)-(25) and (66)-(67) seem to be new, although special cases in the name of augmented Lagrangian methods have been considered for elliptic and Stokes equations [2, 7, 11]. In infinite dimensional spaces for $V \times Q$, these schemes represent continuous differential problems, while in the finite dimensional case, they represent mixed finite element methods. Note that the analysis in [2] covers the case when the bilinear form $c(\cdot,\cdot)$ is zero and B is the divergence operator.

There are some notable differences in the assumptions for the two formulations. In formulation (24)-(25), the symmetry of $c(\cdot,\cdot)$ is required, while $c(\cdot,\cdot)$ may be non-symmetric in (66)-(67); Formulation (66)-(67) requires that $a(\cdot,\cdot)$ be coercive on V while formulation (24)-(25) requires that $a(\cdot,\cdot)$ be invertible on $\mathrm{Ker}(B)$.

3 Implementation Issues

At the matrix level, the mixed finite element methods (16)-(17) and (64)-(65) have the form

$$AU + B^t P = F, \tag{86}$$

$$BU - CP = G, \tag{87}$$

where U and P signify vectors containing the nodal values of the trial functions for u and p, and the iterative schemes (24)-(25) and (66)-(67) have the form

$$AU_m + B^t P_m = F, \tag{88}$$

$$BU_m - (C + \epsilon D)P_m = G - \epsilon D P_{m-1}, \tag{89}$$

where A is a positive semidefinite matrix in the first formulation, and a positive definite matrix in the second, and D is a positive definite matrix. Here U_m and P_m are vectors corresponding to u_m and p_m. Solving for P_m in (89) and substituting it into (88) we obtain

$$AU_m + B^t(C + \epsilon D)^{-1} BU_m = F + B^t(C + \epsilon D)^{-1}(G - \epsilon D P_{m-1}), \tag{90}$$

$$P_m = (C + \epsilon D)^{-1}(\epsilon D P_{m-1} + BU_m - G). \tag{91}$$

The matrices C and D are diagonal or banded with small band width. The system (90) is positive definite and decoupled from (91). Thus our stabilized schemes not only have more stability, but also lead to smaller and positive definite systems (with velocity unknowns u_m only). Compared to penalty methods [8, 9], the stabilized schemes do not require that the parameter ϵ be very small, since the iterative process has geometrical convergence with respect to ϵ; as a result, this improves the condition number of the linear system (90).

At the matrix level, we see that the augmented Lagrangian method (14)-(15) is a special case of our scheme (90)-(91).

4 Applications

In this section, we apply our abstract theory to various physical problems, which include second order elliptic equations, incompressible Stokes equations, linear elasticity problems, and compressible Stokes flow problems.

4.1 Second Order Elliptic Equations

Consider the elliptic problem with Dirichlet boundary condition: find $p \in H^1(\Omega)$ such that

$$-\text{div}\ (K(x)\nabla p) + \phi(x)p = f(x), \quad x \in \Omega, \tag{92}$$

$$p = h(x), \quad x \in \Gamma. \tag{93}$$

We assume that $\phi \in L^\infty(\Omega)$ is non-negative, $h \in H^{1/2}(\Omega)$, and K is a $d \times d$ positive definite matrix with its smallest eigenvalue uniformly bounded away from zero on $\Omega \subset \mathcal{R}^d$. Combining the boundary condition (93) and the flux definition $u = -K\nabla p$ we have

$$(K^{-1}u, v) - (p, \text{div}\ v) = -\langle h, v \cdot \nu \rangle, \quad \forall v \in H(\text{div}; \Omega), \tag{94}$$

where $(\cdot,\cdot)$ denotes the standard inner product in $L^2(\Omega)$ and $\langle \cdot,\cdot \rangle$ denotes the duality in the spaces $H^{1/2}(\Gamma)$ and $H^{-1/2}(\Gamma)$.

Define the Hilbert spaces $V = H(\text{div};\Omega) = \{v : v \in (L^2(\Omega))^d, \text{div } v \in L^2(\Omega)\}$ and $Q = L^2(\Omega)$, and the bilinear forms

$$a(u,v) = \int_\Omega K^{-1}u \cdot v dx, \tag{95}$$

$$b(u,q) = -\int_\Omega q \text{ div } u dx, \tag{96}$$

$$c(p,q) = \int_\Omega \phi(x) p\, q dx. \tag{97}$$

With the definitions (96) and (21), B is the divergence operator from $H(\text{div};\Omega)$ into $L^2(\Omega)$ and is also surjective. By a basic functional analysis result

$$\text{Ker}(B^t)^\perp = \overline{\text{Im}(B)},$$

we see that $\text{Ker}(B^t) = \{0\}$. Thus (23) holds with $\gamma = 1$. Note that $a(\cdot,\cdot)$ is coercive on $\text{Ker}(B)$, which implies (18)-(19). Then Theorem 2.1 applies.

At the discrete level, let $V \times Q$ be some finite dimensional spaces, e.g. Raviart-Thomas-Nedelec space of index k, Brezzi-Douglas-Marini space of index k, or Brezzi-Douglas-Fortin-Marini space of index $k+1$, corresponding to an affine finite element triangulation; see [3] for the specific definition of these spaces. Then B is the restriction to V of the divergence operator, and we still have $\text{Ker}(B^t) = \{0\}$. In the discrete case, $\text{Ker}(B)$ is a subset of the kernel of the divergence operator; thus the coerciveness of $a(\cdot,\cdot)$ on $\text{Ker}(B)$ still holds. Also, the commuting diagram property implies the discrete inf-sup condition. Thus the mixed finite element approximation for the problem (92)-(93) takes the form of (16)-(17) and its stabilized scheme takes the form of (24)-(25), and Theorem 2.1 also applies to the discrete case.

It is trivial to apply the theory to other types of boundary conditions, e.g. Neumann, Robin or mixed boundary conditions.

Different approaches [1, 6] have been explored to circumvent the difficulty of ill-conditioning of the linear systems arising from mixed finite element methods.

4.2 Incompressible Stokes Equations

Following the notation in the introduction section of this paper, we put the stabilized scheme for incompressible Stokes problems (1)-(3) in the fashion: find $\{u_m, p_m\} \in V \times Q$ satisfying, for $m = 0, 1, 2, \cdots$,

$$(\nabla u_m, \nabla v) - (\text{div } v, p_m) = (f, v), \quad \forall v \in V, \tag{98}$$

$$(\text{div } u_m, q) + \epsilon(p_m, q) = \epsilon(p_{m-1}, q), \quad \forall q \in Q, \tag{99}$$

where $V \times Q$ denotes $(H_0^1(\Omega))^d \times \{q : q \in L^2(\Omega), \int_\Omega q dx = 0\}$, or some finite element spaces [3].

It is easy to see that the scheme (98)-(99) is equivalent to the so-called augmented Lagrangian method (8)-(9). Indeed, (99) is the same as (9). Substituting (9) into (98) we have

$$(\nabla u_m, \nabla v) + \frac{1}{\epsilon}(\Pi(\text{div } u_m), \text{div } v) = (f, v) + (\text{div } v, p_{m-1}), \quad \forall v \in V. \tag{100}$$

The equation (100) is nothing but (8).

An analysis at the linear algebra level for inexact Uzawa's algorithms was given in [5]. Uzawa's iterative algorithm also stabilizes the ill-conditioning saddle point linear systems. But its convergence rate can be proved to be always slower than our scheme. Also, the convergence analysis and the convergence rate of Uzawa's algorithms depend on the eigenvalues of some matrices, while our convergence does not.

4.3 Linear Elasticity Problems

Consider the linear elasticity problem: find $\{u, p\} \in V \times Q$ such that

$$a(u,v) + b(v,p) = \langle f, v\rangle_{V'\times V}, \quad \forall v \in V, \tag{101}$$

$$b(u,q) - c(p,q) = 0, \quad \forall q \in Q, \tag{102}$$

where $V = (H_0^1(\Omega))^2, Q = L^2(\Omega)$, and

$$a(u,v) = \int_\Omega 2\mu\Theta(u) : \Theta(v)dx, \tag{103}$$

$$b(u,q) = -\int_\Omega q \text{ div } u dx, \tag{104}$$

$$c(p,q) = \int_\Omega \frac{1}{\lambda} p\, q dx. \tag{105}$$

Here $\Theta(u)$ is the linearized strain tensor corresponding to displacement u, and we have set $p = \lambda \text{div } u$. When the Lame coefficient λ is large, the stability of (101)-(102) is weak. Thus it makes sense to stabilize the problem. Notice that $a(\cdot,\cdot)$ is coercive on V by Korn's first inequality

$$\int_\Omega |\Theta(u)|^2 dx \geq \alpha \|u\|_1^2, \quad \forall u \in (H_0^1(\Omega))^2,$$

and that we can choose $\tau = 0$ in (63). Thus the second formulation and Theorem 2.3 apply.

Various mixed finite element spaces are discussed in [3]. These spaces can be used to construct finite-dimensional approximations to problem (101)-(102).

The schemes here apply unchangeably to nearly-incompressible flow equations.

4.4 Compressible Stokes Flow

Linearizing the steady-state, compressible, viscous Navier-Stokes equations about a given solution, we have [10]

$$-(2\mu+\lambda)u_{xx} - \mu u_{yy} - (\mu+\lambda)v_{xy} + \rho U u_x + \rho V u_y + p_x = f_1, \text{ in } \Omega, \tag{106}$$

$$-2\mu v_{xx} - (2\mu+\lambda)v_{yy} - (\mu+\lambda)u_{xy} + \rho U v_x + \rho V v_y + p_y = f_2, \text{ in } \Omega, \tag{107}$$

$$\rho'(Up_x + Vp_y) + \rho(u_x + v_y) = f_3, \text{ in } \Omega, \tag{108}$$

$$u = v = 0, \quad \text{on } \Gamma, \tag{109}$$

where $\Omega \subset R^2$ with boundary Γ, u, v, p are the velocity and pressure variables, U, V, P are given functions of (x, y) describing the ambient flow, $\rho(P)$ is a given density function and $\rho' = \dfrac{d\rho}{dP}$, μ, λ are two viscosity coefficients.

Following the work in [10], we define the Hilbert spaces $V = H_0^1(\Omega) \times H_0^1(\Omega)$, $M = \{q : q \in L^2(\Omega), \int_\Omega q dx = 0\}$, and Q to be the set of functions q in M such that

$$\|q\|_Q^2 = \|q\|_M^2 + \int_\Omega \left[\frac{\rho' U q_x + \rho' V q_y}{\rho}\right]^2 dxdy < \infty, \tag{110}$$

and the bilinear forms

$$\begin{aligned} & a([u,v],[w,s]) \\ & = \int_\Omega [(2\mu+\lambda)u_x w_x + \mu u_y w_y + (\mu+\lambda)v_x w_y + \mu v_x s_x + (2\mu+\lambda)v_y s_y + (\mu+\lambda)u_x s_y \\ & \quad +\rho(Uu_x + Vu_y)w + \rho(Uv_x + Vv_y)s]dxdy, \ \ \forall\ [u,v] \in V,\ \forall\ [w,s] \in V, \end{aligned} \tag{111}$$

$$b([u,v],q) = -\int_\Omega q(u_x + v_y)dxdy, \quad \forall\ [u,v] \in V,\ \forall q \in M, \tag{112}$$

$$c(p,q) = -\int_\Omega (\rho' U p_x + \rho' V p_y) q/\rho dxdy, \quad \forall p \in Q,\ \forall q \in M. \tag{113}$$

Finite dimensional approximations to the spaces V and Q can be constructed (for example, the MINI element); they are still denoted by V and Q, for simplicity. Thus the mixed finite element method (64)-(65) applies to (106)-(109), and its stabilized scheme is (66)-(67).

5 Numerical Experiments

In this section, we conduct some numerical experiments with our stabilized iterative schemes. It has been demonstrated that these methods work well for Stokes equations [2]. Thus we just consider the following two second order elliptic equations, one with continuous coefficients while the other with discontinuous coefficients.

Example 1:

$$-\text{div}\ (K(x)\nabla p) + \phi(x)p = f(x), \quad x \in \Omega, \tag{114}$$

$$K(x)\frac{\partial p}{\partial \nu} = -h(x), \quad x \in \partial\Omega, \tag{115}$$

where $\Omega = (0,1) \times (0,1)$ with boundary $\partial\Omega$ and outward unit normal ν,

$$K(x) = \begin{bmatrix} e^x & y \\ x & e^y \end{bmatrix}$$

and $\phi(x) = \dfrac{1}{1+x+y}$. The functions f and h were chosen such that the exact solution of the differential problem is

$$p(x,y) = e^{xy}.$$

Example 2: Consider the same equations (114)-(115) as in Example 1, except that we here let $\Omega = (-0.5, 0.5) \times (0, 0.5)$, $\phi(x) = 0.001$, the coefficient matrix

$$K(x) = \begin{bmatrix} 5 + x^2 + y^2 & \sin(x) \\ \sin(x) & 50 + \cos(xy) \end{bmatrix},$$

Table 1: Relative L^∞ errors for Example 1 between iterates at current and previous iterations.

	grid $= \frac{1}{50} \times \frac{1}{50}$			grid $= \frac{1}{110} \times \frac{1}{110}$		
iteration	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-5}$	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-5}$
0	1.00	1.00	1.00	1.00	1.00	1.00
1	8.02E-4	8.05E-5	8.06E-6	7.94E-4	7.96E-5	7.97E-6
2	1.74E-6	1.50E-8	1.41E-10	1.48E-6	1.49E-8	1.49E-10
3	2.86E-9	3.19E-12	2.08E-13	2.83E-9	3.26E-12	6.25E-13

for $x \leq 0$ and

$$K(x) = \begin{bmatrix} \alpha(5 + x^2 + y^2) & \alpha(\sin x) \\ \alpha \sin(x) & \alpha(50 + \cos(xy)) \end{bmatrix},$$

for $x > 0$, where the parameter α represents the strength of discontinuity at $x = 0$. The true solution is

$$p(x, y) = \begin{cases} \alpha(\cos(\pi x) + \cos(\pi y)) + e^x - e^y)/\pi, & x \leq 0 \\ (\cos(\pi x) + \cos(\pi y)) + e^x - e^y)/\pi & x > 0. \end{cases}$$

Partition the domain Ω into a set of rectangles of size $h_1 \times h_2$. We employ the Raviart-Thomas [3, 12] space of index 0. Thus

$$V = (\mathcal{P}_1 \otimes \mathcal{P}_0) \times (\mathcal{P}_0 \otimes \mathcal{P}_1),$$

where $\mathcal{P}_k$ is the set of one variable polynomials of order less than or equal to k. Consequently, the pressure space Q consists of piecewise constants. Partitioning the domain into triangles, or applying higher order approximation polynomials can be treated analogously.

For our numerical experiments, we define relative L^∞ errors between iterates at current and previous iterations as

$$\max \left\{ \frac{\|p_m - p_{m-1}\|_\infty}{\|u_m\|_\infty}, \frac{\|p_m - p_{m-1}\|_\infty}{\|p_m\|_\infty} \right\},$$

and relative L^∞ errors between iterates and the exact solution as

$$\max \left\{ \frac{\|u - u_m\|_\infty}{\|p\|_\infty}, \frac{\|p - p_m\|_\infty}{\|p\|_\infty} \right\},$$

where $\|\cdot\|_\infty$ denotes the discrete L^∞ norm, $\{u, p\}$ denotes the exact solution to the differential problem, and $\{u_m, p_m\}$ denotes the iterative solution to the finite-dimensional problem at the m-th iteration level. The initial guess is always chosen to be zero.

Some results for Example 1 are listed in Tables 1 and 2 for two different grid sizes and three different values for the penalty parameter ϵ. From Table 1 we see that the convergence depends

Table 2: Relative L^∞ errors for Example 1 between iterates and the exact solution.

	grid $= \frac{1}{50} \times \frac{1}{50}$			grid $= \frac{1}{110} \times \frac{1}{110}$		
iteration	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-5}$	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-5}$
1	1.32E-3	5.73E-4	4.98E-4	9.05E-4	1.76E-4	1.03E-4
2	4.91E-4	4.89E-4	4.89E-4	9.72E-5	9.57E-5	9.56E-5
3	4.89E-4	4.89E-4	4.89E-4	9.56E-5	9.56E-5	9.56E-5

Table 3: Relative L^∞ errors for Example 2 between iterates at current and previous iterations with grid size $= \frac{1}{120} \times \frac{1}{70}$. The value for α represents the strength of discontinuity of the coefficient.

	$\alpha = 1$			$\alpha = 100$		
iterations	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-5}$	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-5}$
0	1.0	1.0	1.0	1.0	1.0	1.0
1	6.6E-7	6.6E-8	6.7E-9	3.9E-7	3.9E-8	9.1E-9
2	5.2E-10	9.3E-10	4.0E-10	7.9E-9	9.8E-9	7.7E-9
3	5.2E-10	5.2E-10	3.3E-10	8.9E-9	8.0E-9	7.7E-9

very slightly on the grid size and that a roughly geometrical convergence rate is achieved, as predicted by our theory.

For Example 2, $\alpha = 1$ corresponds to a continuous coefficient problem and $\alpha = 100$ corresponds to a strongly discontinuous coefficient problem. Tables 3 and 4 do not give an exactly the same convergence pattern and accuracy as in Example 1. This is the case partially because the convergence is also limited by the discretization error. Indeed, our final convergence theory should have the form

$$\|u - u_m\|_V + \|p - p_m\|_Q = O(h^{k+1} + \epsilon^m), \tag{116}$$

where $\{u, p\}$ denotes the exact solution to the infinite-dimensional differential problem, $\{u_m, p_m\}$ denotes the iterative solution to our stabilized discrete scheme at the m-th iteration level, and k is the degree of piecewise polynomials used for Q on a mixed finite element mesh with size h. Formula (116) tells us that the iterative convergence accuracy is affected by the discretization term $O(h^{k+1})$.

Overall, the numerical experiments are pretty convincing. They give an approximately geometrical convergence and reasonable accuracy for strongly discontinuous problems. In the implementation, however, our stabilized schemes are much easier to deal with since all linear systems are positive definite, and symmetric if original differential problems are symmetric.

6 Concluding Remarks

We have proposed and analyzed stabilized iterative schemes for two abstract formulations of mixed finite element methods. These schemes overcome the disadvantages of mixed finite element methods, in that they lead to positive definite linear systems and decouple the velocity

Table 4: Relative L^∞ errors for Example 2 between iterates and the exact solution with grid size $= \frac{1}{120} \times \frac{1}{70}$. The value for α represents the strength of discontinuity of the coefficient.

	$\alpha = 1$			$\alpha = 100$		
iterations	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-5}$	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-5}$
1	3.1E-3	3.1E-3	3.1E-3	8.3E-1	8.3E-1	8.3E-1
2	3.1E-3	3.1E-3	3.1E-3	8.3E-1	8.3E-1	8.3E-1
3	3.1E-3	3.1E-3	3.1E-3	8.3E-1	8.3E-1	8.3E-1

variable from the pressure variable. When the given physical problem is symmetric, the resulting linear system is also symmetric. Besides, these schemes keep the advantages of mixed finite element methods. For example, the velocity and pressure variables are approximated at the same accuracy even in the case of rough coefficients (e.g. $K(x)$ in (92) is varying rapidly in its entries and its eigenvalues)

The stabilized iterative schemes considered in this paper are generalizations of the augmented Lagrangian methods in the sense that, when applied to incompressible Stokes equations, it is equivalent to the well known augmented Lagrangian scheme. However, our schemes do not have to come from the context of optimization and thus apply to more general problems.

Compared to penalty methods, our schemes have geometrical convergence. This not only provides very fast convergence, but also improves the conditioning of the resulting linear system since the parameter ϵ does not have to be chosen very small.

Although our presentation has been in general Hilbert spaces, there are notable differences between finite-dimensional and infinite-dimensional cases. For example, in the finite-dimensional case, the conditions (18) and (19) are equivalent.

Our schemes also apply to time-dependent problems in an obvious way.

References

[1] M. B. Allen, R. E. Ewing, and P. Lu, Well-conditioned iterative scheme for mixed finite element models of porous media, SIAM J. Sci. Statist. Comput. 13(1992) 794-814.

[2] S. C. Brenner and L. R. Scott, The Mathematical Theory of Finite Element Methods, Springer-Verlag, New York, 1994.

[3] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, New York, 1991.

[4] J. Douglas, Jr., and J. P. Wang, An absolutely stabilized finite element method for the Stokes problem, Math. Comp. 52(1989) 495-508.

[5] H. C. Elman and G. H. Golub, Inexact and preconditioned Uzawa algorithms for saddle point problems, SIAM J. Numer. Anal. 31(1994) 1645-1661.

[6] R. E. Ewing, R. D. Lazarov, P. Lu, and P. S. Vassilevski, Preconditioning indefinite systems arising from mixed finite element discretization of second-order elliptic problems, Lecture Notes in Math. 1457, Springer, Berlin, 1990.

[7] M. Fortin and R. Glowinski, Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems, Studies in Mathematics and Its Applications 15, North-Holland, 1983.

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

[9] M. D. Gunzburger, Finite element methods for viscous incompressible flows, Academic Press, San Diego, 1989.

[10] R. B. Kellog and B. Liu, A finite element method for the compressible Stokes equations, SIAM J. Numer. Anal. 33(1996) 780-788.

[11] A. Quarteroni and A. Valli, Numerical Approximation of Partial Differential Equations, Springer-Verlag, Berlin, 1994.

[12] P.–A. Raviart and J.–M. Thomas, A mixed finite element method for second order elliptic problems, in: I. Galligani and E. Magenes, Eds., Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics 606, Springer-Verlag, Berlin and New York, 1977, pp. 292-315.