SIAM J. NUMER. ANAL.
Vol. 44, No. 6, pp. 2633–2649

A UNIFIED APPROACH FOR UZAWA ALGORITHMS*

CONSTANTIN BACUTA†

Abstract. We present a unified approach in analyzing Uzawa iterative algorithms for saddle point problems. We study the classical Uzawa method, the augmented Lagrangian method, and two versions of inexact Uzawa algorithms. The target application is the Stokes system, but other saddle point systems, e.g., arising from mortar methods or Lagrange multipliers methods, can benefit from our study. We prove convergence of Uzawa algorithms and find optimal rates of convergence in an abstract setting on finite- or infinite-dimensional Hilbert spaces. The results can be used to design multilevel or adaptive algorithms for solving saddle point problems. The discrete spaces do not have to satisfy the LBB stability condition.

Key words. Uzawa algorithms, saddle point system, multilevel methods, augmented Lagrangian method, Stokes problem

AMS subject classifications. 74S05, 74B05, 65N22, 65N55

DOI. 10.1137/050630714

1. Introduction. In this paper, we provide a unified approach for Uzawa methods for linear saddle point systems. Such systems arise in solving various partial differential equations (PDEs) or systems of PDEs at the continuous level or at the discrete level. Typical examples of such PDEs are second-order elliptic problems, Stokes equations, and elasticity problems. We analyze the classical Uzawa Method (UM) [1], the augmented Lagrangian Uzawa method (ALUM) [14], the inexact Uzawa method (IUM) [7, 13], and a modified (or multilevel) inexact Uzawa method (MIUM) under a general approach on abstract Hilbert spaces. The motivation for considering abstract versions of Uzawa algorithms on infinite-dimensional Hilbert spaces is that the analysis at the continuous level of an algorithm for solving a PDE gives the right strategy for discretizing the PDE. In addition, the convergence factors of certain multilevel or adaptive algorithms for solving saddle point systems depend on the stability parameters of the continuous problem, and in many cases the discrete LBB stability condition is not required to be satisfied (see [4, 12] or section 6). Next, we formulate the general framework of the saddle point problem to be studied in this paper and indicate the way the paper is organized.

We let $\mathbf{V}$ and P be two Hilbert spaces with inner products $a(\cdot,\cdot)$ and $(\cdot,\cdot)$, with the corresponding induced norms $|\cdot|_{\mathbf{V}} = |\cdot| = a(\cdot,\cdot)^{1/2}$ and $\|\cdot\|_P = \|\cdot\| = (\cdot,\cdot)^{1/2}$. The dual parings on $\mathbf{V}^* \times \mathbf{V}$ and $P^* \times P$ are denoted by $\langle\cdot,\cdot\rangle$ and $(\cdot,\cdot)$, respectively. Here, $\mathbf{V}^*$ and P^* denote the dual of $\mathbf{V}$ and P, respectively. We identify P^* and P as Hilbert spaces so that $(\cdot,\cdot)$ represents both the inner product on P and the duality between P^* and P. In applications to Stokes systems, $\mathbf{V} = (H_0^1)^d (d = 2, 3, \dots)$, P is a subspace of L^2 of codimension one and $(\cdot,\cdot)$ is the standard inner product on L^2. Next, we consider that $b(\cdot,\cdot)$ is a continuous bilinear form on $\mathbf{V} \times P$, satisfying the

*Received by the editors May 5, 2005; accepted for publication (in revised form) July 10, 2006; published electronically December 11, 2006. This work was partially supported by the University of Delaware Research Foundation.
http://www.siam.org/journals/sinum/44-6/63071.html

†Department of Mathematical Sciences, University of Delaware, 501 Ewing Hall, Newark, DE 19716 (bacuta@math.udel.edu).

inf-sup condition. More precisely, we assume that

$$\inf_{p\in P} \sup_{\mathbf{v}\in \mathbf{V}} \frac{b(\mathbf{v},p)}{\|p\|\ |\mathbf{v}|} = m > 0 \tag{1.1}$$

and

$$\sup_{p\in P} \sup_{\mathbf{v}\in \mathbf{V}} \frac{b(\mathbf{v},p)}{\|p\|\ |\mathbf{v}|} = M < \infty. \tag{1.2}$$

For $f \in \mathbf{V}^*$, $g \in P^*$, we consider the following variational problem:
Find $(\mathbf{u}, p) \in \mathbf{V} \times P$ such that

$$\begin{array}{llll} a(\mathbf{u},\mathbf{v}) & + \quad b(\mathbf{v},p) & = \langle \mathbf{f},\mathbf{v}\rangle & \text{for all } \mathbf{v}\in \mathbf{V},\\ b(\mathbf{u},q) & & = (g,q) & \text{for all } q \in P. \end{array} \tag{1.3}$$

It is known that the above variational problem has a unique solution for any $f \in \mathbf{V}^*$, $g \in P^*$ (see [9, 10, 15] or Lemma 2.1). With the forms a and b, we associate two linear operators $A : V \to V^*$ and $B : V \to P$ defined by

$$\langle A\mathbf{u},\mathbf{v}\rangle = a(\mathbf{u},\mathbf{v}) \quad \text{ for all } \mathbf{u},\mathbf{v} \in \mathbf{V}$$

and

$$(B\mathbf{u}, q) = b(\mathbf{u}, q) \quad \text{ for all } \mathbf{u} \in \mathbf{V},\ q \in P.$$

Let $B^* : P \to V^*$ be the dual operator of B defined by

$$\langle B^*q,\mathbf{v}\rangle = (q, B\mathbf{v}) = (B\mathbf{v}, q) = b(\mathbf{v}, q) \quad \text{ for all } \mathbf{v} \in \mathbf{V},\ q \in P.$$

The problem (1.3) is equivalent to the following problem:
Find $(\mathbf{u}, p) \in \mathbf{V} \times P$ such that

$$\begin{array}{lll} A\mathbf{u} & + \quad B^*p & = \mathbf{f},\\ B\mathbf{u} & & = g. \end{array} \tag{1.4}$$

In this framework, we analyze Uzawa algorithms for solving the system (1.3) or (1.4). We consider that the form a gives the inner product and the norm on $\mathbf{V}$. A more general case of (1.3) is considered in [9, 10, 15]. Our particular assumptions for the form a give rise to a simplified analysis. For the general case, we obtain sharp convergence estimates only in terms of the two constants m and M.

The rest of the paper is organized as follows. In section 2, we analyze the convergence of the classical Uzawa algorithm. The augmented Lagrangian Uzawa method is analyzed in section 3 (Fortin and Glowinski [14]). In section 4, we shall investigate the convergence of the inexact Uzawa algorithm (Bramble, Pasciak, and Vassilev [7] and Elman and Golub [13]) in the above abstract framework. Applications to discretizations on stable pairs are presented in section 5. A modified inexact Uzawa algorithm with applications in constructing multilevel methods and adaptive methods for solving (1.3) is illustrated in section 6. In section 7, we present applications of our abstract results to the Stokes system.

2. The abstract Uzawa algorithm. We begin this section with two lemmas which provide basic properties of norms and operators introduced in section 1. The proofs are based on the Riesz representation theorem (see, e.g., [20]). For completeness, we include the proofs.

LEMMA 2.1. *The operator $A : V \to V^*$ is invertible and the Schur complement operator $BA^{-1}B^* : P \to P$ is symmetric and a positive definite operator satisfying*

$$(BA^{-1}B^*p, p) = \sup_{\mathbf{v}\in\mathbf{V}} \frac{b(\mathbf{v}, p)^2}{|\mathbf{v}|^2}, \tag{2.1}$$

$$m^2\|p\|^2 \le (BA^{-1}B^*p, p) \le M^2\|p\|^2, \quad p \in P. \tag{2.2}$$

Consequently, the problem (1.3) *(or* (1.4)*) has a unique solution.*

Proof. From the definition of A, we get that A is a bounded injective operator. Using the Riesz representation theorem, it follows that A is also a surjective operator. Let us further note that A satisfies

$$\langle A\mathbf{u}, \mathbf{v}\rangle = a(\mathbf{u}, \mathbf{v}) = a(\mathbf{v}, \mathbf{u}) = \langle A\mathbf{v}, \mathbf{u}\rangle,$$

and the changes of variable $A\mathbf{u} = \mathbf{u}^*$ and $A\mathbf{v} = \mathbf{v}^*$ lead to

$$\langle \mathbf{u}^*, A^{-1}\mathbf{v}^*\rangle = \langle \mathbf{v}^*, A^{-1}\mathbf{u}^*\rangle, \quad \mathbf{u}^*, \mathbf{v}^* \in \mathbf{V}^*. \tag{2.3}$$

Using (2.3), we obtain

$$\begin{aligned}(BA^{-1}B^*p, q) &= \langle B^*q, A^{-1}B^*p\rangle = \langle B^*p, A^{-1}B^*q\rangle \\ &= (BA^{-1}B^*q, p) = (p, BA^{-1}B^*q), \quad p, q \in P.\end{aligned}$$

To prove (2.1), we let $p \in P$ be fixed and consider the following problem:

Find $\mathbf{u} \in \mathbf{V}$ such that

$$a(\mathbf{u}, \mathbf{v}) = b(\mathbf{v}, p) \qquad \text{for all } \mathbf{v} \in \mathbf{V}. \tag{2.4}$$

Since the functional $\mathbf{v} \to b(\mathbf{v}, p)$ is continuous on $\mathbf{V}$, by the Riesz representation theorem we have that the unique solution $\mathbf{u}$ of (2.4) satisfies

$$a(\mathbf{u}, \mathbf{u}) = \|\mathbf{v} \to b(\mathbf{v}, p)\|^2_{\mathbf{V}^*} = \sup_{\mathbf{v}\in\mathbf{V}} \frac{b(\mathbf{v}, p)^2}{|\mathbf{v}|^2}. \tag{2.5}$$

On the other hand, from (2.4) we have

$$A\mathbf{u} = B^*p \;\text{ or }\; \mathbf{u} = A^{-1}B^*p$$

and

$$a(\mathbf{u}, \mathbf{u}) = \langle A\mathbf{u}, \mathbf{u}\rangle = \langle B^*p, A^{-1}B^*p\rangle = (p, BA^{-1}B^*p). \tag{2.6}$$

Thus, (2.1) follows from (2.5) and (2.6). The estimate (2.2) follows imediately from (2.1), (1.1), and (1.2).

To prove the existence and uniqueness of (1.3) (or (1.4)), we substitute $\mathbf{u}$ from the first equation of (1.4) into the second equation of (1.4). The resulting equation in p,

$$BA^{-1}B^*p = BA^{-1}\mathbf{f} - g,$$

has a unique solution due to the fact that $BA^{-1}B^* : P \to P$ is symmetric and a positive definite operator. ∎

Remark 2.2. From the general theory of symmetric operators and Lemma 2.1, we have that $\sigma(BA^{-1}B^*) \subset [m^2, M^2]$ and $m^2, M^2 \in \sigma(BA^{-1}B^*)$. In the finite-dimensional case, m^2 and M^2 are the extreme eigenvalues of the Schur complement $BA^{-1}B^*$.

LEMMA 2.3. *The following norm estimates are valid:*

$$\|\phi\|^2_{\mathbf{V}^*} = a(A^{-1}\phi, A^{-1}\phi) = |A^{-1}\phi|^2, \quad \phi \in \mathbf{V}^*, \tag{2.7}$$

$$\|A\mathbf{u}\|_{\mathbf{V}^*} = |\mathbf{u}|, \quad \mathbf{u} \in \mathbf{V}, \tag{2.8}$$

$$\|B^*q\|_{\mathbf{V}^*} = |A^{-1}B^*q| = (BA^{-1}B^*q, q)^{1/2} \leq M\|q\|, \quad q \in P, \tag{2.9}$$

$$\|B\| = M, \quad \textit{hence} \quad \|B\mathbf{u}\| \leq M\ |\mathbf{u}|, \quad \mathbf{u} \in \mathbf{V}. \tag{2.10}$$

Proof. By the Riesz representation theorem, we have that for any $\phi \in \mathbf{V}^*$ the problem

Find $\mathbf{u} \in \mathbf{V}$ such that

$$\langle A\mathbf{u}, \mathbf{v}\rangle = a(\mathbf{u}, \mathbf{v}) = \langle \phi, \mathbf{v}\rangle \qquad \text{for all} \ \ \mathbf{v} \in \mathbf{V} \tag{2.11}$$

has a unique solution, and the solution $\mathbf{u}$ satisfies

$$a(\mathbf{u}, \mathbf{u}) = \sup_{\mathbf{v}\in\mathbf{V}} \frac{\langle \phi, \mathbf{v}\rangle^2}{a(\mathbf{v}, \mathbf{v})} = \|\phi\|^2_{\mathbf{V}^*}. \tag{2.12}$$

From (2.11), we have that $\mathbf{u} = A^{-1}\phi$, which combined with (2.12) gives (2.7). The equality (2.8) is a consequence of (2.7), and (2.9) follows from (2.8) and (2.2). The last estimate follows from the definition of B and the assumption in (1.2). ∎

Next, we present the Uzawa algorithm [1] for solving the solution of the abstract problem (1.3). Given a parameter $\alpha > 0$, called a relaxation parameter, the Uzawa algorithm for approximating the solution $(\mathbf{u}, p)$ of (1.3) can be described as follows.

ALGORITHM 2.4 (Uzawa method (UM)). *Let p_0 be any approximation for p, and for $k = 1, 2, \ldots,$ construct $(\mathbf{u}_k, p_k)$ by*

$$\begin{aligned} a(\mathbf{u}_k, \mathbf{v}) &= (\mathbf{f}, \mathbf{v}) - b(\mathbf{v}, p_{k-1}), \quad \mathbf{v} \in \mathbf{V}, \\ p_k &= p_{k-1} + \alpha(B\mathbf{u}_k - g). \end{aligned} \tag{2.13}$$

The convergence of the UM is discussed for particular cases in, e.g., [10, 14, 15, 17]. It shows that the UM is convergent for small enough α and that the convergence rate is the same as the convergence rate of the Richardson iterative methods for the Schur complement $BA^{-1}B^*$. For completeness, we include the proof.

THEOREM 2.5. *Let $(\mathbf{u}, p)$ be the solution of* (1.3) *and let $(\mathbf{u}_k, p_k)$ be the sequence of approximations built by the UM* (2.13). *Then, the following holds.*

(i) *The sequences $\mathbf{u} - \mathbf{u}_k$ and $p - p_k$ satisfy*

$$\begin{aligned} a(\mathbf{u} - \mathbf{u}_k, \mathbf{u} - \mathbf{u}_k)^{1/2} &\leq M\ \|p - p_{k-1}\|, \\ \|p - p_k\| &\leq \|I - \alpha BA^{-1}B^*\|\ \|p - p_{k-1}\|. \end{aligned}$$

(ii) *For $\alpha < \frac{2}{M^2}$, the UM is convergent and*

$$\|I - \alpha BA^{-1}B^*\| = \max\{|1 - \alpha m^2|, |1 - \alpha M^2|\} < 1.$$

(iii) *For* $\alpha = \frac{1}{M^2}$, *the convergence factor is* $\|I - \alpha BA^{-1}B^*\| = 1 - \frac{m^2}{M^2}$.
(iv) *The optimal convergence factor is achieved for*

$$\alpha_{opt} = \frac{2}{M^2 + m^2} \quad \textit{and} \quad \|I - \alpha_{opt}BA^{-1}B^*\| = \frac{M^2 - m^2}{M^2 + m^2}.$$

Proof. From the first equation of (1.3) and the first equation of (2.13), we have that

$$a(\mathbf{u} - \mathbf{u}_k, \mathbf{v}) = b(\mathbf{v}, p_{k-1} - p) \qquad \text{for all } \mathbf{v} \in \mathbf{V}. \tag{2.14}$$

The above relation implies

$$\begin{aligned} |\mathbf{u} - \mathbf{u}_k|^2 = a(\mathbf{u} - \mathbf{u}_k, \mathbf{u} - \mathbf{u}_k) &= (BA^{-1}B^*(p_{k-1} - p), (p_{k-1} - p)) \\ &\leq M^2 \|p_{k-1} - p\|^2, \end{aligned}$$

which proves the first part of (i). From the second equation of (1.4) and the second equation of (2.13), we have that

$$p - p_k = p - p_{k-1} + \alpha B(\mathbf{u} - \mathbf{u}_k).$$

Combining with (2.14), we get

$$p - p_k = (I - \alpha BA^{-1}B^*)(p - p_{k-1}), \tag{2.15}$$

which gives the second part of (i). From Lemma 2.1, we have that $(I - \alpha BA^{-1}B^*)$ is a symmetric operator, and for any $p \in$P, $p \neq 0$,

$$1 - \alpha M^2 \leq \frac{((I - \alpha BA^{-1}B^*)p, p)}{\|p\|^2} \leq 1 - \alpha m^2,$$

which justifies part (ii). The rest of the proof follows from (ii). ∎

3. Augmented Lagrangian Uzawa algorithm. The main idea of the augmented Lagrangian method, introduced by Fortin and Glowinski [14], is to use the constraint condition for the variable p and another tuning parameter $\rho > 0$ in order to improve the convergence factor of the Uzawa algorithm. We will consider the approach for abstract Hilbert spaces $\mathbf{V}$ and P and prove sharp convergence estimates for the corresponding Uzawa algorithm.

Let $(\mathbf{u}, p)$ be the solution of the variational problem (1.3). Then, from the second equation of (1.4), we have that

$$(B\mathbf{u}, B\mathbf{v}) = (g, B\mathbf{v}), \quad \mathbf{v} \in \mathbf{V}.$$

Thus, for any $\rho > 0$, $(\mathbf{u}, p)$ is also a solution of

$$\begin{array}{lcl} a(\mathbf{u}, \mathbf{v}) + \rho(B\mathbf{u}, B\mathbf{v}) + b(\mathbf{v}, p) & = & \langle \mathbf{f}, \mathbf{v} \rangle + \rho(g, B\mathbf{v}), \\ b(\mathbf{u}, q) & = & (g, q). \end{array} \tag{3.1}$$

Using the notation

$$a_\rho(\mathbf{u}, \mathbf{v}) := a(\mathbf{u}, \mathbf{v}) + \rho(B\mathbf{u}, B\mathbf{v}) \text{ and } \mathbf{f}_\rho := \mathbf{f} + \rho B^* g,$$

we have that

$$\begin{array}{ll} a_\rho(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) & = \langle \mathbf{f}_\rho, \mathbf{v} \rangle, \quad \mathbf{v} \in \mathbf{V}, \\ b(\mathbf{u}, q) & = (g, q), \quad q \in P. \end{array} \tag{3.2}$$

With the form a_ρ, we associate the linear operator $A_\rho : V \to V^*$,

$$\langle A_\rho \mathbf{u}, \mathbf{v}\rangle = a_\rho(\mathbf{u}, \mathbf{v}) \qquad \text{for all } \mathbf{u}, \mathbf{v} \in \mathbf{V}.$$

Thus, an equivalent form of (3.2) is

$$\begin{aligned} A_\rho \mathbf{u} + B^* p &= \mathbf{f}_\rho, \\ B\mathbf{u} &= g. \end{aligned} \tag{3.3}$$

Since $a_\rho(\cdot,\cdot)$ and $a(\cdot,\cdot)$ give rise to equivalent norms on $\mathbf{V}$, we have that (3.2) (or (3.3)) has a unique solution. Consequently, problems (1.3) and (3.2) are equivalent. In what follows, the Uzawa algorithm applied to (3.2) will be called the augmented Lagrangian Uzawa method (ALUM).

Given a relaxation parameter $\alpha > 0$, the augmented Lagrangian Uzawa algorithm for approximating the solution $(\mathbf{u}, p)$ of (1.3) is as follows.

Algorithm 3.1 (ALUM). *Let p_0 be any approximation for p, and for $k = 1, 2, \ldots,$ construct $(\mathbf{u}_k, p_k)$ by*

$$\begin{aligned} a_\rho(\mathbf{u}_k, \mathbf{v}) &= (\mathbf{f}_\rho, \mathbf{v}) - b(\mathbf{v}, p_{k-1}), \quad \mathbf{v} \in \mathbf{V}, \\ p_k &= p_{k-1} + \alpha(B\mathbf{u}_k - g). \end{aligned}$$

To study the convergence of (3.1), we shall calculate first

$$M_\rho := \sup_{p\in P} \sup_{\mathbf{v}\in\mathbf{V}} \frac{b(\mathbf{v}, p)}{\|p\|(a_\rho(\mathbf{v},\mathbf{v}))^{1/2}} \tag{3.4}$$

and

$$m_\rho := \inf_{p\in P} \sup_{\mathbf{v}\in\mathbf{V}} \frac{b(\mathbf{v}, p)}{\|p\|(a_\rho(\mathbf{v},\mathbf{v}))^{1/2}}. \tag{3.5}$$

Theorem 3.2. *For any $\rho > 0$, we have*

$$BA_\rho^{-1}B^* = \left(\rho I + (BA^{-1}B^*)^{-1}\right)^{-1}, \tag{3.6}$$

$$M_\rho^2 = \frac{1}{\rho + \frac{1}{M^2}} \quad \textit{and} \quad m_\rho^2 = \frac{1}{\rho + \frac{1}{m^2}}. \tag{3.7}$$

Proof. To prove (3.6), we need two identities. First, we note that for any invertible linear operator $C : P \to P$ such that $I + \rho C$ is also invertible, we have

$$\left(\rho I + C^{-1}\right)^{-1} = C - \rho C(I + \rho C)^{-1} C. \tag{3.8}$$

This can be proved by checking that the proposed inverse verifies the algebraic definition of the inverse. The second identity is based on the Sherman–Morrison–Woodbury formula and can be proved again just by algebraic manipulations:

$$(A + \rho B^* B)^{-1} = A^{-1} - \rho A^{-1} B^* (I + \rho BA^{-1}B^*)^{-1} BA^{-1}. \tag{3.9}$$

From (3.9), we get

$$\begin{aligned} B(A + \rho B^* B)^{-1} B^* &= BA^{-1}B^* \\ &\quad - \rho BA^{-1}B^*(I + \rho BA^{-1}B^*)^{-1} BA^{-1}B^*. \end{aligned} \tag{3.10}$$

If we take $C = BA^{-1}B^*$ in (3.8) and combine it with (3.10), we obtain (3.6). To verify (3.7), we notice that by applying Lemma 2.1 with a_ρ instead of a we have

$$\sup_{\mathbf{v}\in\mathbf{V}} \frac{b(\mathbf{v},p)^2}{\|p\|^2\ a_\rho(\mathbf{v},\mathbf{v})} = (BA_\rho^{-1}B^*p,p), \quad p \in P. \tag{3.11}$$

Thus, we get

$$\begin{aligned} M_\rho^2 &= \sup_{p\in P} \frac{(BA_\rho^{-1}B^*p,p)}{(p,p)} = \sup_{p\in P} \frac{\left((\rho I + (BA^{-1}B^*)^{-1}\ )^{-1}p,p\right)}{(p,p)} \\ &= \frac{1}{\inf\limits_{q\in P} \frac{((\rho I+(BA^{-1}B^*)^{-1})q,q)}{(q,q)}} = \frac{1}{\rho + \inf\limits_{q\in P} \frac{((BA^{-1}B^*)^{-1}q,q)}{(q,q)}} \\ &= \left(\rho + \frac{1}{\sup\limits_{r\in P} \frac{(BA^{-1}B^*r,r)}{(r,r)}}\right)^{-1} = \left(\rho + \frac{1}{M^2}\right)^{-1}. \end{aligned}$$

Here, we have used the changes of variable $(\rho I + (BA^{-1}B^*)^{-1}\ )^{-1/2}p = q$ and $(BA^{-1}B^*)^{-1/2}q = r$. The proof for m_ρ is similar. ∎

The above result gives formulas for the inf-sup and sup-sup constants for the ALUM in terms of m, M, and ρ. In applications, the constant m is more difficult to obtain. The following theorem gives the convergence rate of the ALUM.

THEOREM 3.3. *Let* $(\mathbf{u}, p)$ *be the solution of* (1.3) *and let* $(\mathbf{u}_k, p_k)$ *be the sequence of approximations built by Algorithm* 3.1. *Then, the following holds true:*

(i) *The sequences* $\mathbf{u} - \mathbf{u}_k$ *and* $p - p_k$ *satisfy*

$$\begin{aligned} a_\rho(\mathbf{u}-\mathbf{u}_k, \mathbf{u}-\mathbf{u}_k)^{1/2} &\le M_\rho\ \|p - p_{k-1}\|, \\ \|p - p_k\| &\le \|I - \alpha BA_\rho^{-1}B^*\|\ \|p - p_{k-1}\|. \end{aligned}$$

(ii) *For* $\alpha < \frac{2}{M_\rho^2}$, *the ALUM is convergent and*

$$\|I - \alpha BA_\rho^{-1}B^*\| = \max\{|1-\alpha m_\rho^2|, |1-\alpha M_\rho^2|\} < 1.$$

(iii) *For* $\alpha = \frac{1}{M_\rho^2}$, *the convergence factor is*

$$\|I - \alpha BA_\rho^{-1}B^*\| = \left(1 - \frac{m^2}{M^2}\right)\frac{1}{m^2\rho + 1}.$$

(iv) *The optimal convergence factor is achieved for* $\alpha_{opt} = \frac{2}{M_\rho^2 + m_\rho^2}$ *and*

$$\|I - \alpha_{opt}BA_\rho^{-1}B^*\| = \frac{M_\rho^2 - m_\rho^2}{M_\rho^2 + m_\rho^2} = \left(1 - \frac{m^2}{M^2}\right)\frac{1}{2m^2\rho + 1 + m^2/M^2}.$$

Proof. The result is a direct consequence of Theorem 2.5 and (3.7). ∎

A similar result for the discrete version of the Stokes system can be found in [19]. As it was pointed out in [14] and [19], the choice of a very large ρ improves on the rate of convergence of the ALUM, but at the same time, the operator A_ρ becomes more difficult to invert. For the continuous and discrete Stokes system, estimates for the convergence factor of the ALUM were recently obtained by Nochetto and Pyo in [16]. The question raised in [16] on how much we can improve the rate of convergence of the ALUM if information about the spectral value m is available can be easily answered now by comparing part (iii) and part (iv) of Theorem 3.3 or Theorem 7.1.

4. Inexact Uzawa method. Throughout the rest of the paper we will keep the notation and assumptions of section 2. In this section, following the ideas in [7, 13], we shall introduce and investigate the convergence of an abstract inexact Uzawa algorithm where the exact solve of the elliptic problem (the action of A^{-1}) is replaced by an approximation process, which might not be a linear operator. We describe the approximate inverse of A as a map $C : \mathbf{V}^* \to \mathbf{V}$ which, for $\phi \in \mathbf{V}^*$, returns an approximation of $\xi = A^{-1}\phi$ such that

$$|C\phi - A^{-1}\phi|_{\mathbf{V}} \le \delta \|\phi\|_{\mathbf{V}^*} \qquad \text{for all } \phi \in \mathbf{V}^* \tag{4.1}$$

for some $\delta \in (0,1)$. We notice here that (4.1) is a strong condition for the infinite-dimensional case. The condition can be weakened by requiring to be satisfied only for certain values $\phi \in \mathbf{V}^*$. If $\mathbf{V}$ and P are finite-dimensional spaces, then C can be considered as a linear or nonlinear process for inverting A and (4.1) is a reasonable assumption (see [7]). One example of nonlinear process C is the approximate inverse associated with the preconditioned conjugate gradient algorithm. A practical case would be to consider $C\phi = \xi_{num}$, where ξ_{num} is the numerical approximation of ξ defined by

$$a(\xi, \mathbf{v}) = \langle \phi, \mathbf{v} \rangle \qquad \text{for all } v \in \mathbf{V}.$$

In any case, if $A\xi = \phi$ and $C\phi$ is defined by $C\phi = \xi_{ap}$, an approximation of ξ, then, according to (2.7), the assumption (4.1) is equivalent to

$$|\xi_{ap} - \xi|_{\mathbf{V}} \le \delta |\xi|_{\mathbf{V}} \qquad \text{for all } \xi \in \mathbf{V}. \tag{4.2}$$

The inexact Uzawa algorithm for approximating the solution $(\mathbf{u}, p)$ of (1.3) is as follows.

Algorithm 4.1 (inexact Uzawa method (IUM)). *Let $(\mathbf{u}_0, p_0)$ be any approximation for $(\mathbf{u}, p)$, and for $k = 1, 2, \ldots,$ construct $(\mathbf{u}_k, p_k)$ by*

$$\begin{aligned} \mathbf{u}_k &= \mathbf{u}_{k-1} + C(\mathbf{f} - A\mathbf{u}_{k-1} - B^* p_{k-1}), \\ p_k &= p_{k-1} + \alpha(B\mathbf{u}_k - g). \end{aligned}$$

Before we study the stability and convergence rate of Algorithm 4.1 we shall introduce the following notation. For $k = 0, 1, \ldots,$ let $e_k^{\mathbf{u}} = \mathbf{u} - \mathbf{u}_k$, $e_k^p = p - p_k$, and

$$E_k = \begin{pmatrix} |e_k^{\mathbf{u}}| \\ \|e_k^p\| \end{pmatrix}.$$

Let

$$\mathbf{M} := \begin{pmatrix} \delta & M(1+\delta) \\ \alpha M \delta & \gamma + \alpha M^2 \delta \end{pmatrix},$$

where $\gamma := \|I - \alpha B A^{-1} B^*\| = \max\{|1 - \alpha m^2|, |1 - \alpha M^2|\}$. On $\mathbf{R}^2$ we introduce the inner product $[\cdot, \cdot]_w$ defined by

$$\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right]_w = w_1 x_1 y_1 + w_2 x_2 y_2,$$

where w_1, w_2 are any two positive numbers such that

$$\frac{w_1}{w_2} = \frac{\alpha\delta}{1+\delta},$$

and δ is a positive number such that (4.1) is satisfied. We note that $\mathbf{M}$ is symmetric with respect to the $[\cdot,\cdot]_w$ inner product. We will denote the norm induced by $[\cdot,\cdot]_w$ with $\|\cdot\|_w$.

THEOREM 4.2. *Let $0 < \alpha < 2/M^2$ and assume that C satisfies* (4.1) *with*

$$\delta < \frac{1-\gamma}{1-\gamma+2\alpha M^2}. \tag{4.3}$$

Then, the IUM converges. If r is the spectral radius of the matrix $\mathbf{M}$, then $0 < r < 1$ and

$$\|E_k\|_w \le\ r^k\ \|E_0\|_w, \quad k = 1, 2, \ldots . \tag{4.4}$$

Proof. We follow the proof of a similar result in [7] for the finite-dimensional case. From the first equation of (1.4) and the first equation of Algorithm 4.1, we have

$$\begin{aligned} e_k^{\mathbf{u}} &= e_{k-1}^{\mathbf{u}} - C(Ae_{k-1}^{\mathbf{u}} + B^* e_{k-1}^p) \\ &= (A^{-1} - C)(Ae_{k-1}^{\mathbf{u}} + B^* e_{k-1}^p) - A^{-1}B^* e_{k-1}^p. \end{aligned} \tag{4.5}$$

From the second equation of (1.4) and the second equation of Algorithm 4.1, we get

$$e_k^p = e_{k-1}^p + \alpha B e_k^{\mathbf{u}}. \tag{4.6}$$

If we substitute $e_k^{\mathbf{u}}$ from (4.5) into (4.6), then

$$e_k^p = (I - \alpha BA^{-1}B^*)e_{k-1}^p + \alpha B(A^{-1} - C)(Ae_{k-1}^{\mathbf{u}} + B^* e_{k-1}^p). \tag{4.7}$$

From (4.5) and (4.7), by the triangle inequality, and from the estimates (2.9) and (2.10) and the assumption (4.1), we obtain

$$|e_k^{\mathbf{u}}| \le \delta\ |e_{k-1}^{\mathbf{u}}| + M(1+\delta)\ \|e_{k-1}^p\|$$

and

$$\|e_k^p\| \le \alpha M \delta\ |e_{k-1}^{\mathbf{u}}| + (\gamma + \alpha M^2 \delta)\|e_{k-1}^p\|.$$

Using the notation introduced above, we have

$$E_k \le \mathbf{M}\ E_{k-1}, \tag{4.8}$$

where

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \le \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

means $x_1 \le y_1$ and $x_2 \le y_2$. From (4.8), we deduce

$$E_k \le \mathbf{M}^k\ E_0. \tag{4.9}$$

Since $\mathbf{M}$ is symmetric with respect to $[\cdot,\cdot]_w$-inner product, we have

$$\|E_k\|_w^2 = [E_k, E_k]_w \le [\mathbf{M}^k\ E_0, \mathbf{M}^k\ E_0]_w = [\mathbf{M}^{2k}\ E_0, E_0]_w \le r^{2k}\ \|E_0\|_w^2,$$

which proves (4.2). To complete the proof, we have to show that $r \in (0,1)$, provided that $0 < \alpha < 2/M^2$ and (4.3) holds. The characteristic equation of the matrix $\mathbf{M}$ is

$$\lambda^2 - \lambda(\delta + \gamma + \alpha M^2 \delta) + \delta(\gamma - \alpha M^2) = 0.$$

Since $\mathbf{M}$ has positive entries, the characteristic equation has real roots and the largest (positive) root agrees with the spectral radius of $\mathbf{M}$. Consequently,

$$r = \frac{1}{2}\left(\delta + \gamma + \alpha M^2\delta + \sqrt{(\delta + \gamma + \alpha M^2\delta)^2 - 4\delta(\gamma - \alpha M^2)}\right).$$

Using that $\gamma = \max\{|1 - \alpha m^2|, |1 - \alpha M^2|\}$ and $\alpha \in (0, 2/M^2)$, it is easy to verify that the function $\delta \to r = r(\delta)$ is an increasing function on $(0, 1)$ and that $r = 1$ for

$$\delta = \delta_0 := \frac{1 - \gamma}{1 - \gamma + 2\alpha M^2}. \tag{4.10}$$

This completes the proof of the theorem. $\square$

Remark 4.3. For $0 < \alpha \leq \frac{2}{M^2+m^2}$ we have that $\gamma = 1 - \alpha m^2$ and the threshold δ_0 becomes

$$\delta_0 = \frac{m^2}{m^2 + 2M^2},$$

which is independent of α. For $\frac{2}{M^2+m^2} \leq \alpha < \frac{2}{M^2}$ we have that $\gamma = \alpha M^2 - 1$ and the threshold δ_0 becomes

$$\delta_0 = \frac{2 - \alpha M^2}{2 + \alpha M^2}.$$

Nevertheless, the optimal (maximal) value of δ_0 as the function of $\alpha \in [\frac{2}{M^2+m^2}, \frac{2}{M^2})$ is $\delta_0 = \frac{m^2}{m^2+2M^2}$ and is achieved for $\alpha = \frac{2}{M^2+m^2}$. Thus, a good choice for α (independent of m) is $\alpha = 1/M^2$. In this case we still have $\delta_0 = \frac{m^2}{m^2+2M^2}$.

Remark 4.4. We can apply the IUM for the augmented Lagrangian formulation. The only changes in Algorithm 4.1 is that A is replaced by A_ρ and $\mathbf{f}$ is replaced by $\mathbf{f}_\rho$. The convergence analysis follows from Theorem 4.2. Let us further notice that in this case $|e_k^{\mathbf{u}}|^2 = a_\rho(e_k^{\mathbf{u}}, e_k^{\mathbf{u}})$ and for $\alpha = 1/M_\rho^2$ the threshold δ_0 which assures convergence for the IUM is

$$\delta_0(\rho) = \frac{m_\rho^2}{m_\rho^2 + 2M_\rho^2} = \frac{m^2 + \rho m^2 M^2}{m^2 + 2M^2 + 3\rho m^2 M^2} \to \frac{1}{3} \quad \text{as} \quad \rho \to \infty.$$

Thus, if the IUM for the augmented Lagrangian formulation is applied with sufficiently large ρ, $\alpha = 1/M_\rho^2$ and with the approximation operator C satisfying

$$\|C - A_\rho^{-1}\| \leq \delta_0(\rho) < 1/3,$$

then the method converges.

Remark 4.5. A different approach in analyzing the IUM in the finite-dimensional case is presented by Cheng in [11]. From his analysis for $\alpha = 1$ and $M = 1$, it follows that the IUM converges (with a different estimate for the convergence factor), under the weaker assumption that $\delta < \delta_0 = 1/3$. Cheng's result for the infinite-dimensional case seems not to have been investigated. A positive answer for this problem would be an interesting result, since in practice it is difficult to estimate the spectral value m.

5. Discretization with the inf-sup condition. In this section we assume that the variational form of a PDE (or system of PDEs) leads to (1.3) and let $\mathbf{V}_h$ and P_h be two finite-dimensional spaces, $\mathbf{V}_h \subset \mathbf{V}$, $P_h \subset P$, with good approximation properties. We further assume that

$$\inf_{p \in P_h} \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{b(\mathbf{v}, p)}{\|p\| \, |\mathbf{v}|} = m(h) > 0 \text{ and } \sup_{p \in P_h} \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{b(\mathbf{v}, p)}{\|p\| \, |\mathbf{v}|} = M(h). \tag{5.1}$$

For an overview of numerical methods for solving saddle point systems, we refer the reader to the recently published review paper [5] by Benzi, Golub, and Liesen.

From Lemma 2.1 and Remark 2.2 we see that $m(h), M(h)$ are the lowest and the largest eigenvalues of the Schur complement $B_h A_h^{-1} B_h^*$ associated with the discrete spaces $\mathbf{V}_h$ and P_h. Then (see, e.g., [9, 15]), the discrete variational problem

Find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times P_h$ such that

$$\begin{array}{lclll} a(\mathbf{u}_h, \mathbf{v}) & + & b(\mathbf{v}, p_h) & = \langle \mathbf{f}, \mathbf{v} \rangle, & \mathbf{v} \in \mathbf{V}_h, \\ b(\mathbf{u}_h, q) & & & = (g, q), & q \in P_h, \end{array} \tag{5.2}$$

has a unique solution and

$$|\mathbf{u} - \mathbf{u}_h| + \|p - p_h\| \le c \left(\inf_{\mathbf{v} \in \mathbf{V}_h} |\mathbf{u} - \mathbf{v}| + \inf_{q \in P_h} \|p - q\| \right),$$

where c is a constant depending only on $m(h)$ and $M(h)$. In this case, the exact or inexact Uzawa algorithms can be applied for the discrete variational problem (5.2) on $\mathbf{V}_h \times P_h$; see, e.g., [7]. The convergence factors depend on $m(h)$ and $M(h)$ and could deteriorate as $h \to 0$ if the pair $(\mathbf{V}_h, P_h)$ is not stable. We recall here that a pair $(\mathbf{V}_h, P_h)$, or more precisely a family of pairs $\{(\mathbf{V}_h, P_h)\}_h$, is called stable if $m(h)$ defined in (5.1) satisfies

$$m(h) \ge m > 0,$$

with m independent of h. In the next section we use the inexact Uzawa algorithm at the continuous level to construct algorithms which avoid building stable pairs $(\mathbf{V}_h, P_h)$.

6. Modified inexact Uzawa method. Eliminating the discrete inf-sup condition. We shall apply the IUM to construct discrete approximations $(\mathbf{u}_k, p_k) \in (\mathbf{V}_k, P_k)$, where $\mathbf{V}_k \subset \mathbf{V}$ and $P_k \subset P$ are finite-dimensional spaces such that the pairs $(\mathbf{V}_k, P_k)$ do not have to be stable pairs.

The algorithm proposed in this section can be used for building multilevel or adaptive methods for solving the system (1.3). Adaptive methods for saddle point problems have been the subject for recent research in numerical analysis (see, e.g., [12, 4]). Our new approach, combined with standard techniques of a posteriori error estimate theory, could lead to new and efficient adaptive algorithms for solving saddle point systems. To describe our new algorithm, we assume that a sequence of nested subspaces,

$$\mathbf{V}_0 \subset \mathbf{V}_1 \subset \mathbf{V}_2 \subset \cdots \subset \mathbf{V},$$

was determined and for $k = 1, 2, \ldots,$ a linear or nonlinear process $C_k : \mathbf{V}^* \to \mathbf{V}_k$ approximating A^{-1} is available such that for a fixed $\phi \in \mathbf{V}^*$, $C_k \phi \in \mathbf{V}_k$ is an approximation of $\xi = A^{-1}\phi$. To construct a good approximate inverse $C_k : \mathbf{V}^* \to \mathbf{V}_k$

one might need to increase the space $\mathbf{V}_{k-1}$ to a space with better approximation properties using an adaptive method. Thus, the embedding assumption $\mathbf{V}_{k-1} \subset \mathbf{V}_k$ is needed. On the other hand, in the proposed algorithms, the variable p is updated at the continuous level and no inversion is used. Thus, the P_k's are just subsets of the space P and do not have to be nested.

The modified inexact Uzawa algorithm for approximating the solution $(\mathbf{u}, p)$ of (1.3) can be stated now as follows.

ALGORITHM 6.1 (modified inexact Uzawa method (MIUM)). *Let $\mathbf{u}_0 \in \mathbf{V}_0$ be any approximation for $\mathbf{u}$ and let $p_0 \in P$ be any approximation for p. For $k = 1, 2, \ldots$, construct $(\mathbf{u}_k, p_k)$, with $\mathbf{u}_k \in \mathbf{V}_k$, by*

$$\begin{aligned} \mathbf{u}_k &= \mathbf{u}_{k-1} + C_k(\mathbf{f} - A\mathbf{u}_{k-1} - B^* p_{k-1}), \\ p_k &= p_{k-1} + \alpha(B\mathbf{u}_k - g). \end{aligned} \tag{6.1}$$

THEOREM 6.2. *Let $0 < \alpha < 2/M^2$, $\gamma = \max\{|1 - \alpha m^2|, |1 - \alpha M^2|\} = \|I - \alpha B A^{-1} B^*\|$, and assume that for $k = 1, 2, \ldots, C_k$ satisfies*

$$\|(C_k - A^{-1})(\mathbf{f} - A\mathbf{u}_{k-1} - B^* p_{k-1})\|_{\mathbf{V}} \le \delta \|(\mathbf{f} - A\mathbf{u}_{k-1} - B^* p_{k-1})\|_{\mathbf{V}^*}, \tag{6.2}$$

with

$$\delta < \frac{1 - \gamma}{1 - \gamma + 2\alpha M^2}. \tag{6.3}$$

Then, the MIUM converges and the convergence rate is given by (4.4).

Proof. It is similar to the proof of Theorem 4.2. ∎

We notice here that, for a fixed α, the threshold δ which assures the convergence of the MIUM depends only on the constants m and M. In the case $g = 0$, we have $p_k \in P_k := B\mathbf{V}_k$. Nevertheless, no matter the choice of the spaces $\mathbf{V}_k, P_k$, the pairs $(\mathbf{V}_k, P_k)$ do not have to be stable pairs.

For the rest of this section, the first equation in (6.1) will be considered in a variational form as follows. Let $\mathbf{d}_k \in \mathbf{V}_k$ be the solution of

$$a(\mathbf{d}_k, \mathbf{v}) = \langle f, \mathbf{v} \rangle - a(\mathbf{u}_{k-1}, v) - b(\mathbf{v}, p_{k-1}), \quad \mathbf{v} \in \mathbf{V}_k. \tag{6.4}$$

Take $\tilde{\mathbf{d}}_k := C_k(\mathbf{f} - A\mathbf{u}_{k-1} - B^* p_{k-1})$ to be an approximation of $\mathbf{d}_k$. For example, $\tilde{\mathbf{d}}_k$ could be a numerical approximation of $\mathbf{d}_k$. Let us assume that $\mathbf{D}_{k-1} \in \mathbf{V}$ is the solution of the continuous problem

$$a(\mathbf{D}_{k-1}, \mathbf{v}) = \langle f, \mathbf{v} \rangle - a(\mathbf{u}_{k-1}, v) - b(\mathbf{v}, p_{k-1}), \quad \mathbf{v} \in \mathbf{V}. \tag{6.5}$$

From the Riesz representation theorem

$$\|(\mathbf{f} - A\mathbf{u}_{k-1} - B^* p_{k-1})\|_{\mathbf{V}^*} = |\mathbf{D}_{k-1}|_{\mathbf{V}}.$$

Thus, the assumption (6.2) can be rewritten as

$$|\tilde{\mathbf{d}}_k - \mathbf{D}_{k-1}|_{\mathbf{V}} \le \delta \; |\mathbf{D}_{k-1}|_{\mathbf{V}}. \tag{6.6}$$

Since $\mathbf{d}_k \in \mathbf{V}_k$ is the Galerkin approximation of $\mathbf{D}_{k-1} \in \mathbf{V}$, we have that $|\mathbf{d}_k|_{\mathbf{V}} \le |\mathbf{D}_{k-1}|_{\mathbf{V}}$. A sufficient condition for the assumption (6.2) is

$$|\tilde{\mathbf{d}}_k - \mathbf{D}_{k-1}|_{\mathbf{V}} \le \delta \; |\mathbf{d}_k|_{\mathbf{V}}. \tag{6.7}$$

6.1. Multilevel exact Uzawa. In this subsection we assume that the problem (6.4) can be solved exactly on $\mathbf{V}_k$, i.e., $\tilde{\mathbf{d}}_k = \mathbf{d}_k$. Then, $\mathbf{u}_k = \mathbf{u}_{k-1} + \mathbf{d}_k$ and consequently,

$$a(\mathbf{u}_k, \mathbf{v}) = \langle f, \mathbf{v} \rangle - b(\mathbf{v}, p_{k-1}), \quad \mathbf{v} \in \mathbf{V}_k.$$

If $\mathbf{U}_{k-1} \in \mathbf{V}$ satisfies

$$a(\mathbf{U}_{k-1}, \mathbf{v}) = \langle f, \mathbf{v} \rangle - b(\mathbf{v}, p_{k-1}), \quad \mathbf{v} \in \mathbf{V},$$

then $\mathbf{D}_{k-1} = \mathbf{U}_{k-1} - \mathbf{u}_{k-1}$ and (6.6) is equivalent to

$$|\mathbf{u}_k - \mathbf{U}_{k-1}|_{\mathbf{V}} \leq \delta \ |\mathbf{u}_{k-1} - \mathbf{U}_{k-1}|_{\mathbf{V}}. \tag{6.8}$$

If $\eta_k > 0$ is a computable estimator for $|\mathbf{u}_k - \mathbf{U}_{k-1}|_{\mathbf{V}}$, i.e.,

$$|\mathbf{u}_k - \mathbf{U}_{k-1}|_{\mathbf{V}} \leq \eta_k, \tag{6.9}$$

then, using (6.7), we get that a sufficient condition for (6.8) is

$$\eta_k \leq \delta \ |\mathbf{u}_k - \mathbf{u}_{k-1}|_{\mathbf{V}}. \tag{6.10}$$

ALGORITHM 6.3 (multilevel exact Uzawa). *Let $p_0 \in P$ be any approximation for p. For $k = 1, 2, \ldots,$ construct $(\mathbf{u}_k, p_k)$, with $\mathbf{u}_k \in \mathbf{V}_k$, by*

$$a(\mathbf{u}_k, \mathbf{v}) = \langle f, \mathbf{v} \rangle - b(\mathbf{v}, p_{k-1}), \quad \mathbf{v} \in \mathbf{V}_k,$$
$$p_k = p_{k-1} + \alpha(B\mathbf{u}_k - g).$$

As a consequence of Theorem 6.2 we have the following.

COROLLARY 6.4. *Let $0 < \alpha < 2/M^2$, $\gamma = \|I - \alpha BA^{-1}B^*\|$, and assume that* (6.8) *or* (6.9)–(6.10) *are satisfied with $\delta < \frac{1-\gamma}{1-\gamma+2\alpha M^2}$. Then, the multilevel exact Uzawa algorithm converges and the convergence rate is given by* (4.4).

6.2. Multilevel inexact Uzawa. In this subsection, we assume that the problem (6.4) can be solved on each $\mathbf{V}_k$ with an absolute error $\epsilon_k \in [0, \delta)$, i.e.,

$$|\mathbf{d}_k - \tilde{\mathbf{d}}_k|_{\mathbf{V}} \leq \epsilon_k |\mathbf{d}_k|_{\mathbf{V}}. \tag{6.11}$$

If $\eta_k > 0$ is a computable estimator for $|\mathbf{d}_k - \mathbf{D}_{k-1}|_{\mathbf{V}}$, i.e.,

$$|\mathbf{d}_k - \mathbf{D}_{k-1}|_{\mathbf{V}} \leq \eta_k, \tag{6.12}$$

then a computable sufficient condition for (6.6) is

$$\eta_k \leq \frac{\delta - \epsilon_k}{1 + \epsilon_k} \ |\tilde{\mathbf{d}}_k|_{\mathbf{V}}. \tag{6.13}$$

Indeed, from (6.7) and (6.11)–(6.13) and the triangle inequality we have

$$\begin{aligned} |\tilde{\mathbf{d}}_k - \mathbf{D}_{k-1}|_{\mathbf{V}} &\leq |\mathbf{d}_k - \mathbf{D}_{k-1}|_{\mathbf{V}} + |\mathbf{d}_k - \tilde{\mathbf{d}}_k|_{\mathbf{V}} \\ &\leq \eta_k + \epsilon_k |\mathbf{d}_k|_{\mathbf{V}} \leq \frac{\delta_k - \epsilon_k}{1 + \epsilon_k} \ |\tilde{\mathbf{d}}_k|_{\mathbf{V}} + \epsilon_k |\mathbf{d}_k|_{\mathbf{V}} \\ &\leq \frac{\delta_k - \epsilon_k}{1 + \epsilon_k} (1 + \epsilon_k) |\mathbf{d}_k|_{\mathbf{V}} + \epsilon_k |\mathbf{d}_k|_{\mathbf{V}} = \delta_k |\mathbf{d}_k|_{\mathbf{V}} \leq \delta_k \ |\mathbf{D}_{k-1}|_{\mathbf{V}}. \end{aligned}$$

We conclude this subsection with a corollary and some remarks.

COROLLARY 6.5. *Let $0 < \alpha < 2/M^2$, $\gamma = \|I - \alpha BA^{-1}B^*\|$, and let $\mathbf{d}_k, \tilde{\mathbf{d}}_k$ satisfy* (6.11)–(6.13) *with $\delta < \frac{1-\gamma}{1-\gamma+2\alpha M^2}$. Then, the MIUM converges and the convergence rate is given by* (4.4).

6.3. Multilevel and adaptive interpretation of the inexact Uzawa algorithm. We note that the modified inexact Uzawa algorithm can be interpreted as a multilevel algorithm. We consider that a sequence $\{\mathbf{M}_k\}$ of approximating subspaces of $\mathbf{V}$ is constructed such that $\mathbf{M}_k$ is strictly larger than $\mathbf{M}_{k-1}$ and that $\mathbf{M}_k$ is built from $\mathbf{M}_{k-1}$ by a uniform refinement strategy (see, e.g., [3, 6, 8, 19]). Based on this existing sequence of nested subspaces of $\mathbf{V}$, we can now build a sequence $\{\mathbf{V}_k\}$ so that (6.6) holds as follows.

Take $\mathbf{V}_0 = \mathbf{M}_0$, and for any positive integer k, assuming that $\mathbf{V}_{k-1} = \mathbf{M}_j$ is known, define $\mathbf{V}_{k+i} := \mathbf{M}_j$ for $i = 0, 1, \ldots$ as long as (6.13) is satisfied for k replaced by $k+i$. In other words, we update $\mathbf{u}_{k-1}$ without enlarging the space $\mathbf{V}_{k-1}$ as long as (6.13) is satisfied. When (6.13) fails to hold, we solve for the $\mathbf{u}_k$ on the next discrete level space.

The modified inexact Uzawa algorithm can be also interpreted as an adaptive method. We construct the sequence $\{\mathbf{V}_k\}$ (so that (6.6) holds) by starting with a subspace $\mathbf{V}_0$ of $\mathbf{V}$ with good approximation properties and by building the sequence $\{\mathbf{V}_k\}_{k\geq 1}$ in a similar manner. If (6.13) fails to hold for $\mathbf{V}_k = \mathbf{V}_{k-1}$, then the new discrete space $\mathbf{V}_k$ is constructed by using an adaptive strategy which assures that (6.13) and consequently (6.6) hold.

7. Applications to the Stokes system. We consider the stationary Stokes equations

$$\begin{array}{lllll} -\Delta \mathbf{u} & -\nabla p & = \mathbf{f} & \text{in } \Omega, \\ \operatorname{div} \mathbf{u} & & = g & \text{in } \Omega, \end{array} \tag{7.1}$$

with vanishing Dirichlet boundary condition $\mathbf{u} = 0$ on $\partial\Omega$ and g satisfying the constraint

$$\int_\Omega g \; dx = 0.$$

In this section we apply the abstract Uzawa results presented in the previous sections to solve (7.1).

Let $\mathbf{V} := (H_0^1(\Omega))^d$, $d = 2$ or $= 3$, and

$$P = L_0^2(\Omega) := \left\{ h \in L^2(\Omega) | \int_\Omega h \; dx = 0 \right\}.$$

We assume that $\mathbf{f} \in (L^2(\Omega))^d$ and $g \in L^2(\Omega)$. The variational formulation of (7.1) becomes

Find $\mathbf{u} \in \mathbf{V}, p \in P$ such that

$$\begin{array}{llll} (\nabla \mathbf{u}, \nabla \mathbf{v}) & +(\operatorname{div} \mathbf{v}, p) & = (\mathbf{f}, \mathbf{v}), & \mathbf{v} \in \mathbf{V}. \\ (\operatorname{div} \mathbf{u}, q) & & = (g, q), & q \in P, \end{array} \tag{7.2}$$

where $(\cdot, \cdot)$ represents the standard L^2-inner product. We will denote by $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ the bilinear forms

$$a(\mathbf{u}, \mathbf{v}) := (\nabla \mathbf{u}, \nabla \mathbf{v}) = \sum_{i=1}^d (\nabla \mathbf{u}_i, \nabla \mathbf{v}_i)$$

and

$$b(\mathbf{v}, p) := (\operatorname{div} \mathbf{v}, p), \quad \mathbf{v} \in \mathbf{V}, p \in P.$$

We note that, for Ω smooth enough, we have

$$a(\mathbf{u},\mathbf{v}) := (\nabla\mathbf{u},\nabla\mathbf{v}) = (\operatorname{curl}\mathbf{u},\operatorname{curl}\mathbf{v}) + (\operatorname{div}\mathbf{u},\operatorname{div}\mathbf{v}), \quad \mathbf{u},\mathbf{v}\in\mathbf{V}. \tag{7.3}$$

We denote the norm induced by a with $|\cdot|_{\mathbf{V}}$ or $|\cdot|$. The norm on P is the L^2-standard norm and is simply denoted by $\|\cdot\|$. With the above notation, the variational formulation of (7.1) becomes (1.3).

It is known that for Ω smooth enough, the following LBB condition holds. More precisely, we have

$$\inf_{p\in P}\sup_{\mathbf{v}\in\mathbf{V}} \frac{b(\mathbf{v},p)}{\|p\|\,|\mathbf{v}|} = c_0 > 0. \tag{7.4}$$

On the other hand, from (7.3) we get that

$$\sup_{p\in P}\sup_{\mathbf{v}\in\mathbf{V}} \frac{b(\mathbf{v},p)}{\|p\|\,|\mathbf{v}|} = 1. \tag{7.5}$$

We notice that for the Stokes problem the operator $A:\mathbf{V}\to\mathbf{V}^*$ consists of d copies of $-\Delta : H_0^1(\Omega)\to H^{-1}(\Omega)$, $B\mathbf{v} = \operatorname{div}\mathbf{v}$, $B^*p = -\nabla p$, and for $\rho>0$,

$$a_\rho(\mathbf{u},\mathbf{v}) := a(\mathbf{u},\mathbf{v}) + \rho(\operatorname{div}\mathbf{u},\operatorname{div}\mathbf{v}), \quad \mathbf{u},\mathbf{v}\in\mathbf{V}.$$

The next two theorems are direct consequences of Theorems 2.5 and 3.3, respectively.

THEOREM 7.1. *Let* $(\mathbf{u},p)$ *be the solution of* (7.2) *and let* $(\mathbf{u}_k,p_k)$ *be the sequence of approximations built by the UM* (2.13). *Then the statements* (i)–(iv) *of Theorem* 2.5 *hold with* $m=c_0$ *and* $M=1$.

THEOREM 7.2. *Let* $(\mathbf{u},p)$ *be the solution of* (7.2) *and let* $(\mathbf{u}_k,p_k)$ *be the sequence of approximations built by the ALUM* (3.1). *Then the statements* (i)–(iv) *of Theorem* 3.3 *hold with* $m=c_0$ *and* $M=1$.

According to section 5, both the UM and ALUM can be applied to any discretization of (7.2), provided that $(\mathbf{V_h},P_h)$, with $\mathbf{V_h}\subset\mathbf{V}$ and $P_h\subset P$, is a stable pair. Let us assume that a fixed pair $(\mathbf{V}_h,P_h)$ satisfies the discrete inf-sup and sup-sup conditions with constants $m(h)=c_d>0$ and $M(h)=1$. If $Q_h : P\to P_h$ is the L^2-orthogonal projection, then, with the new spaces, the operators associated with the forms a and b are A_h and B_h, respectively, where $A_h:\mathbf{V}_h\to\mathbf{V}_h^*$ consists of d copies of the discrete Laplacian and $B_h\mathbf{v} = Q_h\operatorname{div}\mathbf{v}$. Thus, the update for the pressure becomes

$$p_k = p_{k-1} + \alpha\, Q_h(\operatorname{div}\mathbf{u}_k - g).$$

The analysis of the discrete versions of the UM and the ALUM can be carried on similarly. The only difference in describing the convergence of the two algorithms for the discrete case is that c_0 in Theorems 7.1 and 7.2 is replaced by c_d.

The inexact Uzawa algorithm can be also applied for the discretization of (7.2) on $(\mathbf{V_h},\mathbf{P_h})$ (see, e.g., [7]). Taking for example $C_h:\mathbf{V}_h^*\to\mathbf{V}_h$ to be a preconditioner for A_h such that (4.1) is satisfied with $\delta<\frac{c_d^2}{2+c_d^2}$, we have that the IUM converges for any $\alpha\in(0,2)$. We can also apply the inexact Uzawa algorithm for the augmented Lagrangian Uzawa formulation on $(\mathbf{V_h},\mathbf{P_h})$ (see Remark 4.4).

According to Corollary 6.5, the MIUM for solving (7.2) can be also applied for any $\delta<c_0^2/(2+c_0^2)$. The main difficulty in doing so is to find the sequence of spaces $\{\mathbf{V}_k\}$ such that (6.6) or (6.13) is satisfied. Residual-type a posteriori estimators η_k (see, e.g., [2], [18]) could be involved in finding the right sequence $\{\mathbf{V}_k\}$. Constructing and testing multilevel or adaptive algorithms for solving the Stokes system based on the MIUM remains a challenging new problem and is a subject for future work.

8. Conclusion. The paper gives a unified analysis approach of various Uzawa-like algorithms for solving continuous or discrete saddle point problems. The convergence condition and the convergence factors depend upon the extreme spectral bounds of the Schur complement $BA^{-1}B^*$ only. To the best of our knowledge, the result concerning the optimal convergence factor of the ALUM for the infinite-dimensional case is new. The analysis of the modified inexact Uzawa algorithm at the continuous level, which was introduced in section 6, gives a general strategy for solving saddle point systems. Our inexact Uzawa algorithm is similar to the algorithm for solving the Stokes system presented in [4]. The differences are in the way the error bounds are imposed (see (6.6), (6.13)) and the way the pressure is updated. Our analysis, combined with standard techniques of a posteriori error estimates, could lead to new and efficient adaptive algorithms for solving saddle point systems. The main difficulty in implementing concrete algorithms based on the MIUM is finding error estimators η_k such that the conditions (6.6) or (6.13) are satisfied. Finding spaces $\{\mathbf{V}_k\}$ such that conditions similar to (6.6) are satisfied will be the focus of our future work.

Acknowledgment. The author would like to thank the two reviewers for the valuable suggestions towards improving this paper.

REFERENCES

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

[2] I. Babuška and A. Miller, *A feedback finite element method with a posteriori error estimations: Part* I. *The finite element method and some basic properties of the a posteriori error estimator*, Comput. Methods Appl. Mech. Engrg., 61 (1987), pp. 1–40.

[3] C. Bacuta, J. H. Bramble, and J. Pasciak, *New interpolation results and applications to finite element methods for elliptic boundary value problems*, Numer. Linear Algebra Appl., 10 (2003), pp. 33–64.

[4] E. Bänsch, P. Morin, and R. H. Nocheto, *An adaptive Uzawa FEM for the Stokes problem: Convergence without the inf-sup condition*, SIAM J. Numer. Anal., 40 (2002), pp. 1207–1229.

[5] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solutions of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137.

[6] J. H Bramble, J. E. Pasciak, and P. S. Vassilevski, *Computational scales of Sobolev norms with application to preconditioning*, Math. Comp., 69 (2000), pp. 463–480.

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

[8] J. H. Bramble and X. Zhang, *The analysis of multigrid methods*, in Handbook for Numerical Analysis, P. G. Ciarlet and J .L. Lions, eds., Vol. VII, North–Holland, Amsterdam, 2000, pp. 173–415.

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

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

[11] X.-L. Cheng, *On the nonlinear inexact Uzawa algorithm for saddle-point problems*, SIAM J. Numer. Anal., 37 (2000), pp. 1930–1934.

[12] S. Dahlke, W. Dahmen, and K. Urban, *Adaptive wavelet methods for saddle point problems–optimal convergence rates*, SIAM J. Numer. Anal., 40 (2002), pp. 1230–1262.

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

[14] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solutions of Boundary Value Problems*, Stud. Math. Appl. 15, North–Holland, Amsterdam, 1983.

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

[16] R. H. Nochetto and J. Pyo, *Optimal relaxation parameter for the Uzawa method*, Numer.

Math., 98 (2004), pp. 695–702.
[17] R. Temam, *Navier-Stokes Equations*, North–Holland, Amsterdam, 1984.
[18] R. Verfurth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, Chichester, 1996.
[19] J. Xu, *Multilevel Finite Element Theory*, manuscript.
[20] K. Yosida *Functional Analysis*, Springer-Verlag, Berlin, Heidelberg, 1978.