J. Appl. Math. & Computing Vol. **21***(2006), No. 1 - 2, pp. 1 - 21*
Website: http://jamc.net

NONLINEAR UZAWA METHODS FOR SOLVING NONSYMMETRIC SADDLE POINT PROBLEMS

YANHUA CAO, YIQIN LIN* AND YIMIN WEI

Dedicated to Prof. Zhi-Hao Cao on the occasion of his 70th Birthday

Abstract. In [A new nonlinear Uzawa algorithm for generalized saddle point problems, Appl. Math. Comput., 175(2006), 1432-1454], a nonlinear Uzawa algorithm for solving symmetric saddle point problems iteratively, which was defined by two nonlinear approximate inverses, was considered. In this paper, we extend it to the nonsymmetric case. For the nonsymmetric case, its convergence result is deduced. Moreover, we compare the convergence rates of three nonlinear Uzawa methods and show that our method is more efficient than other nonlinear Uzawa methods in some cases. The results of numerical experiments are presented when we apply them to Navier-Stokes equations discretized by mixed finite elements.

AMS Mathematics Subject Classification : 65F10; 65F15
Key words and phrases : Nonsymmetric saddle point problem; Uzawa algorithm; Schur complement; Navier-Stokes equation; Mixed finite element; Stabilization; GMRES; Preconditioning

1. Introduction

Let H_1 and H_2 be finite-dimensional Hilbert spaces with inner product denoted by $(\cdot,\cdot)$. In this paper we consider the abstract nonsymmetric saddle point problem

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \tag{1}$$

where $f \in H_1$, $g \in H_2$ are given and $x \in H_1$, $y \in H_2$ are unknown. Here $A : H_1 \to H_1$ is assumed to be a linear, nonsymmetric and positive definite operator. $A^T : H_1 \to H_1$ is the adjoint of A with respect to the $(\cdot,\cdot)$-inner

Received November 29, 2005. Revised March 31, 2006. *Corresponding author.
Yiqin Lin is supported by Scientific Research Startup Foundation of Sun Yat-Sen University for Young Teacher under grant 2005-34000-1131042. Yimin Wei is supported by the National Natural Science Foundation of China under grant 10471027 and Shanghai Education Committee.

product. $B : H_1 \to H_2$ is a linear map and $B^T : H_2 \to H_1$ is its adjoint. In addition, $C : H_2 \to H_2$ is linear, symmetric and positive semidefinite. Systems of this form (1) typically result from certain finite element and finite difference discretization of Navier-Stokes equations, Oseen equations, and mixed finite element discretization of second order convection-diffusion problems [5, 8, 15].

A number of solution methods have been proposed for solving saddle point problems, see the nicely written survey [1] and the references therein. The methods studied in this paper are similar to the Uzawa-type methods, see [2, 4, 9, 10].

Bramble et al. [2] have considered the linear system (1) with $C = 0$ and assumed that A is positive definite and the LBB condition [3, 7] holds for the symmetric part of A, i.e., for some positive number c_0,

$$(B(A_S)^{-1}B^T w, w) \equiv \sup_{v \in H_1} \frac{(w, Bv)^2}{(A_S v, v)} \geq c_0 \|w\|^2, \quad \forall \ w \in H_2, \tag{2}$$

where A_S is the symmetric part of A, i.e.,

$$A_S = \frac{1}{2}(A + A^T). \tag{3}$$

Since A is positive definite, A_S is symmetric positive definite. A nonlinear Uzawa algorithm has been proposed by defining the nonlinear approximate inverse of A as a map $\Phi : H_1 \to H_1$, i.e., for any $v \in H_1$, $\Phi(v)$ is an approximation to the solution ξ of $A\xi = v$. They have given the convergence results when the nonlinear Uzawa algorithm is used for iteratively solving the system (1) with $C = 0$.

Cao [4] has considered the iterative solution of linear system (1) which, in the mixed finite element application of Navier-Stokes equations, corresponds to the stabilized case, i.e., the LBB condition (2) is replaced by the following stabilized condition [3, 7, 13, 16]

$$((B(A_S)^{-1}B^T + C)w, w) \geq c_0 \|w\|^2, \quad \forall \ w \in H_2, \tag{4}$$

for some positive number c_0. Another nonlinear Uzawa algorithm has been proposed by defining the nonlinear approximate inverse of approximate Schur complement: $(B\hat{A}^{-1}B^T + C)$ as a map $\Psi : H_2 \to H_2$, i.e., for any $w \in H_2$, $\Psi(w)$ is an approximation to the solution ζ of $(B\hat{A}^{-1}B^T + C)\zeta = w$, where $\hat{A} : H_1 \to H_1$ is linear, nonsymmetric and positive definite. The convergence results have been given when the nonlinear Uzawa algorithm is used for iteratively solving the system (1).

For symmetric saddle point problems, Lin and Cao [9] have proposed a new nonlinear Uzawa method, which was defined by these two nonlinear approximate inverses. In this paper, we extend it to the nonsymmetric case and give out some results on the convergence.

The Uzawa-type algorithms are of interest because they are simple, efficient and have minimal numerical computer memory requirements. This could be important in large-scale scientific applications implemented for today's computing architectures. Therefore, the Uzawa methods are widely used in the engineer community. In the remainder of this paper, a subscript S will be used to denote the symmetric part of various operators as in (3).

The remainder of the paper is organized as follows. In Section 2, we review two nonlinear Uzawa algorithms for iteratively solving the system (1) and their convergence results. In Section 3, we give our new nonlinear Uzawa algorithm and its convengence results. In Section 4, we compare the convergence rates of these nonlinear Uzawa methods. We apply these nonlinear Uzawa algorithms to solve (1) which results from discretization of Navier-Stokes equations by mixed finite elements and the results of the numerical experiments are presented in Section 5. Finally, conclusion is given in Section 6.

2. The nonlinear Uzawa algorithms and their convergence results

Let H be a finite-dimensional Hilbert space and $Q : H \to H$ be a linear, symmetric and positive operator. We define an inner product

$$(v,u)_Q = (Qv,u), \qquad \forall\ v,u \in H.$$

Let $\|\cdot\|_Q$ denote the induced norm

$$\|v\|_Q = (v,v)_Q^{1/2}.$$

Since A in (1) is assumed to be positive definite, its symmetric part A_S is symmetric positive definite,

$$(A_S u,u) \geq \lambda_m \|u\|^2, \qquad \forall\ u \in H_1,$$

where $\lambda_m = \min_{\lambda\in\sigma(A_S)} \lambda$. Thus

$$(Au,v) \leq \alpha (A_S u,u)^{1/2}(A_S v,v)^{1/2}, \qquad \forall\ u,v, \in H_1, \tag{5}$$

for all α, provided $\alpha \geq \|A\|/\lambda_m$. Bramble *et al.* have showed that (cf. Reference [2, Lemma 2.1]) $(A^{-1})_S$ is positive definite and satisfies

$$((A^{-1})_S v,v) \leq ((A_S)^{-1}v,v) \leq \alpha^2((A^{-1})_S v,v), \qquad \forall\ v \in H_1. \tag{6}$$

Let Q_B be a symmetric positive definite linear operator satisfying

$$((B(A_S)^{-1}B^T + C)w,w) \leq (Q_B w,w), \qquad \forall\ w \in H_2. \tag{7}$$

Let

$$\gamma_{opt} = \lambda_{min}(Q_B^{-1}(B(A_S)^{-1}B^T + C)).$$

Here $\lambda_{min}(G)$ denotes the smallest eigenvalue of operator G, then (4) and (7) imply that $\gamma_{opt} \in (0,1]$. We deduce from (7) that (cf.(4))

$$\gamma(Q_B w,w) \leq ((B(A_S)^{-1}B^T + C)w,w), \qquad \forall\ w \in H_2 \tag{8}$$

where $\gamma \in (0,\gamma_{opt}]$.

The first nonlinear inexact Uzawa method is defined as follows [2].

Algorithm 1. The first nonlinear inexact Uzawa method
For $x_0 \in H_1$ and $y_0 \in H_2$ given, the iterative sequence $\{(x_i, y_i)\}$ is defined, for $i = 0, 1, \cdots$, by

$$\begin{aligned} x_{i+1} &= x_i + \Phi(f - Ax_i - B^T y_i), \\ y_{i+1} &= y_i + \tau Q_B^{-1}(Bx_{i+1} - Cy_i - g). \end{aligned}$$

It is assumed that for some positive $\delta < 1$,

$$\|\Phi(v) - A^{-1}v\|_{A_S} \le \delta \|A^{-1}v\|_{A_S} \equiv \delta \|v\|_{(A^{-1})_S}, \qquad \forall\, v \in H_1. \tag{9}$$

Theorem 1. *Suppose that A is positive definite with its symmetric part A_S satisfying (5), and Q_B is a symmetric positive definite operator satisfying (7) and (8). Assume that (9) holds and that τ is chosen such that*

$$0 < \tau \le \frac{\gamma}{\alpha^2}.$$

Set

$$\theta = (1 - \frac{\tau\gamma}{\alpha^2})^{1/2}.$$

Then Algorithm 1 converges if

$$\delta < \frac{1-\theta}{1 + 2\tau - \theta}.$$

Moreover, the iteration errors e_i^x and e_i^y satisfy

$$\frac{\delta\tau}{1+\delta}\|e_{i+1}^x\|_{A_S}^2 + \|e_{i+1}^y\|_{Q_B}^2 \le \rho^{2(i+1)} \left(\frac{\delta\tau}{1+\delta}\|e_0^x\|_{A_S}^2 + \|e_0^y\|_{Q_B}^2 \right)$$

and

$$\|e_{i+1}^x\|_{A_S}^2 \le \tau^{-1}(1+\delta)(1+2\delta)\rho^{2i} \left(\frac{\delta\tau}{1+\delta}\|e_0^x\|_{A_S}^2 + \|e_0^y\|_{Q_B}^2 \right),$$

where

$$\rho = \frac{(1+\tau)\delta + \theta + \sqrt{((1+\tau)\delta + \theta)^2 + 4\delta(\tau - \theta)}}{2} < 1.$$

In order to accelerate the convergence of the Uzawa algorithm, the second nonlinear Uzawa method is defined as follows [4].

Algorithm 2. The second nonlinear inexact Uzawa method
For $x_0 \in H_1$ and $y_0 \in H_2$ given, the iterative sequence $\{(x_i, y_i)\}$ is defined, for $i = 0, 1, \cdots$, by

$$\begin{aligned} x_{i+1} &= x_i + \hat{A}^{-1}(f - Ax_i - B^T y_i), \\ y_{i+1} &= y_i + \Psi(Bx_{i+1} - Cy_i - g), \end{aligned}$$

where $\hat{A} : H_1 \longmapsto H_1$ is positive definite, i.e., its symmetric part $\hat{A}_S$ is symmetric positive definite. $\Psi : H_2 \longmapsto H_2$ is a map. For $v \in H_2$, $\Psi(v)$ is an approximation to the solution ξ of the approximate Schur complement system

$$(B\hat{A}^{-1}B^T + C)\xi = v.$$

It is assumed that for $v \in H_1$,

$$(\hat{A}_S v, v) \leq (A_S v, v) \tag{10}$$

and

$$\|(I - \hat{A}^{-1}A)v\|_{\hat{A}_S} \leq \delta \|v\|_{\hat{A}_S}. \tag{11}$$

Then it is easy to see that the following inequalities hold (cf. (6))

$$((\hat{A}^{-1})_S v, v) \leq ((\hat{A}_S)^{-1} v, v) \leq \hat{\alpha}^2 ((\hat{A}^{-1})_S v, v), \qquad \forall\ v \in H_1,$$

for some $\hat{\alpha} \geq 1$.

It is also assumed that the nonlinear approximation map Ψ satisfies, for any $w \in H_2$,

$$\begin{aligned} &\|\Psi(w) - (B\hat{A}^{-1}B^T + C)^{-1}w\|_{(B\hat{A}^{-1}B^T+C)_S} \\ \leq\ & \varepsilon \|(B\hat{A}^{-1}B^T + C)^{-1}w\|_{(B\hat{A}^{-1}B^T+C)_S}, \end{aligned} \tag{12}$$

for some positive $\varepsilon < 1$.

The result on the convergence of the second nonlinear inexact Uzawa method is given as follows [4].

Theorem 2. *Assume that (10), (11) and (12) hold. Let (x, y) be the solution pair of (1), and $\{(x_i, y_i)\}$ be defined by the sencond nonlinear inexact Uzawa algorithm. Then, x_i and y_i converge to x and y, respectively, if*

$$\delta < \frac{1}{\hat{\alpha}+1} \ \textit{and}\ \varepsilon < \frac{1-\delta(\hat{\alpha}+1)}{1+\delta(\hat{\alpha}-1)}.$$

In this case, the following two inequalities hold

$$\begin{aligned} &\delta(1+\varepsilon)\hat{\alpha}(\hat{A}e_i^x, e_i^x) + ((B\hat{A}^{-1}B^T + C)e_i^y, e_i^y) \\ &\leq \rho^{2i} \left(\ \delta(1+\varepsilon)\hat{\alpha}(\hat{A}e_0^x, e_0^x) + ((B\hat{A}^{-1}B^T + C)e_0^y, e_0^y)\ \right) \end{aligned}$$

and

$$\begin{aligned} (\hat{A}e_i^x, e_i^x) \quad \leq\quad & \left(1 + \frac{\delta}{(1+\varepsilon)\hat{\alpha}}\right) \\ & \cdot \rho^{2i-2} \left(\ \delta(1+\varepsilon)\hat{\alpha}(\hat{A}e_0^x, e_0^x) + ((B\hat{A}^{-1}B^T + C)e_0^y, e_0^y)\ \right), \end{aligned}$$

where

$$\rho = \frac{\delta + \varepsilon + \sqrt{(\delta+\varepsilon)^2 + 4\delta((\hat{\alpha}-1)\varepsilon + \hat{\alpha})}}{2}.$$

3. A new nonlinear Uzawa method and its convergence results

Since A in (1) is assumed to be positive definite, i.e., its symmetric part A_S is symmetric positive definite, A is invertible. Thus, (1) is solve if and only if the Schur complement problem

$$(BA^{-1}B^T + C)y = BA^{-1}f - g \tag{13}$$

is solvable. Since A is positive definite, the following inequality holds

$$(Au, v) \leq \alpha (A_S u, u)^{1/2} (A_S v, v)^{1/2}, \;\; \forall \;\, u, v \in H_1$$

for some number α. Clearly, $\alpha \geq 1$.

In the Lemma 2.1 of [2], Bramble et al. have shown that $(A^{-1})_S$ is positive definite and satisfies

$$((A^{-1})_S v, v) \leq ((A_S)^{-1} v, v) \leq \alpha^2 ((A^{-1})_S v, v), \;\; \forall \;\, v \in H_1. \tag{14}$$

In [4], Cao has pointed out that if A is positive and the stabilized condition (4) holds, then Schur complement problem (13) is solvable, i.e., (1) is solvable.

In Section 2, we review two nonlinear Uzawa algorithms. One (Algorithm 1) is defined by nonlinear approximation of the inverse of A, and the other (Algorithm 2) is defined by nonlinear approximation of the inverse of approximate Schur complement: $B\hat{A}^{-1}B^T + C$. In order to accelerate the convergence, we define a new nonlinear Uzawa algorithm, which is based on both nonlinear approximation of the inverse of A and nonlinear approximation of the inverse of Schur complement: $BA^{-1}B^T + C$.

Algorithm 3. A new nonlinear Uzawa method

`For` $x_0 \in H_1$ `and` $y_0 \in H_2$ `given, the iterative sequence` $\{(x_i, y_i)\}$ `is defined, for` $i = 0, 1, \cdots, by$

$$\begin{aligned} x_{i+1} &= x_i + \Phi(f - Ax_i - B^T y_i), & (15)\\ y_{i+1} &= y_i + \Psi(Bx_{i+1} - Cy_i - g). & (16) \end{aligned}$$

`where` $\Phi : H_1 \to H_1$ `and` $\Psi : H_2 \to H_2$ `are two maps. For` $v \in H_1$`,` $\Phi(v)$ `is an approximation to the solution` ξ `of the system`

$$A\xi = v,$$

`and for` $w \in H_2$`,` $\Psi(w)$ `is an approximation to the solution` ζ `of the system`

$$(BA^{-1}B^T + C)\zeta = w.$$

Remark 1. The nonlinear Uzawa algorithm for solving symmetric saddle point problems iteratively, which was defined by two nonlinear approximate inverses, was considered in [9].

Let

$$S = BA^{-1}B^T + C.$$

We assume that our approximations satisfy

$$\|\Phi(v) - A^{-1}v\|_{A_S} \leq \delta \|A^{-1}v\|_{A_S} \equiv \delta \|v\|_{(A^{-1})_S}, \;\; \forall \; v \in H_1, \tag{17}$$

$$\|\Psi(w) - S^{-1}w\|_{S_S} \leq \varepsilon \|S^{-1}w\|_{S_S} \equiv \varepsilon \|w\|_{(S^{-1})_S}, \;\; \forall \; w \in H_2, \tag{18}$$

for some positive $\delta < 1$ and $\varepsilon < 1$, respectively. Analogous to what was examined in [2], (17) and (18) are two reasonable assumptions which are satisfied by the approximate inverse associated with the GMRES algorithm [12].

First, we give several lemmas, which will be used for proving the following result on the convergence of this new nonlinear Uzawa algorithm.

Lemma 1. *For any $v \in H_1$, we have the following inequality*

$$\|Bv\|_{(S^{-1})_S} \leq \alpha \|v\|_{A_S}.$$

Proof. For any $v \in H_1$ and $w \in H_2$, we can get

$$((S^{-1})_S Bv, w)^2 \leq ((S^{-1})_S Bv, Bv)((S^{-1})_S w, w),$$

with equality if and only if Bv and w are collinear. So,

$$\begin{aligned}\|Bv\|^2_{(S^{-1})_S} &= ((S^{-1})_S Bv, Bv) = \sup_{w \in H_2} \frac{((S^{-1})_S Bv, w)^2}{((S^{-1})_S w, w)} \\ &= \sup_{w \in H_2} \frac{(Bv, w)^2}{([(S^{-1})_S]^{-1} w, w)}.\end{aligned}$$

By (14), we have

$$([(S^{-1})_S]^{-1} w, w) \geq ([(S^{-1})^{-1}]_S w, w) = (S_S w, w).$$

Therefore,

$$\begin{aligned}\|Bv\|^2_{(S^{-1})_S} &\leq \sup_{w \in H_2} \frac{(Bv, w)^2}{(S_S w, w)} = \sup_{w \in H_2} \frac{((A_S)^{1/2} v, (A_S)^{-1/2} B^T w)^2}{(S_S w, w)} \\ &\leq \sup_{w \in H_2} \frac{(A_S v, v)((A_S)^{-1} B^T w, B^T w)}{(S_S w, w)} \\ &\leq \sup_{w \in H_2} \frac{(A_S v, v)((A_S)^{-1} B^T w, B^T w)}{(B(A^{-1})_S B^T w, w)} \\ &= \sup_{w \in H_2} \frac{(A_S v, v)((A_S)^{-1} B^T w, B^T w)}{((A^{-1})_S B^T w, B^T w)} \\ &\leq \alpha^2 (A_S v, v) = \alpha^2 \|v\|^2_{A_S} \quad \text{by (14)}.\end{aligned}$$

□

Lemma 2. *For any $v \in H_1$ and $w \in H_2$, we have the following inequality*

$$\|Av + B^T w\|_{(A^{-1})_S} \leq \|v\|_{A_S} + \|w\|_{S_S}.$$

Proof. It follows from the triangular inequality that

$$\begin{aligned}\|Av + B^T w\|_{(A^{-1})_S} &\leq \|Av\|_{(A^{-1})_S} + \|B^T w\|_{(A^{-1})_S} \\ &= \|v\|_{A_S} + \|w\|_{(BA^{-1}B^T)_S} \\ &\leq \|v\|_{A_S} + \|w\|_{S_S}.\end{aligned}$$

□

We now give the result on the convergence of this new nonlinear Uzawa algorithm, which is one of the main results of the paper.

Theorem 3. *Assume that (17) and (18) hold. Let (x, y) be the solution pair of (1), and $\{(x_i, y_i)\}$ be defined by Algorithm 3. Then, x_i and y_i converge to x and y, respectively, if*

$$0 < \delta < \frac{1}{1+2\alpha} \quad \text{and} \quad 0 < \varepsilon < \frac{1-\delta-2\alpha\delta}{1-\delta+2\alpha\delta}. \tag{19}$$

In this case, the following two inequalities hold

$$\begin{aligned} &\alpha\delta(1+\varepsilon)(A_S e_i^x, e_i^x) + (1+\delta)(S_S e_i^y, e_i^y) \\ &\le \rho^{2i}[\alpha\delta(1+\varepsilon)(A_S e_0^x, e_0^x) + (1+\delta)(S_S e_0^y, e_0^y)] \end{aligned} \tag{20}$$

and

$$\begin{aligned} (A_S e_i^x, e_i^x) \quad \le \quad & \left(1+\delta+\frac{\delta}{\alpha(1+\varepsilon)}\right) \\ & \cdot\rho^{2i-2}\left(\ \alpha\delta(1+\varepsilon)(A_S e_0^x, e_0^x) + (1+\delta)(S_S e_0^y, e_0^y)\ \right), \end{aligned} \tag{21}$$

where

$$\rho = \frac{\delta+\varepsilon+\alpha\delta+\alpha\delta\varepsilon+\sqrt{(\delta+\varepsilon+\alpha\delta+\alpha\delta\varepsilon)^2+4(\alpha\delta+\alpha\delta\varepsilon-\delta\varepsilon)}}{2}. \tag{22}$$

Proof. From (15) and (16), we have the two following equations

$$e_{i+1}^x = e_i^x - \Phi(Ae_i^x + B^T e_i^y), \tag{23}$$

$$e_{i+1}^y = e_i^y - \Psi(Ce_i^y - Be_{i+1}^x). \tag{24}$$

Substituting e_{i+1}^x in the equatinon (24) by the right side of the equation (23) gives

$$\begin{aligned} e_{i+1}^y &= e_i^y - \Psi(Ce_i^y - B(e_i^x - \Phi(Ae_i^x + B^T e_i^y))) \\ &= (S^{-1}-\Psi)(Ce_i^y - Be_i^x + B\Phi(Ae_i^x + B^T e_i^y)) \\ &\quad + e_i^y - S^{-1}(Ce_i^y - Be_i^x + B\Phi(Ae_i^x + B^T e_i^y)) \\ &= (S^{-1}-\Psi)(Ce_i^y - Be_i^x + B\Phi(Ae_i^x + B^T e_i^y)) \\ &\quad + S^{-1}B(A^{-1}-\Phi)(Ae_i^x + B^T e_i^y). \end{aligned}$$

It follows from the triangular inequality that

$$\begin{aligned} \|e_{i+1}^y\|_{S_S} &\le \varepsilon\|Ce_i^y - Be_i^x + B\Phi(Ae_i^x + B^T e_i^y)\|_{(S^{-1})_S} && \text{by (18)} \\ &\quad + \|S^{-1}B(A^{-1}-\Phi)(Ae_i^x + B^T e_i^y)\|_{S_S} \\ &= \varepsilon\|Se_i^y - B(A^{-1}-\Phi)(Ae_i^x + B^T e_i^y)\|_{(S^{-1})_S} \\ &\quad + \|B(A^{-1}-\Phi)(Ae_i^x + B^T e_i^y)\|_{(S^{-1})_S} \\ &\le \varepsilon\|e_i^y\|_{S_S} + \alpha\varepsilon\|(A^{-1}-\Phi)(Ae_i^x + B^T e_i^y)\|_{A_S} \\ &\quad + \alpha\|(A^{-1}-\Phi)(Ae_i^x + B^T e_i^y)\|_{A_S} && \text{by Lemma 1} \\ &\le \varepsilon\|e_i^y\|_{S_S} + \alpha\delta(1+\varepsilon)\|Ae_i^x + B^T e_i^y\|_{(A^{-1})_S} && \text{by (17)} \\ &\le \varepsilon\|e_i^y\|_{S_S} + \alpha\delta(1+\varepsilon)(\|e_i^x\|_{A_S} + \|e_i^y\|_{S_S}) && \text{by Lemma 2} \\ &= \alpha\delta(1+\varepsilon)\|e_i^x\|_{A_S} + (\varepsilon+\alpha\delta+\alpha\delta\varepsilon)\|e_i^y\|_{S_S}. \end{aligned} \tag{25}$$

Using triangular inequality, from the equation (23) we have

$$\begin{aligned}
\|e^x_{i+1}\|_{A_S} &= \|e^x_i - \Phi(Ae^x_i + B^T e^y_i)\|_{A_S} \\
&= \|(A^{-1} - \Phi)(Ae^x_i + B^T e^y_i) - A^{-1}B^T e^y_i\|_{A_S} \\
&\le \delta\|Ae^x_i + B^T e^y_i\|_{(A^{-1})_S} + \|B^T e^y_i\|_{(A^{-1})_S} \qquad \text{by (17)} \\
&\le \delta(\|e^x_i\|_{A_S} + \|e^y_i\|_{S_S}) + \|e^y_i\|_{S_S} \qquad \text{by Lemma 2} \\
&= \delta\|e^x_i\|_{A_S} + (1+\delta)\|e^y_i\|_{S_S}. \qquad (26)
\end{aligned}$$

It follows from (25) and (26) that

$$\begin{pmatrix} \|e^x_i\|_{A_S} \\ \|e^y_i\|_{S_S} \end{pmatrix} \le N^i \begin{pmatrix} \|e^x_0\|_{A_S} \\ \|e^y_0\|_{S_S} \end{pmatrix}, \qquad (27)$$

where N is given by

$$N = \begin{pmatrix} \delta & 1+\delta \\ \alpha\delta(1+\varepsilon) & \varepsilon + \alpha\delta + \alpha\delta\varepsilon \end{pmatrix}.$$

Obviously, N is symmetric with respect to the following inner product of the two-dimensional Euclidean space

$$\left(\begin{pmatrix} \alpha\delta(1+\varepsilon) & 0 \\ 0 & 1+\delta \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = \alpha\delta(1+\varepsilon)x_1x_2 + (1+\delta)y_1y_2.$$

Then, from (27) we have

$$\begin{aligned}
&\alpha\delta(1+\varepsilon)(A_Se^x_i, e^x_i) + (1+\delta)(S_Se^y_i, e^y_i) \\
&= \left(\begin{pmatrix} \alpha\delta(1+\varepsilon) & 0 \\ 0 & 1+\delta \end{pmatrix}\begin{pmatrix} \|e^x_i\|_{A_S} \\ \|e^y_i\|_{S_S} \end{pmatrix}, \begin{pmatrix} \|e^x_i\|_{A_S} \\ \|e^y_i\|_{S_S} \end{pmatrix}\right) \\
&\le \left(\begin{pmatrix} \alpha\delta(1+\varepsilon) & 0 \\ 0 & 1+\delta \end{pmatrix} N^i \begin{pmatrix} \|e^x_0\|_{A_S} \\ \|e^y_0\|_{S_S} \end{pmatrix}, N^i \begin{pmatrix} \|e^x_0\|_{A_S} \\ \|e^y_0\|_{S_S} \end{pmatrix}\right) \\
&\le \rho^{2i}[\alpha\delta(1+\varepsilon)(A_Se^x_0, e^x_0) + (1+\delta)(S_Se^y_0, e^y_0)],
\end{aligned}$$

where ρ is the spectral radius of N. The eigenvalues of N are the roots of

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

Obviously, the spectral radius ρ of N is equal to its positive eigenvalue which is given by (22).

It is easy to see if (19) is satisfied , then $\rho < 1$. This completes the proof of (20).

To prove (21), we apply the following elementary inequality

$$(a+b)^2 \le (1+\eta)a^2 + (1+\eta^{-1})b^2$$

to (26) and get, for any $\eta > 0$,

$$\|e^x_i\|^2_{A_S} \le (1+\eta)\delta^2\|e^x_{i-1}\|^2_{A_S} + (1+\eta^{-1})(1+\delta)^2\|e^y_{i-1}\|^2_{S_S}.$$

Inequality (21) follows from taking $\eta = \frac{\alpha(1+\varepsilon)(1+\delta)}{\delta}$ and applying (20). This completes the proof of the theorem.

□

Remark 2. From Theorem 3, we can know that if $\delta \to 0$ and $\varepsilon \to 0$, ρ, the convergence factor of the new method, $\to 0$. That is to say if two approximate inverses associated with the GMRES algorithm are very approximate to two exact inverses, the new nonlinear Uzawa method will have good convergence property. However, in most cases the Algorithm 6 is impractical in that during applying GMRES algorithm to getting the approximate inverses of Schur complement, too many times iterations may be needed in order to evaluate $A^{-1}u$ very precisely. Thus, we should consider a practical version of the new method, in which the Schur complement system $(BA^{-1}B^T + C)\zeta = w$ is replaced by the approximate Schur complement system $(B\hat{A}^{-1}B^T + C)\zeta = w$. We don't need much work to evaluate $\hat{A}^{-1}u$ during applying GMRES algorithm to getting the approximate inverse of the approximate Schur complement system.

If A is symmetric positive definite, $A_S = A$ and $S_S = S$. At this time, α can be 1. Therefore, we have the following corollary.

Corollary 1. *Assume that A in (1) is symmetric positive definite. And assume that (17) and (18) hold. Let (x, y) be the solution pair of (1), and $\{(x_i, y_i)\}$ be defined by Algorithm 3. Then, x_i and y_i converge to x and y, respectively, if*

$$0 < \delta < \frac{1}{3} \text{ and } 0 < \varepsilon < \frac{1-3\delta}{1+\delta}.$$

In this case, the following two inequalities hold

$$\begin{aligned} &\delta(1+\varepsilon)(Ae_i^x, e_i^x) + (1+\delta)(Se_i^y, e_i^y) \\ &\le \rho^{2i}[\delta(1+\varepsilon)(Ae_0^x, e_0^x) + (1+\delta)(Se_0^y, e_0^y)] \end{aligned}$$

and

$$\begin{aligned} (Ae_i^x, e_i^x) \quad \le \quad & \left(\ 1+\delta+\tfrac{\delta}{1+\varepsilon} \ \right) \\ & \cdot\rho^{2i-2} \left(\ \delta(1+\varepsilon)(Ae_0^x, e_0^x) + (1+\delta)(Se_0^y, e_0^y) \ \right), \end{aligned}$$

where

$$\rho = \frac{\varepsilon + 2\delta + \varepsilon\delta + \sqrt{(\varepsilon + 2\delta + \varepsilon\delta)^2 + 4\delta}}{2}.$$

The practical version of the new nonlinear Uzawa method is given as follows.

Algorithm 4. The practical version of the new nonlinear Uzawa method
`For` $x_0 \in H_1$ `and` $y_0 \in H_2$ `given, the iterative sequence` $\{(x_i, y_i)\}$ `is defined, for` $i = 0, 1, \cdots, by$

$$\begin{aligned} x_{i+1} &= x_i + \Phi(f - Ax_i - B^T y_i), && (28) \\ y_{i+1} &= y_i + \Psi(Bx_{i+1} - Cy_i - g), && (29) \end{aligned}$$

`where` $\Phi : H_1 \to H_1$ `and` $\Psi : H_2 \to H_2$ `are two maps. For` $v \in H_1$`,` $\Phi(v)$ `is an approximation to the solution` ξ `of the system`

$$A\xi = v,$$

and for $w \in H_2$, $\Psi(w)$ is an approximation to the solution ζ of the approximate Schur complement system

$$(B\hat{A}^{-1}B^T + C)\zeta = w,$$

where $\hat{A}$ is a linear positive definite operator.

Let

$$\hat{S} = B\hat{A}^{-1}B^T + C.$$

We assume that our approximations satisfy

$$\|\Phi(v) - A^{-1}v\|_{A_S} \le \delta\|A^{-1}v\|_{A_S} \equiv \delta\|v\|_{(A^{-1})_S}, \ \ \forall\, v \in H_1, \tag{30}$$

$$\|\Psi(w) - \hat{S}^{-1}w\|_{\hat{S}_S} \le \varepsilon\|\hat{S}^{-1}w\|_{\hat{S}_S} \equiv \varepsilon\|w\|_{(\hat{S}^{-1})_S}, \ \ \forall\, w \in H_2, \tag{31}$$

for some positive $\delta < 1$ and some positive $\varepsilon < 1$, respectively.

Since $\hat{A}$ is positive definite, we have the following inequality

$$((\hat{A}^{-1})_S v, v) \le ((\hat{A}_S)^{-1}v, v) \le \hat{\alpha}^2((\hat{A}^{-1})_S v, v), \ \ \forall\, v \in H_1,$$

for some number $\hat{\alpha}$. Clearly, $\hat{\alpha} \ge 1$

We can easily get the following lemma, which is similar with Lemma 1.

Lemma 3. *For any $v \in H_1$, we have the following inequality*

$$\|Bv\|_{(\hat{S}^{-1})_S} \le \hat{\alpha}\|v\|_{\hat{A}_S}.$$

We also assume that

$$\|v\|_{\hat{A}_S} \le \gamma\|v\|_{A_S}, \ \ \forall\, v \in H_1, \tag{32}$$

where $0 < \gamma \le 1$. Then, we easily get by (32)

$$\|v\|_{(A_S)^{-1}} \le \gamma\|v\|_{(\hat{A}_S)^{-1}}, \ \ \forall\, v \in H_1. \tag{33}$$

Lemma 4. *Assume that (32) holds and $0 < \gamma \le 1$. For any $v \in H_1$ and $w \in H_2$, we have the following inequality*

$$\|Av + B^T w\|_{(A^{-1})_S} \le \|v\|_{A_S} + \hat{\alpha}\gamma\|w\|_{\hat{S}_S}.$$

Proof. From (32), the inequality (33) holds. It follows from the triangular inequality that

$$\begin{aligned}
\|Av + B^T w\|_{(A^{-1})_S} &\le \|Av\|_{(A^{-1})_S} + \|B^T w\|_{(A^{-1})_S} \\
&\le \|v\|_{A_S} + \|B^T w\|_{(A_S)^{-1}} \\
&\le \|v\|_{A_S} + \gamma\|B^T w\|_{(\hat{A}_S)^{-1}} \\
&\le \|v\|_{A_S} + \hat{\alpha}\gamma\|B^T w\|_{(\hat{A}^{-1})_S} \\
&\le \|v\|_{A_S} + \hat{\alpha}\gamma\|w\|_{\hat{S}_S}.
\end{aligned}$$

□

We now give the result on the convergenceof this practical version of the new method, which is another of the main results of the paper.

Theorem 4. *Assume that (30), (31) and (32) hold. Let (x, y) be the solution pair of (1), and $\{(x_i, y_i)\}$ be defined by Algorithm 4. Then, x_i and y_i converge to x and y, respectively, if*

$$0 < \delta < \frac{1-\hat{\alpha}\tau}{1-\hat{\alpha}\tau+2\hat{\alpha}^2\gamma^2} \ \textit{and}\ 0 < \varepsilon < \frac{1-\delta-\mu}{1-\delta+\mu}. \tag{34}$$

where $\mu = \hat{\alpha}\tau - \hat{\alpha}\tau\delta + 2\hat{\alpha}^2\gamma^2\delta$. In this case, the following two inequalities hold

$$\begin{aligned} &\delta(1+\varepsilon)(\hat{A}_S e_i^x, e_i^x) + (1+\delta)(\hat{S}_S e_i^y, e_i^y) \\ &\le \rho^{2i}[\delta(1+\varepsilon)(\hat{A}_S e_0^x, e_0^x) + (1+\delta)(\hat{S}_S e_0^y, e_0^y)] \end{aligned} \tag{35}$$

and

$$\begin{aligned} (\hat{A}_S e_i^x, e_i^x) \quad \le \quad & \left(\hat{\alpha}^2\gamma^2(1+\delta) + \frac{\delta}{1+\varepsilon}\right) \\ & \cdot \rho^{2i-2}\left(\ \delta(1+\varepsilon)(\hat{A}_S e_0^x, e_0^x) + (1+\delta)(\hat{S}_S e_0^y, e_0^y)\ \right), \end{aligned} \tag{36}$$

where

$$\rho = \frac{b+\sqrt{b^2-4c}}{2}, \tag{37}$$

$$b = \delta + \varepsilon + \hat{\alpha}^2\gamma^2\delta(1+\varepsilon) + \hat{\alpha}\tau(1+\varepsilon),$$

$$c = \delta[\varepsilon - (1+\varepsilon)(\hat{\alpha}^2\gamma^2 - \hat{\alpha}\tau)],$$

and

$$\tau = \|I - A^{-1}\hat{A}\|_{\hat{A}_S}.$$

Proof. From (28) and (29), we have the two following equations

$$e_{i+1}^x = e_i^x - \Phi(Ae_i^x + B^T e_i^y), \tag{38}$$

$$e_{i+1}^y = e_i^y - \Psi(Ce_i^y - Be_{i+1}^x). \tag{39}$$

Substituting e_{i+1}^x in the equatinon (39) by the right side of the equation (38) gives

$$\begin{aligned} e_{i+1}^y &= e_i^y - \Psi(Ce_i^y - B\left(\ e_i^x - \Phi(Ae_i^x + B^T e_i^y)\ \right)) \\ &= (\hat{S}^{-1} - \Psi)(Ce_i^y - Be_i^x + B\Phi(Ae_i^x + B^T e_i^y)) \\ &\quad + e_i^y - \hat{S}^{-1}(Ce_i^y - Be_i^x + B\Phi(Ae_i^x + B^T e_i^y)) \\ &= (\hat{S}^{-1} - \Psi)(Ce_i^y - Be_i^x + B\Phi(Ae_i^x + B^T e_i^y)) \\ &\quad + \hat{S}^{-1}B(A^{-1} - \Phi)(Ae_i^x + B^T e_i^y) + \hat{S}^{-1}B(\hat{A}^{-1} - A^{-1})B^T e_i^y. \end{aligned}$$

It follows from the triangular inequality that

$$
\begin{aligned}
\|e^y_{i+1}\|_{\hat{S}_S} &\le \varepsilon\|Ce^y_i - Be^x_i + B\Phi(Ae^x_i + B^Te^y_i)\|_{(\hat{S}^{-1})_S} \qquad \text{by (31)}\\
&\quad +\|\hat{S}^{-1}B(A^{-1}-\Phi)(Ae^x_i + B^Te^y_i)\|_{\hat{S}_S}\\
&\quad +\|\hat{S}^{-1}B(\hat{A}^{-1}-A^{-1})B^Te^y_i\|_{\hat{S}_S}\\
&= \varepsilon\|\hat{S}e^y_i + B(\Phi - A^{-1})(Ae^x_i + B^Te^y_i) - B(\hat{A}^{-1}-A^{-1})B^Te^y_i\|_{(\hat{S}^{-1})_S}\\
&+ \|B(A^{-1}-\Phi)(Ae^x_i + B^Te^y_i)\|_{(\hat{S}^{-1})_S} + \|B(\hat{A}^{-1}-A^{-1})B^Te^y_i\|_{(\hat{S}^{-1})_S}\\
&\le \varepsilon\|e^y_i\|_{\hat{S}_S} + (1+\varepsilon)\|B(A^{-1}-\Phi)(Ae^x_i + B^Te^y_i)\|_{(\hat{S}^{-1})_S}\\
&\quad +(1+\varepsilon)\|B(\hat{A}^{-1}-A^{-1})B^Te^y_i\|_{(\hat{S}^{-1})_S}\\
&\le \varepsilon\|e^y_i\|_{\hat{S}} + \hat{\alpha}(1+\varepsilon)\|(A^{-1}-\Phi)(Ae^x_i + B^Te^y_i)\|_{\hat{A}_S}\\
&\quad +\hat{\alpha}(1+\varepsilon)\|(\hat{A}^{-1}-A^{-1})B^Te^y_i\|_{\hat{A}_S} \qquad \text{by Lemma 3}\\
&\le \varepsilon\|e^y_i\|_{\hat{S}_S} + \hat{\alpha}\gamma(1+\varepsilon)\|(A^{-1}-\Phi)(Ae^x_i + B^Te^y_i)\|_{A_S}\\
&\quad +\hat{\alpha}(1+\varepsilon)\|(I-A^{-1}\hat{A})\hat{A}^{-1}B^Te^y_i\|_{\hat{A}_S} \qquad \text{by (32)}\\
&\le \varepsilon\|e^y_i\|_{\hat{S}_S} + \hat{\alpha}\delta\gamma(1+\varepsilon)\|Ae^x_i + B^Te^y_i)\|_{(A^{-1})_S}\\
&\quad +\hat{\alpha}(1+\varepsilon)\|I-A^{-1}\hat{A}\|_{\hat{A}_S}\|\hat{A}^{-1}B^Te^y_i\|_{\hat{A}_S} \qquad \text{by (30)}\\
&\le \varepsilon\|e^y_i\|_{\hat{S}_S} + \hat{\alpha}\delta\gamma(1+\varepsilon)(\|e^x_i\|_{A_S} + \hat{\alpha}\gamma\|e^y_i\|_{\hat{S}_S})\\
&\quad +\hat{\alpha}(1+\varepsilon)\|I-A^{-1}\hat{A}\|_{\hat{A}_S}\|e^y_i\|_{\hat{S}_S} \qquad \text{by Lemma 4}\\
&= \hat{\alpha}\delta\gamma(1+\varepsilon)\|e^x_i\|_{A_S} + [\varepsilon + \hat{\alpha}^2\gamma^2\delta(1+\varepsilon) + \hat{\alpha}\tau(1+\varepsilon)]\|e^y_i\|_{\hat{S}_S}, \qquad (40)
\end{aligned}
$$

where $\tau = \|I - A^{-1}\hat{A}\|_{\hat{A}_S}$.

Using triangular inequality, from the equation (38), we have

$$
\begin{aligned}
\|e^x_{i+1}\|_{A_S} &= \|e^x_i - \Phi(Ae^x_i + B^Te^y_i)\|_{A_S}\\
&= \|(A^{-1}-\Phi)(Ae^x_i + B^Te^y_i) - A^{-1}B^Te^y_i\|_{A_S}\\
&\le \delta\|Ae^x_i + B^Te^y_i\|_{(A^{-1})_S} + \|B^Te^y_i\|_{(A^{-1})_S} \qquad \text{by (30)}\\
&\le \delta(\|e^x_i\|_{A_S} + \hat{\alpha}\gamma\|e^y_i\|_{\hat{S}_S}) + \|B^Te^y_i\|_{(A_S)^{-1}} \qquad \text{by (14) and Lemma 4}\\
&\le \delta(\|e^x_i\|_{A_S} + \hat{\alpha}\gamma\|e^y_i\|_{\hat{S}_S}) + \gamma\|B^Te^y_i\|_{(\hat{A}_S)^{-1}} \qquad \text{by (33)}\\
&\le \delta(\|e^x_i\|_{A_S} + \hat{\alpha}\gamma\|e^y_i\|_{\hat{S}_S}) + \hat{\alpha}\gamma\|B^Te^y_i\|_{(\hat{A}^{-1})_S}\\
&\le \delta(\|e^x_i\|_{A_S} + \hat{\alpha}\gamma\|e^y_i\|_{\hat{S}_S}) + \hat{\alpha}\gamma\|e^y_i\|_{\hat{S}_S}\\
&= \delta\|e^x_i\|_{A_S} + \hat{\alpha}\gamma(1+\delta)\|e^y_i\|_{\hat{S}_S}. \qquad (41)
\end{aligned}
$$

It follows from (40) and (41) that

$$
\begin{pmatrix} \|e^x_i\|_{A_S} \\ \|e^y_i\|_{\hat{S}_S} \end{pmatrix} \le \hat{N}^i \begin{pmatrix} \|e^x_0\|_{A_S} \\ \|e^y_0\|_{\hat{S}_S} \end{pmatrix}, \qquad (42)
$$

where $\hat{N}$ is given by

$$\hat{N} = \begin{pmatrix} \delta & \hat{\alpha}\gamma(1+\delta) \\ \hat{\alpha}\delta\gamma(1+\varepsilon) & \varepsilon + \hat{\alpha}^2\gamma^2\delta(1+\varepsilon) + \hat{\alpha}\tau(1+\varepsilon) \end{pmatrix}.$$

Obviously, $\hat{N}$ is symmetric with respect to the following inner product of the two-dimensional Euclidean space

$$\left(\begin{pmatrix} \delta(1+\varepsilon) & 0 \\ 0 & 1+\delta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right) = \delta(1+\varepsilon)x_1x_2 + (1+\delta)y_1y_2.$$

Then, from (42) we have

$$\begin{aligned} & \delta(1+\varepsilon)(A_S e_i^x, e_i^x) + (1+\delta)(\hat{S}_S e_i^y, e_i^y) \\ = & \left(\begin{pmatrix} \delta(1+\varepsilon) & 0 \\ 0 & 1+\delta \end{pmatrix} \begin{pmatrix} \|e_i^x\|_{A_S} \\ \|e_i^y\|_{\hat{S}_S} \end{pmatrix}, \begin{pmatrix} \|e_i^x\|_{A_S} \\ \|e_i^y\|_{\hat{S}_S} \end{pmatrix} \right) \\ \leq & \left(\begin{pmatrix} \delta(1+\varepsilon) & 0 \\ 0 & 1+\delta \end{pmatrix} \hat{N}^i \begin{pmatrix} \|e_0^x\|_{A_S} \\ \|e_0^y\|_{\hat{S}_S} \end{pmatrix}, \hat{N}^i \begin{pmatrix} \|e_0^x\|_{A_S} \\ \|e_0^y\|_{\hat{S}_S} \end{pmatrix} \right) \\ \leq & \rho^{2i}[\delta(1+\varepsilon)(A_S e_0^x, e_0^x) + (1+\delta)(\hat{S}_S e_0^y, e_0^y)], \end{aligned}$$

where ρ is the spectral radius of $\hat{N}$. The eigenvalues of $\hat{N}$ are the roots of

$$\lambda^2 - b\lambda + c = 0.$$

where $b = \delta + \varepsilon + \hat{\alpha}^2\gamma^2\delta(1+\varepsilon) + \hat{\alpha}\tau(1+\varepsilon)$ and $c = \delta[\varepsilon - (1+\varepsilon)(\hat{\alpha}^2\gamma^2 - \hat{\alpha}\tau)]$.

The spectral radius ρ of $\hat{N}$ is equal to its positive eigenvalue which is given by (37). It is easy to see if (34) is satisfied , then $\rho < 1$. This completes the proof of (35).

To prove (36), we apply the following elementary inequality

$$(a+b)^2 \leq (1+\eta)a^2 + (1+\eta^{-1})b^2$$

to (41) and get, for any $\eta > 0$,

$$\|e_i^x\|_{\hat{A}_S} \leq (1+\eta)\delta^2\|e_{i-1}^x\|_{\hat{A}_S} + (1+\eta^{-1})\hat{\alpha}^2\gamma^2(1+\delta)^2\|e_{i-1}^y\|_{\hat{S}_S}.$$

Inequality (36) follows from taking $\eta = \frac{\hat{\alpha}^2\gamma^2(1+\varepsilon)(1+\delta)}{\delta}$ and applying (35). This completes the proof of the theorem.

□

Remark 3. In the implementation of nonlinear Uzawa algorithms, we often make use of an incomplete LU factorization [6, 11, 17] of $A : A = LU - R$. In order to make $\gamma \leq 1$ hold, we should choose an appropriate scale factor ϑ and set $\hat{A} = \vartheta LU$. In our numerical experiments, Φ results from several iterations of preconditioning GMRES applied to the following system

$$Ax = f - Ax_i - B^T y_i$$

with the preconditioner LU and zero initial guess. In the practical implementation of the new method, one or two iterations are enough. Ψ results from several iterations of GMRES applied to the following approximate Schur complement system

$$(B\hat{A}^{-1}B^T + C)y = Bx_{i+1} - Cy_i - g$$

with zero initial guess.

If A and $\hat{A}$ are symmetric positive definite, $\hat{\alpha}$ can be 1. Therefore, we have the following corollary.

Corollary 2. *Assume that A and $\hat{A}$ are symmetric positive definite. And assume that (30), (31) and (32) hold. Let (x, y) be the solution pair of (1), and $\{(x_i, y_i)\}$ be defined by Algorithm 4. Then, x_i and y_i converge to x and y, respectively, if*

$$0 < \delta < \frac{1-\tau}{1-\tau+2\gamma^2} \text{ and } 0 < \varepsilon < \frac{1-\delta-\mu}{1-\delta+\mu}.$$

where $\mu = \tau - \tau\delta + 2\gamma^2\delta$. In this case, the following two inequalities hold

$$\begin{aligned} &\delta(1+\varepsilon)(\hat{A}e_i^x, e_i^x) + (1+\delta)(\hat{S}e_i^y, e_i^y) \\ &\leq \rho^{2i}[\delta(1+\varepsilon)(\hat{A}e_0^x, e_0^x) + (1+\delta)(\hat{S}e_0^y, e_0^y)] \end{aligned}$$

and

$$\begin{aligned} (\hat{A}e_i^x, e_i^x) \quad \leq \quad & \left(\gamma^2(1+\delta) + \frac{\delta}{1+\varepsilon}\right) \\ & \cdot\rho^{2i-2}\left(\ \delta(1+\varepsilon)(\hat{A}e_0^x, e_0^x) + (1+\delta)(\hat{S}e_0^y, e_0^y)\ \right), \end{aligned}$$

where

$$\begin{aligned} \rho &= \frac{b+\sqrt{b^2-4c}}{2}, \\ b &= \delta + \varepsilon + (\gamma^2\delta + \tau)(1+\varepsilon), \\ c &= \delta[\varepsilon - (1+\varepsilon)(\gamma^2 - \tau)], \end{aligned}$$

and

$$\tau = \|I - \hat{A}^{1/2}A^{-1}\hat{A}^{1/2}\|_2.$$

4. Comparison of the convergence rates of three nonlinear uzawa methods

In this section, we will compare the convergence rates of three nonlinear methods. In this comparison, it is assumed that δ, ε in inequalities (9),(12),(30),(31) are very small. We mainly focus on the convergence factor ρ of three nonlinear Uzawa methods. In order to identify the convergence factor of three nonlinear methods, we use three notations ρ_B, ρ_C,ρ_L to denote the convergence factor of Bramble's nonlinear method, the convergence factor of Cao's nonlinear method and the convergence factor of our new nonlinear method, respectively.

From Theorem 1, the convergence factor of Bramble's nonlinear Uzawa method is

$$\rho_B = \frac{(1+\tau)\delta + \theta + \sqrt{((1+\tau)\delta+\theta)^2 + 4\delta(\tau-\theta)}}{2},$$

where $\gamma \in (0, \gamma_{opt}]$, $\gamma_{opt} = \lambda_{min}(Q_B^{-1}(B(A_S)^{-1}B^T + C))$ and $\theta = (1 - \frac{\tau\gamma}{\alpha^2})^{1/2}$. When δ is very small, the convergence factor ρ_B is approximately equal to θ. So, we have the following relation

$$\rho_B \approx \theta,$$

and the optimal of ρ_B defined as ρ_{Bo} satisfies

$$\rho_{Bo} \approx (1 - \frac{\gamma_{opt}^2}{\alpha^2})^{1/2}. \tag{43}$$

From Theorem 2, the convergence factor of Cao's nonlinear Uzawa method is

$$\rho_C = \frac{\delta + \varepsilon + \sqrt{(\delta+\varepsilon)^2 + 4\delta((\hat{\alpha}-1)\varepsilon + \hat{\alpha})}}{2}.$$

From (11), we know that $\delta \geq \|I - \hat{A}^{-1}A\|_{\hat{A}_S}$. When ε is very small, the convergence factor ρ_C is approximately equal to $\frac{\delta + \sqrt{\delta^2 + 4\delta\hat{\alpha}}}{2}$. So, we have the following relation

$$\rho_C \approx \frac{\delta + \sqrt{\delta^2 + 4\delta\hat{\alpha}}}{2},$$

and the optimal of ρ_C defined as ρ_{Co} satisfies

$$\rho_{Co} \approx \frac{\|I - \hat{A}^{-1}A\|_{\hat{A}_S} + \sqrt{\|I - \hat{A}^{-1}A\|_{\hat{A}_S}^2 + 4\hat{\alpha}\|I - \hat{A}^{-1}A\|_{\hat{A}_S}}}{2}. \tag{44}$$

From Theorem 4, the convergence factor of our nonlinear Uzawa method is

$$\rho_L = \frac{b + \sqrt{b^2 - 4c}}{2},$$

where $b = \delta + \varepsilon + \hat{\alpha}^2\gamma^2\delta(1+\varepsilon) + \hat{\alpha}\tau(1+\varepsilon)$ and $c = \delta[\varepsilon - (1+\varepsilon)(\hat{\alpha}^2\gamma^2 - \hat{\alpha}\tau)]$. When δ and ε are very small, the convergence factor ρ_L is approximately equal to $\hat{\alpha}\tau$. So, we have the following relation

$$\rho_L \approx \hat{\alpha}\tau = \hat{\alpha}\|I - A^{-1}\hat{A}\|_{\hat{A}_S}. \tag{45}$$

From (43), (44) and (45), we should expect that if we choose Q_B and $\hat{A}$ so that they are good approximation to $B(A_S)^{-1}B^T + C$ and A, respectively, then these nonlinear methods may have small convergence factors when their nonlinear approximations are good, i.e., δ, ε in inequalities (9),(12),(30), (31) are small.

We note that it is difficult to choose a matrix Q_B which is a good approximation to $B(A_S)^{-1}B^T + C$. Usually, we choose $Q_B = aI$, where I is a identity matrix and a is a factor. It is not a good approximation in most cases and many times iterations may be needed in order to get a good approximation to the solution of (1). However, as we have pointed out, we may choose $\hat{A} = LU$ in Cao's method and choose $\hat{A} = \vartheta LU$ in our method, where LU is the imcomplete LU factorization of A. At this time, $\hat{A}$ is a good approximation to A.

Table 1. Iteration number for global stabilization

	$\nu = 1$		$\nu = 0.1$		$\nu = 0.01$	
	NUS	**NUAS**	**NUS**	**NUAS**	**NUS**	**NUAS**
h=1/8	9	7	10	10	62	62
h=1/16	22	9	17	13	142	142
h=1/32	53	(2/3)19	43	20	188	190
h=1/64	142	(1/4)36	112	(2/3)39	177	178

5. Numerical experiments

The problem under consideration arises from the linearization of the steady-state Navier-Stokes equations, i.e. the Oseen problems of the following form

$$\begin{cases} -\nu\Delta u + w\cdot\nabla u + \nabla p &= f \text{ in } \Omega \\ divu &= 0 \text{ in } \Omega \end{cases} \tag{46}$$

with suitable boundary conditions on $\partial\Omega$, where $\Omega \subset R^2$ is a bounded domain and w is a given divergence free field. The scalar ν is the viscosity, the vector field u represents the velocity, and p denotes the pressure.

The test problem is a leaky two-dimensional lid-driven cavity problem in a square domain $(0 \le x \le 1) : (0 \le y \le 1)$. The boundary conditions are $u_x = u_y = 0$ on the three fixed walls $(x = 0, y = 0, x = 1)$, and $u_x = 1$,$u_y = 0$ on the moving wall $(y = 1)$. We take the constant wind: $w_x = 1$, $w_y = 2$.

To discretize (46), we take a finite element subdivision based on $n\times n$ uniform grids of square elements. The mixed finite element used is the bilinear-constant velocity-pressure: $Q_1 - P_0$ pair with global stabilization or local stabilization, see [4, 14].

In our numerical experiments, we take $\beta = 1$ and 0.25 for global and local stabilizations, respectively. The incomplete LU factorization of A [6, 11, 17] is: $A = LU - R$ and the drop tolerance for A is 0.01 in all case. Because it has been shown that the second nonlinear Uzawa method (**NUS** for short) converges faster than the first [4], we only compare our method (**NUAS** for short) with **NUS**. In **NUS**, $\hat{A} = LU$ and Ψ defined by five **GMRES** steps. In **NUAS**, Φ defined by two **GMRES** steps with preconditioner $M = LU$ and Ψ defined by five **GMRES** steps with $\hat{A} = \vartheta LU$, where ϑ is a parameter. The stop criterion is

$$\frac{\|r_k\|_2}{\|r_0\|_2} \le 10^{-6}.$$

In all of the figures, the dashdot line and solid line stand for the convergence histories of **NUS** and **NUAS**, respectively.

Table 1 and Table 2 show the number of iterations of **NUS** and **NUAS** in four mesh grids in case $\nu = 1, 0.1$ and 0.01 for global and local stabilizations, respectively. For **NUAS**, we give out the value of ϑ when it is not be one, i.e., the value in the bracket is the value of ϑ using in **NUAS**.

Table 2. Iteration number for local stabilization

	$\nu = 1$		$\nu = 0.1$		$\nu = 0.01$	
	NUS	**NUAS**	**NUS**	**NUAS**	**NUS**	**NUAS**
h=1/8	9	6	11	11	86	92
h=1/16	23	(2/3)13	17	13	178	178
h=1/32	70	(1/3)18	42	(5/6)18	200	201
h=1/64	213	(1/4)39	118	(1/2)42	179	180

From Table 1 and 2, we can see that the number of iterations of **NUAS** is much less than that of **NUS** when the viscosity ν is big ($\nu = 1, 0.1$). This can be seen more obvious when the mesh size are smaller. But the number of iterations of **NUAS** is about the same as that of **NUS** when ($\nu = 0.01$). We show the convergence histories in Figure 1-6.

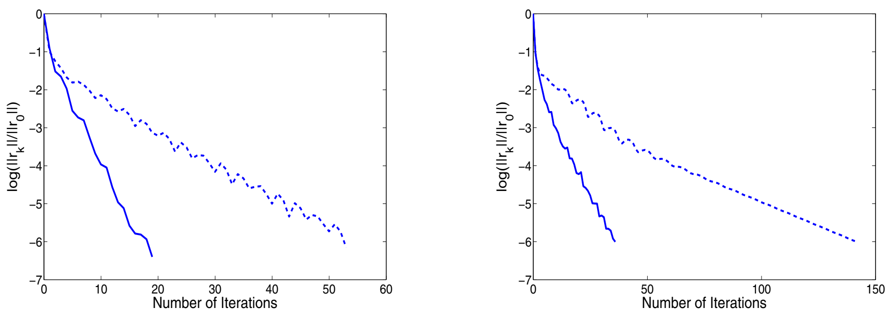

Figure 1. The convergence histories of **NUS** and **NUAS** for global stabilization with $\nu = 1$.Left: grid 32×32; Right: grid 64×64.

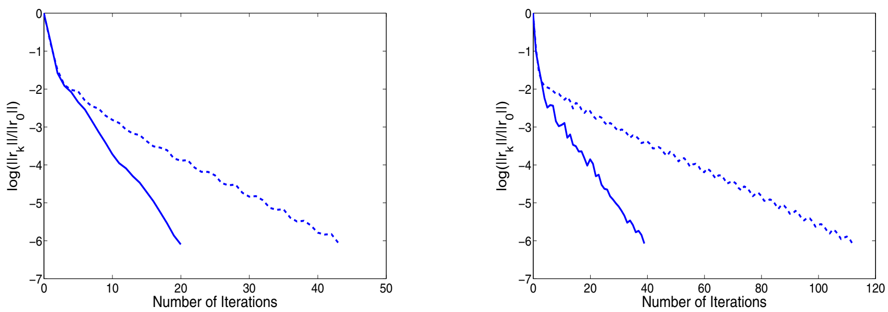

Figure 2. The convergence histories of **NUS** and **NUAS** for global stabilization with $\nu = 0.1$. Left: grid 32×32; Right: grid 64×64.

It can be seen that when $\nu = 1$ and $\nu = 0.1$, **NUAS** converges faster than **NUS** and the convergence curves of **NUAS** are much smoother than those

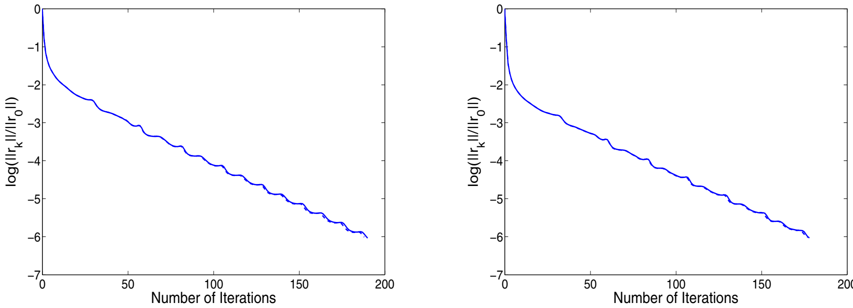

Figure 3. The convergence histories of **NUS** and **NUAS** for global stabilization with $\nu = 0.01$. Left: grid 32×32; Right: grid 64×64.

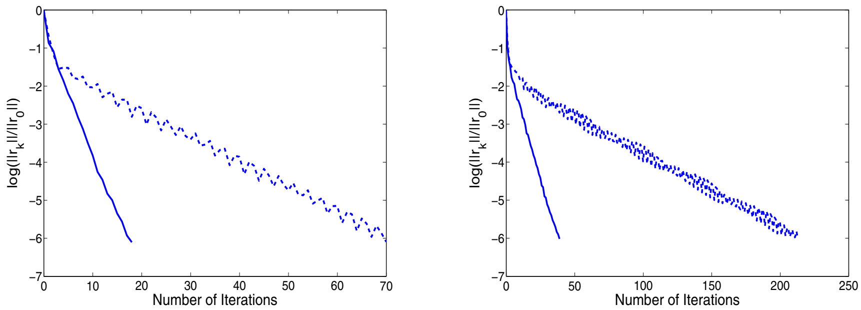

Figure 4. The convergence histories of **NUS** and **NUAS** for local stabilization with $\nu = 1$. Left: grid 32×32; Right: grid 64×64.

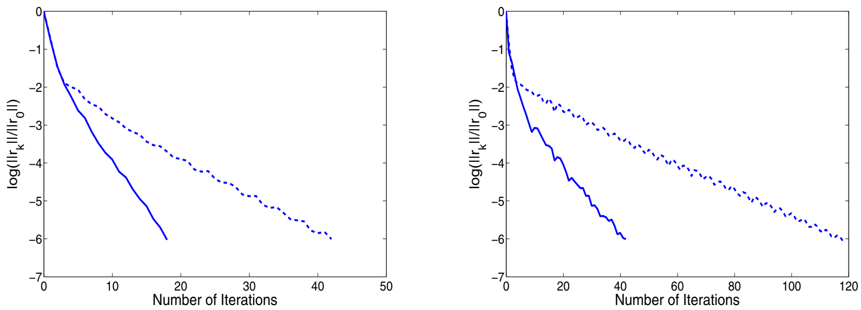

Figure 5. The convergence histories of **NUS** and **NUAS** for local stabilization with $\nu = 0.1$. Left: grid 32×32; Right: grid 64×64.

of **NUS**, but when ν is small ($\nu = 0.01$), all these become similar and the convergence curves of **NUAS** even overlap with those of **NUS**.

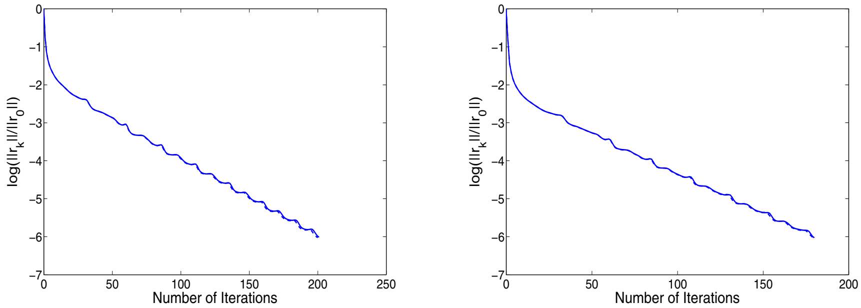

Figure 6. The convergence histories of **NUS** and **NUAS** for local stabilization with $\nu = 0.01$. Left: grid 32×32; Right: grid 64×64.

6. Conclusion

The new nonlinear Uzawa method (**NUAS**), which is defined by two nonlinear approximate inverses, for solving the nonsymmetric saddle point problems is proposed and its convergence results are deduced. Moreover, we compare the convergence rates of three nonlinear Uzawa methods. Numerical experiments show that **NUAS** needs less iterations and has a smoother convergence curve than **NUS** when applied to Navier-Stokes equations with a big viscosity discretized by mixed finite elements. When the viscosity is small, the convergence behavior of **NUAS** is similar to that of **NUS**.

Acknowledgments

The authors would like to thank the anonymous referees for their helpful suggestions, which greatly improve the paper. They also wish to thank their Ph.D supervisor, Prof. Zhi-Hao Cao for his inspired guidance in studying the numerical linear algebra.

References

1. M. Benzi, G. H. Golub and J. Liesen, *Numerical solution of saddle point problems*, Acta Numer. 14(2005), 1-137.
2. J.H. Bramble, J.E. Pasciak and A.T. Vassilev, *Uzawa type algorithms for nonsymmetric saddle point problems*, Math. Comp. 69(1999), 667-689.
3. F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Sprnger-Verlag, 1991.
4. Z-H. Cao, *Fast Uzawa algorithms for solving non-symmetric stabilized saddle point problems* , Numer. Linear Algebra Appl. 11(2004), 1-24.
5. V. Girault and P. A. Raviart, *Finite element approximation of the Navier-Stokes equations*, Lecture Notes in Math. # 749, Sprnger-Verlag, New York, 1981.
6. N. Guessous and O. Souhar, *The effect of block red-black ordering on block ILU preconditioner for sparse matrices*, J. Appl. Math. Comput. 17(2005), 283-296.

7. M. Gunzburger, *Finite element methods for viscous incompressible flows*, Academic Press: San Diego, 1989.
8. O. A. Karakashian, *On a Galerkin-Lagrange multiplier method for the stationary Navier-Stokes equations*, SIAM J. Numer. Anal. 19(1982), 909-923.
9. Y. Lin and Y. Cao, *A new nonlinear Uzawa algorithm for generalized saddle point problems*, Appl. Math. Comput. 175(2006), 1432-1454.
10. Y. Lin and Y. Wei, *Fast corrected Uzawa methods for symmetric saddle point problems*, CALCOLO, to appear.
11. Y. Saad, *Iterative method for sparse linear systems*, PWS, New York, 1996.
12. Y. Saad and M. H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput. 7(1986), 856-869.
13. D.J. Silvester and A. Wathen, *Fast iterative solution of stabilized Stokes systems part II: using general block preconditioner*, SIAM J. Numer. Anal. 31(1994), 1352-1367.
14. D.J. Silvester and N. Kechkar, *Stabilized bilinear-constant velocity-pressure finite elements for the conjugate gradient solution of the stokes problem*, Comput. Methods Appl. Mech. Engrg. 79(1990), 71-86.
15. R. Temam, *Navier-Stokes equations*, North-Holland Publishing Co., York, 1977.
16. A. Wathen and D.J. Silvester, *Fast iterative solution of stabilized Stokes systems part I: using simple diagonal preconditioner*, SIAM J. Numer. Anal. 33(1993), 630-649.
17. J. Yun, S. Oh and E. Kim, *Convergence of Parallel multisplitting methods using ILU factorizations*, J. Appl. Math. Comput. 15(2004), 77-90.

Yanhua Cao received his BS and MS from Xiangtan University, Ph.D from Fudan University and now is a post-doctor of Institute of Applied Physics and Computation of Mathematics. His research interests focus on numerical linear algebra.

Institute of Applied Physics and Computation of Mathematics, Beijing 100088, P.R. China.
`e-mail: 021018037@@fudan.edu.cn`

Yiqin Lin received his BS from Jilin University, Ph.D from Fudan University and now is a lecturer of Sun Yat-Sen University. His research interests focus on numerical linear algebra, computation in control system, saddle point problem and model-order reduction.

School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou 510275, P. R. China.
`e-mail: linyiqin@@mail.sysu.edu.cn`

Yimin Wei received his BS and MS from Shanghai Normal University, Ph.D from Fudan University and now is an associate professor of Fudan University. His research interests focus on numerical linear algebra, especially on the perturbation theory for generalized inverse.

Institute of Mathematics, School of Mathematical Sciences, Fudan University, Shanghai 200433, P. R. China and Key Laboratory of Mathematics for Nonlinear Sciences (Fudan University), Ministry of Education.
`e-mail: ymwei@@fudan.edu.cn`