Numer. Math. (2004) 98: 1–32
Digital Object Identifier (DOI) 10.1007/s00211-004-0521-1

Numerische Mathematik

Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite linear systems

Zhong-Zhi Bai[1,⋆], **Gene H. Golub**[2,⋆⋆], **Jian-Yu Pan**[1]

[1] State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, People's Republic of China; e-mail: {bzz, pjy}@lsec.cc.ac.cn

[2] Scientific Computing and Computational Mathematics Program, Department of Computer Science, Stanford University, Stanford, CA 94305-9025, USA; e-mail: golub@sccm.stanford.edu

Received November 20, 2002 / Revised version received November 5, 2003 /
Published online March 16, 2004 –

Summary. For the positive semidefinite system of linear equations of a block two-by-two structure, by making use of the Hermitian/skew-Hermitian splitting iteration technique we establish a class of preconditioned Hermitian/skew-Hermitian splitting iteration methods. Theoretical analysis shows that the new method converges unconditionally to the unique solution of the linear system. Moreover, the optimal choice of the involved iteration parameter and the corresponding asymptotic convergence rate are computed exactly. Numerical examples further confirm the correctness of the theory and the effectiveness of the method.

Mathematics Subject Classification: 65F10, 65F50, CR: G1.3

1 Introduction

For iterative methods for the solutions of large sparse non-Hermitian and positive *definite* systems of linear equations

$$(1) \qquad Ax = b, \qquad A \in \mathbb{C}^{n\times n} \quad \text{nonsingular}, \quad \text{and} \quad x, b \in \mathbb{C}^n,$$

⋆ Subsidized by The Special Funds For Major State Basic Research Projects G1999032803

⋆⋆ Research supported, in part, by DOE-FC02-01ER4177

Correspondence to: Z.-Z. Bai

based on the *Hermitian/skew-Hermitian* **(HS)** splitting[9,11,15]

$$A = H + S,$$

where

$$H = \frac{1}{2}(A + A^*) \quad \text{and} \quad S = \frac{1}{2}(A - A^*),$$

of the coefficient matrix A, Bai, Golub and Ng[2] have recently proposed the *Hermitian/skew-Hermitian splitting* **(HSS)** iteration method. Theoretical analysis has shown that this HSS iteration converges unconditionally to the exact solution of the system of linear equations (1), with the convergence speed about the same as that of the conjugate gradient method if it is applied to a Hermitian matrix. Moreover, the upper bound of its contraction factor is dependent on the spectrum of the Hermitian part H, but is independent of the spectrum of the skew-Hermitian part S as well as the eigenvalues of the matrices H, S and A. Numerical implementation has also shown that the HSS iteration is very efficient and robust for solving the non-Hermitian and positive definite linear systems.

Recently, Benzi and Golub[3] directly applied the HSS iteration technique to the following system of linear equations with a block coefficient matrix

$$A \begin{bmatrix} y \\ z \end{bmatrix} \equiv \begin{bmatrix} B & E \\ -E^* & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \equiv b, \tag{2}$$

and proved that the resulting iteration method is also convergent for any positive constant α, where α is the involved iteration parameter. This result extends the application region of the original HSS method. Here, the submatrices $B \in \mathbb{C}^{p\times p}$ is Hermitian positive definite and $E \in \mathbb{C}^{p\times q}$ $(p \geq q)$ is of full column rank. Therefore, $A \in \mathbb{C}^{n\times n}$, with $n = p + q$, is a nonsingular, non-Hermitian, and positive *semidefinite* matrix.

The block system of linear equations (2) corresponds to the Kuhn-Tucker conditions for linearly constrained quadratic programming problems or saddle point problems. Such systems typically result from mixed or hybrid finite element approximations of second-order elliptic problems, elasticity problems or the Stokes equations (see, e.g., Brezzi and Fortin[5]) and from Lagrange multiplier methods (see, e.g., Fortin and Glowinski[8]).

We note that in the convergence theorem in [3], it was only proved that the spectral radius of the iteration matrix is less than one. No discussion about other convergence properties such as estimates of the contraction factor and the asymptotic convergence rate, as well as the choice of the optimal iteration parameter in a certain sense, was demonstrated. In fact, it is very difficult to derive such properties for the method proposed in [3], as each eigenvalue of the iteration matrix depends not only on the iteration parameter and all eigenvalues of the Hermitian part of the coefficient matrix, but

also on the eigenvectors corresponding to those eigenvalues. In this paper, by first transforming (or preconditioning) the block system of linear equations (2) into an equivalent one, and then applying the HSS method directly to the preconditioned block linear system, we establish a class of *preconditioned Hermitian/skew-Hermitian splitting* **(PHSS)** iteration methods for the non-Hermitian positive semidefinite system of linear equations (2).

The PHSS method converges unconditionally to the exact solution of the block system of linear equations (2), with a comparable workload and faster convergence speed than the one studied in [3]. In particular, we derive upper bound of its contraction factor and present optimal choice of the involved iteration parameter. Numerical implementation confirms the correctness of the convergence theory, and also shows the robustness and effectiveness of the PHSS method for solving the non-Hermitian positive semidefinite system of linear equations (2).

2 The PHSS iteration

We first review the HSS iteration method for solving the non-Hermitian positive definite system of linear equations (1) presented in Bai, Golub and Ng[2].

The HSS iteration method

Let $A \in \mathbb{C}^{n\times n}$ be a positive definite matrix. Given an initial guess $x^{(0)} \in \mathbb{C}^n$. For $k = 0, 1, 2, \ldots$ until $\{x^{(k)}\}$ converges, compute

$$\begin{cases} (\alpha I + H)x^{(k+\frac{1}{2})} = (\alpha I - S)x^{(k)} + b, \\ (\alpha I + S)x^{(k+1)} \ = (\alpha I - H)x^{(k+\frac{1}{2})} + b, \end{cases}$$

where α is a given positive constant.
In matrix-vector form, the above HSS iteration method can be equivalently rewritten as

$$x^{(k+1)} = \mathcal{M}(\alpha)x^{(k)} + \mathcal{G}(\alpha)b, \qquad k = 0, 1, 2, \ldots, \tag{3}$$

where

$$\mathcal{M}(\alpha) = (\alpha I + S)^{-1}(\alpha I - H)(\alpha I + H)^{-1}(\alpha I - S)$$

and

$$\mathcal{G}(\alpha) = 2\alpha(\alpha I + S)^{-1}(\alpha I + H)^{-1}.$$

Here, $\mathcal{M}(\alpha)$ is the iteration matrix of the HSS iteration. In fact, (3) may also result from the splitting

$$A = F(\alpha) - G(\alpha)$$

of the coefficient matrix A, with

$$\begin{cases} F(\alpha) = \frac{1}{2\alpha}(\alpha I + H)(\alpha I + S), \\ G(\alpha) = \frac{1}{2\alpha}(\alpha I - H)(\alpha I - S). \end{cases}$$

The following theorem describes the convergence property of the HSS iteration.

Theorem 2.1 *[2] Let $A \in \mathbb{C}^{n\times n}$ be a positive definite matrix, $H = \frac{1}{2}(A+A^*)$ and $S = \frac{1}{2}(A-A^*)$ be its Hermitian and skew-Hermitian parts, respectively, and α be a positive constant. Then the spectral radius $\rho(\mathcal{M}(\alpha))$ of the iteration matrix $\mathcal{M}(\alpha)$ of the HSS iteration is bounded by*

$$\sigma(\alpha) = \max_{\lambda_j \in \lambda(H)} \frac{|\alpha - \lambda_j|}{|\alpha + \lambda_j|},$$

where $\lambda(H)$ is the spectral set of the matrix H. Therefore, it follows that

$$\rho(\mathcal{M}(\alpha)) \le \sigma(\alpha) < 1, \qquad \forall \alpha > 0,$$

i.e., the HSS iteration converges to the exact solution $x^* \in \mathbb{C}^n$ of the system of linear equations (1).

Moreover, if $\gamma_{\min}$ and $\gamma_{\max}$ are the lower and the upper bounds of the eigenvalues of the matrix H, respectively, then

$$\alpha^* \equiv \arg\min_{\alpha} \left\{ \max_{\gamma_{\min} \le \lambda \le \gamma_{\max}} \left| \frac{\alpha - \lambda}{\alpha + \lambda} \right| \right\} = \sqrt{\gamma_{\min}\gamma_{\max}}$$

and

$$\sigma(\alpha^*) = \frac{\sqrt{\gamma_{\max}} - \sqrt{\gamma_{\min}}}{\sqrt{\gamma_{\max}} + \sqrt{\gamma_{\min}}} = \frac{\sqrt{\kappa(H)} - 1}{\sqrt{\kappa(H)} + 1},$$

where $\kappa(H)$ is the spectral condition number of H.

To establish the convergence properties of iterative method for the non-Hermitian positive semidefinite system of linear equations (2), we introduce matrices

$$P = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} \in \mathbb{C}^{n\times n} \quad \text{and} \quad \bar{E} = B^{-\frac{1}{2}} E C^{-\frac{1}{2}} \in \mathbb{C}^{p\times q}, \tag{4}$$

where $C \in \mathbb{C}^{q\times q}$ is a Hermitian positive definite submatrix, and define

$$\bar{A} = P^{-\frac{1}{2}} A P^{-\frac{1}{2}} = \begin{bmatrix} I & \bar{E} \\ -\bar{E}^* & 0 \end{bmatrix},$$

$$\begin{bmatrix} \overline{y} \\ \overline{z} \end{bmatrix} = P^{\frac{1}{2}} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} B^{\frac{1}{2}} y \\ C^{\frac{1}{2}} z \end{bmatrix}$$

and

$$\overline{b} = \begin{bmatrix} \bar{f} \\ \overline{g} \end{bmatrix} = P^{-\frac{1}{2}} b = \begin{bmatrix} B^{-\frac{1}{2}} f \\ C^{-\frac{1}{2}} g \end{bmatrix}.$$

Then the system of linear equations (2) can be transformed into the following equivalent one:

$$\bar{A} \begin{bmatrix} \overline{y} \\ \overline{z} \end{bmatrix} = \overline{b}. \tag{5}$$

As a matter of fact, we may also consider (5) as a preconditioned form of (2), with the preconditioning matrix P in (4).

We remark that the matrix $C \in \mathbb{C}^{q\times q}$ is a free-of-choice, and the best choice is to be seen from the convergence theorem and is discussed at the end of Section 3.

Evidently, the Hermitian and skew-Hermitian parts of the matrix $\bar{A} \in \mathbb{C}^{n\times n}$ are, respectively,

$$\overline{H} = \frac{1}{2}(\bar{A} + \bar{A}^*) = \begin{bmatrix} I & \\ & 0 \end{bmatrix} \quad \text{and} \quad \overline{S} = \frac{1}{2}(\bar{A} - \bar{A}^*) = \begin{bmatrix} 0 & \bar{E} \\ -\bar{E}^* & 0 \end{bmatrix}.$$

By straightforwardly applying the HSS iteration technique to (5), we then obtain the iteration scheme

$$\begin{cases} (\alpha I + \overline{H}) \begin{bmatrix} \overline{y}^{(k+\frac{1}{2})} \\ \overline{z}^{(k+\frac{1}{2})} \end{bmatrix} = (\alpha I - \overline{S}) \begin{bmatrix} \overline{y}^{(k)} \\ \overline{z}^{(k)} \end{bmatrix} + \overline{b}, \\ (\alpha I + \overline{S}) \begin{bmatrix} \overline{y}^{(k+1)} \\ \overline{z}^{(k+1)} \end{bmatrix} = (\alpha I - \overline{H}) \begin{bmatrix} \overline{y}^{(k+\frac{1}{2})} \\ \overline{z}^{(k+\frac{1}{2})} \end{bmatrix} + \overline{b}, \end{cases}$$

or equivalently,

$$\begin{bmatrix} \alpha I & \bar{E} \\ -\bar{E}^* & \alpha I \end{bmatrix} \begin{bmatrix} \overline{y}^{(k+1)} \\ \overline{z}^{(k+1)} \end{bmatrix} = \begin{bmatrix} \frac{\alpha(\alpha-1)}{\alpha+1} I & -\frac{\alpha-1}{\alpha+1}\bar{E} \\ \bar{E}^* & \alpha I \end{bmatrix} \begin{bmatrix} \overline{y}^{(k)} \\ \overline{z}^{(k)} \end{bmatrix} + \begin{bmatrix} \frac{2\alpha}{\alpha+1}\overline{f} \\ 2\overline{g} \end{bmatrix}.$$

It then follows immediately that in the original variable we have

$$\begin{bmatrix} \alpha B & E \\ -E^* & \alpha C \end{bmatrix} \begin{bmatrix} y^{(k+1)} \\ z^{(k+1)} \end{bmatrix} = \begin{bmatrix} \frac{\alpha(\alpha-1)}{\alpha+1} B & -\frac{\alpha-1}{\alpha+1} E \\ E^* & \alpha C \end{bmatrix} \begin{bmatrix} y^{(k)} \\ z^{(k)} \end{bmatrix} + \begin{bmatrix} \frac{2\alpha}{\alpha+1} f \\ 2g \end{bmatrix}, \tag{6}$$

which results in the following *preconditioned Hermitian/skew-Hermitian splitting* **(PHSS)** iteration method for solving the block system of linear equations (2).

The PHSS iteration method

Given an initial guess $x^{(0)} = (y^{(0)^*}, z^{(0)^*})^* \in \mathbb{C}^n$. For $k = 0, 1, 2, \ldots$ until $x^{(k)} = (y^{(k)^*}, z^{(k)^*})^* \in \mathbb{C}^n$ converges, compute the next iterate $x^{(k+1)} = (y^{(k+1)^*}, z^{(k+1)^*})^*$ by solving the 2-by-2 block linear system (6), where α is a given positive constant.

Evidently, the PHSS iteration method can be equivalently rewritten as

$$\begin{bmatrix} y^{(k+1)} \\ z^{(k+1)} \end{bmatrix} = \mathcal{L}(\alpha) \begin{bmatrix} y^{(k)} \\ z^{(k)} \end{bmatrix} + \mathcal{N}(\alpha) \begin{bmatrix} f \\ g \end{bmatrix}, \tag{7}$$

where

$$\mathcal{L}(\alpha) = \begin{bmatrix} \alpha B & E \\ -E^* & \alpha C \end{bmatrix}^{-1} \begin{bmatrix} \frac{\alpha(\alpha-1)}{\alpha+1} B & -\frac{\alpha-1}{\alpha+1} E \\ E^* & \alpha C \end{bmatrix} \tag{8}$$

and

$$\mathcal{N}(\alpha) = \begin{bmatrix} \alpha B & E \\ -E^* & \alpha C \end{bmatrix}^{-1} \begin{bmatrix} \frac{2\alpha}{\alpha+1} I & \\ & 2I \end{bmatrix}.$$

Here, $\mathcal{L}(\alpha)$ is the iteration matrix of the PHSS iteration. In fact, (7) may also result from the splitting

$$A = M(\alpha) - N(\alpha) \tag{9}$$

of the coefficient matrix A, with

$$M(\alpha) = \begin{bmatrix} \frac{\alpha+1}{2} B & \frac{\alpha+1}{2\alpha} E \\ -\frac{1}{2} E^* & \frac{\alpha}{2} C \end{bmatrix}, \qquad N(\alpha) = \begin{bmatrix} \frac{\alpha-1}{2} B & -\frac{\alpha-1}{2\alpha} E \\ \frac{1}{2} E^* & \frac{\alpha}{2} C \end{bmatrix}.$$

In actual computations, at each iterate of the PHSS iteration we need to solve a linear system with the coefficient matrix

$$M'(\alpha) = \begin{bmatrix} \alpha B & E \\ -E^* & \alpha C \end{bmatrix}, \quad \text{or equivalently,} \quad M(\alpha). \tag{10}$$

As these matrices are positive definite, we may solve the afore-mentioned linear system inexactly by another iteration procedure, e.g., the HSS iteration. This results in an *inexact preconditioned Hermitian/skew-Hermitian splitting* **(IPHSS)** iteration method for the non-Hermitian positive semidefinite system of linear equations (2), and has been previously discussed in [2].

3 Convergence analysis

By straightforward computations, we can obtain an explicit expression of the iteration matrix $\mathcal{L}(\alpha)$ in (8).

Lemma 3.1 *Consider the system of linear equations (2). Let $B \in \mathbb{C}^{p\times p}$ be Hermitian positive definite, $E \in \mathbb{C}^{p\times q}$ be of full column rank, and $\alpha > 0$ a given constant. Assume that $C \in \mathbb{C}^{q\times q}$ is a Hermitian positive definite matrix. Then we partition $\mathcal{L}(\alpha)$ in (8) as*

$$\mathcal{L}(\alpha) = \begin{bmatrix} \mathcal{L}_{11}(\alpha) & \mathcal{L}_{12}(\alpha) \\ \mathcal{L}_{21}(\alpha) & \mathcal{L}_{22}(\alpha) \end{bmatrix},$$

where

$$\begin{aligned} &\mathcal{L}_{11}(\alpha) = \tfrac{\alpha-1}{\alpha+1} I - \tfrac{2}{\alpha+1} B^{-1} E \mathcal{S}(\alpha)^{-1} E^*, \quad \mathcal{L}_{12}(\alpha) = -\tfrac{2\alpha}{\alpha+1} B^{-1} E \mathcal{S}(\alpha)^{-1} C, \\ &\mathcal{L}_{21}(\alpha) = \tfrac{2\alpha}{\alpha+1} C \mathcal{S}(\alpha)^{-1} E^* B^{-1}, \qquad\qquad \mathcal{L}_{22}(\alpha) = -\tfrac{\alpha-1}{\alpha+1} I + \tfrac{2\alpha^2}{\alpha+1} \mathcal{S}(\alpha)^{-1} C, \end{aligned}$$

and

$$\mathcal{S}(\alpha) = \alpha C + \frac{1}{\alpha} E^* B^{-1} E$$

is the Schur complement of the matrix $M'(\alpha)$ in (10).

Proof. Let

$$\begin{aligned} \overline{M}(\alpha) &= P^{-\frac{1}{2}} M(\alpha) P^{-\frac{1}{2}} = \begin{bmatrix} \frac{\alpha+1}{2} I & \frac{\alpha+1}{2\alpha} \bar{E} \\ -\frac{1}{2} \bar{E}^* & \frac{\alpha}{2} I \end{bmatrix}, \\ \overline{N}(\alpha) &= P^{-\frac{1}{2}} N(\alpha) P^{-\frac{1}{2}} = \begin{bmatrix} \frac{\alpha-1}{2} I & -\frac{\alpha-1}{2\alpha} \bar{E} \\ \frac{1}{2} \bar{E}^* & \frac{\alpha}{2} I \end{bmatrix}, \end{aligned}$$

where the matrices P and $\bar{E}$ are defined in (4). Then

$$\overline{M}(\alpha)^{-1} = \begin{bmatrix} \frac{2}{\alpha+1}(I - \frac{1}{\alpha} \bar{E}\, \overline{\mathcal{S}}(\alpha)^{-1} \bar{E}^*) & -\frac{2}{\alpha} \bar{E}\, \overline{\mathcal{S}}(\alpha)^{-1} \\ \frac{2}{\alpha+1} \overline{\mathcal{S}}(\alpha)^{-1} \bar{E}^* & 2\overline{\mathcal{S}}(\alpha)^{-1} \end{bmatrix}$$

and

$$\overline{\mathcal{L}}(\alpha) = \begin{bmatrix} \overline{\mathcal{L}}_{11}(\alpha) & \overline{\mathcal{L}}_{12}(\alpha) \\ \overline{\mathcal{L}}_{21}(\alpha) & \overline{\mathcal{L}}_{22}(\alpha) \end{bmatrix} = \overline{M}(\alpha)^{-1} \overline{N}(\alpha), \tag{11}$$

with

$$\begin{cases} \overline{\mathcal{L}}_{11}(\alpha) = \frac{\alpha-1}{\alpha+1} I - \frac{2}{\alpha+1} \bar{E}\, \overline{\mathcal{S}}(\alpha)^{-1} \bar{E}^*, & \overline{\mathcal{L}}_{12}(\alpha) = -\frac{2\alpha}{\alpha+1} \bar{E}\, \overline{\mathcal{S}}(\alpha)^{-1}, \\ \overline{\mathcal{L}}_{21}(\alpha) = \frac{2\alpha}{\alpha+1} \overline{\mathcal{S}}(\alpha)^{-1} \bar{E}^*, & \overline{\mathcal{L}}_{22}(\alpha) = -\frac{\alpha-1}{\alpha+1} I + \frac{2\alpha^2}{\alpha+1} \overline{\mathcal{S}}(\alpha)^{-1}, \end{cases} \tag{12}$$

where

$$\overline{\mathcal{S}}(\alpha) = \alpha I + \frac{1}{\alpha}\bar{E}^*\bar{E} \tag{13}$$

is the Schur complement of the matrix

$$\overline{M'}(\alpha) = \begin{bmatrix} \alpha I & \bar{E} \\ -\bar{E}^* & \alpha I \end{bmatrix}.$$

Then from (9) we have

$$\mathcal{L}(\alpha) = M(\alpha)^{-1}N(\alpha) = P^{-\frac{1}{2}}\overline{M}(\alpha)^{-1}\overline{N}(\alpha)P^{\frac{1}{2}} = P^{-\frac{1}{2}}\overline{\mathcal{L}}(\alpha)P^{\frac{1}{2}}; \tag{14}$$

the result follows immediately. □

Based on Lemma 3.1, we can further obtain the eigenvalues of the iteration matrix $\mathcal{L}(\alpha)$ of the PHSS method.

Lemma 3.2 *Let the conditions in Lemma 3.1 be satisfied. If* $\overline{\sigma}_k$ $(k = 1, 2, \ldots, q)$ *are the positive singular values of the matrix* $\bar{E} \in \mathbb{C}^{p\times q}$ *in (4), then the eigenvalues of the iteration matrix* $\mathcal{L}(\alpha)$ *of the PHSS iteration method are* $\frac{\alpha-1}{\alpha+1}$ *with multiplicity* $p - q$, *and*

$$\frac{1}{(\alpha+1)(\alpha^2+\overline{\sigma}_k^2)}\left(\alpha(\alpha^2-\overline{\sigma}_k^2) \pm \sqrt{(\alpha^2+\overline{\sigma}_k^2)^2 - 4\alpha^4\overline{\sigma}_k^2}\right), \quad k=1,2,\ldots,q.$$

Proof. From (14) we know that $\mathcal{L}(\alpha)$ is similar to $\overline{\mathcal{L}}(\alpha)$ of (11). Therefore, we only need to compute the eigenvalues of the matrix $\overline{\mathcal{L}}(\alpha)$.

Let $\bar{E} = \overline{U}^*\overline{\Sigma}_1\overline{V}$ be the singular value decomposition[10] of the matrix $\bar{E} \in \mathbb{C}^{p\times q}$, where $\overline{U} \in \mathbb{C}^{p\times p}$ and $\overline{V} \in \mathbb{C}^{q\times q}$ are unitary matrices, and

$$\overline{\Sigma}_1 = \begin{bmatrix} \overline{\Sigma} \\ 0 \end{bmatrix}, \qquad \overline{\Sigma} = \mathrm{diag}(\overline{\sigma}_1, \overline{\sigma}_2, \ldots, \overline{\sigma}_q) \in \mathbb{C}^{q\times q}.$$

Then after a few computation, we have

$$\overline{\mathcal{S}}(\alpha) = \overline{V}^*\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)\overline{V},$$

and therefore,

$$\overline{\mathcal{L}}_{11}(\alpha) = \overline{U}^* \begin{bmatrix} \frac{\alpha-1}{\alpha+1} I - \frac{2}{\alpha+1}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1}\overline{\Sigma}^2 & 0 \\ 0 & \frac{\alpha-1}{\alpha+1} I \end{bmatrix} \overline{U},$$

$$\overline{\mathcal{L}}_{12}(\alpha) = \overline{U}^* \begin{bmatrix} -\frac{2\alpha}{\alpha+1}\overline{\Sigma}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1} \\ 0 \end{bmatrix} \overline{V},$$

$$\overline{\mathcal{L}}_{21}(\alpha) = \overline{V}^* \left[\frac{2\alpha}{\alpha+1}\overline{\Sigma}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1}, \quad 0 \right] \overline{U},$$

$$\overline{\mathcal{L}}_{22}(\alpha) = \overline{V}^* \left(-\frac{\alpha-1}{\alpha+1} I + \frac{2\alpha^2}{\alpha+1}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1} \right) \overline{V}.$$

Define

$$\overline{Q} = \begin{bmatrix} \overline{U} & 0 \\ 0 & \overline{V} \end{bmatrix} \in \mathbb{C}^{n\times n}.$$

Then $\overline{Q}$ is a unitary matrix and it holds that

$$\overline{Q}\,\overline{\mathcal{L}}(\alpha)\overline{Q}^* = \begin{bmatrix} \frac{\alpha-1}{\alpha+1} I - \frac{2}{\alpha+1}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1}\overline{\Sigma}^2 & 0 & -\frac{2\alpha}{\alpha+1}\overline{\Sigma}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1} \\ 0 & \frac{\alpha-1}{\alpha+1} I & 0 \\ \frac{2\alpha}{\alpha+1}\overline{\Sigma}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1} & 0 & -\frac{\alpha-1}{\alpha+1} I + \frac{2\alpha^2}{\alpha+1}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1} \end{bmatrix}.$$

It follows immediately that the eigenvalues of the matrix $\overline{\mathcal{L}}(\alpha)$ are just $\frac{\alpha-1}{\alpha+1}$ with multiplicity $p - q$, and those of the matrix

$$\begin{bmatrix} \frac{\alpha-1}{\alpha+1} I - \frac{2}{\alpha+1}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1}\overline{\Sigma}^2 & -\frac{2\alpha}{\alpha+1}\overline{\Sigma}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1} \\ \frac{2\alpha}{\alpha+1}\overline{\Sigma}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1} & -\frac{\alpha-1}{\alpha+1} I + \frac{2\alpha^2}{\alpha+1}\left(\alpha I + \frac{1}{\alpha}\overline{\Sigma}^2\right)^{-1} \end{bmatrix}$$

which are the same as the matrices $\frac{1}{(\alpha+1)(\alpha^2+\overline{\sigma}_k^2)}\overline{\mathcal{L}}_k(\alpha)$, $k = 1, 2, \ldots, q$, where

$$\overline{\mathcal{L}}_k(\alpha) = \begin{bmatrix} (\alpha-1)\alpha^2 - (\alpha+1)\overline{\sigma}_k^2 & -2\alpha^2\overline{\sigma}_k \\ 2\alpha^2\overline{\sigma}_k & (\alpha+1)\alpha^2 - (\alpha-1)\overline{\sigma}_k^2 \end{bmatrix}.$$

The two eigenvalues of the matrix $\overline{\mathcal{L}}_k(\alpha)$ are the two roots of the quadratic equations

$$\lambda^2 - 2\alpha(\alpha^2 - \overline{\sigma}_k^2)\lambda + (\alpha^2-1)(\alpha^2+\overline{\sigma}_k^2)^2 = 0,$$

or in other words,

$$\lambda = \alpha(\alpha^2 - \overline{\sigma}_k^2) \pm \sqrt{(\alpha^2+\overline{\sigma}_k^2)^2 - 4\alpha^4\overline{\sigma}_k^2}.$$

We know that the eigenvalues of the matrix $\overline{\mathcal{L}}(\alpha)$ are $\frac{\alpha-1}{\alpha+1}$ with multiplicity $p-q$, and

$$\frac{1}{(\alpha+1)(\alpha^2+\overline{\sigma}_k^2)}\left(\alpha(\alpha^2-\overline{\sigma}_k^2)\pm\sqrt{(\alpha^2+\overline{\sigma}_k^2)^2-4\alpha^4\overline{\sigma}_k^2}\right),\ k=1,2,\ldots,q.$$

This completes our proof. □

We remark that the singular values of the matrix $\bar{E}\in\mathbb{C}^{p\times q}$ are exactly the square roots of the eigenvalues of either the matrix $C^{-1}E^*B^{-1}E$, or equivalently, the matrix $E^*B^{-1}EC^{-1}$.

Lemma 3.3 *Let the conditions in Lemma 3.1 be satisfied. If $\overline{\sigma}_k$ $(k=1,2,\ldots,q)$ are the positive singular values of the matrix $\bar{E}\in\mathbb{C}^{p\times q}$ in (4), and λ is a dominant eigenvalue of the iteration matrix $\mathcal{L}(\alpha)$ of the PHSS iteration, i.e., the spectral radius of $\mathcal{L}(\alpha)$ could be achieved at $|\lambda|$, then it holds that*

$$|\lambda|=\begin{cases}\frac{|\alpha-1|}{\alpha+1}, & \text{or}\\ \frac{\alpha}{\alpha+1}\left(\frac{|\alpha^2-\overline{\sigma}_k^2|}{\alpha^2+\overline{\sigma}_k^2}+\sqrt{\frac{1}{\alpha^2}-\frac{4\alpha^2\overline{\sigma}_k^2}{(\alpha^2+\overline{\sigma}_k^2)^2}}\right), & \text{for}\quad \alpha^2+\overline{\sigma}_k^2>2\alpha^2\overline{\sigma}_k, \quad \text{or}\\ \sqrt{\frac{\alpha-1}{\alpha+1}}, & \text{for}\quad \alpha^2+\overline{\sigma}_k^2\le 2\alpha^2\overline{\sigma}_k,\end{cases}$$

for $k=1,2,\ldots,q$.

Proof. According to Lemma 3.2, we know that when $\alpha^2+\overline{\sigma}_k^2>2\alpha^2\overline{\sigma}_k$,

$$\begin{aligned}|\lambda|&=\frac{1}{(\alpha+1)(\alpha^2+\overline{\sigma}_k^2)}\left|\alpha(\alpha^2-\overline{\sigma}_k^2)\pm\sqrt{(\alpha^2+\overline{\sigma}_k^2)^2-4\alpha^4\overline{\sigma}_k^2}\right|\\&=\frac{1}{(\alpha+1)(\alpha^2+\overline{\sigma}_k^2)}\left(\alpha|\alpha^2-\overline{\sigma}_k^2|+\sqrt{(\alpha^2+\overline{\sigma}_k^2)^2-4\alpha^4\overline{\sigma}_k^2}\right)\\&=\frac{\alpha}{\alpha+1}\left(\frac{|\alpha^2-\overline{\sigma}_k^2|}{\alpha^2+\overline{\sigma}_k^2}+\sqrt{\frac{1}{\alpha^2}-\frac{4\alpha^2\overline{\sigma}_k^2}{(\alpha^2+\overline{\sigma}_k^2)^2}}\right),\end{aligned}$$

and when $\alpha^2+\overline{\sigma}_k^2\le 2\alpha^2\overline{\sigma}_k$,

$$\begin{aligned}|\lambda|&=\frac{1}{(\alpha+1)(\alpha^2+\overline{\sigma}_k^2)}\sqrt{\left(\alpha(\alpha^2-\overline{\sigma}_k^2)\right)^2+\left(\sqrt{4\alpha^4\overline{\sigma}_k^2-(\alpha^2+\overline{\sigma}_k^2)^2}\right)^2}\\&=\frac{1}{(\alpha+1)(\alpha^2+\overline{\sigma}_k^2)}\sqrt{(\alpha^2-1)(\alpha^2+\overline{\sigma}_k^2)^2}\\&=\frac{\sqrt{\alpha^2-1}}{\alpha+1}=\sqrt{\frac{\alpha-1}{\alpha+1}}.\end{aligned}$$

The conclusion then follows directly from the structure of the eigenvalues of the matrix $\mathcal{L}(\alpha)$ described in Lemma 3.2. □

Based on Lemma 3.3, we are now ready to prove the convergence of the PHSS iteration method for solving the system of linear equations (2).

Theorem 3.1 *Consider the system of linear equations (2). Let $B \in \mathbb{C}^{p\times p}$ be Hermitian positive definite, $E \in \mathbb{C}^{p\times q}$ be of full column rank, and $\alpha > 0$ a given constant. Assume that $C \in \mathbb{C}^{q\times q}$ is a Hermitian positive definite matrix. Then*

$$\rho(\mathcal{L}(\alpha)) < 1, \qquad \forall \alpha > 0,$$

i.e., the PHSS iteration converges to the exact solution of the system of linear equations (2).

Proof. Obviously, we have

$$\frac{|\alpha - 1|}{\alpha + 1} < 1 \quad (\forall \alpha > 0) \quad \text{and} \quad \sqrt{\frac{\alpha - 1}{\alpha + 1}} < 1 \quad (\forall \alpha > 1).$$

Table 1. α versus ρ for Example 4.1 ($\mu = 1$)

	m	8	16	24	32
HSS	$\alpha_{\exp}$	17.0	28.2	37.8	46.6
	$\rho(\mathcal{M}(\alpha_{\exp}))$	0.9830	0.9938	0.9967	0.9979
PHSS	$\alpha_{\exp}$	1.30	1.66	1.98	2.24
	$\rho(\mathcal{L}(\alpha_{\exp}))$	0.3612	0.4981	0.5735	0.6186
PHSS	α^*	1.415	1.872	2.245	2.566
	$\rho(\mathcal{L}(\alpha^*))$	0.4146	0.5510	0.6194	0.6626

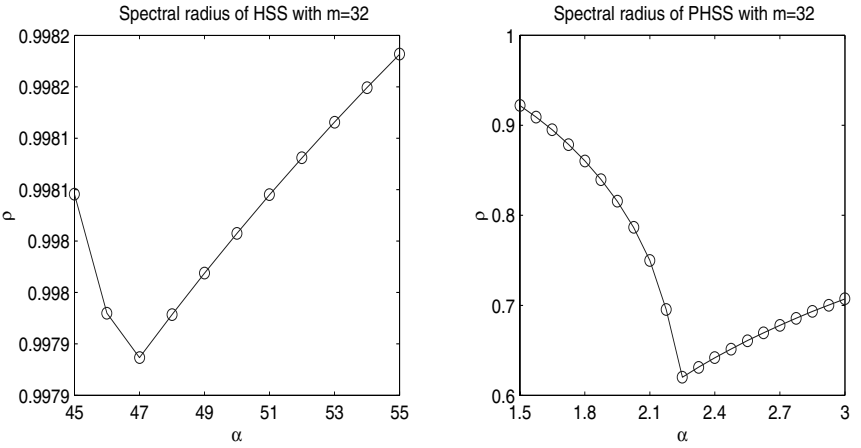

Fig. 1. α versus $\rho(\mathcal{M}(\alpha))$ (left) and $\rho(\mathcal{L}(\alpha))$ (right) when $\mu = 1$

Because for $k = 1, 2, \ldots, q$, when $\alpha^2 + \overline{\sigma}_k^2 > 2\alpha^2\overline{\sigma}_k$, it holds that

$$\begin{aligned}\frac{\alpha}{\alpha+1}\left(\frac{|\alpha^2-\overline{\sigma}_k^2|}{\alpha^2+\overline{\sigma}_k^2}+\sqrt{\frac{1}{\alpha^2}-\frac{4\alpha^2\overline{\sigma}_k^2}{(\alpha^2+\overline{\sigma}_k^2)^2}}\right) &< \frac{\alpha}{\alpha+1}\left(\frac{|\alpha^2-\overline{\sigma}_k^2|}{\alpha^2+\overline{\sigma}_k^2}+\frac{1}{\alpha}\right)\\ &< \frac{\alpha}{\alpha+1}\left(1+\frac{1}{\alpha}\right)\\ &= 1,\end{aligned}$$

by making use of Lemma 3.3 we easily see that $\rho(\mathcal{L}(\alpha)) < 1$ holds for $\forall\alpha > 0$. □

The optimal iteration parameter and the corresponding asymptotic convergence factor of the PHSS iteration method are described in the following theorem.

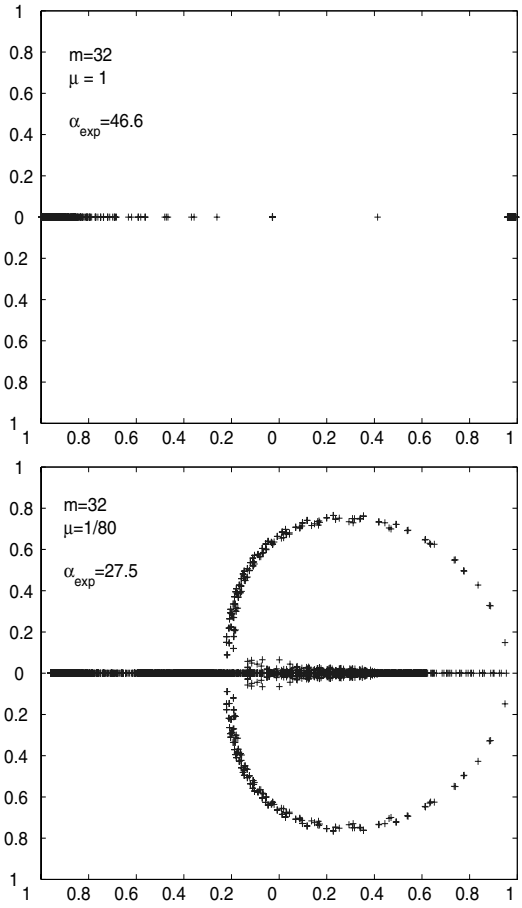

Fig. 2. The spectral distribution of HSS for Example 4.1

Theorem 3.2 *Consider the system of linear equations (2). Let $B \in \mathbb{C}^{p\times p}$ be Hermitian positive definite, $E \in \mathbb{C}^{p\times q}$ be of full column rank, and $\alpha > 0$ a given constant. Assume that $C \in \mathbb{C}^{q\times q}$ is a Hermitian positive definite matrix. If $\sigma_k (k = 1, 2, \ldots, q)$ are the positive singular values of the matrix $B^{-\frac{1}{2}} E C^{-\frac{1}{2}} \in \mathbb{C}^{p\times q}$, and $\sigma_{\min} = \min\limits_{1\le k\le q} \{\sigma_k\}$ and $\sigma_{\max} = \max\limits_{1\le k\le q} \{\sigma_k\}$, then, for the PHSS iteration method of the system of linear equations (2), the optimal value of the iteration parameter α is given by*

$$\alpha^* = \arg\min_{\alpha} \rho(\mathcal{L}(\alpha)) = \sqrt{\sigma_{\min}\sigma_{\max}},$$

and correspondingly,

$$\rho(\mathcal{L}(\alpha^*)) = \frac{\sigma_{\max} - \sigma_{\min}}{\sigma_{\max} + \sigma_{\min}}.$$

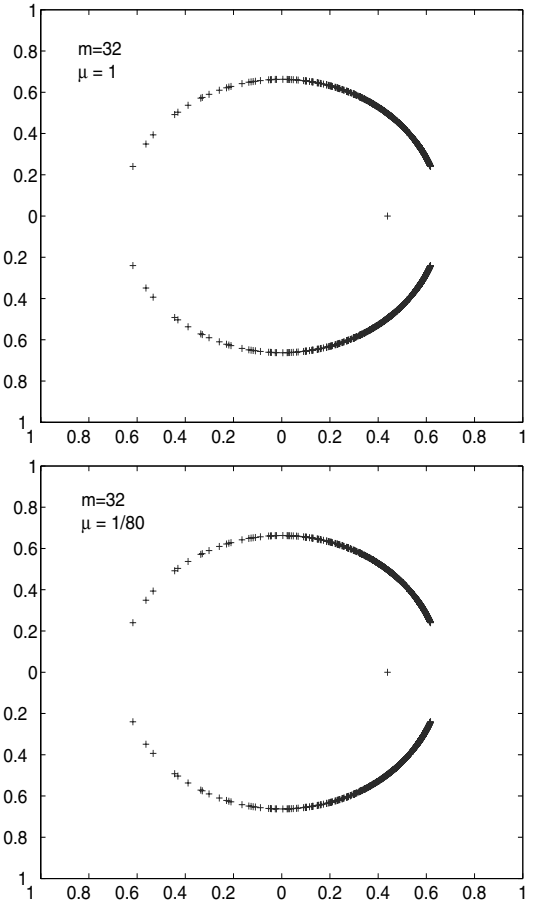

Fig. 3. The spectral distribution of PHSS for Example 4.1

Table 2. IT, CPU and RES for Example 4.1 ($\mu = 1$ and $K = K^{(I)}$)

m		8	16	24	32
Re		4.44444	2.35294	1.6	1.21212
PHSS*	IT	2	2	2	2
	CPU	0.07991	3.10725	29.1659	171.556
PHSS(α^*)	IT	21	31	38	45
	CPU	0.12365	5.38196	58.3807	308.095
UZAWA	IT	46	72	95	117
	CPU	0.11295	5.15491	53.0471	255.453
MINRES	IT	78	163	268	356
	CPU	0.17760	4.60026	65.7691	301.017
PMINRES	IT	82	172	277	380
	CPU	0.33437	13.3612	167.316	817.456
GMRES(20)	IT	–	–	–	–
	CPU	0.69237	31.1923	469.843	2844.59
	RES	4.07E-6	3.54E-6	1.01E-5	3.40E-6
GMRES(100)	IT	65	281	665	1437
	CPU	0.49632	11.3689	175.422	1245.71
GMRES	IT	65	159	258	348
	CPU	0.49161	7.27024	73.1728	318.102
PGMRES(20)	IT	179	–	–	–
	CPU	0.91463	50.5326	922.003	5650.17
	RES		0.09171	0.18437	0.10118
PGMRES(100)	IT	86	231	425	687
	CPU	0.99897	19.5175	249.283	1352.36
PGMRES	IT	86	168	260	357
	CPU	0.99501	16.1059	172.647	837.408

Proof. We first observe that the following two facts hold true:

(F_1) when $\alpha \leq 1$, $\alpha^2 + \sigma_k^2 > 2\alpha^2\sigma_k$, $k = 1, 2, \ldots, q$;
(F_2) when $\alpha > 1$,
(a) $\alpha^2 + \sigma_k^2 > 2\alpha^2\sigma_k$ iff[1] $\sigma_k \in (0, \alpha_-) \cup (\alpha_+, +\infty)$, $k \in \{1, 2, \ldots, q\}$;
(b) $\alpha^2 + \sigma_k^2 \leq 2\alpha^2\sigma_k$ iff $\sigma_k \in [\alpha_-, \alpha_+]$, $k \in \{1, 2, \ldots, q\}$;
(c) $\frac{\alpha-1}{\alpha+1} < \sqrt{\frac{\alpha-1}{\alpha+1}}$,
where $\alpha_- = \alpha^2 - \alpha\sqrt{\alpha^2 - 1}$ and $\alpha_+ = \alpha^2 + \alpha\sqrt{\alpha^2 - 1}$.

Let

$$\theta(\alpha, \sigma) = \frac{\alpha}{\alpha + 1}\left(\frac{|\alpha^2 - \sigma^2|}{\alpha^2 + \sigma^2} + \sqrt{\frac{1}{\alpha^2} - \frac{4\alpha^2\sigma^2}{(\alpha^2 + \sigma^2)^2}}\right).$$

[1] "iff" is used to represent "if and only if"

Table 3. IT, CPU and RES for Example 4.1 ($\mu = \frac{1}{80}$ and $K = K^{(I)}$)

m		8	16	24	32
Re		4.44444	2.35294	1.6	1.21212
PHSS*	IT	2	2	2	2
	CPU	0.084689	3.24394	29.3525	172.718
PHSS(α^*)	IT	23	33	40	46
	CPU	0.139401	6.0266	60.4594	311.438
UZAWA	IT	63	101	136	169
	CPU	0.177551	7.80024	71.9177	340.822
MINRES	IT	95	218	324	428
	CPU	0.226446	7.78713	81.3058	355.773
PMINRES	IT	146	359	591	819
	CPU	0.589088	30.177	327.432	1538.31
GMRES(20)	IT	–	–	1714	2099
	CPU	0.760219	34.302	476.273	1932.63
	RES	0.00127	3.87E-6		
GMRES(100)	IT	94	692	960	1017
	CPU	1.05663	35.2611	258.438	884.403
GMRES	IT	94	248	365	449
	CPU	1.057	16.7705	111.241	416.94
PGMRES(20)	IT	–	–	–	–
	CPU	1.01018	62.225	938.402	5512.51
	RES	8.90E-4	2.10E-4	1.88E-4	5.13E-3
PGMRES(100)	IT	–	–	–	–
	CPU	2.46866	64.9566	931.919	5384.18
	RES	1.11E-8	8.14E-8	1.21E-8	3.54E-8
PGMRES	IT	141	324	538	760
	CPU	2.45599	38.5959	353.267	1611.72

Then based on the facts (F_1) and (F_2) we easily see that

(15)
$$\rho(\mathcal{L}(\alpha)) = \begin{cases} \max\left\{\frac{1-\alpha}{1+\alpha}, \quad \max\limits_{1\le k\le q} \theta(\alpha,\sigma_k)\right\}, & \text{for} \quad \alpha \le 1, \\ \max\left\{\sqrt{\frac{\alpha-1}{\alpha+1}}, \quad \max\limits_{\substack{\sigma_k<\alpha_- \text{ or } \sigma_k>\alpha_+ \\ k\in\{1,2,\ldots,q\}}} \theta(\alpha,\sigma_k)\right\}, & \text{for} \quad \alpha > 1. \end{cases}$$

For any fixed $\beta > 0$, we define two functions $\theta_1, \theta_2 : (0, +\infty) \to (0, +\infty)$ by

$$\theta_1(t) = \frac{\beta - t}{\beta + t}, \qquad \theta_2(t) = \frac{1}{\beta} - \frac{4\beta t}{(\beta + t)^2}.$$

After straightforward computations we obtain

$$\frac{d\theta_1(t)}{dt} = -\frac{2\beta}{(\beta + t)^2}, \qquad \frac{d\theta_2(t)}{dt} = \frac{4\beta(t - \beta)}{(\beta + t)^3}.$$

Table 4. IT, CPU and RES for Example 4.1 ($m = 32$ and $K = K^{(I)}$)

μ		1	1/20	1/40	1/80	1/160	1/1600
Re		0.01515	0.30303	0.60606	1.21212	2.42424	24.2424
PHSS*	IT	2	2	2	2	2	2
	CPU	171.397	173.229	172.429	172.718	173.407	177.104
PHSS(α^*)	IT	45	45	45	46	47	52
	CPU	308.898	310.328	309.926	311.438	314.137	326.847
UZAWA	IT	117	150	160	169	177	201
	CPU	255.138	309.913	326.741	340.822	354.109	392.042
MINRES	IT	356	329	361	428	529	1312
	CPU	297.128	273.486	301.351	355.773	441.568	1093.07
PMINRES	IT	380	650	745	819	900	1193
	CPU	811.96	1263.05	1430.02	1538.31	1675.78	2152.56
GMRES(20)	IT	–	726	944	2099	–	–
	CPU	2959.95	668.274	870.773	1932.63	2815.85	2974.07
	RES	3.40E-6				8.76E-8	5.13E-4
GMRES(100)	IT	1437	498	611	1017	2268	–
	CPU	1253.73	433.621	533.781	884.403	1977.28	2843.8
	RES						2.72E-4
GMRES	IT	348	260	305	449	614	1071
	CPU	317.015	232.84	275.363	416.94	590.06	1130.32
PGMRES(20)	IT	–	–	–	–	–	–
	CPU	5515.38	5661.13	6066.41	5512.51	5490.68	5515.1
	RES	0.1012	0.13445	0.09506	0.00513	6.61E-4	2.90E-4
PGMRES(100)	IT	687	2465	2791	–	–	–
	CPU	1345.25	4347.23	4835.46	5384.18	5357.32	6257.99
	RES				3.54E-8	4.83E-6	2.24E-5
PGMRES	IT	357	674	722	760	790	849
	CPU	832.324	1448.8	1522.54	1611.72	1808.09	1774.56

It then follows that:

(i) $\max\limits_{1\le k\le q} \theta(\alpha, \sigma_k) = \max\{\theta(\alpha, \sigma_{\min}), \quad \theta(\alpha, \sigma_{\max})\}$, for $\alpha \le 1$, and $\sigma_{\min} \le \alpha \le \sigma_{\max}$;

(ii) $\max\limits_{\substack{\sigma_k<\alpha_- \text{ or } \sigma_k>\alpha_+ \\ k\in\{1,2,\ldots,q\}}} \theta(\alpha, \sigma_k) = \max\{\theta(\alpha, \sigma_{\min}), \quad \theta(\alpha, \sigma_{\max})\}$, for $\alpha > 1$, and $\sigma_{\min} < \alpha_-$ or $\sigma_{\max} > \alpha_+$.

Therefore, when $\alpha \le 1$, the optimal parameter α^* must satisfy $\sigma_{\min} \le \alpha^* \le \sigma_{\max}$, and either of the following three conditions:

(A_1) $\frac{1-\alpha^*}{1+\alpha^*} = \theta(\alpha^*, \sigma_{\min}) \ge \theta(\alpha^*, \sigma_{\max})$;
(A_2) $\frac{1-\alpha^*}{1+\alpha^*} = \theta(\alpha^*, \sigma_{\max}) \ge \theta(\alpha^*, \sigma_{\min})$;
(A_3) $\theta(\alpha^*, \sigma_{\min}) = \theta(\alpha^*, \sigma_{\max}) \ge \frac{1-\alpha^*}{1+\alpha^*}$,

and when $\alpha > 1$, the optimal parameter α^* must satisfy $\sigma_{\min} < \alpha_-^*$ or $\sigma_{\max} > \alpha_+^*$, and either of the following three conditions:

Table 5. IT, CPU and RES for Example 4.1 ($\mu = \frac{1}{80}$ and $K = K^{(II)}$)

m		8	16	24	32
Re		4.44444	2.35294	1.6	1.21212
PHSS*	IT	2	2	2	2
	CPU	0.082536	3.1834	29.21	178.831
PHSS(α^*)	IT	23	33	40	46
	CPU	0.135786	6.23739	59.9469	314.107
UZAWA	IT	63	101	136	169
	CPU	0.165623	7.69543	71.4379	342.64
MINRES	IT	95	218	324	428
	CPU	0.221924	7.69403	79.9844	356.747
PMINRES	IT	65	118	166	209
	CPU	0.275354	10.655	113.691	537.217
GMRES(20)	IT	–	–	1714	2099
	CPU	0.726388	32.8369	469.544	1915.71
	RES	0.001269	3.87E-6		
GMRES(100)	IT	94	692	960	1017
	CPU	0.972266	32.2305	255.44	893.764
GMRES	IT	94	248	365	449
	CPU	0.967291	15.839	111.638	458.087
PGMRES(20)	IT	108	178	252	345
	CPU	0.596745	15.953	162.784	782.645
PGMRES(100)	IT	67	131	184	225
	CPU	0.671826	12.874	126.151	567.577
PGMRES	IT	67	121	169	212
	CPU	0.672365	13.2176	119.63	557.015

(B_1) $\sqrt{\frac{\alpha^*-1}{\alpha^*+1}} = \theta(\alpha^*, \sigma_{\min}) \geq \theta(\alpha^*, \sigma_{\max})$;

(B_2) $\sqrt{\frac{\alpha^*-1}{\alpha^*+1}} = \theta(\alpha^*, \sigma_{\max}) \geq \theta(\alpha^*, \sigma_{\min})$;

(B_3) $\theta(\alpha^*, \sigma_{\min}) = \theta(\alpha^*, \sigma_{\max}) \geq \sqrt{\frac{\alpha^*-1}{\alpha^*+1}}$,

where $\alpha^*_- = (\alpha^*)^2 - \alpha^*\sqrt{(\alpha^*)^2 - 1}$ and $\alpha^*_+ = (\alpha^*)^2 + \alpha^*\sqrt{(\alpha^*)^2 - 1}$. By straightforwardly solving the inequalities (A_1)-(A_3) and (B_1)-(B_3), we obtain the concrete expression of α^*. Then by substituting α^* into (15), we obtain $\rho(\mathcal{L}(\alpha^*))$. □

From Theorem 3.2 we see that the Hermitian positive definite matrix $C \in \mathbb{C}^{q\times q}$ should be chosen such that the linear system with coefficient matrix C is easily solvable and the singular values of the matrix $B^{-\frac{1}{2}}EC^{-\frac{1}{2}} \in \mathbb{C}^{p\times q}$ are tightly clustered, or in other words, C should be a good preconditioner to the matrix $E^*B^{-1}E \in \mathbb{C}^{q\times q}$. Finding a good matrix C is fundamental to accelerate the convergence speed of the PHSS iteration method. According to the computation of the optimal parameter $\alpha^* = \sqrt{\sigma_{\min}\sigma_{\max}}$ in actual appli-

Table 6. IT, CPU and RES for Example 4.1 ($m = 32$ and $K = K^{(II)}$)

μ		1	1/20	1/40	1/80	1/160	1/1600
Re		0.01515	0.30303	0.60606	1.21212	2.42424	24.2424
PHSS*	IT	2	2	2	2	2	2
	CPU	181.89	178.584	185.672	178.831	175.69	181.115
PHSS(α^*)	IT	45	45	45	46	47	52
	CPU	319.35	312.553	318.332	314.107	317.886	328.123
UZAWA	IT	117	150	160	169	177	201
	CPU	257.461	309.91	328.605	342.64	355.508	478.421
MINRES	IT	356	329	361	428	529	1312
	CPU	301.564	274.497	301.299	356.747	439.56	1093.27
PMINRES	IT	207	210	210	209	209	214
	CPU	535.387	537.073	541.215	537.217	539.828	546.914
GMRES(20)	IT	–	726	944	2099	–	–
	CPU	2992.2	670.1	869.523	1915.71	2840.78	2831.26
	RES	3.40E-6				8.76E-8	5.13E-4
GMRES(100)	IT	1437	498	611	1017	2268	–
	CPU	1274.53	435.03	533.364	893.764	2009.13	2682.07
	RES						2.72E-4
GMRES	IT	348	260	305	449	614	1071
	CPU	324.33	235.196	276.626	458.087	595.624	1259.71
PGMRES(20)	IT	319	327	331	345	345	432
	CPU	732.428	756.61	765.343	782.645	787.287	939.448
PGMRES(100)	IT	252	225	222	225	231	276
	CPU	610.644	564.61	577.20	567.577	581.029	657.516
PGMRES	IT	215	210	210	212	219	238
	CPU	584.928	591.877	570.831	557.015	590.012	613.175

Table 7. α versus ρ for Example 4.2 ($\ell = 1$)

	m	8	16	24	32	40
HSS	$\alpha_{\exp}$	3.25	4.625	5.875	7.00	7.75
	$\rho(\mathcal{M}(\alpha_{\exp}))$	0.9903	0.9977	0.9990	0.9995	0.9997
PHSS	$\alpha_{\exp}$	1.25	1.25	1.25	1.50	1.50
	$\rho(\mathcal{L}(\alpha_{\exp}))$	0.3333	0.3333	0.3333	0.4472	0.4472
PHSS	α^*	1.228	1.448	1.575	1.702	1.829
	$\rho(\mathcal{L}(\alpha^*))$	0.3200	0.4279	0.4727	0.5098	0.5412

cations, because the smallest and the largest singular values $\sigma_{\min}$ and $\sigma_{\max}$ of the matrix $B^{-\frac{1}{2}}EC^{-\frac{1}{2}}$ are equal to the square roots of the smallest and the largest eigenvalues $\lambda_{\min}$ and $\lambda_{\max}$ of the matrix $C^{-1}E^*B^{-1}E$, respectively, i.e., $\sigma_{\min} = \sqrt{\lambda_{\min}}$ and $\sigma_{\max} = \sqrt{\lambda_{\max}}$, we can obtain $\alpha^* = \sqrt[4]{\lambda_{\min}\lambda_{\max}}$ through computing $\lambda_{\min}$ and $\lambda_{\max}$.

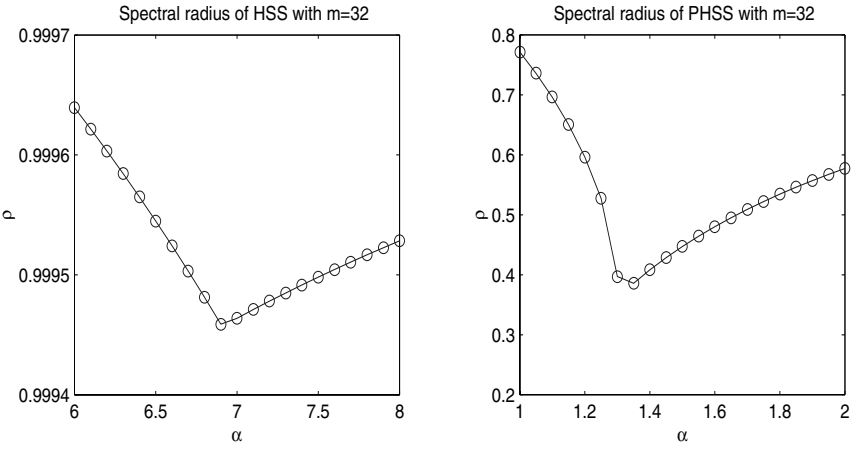

Fig. 4. α versus $\rho(\mathcal{M}(\alpha))$ (left) and $\rho(\mathcal{L}(\alpha))$ (right) when $\ell = 1$

4 Numerical examples

In this section, we use several examples to further examine the effectiveness, and show the advantages of the PHSS method over the HSS method, as well as some typical methods for solving the system of linear equations (2) such as UZAWA[1,4,10] iteration, MINRES[12], GMRES and GMRES($\sharp$)[14] methods and their preconditioned variants **PMINRES**, **PGMRES** and **PGMRES**($\sharp$), from aspects of spectral radius $\rho(\cdot)$ of iteration matrix, number of total iteration steps (denoted by "IT"), and elapsed CPU time in seconds (denoted by "CPU"). Here, the integer $\sharp$ in GMRES($\sharp$) and PGMRES($\sharp$) denotes the number of restarting steps.

To this end, we need to choose the matrix C in PHSS, and suitable preconditioners for PMINRES, PGMRES and PGMRES($\sharp$). There are two natural choices of the matrix C: the first is $C = E^*B^{-1}E$, and the second is $C = E^*\hat{B}^{-1}E$, where $\hat{B}$ is a good approximation to the matrix block B. We remark that the theoretical optimal parameter for the first case is given by $\alpha^* = 1$ which makes PHSS become a direct method, and we particularly denote this PHSS by PHSS*. According to the preconditioner, say K, we also have two typical choices[6,7,13,16]: the first is

$$K \equiv K^{(I)} = \begin{bmatrix} \hat{B} & 0 \\ 0 & I \end{bmatrix},$$

and the second is

$$K \equiv K^{(II)} = \begin{bmatrix} \hat{B} & 0 \\ 0 & C \end{bmatrix}, \quad \text{with} \quad C = E^*\hat{B}^{-1}E.$$

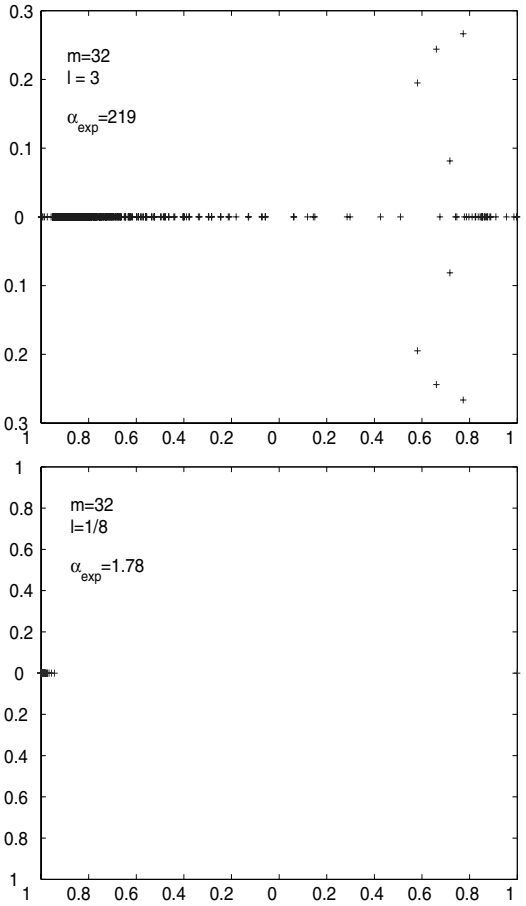

Fig. 5. The spectral distribution of HSS for Example 4.2

We should mention that in our numerical implementations, UZAWA, MINRES and PMINRES are actually applied to the equivalently transformed block two-by-two linear system

$$A \begin{bmatrix} y \\ z \end{bmatrix} \equiv \begin{bmatrix} B & E \\ E^* & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} f \\ -g \end{bmatrix} \equiv b.$$

To give intuition about the asymptotic convergence rates of HSS and PHSS, we compute the experimental optimal parameter α (denoted by "$\alpha_{\exp}$") and the corresponding spectral radii of the iteration matrices $\mathcal{M}(\alpha_{\exp})$ and $\mathcal{L}(\alpha_{\exp})$ of the HSS and the PHSS iterations, as well as the theoretical optimal parameter α (denoted by α^*) and the corresponding spectral radius of the iteration matrix $\mathcal{L}(\alpha^*)$ of the PHSS iteration (see Theorem 3.2). Here, we take $C = E^*\hat{B}^{-1}E$ in PHSS.

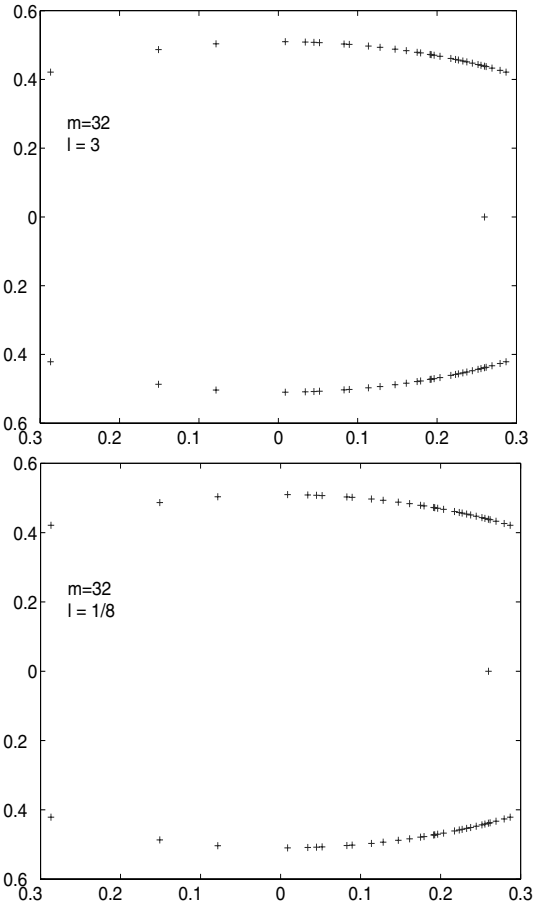

Fig. 6. The spectral distribution of PHSS for Example 4.2

In order to emphasize the dependence of the PHSS iteration upon the acceleration parameter α, we sometimes also use the notation PHSS(α) instead of PHSS.

In actual computations, we choose the right-hand-side vector b such that the exact solution of the system of linear equations (2) is $(1, 1, \ldots, 1)^T \in \mathbb{R}^n$. In addition, all runs are initiated with the initial vector $x^{(0)} = 0$, terminated if the current iterations satisfy

$$\text{RES} \equiv \frac{\|b - Ax^{(k)}\|_2}{\|b - Ax^{(0)}\|_2} \leq 10^{-8}$$

or if the numbers of the prescribed iteration $k_{\max} = n$ are exceeded, and performed in MATLAB (version 6.1) with a machine precision 10^{-16}. The machine used is a Pentium-III 500 personal computer with 256M memory.

Table 8. IT, CPU and RES for Example 4.2 ($m = 48$ and $K = K^{(I)}$)

ℓ		1/8	1/5	1/2	1	3
PHSS*	IT	2	2	2	2	2
	CPU	65.1285	65.1653	65.91	66.8038	70.6504
PHSS(α^*)	IT	33	33	33	33	37
	CPU	123.409	124.256	123.983	125.277	131.695
UZAWA	IT	83	131	962	–	–
	CPU	151.206	195.229	1034.93	2251.12	2189.58
	RES				5.97E-6	3.5764
MINRES	IT	397	462	838	2233	–
	CPU	181.325	209.816	378.351	1024.76	1048.91
	RES					7.86E-8
PMINRES	IT	296	331	562	593	–
	CPU	349.963	383.387	593.278	626.536	2256.32
	RES					9.00E-6
GMRES(20)	IT	–	–	–	–	–
	CPU	1186.13	1187.07	1183.66	1198.32	1223.81
	RES	1.17E-7	1.94E-7	1.60E-6	3.12E-6	2.61E-4
GMRES(100)	IT	–	–	–	–	–
	CPU	1132.14	1134.47	1133.98	1138.5	1163.07
	RES	4.30E-8	3.91E-8	6.97E-8	6.93E-7	5.64E-7
GMRES	IT	371	416	634	1111	404
	CPU	193.524	220.94	360.721	743.155	229.986
PGMRES(20)	IT	–	–	–	–	–
	CPU	2348.5	2349.06	2352.63	2415.36	2412.77
	RES	9.77E-3	0.01431	0.03473	7.82E-8	0.1421
PGMRES(100)	IT	901	1047	1322	1329	–
	CPU	948.134	1088.35	1356.21	1404.89	2351.7
	RES					1.68E-5
PGMRES	IT	374	402	479	299	197
	CPU	475.038	496.569	597.664	400.88	280.738

The numerical results clearly show that PHSS* always converges very faster than the other methods in terms of both IT and CPU, because it is substantially a direct method.

According to PHSS(α^*), we see that it is more robust than UZAWA, and MINRES, GMRES, GMRES(♯) as well as their preconditioned variants, because it can always successfully produce an accurate approximation to the exact solution of the system of linear equations (2), while UZAWA and the abovementioned Krylov subspace methods may be breakdown, may be divergent, or may be unable to attain the stopping criterion within the largest admissible number of iteration steps. For the convergence cases, PHSS(α^*) always shows smaller number of iteration steps than the other testing methods. However, the computing time of PHSS(α^*) is, roughly speaking, comparable

Table 9. IT, CPU and RES for Example 4.2 ($m = 48$ and $K = K^{(II)}$)

ℓ		1/8	1/5	1/2	1	3
PHSS*	IT	2	2	2	2	2
	CPU	65.6766	65.4644	66.2715	65.8774	65.0591
PHSS(α^*)	IT	33	33	33	33	37
	CPU	124.8	124.991	125.149	123.423	126.704
UZAWA	IT	83	131	962	–	–
	CPU	151.786	196.564	959.969	2196.48	2192.05
	RES				5.97E-6	3.5764
MINRES	IT	397	462	838	2233	–
	CPU	183.855	216.571	381.613	1002.48	1068.27
	RES					7.86E-8
PMINRES	IT	143	143	143	143	135
	CPU	212.416	214.127	212.909	209.366	202.441
GMRES(20)	IT	–	–	–	–	–
	CPU	1203.0	1209.65	1188.81	1176.71	1177.99
	RES	1.17E-7	1.94E-7	1.60E-6	3.12E-6	2.61E-4
GMRES(100)	IT	–	–	–	–	–
	CPU	1147.72	1149.35	1136.65	1119.51	1122.18
	RES	4.30E-8	3.90E-8	6.97E-8	6.93E-7	5.64E-7
GMRES	IT	371	416	634	1111	404
	CPU	196.632	224.835	370.987	712.478	210.526
PGMRES(20)	IT	–	–	–	311	191
	CPU	2370.22	2374.16	2347.99	376.944	263.938
	RES	0.06356	0.06315	0.05952		
PGMRES(100)	IT	166	166	166	166	143
	CPU	239.12	238.959	238.84	236.252	214.846
PGMRES	IT	146	146	146	146	–
	CPU	224.254	225.235	228.736	222.856	5138.72
	RES					2.14E-8

to UZAWA and MINRES, and sometimes, even to GMRES, for Example 4.1, and it is less than the computing times of UZAWA, MINRES and GMRES for Examples 4.2 and 4.3; the reason could be that PHSS(α^*) is less efficient to exploit the special structures of the matrices, and its cost may be larger than the costs of the other testing methods at each iteration step. Obviously, Example 4.1 is sparser and more structured than both Examples 4.2 and 4.3. For all these three examples, the computing time of PHSS(α^*) is less than those of PMINRES, PGMRES, GMRES(♯) and PGMRES(♯); this could be because the preconditioning process costs time, and the best preconditioners and the optimal restarting numbers are not easily obtainable for these examples. Therefore, PHSS(α^*) is a competitive iteration method for solving the large sparse system of linear equations (2), in particular, when some efficient inner iterations are technically incorporated in it to obtain economical, stable and fast inexact iteration solvers.

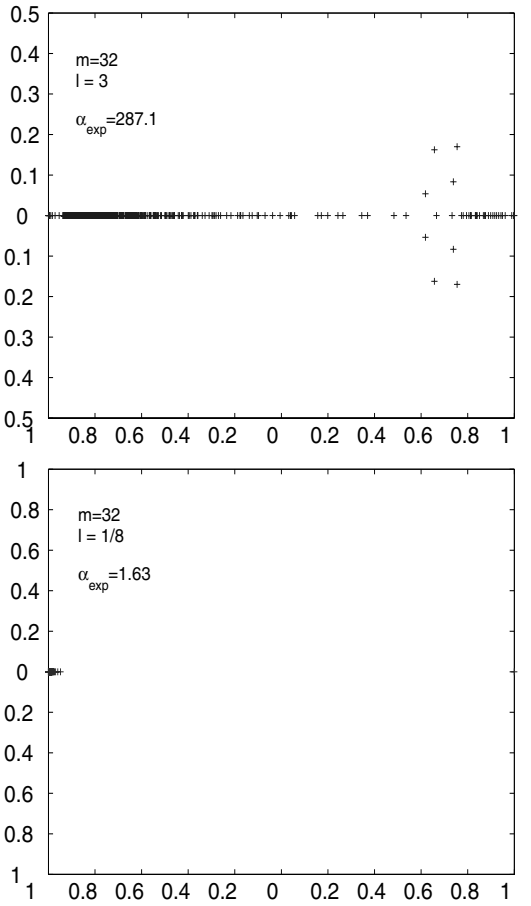

Fig. 7. The spectral distribution of HSS for Example 4.3

Example 4.1 Consider the Stokes problem: Find $\mathbf{u}$ and w such that

$$\begin{cases} -\mu\,\triangle\mathbf{u} + \nabla w = \tilde{f}, & \text{in } \Omega, \\ \nabla\cdot\mathbf{u} = \tilde{g}, & \text{in } \Omega, \\ \mathbf{u} = 0, & \text{on } \partial\Omega, \\ \displaystyle\int_{\Omega} w(x)dx = 0, \end{cases} \tag{16}$$

where $\Omega = (0,1)\times(0,1) \subset \mathbb{R}^2$, $\partial\Omega$ is the boundary of Ω, $\triangle$ is the componentwise Laplace operator, $\mathbf{u}$ is a vector-valued function representing the velocity, and w is a scalar function representing the pressure. By discretizing (16) with the upwind scheme, we obtain the system of linear equations (2),

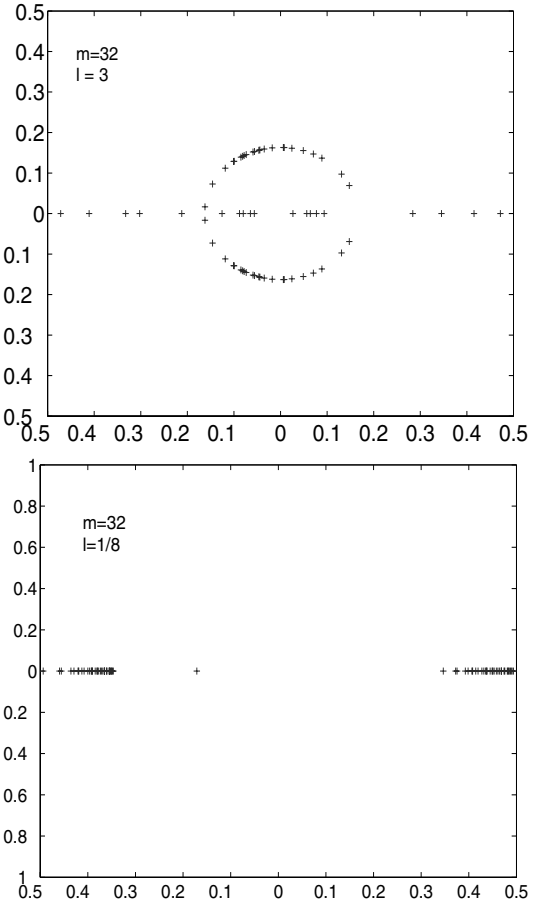

Fig. 8. The spectral distribution of PHSS for Example 4.3

in which

$$B = \begin{bmatrix} I \otimes T + T \otimes I & 0 \\ 0 & I \otimes T + T \otimes I \end{bmatrix} \in \mathbb{R}^{2m^2 \times 2m^2},$$

$$E = \begin{bmatrix} I \otimes F \\ F \otimes I \end{bmatrix} \in \mathbb{R}^{2m^2 \times m^2},$$

and

$$T = \frac{\mu}{h^2} \cdot \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{m \times m},$$

$$F = \frac{1}{h} \cdot \text{tridiag}(-1, 1, 0) \in \mathbb{R}^{m \times m},$$

with $h = \frac{1}{m+1}$ being the discretization meshsize, $\otimes$ the Kronecker product symbol, and $Re = \frac{h}{2\mu}$.

For this example, we take $\hat{B} = D_B$, the block diagonal matrix of B. We remark that the number of variables is $n = 3m^2$.

In Table 1, we list $\alpha_{\exp}$, and the corresponding $\rho(\mathcal{M}(\alpha_{\exp}))$ and $\rho(\mathcal{L}(\alpha_{\exp}))$ of the HSS and the PHSS iterations, as well as α^* and the corresponding $\rho(\mathcal{L}(\alpha^*))$ of the PHSS iteration, for various meshsizes h, or in other words, m. It is clear that the asymptotic convergence rate of PHSS is much faster than that of HSS. However, due to round-off error, $\alpha_{\exp}$ always results in smaller convergence factor than α^* does.

Figure 1 depicts the curves of the spectral radii versus the parameter α of HSS and PHSS when $m = 32$ and $\mu = 1$. Again, we see that PHSS possesses faster convergence speed than HSS.

In Fig. 2 and 3 we respectively display the eigenvalues of the iteration matrices $\mathcal{M}(\alpha_{\exp})$ and $\mathcal{L}(\alpha^*)$ of the HSS and the PHSS iterations in the cases of $\mu = 1$ and $\mu = \frac{1}{80}$ when $m = 32$. Note that the spectral radius of the iteration matrix $\mathcal{L}(\alpha^*)$ is much smaller than that of $\mathcal{M}(\alpha_{\exp})$.

In Tables 2-6, we list numerical results with respect to IT, CPU or RES (if the convergence criterion is not achieved within $k_{\max}$ iteration steps) for the testing methods. In particular, we use $K = K^{(I)}$ as preconditioner to PMINRES, PGMRES and PGMRES($\sharp$) in Tables 2-4, and $K = K^{(II)}$ in Tables 5-6.

Tables 2-3 list numerical results for a fixed μ ($\mu = 1$ for Table 2 and $\mu = \frac{1}{80}$ for Table 3, respectively) and varying m, while Table 4 lists those for a fixed $m = 32$ and varying μ, when the preconditioner $K = K^{(I)}$ is employed to PMINRES, PGMRES and PGMRES($\sharp$). In addition, Table 5 list numerical results for a fixed $\mu = \frac{1}{80}$ and varying m, while Table 6 lists those for a fixed $m = 32$ and varying μ, when the preconditioner $K = K^{(II)}$ is employed to PMINRES, PGMRES and PGMRES($\sharp$).

Comparisons between Table 3 and Table 5, and Table 4 and Table 6 show that $K^{(II)}$ is more efficient than $K^{(I)}$ when they are applied to precondition the Krylov subspace methods for solving this class of linear systems.

Example 4.2 Consider the system of linear equations (2), in which the matrix block $B \in \mathbb{R}^{m^2 \times m^2}$ is the five-point difference matrix of the two-dimensional Laplace operator, and $E = U\Lambda V$ with $U \in \mathbb{R}^{m^2 \times m}$ being a column orthogonal matrix, $V \in \mathbb{R}^{m \times m}$ being an orthogonal matrix, $\Lambda = \mathrm{diag}(1, 2^\ell, \ldots, m^\ell) \in \mathbb{R}^{m \times m}$ and ℓ a given positive real number. Here, U and V are generated randomly with normal distribution by the MATLAB code *randn*.

This example could be considered as a KKT system corresponding to a least squares problem with linear constraints. For this example, we also take

$\hat{B} = D_B$, the block diagonal matrix of B. We remark that the number of variables is $n = 2m^2$.

In Table 7, we list $\alpha_{\exp}$, and the corresponding $\rho(\mathcal{M}(\alpha_{\exp}))$ and $\rho(\mathcal{L}(\alpha_{\exp}))$ of the HSS and the PHSS iterations, as well as α^* and the corresponding $\rho(\mathcal{L}(\alpha^*))$ of the PHSS iteration, for various meshsizes h, or in other words, m. It is clear that the asymptotic convergence rate of PHSS is much faster than that of HSS. However, due to round-off error, $\alpha_{\exp}$ always results in smaller convergence factor than α^* does.

Figure 4 depicts the curves of the spectral radii versus the parameter α of HSS and PHSS when $m = 32$ and $\ell = 1$. Again, we see that PHSS possesses faster convergence speed than HSS.

In Fig. 5 and 6 we respectively display the eigenvalues of the iteration matrices $\mathcal{M}(\alpha_{\exp})$ and $\mathcal{L}(\alpha^*)$ of the HSS and the PHSS iterations in the cases of $\ell = 3$ and $\ell = \frac{1}{8}$ when $m = 32$. Note that the spectral radius of the iteration matrix $\mathcal{L}(\alpha^*)$ is considerably smaller than that of $\mathcal{M}(\alpha_{\exp})$.

In Tables 8 and 9, we list numerical results with respect to IT, CPU or RES (when the convergence criterion is not achieved within $k_{\max}$ iteration steps) for the testing methods. In particular, we use $K = K^{(I)}$ in Table 8 and $K = K^{(II)}$ in Table 9 as preconditioners to PMINRES, PGMRES and PGMRES($\sharp$), respectively.

Example 4.3 Consider the system of linear equations (2), in which the matrix block $B = \tilde{B} + \gamma\, EE^T \in \mathbb{R}^{m^2 \times m^2}$ with $\tilde{B}$ being the five-point difference matrix of the two-dimensional Laplace operator, $\gamma = \frac{\|D_{\tilde{B}}\|_2}{\|E\|_2^2}$ being the balancing number and $D_{\tilde{B}}$ the block diagonal matrix of $\tilde{B}$; and $E = U\Lambda V$ with $U \in \mathbb{R}^{m^2 \times m}$ being a column orthogonal matrix, $V \in \mathbb{R}^{m \times m}$ being an orthogonal matrix, $\Lambda = \mathrm{diag}(1, 2^\ell, \ldots, m^\ell) \in \mathbb{R}^{m \times m}$ and ℓ a given positive real number. Here, U and V are generated randomly with normal distribution by the MATLAB code randn.

This example could be considered as a KKT system from the augmented Lagrangian approach of a least squares problem with linear constraints. For this example, we first take $\hat{B} = D_{\tilde{B}}$ (in Table 10) , and then $\hat{B} = D_{\tilde{B}} + \gamma\ \mathrm{diag}(EE^T)$ (in Tables 11 and 12). The numerical results corresponding to these two cases are listed in Tables 10-12, respectively. We remark that the number of variables is $n = 2m^2$.

In Fig. 7 and 8 we respectively display the eigenvalues of the iteration matrices $\mathcal{M}(\alpha_{\exp})$ and $\mathcal{L}(\alpha^*)$ of the HSS and the PHSS iterations in the cases of $\ell = 1$ and $\ell = \frac{1}{8}$ when $m = 32$ and $\hat{B} = D_{\tilde{B}} + \gamma\ \mathrm{diag}(EE^T)$. Note that the spectral radius of the iteration matrix $\mathcal{L}(\alpha^*)$ is considerably smaller than that of $\mathcal{M}(\alpha_{\exp})$.

Table 10. IT, CPU and RES for Example 4.3 ($\ell = 3$ and $K = K^{(I)}$)

m		8	16	24	32	40	48
γ		0.0024	0.00014	2.61E-5	8.10E-6	3.28E-6	1.57E-6
PHSS*	IT	2	2	2	2	2	2
	CPU	0.009543	0.223665	1.76583	8.58689	27.2725	76.8019
PHSS(α^*)	IT	34	40	39	55	45	52
	CPU	0.014443	0.392779	3.49424	19.9228	53.2951	146.403
UZAWA	IT	–	–	–	–	–	–
	CPU	0.029453	1.63741	26.1196	179.791	745.104	2233.81
	RES	1.40708	1.05318	0.41129	0.92124	0.06156	0.4821
MINRES	IT	72	237	582	919	1600	–
	CPU	0.077291	0.827095	13.5156	87.3604	359.809	1142.93
	RES						7.81E-8
PMINRES	IT	23	71	135	250	388	646
	CPU	0.0279	0.565633	8.03456	56.9218	202.878	794.881
GMRES(20)	IT	–	–	–	–	–	–
	CPU	0.151788	1.37608	15.4654	111.168	409.29	1233.58
	RES	2.28E-4	8.37E-5	2.93E-4	9.12E-5	1.06E-4	1.82E-4
GMRES(100)	IT	57	–	–	–	–	–
	CPU	0.300949	3.21144	19.1324	114.099	399.165	1178.38
	RES		7.48E-8	1.23E-7	5.67E-7	2.03E-7	5.04E-7
GMRES	IT	57	132	215	257	328	403
	CPU	0.302064	2.06646	9.74037	32.6371	90.3356	221.457
PGMRES(20)	IT	17	–	–	–	–	–
	CPU	0.041317	2.10002	31.4317	218.74	795.149	2378.11
	RES		0.09668	0.15299	0.08502	0.07034	0.08959
PGMRES(100)	IT	17	33	49	65	81	97
	CPU	0.041203	0.439273	4.02187	20.6805	65.4712	177.979
PGMRES	IT	17	33	49	65	81	97
	CPU	0.041215	0.441574	3.81579	20.6276	65.4895	212.651

5 A further generalization and concluding remarks

More generally, we consider the block two-by-two system of linear equations (2), where $B \in \mathbb{C}^{p\times p}$ is Hermitian, but is now only required to be positive semidefinite. In this case, we know that the coefficient matrix

$$A = \begin{bmatrix} B & E \\ -E^* & 0 \end{bmatrix}$$

is nonsingular if and only if

$$\begin{cases} \text{rank}(E) = q, \\ \text{null}(B) \cap \text{null}(E^*) = \{0\}, \end{cases} \tag{17}$$

Table 11. IT, CPU and RES for Example 4.3 ($\ell = 3$ and $K = K^{(I)}$)

m		8	16	24	32	40	48
γ		0.002397	1.37E-4	2.61E-5	8.10E-6	3.28E-6	1.57E-6
PHSS*	IT	2	2	2	2	2	2
	CPU	0.008943	0.224577	1.78463	8.37769	27.2815	74.771
PHSS(α^*)	IT	29	33	35	40	40	43
	CPU	0.013059	0.373465	3.4772	16.9336	51.128	132.912
UZAWA	IT	–	–	–	–	–	–
	CPU	0.030552	1.60072	25.895	175.803	746.558	2213.46
	RES	1.40708	1.05318	0.41129	0.9212	0.0616	0.4821
MINRES	IT	72	237	582	919	1600	–
	CPU	0.075279	0.89353	14.3642	84.334	359.405	1070.84
	RES						7.81E-8
PMINRES	IT	39	106	245	322	529	851
	CPU	0.043463	0.861695	14.3512	69.6912	280.163	898.92
GMRES(20)	IT	–	–	–	–	–	–
	CPU	0.1512	1.37837	16.4289	108.176	408.834	1181.41
	RES	2.28E-4	8.37E-5	2.93E-4	9.12E-5	1.06E-4	1.82E-4
GMRES(100)	IT	57	–	–	–	–	–
	CPU	0.302395	3.20822	20.4787	112.11	399.333	1129.2
	RES		7.48E-8	1.23E-7	5.67E-7	2.03E-7	5.04E-7
GMRES	IT	57	132	215	257	328	403
	CPU	0.301996	2.02994	9.91605	32.1331	90.4633	214.495
PGMRES(20)	IT	55	–	–	–	–	–
	CPU	0.141147	2.23017	31.5461	216.782	796.187	2333.17
	RES		0.111914	0.1686	0.09104	0.08306	0.1001
PGMRES(100)	IT	24	40	58	74	90	293
	CPU	0.071506	0.537747	4.51297	22.7208	71.0793	463.723
PGMRES	IT	24	40	58	74	90	107
	CPU	0.071113	0.557394	4.50929	22.7856	71.103	225.466

i.e., the submatrix E has full column rank, and the null spaces of the submatrices B and E^* have only intersection at the origin. Under these two conditions, an analogous derivation to the PHSS iteration formula (6), incorporated with matrix perturbation technique, immediately leads to the iteration scheme which is of the same form as (6) for this generalized block two-by-two system of linear equations. Moreover, theoretical analysis similar to Theorems 3.1 and 3.2, also incorporated with matrix perturbation technique, straightforwardly results in the following convergence theorem of the PHSS iteration method for solving the generalized block two-by-two system of linear equations.

Theorem 5.1 *Consider the system of linear equations (2). Let $B \in \mathbb{C}^{p\times p}$ be Hermitian positive semidefinite and $E \in \mathbb{C}^{p\times q}$ be of full column rank satisfying (17), and $\alpha > 0$ a given constant. Assume that $C \in \mathbb{C}^{q\times q}$ is a Hermitian*

Table 12. IT, CPU and RES for Example 4.3 ($\ell = 3$ and $K = K^{(II)}$)

m		8	16	24	32	40	48
γ		2.40E-3	1.37E-4	2.61E-5	8.10E-6	3.28E-6	1.57E-6
PHSS*	IT	2	2	2	2	2	2
	CPU	0.009187	0.223683	1.7876	8.35953	27.2458	74.9208
PHSS(α^*)	IT	29	33	35	40	40	43
	CPU	0.013687	0.376879	3.40044	17.0004	51.1166	134.748
UZAWA	IT	–	–	–	–	–	–
	CPU	0.028926	1.65954	25.2599	176.793	749.659	2233.14
	RES	1.4071	1.0532	0.4113	0.9212	0.06156	0.4821
MINRES	IT	72	237	582	919	1600	–
	CPU	0.077162	0.926572	13.8495	87.0291	360.521	1087.09
	RES						7.81E-8
PMINRES	IT	25	37	49	60	71	81
	CPU	0.032469	0.415089	3.80665	20.4088	63.5426	174.001
GMRES(20)	IT	–	–	–	–	–	–
	CPU	0.149882	1.39643	16.7432	110.525	409.392	1202.19
	RES	2.28E-4	8.37E-5	2.93E-4	9.12E-5	1.06E-4	1.82E-4
GMRES(100)	IT	57	–	–	–	–	–
	CPU	0.301634	3.22465	20.5963	113.573	399.722	1150.02
	RES		7.48E-8	1.23E-7	5.67E-7	2.03E-7	5.04E-7
GMRES	IT	57	132	215	257	328	403
	CPU	0.297637	2.03342	9.9284	32.5943	90.3452	216.455
PGMRES(20)	IT	36	69	90	107	116	135
	CPU	0.094835	0.948188	7.46312	35.1844	99.6631	260.322
PGMRES(100)	IT	25	34	45	54	61	115
	CPU	0.077722	0.503869	4.18165	20.7028	63.3704	238.675
PGMRES	IT	25	34	45	54	61	–
	CPU	0.078288	0.505027	4.1522	20.645	63.2949	45622.9
	RES						4.32E-8

positive semidefinite matrix. Denote by $\sigma_k (k = 1, 2, \ldots, q)$ *the positive singular values of the matrix* $(B^\dagger)^{\frac{1}{2}} E (C^\dagger)^{\frac{1}{2}} \in \mathbb{C}^{p\times q}$, *and* $\sigma_{\min} = \min_{1\le k\le q} \{\sigma_k\}$ *and* $\sigma_{\max} = \max_{1\le k\le q} \{\sigma_k\}$, *where* $(\cdot)^\dagger$ *represents the Moore-Penrose generalized inverse of the corresponding matrix. Then*

(i) $\rho(\mathcal{L}(\alpha)) < 1$, $\forall \alpha > 0$, *i.e., the PHSS iteration converges to the exact solution of the generalized block two-by-two system of linear equations;*

(ii) the optimal value of the iteration parameter α *is given by*

$$\alpha^* = \arg\min_{\alpha} \rho(\mathcal{L}(\alpha)) = \sqrt{\sigma_{\min}\sigma_{\max}},$$

and correspondingly,

$$\rho(\mathcal{L}(\alpha^*)) = \frac{\sigma_{\max} - \sigma_{\min}}{\sigma_{\max} + \sigma_{\min}}.$$

We remark that the matrix $P^{-\frac{1}{2}} \in \mathbb{C}^{n\times n}$ introduced in (4) is actually not needed. All we need is a nonsingular matrix

$$\bar{P} = \begin{bmatrix} \bar{B} & 0 \\ 0 & \bar{C} \end{bmatrix} \in \mathbb{C}^{n\times n}$$

such that $\bar{B}^* B \bar{B} = I$ and

$$\bar{A} = \bar{P}^* A \bar{P} = \begin{bmatrix} I & \bar{E} \\ -\bar{E}^* & 0 \end{bmatrix},$$

where

$$\bar{E} = \bar{B}^* E \bar{C} \in \mathbb{C}^{p\times q}.$$

For example, we may choose $\bar{B} = B^{-\frac{1}{2}} \in \mathbb{C}^{p\times p}$ as we have done, or $\bar{B} = L^{-1} \in \mathbb{C}^{p\times p}$ when $B = L^*L$ is the Cholesky factorization of the Hermitian positive definite matrix B. For this case, we can still transform the system of linear equations (2) into an equivalent one of the form (5), with

$$\begin{bmatrix} \overline{y} \\ \overline{z} \end{bmatrix} = \bar{P}^{-1} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} \bar{B}^{-1} & y \\ \bar{C}^{-1} & z \end{bmatrix}$$

and

$$\overline{b} = \begin{bmatrix} \overline{f} \\ \overline{g} \end{bmatrix} = \bar{P}^* b = \begin{bmatrix} \bar{B}^* & f \\ \bar{C}^* & g \end{bmatrix}.$$

Following an analogous derivation to the PHSS iteration formula (6) and similar demonstrations to Theorems 3.1, 3.2 and 5.1, we can immediately obtain the PHSS iteration formula and its corresponding convergence theory with respect to this equivalent transformation of the system of linear equations (2) and its generalized version.

In summary, for the block two-by-two system of linear equations (2), we have presented a class of *preconditioned Hermitian and skew-Hermitian splitting* (**PHSS**) iteration methods, established its convergence theory, and computed its optimal iteration parameter and the corresponding optimal convergence rate. Numerical results have shown that PHSS is more effective, or at least comparable to some traditional iteration methods such as UZAWA, MINRES, GMRES, GMRES(♯) and their preconditioned variants for solving this kind of linear systems. In addition, we have also extended the techniques of designing and analyzing the PHSS method to the block two-by-two system of linear equations (2) whose (1, 1)-block matrix B is Hermitian and positive semidefinite.

Acknowledgements. The first author sincerely thanks professor Gene H. Golub for his generous and kind invitation for a visit of SCCM, Department of Computer Science, Stanford University, during January to April in 2002. The authors are indebted to the referees for their valuable comments and suggestions which greatly improved the original manuscript of this paper.

References

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

[2] Bai, Z.-Z., Golub, G.H., Ng, M.K.: Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. SIAM J. Matrix. Anal. Appl. **24**, 603–626 (2003)

[3] Benzi, M., Golub, G.H.: An iterative method for generalized saddle point problems. Technical Report SCCM-02-14, Scientific Computing and Computational Mathematics Program, Department of Computer Science, Stanford University, 2002

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

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

[6] Elman, H.C., Schultz, M.H.: Preconditioning by fast direct methods for nonself-adjoint nonseparable elliptic equations. SIAM J. Numer. Anal. **23**, 44–57 (1986)

[7] Elman, H.C., Silvester, D.J., Wathen, A.J.: Performance and analysis of saddle point preconditioners for the discrete steady-state Navier-Stokes equations. Numer. Math. **90**, 665–688 (2002)

[8] Fortin, M., Glowinski, R.: Augmented Lagrangian Methods, Applications to the Numerical Solution of Boundary Value Problems. North-Holland, Amsterdam, 1983

[9] Golub, G.H., Vanderstraeten, D.: On the preconditioning of matrices with skew-symmetric splittings. Numer. Algorithms **25**, 223–239 (2000)

[10] Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd Edition. The Johns Hopkins University Press, Baltimore, 1996

[11] Golub, G.H., Wathen, A.J.: An iteration for indefinite systems and its application to the Navier-Stokes equations. SIAM J. Sci. Comput. **19**, 530–539 (1998)

[12] Greenbaum, A.: Iterative Methods for Solving Linear Systems. SIAM, Philadelphia, 1997

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

[14] Saad, Y.: Iterative Methods for Sparse Linear Systems, *PWS Publishing Company,* Boston, 1996

[15] Wang, C.-L., Bai, Z.-Z.: Sufficient conditions for the convergent splittings of non-Hermitian positive definite matrices. Linear Algebra Appl. **330**, 215–218 (2001)

[16] Wathen, A.J., Silvester, D.J.: Fast iterative solution of stabilized Stokes systems. Part I: Using simple diagonal preconditioners. SIAM J. Numer. Anal. **30**, 630–649 (1993)