Journal of Computational Mathematics
Vol.29, No.4, 2011, 441–457.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1103-m3422

HSS METHOD WITH A COMPLEX PARAMETER FOR THE SOLUTION OF COMPLEX LINEAR SYSTEM*

Guiding Gu
Department of Applied Mathematics, Shanghai University of Finance and Economics, Shanghai 200433, China
Email: guiding@mail.shufe.edu.cn

Abstract

In this paper, a complex parameter is employed in the Hermitian and skew-Hermitian splitting (HSS) method (Bai, Golub and Ng: SIAM J. Matrix Anal. Appl., 24(2003), 603-626) for solving the complex linear system $Ax = f$. The convergence of the resulting method is proved when the spectrum of the matrix A lie in the right upper (or lower) part of the complex plane. We also derive an upper bound of the spectral radius of the HSS iteration matrix, and a estimated optimal parameter α (denoted by α_{est}) of this upper bound is presented. Numerical experiments on two modified model problems show that the HSS method with α_{est} has a smaller spectral radius than that with the real parameter which minimizes the corresponding upper bound. In particular, for the $'$dominant$'$ imaginary part of the matrix A, this improvement is considerable. We also test the GMRES method preconditioned by the HSS preconditioning matrix with our parameter α_{est}.

Mathematics subject classification: 65F10,65Y20.
Key words: Hermitian matrix, Skew-Hermitian matrix, Splitting iteration method, Complex linear system, Complex parameter.

1. Introduction

We are interested in the iterative solution of the following complex linear system

$$Ax = f. \tag{1.1}$$

We consider the case in which $A \in C^{n\times n}$ is large, sparse, non-Hermitian and positive definite and $f \in C^n$; see several applications in [12,17,21].

Bai, Golub and Ng [6] proposed the Hermitian/skew-Hermitian splitting (HSS) method based on the fact that the matrix A naturally possesses the Hermitian/skew-Hermitian splitting

$$A = H + S,$$

where $H = \frac{1}{2}(A + A^H)$ is the Hermitian matrix, $S = \frac{1}{2}(A - A^H)$ is the skew-Hermitian matrix, and A^H is the conjugate transpose of the matrix A. The HSS method has the following form:

$$\begin{cases} (\alpha I + H)x^{(k+\frac{1}{2})} = (\alpha I - S)x^{(k)} + f, \\ (\alpha I + S)x^{(k+1)} = (\alpha I - H)x^{(k+\frac{1}{2})} + f, \end{cases} \tag{1.2}$$

where the parameter $\alpha > 0$ can be chosen. The above form can be equivalently rewritten as

$$x^{(k+1)} = T(\alpha)x^{(k)} + G(\alpha)f, \qquad k = 0, 1, 2, \cdots, \tag{1.3}$$

* Received June 2, 2010 / Revised version received October 20, 2010 / Accepted December 28, 2010 / Published online June 27, 2011 /

where $T(\alpha) = (\alpha I + S)^{-1}(\alpha I - H)(\alpha I + H)^{-1}(\alpha I - S)$ is the iteration matrix, and $G(\alpha) = 2\alpha(\alpha I + S)^{-1}(\alpha I + H)^{-1}$.

The following theorem [6] gives the convergence property of the HSS iteration.

Theorem 1.1. *Suppose that $A \in C^{n\times n}$ is a positive definite matrix, $H = \frac{1}{2}(A + A^H), S = \frac{1}{2}(A - A^H)$ are the Hermitian and Skew-Hermitian parts of A respectively, and the parameter $\alpha > 0$. Then the spectral radius $\rho(T(\alpha))$ of the iteration matrix $T(\alpha)$ of the HSS iteration is bounded by*

$$\rho(T(\alpha)) \leq \sigma(\alpha) = \max_{\lambda_j \in \Lambda(H)} \left| \frac{\alpha - \lambda_j}{\alpha + \lambda_j} \right|, \tag{1.4}$$

where $\Lambda(\cdot)$ represents the spectrum of the corresponding matrix. Since A is positive definite ($\lambda_j > 0$), we have

$$\rho(T(\alpha)) \leq \sigma(\alpha) < 1, \qquad \text{for all } \alpha > 0,$$

i.e., the HSS iteration converges.

Furthermore, let $\lambda_1 \geq \cdots \geq \lambda_n > 0$ be the eigenvalues of H. Then the upper bound $\sigma(\alpha)$ has the optimal parameter

$$\tilde{\alpha} = \sqrt{\lambda_1 \lambda_n} \tag{1.5}$$

and

$$\sigma(\tilde{\alpha}) = \min_{\alpha > 0} \sigma(\alpha) = \frac{\sqrt{\kappa(H)} - 1}{\sqrt{\kappa(H)} + 1},$$

where $\kappa(H) = \frac{\lambda_1}{\lambda_n}$ is the spectral condition number of H.

However, we have the following observations:

(1) $\tilde{\alpha}$ is usually different from the optimal parameter

$$\alpha^* = arg \min_{\alpha > 0} \rho(T(\alpha)).$$

(2) Numerical experiments in [5,6,10,11] have shown that in most situations,

$$\rho(T(\alpha^*)) \ll \rho(T(\tilde{\alpha})).$$

(3) $\tilde{\alpha}$ and $\sigma(\tilde{\alpha})$ do not include any information of S.

To further improve the efficiency of the HSS method, it is desirable to determine or find a good estimate for the optimal parameter α^*. For some special constructed matrices, in particular, for saddle-point problems, the optimal parameter, or the quasi-optimal parameter [2], has been extensively discussed [2,4,5,8,10], and the results show that the optimal parameter does include the information of S.

The matrix

$$P = \frac{1}{2\alpha}(H + \alpha I)(S + \alpha I) \tag{1.6}$$

can also be employed as a preconditioner [2,8,11,23], where α is referred to as the preconditioning parameter. The idea of HSS preconditioner is motivated from the HSS method.

More generally, the coefficient matrix $A \in C^{n\times n}$ can be splitted into

$$A = N + S_0,$$

where N is a normal matrix and S_0 is a skew-Hermitian matrix. Similarly to the HSS method, *normal/skew-Hermitian splitting* (NSS) method with a real parameter could be formed [7].

When A is a positive definite matrix and α is a positive parameter, the NSS method converges, and the spectral radius $\rho(M_0(\alpha))$ of the iteration matrix $M_0(\alpha)$ of the NSS iteration is bounded by

$$\rho(M_0(\alpha)) \leq \sigma_0(\alpha) = \max_{\gamma_j + i\eta_j \in \Lambda(N)} \sqrt{\frac{(\alpha-\gamma_j)^2+\eta_j^2}{(\alpha+\gamma_j)^2+\eta_j^2}}.$$

In [7], the optimal real parameter of the upper bound $\sigma_0(\alpha)$ is also discussed, and the optimal upper bound of the contraction factor of the HSS iteration is the smallest among all NSS iterations.

In this paper, we employ a complex parameter α in the HSS iteration (1.2) for solving the complex linear system (1.1). This idea is natural as it does not increase the computational complexity of the HSS method for the complex linear systems. We show that the resulting method converges when the spectrum of the matrix A lie in the right upper (or lower) part of the complex plane. An upper bound of the spectral radius $\rho(T(\alpha))$ of the HSS iteration matrix $T(\alpha)$ is given. This upper bound includes the spectral information of the matrix S. Moreover, a estimated optimal parameter α_{est} of this upper bound is presented. Numerical experiments on two modified model problems show that the HSS method with α_{est} has a smaller spectral radius than that with the real parameter that minimizes the upper bound (1.4). In particular, for the $'$dominant$'$ imaginary part of the matrix A (see Experiment 2), this improvement is considerable. In Experiment 2, we also test the GMRES method preconditioned by the HSS preconditioner with α_{est}, and investigate how sensitive are the estimated parameter α_{est} and $\rho(T(\alpha_{est}))$ with respect to the spectral information of H and S.

2. HSS Method with Complex Parameter

We still consider the HSS iteration (1.2) for the solution of the complex linear system (1.1), but now the parameter α is complex.

Since the matrix S is skew-Hermitian, its eigenvalues are imaginary numbers or zero [22], denoted by $\lambda(S) = i\tau_j, i = \sqrt{-1}, \tau_j \in R,\ j = 1, 2, \cdots, n$. Suppose that the spectrum of the $n \times n$ complex matrix A lie in the right upper (or lower) part of the complex plane, *i.e.*, A is a positive definite matrix, and all $\tau_j \geq 0$ (or $\tau_j \leq 0$). This assumption is needed for the convergence of the HSS method with a complex parameter α.

Theorem 2.1. *Suppose that the spectrum of the $n \times n$ complex matrix A lie in the right upper (or lower) part of the complex plane. The parameter $\alpha = a + ib,\ a, b \in R$ is chosen such that $a > 0$ and $b \cdot \tau_j \geq 0,\ j = 1, \cdots, n$. Then the HSS iteration* (1.2) *with the complex parameter α converges. Furthermore, it holds that*

$$\rho(T(\alpha)) \leq \omega(\alpha) \equiv \max_{\lambda_j \in \Lambda(H)} \left| \frac{\alpha - \lambda_j}{\alpha + \lambda_j} \right| \max_{i\tau_j \in \Lambda(S)} \left| \frac{\alpha - i\tau_j}{\alpha + i\tau_j} \right| < 1. \tag{2.1}$$

Proof. Let $\tilde{T}(\alpha) = (\alpha I + S)T(\alpha)(\alpha I + S)^{-1} = (\alpha I - H)(\alpha I + H)^{-1}(\alpha I - S)(\alpha I + S)^{-1}$. Then

$$\begin{aligned}\rho(T(\alpha)) =& \rho((\alpha I - H)(\alpha I + H)^{-1}(\alpha I - S)(\alpha I + S)^{-1}) \\ \leq& \|(\alpha I - H)(\alpha I + H)^{-1}\|_2 \|(\alpha I - S)(\alpha I + S)^{-1}\|_2.\end{aligned}$$

Since S is skew-Hermitian, there exists a unitary matrix U, such that $S = U\Lambda U^H$, where $\Lambda = diag(i\tau_1, \cdots, i\tau_n)$. It follows that

$$\begin{aligned}\|(\alpha I - S)(\alpha I + S)^{-1}\|_2 =& \|U(\alpha I - \Lambda)(\alpha I + \Lambda)^{-1}U^H\|_2\\ =& \|(\alpha I - \Lambda)(\alpha I + \Lambda)^{-1}\|_2\\ =& \max_{i\tau_j \in \Lambda(S)} \left|\frac{\alpha - i\tau_j}{\alpha + i\tau_j}\right|.\end{aligned}$$

Similarly, it holds that $\|(\alpha I - H)(\alpha I + H)^{-1}\|_2 = \max_{\lambda_j \in \Lambda(H)} \left|\frac{\alpha-\lambda_j}{\alpha+\lambda_j}\right|$. Thus we obtain

$$\rho(T(\alpha)) \leq \max_{\lambda_j \in \Lambda(H)} \left|\frac{\alpha - \lambda_j}{\alpha + \lambda_j}\right| \max_{i\tau_j \in \Lambda(S)} \left|\frac{\alpha - i\tau_j}{\alpha + i\tau_j}\right|.$$

Under the assumption that $\lambda_j > 0$ nd $Re(\alpha) > 0, Im(\alpha)\tau_j \geq 0$, we have

$$\rho(T(\alpha)) \leq \omega(\alpha) < 1;$$

i.e., the HSS iteration converges. □

In particular, if the parameter $\alpha = a > 0$ is chosen to be a real number (i.e., $b = 0$), then when the matrix A is positive definite, the HSS iteration (1.2) converges; and it holds that

$$\rho(T(\alpha)) \leq \max_{\lambda_j \in \Lambda(H)} \left|\frac{a - \lambda_j}{a + \lambda_j}\right| < 1.$$

This is the conclusion of Theorem 1.1.

The upper bound (2.1) shows that $\omega(\alpha)$ includes the spectral information of the matrix S, and if a suitable complex parameter α is chosen, it is possible that $\omega(\alpha) < \sigma(\alpha)$.

Remark 2.1. For a complex parameter $\alpha = a + ib$, the HS splitting can be viewed as the NS splitting with a real parameter a:

$$A = \alpha I + H - (\alpha I - S) = aI + (ibI + H) - (aI - (-ibI + S))$$

and

$$A = \alpha I + S - (\alpha I - H) = aI + (ibI + S) - (aI - (-ibI + H)),$$

where $H_1 \equiv ibI + H$ and $H_2 \equiv -ibI + H$ are normal matrices but $H_1 \neq H_2$, and $S_1 \equiv -ibI + S$, $S_2 \equiv ibI + S$ are skew-Hermitian matrices but $S_1 \neq S_2$ too. Thus the HSS method with a complex parameter cannot be viewed as the NSS method with a real parameter [7].

Next we discuss the estimation to the optimal parameter of the upper bound $\omega(\alpha)$.

Let

$$\omega_1(\alpha) = \max_{\lambda_j \in \Lambda(H)} \left|\frac{\alpha - \lambda_j}{\alpha + \lambda_j}\right|, \qquad \omega_2(\alpha) = \max_{i\tau_j \in \Lambda(S)} \left|\frac{\alpha - i\tau_j}{\alpha + i\tau_j}\right|.$$

According to the condition for the convergence of HSS with a complex parameter in Theorem 2.1, we assume the eigenvalues $i\tau_j$ of the matrix S satisfying

$$\tau_1 \geq \cdots \geq \tau_n \geq 0,$$

and choose the parameter $\alpha = a + ib$ such that $a > 0$ and $b \geq 0$.

Lemma 2.1. *Let* $\lambda_1 \geq \cdots \geq \lambda_n > 0, \quad \tau_1 \geq \cdots \geq \tau_n \geq 0,$ *and* $a > 0, \quad b \geq 0$*. Then it holds that*

$$\omega_1(\alpha) = \max\left(\frac{|a-\lambda_1+ib|}{|a+\lambda_1+ib|}, \frac{|a-\lambda_n+ib|}{|a+\lambda_n+ib|}\right) \tag{2.2}$$

and

$$\omega_2(\alpha) = \max\left(\frac{|a+i(b-\tau_1)|}{|a+i(b+\tau_1)|}, \frac{|a+i(b-\tau_n)|}{|a+i(b+\tau_n)|}\right). \tag{2.3}$$

Proof. We consider $\omega_1^2(\alpha) = \max_{\lambda_j\in\Lambda(H)} \frac{(a-\lambda_j)^2+b^2}{(a+\lambda_j)^2+b^2}$. Let

$$g(\lambda) \equiv \frac{(a-\lambda)^2+b^2}{(a+\lambda)^2+b^2}, \qquad 0<\lambda_n \leq \lambda \leq \lambda_1.$$

It is clear that $\lambda = \sqrt{a^2+b^2}$ is the unique positive minimizer of $g(\lambda)$. Therefore the maximum of $g(\lambda)$ reaches at the point λ_1 or λ_n. This proves the result (2.2). Similarly, the result (2.3) can be shown in the same way. □

Lemma 2.2. *Under the assumption of Lemma* 2.1*, for a fixed* $b \geq 0$*,*

(1) *if* $b^2 \leq \frac{\lambda_n}{2}(\lambda_1-\lambda_n)$*, there is a minimizer of* ω_1*:* $a^* \equiv argmin_{a>0}\omega_1(a+ib) = \sqrt{\lambda_1\lambda_n - b^2}$*, and it holds that*

$$\omega_1(a^*+ib) = \sqrt{\frac{\lambda_1+\lambda_n-2a^*}{\lambda_1+\lambda_n+2a^*}} = \sqrt{\frac{\lambda_1+\lambda_n-2\sqrt{\lambda_1\lambda_n-b^2}}{\lambda_1+\lambda_n+2\sqrt{\lambda_1\lambda_n-b^2}}}; \tag{2.4$'$}$$

(2) *if* $b^2 > \frac{\lambda_n}{2}(\lambda_1-\lambda_n)$*, there is a minimizer of* ω_1*:* $a_* \equiv argmin_{a>0}\omega_1(a+ib) = \sqrt{\lambda_n^2+b^2}$*, and it holds that*

$$\omega_1(a_*+ib) = \sqrt{\frac{a_*-\lambda_n}{a_*+\lambda_n}} = \sqrt{\frac{\sqrt{\lambda_n^2+b^2}-\lambda_n}{\sqrt{\lambda_n^2+b^2}+\lambda_n}}. \tag{2.4$''$}$$

Similarly, for a fixed $a > 0$*,*

(3) *if* $a^2 \leq \frac{\tau_n}{2}(\tau_1-\tau_n)$*, there is a minimizer of* ω_2*:* $b^* \equiv argmin_{b\geq 0}\omega_2(a+ib) = \sqrt{\tau_1\tau_n - a^2}$*, and it holds that*

$$\omega_2(a+ib^*) = \sqrt{\frac{\tau_1+\tau_n-2b^*}{\tau_1+\tau_n+2b^*}} = \sqrt{\frac{\tau_1+\tau_n-2\sqrt{\tau_1\tau_n-a^2}}{\tau_1+\tau_n+2\sqrt{\tau_1\tau_n-a^2}}}; \tag{2.5$'$}$$

(4) *if* $a^2 > \frac{\tau_n}{2}(\tau_1-\tau_n)$*, there is a minimizer of* ω_2*:* $b_* \equiv argmin_{b\geq 0}\omega_2(a+ib) = \sqrt{\tau_n^2+a^2}$*, and it holds that*

$$\omega_2(a+ib_*) = \sqrt{\frac{b_*-\tau_n}{b_*+\tau_n}} = \sqrt{\frac{\sqrt{\tau_n^2+a^2}-\tau_n}{\sqrt{\tau_n^2+a^2}+\tau_n}}. \tag{2.5$''$}$$

Proof. We only prove the result (2.4). The conclusion (2.5) can be proven in the same way. We consider the minimum of $\omega_1^2(a+ib)$ for a fixed b. For $j=1$ and $j=n$, let

$$\xi_j(a) \equiv \frac{(a-\lambda_j)^2+b^2}{(a+\lambda_j)^2+b^2}.$$

Then

$$\xi_j'(a) = \frac{4\lambda_j(a^2-\lambda_j^2-b^2)}{[(a+\lambda_j)^2+b^2]^2}.$$

It is clare that $a = \sqrt{\lambda_j^2 + b^2}$ is the unique positive minimizer of $\xi_j(a)$.

(i) In the case of $b^2 \geq \lambda_1\lambda_n \geq \frac{\lambda_n}{2}(\lambda_1 - \lambda_n)$, since for any $a > 0$,

$$\xi_1(a) - \xi_n(a) = \frac{4a}{((a+\lambda_1)^2 + b^2)((a+\lambda_n)^2 + b^2)}(\lambda_1\lambda_n - b^2 - a^2)(\lambda_1 - \lambda_n) < 0, \tag{2.6}$$

it holds that $\omega_1^2(a + ib) = \xi_n(a)$ for a fixed $b \geq \sqrt{\lambda_1\lambda_n}$. Therefore, the minimizer of $\omega_1(a + ib)$ with respect to a is the minimizer of $\xi_n(a)$: $a_* = \sqrt{\lambda_n^2 + b^2}$.

(ii) In the case of $b^2 < \lambda_1\lambda_n$, when $a > 0$, ξ_1 and ξ_n have the unique intersection point $a^* = \sqrt{\lambda_1\lambda_n - b^2}$, and $a^* < \lambda_1$.

When $a^* = \sqrt{\lambda_1\lambda_n - b^2} \geq \sqrt{\lambda_n^2 + b^2}$, or equivalently, $b^2 \leq \frac{\lambda_n}{2}(\lambda_1 - \lambda_n)$, for $a \in (\sqrt{\lambda_n^2 + b^2},\ \sqrt{\lambda_1^2 + b^2})$, $\xi_1(a)$ decreases, while ξ_n increases, so the minimum of $\omega_1(a + ib)$ with respect to a is reached at the intersection point $a^* = \sqrt{\lambda_1\lambda_n - b^2}$, which is the result of (2.4′).

When $a^* = \sqrt{\lambda_1\lambda_n - b^2} < \sqrt{\lambda_n^2 + b^2}$, or equivalently, $b^2 > \frac{\lambda_n}{2}(\lambda_1 - \lambda_n)$, for $a \in (0,\ \sqrt{\lambda_n^2 + b^2})$, both ξ_1 and ξ_n decrease, and from (2.6), we have

$$\xi_1(a_*) - \xi_n(a_*) = \frac{4a_*}{((a_*+\lambda_1)^2 + b^2)((a_*+\lambda_n)^2 + b^2)}(\lambda_1\lambda_n - b^2 - (\lambda_n^2 + b^2))(\lambda_1 - \lambda_n) < 0,$$

so the minimum value of $\omega_1(a + ib)$ with respect to a is reached at $a_* = \sqrt{\lambda_n^2 + b^2}$ of ξ_n, which is the result (2.4″). □

According to the expression (2.1) of $\omega(a + bi)$, we consider the minimization of $\omega(a + ib)$ in the area $\Omega = [0 < a \leq \lambda_1, 0 \leq b \leq \tau_1]$:

$$\min_{(a,b)\in\Omega} \omega(a + ib) = \min_{(a,b)\in\Omega} \omega_1(a + ib)\omega_2(a + ib). \tag{2.7}$$

Clearly, the following lemma holds.

Lemma 2.3. *Let $f_1(a, b) \geq 0,\ f_2(a, b) \geq 0$ for $(a, b) \in \Omega$. Then for any non-negative function $b = s(a) \leq \tau_1$, it holds that*

$$\min_{(a,b)\in\Omega} f_1(a, b)f_2(a, b) \leq \min_{0<a\leq\lambda_1} [f_1(a, s(a)) \min_{0\leq b\leq\tau_1} f_2(a, b)]. \tag{2.8′}$$

Similarly, for any positive function $a = t(b) \leq \lambda_1$, it holds that

$$\min_{(a,b)\in\Omega} f_1(a, b)f_2(a, b) \leq \min_{0\leq b\leq\tau_1} [f_2(t(b), b) \min_{0<a\leq\lambda_1} f_1(a, b)]. \tag{2.8″}$$

Set

$$a_0 = \min\left(\sqrt{\frac{\tau_n}{2}(\tau_1 - \tau_n)}, \lambda_1\right), \qquad b_0 = \min\left(\sqrt{\frac{\lambda_n}{2}(\lambda_1 - \lambda_n)}, \tau_1\right).$$

According to Lemma 2.2, when $0 < a < a_0$, $\omega_2(a, b)$ reaches its minimum at the point $b^* = \sqrt{\tau_1\tau_n - a^2}$, so we set $s(a) = \sqrt{\tau_1\tau_n - a^2}$ in Lemma 2.3; when $a_0 < a \leq \lambda_1$, $\omega_2(a, b)$ reaches its minimum at the point $b_* = \sqrt{\tau_n^2 + a^2}$, so we set $s(a) = \sqrt{\tau_1\tau_n - a^2}$ in Lemma 2.3. Thus we have the following result (2.9). Similarly, by setting $t(b) = \sqrt{\lambda_1\lambda_n - b^2}$ and $t(b) = \sqrt{\lambda_n^2 + b^2}$ in Lemma 2.3, we can obtain the result (2.10). We summarize these in the following theorem.

Theorem 2.2. *Under the assumption of Lemma* 2.1*, the following results hold true*

$$\min_{(a,b)\in\Omega}\omega(a+ib)\le\min(\min_{0<a\le a_0}\phi_1(a),\ \min_{a_0\le a\le\lambda_1}\phi_2(a)) \tag{2.9}$$

and

$$\min_{(a,b)\in\Omega}\omega(a+ib)\le\min(\min_{0\le b\le b_0}\psi_1(b),\ \min_{b_0\le b\le\tau_1}\psi_2(b)), \tag{2.10}$$

where

$$\begin{aligned}
\phi_1(a)=&\omega_1(a+i\sqrt{\tau_1\tau_n-a^2})\omega_2(a+i\sqrt{\tau_1\tau_n-a^2}),\\
\phi_2(a)=&\omega_1(a+i\sqrt{\tau_n^2+a^2})\omega_2(a+i\sqrt{\tau_n^2+a^2}),\\
\psi_1(b)=&\omega_1(\sqrt{\lambda_1\lambda_n-b^2}+ib)\omega_2(\sqrt{\lambda_1\lambda_n-b^2}+ib),\\
\psi_2(b)=&\omega_1(\sqrt{\lambda_n^2+b^2}+ib)\omega_2(\sqrt{\lambda_n^2+b^2}+ib).
\end{aligned}$$

Next, we discuss the problem $\min_{0\le b\le b_0}\psi_1(b)$ by considering the minimization of $\psi_1^2(b)$. The problem $\min_{0<a\le a_0}\phi_1(a)$ can be discussed in the same way.

From (2.4′) in Lemma 2.2 and by setting $a^*(b)=\sqrt{\lambda_1\lambda_n-b^2}$, we have

$$\omega_1^2(a^*(b)+ib)=1-\frac{4a^*(b)}{\lambda_1+\lambda_n+2a^*(b)}$$

and

$$\omega_2^2(a^*(b)+ib)=max\left(\frac{(a^*(b))^2+(b-\tau_1)^2}{(a^*(b))^2+(b+\tau_1)^2},\frac{(a^*(b))^2+(b-\tau_n)^2}{(a^*(b))^2+(b+\tau_n)^2}\right).$$

For $j=1$ and $j=n$, let

$$h_j(b)\equiv\frac{(a^*(b))^2+(b-\tau_j)^2}{(a^*(b))^2+(b+\tau_j)^2}=1-\frac{4b\tau_j}{\tau_j^2+\lambda_1\lambda_n+2b\tau_j}.$$

Then

$$h_1(b)-h_n(b)=\frac{4b(\tau_1-\tau_n)(\tau_1\tau_n-\lambda_1\lambda_n)}{(\lambda_1\lambda_n+\tau_1^2+2b\tau_1)(\lambda_1\lambda_n+\tau_n^2+2b\tau_n)}.$$

If $\tau_1\tau_n\ge\lambda_1\lambda_n$, then $h_1(b)\ge h_n(b)$. Therefore,

$$\omega_2^2(a^*(b)+ib)=h_1(b)=1-\frac{4b\tau_1}{\tau_1^2+\lambda_1\lambda_n+2b\tau_1}.$$

If $\tau_1\tau_n\le\lambda_1\lambda_n$, then $h_1(b)\le h_n(b)$. Therefore,

$$\omega_2^2(a^*(b)+ib)=h_n(b)=1-\frac{4b\tau_n}{\tau_n^2+\lambda_1\lambda_n+2b\tau_n}.$$

Thus, $\psi_1^2(b)$ has the following expression in $[0,b_0]$:

$$\psi_1^2(b)=\begin{cases}(1-\frac{4a^*(b)}{\lambda_1+\lambda_n+2a^*(b)})(1-\frac{4b\tau_1}{\tau_1^2+\lambda_1\lambda_n+2b\tau_1}), & \tau_1\tau_n\ge\lambda_1\lambda_n,\\ (1-\frac{4a^*(b)}{\lambda_1+\lambda_n+2a^*(b)})(1-\frac{4b\tau_n}{\tau_n^2+\lambda_1\lambda_n+2b\tau_n}), & \tau_1\tau_n\le\lambda_1\lambda_n.\end{cases}$$

Similarly, $\phi_1^2(a)$ has the following expression in $(0,a_0]$:

$$\phi_1^2(a)=\begin{cases}(1-\frac{4b^*(a)}{\tau_1+\tau_n+2b^*(a)})(1-\frac{4a\lambda_1}{\lambda_1^2+\tau_1\tau_n+2a\lambda_1}), & \tau_1\tau_n\le\lambda_1\lambda_n,\\ (1-\frac{4b^*(a)}{\tau_1+\tau_n+2b^*(a)})(1-\frac{4a\lambda_n}{\lambda_n^2+\tau_1\tau_n+2a\lambda_n}), & \tau_1\tau_n\ge\lambda_1\lambda_n,\end{cases}$$

where $b^*(a)=\sqrt{\tau_1\tau_n-a^2}$.

The following theorem provides the minima of $\phi_1(a)$ and $\psi_1(b)$.

Theorem 2.3. *Under the assumption of Lemma* 2.1 *and* $\tau_1 \neq \tau_n$*, there is the minimizer* $a'_{est} = \sqrt{\tilde{a}^*}$ *of* $\phi_1(a)$ *in* $(0, \sqrt{\tau_1\tau_n}]$*, and* $\tilde{a}^*$ *is a positive root of either*

$$p_1(\tilde{a}) = 0, \qquad \lambda_1\lambda_n \geq \tau_1\tau_n, \tag{2.11$'$}$$

or

$$p_n(\tilde{a}) = 0, \qquad \lambda_1\lambda_n \leq \tau_1\tau_n, \tag{2.11$''$}$$

where for $j = 1$ *and* $j = n$*,*

$$\begin{aligned} p_j(\tilde{a}) \equiv& 16\lambda_j^2(\lambda_j^2\tau_c^2 + u_j^2)\tilde{a}^3 - 48\tau_1\tau_n\lambda_j^2u_j^2\tilde{a}^2 \\ &+ u_j^2[\tau_c^2u_j^2 + \lambda_j^2(\tau_1 - \tau_n)^2(\tau_1^2 + \tau_n^2 - 10\tau_1\tau_n)]\tilde{a} - \lambda_j^2u_j^2\tau_1\tau_n(\tau_1 - \tau_n)^4, \end{aligned}$$

$u_j = \lambda_j^2 + \tau_1\tau_n,\ \tau_c = \tau_1 + \tau_n,\ \tilde{a} = a^2.$

Similarly, there is a minimizer $b''_{est} = \sqrt{\tilde{b}^*}$ *of* $\psi_1(b)$ *in* $[0, \sqrt{\lambda_1\lambda_n}]$ *, and* $\tilde{b}^*$ *is a non-negative root of either*

$$q_1(\tilde{b}) = 0, \qquad \tau_1\tau_n \geq \lambda_1\lambda_n, \tag{2.12$'$}$$

or

$$q_n(\tilde{b}) = 0, \qquad \tau_1\tau_n \leq \lambda_1\lambda_n, \tag{2.12$''$}$$

where for $j = 1$ *and* $j = n$*,*

$$\begin{aligned} q_j(\tilde{b}) \ \equiv\ & 16\tau_j^2(\tau_j^2\lambda_c^2 + v_j^2)\tilde{b}^3 - 48\lambda_1\lambda_n\tau_j^2v_j^2\tilde{b}^2 \\ +\ & v_j^2[\lambda_c^2v_j^2 + \tau_j^2(\lambda_1 - \lambda_n)^2(\lambda_1^2 + \lambda_n^2 - 10\lambda_1\lambda_n)]\tilde{b} - \tau_j^2v_j^2\lambda_1\lambda_n(\lambda_1 - \lambda_n)^4, \end{aligned}$$

$v_j = \tau_j^2 + \lambda_1\lambda_n,\ \lambda_c = \lambda_1 + \lambda_n,\ \tilde{b} = b^2.$

Proof. We prove the result (2.12$'$). The results (2.12$''$) and (2.11) can be shown in the same way.

For simplicity, we now omit the superscript of $a^*(b)$ and let $a(b) = \sqrt{\lambda_1\lambda_n - b^2}$. Then

$$\psi_1^2(b) = \frac{\lambda_c - 2a(b)}{\lambda_c + 2a(b)}\frac{v_1 - 2b\tau_1}{v_1 + 2b\tau_1}.$$

By setting $\frac{d\psi_1^2}{db} = 0$, we have

$$q(b) \equiv 4\lambda_c\tau_1^2a'(b)b^2 - \lambda_cv_1^2a'(b) + 4\tau_1v_1a(b)^2 - \tau_1v_1\lambda_c^2 = 0. \tag{2.13}$$

Consider the value of $q(b)$ at $b_\epsilon = \sqrt{\lambda_1\lambda_n - \epsilon^2}$, where $\epsilon > 0$ is a small positive number. Note that $a(b_\epsilon) = \epsilon > 0$ and $a'(b_\epsilon) = -\frac{\sqrt{\lambda_1\lambda_n - \epsilon^2}}{\epsilon}$. Thus we have

$$\begin{aligned} q(b_\epsilon) =& 4\lambda_c\tau_1^2(\lambda_1\lambda_n - \epsilon^2)(-\frac{\sqrt{\lambda_1\lambda_n - \epsilon^2}}{\epsilon}) - \lambda_cv_1^2(-\frac{\sqrt{\lambda_1\lambda_n - \epsilon^2}}{\epsilon}) + 4\tau_1v_1\epsilon^2 - \tau_1v_1\lambda_c^2 \\ =& \frac{1}{\epsilon}[\lambda_c\sqrt{\lambda_1\lambda_n - \epsilon^2}(v_1^2 - 4\tau_1^2\lambda_1\lambda_n) + \mathcal{O}(\epsilon)] \\ =& \frac{1}{\epsilon}[\lambda_c\sqrt{\lambda_1\lambda_n - \epsilon^2}(\tau_1^2 - \lambda_1\lambda_n)^2 + \mathcal{O}(\epsilon)]. \end{aligned}$$

For a suitable small $\epsilon > 0$, $q(b_\epsilon) \geq 0$. Also for a suitable small $\epsilon > 0$,

$$q(\epsilon) = -\tau_1v_1(\lambda_1 - \lambda_n)^2 + \mathcal{O}(\epsilon) \leq 0.$$

This shows that as b increases in $[0, \sqrt{\lambda_1\lambda_n}]$, $\psi_1^2(b)$ firstly decreases and then increases, and therefore, $\psi_1^2(b)$ has the minimizer b''_{est} in $[0, \sqrt{\lambda_1\lambda_n}]$, which satisfies $q(b''_{est}) = 0$. If $\lambda_1 \neq \lambda_n$, $b''_{est} \in (0, \sqrt{\lambda_1\lambda_n}]$.

Next, we simplify the expression (2.13). By substituting $a(b)^2 = \lambda_1\lambda_n - b^2$ and $a'(b) = -\frac{b}{\sqrt{\lambda_1\lambda_n - b^2}}$ into (2.13), we have

$$-4\lambda_c\tau_1^2\frac{b^3}{\sqrt{\lambda_1\lambda_n - b^2}} + \lambda_c v_1^2\frac{b}{\sqrt{\lambda_1\lambda_n - b^2}} - 4\tau_1 v_1 b^2 + \tau_1 v_1(4\lambda_1\lambda_n - \lambda_c^2) = 0,$$

or equivalently,

$$\lambda_c^2(v_1^2 - 4\tau_1^2 b^2)^2 b^2 = \tau_1^2 v_1^2[4b^2 + (\lambda_1 - \lambda_n)^2]^2(\lambda_1\lambda_n - b^2).$$

Let $\tilde{b} = b^2$. Then we have

$$\begin{aligned}&(16\lambda_c^2\tau_1^4 + 16\tau_1^2 v_1^2)\tilde{b}^3 + [-8\lambda_c^2\tau_1^2 v_1^2 + 8\tau_1^2 v_1^2(\lambda_1 - \lambda_n)^2 - 16\tau_1^2 v_1^2\lambda_1\lambda_n]\tilde{b}^2\\ &\quad + [\lambda_c^2 v_1^4 - 8\tau_1^2 v_1^2(\lambda_1 - \lambda_n)^2\lambda_1\lambda_n + \tau_1^2 v_1^2(\lambda_1 - \lambda_n)^4]\tilde{b} - \tau_1^2 v_1^2(\lambda_1 - \lambda_n)^4\lambda_1\lambda_n = 0,\end{aligned}$$

which leads to (2.12′). □

The minimization problems of $\phi_2(a)$ and $\psi_2(b)$ can be discussed in a similar way, and hence are omitted. It is worthy mentioned that our numerical experiments indicate that the minimum value of $\phi_1(a)$ or $\psi_1(b)$ is smaller than the minimum value of $\phi_2(a)$ or $\psi_2(b)$ in (2.9) and (2.10).

Consequently by Theorem 2.2, we have the following upper bound for $min\ \omega(a + ib)$,

$$\min_{0<a\leq\lambda_1, 0\leq b\leq\tau_1} \omega(a + ib) \leq min(\min_{0<a\leq\sqrt{\tau_1\tau_2}} \phi_1(a), \min_{0\leq b\leq\sqrt{\lambda_1\lambda_2}} \psi_1(b)). \tag{2.14}$$

This minimization problem, by Theorem 2.3, can be solved via solving equations of (2.11) and (2.12).

We now summarize our discussions and provide our parameter α_{est} which estimates the optimal parameter of the upper bound $\omega(\alpha)$ as follows.

Given $\lambda_1 \geq \lambda_n > 0$ and $\tau_1 > \tau_n \geq 0$,

(1) for the case $\lambda_1\lambda_n \geq \tau_1\tau_n$:

- find the minimizer a'_{est} of $\phi_1(a)$ by solving the positive root $p_1(\tilde{a}) = 0$ in (2.11′), and the minimizer b''_{est} of $\psi_1(b)$ by solving the positive root $q_n(\tilde{b}) = 0$ in (2.12″);
- if $\phi_1(a'_{est}) \leq \psi_1(b''_{est})$, let $a_{est} = a'_{est}, b_{est} = \sqrt{\tau_1\tau_n - a_{est}^2}$ to form $\alpha_{est} = a_{est} + ib_{est}$; otherwise
- if $\phi_1(a'_{est}) > \psi_1(b''_{est})$, let $b_{est} = b''_{est}, a_{est} = \sqrt{\lambda_1\lambda_n - b_{est}^2}$ to form $\alpha_{est} = a_{est} + ib_{est}$.

(2) for the case $\lambda_1\lambda_n < \tau_1\tau_n$:

- find the minimizer a'_{est} of $\phi_1(a)$ by solving the positive root $p_n(\tilde{a}) = 0$ in (2.11″), and the minimizer b''_{est} of $\psi_1(b)$ by solving the positive root $q_1(\tilde{b}) = 0$ in (2.12′);
- if $\phi_1(a'_{est}) \leq \psi_1(b''_{est})$, let $a_{est} = a'_{est}, b_{est} = \sqrt{\tau_1\tau_n - a_{est}^2}$ to form $\alpha_{est} = a_{est} + ib_{est}$; otherwise
- if $\phi_1(a'_{est}) > \psi_1(b''_{est})$, let $b_{est} = b''_{est}, a_{est} = \sqrt{\lambda_1\lambda_n - b_{est}^2}$ to form $\alpha_{est} = a_{est} + ib_{est}$.

3. Implementation Aspects

In the HSS method, two shifted sub-systems with respect to $\alpha I + H$ and $\alpha I + S$ must be solved in each iteration.

For a real parameter $\alpha > 0$ and a Hermitian positive definite matrix H, the shifted matrix $\alpha I + H$ remains a Hermitian positive definite, so the $'$exact$'$ solution could be solved by the (complex) Cholesky factorization method, while its $'$inexact$'$ solution could be solved by the (complex) CG method; however, for the real shifted skew-Hermitian matrix $\alpha I + S$, when S is structured [3], the difficulty in solving the relevant system varies from case to case. But in general, this will not be the case.

For a complex parameter $\alpha = a + ib$, since the sub-system $(\alpha I + S)u = g$ is equivalent to $(i\alpha I + iS)u = ig$, both shifted sub-systems become systems with respect to a complex shifted Hermitian positive definite matrix $\alpha I + H$ or $i\alpha I + iS$ under the assumption of Lemma 2.1. For large sparse matrices, these shifted sub-systems could be solved by certain shifted Krylov subspace method like the Lanczos method, or the MINRES method *et al.* [13-16,18-20]. Since the Krylov method keeps shift invariance, the basis vector can be constructed by the matrix H or iS (using short recurrence), irrelative to the shift α or $i\alpha$. Thus the computational cost with a complex parameter could not be significantly higher than that with a real parameter. This deserves a further and detailed investigation.

The HSS preconditioner is nowadays mostly discussed. The matrix

$$P = \frac{1}{2\alpha}(H + \alpha I)(S + \alpha I)$$

is called the HSS preconditioner, where α is referred to as the preconditioning parameter. It is known that in general, the smaller the spectral radius $\rho(T(\alpha))$ of the HSS iterative matrix is, the better the gathering of the spectrum the preconditioned matrix $P^{-1}A$. This will make the Krylov subspace method converge faster; see Experiment 2 in the next section. Also, two shifted systems with respect to $H + \alpha I$ and $S + \alpha I$ (or $iS + i\alpha I$) have to be solved in solving the preconditioning sub-system each iteration. We can still use the shifted Krylov subspace method to solve them as a polynomial preconditioner. In addition, the incomplete factorization [14] of these shifted matrix may be used as a proconditioner, instead of $H + \alpha I$ and $S + \alpha I$, which could probably keep the computational cost economic, even for the complex parameter.

How to solve the systems with respect to a shifted Hermitian or skew-Hermitian efficiently is a further problem. We will leave it for our further work.

4. Numerical Experiments

In this section, we present two numerical experiments in solving the complex linear systems (1.1) from two modified model problems by the HSS iteration with our estimated parameter α_{est} and with the real parameter $\tilde{\alpha}$ in (1.5); moreover, for comparison purpose, the HSS method with the experimental $'$*optimal*$'$ parameter $\alpha_{exp} = a_{exp} + ib_{exp}$ is also tested, which is obtained by computing the spectral radius $\rho(T(\alpha))$ of the iterative matrix $T(\alpha)$ with all parameter α on 100×100 mesh points in $[0, \lambda_1; 0, \tau_1]$. In Experiment 2, we also report results from full GMRES method, full GMRES preconditioned by HSS(α) preconditioner (see Table 4.5), and the sensitivity of the estimated parameter α_{est} and $\rho(T(\alpha_{est}))$ with respect to the spectral information of H and S (see Table 4.6).

In our test, the eigenvalues of a matrix are solved by the function *eig* and the root of $\pi(x)=0$ is solved by the function *root* in *Matlab(7.4 ed)*.

We report the results of our numerical experiments with a Fortran 77 implementation of the HSS method based on the iteration (1.2). Two shifted sub-systems are solved by exact factorization method. The right-hand side of the linear system (1.1) is formed by $f = Ax$, where

$$x = (1-i, 1-i, \cdots, 1-i)^T. \tag{4.1}$$

The initial value is $x^{(0)} = 0$, and the stopping criterion is based on the residual of system (1.1) $\|f - Ax^{(k)}\| < 10^{-6}$.

Experiment 1. Consider a differential equation with complex coefficient which forms a non-Hermitian complex matrix,

$$-u_{xx} + i\delta x u_x + \gamma u = g, \qquad x \in [0,1],$$

where $\delta \in R$, $\gamma \in C$, $i = \sqrt{-1}$. This equation is modified to the model problem

$$-u_{xx} + \delta x u_x + \gamma u = g, \qquad \delta,\ \gamma \in R.$$

When the centered difference to u_{xx} and centered differences or the forward difference to u_x are applied to the above model, we get the linear system (1.1) with the tridiagonal complex coefficient matrices, denoted by A_c and A_f respectively. The kth rows of A_c and A_f are:

$$(A_c)_k = \mathrm{tridiag}(-1 - ikP, 2 + \gamma h^2, -1 + ikP),$$

or

$$(A_f)_k = \mathrm{tridiag}(-1, 2 + (\gamma - ik\delta)h^2, -1 + ik\delta h^2),$$

where $h = \frac{1}{n+1}$ is the step-size and $P = \frac{\delta h^2}{2}$. In our numerical tests, we set $n = 100$ (since we want to seek the experimental $'$optimal$'$ parameter α_{exp} on 100×100 mesh points, it will take too much time if n is large) and $\gamma = 2000 + i20000$. The numerical results with A_c and A_f are listed in Table 4.1 and Table 4.2 respectively.

In these tables, the experimental $'$optimal$'$ parameter α_{exp}, the estimated parameter α_{est}, the upper-bound minimizer $\tilde{\alpha}$ for real parameter, and the corresponding spectral radii $\rho(\alpha) \equiv \rho(T(\alpha))$ of the HSS iteration matrix $T(\alpha)$ are presented; moreover the number of iterations (denoted by IT) for the convergence of the HSS iteration method are also listed with different parameter δ.

Table 4.1: α, $\rho(\alpha)$ and iteration number(IT) of HSS with A_c for Ex. 1.

δ	5	10	100
α_{exp}	$0.0000 + 1.9608i$	$0.0000 + 1.9611i$	$0.0000 + 1.9655i$
$\rho(\alpha_{exp})$	1.1549×10^{-4}	2.5441×10^{-4}	2.4646×10^{-3}
IT	3	3	4
α_{est}	$6.3362 \times 10^{-9} + 1.9605i$	$2.5280 \times 10^{-8} + 1.9606i$	$3.3174 \times 10^{-7} + 1.9606i$
$\rho(\alpha_{est})$	8.5953×10^{-5}	1.2693×10^{-4}	1.2431×10^{-3}
IT	3	3	4
$\tilde{\alpha}$	0.9087	0.9075	0.3406
$\rho(\tilde{\alpha})$	0.6439	0.6443	0.8553
IT	39	39	×

From Table 4.1 we see that the numerical results produced with α_{est} coincide with those using α_{exp}, but it does not for $\tilde{\alpha}$.

Remark 4.1. (1) The situation $\rho(\alpha_{exp}) > \rho(\alpha_{est})$ occurs. This is because the mesh 100×100 over $[0, \lambda_1; 0, \tau_1]$ is rough, *e.g.*, for $\delta = 5$, $\lambda_1 = 4.1953, \tau_1 = 1.9608$, and the mesh size is 0.04195×0.01961, which cannot reach the precision $\mathcal{O}(10^{-3})$. Therefore it is possible that α_{exp} is not an exact optimal parameter.

(2) The symbol $' \times '$ in Table 4.1 (also in Table 4.2) means that the HSS method could not meet the stopping criterion after 200 iteration steps.

Table 4.2: α, $\rho(\alpha)$ and iteration number(IT) of HSS with A_f for Ex. 1.

δ	5	10	100
α_{exp}	$0.0000 + 1.9211i$	$0.0000 + 1.8817i$	$1.9200 + 1.9556i$
$\rho(\alpha_{exp})$	0.0139	0.0288	0.4373
IT	5	6	19
α_{est}	$0.2246 \times 10^{-3} + 1.9138i$	$0.9409 \times 10^{-3} + 1.8659i$	$0.0957 + 0.4701i$
$\rho(\alpha_{est})$	0.0120	0.0246	0.5855
IT	5	6	49
$\tilde{\alpha}$	0.9087	0.9075	0.3528
$\rho(\tilde{\alpha})$	0.6439	0.6444	0.8501
IT	39	39	$\times$

The results in Table 4.2 indicate that that α_{est} is a good approximation to α_{exp} for small δ, and the iteration number for convergence with α_{est} is the same as that with α_{exp}. On the other hand for large δ, α_{est} is not close to α_{exp}, but is still better than $\tilde{\alpha}$.

Experiment 2. (See [1,3]) The linear systems (1.1) is of the form

$$(W + iZ)x = f, \tag{4.2}$$

where $W = \tilde{K} + w_1 I,\ Z = \tilde{K} + w_2 I$ and $w_1 = \frac{3+\sqrt{3}}{\tau}, w_2 = \frac{3-\sqrt{3}}{\tau}$, τ is the time step-size and $\tilde{K}$ is the five-point centered difference matrix approximating the negative Laplacian operator $L = -\Delta$ with homogeneous Dirichlet boundary conditions, on a uniform mesh in the unit square $[0, 1] \times [0, 1]$ with the mesh-size $h = \frac{1}{m+1}$. Thus, the systems is a complex symmetric system with the $n \times n$ coefficient matrix $A = W + iZ$ and $n = m \times m$. For the complex symmetric systems, the modified HSS method [3] has the considerable advantage.

In our experiment, in order to show the advantage of the HSS with a complex parameter, we modify the above so that the resulting system is a complex non-symmetric system, but satisfying with the convergence condition of Theorem 2.1.

First, the matrix $\tilde{K}$ is changed to K, which is the five-point centered difference matrix approximating the operator $L = -\Delta + \gamma(\partial_x + \partial_y),\ \gamma \in R$. Let

$$W = K + w_1 I \quad and \quad Z = K + w_2 I. \tag{4.3}$$

Thus, the matrix $A = W + iZ$ is complex non-symmetric. In our tests, we also take $\tau = h$ and normalize the coefficient matrix and the right-hand side by multiplying both by h^2; see [3].

Let $H = \frac{1}{2}(A + A^H)$ and $S = \frac{1}{2}(A - A^H)$. We use the function *eig* in *Matlab* to solve the eigenvalues of the matrices H and S: $\lambda_1 = \lambda_{max}(H) = 8.2119,\ \lambda_n = \lambda_{min}(H) = 0.3448;\ \tau_1 =$

$\tau_{max}(S) = 8.0082$, $\tau_n = \tau_{min}(S) = 0.1410$. Our estimated parameter α_{est} can be derived by (2.11), (2.12). The experimental result for Ex. 2 with (4.2), (4.3) and $m = 16, \gamma = 1$ is listed in Table 4.3. Fig.1 shows the spectral radii $\rho(\alpha)$ with all $\alpha = a + ib$ on the 100×100 mesh points in $[0, \lambda_1; 0, \tau_1]$.

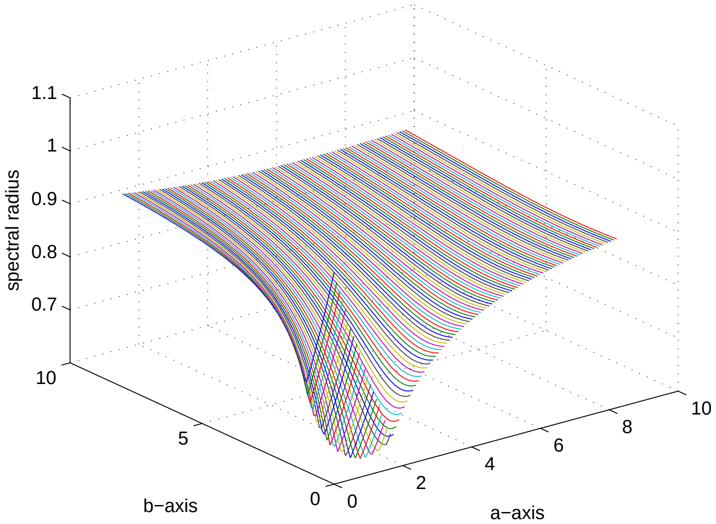

Fig. 4.1. $\rho(\alpha)$ with all $\alpha = a + ib$ on 100×100 mesh points in $[0, \lambda_1; 0, \tau_1]$ for Ex. 2 with (4.2), (4.3) and $m = 16, \gamma = 1$.

Table 4.3: α, $\rho(\alpha)$, $\omega(\alpha)$ or $\sigma(\alpha)$ and iteration number(IT) of HSS for Ex. 2 with (4.2), (4.3) and $m = 16, \gamma = 1$.

parameter	$\alpha_{est} = 1.5799 + 0.5792i$	$\tilde{\alpha} = 1.6827$	$\alpha_{exp} = 1.3139 + 0.7207i$
spectral radius	$\rho(\alpha_{est}) = 0.6375$	$\rho(\tilde{\alpha}) = 0.6598$	$\rho(\alpha_{exp}) = 0.6089$
upper bound	$\omega(\alpha_{est}) = 0.6409$	$\sigma(\tilde{\alpha}) = 0.6599$	-
IT	37	39	33

From Table 4.3 we see that α_{est} yields a better approximation to α_{exp} than $\tilde{\alpha}$ does; however α_{est} makes improvement on the spectral radius a little bit on $\tilde{\alpha}$. We note that $\tilde{\alpha}$ is close to the real part of α_{exp} (or α_{est}) and this real part is $'dominant'$ to the imaginary part.

Now we exchange w_1 with w_2 in (4.3), such that

$$W = K + w_2 I \quad \textit{and} \quad Z = K + w_1 I. \tag{4.4}$$

Note that $w_1 > w_2$, which means that the imaginary part Z of the matrix A is $'dominant'$ to the real part W. In this case, the eigenvalues of the matrices H and S are $\lambda_1 = 8.0082$, $\lambda_n = 0.1410$; $\tau_1 = 8.2119$, $\tau_n = 0.3448$. The experimental result for Ex. 2 with (4.2), (4.4) and $m = 16, \gamma = 1$ is listed in Table 4.4.

Fig. 2 shows that the optimal parameter is complex indeed, and $\rho(\alpha)$ can not reach its minimum on the real axis.

The results in Table 4.4 show that α_{est} yields a good approximation to α_{exp}, and a considerable improvement on the spectral radius, as well as the iteration number for convergence on the parameter $\tilde{\alpha}$. Also, We note that the imaginary part of α_{exp} is $'dominant'$ to the real part.

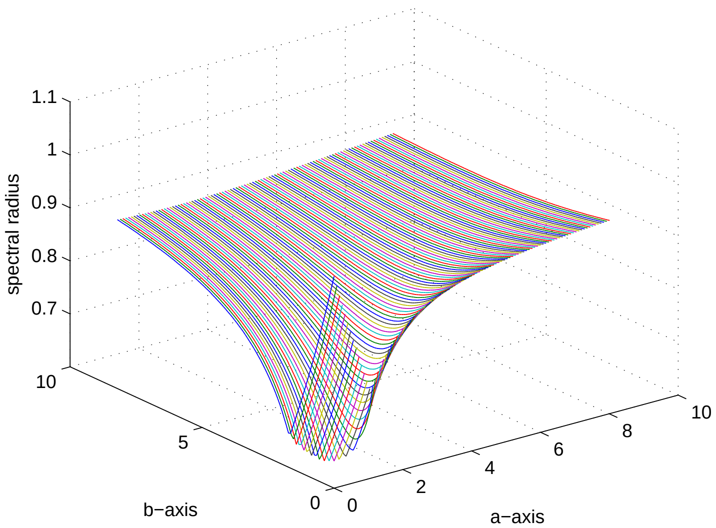

Fig. 4.2. $\rho(\alpha)$ with all $\alpha = a + ib$ on 100×100 mesh points in $[0, \lambda_1; 0, \tau_1]$ for Ex.2 with (4.2), (4.4) and $m = 16, \gamma = 1$.

Next we make the imaginary part Z more $'dominant'$ to the real part W by setting

$$W = K + \frac{w_2}{2} I \quad \text{and} \quad Z = K + 2w_1 I. \tag{4.5}$$

The eigenvalues of the matrices H and S are $\lambda_1 = 7.9709,\ \lambda_n = 0.1037;\ \tau_1 = 8.4903,\ \tau_n = 0.6231$. The experimental results for Ex. 2 with (4.2), (4.5) and $m = 16, \gamma = 1$ are also shown in Table 4.4. Our estimated parameter α_{est} makes more improvement.

Other mesh-size m and the parameter γ are also tested; see the following Table 4.5. However we do not seek α_{exp}, since it will take much time for large m. Our estimated parameter α_{est} makes the HSS iteration a considerable improvement on the parameter $\tilde{\alpha}$.

Table 4.4: α, $\rho(\alpha)$, $\omega(\alpha)$ or $\sigma(\alpha)$ and iteration number(IT) of HSS for Ex. 2 with (4.2), (4.4), (4.5) and $m = 16, \gamma = 1$.

Ex.2 with (4.2), (4.4)			
parameter	$\alpha_{est} = 0.5792 + 1.5799i$	$\tilde{\alpha} = 1.0626$	$\alpha_{exp} = 0.7207 + 1.3139i$
spectral radius	$\rho(\alpha_{est}) = 0.6375$	$\rho(\tilde{\alpha}) = 0.7656$	$\rho(\alpha_{exp}) = 0.6089$
upper bound	$\omega(\alpha_{est}) = 0.6409$	$\sigma(\tilde{\alpha}) = 0.7657$	-
IT	37	61	33
Ex.2 with (4.2), (4.5)			
parameter	$\alpha_{est} = 0.2088 + 2.2906i$	$\tilde{\alpha} = 0.9092$	$\alpha_{exp} = 0.8768 + 1.7830i$
spectral radius	$\rho(\alpha_{est}) = 0.5683$	$\rho(\tilde{\alpha}) = 0.7952$	$\rho(\alpha_{exp}) = 0.5395$
upper bound	$\omega(\alpha_{est}) = 0.5703$	$\sigma(\tilde{\alpha}) = 0.7952$	-
IT	30	74	28

Table 4.5: α, $\rho(\alpha)$, $\omega(\alpha)$ or $\sigma(\alpha)$, and iteration number(IT) of HSS,GMRES, PGMRES(α) for Ex. 2 with (4.2), (4.4) and different m, γ.

	$m = 32$		$m = 48$	
	$\gamma = 2$	$\gamma = 8$	$\gamma = 3$	$\gamma = 12$
α_{est}	$0.3520 + 1.0835i$	$0.2012 + 1.0194i$	$0.2640 + 0.8734i$	$0.0436 + 0.7791i$
$\rho(\alpha_{est})$	0.7368	0.7389	0.7809	0.8148
$\omega(\alpha_{est})$	0.7428	0.7700	0.7891	0.8244
IT	55	47	68	59
$\tilde{\alpha}$	0.6624	0.4696	0.5082	0.1860
$\rho(\tilde{\alpha})$	0.8474	0.8890	0.8808	0.9545
$\sigma(\tilde{\alpha})$	0.8474	0.8897	0.8808	0.9548
IT	97	100	123	192
GMRES	54	62	71	90
PGMRES(α_{est})	14	17	17	23
PGMRES($\tilde{\alpha}$)	21	23	26	30

With these parameters, in Table 4.5 we also report results for full GMRES method and full preconditioned GMRES with an HSS(α) preconditioner (1.6), denoted by PGMRES(α). The code of GMRES we use is from the function *gmres* in *Mablab(7.4 ed)* and the preconditioning systems with respect to $\alpha I + H$ and $\alpha I + S$ are solved by the exact factorization method. We test the real parameter $\tilde{\alpha}$ as well as the complex parameter α_{est}. From these results, we observe that the preconditioner $HSS(\alpha_{est})$ performs much better than the preconditioner $HSS(\tilde{\alpha})$.

Since our estimated parameter α_{est} depends on the extreme eigenvalues of H and S, one may ask how sensitive is the performance of the HSS iteration with respect to the spectrum of H and S. We last use an ′approximation′, denoted by $\tilde{\alpha}_{est}$, to our estimated parameter α_{est} to test the HSS iteration. The ′approximation′ $\tilde{\alpha}_{est}$ is derived by our approach from the approximating eigenvalues of H and S containing 10% noise of the exact eigenvalues. We test the HSS iteration and PGMRES($\tilde{\alpha}_{est}$) for Ex. 2 with (4.2), (4.4) and $m = 32, \gamma = 2$; $m = 48, \gamma = 3$.

For $m = 32, \gamma = 2$, the exact eigenvalues are $\lambda_1 = 8.0221$, $\lambda_n = 0.0547$, $\tau_1 = 8.1271$, $\tau_n = 0.1597$. The approximating eigenvalues are derived by putting +10% error: $\tilde{\lambda}_1 = 8.8243$, $\tilde{\lambda}_n = 0.0602$, $\tilde{\tau}_1 = 8.9398$, $\tilde{\tau}_n = 0.1757$. By these approximating eigenvalues, we can derive the approximation $\tilde{\alpha}_{est} = 0.3545 + 1.2021i$. The detailed numerical results are listed in Table 4.6. Also we put -10% error of the exact eigenvalues to derive an approximation $\tilde{\alpha}_{est} = 0.3167 + 0.9751i$.

The results for this example show that the estimated parameter α_{est} has a normal sensitivity (about 10% error) to the eigenvalues, while the spectral radius $\rho(\alpha_{est})$ is less sensitive (about 2%).

Table 4.6: $\tilde{\alpha}_{est}, \rho(\tilde{\alpha}_{est})$ and iteration number(IT) of HSS($\tilde{\alpha}_{est}$), PGMRES($\tilde{\alpha}_{est}$) for Ex. 2 with (4.2), (4.4) and $m = 32, \gamma = 2$, $m = 48, \gamma = 3$.

	$m = 32, \gamma = 2$		$m = 48, \gamma = 3$	
	+10%	−10%	+10%	−10%
$\tilde{\alpha}_{est}$	$0.3874 + 1.1919i$	$0.3167 + 0.9751i$	$0.2902 + 0.9608i$	$0.2377 + 0.7860i$
$\rho(\tilde{\alpha}_{est})$	0.7577	0.7271	0.7988	0.7760
IT	60	50	75	62
PGMRES($\tilde{\alpha}_{est}$)	15	13	18	16

Achnowledgments. We would like to thank Professor Zhong-Zhi Bai for his careful reading of the manuscript and the comment on the NSS method [7] and on the optimal parameter of the HSS method [2]; and also we are grateful to Dr. Lei-Hong Zhang for polishing the English of the manuscript. This work is supported by the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China [No708040], and the Leading Academic Discipline Program, 211 Project for Shanghai University of Finance and Economics (the 3rd phase).

References

[1] O. Axelsson and A. Kucherov, Real valued iterative methods for solving complex symmetric linear systems, *Numer. Linear Algebra Appl.*, **7** (2000), 197-218.

[2] Z.-Z. Bai, Optimal parameters in the HSS-like methods for saddle-point problems, *Numer. Linear Algebr.*, **16** (2009), 447-479.

[3] Z.-Z. Bai, M. Benzi and F. Chen, Modified HSS iteration methods for a class of complex symmetric linear systems, *Computing*, **87** (2010), 93-111.

[4] Z.-Z. Bai and G.H. Golub, Accelerated Hermitian and skew-Hermitian splitting methods for saddle-point problems, *IMA J. Numer. Anal.*, **27** (2007), 1-23.

[5] Z.-Z. Bai, G.H. Golub and C.-K. Li, Optimal parameter in Hermitian and skew-Hermitian splitting method for certain two-by-two block matrices, *SIAM J. Sci. Comput.*, **28** (2006), 583-603.

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

[7] Z.-Z. Bai, G.H. Golub and M.K. Ng, On successive-overrelaxation acceleration of the Hermitian and skew-Hermitian splitting iterations, *Numer. Linear Algebr.*, **14** (2007), 319-335.

[8] Z.-Z. Bai, G.H. Golub and J.-Y. Pan, Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite linear systems, *Numer. Math.*, **98** (2004), 1-32.

[9] M. Benzi, A generalization of the Hermitian and skew-Hermitian splitting iteration, *SIAM J. Matrix Anal. A.*, **31** (2009), 360-374.

[10] M. Benzi, M.J. Gander and G.H. Golub, Optimization of the Hermitian and skew-Hermitian splitting iteration for saddle-point problems, *BIT Numer. Math.*, **43** (2003), 881-900.

[11] D. Bertaccini, G.H. Golub, S. Serra-Capizzano and C. Tablino Possio, Preconditioned HSS method for the solution of non-Hermitian positive definite linear systems and applications to the discrete convection-diffusion equation, *Numer. Math.*, **99** (2005), 441-484.

[12] C. Cabos and F. Ihlenburg, Vibrational analysis of ships with coupled finite and boundary elements, *J. Comput. Acoust.*, **11** (2003), 91-114.

[13] A. Frommer, BiCGStab(l) for families of shifted linear systems, *Preprint, BUGHW-SC 02/04*, **November** (2002).

[14] C. Greif and J.M. Varah, Iterative solution of skew-symmetric linear systems, *SIAM J. Matrix Anal. A.*, **31** (2009), 584-601.

[15] C. Gu and H. Qian, Skew-symmetric methods for nonsymmetric linear systems with multiple right-hand sides, *J. Comput. Appl. Math.*, **23** (2009), 567-577.

[16] G.-D. Gu, Restarted GMRES augmented with Harmonic-Ritz vectors for shifted linear systems, Intern. J. Comput. Math., **82** (2005), 837-849.

[17] G.-D. Gu and V. Simoncini, Numerical solution of parameter-dependent linear systems, *Numer. Linear Algebr.*, **12** (2005), 923-940.

[18] G.-D. Gu, X.-L. Zhou and L. Lin, A flexible preconditioned Arnoldi method for shifted linear systems, *J. Comput. Math.*, **25** (2007), 522-530.

[19] R. Idema and C. Vuik, A minimal residual method for shifted skew-symmetric systems, *Tech. report 07-09, Delft Univ. Technology, Delft, The Netherlands*, (2007).

[20] E. Jiang, Algorithm for solving shifted skew-symmetric linear system, *Front. Math. China*, **2** (2007), 227-242.

[21] M. Kuzuoglu and R. Mittra, Finite element solution of electromagnetic problem over a wide frequency range via the Pade approximation, *Comput. Method. Appl. M.*, **169** (1999), 263-277.

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

[23] V. Simoncini and M. Benzi, Spectral properties of the Hermitian and skew-Hermitian splitting preconditioner for saddle point problems, *SIAM J. Matrix Anal. A.*, **26** (2004), 377-389.