Convergence Properties of Preconditioned Hermitian and Skew-Hermitian Splitting Methods for Non-Hermitian Positive Semidefinite Matrices

Zhong-Zhi Bai [1]
Department of Mathematics, Fudan University
Shanghai 200433, P.R. China
and
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, P.R. China [2]
bzz@lsec.cc.ac.cn

Gene H. Golub [3]
Scientific Computing and Computational Mathematics Program
Department of Computer Science
Stanford University, Stanford, CA 94305-9025, USA
golub@sccm.stanford.edu

Chi-Kwong Li [4]
Department of Mathematics
The College of William & Mary
P.O. Box 8795, Williamsburg, VA 23187-8795, USA
ckli@math.wm.edu

September 6, 2004

[1]Subsidized by The Special Funds For Major State Basic Research Projects (No. G1999032803) and The National Natural Science Foundation (No. 10471146), P.R. China; and The 2004 Ky and Yu-Fen Fan Fund Travel Grant of American Mathematical Society
[2]Correspondence and permanent address
[3]The work of this author was in part supported by the Department of Energy: DE-FC02-01ER41177
[4]Research partially supported by an NSF grant

Abstract

For the non-Hermitian and positive semidefinite systems of linear equations, we derive sufficient and necessary conditions for guaranteeing the unconditional convergence of the preconditioned Hermitian and skew-Hermitian splitting iteration methods. These result is specifically applied to linear systems of block tridiagonal form to obtain convergence conditions for the corresponding block variants of the preconditioned Hermitian and skew-Hermitian splitting iteration methods.

Keywords: Non-Hermitian matrix, positive semidefinite matrix, Hermitian and skew-Hermitian splitting, splitting iteration method, convergence.

AMS(MOS) Subject Classifications: 65F10, 65F50; CR: G1.3.

1 Introduction

We consider iterative solution of the large sparse non-Hermitian system of linear equations

$$Ax = b,\ A \in \mathbf{C}^{n\times n} \text{ nonsingular},\ A \neq A^*,\ \text{and } x, b \in \mathbf{C}^n, \tag{1.1}$$

where A^* denotes the conjugate transpose of the complex matrix A.

Based on the Hermitian and skew-Hermitian (HS) splitting

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

Bai, Golub and Ng recently established a class of Hermitian and skew-Hermitian splitting (HSS) iteration methods in [3] for solving the non-Hermitian system of linear equations (1.1).

When the coefficient matrix $A \in \mathbf{C}^{n\times n}$ is positive definite, i.e., its Hermitian part $\mathcal{H}(A) \in \mathbf{C}^{n\times n}$ is Hermitian positive definite, they proved in [3] that the HSS iteration converges unconditionally to the exact solution of the system of linear equations (1.1), with the bound on the rate of convergence about the same as that of the conjugate gradient method when applied to the Hermitian matrix $\mathcal{H}(A)$. Moreover, **the upper bound** of the contraction factor is dependent on the spectrum of the Hermitian part $\mathcal{H}(A)$, but is independent of the spectrum of the skew-Hermitian part $\mathcal{S}(A)$ as well as the eigenvalues of the matrices $\mathcal{H}(A)$, $\mathcal{S}(A)$ and A. Numerical experiments have shown that the HSS iteration method is very efficient and robust for solving the non-Hermitian and positive definite linear systems, see [3].

When the coefficient matrix $A \in \mathbf{C}^{n\times n}$ has the two-by-two block structure

$$A = \begin{pmatrix} B & E \\ -E^* & C \end{pmatrix}, \tag{1.2}$$

with $B \in \mathbf{C}^{p\times p}$ being positive definite (i.e., $\mathcal{H}(B)$ is Hermitian positive definite), $C \in \mathbf{C}^{q\times q}$ being Hermitian positive semidefinite and $E \in \mathbf{C}^{p\times q}$ being of full column rank, Benzi and Golub further proved in [8] that the HSS iteration method for the corresponding saddle-point problem

$$Ax \equiv \begin{pmatrix} B & E \\ -E^* & C \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \equiv b \tag{1.3}$$

also converges unconditionally to its exact solution. Note that the matrix A is now only positive semidefinite with some special structure, namely, its Hermitian part

$$\mathcal{H}(A) = \begin{pmatrix} \mathcal{H}(B) & 0 \\ 0 & C \end{pmatrix} \in \mathbf{C}^{n\times n}$$

is such that $\mathcal{H}(B)$ is positive definite and C is Hermitian positive semidefinite.

In this paper, we give a necessary and sufficient condition for an arbitrary non-Hermitian positive **semidefinite** linear system so that the *preconditioned Hermitian and skew-Hermitian splitting* (PHSS) iteration method will lead to an unconditionally convergent iteration sequence. This result is further specialized to linear systems of block tridiagonal form to obtain unconditional convergence conditions for the corresponding *block PHSS* (BPHSS) iteration method.

2 The Preconditioned HSS Method

Instead of applying the HSS iteration technique directly to the system of linear equations (1.1), we may apply it to the systematically preconditioned linear system

$$\widehat{A}\widehat{x} = \widehat{b}, \text{ with } \widehat{A} = R^{-*}AR^{-1}, \ \widehat{x} = Rx \text{ and } \widehat{b} = R^{-*}b, \tag{2.1}$$

where $R \in \mathbf{C}^{n\times n}$ is a prescribed nonsingular matrix and $R^{-*} = (R^{-1})^* = (R^*)^{-1}$. Let $P = R^*R$. Then $P \in \mathbf{C}^{n\times n}$ is a Hermitian positive definite matrix. This leads to the preconditioned Hermitian and skew-Hermitian splitting (PHSS) iteration method as follows. See also [2, 3, 4] and [8, 9].

The PHSS Iteration Method.

Let $P \in \mathbf{C}^{n\times n}$ be a prescribed Hermitian positive definite matrix. Given an initial guess $x^{(0)} \in \mathbf{C}^n$, compute $x^{(k)}$ for $k = 0, 1, 2, \ldots$ using the following iteration scheme until $\{x^{(k)}\}$ satisfies the stopping criterion:

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

where α is a given positive constant.

Clearly, when $P = I$, the identity matrix, the PHSS iteration method reduces to the HSS iteration method studied in Bai, Golub and Ng in [3]. When $P \neq I$, we can suitably choose P and α such that the induced PHSS iteration method possesses fast convergence and high computing efficiency. In addition, the Hermitian positive definite matrix P and the positive constant α should be judiciously selected so that the two sub-systems of linear equations with the coefficient matrices $\alpha P + \mathcal{H}(A)$ and $\alpha P + \mathcal{S}(A)$ can be solved economically and rapidly.

In matrix-vector form, the above PHSS iteration method can be rewritten as

$$x^{(k+1)} = \mathcal{L}(\alpha, P)x^{(k)} + \mathcal{G}(\alpha, P)b, \qquad k = 0, 1, 2, \ldots, \tag{2.2}$$

where

$$\mathcal{L}(\alpha, P) = (\alpha P + \mathcal{S}(A))^{-1}(\alpha P - \mathcal{H}(A))(\alpha P + \mathcal{H}(A))^{-1}(\alpha P - \mathcal{S}(A))$$

and

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

Here, $\mathcal{L}(\alpha, P)$ is the iteration matrix of the PHSS iteration method. In fact, (2.2) may also result from the splitting

$$A = \mathcal{M}(\alpha, P) - \mathcal{N}(\alpha, P)$$

of the coefficient matrix A, with

$$\begin{cases} \mathcal{M}(\alpha, P) &= \frac{1}{2\alpha}(\alpha P + \mathcal{H}(A))(\alpha P + \mathcal{S}(A)), \\ \mathcal{N}(\alpha, P) &= \frac{1}{2\alpha}(\alpha P - \mathcal{H}(A))(\alpha P - \mathcal{S}(A)). \end{cases}$$

Therefore, the PHSS iteration method can naturally induce a preconditioner $\mathcal{M}(\alpha, P)$ to the matrix A. This preconditioner is called as the PHSS preconditioner. See [2, 4, 8, 9].

When $A \in \mathbf{C}^{n\times n}$ is a positive definite matrix, duplicating the proofs of Theorem 2.2 and Corollary 2.3 in [3] we can establish the following convergence theorem for the PHSS iteration method. In the sequel, $\mathrm{sp}(X)$ represents the spectrum of the square matrix X.

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

$$\sigma(\alpha, P) = \max_{\lambda_j \in \mathrm{sp}(P^{-1}\mathcal{H}(A))} \frac{|\alpha - \lambda_j|}{|\alpha + \lambda_j|}.$$

Consequently, we have

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

i.e., the PHSS iteration unconditionally converges to the exact solution of the system of linear equations (1.1).

Moreover, if $\gamma_{\min}$ and $\gamma_{\max}$ are the lower and the upper bounds of the eigenvalues of the matrix $P^{-1}\mathcal{H}(A)$, respectively, then

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

and

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

where $\kappa(P^{-1}\mathcal{H}(A))$ is the spectral condition number of the matrix $P^{-1}\mathcal{H}(A)$.

¿From Theorem 2.1 we see that the Hermitian positive definite matrix $P \in \mathbf{C}^{n\times n}$ should be chosen such that it is at least a good approximate to the matrix $\mathcal{H}(A)$. In this situation, $\kappa(P^{-1}\mathcal{H}(A))$ may be reasonably small so that the PHSS iteration method may achieve a fast convergence speed. On the other hand, since we often have to solve the two half-iterates $x^{(k+\frac{1}{2})}$ and $x^{(k+1)}$ inexactly by some iteration schemes, P and α should be chosen such that both matrices $\alpha P + \mathcal{H}(A)$ and $\alpha P + \mathcal{S}(A)$ are well conditioned and economically invertible. Hence, in a practical computation, it is crucial but a difficult problem to determine a good preconditioner P and choose an optimal iteration parameter α. For some discussions on this aspect, we refer the readers to [2, 4, 8, 9].

3 Convergence Theorems

In this section, we study the convergence properties of the PHSS iteration method when the coefficient matrix $A \in \mathbf{C}^{n\times n}$ is positive semidefinite. To this end, we call an eigenvalue λ of a matrix $W \in \mathbf{C}^{n\times n}$ a reducing eigenvalue if $Wx = \lambda x$ and $W^*x = \lambda^* x$. Equivalently, W is unitarily similar to $[\lambda] \oplus W_0$ where $W_0 \in \mathbf{C}^{(n-1)\times(n-1)}$.

The following theorem describes the convergence property of the PHSS iteration method when the coefficient matrix $A \in \mathbf{C}^{n\times n}$ is positive **semidefinite**.

Theorem 3.1. *Let $A \in \mathbf{C}^{n\times n}$ be a positive semidefinite matrix, $\mathcal{H}(A) = \frac{1}{2}(A + A^*)$ and $\mathcal{S}(A) = \frac{1}{2}(A - A^*)$ be its Hermitian and skew-Hermitian parts, respectively, and α be a positive constant. Let $P \in \mathbf{C}^{n\times n}$ be a Hermitian positive definite matrix. Then the spectral radius $\rho(\mathcal{L}(\alpha, P))$ of the iteration matrix $\mathcal{L}(\alpha, P)$ of the PHSS iteration is bounded by 1, i.e.,*

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

The inequality becomes an equality if and only if the matrix $\widehat{A} := R^{-*}AR^{-1}$ has an (reducing) eigenvalue of the form $i\xi$ with $\xi \in \mathbf{R}$ and i the imaginary unit, i.e., the null space of $\mathcal{H}(\widehat{A})$ contains an eigenvector of $\mathcal{S}(\widehat{A})$. Here, $P = R^*R$ and $R \in \mathbf{C}^{n\times n}$ is a prescribed nonsingular matrix.

Proof. Evidently, we only need to consider the case when $P = I$, as otherwise, we can turn to the preconditioned linear system (2.1) instead. Denote by

$$\mathcal{L}(\alpha) := \mathcal{L}(\alpha, I) = (\alpha I + \mathcal{S}(A))^{-1}(\alpha I - \mathcal{H}(A))(\alpha I + \mathcal{H}(A))^{-1}(\alpha I - \mathcal{S}(A)),$$

which is similar to the matrix

$$\overline{\mathcal{L}}(\alpha) := (\alpha I + \mathcal{H}(A))^{-1}(\alpha I - \mathcal{H}(A))(\alpha I + \mathcal{S}(A))^{-1}(\alpha I - \mathcal{S}(A)).$$

Therefore, we only need to investigate the property of the eigenvalues of $\overline{\mathcal{L}}(\alpha)$.

Suppose that $\mathcal{H}(A)$ has eigenvalues

$$\mu_1 \geq \cdots \geq \mu_r > 0 = \mu_{r+1} = \cdots = \mu_n = 0.$$

Then $(\alpha I + \mathcal{H}(A))^{-1}(\alpha I - \mathcal{H}(A))$ is Hermitian and has eigenvalues

$$\nu_j = (\alpha - \mu_j)/(\alpha + \mu_j), \qquad j = 1, 2, \ldots, n,$$

so that

$$-1 < \nu_1 \leq \cdots \leq \nu_r < 1 = \nu_{r+1} = \cdots = \nu_n. \tag{3.1}$$

Hence $\overline{\mathcal{L}}(\alpha)$ has singular values

$$|\nu_j| \leq 1, \qquad j = 1, 2, \ldots, n.$$

Consequently,

$$\rho(\overline{\mathcal{L}}(\alpha)) \leq \|\overline{\mathcal{L}}(\alpha)\| = 1, \qquad \forall \alpha > 0.$$

Suppose that A has an eigenvalue of the form $i\xi$ with $\xi \in \mathbf{R}$ corresponding to a unit eigenvector v. We show that $i\xi$ is in fact a reducing eigenvalue of A. To see this, let V be a unitary matrix such that V^*AV is in lower triangular form with $i\xi$ in the $(1, 1)$ entry, and w as the first column. Then

$$\mathcal{H}(V^*AV) = V^*\mathcal{H}(A)V$$

is positive semidefinite with 0 in the $(1, 1)$ entry and $\frac{1}{2}w$ as the first column. It follows that $w = 0$ and $U^*AU = [i\xi] \oplus A_0$ for some $A_0 \in \mathbf{C}^{(n-1)\times(n-1)}$, i.e., $i\xi$ is a reducing eigenvalue of A. Now,

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

with $\lambda = (\alpha - i\xi)/(\alpha + i\xi)$ such that $|\lambda| = 1$.

Conversely, if $\overline{\mathcal{L}}(\alpha)$ has an eigenvalue λ of modulus 1, then λ is an eigenvalue of the matrix

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

Thus, there is a unit vector $v \in \mathbf{C}^n$ such that

$$(\alpha I + \mathcal{S}(A))^{-1}(\alpha I - \mathcal{S}(A))(\alpha I + \mathcal{H}(A))^{-1}(\alpha I - \mathcal{H}(A))v = \lambda v.$$

Since $|\lambda| = 1$ and $(\alpha I + \mathcal{S}(A))^{-1}(\alpha I - \mathcal{S}(A))$ is unitary, we see that

$$\|(\alpha I + \mathcal{H}(A))^{-1}(\alpha I - \mathcal{H}(A))v\| = \|v\|.$$

Suppose that $\{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n\}$ is an orthonormal basis for $\mathbf{C}^n$ consisting of eigenvectors of $\mathcal{H}(A)$ such that $\mathcal{H}(A)\mathrm{x}_j = \mu_j \mathrm{x}_j$ for $j = 1, 2, \ldots, n$. Let $v = \sum_{j=1}^n \theta_j \mathrm{x}_j$ with $\theta_j \in \mathbf{C}$. Then

$$1 = \sum_{j=1}^n |\theta_j|^2 = \|v\|^2 = \|(\alpha I + \mathcal{H}(A))^{-1}(\alpha I - \mathcal{H}(A))v\|^2 = \sum_{j=1}^n |\theta_j|^2 \nu_j^2.$$

By (3.1), we know that $\theta_j = 0$ for $j = 1, 2, \ldots, r$. It follows that $v = \sum_{j=r+1}^n \theta_j \mathrm{x}_j$ and

$$\mathcal{H}(A)v = \sum_{j=r+1}^n \theta_j \mathcal{H}(A)\mathrm{x}_j = \sum_{j=r+1}^n \theta_j \mu_j \mathrm{x}_j = 0.$$

Furthermore,

$$\begin{aligned} \lambda v &= (\alpha I + \mathcal{S}(A))^{-1}(\alpha I - \mathcal{S}(A))(\alpha I + \mathcal{H}(A))^{-1}(\alpha I - \mathcal{H}(A))v \\ &= (\alpha I + \mathcal{S}(A))^{-1}(\alpha I - \mathcal{S}(A))v. \end{aligned}$$

Thus, v is an eigenvector of $(\alpha I + \mathcal{S}(A))^{-1}(\alpha I - \mathcal{S}(A))$, and hence v is an eigenvector of $\mathcal{S}(A)$ such that $\mathcal{S}(A)v = i\xi v$ with $\xi \in \mathbf{R}$ satisfying $\lambda = (\alpha - i\xi)/(\alpha + i\xi)$. As a result, $Av = i\xi v$ and $A^*v = -i\xi v$. So, $i\xi$ is a reducing eigenvalue of A. □

Corollary 3.2. *Suppose that $A \in \mathbf{C}^{n\times n}$ satisfies the hypothesis of Theorem 3.1. If $\rho(\mathcal{L}(\alpha, P)) < 1$, then A is nonsingular.*

The contra-positive of the above corollary asserts that if a matrix A satisfying the hypothesis of Theorem 3.1 is singular, then $\rho(\mathcal{L}(\alpha, P)) = 1$. Note also that such an A is singular if and only if 0 is a reducing eigenvalue. This happens if and only if $\mathcal{H}(A)$ and $\mathcal{S}(A)$ have a common null vector.

Note that in general, a matrix may have an eigenvalue of the form $i\xi$ which is not a reducing eigenvalue. However, this cannot happen for matrices A such that $\mathcal{H}(A)$ is positive semidefinite.

For matrices A such that $\mathcal{H}(A)$ is positive semidefinite, we need to determine whether it has no (reducing) eigenvalue of the form $i\xi$ with $\xi \in \mathbf{R}$. The next proposition gives some information along this direction.

Proposition 3.3. *Suppose that $A \in \mathbf{C}^{n\times n}$ satisfies the hypothesis of Theorem 3.1. Then the following statements are equivalent:*

(a) A does not have an (reducing) eigenvalue of the form $i\xi$ with $\xi \in \mathbf{R}$;

(b) The null space of $\mathcal{H}(A)$ does not contain an eigenvector of $\mathcal{S}(A)$;

(c) If v is an eigenvector of $\mathcal{S}(A)$, then $v^*\mathcal{H}(A)v > 0$;

(d) Let V be unitary such that $V^*\mathcal{H}(A)V = H_1 \oplus 0_\ell$ where H_1 is nonsingular, and let

$$V^*\mathcal{S}(A)V = \begin{pmatrix} S_1 & E \\ -E^* & S_2 \end{pmatrix}.$$

Then the null space of E does not contain an eigenvector of S_2.

Proof. The equivalence of (a), (b) and (c) are straightforward. Now we consider (d). Suppose that V is unitary such that $V^*\mathcal{H}(A)V = H_1 \oplus 0_\ell$, where H_1 is nonsingular, and

$$V^*\mathcal{S}(A)V = \begin{pmatrix} S_1 & E \\ -E^* & S_2 \end{pmatrix}.$$

Then a vector in the null space of $\mathcal{H}(A)$ must be of the form

$$V \begin{pmatrix} 0 \\ x \end{pmatrix}, \quad \text{with} \quad x \in \mathbf{C}^\ell.$$

Furthermore, it is an eigenvector of $\mathcal{S}(A)$ corresponding to the eigenvalue $i\xi$ with $\xi \in \mathbf{R}$ if and only if $Ex = 0$ and $S_2x = i\xi x$. Thus, (a) and (d) are equivalent. □

¿From Theorem 3.1 and Proposition 3.3, one can easily deduce the convergence results on the HSS iteration method for positive definite matrices in [3] and on the PHSS iteration methods for special positive semidefinite saddle-point matrices (1.2) in [2, 4, 8]. In the former case, it is clear that no eigenvalue has the form $i\xi$ with $\xi \in \mathbf{R}$. In the lattter case, under the assumption that E has full column rank, condition (d) in Proposition 3.3 cannot hold.

The following example shows that even if A is in two-by-two block form with

$$\mathcal{H}(A) = \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix} \quad \text{and} \quad \mathcal{S}(A) = \begin{pmatrix} S_1 & E \\ -E^* & S_2 \end{pmatrix},$$

the relation between the null spaces of the matrices H_1, H_2, E and E^* may not be too useful in determining whether $\rho(\mathcal{L}(\alpha, P)) < 1$. See [8].

Example 3.4. Suppose that $A \in \mathbf{C}^{n\times n}$ is such that

$$\mathcal{H}(A) = \begin{pmatrix} H_0 & 0 \\ 0 & H_0 \end{pmatrix} \quad \text{with} \quad H_0 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

(a) If

$$\mathcal{S}(A) = \begin{pmatrix} 0 & E \\ -E^* & 0 \end{pmatrix} \quad \text{with} \quad E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

then A has eigenvalues $\pm i$. So, $\rho(\mathcal{L}(\alpha, I)) = 1$. However, it holds that $\text{null}(H_0) \cap \text{null}(E) = \{0\}$;

(b) If

$$\mathcal{S}(A) = \begin{pmatrix} 0 & E \\ -E^* & 0 \end{pmatrix} \quad \text{with} \quad E = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

then A has no eigenvalue of the form $i\xi$ with $\xi \in \mathbf{R}$. So, $\rho(\mathcal{L}(\alpha, I)) < 1$. However, it holds that $\text{null}(H_0) \cap \text{null}(E) = \{0\}$;

(c) If

$$\mathcal{S}(A) = \begin{pmatrix} S_0 & 0 \\ 0 & S_0 \end{pmatrix} \quad \text{with} \quad S_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

then A has no eigenvalue of the form $i\xi$ with $\xi \in \mathbf{R}$. So, $\rho(\mathcal{L}(\alpha, I)) < 1$. However, it holds that $\text{null}(H_0) \cap \text{null}(E) = \{0\}$;

(d) If

$$\mathcal{S}(A) = \begin{pmatrix} S_0 & E \\ -E^* & S_0 \end{pmatrix} \quad \text{with} \quad S_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

so that $\|E\|$ is small, then A has eigenvalues close to 1. Thus, $\rho(\mathcal{L}(\alpha, I)) = 1$. However, we can choose E such that either $\text{null}(H_0) \cap \text{null}(E) = \{0\}$ or $\text{null}(H_0) \cap \text{null}(E) \neq \{0\}$ holds.

If $\mathcal{H}(A)$ or $\mathcal{S}(A)$ is in diagonal block form

$$B_1 \oplus B_2 \oplus \cdots \oplus B_\ell, \quad \text{with} \quad B_j \in \mathbf{C}^{n_j \times n_j} \, (j = 1, 2, \ldots, \ell),$$

then one can consider

$$P = \alpha_1 I_{n_1} \oplus \alpha_2 I_{n_2} \oplus \cdots \oplus \alpha_\ell I_{n_\ell}, \quad \text{with} \quad \alpha_j > 0 \, (j = 1, 2, \ldots, \ell).$$

As long as $\widehat{A} := P^{-1/2} A P^{-1/2}$ does not have an eigenvalue of the form $i\xi$ with $\xi \in \mathbf{R}$, then the iteration matrix $\mathcal{L}(\alpha, P)$ of the PHSS iteration method has spectral radius less than one, i.e., the PHSS iteration scheme converges. In particular, when $\ell = 1$, this conclusion recovers the convergence theorem established in [8].

We remark that one may relax the condition that $\mathcal{H}(A) = \frac{1}{2}(A + A^*)$ is positive semidefinite. In fact, if there exists a $\theta \in [0, 2\pi)$ such that $\mathcal{H}(e^{i\theta} A) = \frac{1}{2}(e^{i\theta} A + e^{-i\theta} A^*)$ is positive semidefinite, then one can apply Theorem 3.1 to $e^{i\theta} A$. This latter condition is equivalent to the fact that the numerical range of A defined by

$$\mathcal{W}(A) := \{v^* A v : v \in \mathbf{C}^n, \ v^* v = 1\}$$

lies on a closed half plane defined by a line passing through the origin. In particular, if $\theta = \frac{\pi}{2}$ is such that $\mathcal{H}(e^{i\theta} A)$ is positive semidefinite, then the PHSS iteration method resulted from interchanging the Hermitian matrix $\mathcal{H}(A)$ and the skew-Hermitian matrix $\mathcal{S}(A)$ may still converge. Note that in this PHSS iteration method the right-hand-side vector b is replaced by ib, correspondingly.

4 Applications

We consider the non-Hermitian system of linear equations (1.1) whose coefficient matrix $A \in \mathbf{C}^{n\times n}$ is in the block tridiagonal form, i.e.,

$$Ax \equiv \begin{pmatrix} A_1 & E_1 & & & \\ -E_1^* & A_2 & E_2 & & \\ & \ddots & \ddots & \ddots & \\ & & -E_{\ell-2}^* & A_{\ell-1} & E_{\ell-1} \\ & & & -E_{\ell-1}^* & C \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\ell-1} \\ x_\ell \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{\ell-1} \\ b_\ell \end{pmatrix} \equiv b, \qquad (4.1)$$

where $A_j \in \mathbf{C}^{n_j\times n_j}$ $(j = 1, 2, \ldots, \ell-1)$ are non-Hermitian matrices, $E_j \in \mathbf{C}^{n_j\times n_{j+1}}$ $(j = 1, 2, \ldots, \ell-1)$, $C \in \mathbf{C}^{n_\ell\times n_\ell}$ is a Hermitian matrix, $x_j, b_j \in \mathbf{C}^{n_j}$, and n_j $(j = 1, 2, \ldots, \ell)$ are positive integers satisfying $n_1 \geq n_2 \geq \cdots \geq n_\ell$ and $\sum_{j=1}^{\ell} n_j = n$.

The block tridiagonal systems of linear equations may arise from many applications, e.g., the remaining (linearized) Euler-Lagrange equations [24, 25] and a coupled DEM-FEM formulation combined with Lagrange multipliers in the imperious porous material with an incompressible pore fluid[20].

In particular, when $\ell = 2$, the system of linear equations (4.1) reduces to the generalized saddle-point problem (1.3). As is known, saddle-point problems correspond to the Kuhn-Tucker conditions for linearly constrained quadratic programming problems, which 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[10]) and from Lagrange multiplier methods (see, e.g., Fortin and Glowinski[14]). A number of structured preconditioners [12, 13, 21, 8] and iterative methods [11, 19, 4, 2] have been studied in the literature for these problems. See also [23, 18, 17, 22, 16, 15, 6] and the references therein.

In this section we consider the block tridiagonal systems of linear equations satisfying all of the following assumptions:

- A_1 is positive definite, i.e., $\mathcal{H}(A_1)$ is Hermitian positive definite;
- A_j $(j = 2, 3, \ldots, \ell-1)$ are positive semidefinite, i.e., $\mathcal{H}(A_j)$ $(j = 2, 3, \ldots, \ell-1)$ are Hermitian positive semidefinite;
- E_j $(j = 1, 2, \ldots, \ell-2)$ are of full column rank;
- C is Hermitian positive semidefinite;
- $\mathrm{null}(C) \cap \mathrm{null}(E_{\ell-1}) = \{0\}$.

As shown below, these assumptions guarantee existence and uniqueness of the solution.

Proposition 4.1. *Let $A \in \mathbf{C}^{n\times n}$ be the coefficient matrix of the system of linear equations (4.1). Assume that A_1 is positive definite, A_j $(j = 2, 3, \ldots, \ell-1)$ are positive semidefinite, E_j $(j = 1, 2, \ldots, \ell-2)$ are of full column ranks, C is Hermitian positive semidefinite, and* $\mathrm{null}(C) \cap \mathrm{null}(E_{\ell-1}) = \{0\}$. *Then A is nonsingular.*

Proof. Let $x = (x_1^*, x_2^*, \ldots, x_\ell^*)^* \in \mathbf{C}^n$ be such that $Ax = 0$, where $x_j \in \mathbf{C}^{n_j}$ for $j = 1, 2, \ldots, \ell$. Then

$$\begin{cases} A_1 x_1 + E_1 x_2 = 0, \\ -E_{j-1}^* x_{j-1} + A_j x_j + E_j x_{j+1} = 0, \quad j = 2, 3, \ldots, \ell - 1, \\ -E_{\ell-1}^* x_{\ell-1} + C_\ell x_\ell = 0. \end{cases} \tag{4.2}$$

Because $Ax = 0$ implies both $x^* A x = 0$ and $x^* A^* x = 0$, we know that $x^* \mathcal{H}(A) x = 0$. As $\mathcal{H}(A_j)$ $(j = 1, 2, \ldots, \ell - 1)$ and C are Hermitian positive semidefinite, $\mathcal{H}(A)$ is Hermitian positive semidefinite, too. Hence, $x \in \text{null}(\mathcal{H}(A))$, or equivalently,

$$Cx_\ell = 0 \quad \text{and} \quad \mathcal{H}(A_j) x_j = 0 \quad \text{for } j = 1, 2, \ldots, \ell - 1.$$

The system of linear equations (4.2) then reduces to the following:

$$\begin{cases} \mathcal{S}(A_1) x_1 + E_1 x_2 = 0, \\ -E_{j-1}^* x_{j-1} + \mathcal{S}(A_j) x_j + E_j x_{j+1} = 0, \quad j = 2, 3, \ldots, \ell - 1, \\ -E_{\ell-1}^* x_{\ell-1} = 0. \end{cases} \tag{4.3}$$

Since $\mathcal{H}(A_1)$ is Hermitian positive definite, we see that $x_1 = 0$. Based on (4.3) and the assumption that E_j has full column rank for $j = 1, 2, \ldots, \ell - 2$, we can successively obtain $x_j = 0$ for $j = 1, 2, \ldots, \ell - 1$. Thereby, (4.3) can be further reduced to

$$E_{\ell-1} x_\ell = 0.$$

Since $Cx_\ell = 0$, we conclude that $x_\ell \in \text{null}(C) \cap \text{null}(E_{\ell-1})$, which is $\{0\}$ by our assumption. Hence, $x_\ell = 0$. Therefore, the only solution for $Ax = 0$ is the trivial solution, and A is nonsingular. □

For the PHSS iteration method described in § 2, if we first specifically take the Hermitian positive definite matrix $P \in \mathbf{C}^{n \times n}$ to be of block diagonal form, i.e.,

$$P = \text{Diag}\left(\frac{\alpha_1}{\alpha} P_1, \frac{\alpha_2}{\alpha} P_2, \ldots, \frac{\alpha_\ell}{\alpha} P_\ell\right)$$

with

$$\alpha_j > 0 \quad \text{and} \quad P_j \in \mathbf{C}^{n_j \times n_j} \text{ Hermitian positive definite}, \quad j = 1, 2, \ldots, \ell,$$

and then directly apply it to the block tridiagonal system of linear equations (4.1), the following iteration scheme, called the *block preconditioned Hermitian and skew-Hermitian splitting* (BPHSS) iteration method, can be obtained immediately.

The BPHSS Iteration Method.

Let $P_j \in \mathbf{C}^{n_j \times n_j}$ $(j = 1, 2, \ldots, \ell)$ be prescribed Hermitian positive definite matrices and α_j $(j = 1, 2, \ldots, \ell)$ be given positive constants. Given an initial guess

$$x^{(0)} = \left(x_1^{(0)^*}, x_2^{(0)^*}, \ldots, x_\ell^{(0)^*}\right)^* \in \mathbf{C}^n \quad \text{with} \quad x_j^{(0)} \in \mathbf{C}^{n_j \times n_j},$$

compute

$$x^{(k)} = \left(x_1^{(k)^*}, x_2^{(k)^*}, \ldots, x_\ell^{(k)^*}\right)^* \in \mathbf{C}^n \quad \text{with} \quad x_j^{(k)} \in \mathbf{C}^{n_j \times n_j}$$

for $k = 0, 1, 2, \ldots$ using the following iteration scheme until $\{x^{(k)}\}$ satisfies the stopping criterion:

- `Solve` $x_j^{(k+\frac{1}{2})}$ `(`$j = 1, 2, \ldots, \ell$`) successively from the sub-systems of linear equations`

$$(\alpha_j P_j + \mathcal{H}(A_j))x_j^{(k+\frac{1}{2})} = (\alpha_j P_j - \mathcal{S}(A_j))x_j^{(k)} + E_{j-1}^* x_{j-1}^{(k)} - E_{j+1}x_{j+1}^{(k)} + b_j, \quad j = 1, 2, \ldots, \ell,$$

- `Solve` $x_j^{(k+1)}$ `(`$j = 1, 2, \ldots, \ell$`) from the system of linear equations`

$$(\alpha_j P_j + \mathcal{S}(A_j))x_j^{(k+1)} - E_{j-1}^* x_{j-1}^{(k+1)} + E_{j+1}x_{j+1}^{(k+1)} = (\alpha_j P_j - \mathcal{H}(A_j))x_j^{(k+\frac{1}{2})} + b_j, \quad j = 1, 2, \ldots, \ell.$$

`Here, we have stipulated that` $x_0^{(k)} = x_0^{(k+1)} = 0$ `and` $x_{\ell+1}^{(k)} = x_{\ell+1}^{(k+1)} = 0$.

Note that in the BPHSS iteration method, for each fixed iteration index k, the block vectors $x_j^{(k+\frac{1}{2})}$ $(j = 1, 2, \ldots, \ell)$ can be computed independently and, hence, the vector $x^{(k+\frac{1}{2})}$ can be easily obtained on a multiprocessor system. Comparatively, the block vectors $x_j^{(k+1)}$ $(j = 1, 2, \ldots, \ell)$ are more dependent, which may cause difficulty in solving the second-half iterate $x^{(k+1)}$ in parallel. However, there are efficient direct and iterative methods for solving this special class of block tridiagonal systems of linear equations, see [18, 5, 1, 7]. Therefore, the BPHSS iteration method can be easily and effectively implemented in parallel on a multiprocessor system.

In addition, in actual computing it may be beneficial in solving the second-half iterate $x^{(k+1)}$ if we first execute block re-ordering for the system of linear equations (4.1), although this does not change the sub-system of linear equations defining the first-half iterate $x^{(k+\frac{1}{2})}$.

The following theorem describes the convergence property of the BPHSS iteration method.

Theorem 4.2. *Let all conditions of Proposition 4.1 be satisfied. Then the BPHSS iteration scheme is unconditionally convergent; that is, the spectral radius of its iteration matrix* $\mathcal{L}(\alpha_1, \alpha_2, \ldots, \alpha_\ell; P)$ *satisfies*

$$\rho(\mathcal{L}(\alpha_1, \alpha_2, \ldots, \alpha_\ell; P)) < 1, \quad \textit{for all} \quad \alpha_1, \alpha_2, \ldots, \alpha_\ell > 0.$$

Proof. Without loss of generality, we only need to consider the case that $P = I$, as otherwise, we can turn to the preconditioned linear system (2.1) instead.

To prove the unconditional convergence of the BPHSS iteration method, according to Theorem 3.1 we only need to show that the null space of $\mathcal{H}(A)$ does not contain an eigenvector of $\mathcal{S}(A)$. In fact, if there exists a nonzero vector $x = (x_1^*, x_2^*, \ldots, x_\ell^*)^* \in \mathbf{C}^n$, with $x_j \in \mathbf{C}^{n_j}$ $(j = 1, 2, \ldots, \ell)$, such that $\mathcal{H}(A)x = 0$ and $\mathcal{S}(A)x = i\xi x$ hold for some $\xi \in \mathbf{R}$, i.e.,

$$Cx_\ell = 0, \qquad \mathcal{H}(A_j)x_j = 0 \quad \text{for} \quad j = 1, 2, \ldots, \ell - 1,$$

and

$$\begin{cases} \mathcal{S}(A_1)x_1 + E_1 x_2 = i\xi x_1, \\ -E_{j-1}^* x_{j-1} + \mathcal{S}(A_j)x_j + E_j x_{j+1} = i\xi x_j \quad \text{for} \quad j = 2, 3, \ldots, \ell - 1, \\ -E_{\ell-1}^* x_{\ell-1} = i\xi x_\ell, \end{cases} \tag{4.4}$$

then from the Hermitian positive definiteness of the matrix $\mathcal{H}(A_1)$ we know that $x_1 = 0$. It then follows from the full-rank assumption of the matrices E_j $(j = 1, 2, \ldots, \ell - 2)$ that $x_j = 0$ $(j = 1, 2, \ldots, \ell - 1)$. Thereby, (4.4) can be further reduced to

$$E_{\ell-1}x_\ell = 0 \quad \text{and} \quad i\xi x_\ell = 0.$$

Evidently, whether $\xi = 0$ or not we can obtain $x_\ell = 0$ due to the assumption $\text{null}(C) \cap \text{null}(E_{\ell-1}) = \{0\}$. Therefore, $x = 0$, a contradiction. □

By suitable re-ordering or re-decomposing the block tridiagonal system of linear equations (4.1) can be reformulated as one with a two-by-two block coefficient matrix, or in the form of saddle-point problems. But now the $(1, 1)$-block of the newly obtained two-by-two block matrix is not positive definite, even though its $(2, 2)$-block is Hermitian positive semidefinite and the overlapping set between the null spaces of its $(1, 2)$-block and $(2, 2)$-block is $\{0\}$. Therefore, Theorem 3.1 in [8] can not guarantee the convergence of the BPHSS iteration sequence. However Theorem 4.2 shows that the BPHSS iteration method is convergent unconditionally to the exact solution of the block tridiagonal system of linear equations (4.1).

Acknowledgement: Part of this work was done during the first author was visiting Department of Mathematics of The College of William and Mary during May–June 2004, and Shanghai Key Laboratory of Contemporary Applied Mathematics of Fudan University.

References

[1] Z.-Z. Bai, A class of parallel decomposition-type relaxation methods for large sparse systems of linear equations, *Linear Algebra Appl.*, 282(1998), 1-24.

[2] Z.-Z. Bai and G.H. Golub, Generalized preconditioned Hermitian and skew-Hermitian splitting methods for saddle-point problems, *Technical Report SCCM-04-07*, Scientific Computing and Computational Mathematics Program, Department of Computer Science, Stanford University, 2004.

[3] 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. Appl.*, 24(2003), 603-626.

[4] 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.

[5] Z.-Z. Bai and Y.-F. Su, On the convergence of a class of parallel decomposition-type relaxation methods, *Appl. Math. Comput.*, 81:1(1997), 1-21.

[6] Z.-Z. Bai and G.-Q. Li, Restrictively preconditioned conjugate gradient methods for systems of linear equations, *IMA J. Numer. Anal.*, 23:4(2003), 561-580.

[7] Z.-Z. Bai and R. Nabben, Some properties of the block matrices in the parallel decomposition-type relaxation methods, *Appl. Numer. Math.*, 29:2(1999), 167-170.

[8] M. Benzi and G.H. Golub, A preconditioner for generalized saddle point problems, *SIAM J. Matrix. Anal. Appl.*, 26(2004), 20-41.

[9] M. Bertaccini, G.H. Golub, S. Serra-Capizzano and C.T. Possio, Preconditioned HSS methods for the solution of non-Hermitian positive definite linear systems, *Technical Report SCCM-02-11*, Scientific Computing and Computational Mathematics Program, Department of Computer Science, Stanford University, 2002.

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

[11] S.C. Eisenstat, H.C. Elman and M.H. Schultz, Variational iterative methods for nonsymmetric systems of linear equations, *SIAM J. Numer. Anal.*, 20(1983), 345-357.

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

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

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

[15] G.H. Golub and C. Greif, On solving block-structured indefinite linear systems, *SIAM J. Sci. Comput.*, 24(2003), 2076-2092.

[16] G.H. Golub, C. Greif and J.M. Varah, Block orderings for tensor-product grids in two and three dimensions, *Numer. Algorithms*, 30(2002), 93-111.

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

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

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

[20] H. Modaressi and P. Aubert, A diffuse element-finite element technique for transient coupled analysis, *Intern. J. Numer. Methods Engrg.*, 39(1996), 3809-3838.

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

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

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

[24] S.L. Weissman, High-accuracy low-order three-dimensional brick elements, *Intern. J. Numer. Methods Engrg.*, 39(1996), 2337-2361.

[25] J.-X. Zhao, W.-G. Wang and W.-Q. Ren, Stability of the matrix factorization for solving block tridiagonal symmetric indefinite linear systems, *BIT*, 44(2004), 181-188.