East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.201015.140316a

Vol. **6**, No. 2, pp. 211-221
May 2016

# Perturbation Bounds and Condition Numbers for a Complex Indefinite Linear Algebraic System

Lei Zhu[1,*], Wei-Wei Xu[2] and Xing-Dong Yang[2]

[1] *College of Engineering, Nanjing Agricultural University, Nanjing 210031, P.R. China.*
[2] *School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, P.R. China.*

*Received 20 October 2015; Accepted (in revised version) 14 March 2016.*

**Abstract.** We consider perturbation bounds and condition numbers for a complex indefinite linear algebraic system, which is of interest in science and engineering. Some existing results are improved, and illustrative numerical examples are provided.

**AMS subject classifications**: 65F10, 65F15, 65H10, 65L15.

**Key words**: Complex indefinite linear algebraic system, perturbation bounds, condition numbers.

## 1. Introduction

Complex indefinite linear algebraic systems arise in many scientific computing and engineering applications, including structural dynamics [1,6,12], electromagnetism [11], wave propagation [5,10], electrical power system modelling and the discretisation of different types of Helmholtz equation [7,8,9]. Many authors have recently considered iterative methods for solving complex indefinite linear algebraic systems, including the Modified HSS (MHSS) and preconditioned Modified HSS (PMHSS) methods [2,3,4].

The complex indefinite linear algebraic system considered is

$$\mathscr{A}x \equiv (W + iT)x = b\,, \tag{1.1}$$

where $i = \sqrt{-1}$, $W$ and $T \in \mathbb{R}^{(m+n)\times(m+n)}$, $x$ and $b \in \mathbb{C}^{(m+n)}$, and we assume the coefficient matrix $\mathscr{A}$ is nonsingular. One case eample of this complex indefinite linear algebraic system is the finite difference discretisation of the Helmholtz partial differential equation

$$-\nabla\cdot(c\nabla\mu) - \sigma_1\mu + i\sigma_2\mu = f\,, \tag{1.2}$$

where the coefficients $c$, $\sigma_1$ and $\sigma_2$ are real-valued functions. Discretising Eq. (1.2) with Dirichlet boundary conditions and both $\sigma_1$ and $\sigma_2$ strictly positive produces Eq. (1.1) with an indefinite real part $W$ and a definite imaginary part $T$ of the discrete operator $A =$

*Corresponding author. *Email address:* zhulei@njau.edu.cn (L. Zhu)

 ©2016 Global-Science Press

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

$W + iT$. Another application concerns the interior penalty discontinuous Galerkin (IPDG) approximation of the prototypical Helmholtz problem

$$-\triangle\mu - k^2\mu = f \ in\, \Omega := \Omega_1 \setminus D\,, \tag{1.3}$$

$$\mu = 0 \ \text{ on } \ \Gamma_D\,, \tag{1.4}$$

$$\frac{\partial\mu}{\partial n_\Omega} + ik\mu = g \ \text{ on } \ \Gamma_R\,. \tag{1.5}$$

Here $i = \sqrt{-1}$, $k \in R$ is a given positive large number (wave number), $\Omega_1 \subset R^d$, $d = 2,3$ is a polygonal/polyhedral domain (often a d-rectangle), $D \subset \Omega_1$ represents a scatterer, $\Gamma_R := \partial D$ hence $\partial\Omega = \Gamma_R \cup \Gamma_D$, and $n_\Omega$ is the unit outward normal to $\partial\Omega$. The Robin boundary condition (1.3) is known as the first order absorbing boundary condition, and implies that the scatterer is sound-soft. The finite element method has also been widely used to discretise Helmholtz problems involving various boundary conditions.

We focus on how the iterative solution of the complex indefinite linear algebraic system (1.1) changes when the coefficient matrix of the complex indefinite linear algebraic system undergoes a small perturbation — and in particular, we discuss perturbation bounds and condition numbers that to the best of our knowledge have not been analysed previously. The system (1.1) can be rewritten in the 2-by-2 block form

$$\mathscr{A}x \equiv \left[\begin{pmatrix} W_1 & W_2 \\ W_3 & W_4 \end{pmatrix} + i\begin{pmatrix} T_1 & T_2 \\ T_3 & T_4 \end{pmatrix}\right]\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{1.6}$$

where $W_1, T_1 \in \mathbb{R}^{m\times m}$; $W_4, T_4 \in \mathbb{R}^{n\times n}$; $W_2, T_2 \in \mathbb{R}^{m\times n}$; $W_3, T_3 \in \mathbb{R}^{n\times m}$ and $x_1, b_1 \in \mathbb{C}^m$; $x_2, b_2 \in \mathbb{C}^n$. The problem can be specified as follows: $W$, $T$, $b$ in Eq. (1.1) or $W_i$, $T_i$ $(i = 1,2,3,4)$ and $b_j$ $(j = 1,2)$ in the corresponding Eq. (1.6) are subjected to small perturbations denoted by $\Delta W$, $\Delta T$, $\Delta b$, $\Delta W_i$, $\Delta T_i$ and $\Delta b_j$, respectively. We consider perturbation bounds denoted by $\|\Delta x\|_2$, where $\Delta x$ is the error in the solution $x$. The relevant perturbed system respectively obtained from Eq. (1.1) or Eq. (1.6) is

$$(\mathscr{A} + \Delta\mathscr{A})(x + \Delta x) = (W + \Delta W + iT + i\Delta T)(x + \Delta x) = b + \Delta b$$

or

$$\begin{aligned}&\left[\begin{pmatrix} W_1 + \Delta W_1 & W_2 + \Delta W_2 \\ W_3 + \Delta W_3 & W_4 + \Delta W_4 \end{pmatrix} + i\begin{pmatrix} T_1 + \Delta T_1 & T_2 + \Delta T_2 \\ T_3 + \Delta T_3 & T_4 + \Delta T_4 \end{pmatrix}\right]\begin{pmatrix} x_1 + \Delta x_1 \\ x_2 + \Delta x_2 \end{pmatrix} \\ &= b + \Delta b \\ &= \begin{pmatrix} b_1 + \Delta b_1 \\ b_2 + \Delta b_2 \end{pmatrix},\end{aligned} \tag{1.7}$$

where $\Delta W_1, \Delta T_1 \in \mathbb{R}^{m\times m}$, $\Delta W_4, \Delta T_4 \in \mathbb{R}^{n\times n}$, $\Delta W_2, \Delta T_2 \in \mathbb{R}^{m\times n}$, $\Delta W_3, \Delta T_3 \in \mathbb{R}^{m\times n}$, $\Delta x_1, \Delta b_1 \in \mathbb{C}^m$ and $\Delta x_2, \Delta b_2 \in \mathbb{C}^n$.

Some definitions, notation and useful lemmas to deduce the main results are given in Section 2. Perturbation bounds for the complex linear algebraic system are derived in Section 3, and the corresponding condition numbers are discussed in Section 4. We present some illustrative numerical examples in Section 5, and briefly summarise our results in Section 6.

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

## 2. Preliminaries

The necessary notation, definitions and useful lemmas are as follows.

The set of $m\times n$ real matrices is denoted by $\mathbb{R}^{m\times n}$, and the set of $m\times n$ complex matrices by $\mathbb{C}^{m\times n}$. For any matrix $A=(a_1,a_2,\cdots,a_n)\in\mathbb{R}^{m\times n}$, $A^T$ denotes the transpose of $A$ and $Vec(A)=\begin{pmatrix} a_1\\ a_2\\ \vdots\\ a_n\end{pmatrix}$, $\|\cdot\|_F$ is the Frobenius norm and $\otimes$ denotes the Kronecker product. For $\varepsilon>0$ sufficiently small, we assume that

$$\|\Delta W_i\|_F\le\varepsilon\delta_{W_i},\qquad \|\Delta T_i\|_F\le\varepsilon\delta_{T_i},\qquad \|\Delta b_j\|_2\le\varepsilon\delta_{b_j},$$
$$i=1,2,3,4,\quad j=1,2\,(\delta_{W_i},\delta_{T_i},\delta_{b_j}>0). \tag{2.1}$$

We also write:

$$\delta=\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix}\equiv\begin{pmatrix}\sqrt{\sum_{i=1}^{4}(\delta_{W_i}^2+\delta_{T_i}^2)}\\ \sqrt{\sum_{i=1}^{2}\delta_{b_i}^2}\end{pmatrix},$$

$$\mathscr{P}=\mathscr{A}^{-1}(x_1^T\otimes I_{m+n},x_2^T\otimes I_{m+n},I_{m+n}),$$

$$\omega=(\omega_1^T,\omega_2^T,\omega_3^T)^T\quad\text{where }\omega_1=Vec(G),\quad\omega_2=Vec(H),\quad\omega_3=L=\begin{pmatrix}\Delta b_1\\ \Delta b_2\end{pmatrix},$$

$$G=\begin{pmatrix}-(\Delta W_1+i\Delta T_1)\\ -(\Delta W_3+i\Delta T_3)\end{pmatrix},\qquad H=\begin{pmatrix}-(\Delta W_2+i\Delta T_2)\\ -(\Delta W_4+i\Delta T_4)\end{pmatrix},$$

$$\mathscr{F}=(F_1,F_2)\equiv(x^T\otimes\mathscr{A}^{-1},\mathscr{A}^{-1})=\left(\begin{pmatrix}x_1\\ x_2\end{pmatrix}^T\otimes\mathscr{A}^{-1},\mathscr{A}^{-1}\right),$$

$$\eta=(\eta_1,\eta_2,\eta_3)^T,\qquad\eta_1=\sqrt{\delta_{W_1}^2+\delta_{T_1}^2+\delta_{W_3}^2+\delta_{T_3}^2},$$
$$\eta_2=\sqrt{\delta_{W_2}^2+\delta_{T_2}^2+\delta_{W_4}^2+\delta_{T_4}^2},\qquad\eta_3=\delta_2\,. \tag{2.2}$$

**Lemma 2.1.** *If coefficient matrix $\mathscr{A}$ in* (1.6) *is nonsingular, then*

$$\begin{pmatrix}\Delta x_1\\ \Delta x_2\end{pmatrix}=\mathscr{P}\omega+\mathscr{A}^{-1}(G,H)\begin{pmatrix}\Delta x_1\\ \Delta x_2\end{pmatrix}, \tag{2.3}$$

*where $\mathscr{P}$, $\omega$, $G$ and $H$ is given by* (2.3).

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

*Proof.* Let

$$\begin{aligned}
&\mathscr{P} = \mathscr{A}^{-1}\left(x_1^T \otimes I_{m+n}, x_2^T \otimes I_{m+n}, I_{m+n}\right),\\
&\omega = (\omega_1^T, \omega_2^T, \omega_3^T)^T, \qquad \omega_1 = Vec(G), \qquad \omega_2 = Vec(H), \qquad \omega_3 = L,\\
&G = \begin{pmatrix} -(\Delta W_1 + i\Delta T_1) \\ -(\Delta W_3 + i\Delta T_3) \end{pmatrix}, \qquad H = \begin{pmatrix} -(\Delta W_2 + i\Delta T_2) \\ -(\Delta W_4 + i\Delta T_4) \end{pmatrix}, \qquad L = \begin{pmatrix} \Delta b_1 \\ \Delta b_2 \end{pmatrix}.
\end{aligned}$$

From (1.7),

$$\begin{aligned}
&\Big[(W_1 + iT_1) + (\Delta W_1 + i\Delta T_1)\Big](x_1 + \Delta x_1) + \Big[(W_2 + iT_2) + (\Delta W_2 + i\Delta T_2)\Big](x_2 + \Delta x_2)\\
=&b_1 + \Delta b_1,\\
&\Big[(W_3 + iT_3) + (\Delta W_3 + i\Delta T_3)\Big](x_1 + \Delta x_1) + \Big[(W_4 + iT_4) + (\Delta W_4 + i\Delta T_4)\Big](x_2 + \Delta x_2)\\
=&b_2 + \Delta b_2,
\end{aligned}$$

which can be rewritten as

$$\begin{aligned}
&\begin{pmatrix} W_1 + iT_1 & W_2 + iT_2 \\ W_3 + iT_3 & W_4 + iT_4 \end{pmatrix}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}\\
=&\begin{pmatrix} \Delta b_1 \\ \Delta b_2 \end{pmatrix} - \begin{pmatrix} \Delta W_1 + i\Delta T_1 & \Delta W_2 + i\Delta T_2 \\ \Delta W_3 + i\Delta T_3 & \Delta W_4 + i\Delta T_4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\\
&- \begin{pmatrix} \Delta W_1 + i\Delta T_1 & \Delta W_2 + i\Delta T_2 \\ \Delta W_3 + i\Delta T_3 & \Delta W_4 + i\Delta T_4 \end{pmatrix}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}.
\end{aligned}$$

If $\mathscr{A}$ is nonsingular, it follows that

$$\begin{aligned}
\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} =& \mathscr{A}^{-1}\begin{pmatrix} \Delta b_1 \\ \Delta b_2 \end{pmatrix} - \mathscr{A}^{-1}\begin{pmatrix} \Delta W_1 + i\Delta T_1 & \Delta W_2 + i\Delta T_2 \\ \Delta W_3 + i\Delta T_3 & \Delta W_4 + i\Delta T_4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\\
&- \mathscr{A}^{-1}\begin{pmatrix} \Delta W_1 + i\Delta T_1 & \Delta W_2 + i\Delta T_2 \\ \Delta W_3 + i\Delta T_3 & \Delta W_4 + i\Delta T_4 \end{pmatrix}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}\\
=& \mathscr{A}^{-1}\begin{pmatrix} \Delta b_1 \\ \Delta b_2 \end{pmatrix} - \mathscr{A}^{-1}\left(\begin{pmatrix} \Delta W_1 + i\Delta T_1 \\ \Delta W_3 + i\Delta T_3 \end{pmatrix}x_1 + \begin{pmatrix} \Delta W_2 + i\Delta T_2 \\ \Delta W_4 + i\Delta T_4 \end{pmatrix}x_2\right)\\
&- \mathscr{A}^{-1}\begin{pmatrix} \Delta W_1 + i\Delta T_1 & \Delta W_2 + i\Delta T_2 \\ \Delta W_3 + i\Delta T_3 & \Delta W_4 + i\Delta T_4 \end{pmatrix}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}\\
=& \mathscr{A}^{-1}(x_1^T \otimes I_{m+n}, x_2^T \otimes I_{m+n}, I_{m+n})\begin{pmatrix} Vec(G) \\ Vec(H) \\ L \end{pmatrix}\\
&- \mathscr{A}^{-1}\begin{pmatrix} \Delta W_1 + i\Delta T_1 & \Delta W_2 + i\Delta T_2 \\ \Delta W_3 + i\Delta T_3 & \Delta W_4 + i\Delta T_4 \end{pmatrix}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}.
\end{aligned}$$

The result follows immediately from the definitions of $\mathscr{P}$, $\omega$, $G$, $H$ and $L$. □

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

## 3. Perturbation bounds

We now present perturbation bounds for the complex indefinite algebraic linear system (1.1).

**Theorem 3.1.** *Let $\varepsilon$, $\delta_1$, $\delta$ and $\mathscr{F} = (F_1, F_2)$ be given in* (2.3). *If $\varepsilon\delta_1 \le \sigma_{\min}(\mathscr{A})$, then*

$$\|\Delta x\|_2 \le \frac{\varepsilon \min\left\{\|\mathscr{F}\|_2\|\delta\|_2, \sqrt{\delta^T \tilde{F}\delta}\right\}}{1-\varepsilon\delta_1\|\mathscr{A}^{-1}\|_2},$$

*where $\sigma_{\min}(\mathscr{A})$ is the smallest singular value of $\mathscr{A}$ and $\tilde{F} = (\tilde{f}_{ij})$ is a $2\times 2$ matrix with entries $\tilde{f}_{ij} = \|F_i^T F_j\|_2$, $i, j = 1, 2$.*

*Proof.* It follows from (1.7) that

$$(\mathscr{A} + \Delta\mathscr{A})(x + \Delta x) = (W + \Delta W + iT + i\Delta T)(x + \Delta x) = b + \Delta b\,.$$

Since $\mathscr{A} = W + iT$ is nonsingular,

$$\begin{aligned}\Delta x &= \mathscr{A}^{-1}\Delta b - \mathscr{A}^{-1}(\Delta W + i\Delta T)x - \mathscr{A}^{-1}(\Delta W + i\Delta T)\Delta x\\ &= \mathscr{A}^{-1}((x^T \otimes I_n),\, I_n)\begin{pmatrix} Vec(-(\Delta W + i\Delta T)) \\ \Delta b \end{pmatrix} - \mathscr{A}^{-1}(\Delta W + i\Delta T)\Delta x\\ &= (x^T \otimes \mathscr{A}^{-1}, \mathscr{A}^{-1})\begin{pmatrix} Vec(-(\Delta W + i\Delta T)) \\ \Delta b \end{pmatrix} - \mathscr{A}^{-1}(\Delta W + i\Delta T)\Delta x\,.\end{aligned}$$

Let $\mathscr{F} = (x^T \otimes \mathscr{A}^{-1}, \mathscr{A}^{-1})$ and $\upsilon = \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \end{pmatrix} \equiv \begin{pmatrix} Vec(-(\Delta W + i\Delta T)) \\ \Delta b \end{pmatrix}$. Then from (3.1) we have

$$\Delta x = \mathscr{F}\upsilon - \mathscr{A}^{-1}(\Delta W + i\Delta T)\Delta x\,,$$

whence

$$\|\Delta x\|_2 \le \|\mathscr{F}\upsilon\|_2 + \|\mathscr{A}^{-1}\|_2\|\Delta W + i\Delta T\|_F\|\Delta x\|_2\,,$$

which together with (2.1)-(2.3) yields $\|\Delta W + i\Delta T\|_F \le \varepsilon\delta_1$ and $\|\upsilon\|_2 \le \varepsilon\delta$. Then

$$\|\mathscr{F}\upsilon\|_2 \le \varepsilon\|\mathscr{F}\|_2\delta\,, \tag{3.1}$$

$$\begin{aligned}\|\mathscr{F}\upsilon\|_2^2 &= \upsilon^T\mathscr{F}^T\mathscr{F}\upsilon = \sum_{i=1}^{2}\sum_{j=1}^{2}\upsilon_i^T F_i^T F_j \upsilon_j\\ &\le \sum_{i=1}^{2}\sum_{j=1}^{2}\|F_i^T F_j\|_2\|\upsilon_i\|_2\|\upsilon_j\|_2\\ &\le \varepsilon^2\sum_{i=1}^{2}\sum_{j=1}^{2}\|F_i^T F_j\|_2\delta_i\delta_j := \varepsilon^2\delta^T\tilde{F}\delta\,,\end{aligned} \tag{3.2}$$

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

where $\tilde{F} = (\tilde{f}_{ij})$ is a $2 \times 2$ matrix with entries $\tilde{f}_{ij} = \|F_i^T F_j\|_2$, $i, j = 1, 2$ and $\delta$, $\delta_1$, $\delta_2$, $F_1$, $F_2$ are given in (2.1) and (2.2). Combining (3.1) with (3.2) and (3.3) gives

$$\|\Delta x\|_2 \le \varepsilon \min\left\{\|\mathscr{F}\|_2\|\delta\|_2, \sqrt{\delta^T \tilde{F} \delta}\right\} + \varepsilon\delta_1 \|\mathscr{A}^{-1}\|_2 \|\Delta x\|_2 . \tag{3.3}$$

On assuming that $\varepsilon\delta_1 \le \sigma_{\min}(\mathscr{A})$, from (3.3) we obtain

$$\|\Delta x\|_2 \le \frac{\varepsilon \ \min\left\{\|\mathscr{F}\|_2\|\delta\|_2, \sqrt{\delta^T \tilde{F} \delta}\right\}}{1 - \varepsilon\delta_1 \|\mathscr{A}^{-1}\|_2} .$$

□

**Theorem 3.2.** *Let $\varepsilon$, $\eta_1$, $\eta_2$, $\eta$ and $\mathscr{P} = (P_1, P_2, P_3)$ be as given in* (2.1) *and* (2.2). *If $\epsilon\sqrt{\eta_1^2 + \eta_2^2} \le \sigma_{\min}(\mathscr{A})$, then*

$$\|\Delta x\|_2 \le \frac{\varepsilon \ \min\left\{\|\mathscr{P}\|_2\|\eta\|_2, \sqrt{\eta^T \tilde{P} \eta}\right\}}{1 - \varepsilon\sqrt{\eta_1^2 + \eta_2^2} \|\mathscr{A}^{-1}\|_2} ,$$

*where $\sigma_{\min}(\mathscr{A})$ is the smallest singular value of $\mathscr{A}$ and $\tilde{P} = (\tilde{p}_{ij})$ is a $2 \times 2$ matrix with entries $\tilde{p}_{ij} = \|P_i^T P_j\|_2$, $i, j = 1, 2$.*

*Proof.* Let $P_1 = \mathscr{A}^{-1}(x^T \otimes I_{m+n}) = x^T \otimes \mathscr{A}^{-1}$, $P_2 = \mathscr{A}^{-1}(y^T \otimes I_{m+n}) = y^T \otimes \mathscr{A}^{-1}$ and $P_3 = \mathscr{A}^{-1}$. Then from (2.2), $\mathscr{P} = (P_1, P_2, P_3)$, and

$$\|\omega_1\|_2 \le \epsilon\eta_1, \qquad \|\omega_2\|_2 \le \epsilon\eta_2 , \qquad \|\omega_3\|_2 \le \epsilon\eta_3 ,$$

so

$$\|\mathscr{P}\eta\|_2 \le \|\mathscr{P}\|_2 \left(\|G\|_F^2 + \|H\|_F^2 + \|L\|_F^2\right)^{\frac{1}{2}} \le \epsilon\|\mathscr{P}\|_2\|\eta\|_2 , \tag{3.4}$$

$$\begin{aligned}
\|\mathscr{P}\omega\|_2^2 &= \omega^T \mathscr{P}^T \mathscr{P} \omega = \sum_{i=1}^{3}\sum_{j=1}^{3} \omega_i^T P_i^T P_j \omega_j \\
&\le \sum_{i=1}^{3}\sum_{j=1}^{3} \|P_i^T P_j\|_2 \|\omega_i\|_2 \|\omega_j\|_2 \\
&\le \epsilon^2 \sum_{i=1}^{3}\sum_{j=1}^{3} \|P_i^T P_j\|_2 \eta_i \eta_j := \epsilon^2 \eta^T \tilde{P} \eta ,
\end{aligned} \tag{3.5}$$

where $\tilde{P} = (p_{ij})$ is a $3 \times 3$ matrix with entries $\tilde{p}_{ij} = \|P_i^T P_j\|_2$. Then from Lemma 2.1, (3.4) and (3.5) we have

$$\begin{aligned}
\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} &\le \|P\omega\|_2 + \|\mathscr{A}^{-1}\|_2 \sqrt{\|G\|_F^2 + \|H\|_F^2} \left\| \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} \right\|_2 \\
&\le \frac{\epsilon \min\left\{\|\mathscr{P}\|_2\|\eta\|_2, \sqrt{\eta^T \tilde{P} \eta}\right\}}{1 - \epsilon\sqrt{\eta_1^2 + \eta_2^2} \|\mathscr{A}^{-1}\|_2} ,
\end{aligned}$$

which proves the desired inequality. □

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

**Remark 3.1.** Theorems 3.1 and 3.2 provide the upper bounds of $\|\Delta x\|_2$ in terms of $\varepsilon$, $\delta_{W_i}$, $\delta_{T_i}$ and $\delta_{b_j}$, $i = 1,2,3,4$, $j = 1,2$.
Thus when $W_i$, $T_i$ and $b_j$ undergo small perturbations $\Delta W_i$, $\Delta T_i$ and $\Delta b_j$, we can assess the bound of $\|\Delta x\|_2$, and also see that the upper bounds of $\|\Delta x\|_2$ are linear about $\varepsilon$. In general, $\varepsilon$ is thus enough small such that the upper bounds of $\|\Delta x\|_2$ are very small, especially in (1) when $T_i = \mathbf{0}$ and $W$ has the structure

$$W = \begin{pmatrix} W_1 & W_2 \\ W_3 & W_4 \end{pmatrix},$$

so system (1.1) reduces to

$$\begin{pmatrix} W_1 & W_2 \\ W_3 & W_4 \end{pmatrix} x = b\,. \tag{3.6}$$

System (3.6) is an important generalised saddle point system [1]. The error bound on the iterative solution due to perturbations, and therefore the upper bound of $\|\Delta x\|_2$ of complex indefinite linear algebraic system (1.1), are of interest. Furthermore, when $T_i = \mathbf{0}$ and $W_2 = W_3^T$, the results of Theorems 3.1 and 3.2 reduce to Theorem 3.1 of Ref. [15]. Indeed, the assumptions are the same — i.e., $\|\mathscr{F}\|_2\|\delta\|_2 = \|\mathscr{P}\|_2\|\eta\|_2$. However, sometimes $\delta^T\tilde{F}\delta$ is smaller and sometimes larger than $\eta^T\tilde{P}\eta$, as illustrated by a numerical example in Section 5.

## 4. Condition numbers

It is well known that the condition number of a matrix plays an important role in numerical analysis [13–15], and we now present relevant joint condition numbers.

**Definition 4.1** (cf. Ref. [13])**.**

$$cond_{x_1,x_2} := \lim_{\epsilon\to 0}\sup_{S_\epsilon}\frac{\|(\Delta x_1^T, \Delta x_2^T)^T\|_2}{\epsilon\|(x_1^T, x_2^T)^T\|_2},$$

$$\tilde{cond}_{x_1,x_2} := \lim_{\epsilon\to 0}\sup_{\hat{S}_\epsilon}\frac{\|(\Delta x_1^T, \Delta x_2^T)^T\|_2}{\epsilon\|(x_1^T, x_2^T)^T\|_2},$$

where

$$S_\epsilon = \left\{\mathscr{R} : \|\mathscr{R}\|_F \le \epsilon \left\| \begin{pmatrix} W_1 + iT_1 & W_2 + iT_2 & b_1 \\ W_3 + iT_3 & W_4 + iT_4 & b_2 \end{pmatrix} \right\|_F, \right.$$
$$\left. \mathscr{R} = \begin{pmatrix} \Delta W_1 + i\Delta T_1 & \Delta W_2 + i\Delta T_2 & \Delta b_1 \\ \Delta W_3 + i\Delta T_3 & \Delta W_4 + i\Delta T_4 & \Delta b_2 \end{pmatrix} \right\},$$

$$\hat{S}_\epsilon = \big\{(\Delta W_i, \Delta T_i, \Delta b_1, \Delta b_2) : (\|\Delta W_i\|_F, \|\Delta T_i\|_F, \|\Delta b_1\|_2, \|\Delta b_2\|_2)$$
$$\le \epsilon(\alpha_i^{(1)}, \alpha_i^{(2)}, \beta_1, \beta_2),\ i = 1,2,3,4\big\}\,.$$

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

The following theorem gives explicit expressions for the upper bounds of the joint condition numbers.

**Theorem 4.1.** *With the above notation and $\tilde{P}$ is given by Theorem* 3.2*,*

$$cond_{x_1,x_2} = \left\|\mathscr{A}^{-1}\mathscr{Y}\right\|_2 \frac{\sqrt{\sum_{i=1}^{4}\|W_i\|_F^2 + \sum_{i=1}^{4}\|T_i\|_F^2 + \|b_1\|_2^2 + \|b_2\|_2^2}}{\sqrt{\|x_1\|_2^2 + \|x_2\|_2^2}}, \tag{4.1}$$

$$\tilde{cond}_{x_1,x_2} \le \frac{\left[\begin{pmatrix} \sqrt{\alpha_1^{(1)^2}+\alpha_1^{(2)^2}+\alpha_3^{(1)^2}+\alpha_3^{(2)^2}} \\ \sqrt{\alpha_2^{(1)^2}+\alpha_2^{(2)^2}+\alpha_4^{(1)^2}+\alpha_4^{(2)^2}} \\ \sqrt{\beta_1^2+\beta_2^2} \end{pmatrix}^T \tilde{P} \begin{pmatrix} \sqrt{\alpha_1^{(1)^2}+\alpha_1^{(2)^2}+\alpha_3^{(1)^2}+\alpha_3^{(2)^2}} \\ \sqrt{\alpha_2^{(1)^2}+\alpha_2^{(2)^2}+\alpha_4^{(1)^2}+\alpha_4^{(2)^2}} \\ \sqrt{\beta_1^2+\beta_2^2} \end{pmatrix}\right]^{\frac{1}{2}}}{\sqrt{\|x_1\|_2^2 + \|x_2\|_2^2}}, \tag{4.2}$$

*where*

$$\mathscr{Y} = \begin{pmatrix} x_1^T \otimes I_m & 0 & x_2^T \otimes I_m & 0 & -I_m & 0 \\ 0 & x_1^T \otimes I_n & 0 & x_2^T \otimes I_n & 0 & -I_n \end{pmatrix}.$$

*Proof.* Neglecting higher orders in (2.3), we have the first-order expansion

$$\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = -\mathscr{A}^{-1}\mathscr{Y} \begin{pmatrix} Vec(\Delta W_1 + i\Delta T_1) \\ Vec(\Delta W_3 + i\Delta T_3) \\ Vec(\Delta W_2 + i\Delta T_2) \\ Vec(\Delta W_4 + i\Delta T_4) \\ \Delta b_1 \\ \Delta b_2 \end{pmatrix}, \tag{4.3}$$

such that

$$\begin{aligned} \left\|\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}\right\|_2 &= \left\|\mathscr{A}^{-1}\mathscr{Y} \begin{pmatrix} Vec(\Delta W_1 + i\Delta T_1) \\ Vec(\Delta W_3 + i\Delta T_3) \\ Vec(\Delta W_2 + i\Delta T_2) \\ Vec(\Delta W_4 + i\Delta T_4) \\ \Delta b_1 \\ \Delta b_2 \end{pmatrix}\right\|_2 \le \|\mathscr{A}^{-1}\mathscr{Y}\|_2 \left\|\begin{pmatrix} Vec(\Delta W_1 + i\Delta T_1) \\ Vec(\Delta W_3 + i\Delta T_3) \\ Vec(\Delta W_2 + i\Delta T_2) \\ Vec(\Delta W_4 + i\Delta T_4) \\ \Delta b_1 \\ \Delta b_2 \end{pmatrix}\right\|_2 \\ &= \|\mathscr{A}^{-1}\mathscr{Y}\|_2 \left\|\begin{pmatrix} Vec(\Delta W_1 + i\Delta T_1) & Vec(\Delta W_2 + i\Delta T_2) & \Delta b_1 \\ Vec(\Delta W_3 + i\Delta T_3) & Vec(\Delta W_4 + i\Delta T_4) & \Delta b_2 \end{pmatrix}\right\|_F \\ &\le \epsilon \|\mathscr{A}^{-1}\mathscr{Y}\|_2 \left\|\begin{pmatrix} W_1 + iT_1 & W_2 + iT_2 & b_1 \\ W_3 + iT_3 & W_4 + iT_4 & b_2 \end{pmatrix}\right\|_F. \end{aligned}$$

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

Now there exists $x_0$ with $\|x_0\|_2 = 1$ such that $\|\mathscr{A}^{-1}\mathscr{Y}\|_2 = \max_{\|x\|_2=1} \|\mathscr{A}^{-1}\mathscr{Y}x\|_2 = \|\mathscr{A}^{-1}\mathscr{Y}x_0\|_2$, so since the perturbations are arbitrary we can choose

$$\begin{pmatrix} Vec(\Delta W_1 + i\Delta T_1) \\ Vec(\Delta W_3 + i\Delta T_3) \\ Vec(\Delta W_2 + i\Delta T_2) \\ Vec(\Delta W_4 + i\Delta T_4) \\ \Delta b_1 \\ \Delta b_2 \end{pmatrix} = \epsilon \left\| \begin{pmatrix} W_1 + iT_1 & W_2 + iT_2 & b_1 \\ W_3 + iT_3 & W_4 + iT_4 & b_2 \end{pmatrix} \right\|_F x_0 ,$$

which implies (4.1) holds. The first-order expansion of $\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}$ in Lemma 2.1 is

$$\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = \mathscr{P}\omega ,$$

so from (3.6) and the definition of $\tilde{cond}_{x_1,x_2}$ we get (4.2). □

**Remark 4.1.** From Theorem 4.1 we can compute the joint condition numbers $cond_{x_1,x_2}$ and $\tilde{cond}_{x_1,x_2}$ for the complex indefinite linear algebraic systems (1.1), as illustrated by numerical example in Section 5. In particular, when $T_i = \mathbf{0}$ and $W_2 = W_3^T$ the results of Theorem 4.1 reduce to Theorem 4.2 of Ref. [13].

## 5. Numerical example

We consider the discretisation of different types of Helmholtz equations [5,13,14] to assess perturbation bounds of $\|\Delta x\|_2$ and the joint condition numbers $cond_{x_1,x_2}, \tilde{cond}_{x_1,x_2}$ for complex indefinite linear algebraic systems (1.1).

Helmholtz equations can arise in modelling wave propagation [5,10]. In particular, let us consider the two-dimensional Helmholtz problem

$$-\Delta\mu - k^2\mu = f := \frac{\sin(kr)}{r} \quad in\ \Omega , \tag{5.1}$$

$$\frac{\partial\mu}{\partial n_\Omega} + ik\mu = g \quad on\ \Gamma_R := \partial\Omega , \tag{5.2}$$

where $\Delta$ denotes the Laplacian, $\Omega$ is the unit regular hexagon with centre (0,0), and $g$ is chosen so that the exact solution is

$$\mu = \frac{\cos(kr)}{k} - \frac{\cos k + i\sin k}{k(J_0(k) + iJ_1(k))} J_0(kr)$$

in polar coordinates, where $J_v(z)$ are Bessel function of the first kind. In discretising (5.1) and (5.2) using finite elements, the complex indefinite linear algebraic systems (1.1) is obtained. To ensure the coefficient matrix is nonsingular, sometimes blocks in the coefficient

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

Table 1: Perturbation bounds of $\|\Delta x\|_2$ with $x_1 = 0$.

| $x_2$ | $d_1$ | $d_2$ |
|---|---|---|
| $(1,0,0,\cdots,0,0)^T$ | $2.5184e-8$ | $2.5184e-8$ |
| $(0,1,0,\cdots,0,0)^T$ | $1.9937e-8$ | $2.0185e-8$ |
| $(0,0,0,\cdots,0,1)^T$ | $1.6351e-8$ | $1.6351e-8$ |
| $(1,1,0,\cdots,0,0)^T$ | $2.7820e-8$ | $2.6691e-8$ |
| $(1,1,0,\cdots,0,0)^T$ | $1.2708e-8$ | $1.2536e-8$ |

Table 2: Perturbation bounds of $\|\Delta x\|_2$ with $x_2 = 0$.

| $x_1$ | $d_1$ | $d_2$ |
|---|---|---|
| $(1,0,0,\cdots,0,0)^T$ | $1.4782e-8$ | $1.3972e-8$ |
| $(0,1,0,\cdots,0,0)^T$ | $2.0257e-8$ | $2.0257e-8$ |
| $(0,0,0,\cdots,0,1)^T$ | $3.7358e-8$ | $3.7358e-8$ |
| $(0,1,0,\cdots,0,1)^T$ | $2.8264e-8$ | $2.8136e-8$ |
| $(1,1,0,\cdots,0,0)^T$ | $3.6007e-8$ | $3.7040e-8$ |
| $(0,0,\cdots,0,1,1)^T$ | $1.4631e-8$ | $1.4893e-8$ |

Table 3: $cond_{x_1,x_2}$ and upper bounds of $\tilde{cond}_{x_1,x_2}$.

| $x_1$ | $x_2$ | $cond_{x_1,x_2}$ | $\tilde{cond}_{x_1,x_2}$ |
|---|---|---|---|
| $(1,0,0,\cdots,0,0)^T$ | $(1,0,0,\cdots,0,0)^T$ | 113.8204 | 7.3992 |
| $(1,1,0,\cdots,0,0)^T$ | $(0,1,0,\cdots,0,0)^T$ | 118.4619 | 8.3580 |
| $(0,0,0,\cdots,0,1)^T$ | $(0,0,0,\cdots,0,1)^T$ | 114.7253 | 8.0068 |
| $(1,0,0,\cdots,0,1)^T$ | $(1,0,0,\cdots,0,1)^T$ | 117.6485 | 7.8101 |
| $(0,1,0,\cdots,0,1)^T$ | $(1,1,0,\cdots,0,0)^T$ | 120.0079 | 8.2946 |

matrix must be modified — e.g. by taking the (1,1) block of the coefficient matrix to be $W_1 + \epsilon Im + i(T_1 + \epsilon Im)$.

We set $\epsilon = 10^{-8}$, $\|\Delta W_i\|_F \le \epsilon$, $\|\Delta T_i\|_F \le \epsilon$, $\|\Delta b_j\|_2 \le \epsilon$, $i = 1,2,3,4$; $j = 1,2$. All the runs were done in MATLAB 7.0 on a CPU 1.86GHZ and 1022MB memory computer. We list perturbation bounds of $\|\Delta x\|_2$ in Tables 1 and 2, denoting the bounds from Theorem 3.1 by $d_1$ and those from Theorem 3.2 by $d_2$.

Tables 1 and 2 show that the perturbation bounds of $\|\Delta x\|_2$ are linear in $\varepsilon$ and about $\mathcal{O}(10^{-8})$. Sometimes the bounds derived from Theorem 3.1 are smaller than those derived from Theorem 3.2, and sometimes larger.

Table 3 shows the computed values for $cond_{x_1,x_2}$ and its upper bounds for various $x_1$ and $x_2$.

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press

## 6. Conclusion

The perturbation bounds and joint condition numbers obtained in this article for a complex indefinite linear algebraic system arising in various scientific and engineering applications improve some existing results.

## Acknowledgments

This work was supported by the Jiangsu Provincial Natural Science Foundation of Jiangsu Province of China under Grant No. BK20130985, SIT Foundation under Grant No. 12140715 01279, the Natural Science Foundation of China under Grant No. 11501300, the National Science Foundation of China under Grant No. 51008160 and the Fund for Social Science Research of Fundamental Research Funds at the Central Universities of Nanjing Agricultural University under Grant No. SKZK2015005.

## References

[1] Z.Z. Bai, *Optimal parameters in the HSS-like methods for saddle-point problems*, Num. Linear Algebra Appl. **16**, 447-479 (2009).
[2] Z.Z. Bai, M. Benzi and F. Chen, *On preconditioned MHSS iteration methods for complex symmetric linear systems*, Num. Algor. **56**, 297-317 (2011).
[3] R. Hiptmair, *Finite elements in computational electromagnetism*, Acta Numer. **11**, 237-339 (2003).
[4] Z.Z. Bai, M. Benzi and F. Chen, *Modified HSS iteration methods for a class of complex symmetric linear systems*, Computing **87**, 93-111 (2010).
[5] Z.Z. Bai, G.H. Golub and J.Y. Pan, *Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite linear systems*, Num. Math. **98**, 1-32 (2004).
[6] B. Poirier, *Efficient preconditioning scheme for block partitioned matrices with structured sparsity*, Num. Linear Algebra Appl. **7**, 715-726 (2000).
[7] M. Benzi and D. Bertaccini, *Block preconditioning of real-valued iterative algorithms for complex linear systems*, IMA J. Numer. Anal. **28**, 598-618 (2008).
[8] A. Feriani, F. Perotti and V. Simoncini, *Iterative system solvers for the frequency analysis of linear mechanical systems*, Comp. Methods Appl. Mech. Eng. **190**, 1719-1739 (2000).
[9] R.W. Freund, *Conjugate gradient-type methods for linear systems with complex symmetric coefficient matrices*, SIAM J. Sci. Stat. Comp. **13**, 425-448 (1992).
[10] A. Sommerfeld, *Partial Differential Equations*, Academic Press (1949).
[11] O. Axelsson and A. Kucherov, *Real valued iterative methods for solving complex symmetric linear systems*, Num. Linear Algebra Appl. **7**, 197-218 (2000).
[12] X.X. Guo and S. Wang, *Modified HSS iteration methods for a class of non-Hermitian positive-definite linear systems*, Appl. Math. Comput. **218**, 10122-10128 (2012).
[13] W.W. Xu and W. Li, *New perturbation analysis for generalized saddle point systems*, Calcolo **46**, 25-36 (2009).
[14] X.S. Chen, W. Li and M.K. Ng, *On condition numbers of the spectral projections associated with periodic eigenproblems*, J. Comp. Appl. Math. **272**, 417-429 (2014).
[15] Z.J. Xie, W. Li and X.Q. Jin, *On condition numbers for the canonical generalized Polar decomposition of real matrices*, Electr. J. Linear Algebra. **26**, 842-857 (2013).

https://doi.org/10.4208/eajam.201015.140316a Published online by Cambridge University Press