

Gradient-Based Iterative Algorithm for a Coupled Complex Conjugate and Transpose Matrix Equations

Hongcai Yin¹, Huamin Zhang²

¹School of Management Science and Engineering, Anhui University of Finance & Economics, Bengbu, China

²College of Information & Network Engineering, Anhui Science & Technology University, Bengbu, China

Email: hcyhy@163.com, zhangeasymail@126.com

How to cite this paper: Yin, H.C. and Zhang, H.M. (2021) Gradient-Based Iterative Algorithm for a Coupled Complex Conjugate and Transpose Matrix Equations. *Advances in Linear Algebra & Matrix Theory*, 11, 92-107.

<https://doi.org/10.4236/alamt.2021.113007>

Received: July 10, 2021

Accepted: August 27, 2021

Published: August 30, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Gradient-based iterative algorithm is suggested for solving a coupled complex conjugate and transpose matrix equations. Using the hierarchical identification principle and the real representation of a complex matrix, a convergence proof is offered. The necessary and sufficient conditions for the optimal convergence factor are determined. A numerical example is offered to validate the efficacy of the suggested algorithm.

Keywords

Hierarchical Identification Principle, Complex Conjugate and Transpose Matrix Equation, Real Representation of a Complex

1. Introduction

How to solve a matrix equation is a focus question of the applied mathematics and engineering fields [1] [2]. Based on the classical iterative algorithm and the hierarchical identification principle, some gradient-based and least squares based iterative algorithms were established. This method has been developed in solving other coupled matrix equations [3]. For example, to speed up the convergence behavior, a shift-splitting hierarchical identification method was suggested [4]. Different from the above work, Zhou and his coworkers determined the optimal convergence factor using the object function and gradient search method [5] [6].

Although achieving some fruitful results, there are still some problems that need to be solved. For example, the gradient-based iterative algorithm for solving a class of complex matrix equations was established and the sufficient condi-

tion for the convergence of the algorithm was obtained [3]. So far this conjecture is still open. Although the convergence theory has been established for the real linear matrix equations [5] [7] [8] [9], the corresponding convergence theory for the complex matrix equations is still open especially for the complex matrix equations with conjugate transpose unknown matrices.

Inspired by the above work, this paper discusses gradient-based iterative algorithm for a coupled complex conjugate and transpose matrix equation. Using the real representation of a complex matrix, the convergence proof is given and the optimal convergence factor is determined. An numerical test is offered to validate the efficacy of the suggested algorithm.

We organized this paper as follows. Some preliminary lemmas and symbols are listed in Section 2. Section 3 proves convergence of the gradient-based iterative algorithm for solving a coupled complex conjugate and transpose matrix equation. Section 4 brings a numerical test to validate the efficacy of the proposed algorithm. Some concluding remarks are given in Section 5.

2. Preliminary Symbols and Lemmas

Let A be an $m \times n$ complex matrix, symbols \bar{A} , A^H , A^T , $\lambda_{\max}[A]$, $\lambda_{\min}[A]$, $\|A\|_2$, $\|A\|$ represents the conjugate, the conjugate transpose, the transpose, the maximal eigenvalue, the minimal nonzero eigenvalue, the spectral norm, the Frobenius norm, respectively. I_n denotes an identity matrix of size $n \times n$. Let A be $m \times n$ real matrix and $\text{rank}[A] = r$, symbol $\sigma_i(A)$, $i = 1, 2, \dots, r$ represents the nonzero singular values that satisfy $\sigma_1(A) \geq \sigma_2(A) \geq \dots \geq \sigma_r(A)$. $\text{col}[X]$ denotes an mn -dimensional vector defined by the formula $\text{col}[X] = [x_1^T, x_2^T, \dots, x_n^T]^T$. Here, $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$. The Kronecker product $A \otimes B$ of matrices A and B is defined by formula $A \otimes B = [a_{ij}B] \in \mathbb{R}^{(mp) \times (nq)}$. Here $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{p \times q}$.

Next lemma is a conclusion of the singular values decomposition of a complex matrix, the proof can be found in a standard matrix theory book, we omit the proof here.

Lemma 1. A is an $m \times n$ complex matrix and $\text{rank}[A] = r$. There exist two unitary matrices $U \in \mathbb{C}^{n \times n}$ and $V \in \mathbb{C}^{m \times m}$ such that

$$V^H A U = \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (1)$$

where $\Sigma = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_r\}$ with order $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$.

Refer to [10], we have the real representation of a complex matrix.

For a complex matrix $A \in \mathbb{C}^{m \times n}$, it can be exclusive signified as $A = A_1 + iA_2$, where $A_1, A_2 \in \mathbb{R}^{m \times n}$ are two real matrices. Now we use symbol A_σ to represent it as

$$A_\sigma := \begin{bmatrix} A_1 & A_2 \\ A_2 & -A_1 \end{bmatrix} \in \mathbb{R}^{2m \times 2n}. \quad (2)$$

with this symbol, in the following we use the following notation $\bar{A}_\sigma := (\bar{A})_\sigma$, $A_\sigma^H := (A^H)_\sigma$ and $A_\sigma^{-1} := (A_\sigma)^{-1}$. Because

$$(A^\top)_\sigma = (A_1^\top + iA_2^\top)_\sigma = \begin{bmatrix} A_1^\top & A_2^\top \\ A_2^\top & -A_1^\top \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ A_2 & -A_1 \end{bmatrix}^\top = (A_\sigma)^\top,$$

we use A_σ^\top to represent $(A^\top)_\sigma$ and have symbols:

$$\begin{cases} A_\sigma^\top := (A^\top)_\sigma, \\ \bar{A}_\sigma := (\bar{A})_\sigma, \\ A_\sigma^H := (A^H)_\sigma, \\ A_\sigma^{-1} := (A_\sigma)^{-1}. \end{cases} \tag{3}$$

Using symbols [10]

$$P_m := \begin{bmatrix} I_m & \mathbf{0} \\ \mathbf{0} & -I_m \end{bmatrix}, \quad Q_m := \begin{bmatrix} \mathbf{0} & I_m \\ -I_m & \mathbf{0} \end{bmatrix},$$

We have the following lemma [10].

Lemma 2. The real representation possesses the following properties:

1) If $A, B \in \mathbb{C}^{m \times n}$, $a \in \mathbb{R}$, then

$$\begin{cases} (A+B)_\sigma = A_\sigma + B_\sigma, \\ (aA)_\sigma = aA_\sigma, \\ (A_\sigma^\top)^\top = A_\sigma. \end{cases}$$

2) If $A \in \mathbb{C}^{m \times n}$, then

$$\begin{cases} A_\sigma = Q_m A_\sigma Q_n, \\ \bar{A}_\sigma = P_m A_\sigma P_n, \\ A_\sigma^H = P_m A_\sigma^\top P_n. \end{cases}$$

3) If $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times r}$, $C \in \mathbb{C}^{r \times p}$, then

$$\begin{cases} (AB)_\sigma = A_\sigma P_n B_\sigma = A_\sigma \bar{B}_\sigma P_r, \\ (ABC)_\sigma = A_\sigma \bar{B}_\sigma C_\sigma. \end{cases}$$

The proof one can refer to [10] [11], we omit the proof here.

Next lemma describes the properties of the spectral norm and the Frobenius norm.

Lemma 3. Let A be a complex matrix, we have

- 1) $\|A_\sigma\|^2 = 2\|A\|^2$,
- 2) $\|A_\sigma\|_2 = \|A\|_2$.

For the complex matrix equation $AXB = F$, referring to [12], we have the following gradient-based iterative algorithm.

Lemma 4. For the complex matrix equation $AXB = F$, if A is a column-full rank matrix and B is a row-full rank matrix, then the iterative solution $X(k)$ given by the gradient-based iterative algorithm

$$X(k) = X(k-1) + \mu A^H [F - AX(k-1)B] B^H$$

converges to the exact solution X_* (i.e., $\lim_{k \rightarrow \infty} X(k) = X_*$) for any initial values $X(0)$ if and only if

$$0 < \mu < \frac{2}{\lambda_{\max} [A^H A] \lambda_{\max} [B B^H]}.$$

Moreover, the best convergence factor μ_0 is

$$\mu_0 = \frac{2}{\lambda_{\max} [A^H A] \lambda_{\max} [B B^H] + \lambda_{\min} [A^H A] \lambda_{\min} [B B^H]}.$$

3. Complex Coupled Conjugate and Transpose Matrix Equations

In this section, we suggest the gradient-based iterative algorithm for solving the complex coupled conjugate and transpose matrix equations. For convenience, we introduce the following operator

$$\mathcal{L}_{ij}(X_j) := A_{ij} X_j B_{ij} + C_{ij} \bar{X}_j D_{ij} + G_{ij} X_j^T H_{ij} + M_{ij} X_j^H N_{ij}, \quad i, j = 1, 2. \quad (4)$$

Using these operators, we consider the following complex coupled matrix equations

$$\begin{cases} \mathcal{L}_{11}(X_1) + \mathcal{L}_{12}(X_2) = F_1, \\ \mathcal{L}_{21}(X_1) + \mathcal{L}_{22}(X_2) = F_2, \end{cases} \quad (5)$$

where $A_{ij}, B_{ij}, C_{ij}, D_{ij}, G_{ij}, H_{ij}, M_{ij}, N_{ij}, F_i \in \mathbb{C}^{n \times n}$, $i, j = 1, 2$, are known matrices and $X_1, X_2 \in \mathbb{C}^{n \times n}$ are the matrices to be determined.

3.1. The Exact Solution

According to the definition of the real representation of a complex matrix and Lemma 2, the real representation of matrix $(\mathcal{L}_{ij}(X_j))_\sigma$ denoted as $\mathcal{L}_{ij\sigma}(X_j)$ can be rewritten as

$$\begin{aligned} & \mathcal{L}_{ij\sigma}(X_j) \\ & := (\mathcal{L}_{ij}(X_j))_\sigma = (A_{ij} X_j B_{ij} + C_{ij} \bar{X}_j D_{ij} + G_{ij} X_j^T H_{ij} + M_{ij} X_j^H N_{ij})_\sigma \\ & = (A_{ij} X_j B_{ij})_\sigma + (C_{ij} \bar{X}_j D_{ij})_\sigma + (G_{ij} X_j^T H_{ij})_\sigma + (M_{ij} X_j^H N_{ij})_\sigma \\ & = A_{ij\sigma} \bar{X}_{j\sigma} B_{ij\sigma} + C_{ij\sigma} X_{j\sigma} D_{ij\sigma} + G_{ij\sigma} X_{j\sigma}^H H_{ij\sigma} + M_{ij\sigma} X_{j\sigma}^T N_{ij\sigma} \\ & = A_{ij\sigma} P_n X_{j\sigma} P_n B_{ij\sigma} + C_{ij\sigma} X_{j\sigma} D_{ij\sigma} + G_{ij\sigma} P_n X_{j\sigma}^T P_n H_{ij\sigma} + M_{ij\sigma} X_{j\sigma}^T N_{ij\sigma}. \end{aligned} \quad (6)$$

Taking the vec-operator of the above equation gives

$$\begin{aligned} & \text{col}[\mathcal{L}_{ij\sigma}(X_j)] \\ & = \text{col}[A_{ij\sigma} P_n X_{j\sigma} P_n B_{ij\sigma}] + \text{col}[C_{ij\sigma} X_{j\sigma} D_{ij\sigma}] \\ & \quad + \text{col}[G_{ij\sigma} P_n X_{j\sigma}^T P_n H_{ij\sigma}] + \text{col}[M_{ij\sigma} X_{j\sigma}^T N_{ij\sigma}] \\ & = \left((P_n B_{ij\sigma})^T \otimes A_{ij\sigma} P_n \right) \text{col}[X_{j\sigma}] + (D_{ij\sigma}^T \otimes C_{ij\sigma}) \text{col}[X_{j\sigma}] \\ & \quad + \left((P_n H_{ij\sigma})^T \otimes G_{ij\sigma} P_n \right) \text{col}[X_{j\sigma}^T] + (N_{ij\sigma}^T \otimes M_{ij\sigma}) \text{col}[X_{j\sigma}^T] \end{aligned}$$

$$\begin{aligned}
 &= \left((\mathbf{P}_n \mathbf{B}_{ij\sigma})^T \otimes \mathbf{A}_{ij\sigma} \mathbf{P}_n \right) \text{col}[\mathbf{X}_{j\sigma}] + (\mathbf{D}_{ij\sigma}^T \otimes \mathbf{C}_{ij\sigma}) \text{col}[\mathbf{X}_{j\sigma}] \\
 &\quad + \left((\mathbf{P}_n \mathbf{H}_{ij\sigma})^T \otimes \mathbf{G}_{ij\sigma} \mathbf{P}_n \right) \mathbf{P}(2n, 2n) \text{col}[\mathbf{X}_{j\sigma}] \\
 &\quad + (\mathbf{N}_{ij\sigma}^T \otimes \mathbf{M}_{ij\sigma}) \mathbf{P}(2n, 2n) \text{col}[\mathbf{X}_{j\sigma}] \\
 &= \mathbf{\Phi}_{ij\sigma} \text{col}[\mathbf{X}_{j\sigma}].
 \end{aligned} \tag{7}$$

Here, symbol $\mathbf{\Phi}_{ij\sigma}$ is defined as

$$\begin{aligned}
 \mathbf{\Phi}_{ij\sigma} := & \left((\mathbf{P}_n \mathbf{B}_{ij\sigma})^T \otimes \mathbf{A}_{ij\sigma} \mathbf{P}_n \right) + \left((\mathbf{P}_n \mathbf{H}_{ij\sigma})^T \otimes \mathbf{G}_{ij\sigma} \mathbf{P}_n \right) \mathbf{P}(2n, 2n) \\
 & + (\mathbf{D}_{ij\sigma}^T \otimes \mathbf{C}_{ij\sigma}) + (\mathbf{N}_{ij\sigma}^T \otimes \mathbf{M}_{ij\sigma}) \mathbf{P}(2n, 2n).
 \end{aligned} \tag{8}$$

Taking the real representation of Equation (5) gives

$$\begin{cases} \mathcal{L}_{11\sigma}(\mathbf{X}_1) + \mathcal{L}_{12\sigma}(\mathbf{X}_2) = \mathbf{F}_{1\sigma}, \\ \mathcal{L}_{21\sigma}(\mathbf{X}_1) + \mathcal{L}_{22\sigma}(\mathbf{X}_2) = \mathbf{F}_{2\sigma}, \end{cases} \tag{9}$$

Taking the vec-operator of the above system, we have

$$\begin{cases} \text{col}[\mathcal{L}_{11\sigma}(\mathbf{X}_1)] + \text{col}[\mathcal{L}_{12\sigma}(\mathbf{X}_2)] = \text{col}[\mathbf{F}_{1\sigma}], \\ \text{col}[\mathcal{L}_{21\sigma}(\mathbf{X}_1)] + \text{col}[\mathcal{L}_{22\sigma}(\mathbf{X}_2)] = \text{col}[\mathbf{F}_{2\sigma}]. \end{cases} \tag{10}$$

Using notation $\mathbf{\Phi}_{ij\sigma}$ gives

$$\begin{cases} \mathbf{\Phi}_{11\sigma} \text{col}[\mathbf{X}_{1\sigma}] + \mathbf{\Phi}_{12\sigma} \text{col}[\mathbf{X}_{2\sigma}] = \text{col}[\mathbf{X}_{1\sigma}], \\ \mathbf{\Phi}_{21\sigma} \text{col}[\mathbf{X}_{1\sigma}] + \mathbf{\Phi}_{22\sigma} \text{col}[\mathbf{X}_{2\sigma}] = \text{col}[\mathbf{X}_{2\sigma}]. \end{cases} \tag{11}$$

Setting

$$\mathbf{\Phi}_\sigma := \begin{bmatrix} \mathbf{\Phi}_{11\sigma} & \mathbf{\Phi}_{12\sigma} \\ \mathbf{\Phi}_{21\sigma} & \mathbf{\Phi}_{22\sigma} \end{bmatrix},$$

Equation (11) can be rewritten in a compact one as

$$\mathbf{\Phi}_\sigma \text{col}[\mathbf{X}_{1\sigma}, \mathbf{X}_{2\sigma}] = \text{col}[\mathbf{F}_{1\sigma}, \mathbf{F}_{2\sigma}]. \tag{12}$$

For Equation (12), we have the following result.

Lemma 5. If matrix $\mathbf{\Phi}_\sigma$ is invertible, then Equation (12) has a unique solution and it can be given by formula

$$\text{col}[\mathbf{F}_{1\sigma}, \mathbf{F}_{2\sigma}] = \mathbf{\Phi}_\sigma^{-1} \text{col}[\mathbf{X}_{1\sigma}, \mathbf{X}_{2\sigma}].$$

According to Lemma 4, the gradient-based iterative algorithm

$$\begin{aligned}
 &\text{col}[\mathbf{X}_{1\sigma}(k+1), \mathbf{X}_{2\sigma}(k+1)] \\
 &= (\mathbf{I}_{8n^2} - \mu \mathbf{\Phi}_\sigma^T \mathbf{\Phi}_\sigma) \text{col}[\mathbf{X}_{1\sigma}(k), \mathbf{X}_{2\sigma}(k)] + \mu \mathbf{\Phi}_\sigma^T \text{col}[\mathbf{X}_{1\sigma}, \mathbf{X}_{2\sigma}]
 \end{aligned} \tag{13}$$

converges to the exact solution pair $\text{col}[\mathbf{X}_{1*}]$ and $\text{col}[\mathbf{X}_{2*}]$ for any initial values $\mathbf{X}_1(0)$ and $\mathbf{X}_2(0)$ if and only if

$$0 < \mu < \frac{2}{\lambda_{\max}[\mathbf{\Phi}_\sigma^T \mathbf{\Phi}_\sigma]}.$$

Moreover, the best convergence factor μ_0 is

$$\mu_0 = \frac{2}{\lambda_{\max} [\Phi_\sigma^T \Phi_\sigma] + \lambda_{\min} [\Phi_\sigma^T \Phi_\sigma]}.$$

Clearly, if $X_1, X_2 \in \mathbb{C}^{n \times n}$ then we have $\mathcal{L}_{ij\sigma} \in \mathbb{R}^{2n \times 2n}$, $\Phi_{ij\sigma} \in \mathbb{R}^{4n^2 \times 4n^2}$ and $\Phi \in \mathbb{R}^{8n^2 \times 8n^2}$. The increasing dimension leads to a computational difficulty due to an excessive computer memory. To overcome this problem, next we construct a gradient-based iterative algorithm by using the hierarchical identification principle.

3.2. The Iterative Solution

For convenience, we introduce notation Ψ_1 and Ψ_2 as follows:

$$\Psi_1 := F_1 - \mathcal{L}_{11}(X_1(k)) - \mathcal{L}_{12}(X_2(k)), \tag{14}$$

$$\Psi_2 := F_2 - \mathcal{L}_{21}(X_1(k)) - \mathcal{L}_{22}(X_2(k)). \tag{15}$$

According to the hierarchical identification principle, we introduce the intermediate matrices:

$$F_{11} := \Psi_1 + A_{11}X_1B_{11}, \quad F_{12} := \Psi_1 + C_{11}\bar{X}_1D_{11}, \tag{16}$$

$$F_{13} := \Psi_1 + G_{11}X_1^T H_{11}, \quad F_{14} := \Psi_1 + M_{11}X_1^H N_{11}, \tag{17}$$

$$F_{15} := \Psi_1 + A_{12}X_2B_{12}, \quad F_{16} := \Psi_1 + C_{12}\bar{X}_2D_{12}, \tag{18}$$

$$F_{17} := \Psi_1 + G_{12}X_2^T H_{12}, \quad F_{18} := \Psi_1 + M_{12}X_2^H N_{12}, \tag{19}$$

$$F_{21} := \Psi_2 + A_{21}X_1B_{21}, \quad F_{22} := \Psi_2 + C_{21}\bar{X}_1D_{21}, \tag{20}$$

$$F_{23} := \Psi_2 + G_{21}X_1^T H_{21}, \quad F_{24} := \Psi_2 + M_{21}X_1^H N_{21}, \tag{21}$$

$$F_{25} := \Psi_2 + A_{22}X_2B_{22}, \quad F_{26} := \Psi_2 + C_{22}\bar{X}_2D_{22}, \tag{22}$$

$$F_{27} := \Psi_2 + G_{22}X_2^T H_{22}, \quad F_{28} := \Psi_2 + M_{22}X_2^H N_{22}. \tag{23}$$

Equation (5) can be decomposed into the following matrix equations

$$A_{11}X_1B_{11} = F_{11}, \quad C_{11}\bar{X}_1D_{11} = F_{12}, \tag{24}$$

$$G_{11}X_1^T H_{11} = F_{13}, \quad M_{11}X_1^H N_{11} = F_{14}, \tag{25}$$

$$A_{12}X_2B_{12} = F_{15}, \quad C_{12}\bar{X}_2D_{12} = F_{16}, \tag{26}$$

$$G_{12}X_2^T H_{12} = F_{17}, \quad M_{12}X_2^H N_{12} = F_{18}, \tag{27}$$

$$A_{21}X_1B_{21} = F_{21}, \quad C_{21}\bar{X}_1D_{21} = F_{22}, \tag{28}$$

$$G_{21}X_1^T H_{21} = F_{23}, \quad M_{21}X_1^H N_{21} = F_{24}, \tag{29}$$

$$A_{22}X_2B_{22} = F_{25}, \quad C_{22}\bar{X}_2D_{22} = F_{26}, \tag{30}$$

$$G_{22}X_2^T H_{22} = F_{27}, \quad M_{22}X_2^H N_{22} = F_{28}. \tag{31}$$

According to Lemma 4, with proper manipulations, one can construct the following iterative algorithms for matrix equations in Equations (24)-(30) respectively,

$$X_{11}(k+1) = X_{11}(k) + \mu A_{11}^H (F_{11} - A_{11}X_{11}(k)B_{11})B_{11}^H, \tag{32}$$

$$X_{12}(k+1) = X_{12}(k) + \mu C_{11}^T (\bar{F}_{12} - \bar{C}_{11} X_{12}(k) \bar{D}_{11}) D_{11}^T, \tag{33}$$

$$X_{13}(k+1) = X_{13}(k) + \mu \bar{H}_{11} (F_{13}^T - H_{11}^T X_{13}(k) G_{11}^T) \bar{G}_{11}, \tag{34}$$

$$X_{14}(k+1) = X_{14}(k) + \mu N_{11} (F_{14}^H - N_{11}^H X_{14}(k) M_{11}^H) M_{11}, \tag{35}$$

$$X_{21}(k+1) = X_{21}(k) + \mu A_{12}^H (F_{15} - A_{12} X_{21}(k) B_{12}) B_{12}^H, \tag{36}$$

$$X_{22}(k+1) = X_{22}(k) + \mu C_{12}^T (\bar{F}_{16} - \bar{C}_{12} X_{22}(k) \bar{D}_{12}) D_{12}^T, \tag{37}$$

$$X_{23}(k+1) = X_{23}(k) + \mu \bar{H}_{12} (F_{17}^T - H_{12}^T X_{23}(k) G_{12}^T) \bar{G}_{12}, \tag{38}$$

$$X_{24}(k+1) = X_{24}(k) + \mu N_{12} (F_{18}^H - N_{12}^H X_{24}(k) M_{12}^H) M_{12}, \tag{39}$$

$$X_{15}(k+1) = X_{15}(k) + \mu A_{21}^H (F_{21} - A_{21} X_{15}(k) B_{21}) B_{21}^H, \tag{40}$$

$$X_{16}(k+1) = X_{16}(k) + \mu C_{21}^T (\bar{F}_{22} - \bar{C}_{21} X_{16}(k) \bar{D}_{21}) D_{21}^T, \tag{41}$$

$$X_{17}(k+1) = X_{17}(k) + \mu \bar{H}_{21} (F_{23}^T - H_{21}^T X_{17}(k) G_{21}^T) \bar{G}_{21}, \tag{42}$$

$$X_{18}(k+1) = X_{18}(k) + \mu N_{21} (F_{24}^H - N_{21}^H X_{18}(k) M_{21}^H) M_{21}, \tag{43}$$

$$X_{25}(k+1) = X_{25}(k) + \mu A_{22}^H (F_{25} - A_{22} X_{25}(k) B_{22}) B_{22}^H, \tag{44}$$

$$X_{26}(k+1) = X_{26}(k) + \mu C_{22}^T (\bar{F}_{26} - \bar{C}_{22} X_{26}(k) \bar{D}_{22}) D_{22}^T, \tag{45}$$

$$X_{27}(k+1) = X_{27}(k) + \mu \bar{H}_{22} (F_{27}^T - H_{22}^T X_{27}(k) G_{22}^T) \bar{G}_{22}, \tag{46}$$

$$X_{28}(k+1) = X_{28}(k) + \mu N_{22} (F_{28}^H - N_{22}^H X_{28}(k) M_{22}^H) M_{22}. \tag{47}$$

Substituting Equations (16)-(23) into Equations (32)-(47) respectively gives

$$X_{11}(k+1) = X_{11}(k) + \mu A_{11}^H (\Psi_1 + A_{11} X_1 B_{11} - A_{11} X_{11}(k) B_{11}) B_{11}^H,$$

$$X_{12}(k+1) = X_{12}(k) + \mu C_{11}^T (\bar{\Psi}_1 + \bar{C}_{11} X_1 \bar{D}_{11} - \bar{C}_{11} X_{12}(k) \bar{D}_{11}) D_{11}^T,$$

$$X_{13}(k+1) = X_{13}(k) + \mu \bar{H}_{11} (\Psi_1^T + H_{11}^T X_1 G_{11}^T - H_{11}^T X_{13}(k) G_{11}^T) \bar{G}_{11},$$

$$X_{14}(k+1) = X_{14}(k) + \mu N_{11} (\Psi_1^H + N_{11}^H X_1 M_{11}^H - N_{11}^H X_{14}(k) M_{11}^H) M_{11},$$

$$X_{21}(k+1) = X_{21}(k) + \mu A_{12}^H (\Psi_1 + A_{12} X_2 B_{12} - A_{12} X_{21}(k) B_{12}) B_{12}^H,$$

$$X_{22}(k+1) = X_{22}(k) + \mu C_{12}^T (\bar{\Psi}_1 + \bar{C}_{12} X_2 \bar{D}_{12} - \bar{C}_{12} X_{22}(k) \bar{D}_{12}) D_{12}^T,$$

$$X_{23}(k+1) = X_{23}(k) + \mu \bar{H}_{12} (\Psi_1^T + H_{12}^T X_2 G_{12}^T - H_{12}^T X_{23}(k) G_{12}^T) \bar{G}_{12},$$

$$X_{24}(k+1) = X_{24}(k) + \mu N_{12} (\Psi_1^H + N_{12}^H X_2 M_{12}^H - N_{12}^H X_{24}(k) M_{12}^H) M_{12},$$

$$X_{15}(k+1) = X_{15}(k) + \mu A_{21}^H (\Psi_2 + A_{21} X_1 B_{21} - A_{21} X_{15}(k) B_{21}) B_{21}^H,$$

$$X_{16}(k+1) = X_{16}(k) + \mu C_{21}^T (\bar{\Psi}_2 + \bar{C}_{21} X_1 \bar{D}_{21} - \bar{C}_{21} X_{16}(k) \bar{D}_{21}) D_{21}^T,$$

$$X_{17}(k+1) = X_{17}(k) + \mu \bar{H}_{21} (\Psi_2^T + H_{21}^T X_1 G_{21}^T - H_{21}^T X_{17}(k) G_{21}^T) \bar{G}_{21},$$

$$X_{18}(k+1) = X_{18}(k) + \mu N_{21} (\Psi_2^H + N_{21}^H X_1 M_{21}^H - N_{21}^H X_{18}(k) M_{21}^H) M_{21},$$

$$X_{25}(k+1) = X_{25}(k) + \mu A_{22}^H (\Psi_2 + A_{22} X_2 B_{22} - A_{22} X_{25}(k) B_{22}) B_{22}^H,$$

$$\begin{aligned}
 X_{26}(k+1) &= X_{26}(k) + \mu C_{22}^T (\bar{\Psi}_2 + \bar{C}_{22} X_2 \bar{D}_{22} - \bar{C}_{22} X_{26}(k) \bar{D}_{22}) D_{22}^T, \\
 X_{27}(k+1) &= X_{27}(k) + \mu \bar{H}_{22} (\Psi_2^T + H_{22}^T X_2 G_{22}^T - H_{22}^T X_{27}(k) G_{22}^T) \bar{G}_{21}, \\
 X_{28}(k+1) &= X_{28}(k) + \mu N_{22} (\Psi_2^H + N_{22}^H X_2 M_{22}^H - N_{22}^H X_{28}(k) M_{22}^H) M_{22}.
 \end{aligned}$$

The unknown matrices X_1 and X_2 appear in the right-hand sides of the above equations, so it is impossible to realize the above algorithms. According to the hierarchical identification principle, the unknown matrices X_1 and X_2 can be replaced by their estimates $X_1(k)$ and $X_2(k)$. Since only two iterative values $X_1(k)$ and $X_2(k)$ are needed, we take the average of $X_{11}(k)$, $X_{12}(k)$, $X_{13}(k)$, $X_{14}(k)$, $X_{15}(k)$, $X_{16}(k)$, $X_{17}(k)$ and $X_{18}(k)$ as the iterative value $X_1(k)$ and $X_{21}(k)$, $X_{22}(k)$, $X_{23}(k)$, $X_{24}(k)$, $X_{25}(k)$, $X_{26}(k)$, $X_{27}(k)$ and $X_{28}(k)$ as the iterative value $X_2(k)$, respectively. We have the following algorithm:

$$\Psi_1 = F_1 - \mathcal{L}_{11}(X_1(k)) - \mathcal{L}_{12}(X_2(k)), \tag{48}$$

$$\Psi_2 = F_2 - \mathcal{L}_{21}(X_1(k)) - \mathcal{L}_{22}(X_2(k)), \tag{49}$$

$$X_{11}(k+1) = X_1(k) + \mu A_{11}^H \Psi_1 B_{11}^H, \tag{50}$$

$$X_{12}(k+1) = X_1(k) + \mu C_{11}^T \bar{\Psi}_1 D_{11}^T, \tag{51}$$

$$X_{13}(k+1) = X_1(k) + \mu \bar{H}_{11} \Psi_1^T \bar{G}_{11}, \tag{52}$$

$$X_{14}(k+1) = X_1(k) + \mu N_{11} \Psi_1^H M_{11}, \tag{53}$$

$$X_{21}(k+1) = X_2(k) + \mu A_{12}^H \Psi_1 B_{12}^H, \tag{54}$$

$$X_{22}(k+1) = X_2(k) + \mu C_{12}^T \bar{\Psi}_1 D_{12}^T, \tag{55}$$

$$X_{23}(k+1) = X_2(k) + \mu \bar{H}_{12} \Psi_1^T \bar{G}_{12}, \tag{56}$$

$$X_{24}(k+1) = X_2(k) + \mu N_{12} \Psi_1^H M_{12}, \tag{57}$$

$$X_{15}(k+1) = X_1(k) + \mu A_{21}^H \Psi_2 B_{21}^H, \tag{58}$$

$$X_{16}(k+1) = X_1(k) + \mu C_{21}^T \bar{\Psi}_2 D_{21}^T, \tag{59}$$

$$X_{17}(k+1) = X_1(k) + \mu \bar{H}_{21} \Psi_2^T \bar{G}_{21}, \tag{60}$$

$$X_{18}(k+1) = X_1(k) + \mu N_{21} \Psi_2^H M_{21}, \tag{61}$$

$$X_{25}(k+1) = X_2(k) + \mu A_{22}^H \Psi_2 B_{22}^H, \tag{62}$$

$$X_{26}(k+1) = X_2(k) + \mu C_{22}^T \bar{\Psi}_2 D_{22}^T, \tag{63}$$

$$X_{27}(k+1) = X_2(k) + \mu \bar{H}_{22} \Psi_2^T \bar{G}_{21}, \tag{64}$$

$$X_{28}(k+1) = X_2(k) + \mu N_{22} \Psi_2^H M_{22}. \tag{65}$$

$$X_1(k) = \frac{1}{8} \sum_{p=1}^8 X_{1p}(k), \tag{66}$$

$$X_2(k) = \frac{1}{8} \sum_{q=1}^8 X_{2q}(k). \tag{67}$$

Theorem 6. If the complex matrix Equation (5) has a unique solution pair X_1^* and X_2^* , then the iterative solution pair $X_1(k)$ and $X_2(k)$ given by the algorithm in Equations (50)-(66), converges to the unique solution pair X_1^* and X_2^* if and only if

$$0 < \mu < \frac{16}{\lambda_{\max}^2[\Phi_\sigma]} =: \mu_{\max}. \tag{68}$$

Under these conditions, the best convergence factor μ_0 is

$$\mu_0 = \frac{16}{\lambda_{\max}^2[\Phi_\sigma] + \lambda_{\min}^2[\Phi_\sigma]}. \tag{69}$$

Proof. Similar to symbols in Equation (3), we use the following notation:

$$\begin{cases} \Psi_{1\sigma} := (\Psi_1)_\sigma, \\ \Psi_{2\sigma} := (\Psi_2)_\sigma, \\ \mathcal{L}_{ij\sigma}^T(X) := (\mathcal{L}_{ij}^T(X))_\sigma, \\ \overline{\mathcal{L}_{ij\sigma}(X)} := (\overline{\mathcal{L}_{ij}(X)})_\sigma, \\ \mathcal{L}_{ij\sigma}^H(X) := (\mathcal{L}_{ij}^H(X))_\sigma. \end{cases} \tag{70}$$

For convenience, we set $X_{j\sigma}(k) := (X_j(k))_\sigma$. Taking the real representation of the both sides of Equation (50) gives

$$\begin{aligned} X_{11\sigma}(k+1) &= X_{1\sigma}(k) + \mu(A_{11}^H \Psi_1 B_{11}^H)_\sigma \\ &= X_{1\sigma}(k) + \mu A_{11\sigma}^H (\overline{\Psi_1})_\sigma B_{11\sigma}^H \\ &= X_{1\sigma}(k) + \mu A_{11\sigma}^H P_n \Psi_{1\sigma} P_n B_{11\sigma}^H \\ &= X_{1\sigma}(k) + \mu P_n A_{11\sigma}^T \Psi_{1\sigma} B_{11\sigma}^T P_n. \end{aligned} \tag{71}$$

Similarly, taking the real representation of Equations (50)-(64) give

$$X_{12\sigma}(k+1) = X_{1\sigma}(k) + \mu C_{11\sigma}^T \Psi_{1\sigma} D_{11\sigma}^T, \tag{72}$$

$$X_{13\sigma}(k+1) = X_{1\sigma}(k) + \mu P_n H_{11\sigma} \Psi_{1\sigma}^T G_{11\sigma} P_n, \tag{73}$$

$$X_{14\sigma}(k+1) = X_{1\sigma}(k) + \mu N_{11\sigma} \Psi_{1\sigma}^T M_{11\sigma}, \tag{74}$$

$$X_{21\sigma}(k+1) = X_{2\sigma}(k) + \mu P_n A_{12\sigma}^T \Psi_{1\sigma} B_{12\sigma}^T P_n, \tag{75}$$

$$X_{22\sigma}(k+1) = X_{2\sigma}(k) + \mu C_{12\sigma}^T \Psi_{1\sigma} D_{12\sigma}^T, \tag{76}$$

$$X_{23\sigma}(k+1) = X_{2\sigma}(k) + \mu P_n H_{12\sigma} \Psi_{1\sigma}^T G_{12\sigma} P_n, \tag{77}$$

$$X_{24\sigma}(k+1) = X_{2\sigma}(k) + \mu N_{12\sigma} \Psi_{1\sigma}^T M_{12\sigma}, \tag{78}$$

$$X_{15\sigma}(k+1) = X_{1\sigma}(k) + \mu P_n A_{21\sigma}^T \Psi_{2\sigma} B_{21\sigma}^T P_n, \tag{79}$$

$$X_{16\sigma}(k+1) = X_{1\sigma}(k) + \mu C_{21\sigma}^T \Psi_{2\sigma} D_{21\sigma}^T, \tag{80}$$

$$X_{17\sigma}(k+1) = X_{1\sigma}(k) + \mu P_n H_{21\sigma} \Psi_{2\sigma}^T G_{21\sigma} P_n, \tag{81}$$

$$X_{18\sigma}(k+1) = X_{1\sigma}(k) + \mu N_{21\sigma} \Psi_{2\sigma}^T M_{21\sigma}, \tag{82}$$

$$X_{25\sigma}(k+1) = X_{2\sigma}(k) + \mu P_n A_{22\sigma}^T \Psi_{2\sigma} B_{22\sigma}^T P_n, \tag{83}$$

$$X_{26\sigma}(k+1) = X_{2\sigma}(k) + \mu C_{22\sigma}^T \Psi_{2\sigma} D_{22\sigma}^T, \tag{84}$$

$$X_{27\sigma}(k+1) = X_{2\sigma}(k) + \mu P_n H_{22\sigma} \Psi_{2\sigma}^T G_{22\sigma} P_n, \tag{85}$$

$$X_{28\sigma}(k+1) = X_{2\sigma}(k) + \mu N_{22\sigma} \Psi_{2\sigma}^T M_{22\sigma}, \tag{86}$$

$$X_{1\sigma}(k) = \frac{1}{8} \sum_{p=1}^8 X_{1p\sigma}(k), \quad X_{2\sigma}(k) = \frac{1}{8} \sum_{q=1}^8 X_{2q\sigma}(k). \tag{87}$$

Taking the vec-operator of Equations (71)-(86) respectively gives

$$\text{col}[X_{11\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \mu \left((B_{11\sigma}^T P_n)^T \otimes P_n A_{11\sigma}^T \right) \text{col}[\Psi_{1\sigma}], \tag{88}$$

$$\text{col}[X_{12\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \mu (D_{11\sigma} \otimes C_{11\sigma}^T) \text{col}[\Psi_{1\sigma}], \tag{89}$$

$$\text{col}[X_{13\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \mu \left((G_{11\sigma} P_n)^T \otimes P_n H_{11\sigma} \right) P(2n, 2n) \text{col}[\Psi_{1\sigma}], \tag{90}$$

$$\text{col}[X_{14\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \mu (M_{11\sigma}^T \otimes N_{11\sigma}) P(2n, 2n) \text{col}[\Psi_{1\sigma}], \tag{91}$$

$$\text{col}[X_{21\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \mu \left((B_{12\sigma}^T P_n)^T \otimes P_n A_{12\sigma}^T \right) \text{col}[\Psi_{1\sigma}], \tag{92}$$

$$\text{col}[X_{22\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \mu (D_{12\sigma} \otimes C_{12\sigma}^T) \text{col}[\Psi_{1\sigma}], \tag{93}$$

$$\text{col}[X_{23\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \mu \left((G_{12\sigma} P_n)^T \otimes P_n H_{12\sigma} \right) P(2n, 2n) \text{col}[\Psi_{1\sigma}], \tag{94}$$

$$\text{col}[X_{24\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \mu (M_{12\sigma}^T \otimes N_{12\sigma}) P(2n, 2n) \text{col}[\Psi_{1\sigma}], \tag{95}$$

$$\text{col}[X_{15\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \mu \left((B_{21\sigma}^T P_n)^T \otimes P_n A_{21\sigma}^T \right) \text{col}[\Psi_{2\sigma}], \tag{96}$$

$$\text{col}[X_{16\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \mu (D_{21\sigma} \otimes C_{21\sigma}^T) \text{col}[\Psi_{2\sigma}], \tag{97}$$

$$\text{col}[X_{17\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \mu \left((G_{21\sigma} P_n)^T \otimes P_n H_{21\sigma} \right) P(2n, 2n) \text{col}[\Psi_{2\sigma}], \tag{98}$$

$$\text{col}[X_{18\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \mu (M_{21\sigma}^T \otimes N_{21\sigma}) P(2n, 2n) \text{col}[\Psi_{2\sigma}], \tag{99}$$

$$\text{col}[X_{25\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \mu \left((B_{22\sigma}^T P_n)^T \otimes P_n A_{22\sigma}^T \right) \text{col}[\Psi_{2\sigma}], \tag{100}$$

$$\text{col}[X_{26\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \mu (D_{22\sigma} \otimes C_{22\sigma}^T) \text{col}[\Psi_{2\sigma}], \tag{101}$$

$$\text{col}[X_{27\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \mu \left((G_{22\sigma} P_n)^T \otimes P_n H_{22\sigma} \right) P(2n, 2n) \text{col}[\Psi_{2\sigma}], \tag{102}$$

$$\text{col}[X_{28\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \mu (M_{22\sigma}^T \otimes N_{22\sigma}) P(2n, 2n) \text{col}[\Psi_{2\sigma}], \tag{103}$$

$$\text{col}[X_{1\sigma}(k)] = \frac{1}{8} \sum_{p=1}^8 \text{col}[X_{1p\sigma}(k)], \tag{104}$$

$$\text{col}[X_{2\sigma}(k)] = \frac{1}{8} \sum_{q=1}^8 \text{col}[X_{2q\sigma}(k)]. \tag{105}$$

Using notation in Equations (8) and Equations (104) and (105), the above equations can be equivalently written as

$$\text{col}[X_{1\sigma}(k+1)] = \text{col}[X_{1\sigma}(k)] + \frac{\mu}{8} [\Phi_{11\sigma}^T, \Phi_{21\sigma}^T] \text{col}[\Psi_{1\sigma}, \Psi_{2\sigma}], \tag{106}$$

$$\text{col}[X_{2\sigma}(k+1)] = \text{col}[X_{2\sigma}(k)] + \frac{\mu}{8} [\Phi_{12\sigma}^T, \Phi_{22\sigma}^T] \text{col}[\Psi_{1\sigma}, \Psi_{2\sigma}]. \tag{107}$$

Combining these two equations gives

$$\begin{aligned} & \text{col}[X_{1\sigma}(k+1), X_{2\sigma}(k+1)] \\ &= \text{col}[X_{1\sigma}(k), X_{2\sigma}(k)] + \frac{\mu}{8} \begin{bmatrix} \Phi_{11\sigma}^T & \Phi_{21\sigma}^T \\ \Phi_{12\sigma}^T & \Phi_{22\sigma}^T \end{bmatrix} \text{col}[\Psi_{1\sigma}, \Psi_{2\sigma}] \\ &= \text{col}[X_{1\sigma}(k), X_{2\sigma}(k)] + \frac{\mu}{8} \begin{bmatrix} \Phi_{11\sigma} & \Phi_{12\sigma} \\ \Phi_{21\sigma} & \Phi_{22\sigma} \end{bmatrix}^T \text{col}[\Psi_{1\sigma}, \Psi_{2\sigma}] \\ &= \text{col}[X_{1\sigma}(k), X_{2\sigma}(k)] + \frac{\mu}{8} \Phi_{\sigma}^T \text{col}[\Psi_{1\sigma}, \Psi_{2\sigma}]. \end{aligned} \tag{108}$$

According to notation in Equations (48) and (49) and using the conclusions in Equations (12) and (7), we have

$$\begin{aligned} & \text{col}[\Psi_{1\sigma}, \Psi_{2\sigma}] \\ &= \text{col} \left[\begin{pmatrix} F_1 - \mathcal{L}_{11}(X_1(k)) - \mathcal{L}_{12}(X_2(k)) \\ F_2 - \mathcal{L}_{21}(X_1(k)) - \mathcal{L}_{22}(X_2(k)) \end{pmatrix}_{\sigma} \right] \\ &= \text{col}[F_{1\sigma}, F_{2\sigma}] - \text{col} \left[\begin{pmatrix} \mathcal{L}_{11}(X_1(k)) + \mathcal{L}_{12}(X_2(k)) \\ \mathcal{L}_{21}(X_1(k)) + \mathcal{L}_{22}(X_2(k)) \end{pmatrix}_{\sigma} \right] \\ &= \text{col}[F_{1\sigma}, F_{2\sigma}] - \text{col} \left[\begin{pmatrix} \mathcal{L}_{11\sigma}(X_1(k)) + \mathcal{L}_{12\sigma}(X_2(k)) \\ \mathcal{L}_{21\sigma}(X_1(k)) + \mathcal{L}_{22\sigma}(X_2(k)) \end{pmatrix} \right] \\ &= \begin{bmatrix} \text{col}[F_{1\sigma}] \\ \text{col}[F_{2\sigma}] \end{bmatrix} - \begin{bmatrix} \text{col}[\mathcal{L}_{11\sigma}(X_1(k)) + \mathcal{L}_{12\sigma}(X_2(k))] \\ \text{col}[\mathcal{L}_{21\sigma}(X_1(k)) + \mathcal{L}_{22\sigma}(X_2(k))] \end{bmatrix} \\ &= \begin{bmatrix} \text{col}[F_{1\sigma}] \\ \text{col}[F_{2\sigma}] \end{bmatrix} - \begin{bmatrix} \Phi_{11\sigma} \text{col}[X_{1\sigma}(k)] + \Phi_{12\sigma} \text{col}[X_{2\sigma}(k)] \\ \Phi_{21\sigma} \text{col}[X_{1\sigma}(k)] + \Phi_{22\sigma} \text{col}[X_{2\sigma}(k)] \end{bmatrix} \\ &= \begin{bmatrix} \text{col}[F_{1\sigma}] \\ \text{col}[F_{2\sigma}] \end{bmatrix} - \begin{bmatrix} \Phi_{11\sigma} & \Phi_{12\sigma} \\ \Phi_{21\sigma} & \Phi_{22\sigma} \end{bmatrix} \begin{bmatrix} \text{col}[X_{1\sigma}(k)] \\ \text{col}[X_{2\sigma}(k)] \end{bmatrix} \end{aligned} \tag{109}$$

Substituting Equation (109) into Equation (108) gives

$$\begin{aligned} & \text{col}[X_{1\sigma}(k+1), X_{2\sigma}(k+1)] \\ &= \left(I_{8n^2} - \frac{\mu}{8} \Phi_{\sigma}^T \Phi_{\sigma} \right) \text{col}[X_{1\sigma}(k), X_{2\sigma}(k)] + \frac{\mu}{8} \Phi_{\sigma}^T \text{col}[F_{1\sigma}, F_{2\sigma}]. \end{aligned} \tag{110}$$

According to Lemma 4, if the algorithm in Equation (110) is convergent if and only if the convergence factor μ satisfies

$$0 < \frac{\mu}{8} < \frac{2}{\lambda_{\max}[\Phi_{\sigma}^T \Phi_{\sigma}]}. \tag{111}$$

Namely, we have

$$0 < \mu < \frac{16}{\lambda_{\max}^2 [\Phi_\sigma]}.$$

The remainder of this theorem can be proved by referencing to Lemma 4. The proof is completed. \square

Remark 1. Although some similar iterative algorithms have been established, in the preceding proof [3], there is only the proof about the sufficient condition of the convergence factor for the convergence. Using the real representation tool, Theorem 6 offers the sufficient and necessary condition proof for the first time.

Remark 2. Although the sufficient and necessary condition of the convergence for algorithm in Equations (50)-(66) is given in Theorem 6, the Kronecker product and the real presentation of the coefficient matrices are involved in computing μ_{\max} , this results in the high dimensionality problem. To overcome this drawback, we give a sufficient condition for the suggested iterative algorithm.

Corollary 1. If the system of the complex coupled matrix Equations (5) has a unique solution pair X_{1*} and X_{2*} , then the iterative algorithm in Equations (50)-(66) will converge to the solution pair X_{1*} and X_{2*} if

$$0 < \mu \leq \frac{1}{\Sigma} =: \mu_1. \tag{112}$$

Here

$$\Sigma := \sum_{i,j=1}^2 \left(\|B_{ij}\|^2 \|A_{ij}\|^2 + \|D_{ij}\|^2 \|C_{ij}\|^2 + \|H_{ij}\|^2 \|G_{ij}\|^2 + \|N_{ij}\|^2 \|M_{ij}\|^2 \right).$$

Proof. Recall $\sigma_{\max}[A] \leq \|A\|$ and $A \otimes B = \|A\| \|B\|$ and refer to Lemma 3, we have

$$\begin{aligned} \lambda_{\max}^2 [\Phi_\sigma] &= \sigma_{\max}^2 [\Phi_\sigma] \leq \|\Phi_\sigma\|^2 \\ &= \left\| \begin{bmatrix} \Phi_{11\sigma} & \Phi_{12\sigma} \\ \Phi_{21\sigma} & \Phi_{22\sigma} \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \|\Phi_{11\sigma}\| & \|\Phi_{12\sigma}\| \\ \|\Phi_{21\sigma}\| & \|\Phi_{22\sigma}\| \end{bmatrix} \right\|^2 \\ &= \left(\sum_{i,j=1}^2 \|\Phi_{ij\sigma}\| \right)^2 \leq 4 \sum_{i,j=1}^2 \|\Phi_{ij\sigma}\|^2. \end{aligned} \tag{113}$$

Noting the notation in Equation (8), we have

$$\begin{aligned} \|\Phi_{ij\sigma}\| &\leq \left\| \left((P_n B_{ij\sigma})^T \otimes A_{ij\sigma} P_n \right) \right\| + \left\| (D_{ij\sigma}^T \otimes C_{ij\sigma}) \right\| \\ &\quad + \left\| \left((P_n H_{ij\sigma})^T \otimes G_{ij\sigma} P_n \right) P(2n, 2n) \right\| + \left\| (N_{ij\sigma}^T \otimes M_{ij\sigma}) P(2n, 2n) \right\| \\ &= \left\| (P_n B_{ij\sigma})^T \right\| \|A_{ij\sigma} P_n\| + \|D_{ij\sigma}^T\| \|C_{ij\sigma}\| + \left\| (P_n H_{ij\sigma})^T \right\| \|G_{ij\sigma} P_n\| + \|N_{ij\sigma}^T\| \|M_{ij\sigma}\| \tag{114} \\ &= \|P_n B_{ij\sigma}\| \|A_{ij\sigma} P_n\| + \|D_{ij\sigma}\| \|C_{ij\sigma}\| + \|P_n H_{ij\sigma}\| \|G_{ij\sigma} P_n\| + \|N_{ij\sigma}\| \|M_{ij\sigma}\| \\ &= \|B_{ij\sigma}\| \|A_{ij\sigma}\| + \|D_{ij\sigma}\| \|C_{ij\sigma}\| + \|H_{ij\sigma}\| \|G_{ij\sigma}\| + \|N_{ij\sigma}\| \|M_{ij\sigma}\| \\ &= 4 \left(\|B_{ij}\| \|A_{ij}\| + \|D_{ij}\| \|C_{ij}\| + \|H_{ij}\| \|G_{ij}\| + \|N_{ij}\| \|M_{ij}\| \right). \end{aligned}$$

Combining Equations (113) and (114) gives

$$\lambda_{\max}^2 [\Phi_\sigma] \leq 16 \sum_{i,j=1}^2 \left(\|B_{ij}\| \|A_{ij}\| + \|D_{ij}\| \|C_{ij}\| + \|H_{ij}\| \|G_{ij}\| + \|N_{ij}\| \|M_{ij}\| \right)^2. \quad (115)$$

Substituting Equation (115) into Equation (112) gives Equation (112). The proof is completed. \square

In the practical work, condition

$$\lambda_{\max}^2 [\Phi_\sigma] = 16 \sum_{i,j=1}^2 \left(\|B_{ij}\| \|A_{ij}\| + \|D_{ij}\| \|C_{ij}\| + \|H_{ij}\| \|G_{ij}\| + \|N_{ij}\| \|M_{ij}\| \right)^2$$

almost does not hold, so we use symbol “ \leq ” to replace “ $<$ ” in Corollary 1. Clearly, the Frobenius norm

$$\sum_{i,j=1}^2 \left(\|B_{ij}\| \|A_{ij}\| + \|D_{ij}\| \|C_{ij}\| + \|H_{ij}\| \|G_{ij}\| + \|N_{ij}\| \|M_{ij}\| \right)^2$$

is easier to compute than $\lambda_{\max}^2 [\Phi_\sigma]$ in practical work.

4. Numerical Test

In this section, we bring a numerical test to validate the efficacy of the related conclusions. Equation (5) includes some coupled complex matrix equations as its special cases. For convenience, we consider the following equation

$$\begin{cases} AXB + C\bar{Y}D = F_1, \\ GX^T H + MY^H N = F_2. \end{cases}$$

with

$$\begin{aligned} A &= \begin{bmatrix} 19+11i & 1 \\ 4+9i & 19+13i \end{bmatrix}, & B &= \begin{bmatrix} 14+11i & 10+7i \\ 1+7i & 20+10i \end{bmatrix}, \\ C &= \begin{bmatrix} 17+20i & 0 \\ 5+9i & 20+20i \end{bmatrix}, & D &= \begin{bmatrix} 16+17i & 9+2i \\ 7+8i & 11+14i \end{bmatrix}, \\ G &= \begin{bmatrix} 14+20i & 5+5i \\ 5+3i & 13+16i \end{bmatrix}, & H &= \begin{bmatrix} 20+13i & 1+i \\ 2+8i & 15+16i \end{bmatrix}, \\ M &= \begin{bmatrix} 17+17i & 8+i \\ 4+4i & 10+16i \end{bmatrix}, & N &= \begin{bmatrix} 16+10i & 5+7i \\ 6+9i & 11+11i \end{bmatrix}, \\ F_1 &= \begin{bmatrix} 1223+2480i & 2262+1209i \\ -737+1962i & -2060+1125i \end{bmatrix}, \\ F_2 &= \begin{bmatrix} 643+2484i & -509+453i \\ 1132+43i & -1455-538i \end{bmatrix}. \end{aligned}$$

The unique solution pair of this system is

$$X_* = \begin{bmatrix} 3 & -i \\ -1+i & -2+3i \end{bmatrix}, \quad Y_* = t \begin{bmatrix} 2 & 3i \\ 2+i & 1-2i \end{bmatrix}.$$

According to the suggested algorithm in Equations (50)-(66), we can construct the gradient-based algorithm for this system as follows:

$$X_1(k+1) = X(k) + A^H (F_1 - AX(k)B - C\bar{Y}(k)D) B^H, \quad (116)$$

$$X_2(k+1) = X(k) + \bar{H} (F_2 - GX^T(k)H - MY^H(k)N)^T \bar{G}, \tag{117}$$

$$Y_1(k+1) = Y(k) + C^T (F_1 - AX(k)B - CY(k)D)D^T, \tag{118}$$

$$Y_2(k+1) = Y(k) + N (F_2 - GX^T(k)H - MY^H(k)N)^H M, \tag{119}$$

$$X(k) = \frac{X_1(k) + X_2(k)}{2}, \quad Y(k) = \frac{Y_1(k) + Y_2(k)}{2}. \tag{120}$$

According to Corollary 1, a sufficient condition for the convergence factor μ is

$$0 < \mu \leq \frac{4}{\|A\|^2 \|B\|^2 + \|C\|^2 \|D\|^2 + \|G\|^2 \|H\|^2 + \|M\|^2 \|N\|^2} = 8.2144 \times 10^{-7} =: \mu_1.$$

We take $X(0) = Y(0) = \mathbf{1}_{2 \times 2} \times 10^{-6}$ as the initial iterative value and use the iterative technique in Equations (116)-(120) to compute $X(k)$ and $Y(k)$. The relative error with different convergence factors is shown in **Figure 1**, where

$$\delta := \sqrt{\frac{\|X(k) - X_*\|^2 + \|Y(k) - Y_*\|^2}{\|X_*\|^2 + \|Y_*\|^2}}$$

is the relative error.

From **Figure 1**, we find that the relative error δ becomes smaller and approaches zero as the iterative step k increases. This denotes that the suggested algorithm is efficacy and convergent. From the three convergence curves, we find that the convergence effectiveness is improved when the convergence factor is changed from $1.0\mu_1$ to $1.3\mu_1$ via $1.7\mu_1$. In this example, we have

$$\begin{aligned} \lambda_{\max}^2[\Phi] &= 2.9622 \times 10^6 < 4.8695 \times 10^6 \\ &= \|B\|_2^2 \|A\|_2^2 + \|D\|_2^2 \|C\|_2^2 + \|G\|_2^2 \|H\|_2^2 + \|M\|_2^2 \|N\|_2^2. \end{aligned}$$

so the range of the convergence factor of Corollary 1 is conservative. Actually, $1.7\mu_1$ plays the role of the best convergence factor in this numerical example.

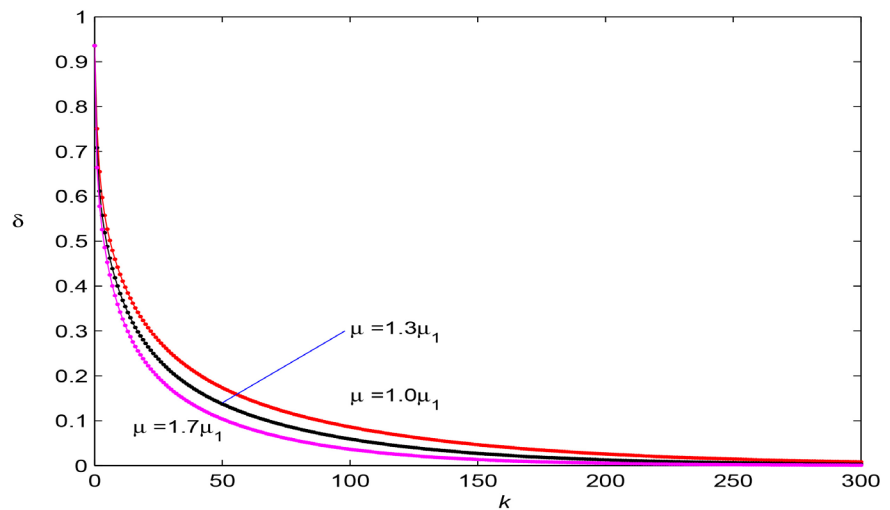


Figure 1. The relative error δ versus iteration k .

5. Conclusion

This paper brings the gradient-based iterative technique for a coupled complex conjugate and transpose matrix equation. The convergence proof is offered. In the proof process, the necessary and sufficient conditions for the convergence factor are determined to guarantee the convergence of the algorithm. And the optimal choice of the convergence factor is settled theoretically. A numerical test is offered to validate the efficacy of the suggested algorithm. The iterative strategy used in this paper can be used in the system identification.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Young, D.M. and Young, M.M. (2017) A General Hermitian Nonnegative-Definite Solution to the Matrix Equation $AXB = C$. *Advances in Linear Algebra & Matrix Theory*, **7**, 7-17. <https://doi.org/10.4236/alamt.2017.71002>
- [2] Gao, D.J. (2017) Iterative Methods for Solving the Nonlinear Matrix Equation $A - A^* X^p A - B^* X^{-q} B = I (0 < p, q < 1)$. *Advances in Linear Algebra & Matrix Theory*, **7**, 72-78.
- [3] Wu, A.G., Lv, L.L. and Duan, G.R. (2011) Iterative Algorithms for Solving a Class of Complex Conjugate and Transpose Matrix Equations. *Applied Mathematics and Computation*, **217**, 8343-8353. <https://doi.org/10.1016/j.amc.2011.02.113>
- [4] Gu, C.Q. and Xue, H.Y. (2009) A Shift-Splitting Hierarchical Identification Method for Solving Lyapunov Matrix Equations. *Linear Algebra and its Applications*, **430**, 1517-1530. <https://doi.org/10.1016/j.laa.2008.01.010>
- [5] Zhou, B., Li, Z.Y., Duan, G.R. and Wang, Y. (2009) Weighted Least Squares Solutions to General Coupled Sylvester Matrix Equations. *Journal of Computational and Applied Mathematics*, **224**, 759-776. <https://doi.org/10.1016/j.cam.2008.06.014>
- [6] Zhou, B., Lam, J. and Duan, G.R. (2010) Gradient-Based Maximal Convergence Rate Iterative Method for Solving Linear Matrix Equations. *International Journal of Computer Mathematics*, **87**, 515-527. <https://doi.org/10.1080/00207160802123458>
- [7] Zhou, B., Lam J. and Duan, G.R. (2008) Convergence of Gradient-Based Iterative Solution of the Coupled Markovian Jump Lyapunov Equations. *Computers & Mathematics with Applications*, **56**, 3070-3078. <https://doi.org/10.1016/j.camwa.2008.07.037>
- [8] Zhou, B., Duan, G.R. and Li, Z.Y. (2009) Gradient Based Iterative Algorithm for Solving Coupled Matrix Equations. *Systems & Control Letters*, **58**, 327-333. <https://doi.org/10.1016/j.sysconle.2008.12.004>
- [9] Li, Z.Y., Wang, Y., Zhou, B. and Duan, G.R. (2010) Least Squares Solution with the Minimum-Norm to General Matrix Equations via Iteration. *Applied Mathematics and Computation*, **215**, 3547-3562. <https://doi.org/10.1016/j.amc.2009.10.052>
- [10] Jian, T.S. and Wei, M.S. (2003) On Solutions of the Matrix Equations $X - AXB = C$ and $X - A\bar{X}B = C$. *Linear Algebra and its Applications*, **367**, 225-233. [https://doi.org/10.1016/S0024-3795\(02\)00633-X](https://doi.org/10.1016/S0024-3795(02)00633-X)

- [11] Zhou, B., Lam, J. and Duan, G.R. (2011) Toward Solution of Matrix Equations $X = Af(X)B + C$. *Linear Algebra and its Applications*, **435**, 1370-1398.
- [12] Ding, F., Liu, X.P. and Ding, J. (2008) Iterative Solutions of the Generalized Sylvester Matrix Equations by Using the Hierarchical Identification Principle. *Applied Mathematics and Computation*, **197**, 41-50.
<https://doi.org/10.1016/j.amc.2007.07.040>