axioms

MDPI

Article

The Weighted, Relaxed Gradient-Based Iterative Algorithm for the Generalized Coupled Conjugate and Transpose Sylvester Matrix Equations

Xiaowen Wu, Zhengge Huang *, Jingjing Cui and Yanping Long

College of Mathematics and Physics, Center for Applied Mathematics of Guangxi, Guangxi Minzu University, Nanning 530006, China; xiaowenwu1998@163.com (X.W.); jingjingcui1990@163.com (J.C.); yanpinglong1999@163.com (Y.L.)
* Correspondence: zhenggehuang@163.com

Abstract: By applying the weighted relaxation technique to the gradient-based iterative (GI) algorithm and taking proper weighted combinations of the solutions, this paper proposes the weighted, relaxed gradient-based iterative (WRGI) algorithm to solve the generalized coupled conjugate and transpose Sylvester matrix equations. With the real representation of a complex matrix as a tool, the necessary and sufficient conditions for the convergence of the WRGI algorithm are determined. Also, some sufficient convergence conditions of the WRGI algorithm are presented. Moreover, the optimal step size and the corresponding optimal convergence factor of the WRGI algorithm are given. Lastly, some numerical examples are provided to demonstrate the effectiveness, feasibility and superiority of the proposed algorithm.

Keywords: generalized coupled conjugate and transpose Sylvester matrix equations; weighted relaxed gradient-based iterative algorithm; real representation; relaxation parameter; convergence condition; optimal convergence factor

MSC: 15A06; 15A24; 65F45

check for updates

Citation: Wu, X.; Huang, Z.; Cui, J.; Long, Y. The Weighted, Relaxed Gradient-Based Iterative Algorithm for the Generalized Coupled Conjugate and Transpose Sylvester Matrix Equations. *Axioms* **2023**, *12*, 1062. https://doi.org/10.3390/axioms12111062

Academic Editor: Simeon Reich

Received: 8 October 2023
Revised: 10 November 2023
Accepted: 17 November 2023
Published: 20 November 2023

1. Introduction

Matrix equations are often used in mathematics and engineering applications, such as control theory, signal processing and computational mathematics [1–4]. For example, the forward and backward periodic Sylvester matrix equations (PSMEs) with the following forms

$$A_iX_iB_i + C_iX_{i+1}D_i = F_i, i = 1, 2, \cdots, \omega$$

and

$$A_iX_{i+1}B_i + C_iX_iD_i = F_i, i = 1, 2, \cdots, \omega,$$

are an indispensable part of pole assignment and the design of state observers for linear discrete periodic systems [5]. Thus, studying the computational methods of matrix equations has become an important subject in the field of computational mathematics and control. For matrix equations, computing their exact solutions is very meaningful for many practical problems. However, in many applications, such as stability analysis of control systems, it is usually not necessary to calculate the exact solution, as the approximate solution is sufficient. Therefore, the research of iterative solutions to matrix equations has attracted many researchers [6–12].

One of the important ways to study the approximate solutions of matrix equations is to establish iterative methods. In recent years, many researchers have proposed a great deal of iterative methods to solve different kinds of Sylvester matrix equations. For the generalized coupled Sylvester matrix equations

$$\sum_{j=1}^{q} A_{ij}X_jB_{ij} = F_i, i = 1, 2, \cdots, p, \tag{1}$$

where $A_{ij} \in \mathbb{R}^{r_i \times m_j}, B_{ij} \in \mathbb{R}^{n_j \times s_i}$, $F_i \in \mathbb{R}^{r_i \times s_i}$ and $X_j \in \mathbb{R}^{m_j \times n_j}, i = 1, 2, \cdots, p, j = 1, 2, \cdots, q$, Ding and Chen [13] applied the hierarchical identification principle and introduced the block matrix inner product to propose the GI algorithm. Based on the idea of the GI algorithm, some researchers established many improved versions of the GI algorithm and investigated their convergence properties [14,15]. To improve the convergence rate of the GI algorithm, Zhang [16] proposed the residual norm steepest descent (RNSD), conjugate gradient normal equation (CGNE) and biconjugate gradient stabilized (Bi-CGSTAB) algorithms to solve the matrix in Equation (1). Subsequently, Zhang [17] constructed the full-column rank, full-row rank and reduced-rank gradient-based algorithms, with the main idea of them being to construct an objective function and use the gradient search. Afterward, Zhang and Yin [18] developed the conjugate gradient least squares (CGLS) algorithm for Equation (1), which can be convergent within the finite iteration steps in the absence of round off errors.

The generalized coupled Sylvester-conjugate matrix equations have the following form:

$$\sum_{j=1}^{q} \left(A_{ij}X_jB_{ij} + C_{ij}\overline{X_j}D_{ij}\right) = F_i, i = 1, 2, \cdots, p, \tag{2}$$

where $A_{ij}, C_{ij} \in \mathbb{C}^{m_i \times r_j}, B_{ij}, D_{ij} \in \mathbb{C}^{t_j \times n_i}$, $F_i \in \mathbb{C}^{m_i \times n_i}$ and $X_j \in \mathbb{C}^{r_j \times t_j}, i = 1, 2, \cdots, p$, $j = 1, 2, \cdots, q$. The matrix in Equation (2) can be regarded as the generalization of Equation (1) in the complex field. For the matrix in Equation (2), Wu et al. [19] extended the GI algorithm to solve it and derived the sufficient condition for the convergence of the GI algorithm. Due to the fact that the sufficient condition in [19] is somewhat conservative, Huang and Ma [20] established the sufficient and necessary conditions for convergence of the GI algorithm based on the properties of the real representation of a complex matrix and the vec operator. Also, they made use of different definitions of the real representation to derive another sufficient and necessary condition for the convergence of the GI algorithm. In [21], Huang and Ma introduced l relaxation factors into the GI algorithm and proposed two relaxed gradient-based iterative (RGI) algorithms. They proved that the RGI algorithms are convergent under suitable restrictions in light of the real representation of a complex matrix and the vec operator. Quite recently, Wang et al. [22] developed a cyclic gradient-based iterative (CGI) algorithm by introducing a modular operator, which is different from previous iterative methods. The most remarkable advantage of the CGI algorithm is that less information is used in each iteration update, which helps save memory and improve efficiency.

In addition, the generalized, coupled Sylvester-transpose matrix equations

$$\sum_{j=1}^{l} \left(A_{ij}X_jB_{ij} + C_{ij}X_j^TD_{ij}\right) = F_i, i = 1, 2, \cdots, s, \tag{3}$$

where $A_{ij} \in \mathbb{R}^{m_i \times r_j}, C_{ij} \in \mathbb{R}^{m_i \times t_j}, B_{ij} \in \mathbb{R}^{t_j \times n_i}, D_{ij} \in \mathbb{R}^{r_j \times n_i}$, $F_i \in \mathbb{R}^{m_i \times n_i}$ and $X_j \in \mathbb{R}^{r_j \times t_j}$, $i = 1, 2, \cdots, s$, $j = 1, 2, \cdots, l$ are related to fault detection, observer design and so forth. Due to the important role of the generalized, coupled Sylvester-transpose matrix equations in several applied problems, numerous methods have been developed to solve them. For example, Song et al. [23] constructed the GI algorithm for the matrix in Equation (3) by using the principle of hierarchical identification. According to the rank of the related matrices of the matrix in Equation (3), Huang and Ma [24] developed three relaxed gradient-based iterative (RGI) algorithms recently by minimizing the objective functions.

In [25], Beik et al. considered the following matrix equations:

$$\mathrm{T}_v(X) = \sum_{i=1}^{p}\left(\sum_{\mu=1}^{s_1} A_{vi\mu}X_iB_{vi\mu} + \sum_{\mu=1}^{s_2} C_{vi\mu}X_i^TD_{vi\mu} + \sum_{\mu=1}^{s_3} M_{vi\mu}\overline{X_i}N_{vi\mu} + \sum_{\mu=1}^{s_4} H_{vi\mu}X_i^HG_{vi\mu}\right) = F_v, \quad (4)$$

where $A_{vi\mu}, B_{vi\mu}, C_{vi\mu}, D_{vi\mu}, M_{vi\mu}, N_{vi\mu}, H_{vi\mu}, G_{vi\mu}$ and F_v, $v = 1, 2, \cdots, N$ are the known matrices with suitable dimensions in a complex number field, $X = (X_1, X_2, \cdots, X_p)$ is a group of unknown matrices and X_j^H represents the conjugate transpose of the matrix X_j. The matrix in Equation (4) is quite general and includes several kinds of Sylvester matrix equations, and it can be viewed as a general form of the aforementioned matrices in Equations (1)–(3). By using the hierarchical identification principle, the authors in [25] put forward the GI algorithm over a group of reflexive (anti-reflexive) matrices.

Inspired by the above work, this paper focuses on solving the iterative solution of the generalized coupled conjugate and transpose Sylvester matrix equations:

$$\sum_{j=1}^{l}\left(A_{ij}X_jB_{ij} + C_{ij}\overline{X_j}D_{ij} + E_{ij}X_j^TF_{ij} + G_{ij}X_j^HH_{ij}\right) = M_i, i = 1, 2, \cdots, s, \quad (5)$$

where $A_{ij}, C_{ij} \in \mathbb{C}^{m_i \times r_j}, B_{ij}, D_{ij} \in \mathbb{C}^{s_j \times n_i}, E_{ij}, G_{ij} \in \mathbb{C}^{m_i \times s_j}, F_{ij}, H_{ij} \in \mathbb{C}^{r_j \times n_i}$ and $M_i \in \mathbb{C}^{m_i \times n_i}$, $i = 1, 2, \cdots, s$, $j = 1, 2, \cdots, l$ are the known matrices and $X_j \in \mathbb{C}^{r_j \times s_j}$ for $j = 1, 2, \cdots, l$ are the unknown matrices that need to be determined. When $l = 1$, Wang et al. [26] presented the relaxed gradient iterative (RGI) algorithm for Equation (5), which has four system parameters. Note that the matrices in Equations (1)–(3) are special cases of the matrix in Equation (5), and thus the results obtained in this work contain the existing ones in [20,23,24,27]. Owing to the fact that the convergence speed of the GI algorithm is slow in many cases, it is quite meaningful to further improve the numerical performance of the GI algorithm for the matrix in Equation (5). According to [21,24,26,28,29], it can be seen that the relaxation technique can ameliorate the numerical behaviors of the existing GI-like algorithms. Then, this motivated us to apply the weighted relaxation technique to the GI algorithm. By using different step size factors and the weighted technique, we construct the weighted, relaxed gradient-based iterative (WRGI) algorithm for solving the matrix in Equation (5). The WRGI algorithm contains s relaxation factors. When all relaxation factors are equal, the WRGI algorithm will reduce to the GI one proposed in [25]. In [25], the optimal convergence factor of the GI algorithm was not derived. Compared with the GI algorithm in [25], the WRGI algorithm proposed in this paper can own a higher computational efficiency, and its convergence properties are analyzed in detail, including the convergence conditions, optimal parameters and corresponding optimal convergence factor. The proposed WRGI algorithm has a faster convergence rate than the GI one by adjusting the values of the relaxation factors so it is more conducive to solving matrix equations in control theory, signal processing and computational mathematics, among other applications. The main contributions of this paper are given below:

- By using a series of step size factors and the weighted relaxation technique, we establish the weighted, relaxed gradient-based iterative (WRGI) algorithm for solving the matrix in Equation (5), which generalizes and improves the existing GI one in [25]. Aside from that, the proposed WRGI algorithm contains the RGI ones in [21,24].
- We analytically provide the necessary and sufficient condition for the convergence of the proposed WRGI algorithm. Also, the expressions of the optimal step size and the corresponding optimal convergence factor of the WRGI algorithm are derived.

The rest of this paper is organized as follows. In Section 2, some definitions and previous results are given, and the GI algorithm that has been proposed before is reviewed. In Section 3, we first design a new algorithm referred to as the WRGI algorithm to solve the matrix in Equation (5). Then, we prove that the WRGI algorithm is convergent for any started matrices under proper conditions, and we explicitly give the optimal step factor such that the convergence rate of the WRGI algorithm is maximized. Section 4 gives two numerical examples to demonstrate the effectiveness and advantages of the proposed WRGI algorithm. Finally, we give some concluding remarks to end this paper in Section 5.

2. Preliminaries

In this section, as a matter of convenience to discuss the main results of this paper, we describe the following notations which will be used throughout this paper. Let $\mathbb{C}^{n\times n}$ be the set of all $n \times n$ complex matrices. For a given matrix $A \in \mathbb{C}^{n\times n}$, some related notations are the following:

- $\overline{A}$ denotes the conjugate of the matrix A;
- A^T stands for the transpose of the matrix A;
- A^H represents the conjugate transpose of the matrix A;
- A^{-1} denotes the inverse of the matrix A;
- $\|A\|$ denotes the Frobenius norm of the matrix A;
- $\|A\|_2$ denotes the spectrum norm of the matrix A;
- $\sigma_{\max}(A)$ indicates the maximum singular value of the matrix A;
- $\sigma_{\min}(A)$ indicates the minimum singular value of the matrix A;
- $\rho(A)$ stands for the spectral radius of the matrix A;
- $\lambda(A)$ stands for the spectrum of the matrix A;
- $\mathrm{rank}(A)$ stands for the rank of the matrix A.

Moreover, some useful definitions and lemmas are given below:

Definition 1 ([30]). *For two matrices $E \in \mathbb{C}^{m\times n}$ and $F \in \mathbb{C}^{k\times l}$, the Kronecker product of the matrices E and F is defined as follows:*

$$E \otimes F = \begin{bmatrix} e_{11}F & e_{12}F & \cdots & e_{1n}F \\ e_{21}F & e_{22}F & \cdots & e_{2n}F \\ \vdots & \vdots & & \vdots \\ e_{m1}F & e_{m2}F & \cdots & e_{mn}F \end{bmatrix} = [e_{ij}F]_{m\times n} \in \mathbb{C}^{mk\times nl}. \tag{6}$$

Definition 2 ([27]). *Let e_{in} be the n-dimensional column vector whose ith element of e_{in} is one, and other elements are zero. Then, the vec permutation matrix $P(m,n)$ can be defined as follows:*

$$P(m,n) = \begin{bmatrix} I_m \otimes e_{1n}^T \\ I_m \otimes e_{2n}^T \\ \vdots \\ I_m \otimes e_{nn}^T \end{bmatrix}. \tag{7}$$

Definition 3 ([30]). *Let $A = [a_1, a_2, \cdots, a_n] \in \mathbb{C}^{m\times n}$, with a_i being the ith column of A. The vector-stretching function of A is defined as follows:*

$$\mathrm{vec}(A) = [a_1^T, a_2^T, \cdots, a_n^T]^T \in \mathbb{C}^{mn}. \tag{8}$$

Lemma 1 ([20]). *Let $CYD = G$ be a matrix equation, with the matrices C and D being full-column rank and full-row rank, respectively. Then, the iterative solution $Y(l)$ produced by the GI iteration process*

$$Y(l+1) = Y(l) + \tau C^H(G - CY(l)D)D^H \tag{9}$$

converges to the exact solution Y^* of $CYD = G$ for any initial matrix $Y(0)$ if and only if

$$0 < \tau < \frac{2}{\sigma_{\max}(C)^2\sigma_{\max}(D)^2}. \tag{10}$$

In addition, the optimal step size τ_0 is

$$\tau_0 = \frac{2}{\lambda_{\max}(C^H C)\lambda_{\max}(DD^H) + \lambda_{\min}(C^H C)\lambda_{\min}(DD^H)}.$$

Lemma 2 ([30]). *Let $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{s\times t}$ and $X \in \mathbb{C}^{n\times s}$. Then, we have the following:*
(i) $\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X)$;
(ii) $\mathrm{vec}(X^T) = P(n,s)\mathrm{vec}(X)$.

Lemma 3 ([19]). *Let $A, B \in \mathbb{C}^{n\times n}$. If $\mathrm{tr}(A) + \mathrm{tr}(B) \in \mathbb{R}$, then*

$$\mathrm{tr}(A) + \mathrm{tr}(B) = \overline{\mathrm{tr}(A) + \mathrm{tr}(B)} = \mathrm{tr}(\overline{A}) + \mathrm{tr}(\overline{B}).$$

Next, we introduce two definitions of the real representation of a complex matrix. For $A \in \mathbb{C}^{m\times n}$, A can be uniquely expressed as $A = A_1 + iA_2 \in \mathbb{C}^{m\times n}$, where $A_1, A_2 \in \mathbb{R}^{m\times n}$. We define the operators $(\cdot)^{\nabla}$ and $(\cdot)^{\blacktriangledown}$ as follows:

$$A^{\nabla} = \begin{bmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{bmatrix}, A^{\blacktriangledown} = \begin{bmatrix} A_2 & A_1 \\ A_1 & -A_2 \end{bmatrix}. \tag{11}$$

It can be seen from Equation (11) that the sizes of A^{∇} and $A^{\blacktriangledown}$ are two times that of A. Then, by combining Equation (11) with the definition of the Frobenius norm, we can obtain

$$\|A^{\nabla}\|^2 = 2\|A\|^2. \tag{12}$$

Aside from that, for the identity matrix I_n with the matrix order n, we define the following matrices:

$$Q_n = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}, P_n = \frac{\sqrt{2}}{2}\begin{bmatrix} iI_n & I_n \\ I_n & iI_n \end{bmatrix}.$$

The properties of the real representation of several complex matrices are given by the following lemma:

Lemma 4 ([21]). *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times r}$. Then, the following statements hold:*
(1)

$$(AB)^{\nabla} = A^{\nabla}B^{\nabla}, (A^T)^{\nabla} = Q_n(A^{\nabla})^T Q_m,$$
$$(A^H)^{\nabla} = (A^{\nabla})^T, (\bar{A})^{\nabla} = Q_m A^{\nabla} Q_n, A^{\blacktriangledown} = Q_m A^{\nabla}.$$

(2) If $A \in \mathbb{C}^{n\times n}$ is nonsingular, then $(A^{-1})^{\nabla} = (A^{\nabla})^{-1}$.
(3) $\|B^{\nabla}\|_2 = \|B^{\blacktriangledown}\|_2 = \|B\|_2$.
(4) If $n = r$, then it holds that $\rho(B) = \rho(B^{\nabla})$ and $\rho(B^{\blacktriangledown}) = \sqrt{\rho(B\overline{B})}$.

The lemma below gives the norm relationship between the block matrices and its submatrices:

Lemma 5 ([31]). *Let B be a block-partitioned matrix with*

$$B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & \vdots & & \vdots \\ B_{m1} & B_{m2} & \cdots & B_{mn} \end{bmatrix} \tag{13}$$

and B_{ij} be a matrix with a proper size. Then, it holds that

$$\|B\|_2 = \sigma_{\max}(B) \leq \|B\| = \left\| \begin{matrix} \|B_{11}\| & \|B_{12}\| & \cdots & \|B_{1n}\| \\ \|B_{21}\| & \|B_{22}\| & \cdots & \|B_{2n}\| \\ \vdots & \vdots & & \vdots \\ \|B_{m1}\| & \|B_{m2}\| & \cdots & \|B_{mn}\| \end{matrix} \right\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} \|B_{ij}\|^2}.$$

If $\operatorname{rank}(B) = 1$*, or if* B *is a vector, then we have*

$$\|B\|_2 = \|B\| = \sigma_{\max}(B).$$

The following result gives the norm inequality for the p norms of the block matrix and its submatrices, which is essential for analyzing the convergence of the WRGI algorithm:

Lemma 6 ([21])**.** *Let* $B = [B_{ij}]_{m\times n}$ *be a partitioned matrix with the form in Equation (13), and let the orders of the matrices* B_{ij} $(i = 1, \cdots, m, j = 1, \cdots, n)$ *be compatible. Then, for any deduced* q *norm, we have*

$$\|B\|_q \leq \left\| \begin{matrix} \|B_{11}\|_q & \|B_{12}\|_q & \cdots & \|B_{1n}\|_q \\ \|B_{21}\|_q & \|B_{22}\|_q & \cdots & \|B_{2n}\|_q \\ \vdots & \vdots & & \vdots \\ \|B_{m1}\|_q & \|B_{m2}\|_q & \cdots & \|B_{mn}\|_q \end{matrix} \right\|_q.$$

3. The Weighted, Relaxed Gradient-Based Iterative (WRGI) Algorithm and Its Convergence Analysis

In this section, we first propose the weighted, relaxed gradient-based iterative (WRGI) algorithm to solve the matrix in Equation (5) based on the hierarchical identification principle. Then, we discuss the convergence properties of the WRGI algorithm, which include the convergence conditions, optimal step size and the corresponding optimal convergence factor of the WRGI algorithm.

First, we define the following intermediate matrices for $i = 1, \cdots, s, p = 1, \cdots, l$ by applying the hierarchical identification principle:

$$\Pi_{ip} = M_i - \sum_{j=1}^{l} \left(A_{ij}X_jB_{ij} + C_{ij}\overline{X_j}D_{ij} + E_{ij}X_j^TF_{ij} + G_{ij}X_j^HH_{ij} \right) + A_{ip}X_pB_{ip}, \tag{14}$$

$$\gamma_{ip} = M_i - \sum_{j=1}^{l} \left(A_{ij}X_jB_{ij} + C_{ij}\overline{X_j}D_{ij} + E_{ij}X_j^TF_{ij} + G_{ij}X_j^HH_{ij} \right) + C_{ip}\overline{X_p}D_{ip}, \tag{15}$$

$$\Phi_{ip} = \left(M_i - \sum_{j=1}^{l} \left(A_{ij}X_jB_{ij} + C_{ij}\overline{X_j}D_{ij} + E_{ij}X_j^TF_{ij} + G_{ij}X_j^HH_{ij} \right) + E_{ip}X_p^TF_{ip} \right)^T, \tag{16}$$

$$\Omega_{ip} = \left(M_i - \sum_{j=1}^{l} \left(A_{ij}X_jB_{ij} + C_{ij}\overline{X_j}D_{ij} + E_{ij}X_j^TF_{ij} + G_{ij}X_j^HH_{ij} \right) + G_{ip}X_p^HH_{ip} \right)^H. \tag{17}$$

For the sake of the following discussions, we define

$$\Gamma_{ij} = A_{ij}X_jB_{ij} + C_{ij}\overline{X_j}D_{ij} + E_{ij}X_j^TF_{ij} + G_{ij}X_j^HH_{ij},$$
$$\Gamma_{ij}(k) = A_{ij}X_j(k)B_{ij} + C_{ij}\overline{X_j(k)}D_{ij} + E_{ij}X_j^T(k)F_{ij} + G_{ij}X_j^H(k)H_{ij}.$$

Then, Equations (14)–(17) can be written as

$$\Pi_{ip} = M_i - \sum_{j=1}^{l} \Gamma_{ij} + A_{ip} X_p B_{ip}, \tag{18}$$

$$\gamma_{ip} = \overline{M_i - \sum_{j=1}^{l} \Gamma_{ij} + C_{ip} \overline{X_p} D_{ip}}, \tag{19}$$

$$\Phi_{ip} = \left(M_i - \sum_{j=1}^{l} \Gamma_{ij} + E_{ip} X_p^T F_{ip} \right)^T, \tag{20}$$

$$\Omega_{ip} = \left(M_i - \sum_{j=1}^{l} \Gamma_{ij} + G_{ip} X_p^H H_{ip} \right)^H. \tag{21}$$

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

$$\Pi_{ip} = A_{ip} X_p B_{ip}, \tag{22}$$

$$\gamma_{ip} = \overline{C_{ip}} X_p \overline{D_{ip}}, \tag{23}$$

$$\Phi_{ip} = F_{ip}^T X_p E_{ip}^T, \tag{24}$$

$$\Omega_{ip} = H_{ip}^H X_p G_{ip}^H, \tag{25}$$

for $i = 1, \cdots, s, p = 1, \cdots, l$. By applying Lemma 1 to Equations (22)–(25), we can construct the recursive forms as follows

$$X_p^{1,i}(k+1) = X_p^{1,i}(k) + \mu A_{ip}^H \left[\Pi_{ip} - A_{ip} X_p^{1,i}(k) B_{ip} \right] B_{ip}^H, \tag{26}$$

$$X_p^{2,i}(k+1) = X_p^{2,i}(k) + \mu C_{ip}^T \left[\gamma_{ip} - \overline{C_{ip}} X_p^{2,i}(k) \overline{D_{ip}} \right] D_{ip}^T, \tag{27}$$

$$X_p^{3,i}(k+1) = X_p^{3,i}(k) + \mu \overline{F_{ip}} \left[\Phi_{ip} - F_{ip}^T X_p^{3,i}(k) E_{ip}^T \right] \overline{E_{ip}}, \tag{28}$$

$$X_p^{4,i}(k+1) = X_p^{4,i}(k) + \mu H_{ip} \left[\Omega_{ip} - H_{ip}^H X_p^{4,i}(k) G_{ip}^H \right] G_{ip}, \tag{29}$$

for $i = 1, \cdots, s, p = 1, \cdots, l$, where μ is a step size factor.

For convenience, we define the following notations:

$$\Gamma_{ij}^{r,i}(k) = A_{ij} X_j^{r,i}(k) B_{ij} + C_{ij} \overline{X_j^{r,i}(k)} D_{ij} + E_{ij} X_j^{r,i}(k)^T F_{ij} + G_{ij} X_j^{r,i}(k)^H H_{ij}, r = 1, 2, 3, 4.$$

Substituting Equations (18)–(21) into Equations (26)–(29), respectively, and then using $X_p^{1,i}(k), X_p^{2,i}(k), X_p^{3,i}(k)$ and $X_p^{4,i}(k)$ to replace X_p for $i = 1 \cdots, s, p = 1, \cdots, l$ gives

$$X_p^{1,i}(k+1) = X_p^{1,i}(k) + \mu A_{ip}^H \left[M_i - \sum_{j=1}^{l} \Gamma_{ij}^{1,i}(k) \right] B_{ip}^H,$$

$$X_p^{2,i}(k+1) = X_p^{2,i}(k) + \mu C_{ip}^T \overline{\left[M_i - \sum_{j=1}^{l} \Gamma_{ij}^{2,i}(k) \right]} D_{ip}^T,$$

$$X_p^{3,i}(k+1) = X_p^{3,i}(k) + \mu \overline{F_{ip}} \left[M_i - \sum_{j=1}^{l} \Gamma_{ij}^{3,i}(k) \right]^T \overline{E_{ip}},$$

$$X_p^{4,i}(k+1) = X_p^{4,i}(k) + \mu H_{ip} \left[M_i - \sum_{j=1}^{l} \Gamma_{ij}^{4,i}(k) \right]^H G_{ip}.$$

By taking the average of $X_p^{1,i}(k+1), X_p^{2,i}(k+1), X_p^{3,i}(k+1)$ and $X_p^{4,i}(k+1)$ in the above equations, one can obtain the following iterative algorithm:

$$\begin{aligned} X_p^{(1)}(k+1) &= X_p^{(1)}(k) + \frac{\mu}{4}\left\{ A_{1p}^H\left[M_1 - \sum_{j=1}^{l}\Gamma_{1j}(k)\right]B_{1p}^H + C_{1p}^T\overline{\left[M_1 - \sum_{j=1}^{l}\Gamma_{1j}(k)\right]}D_{1p}^T \right. \\ &\quad \left. + \overline{F_{1p}}\left[M_1 - \sum_{j=1}^{l}\Gamma_{1j}(k)\right]^T\overline{E_{1p}} + H_{1p}\left[M_1 - \sum_{j=1}^{l}\Gamma_{1j}(k)\right]^H G_{1p} \right\}, \\ &\cdots\cdots\cdots \\ X_p^{(s)}(k+1) &= X_p^{(s)}(k) + \frac{\mu}{4}\left\{ A_{sp}^H\left[M_s - \sum_{j=1}^{l}\Gamma_{sj}(k)\right]B_{sp}^H + C_{sp}^T\overline{\left[M_s - \sum_{j=1}^{l}\Gamma_{sj}(k)\right]}D_{sp}^T \right. \\ &\quad \left. + \overline{F_{sp}}\left[M_s - \sum_{j=1}^{l}\Gamma_{sj}(k)\right]^T\overline{E_{sp}} + H_{sp}\left[M_s - \sum_{j=1}^{l}\Gamma_{sj}(k)\right]^H G_{sp} \right\}. \end{aligned}$$

Then, we construct the calculation forms of $X_p(k+1)$ $(p = 1, \cdots, l)$ by introducing the suitable relaxation parameters and using the balanced strategies

$$X_p(k+1) = \alpha_1 X_p^{(1)}(k+1) + \alpha_2 X_p^{(2)}(k+1) + \cdots + \alpha_s X_p^{(s)}(k+1),$$

with $\alpha_i > 0$ $(i = 1, \cdots, s)$ being the weighted relaxation factors.

Remark 1. *We apply Lemma 1 to Equations* (22)–(25) *and then use the same step size factor* μ *to establish the iterative sequences in Equations* (26)–(29)*, which is conducive to deducing the convergence conditions of the proposed algorithm. It is noteworthy that we also can use different step size factors in Equations* (26)–(29) *and design a new algorithm, but it may be difficult to derive the convergence conditions of the new algorithm this way. Using different step size factors to construct iterative sequences and proposing a new algorithm will be the direction and focus in our future work.*

Based on the above discussions, we obtain the following weighted, relaxed gradient-based iterative (WRGI) algorithm (Algorithm 1) to solve the generalized coupled conjugate and transpose Sylvester matrix equations.

Algorithm 1: The weighted, relaxed gradient-based iterative (WRGI) algorithm

Step 1: Given the matrices $A_{ij}, C_{ij} \in \mathbb{C}^{m_i \times r_j}, B_{ij}, D_{ij} \in \mathbb{C}^{s_j \times n_i}, E_{ij}, G_{ij} \in \mathbb{C}^{m_i \times s_j}, F_{ij}, H_{ij} \in \mathbb{C}^{r_j \times n_i}$ and $M_i \in \mathbb{C}^{m_i \times n_i}, i = 1, 2, \cdots, s, j = 1, 2, \cdots, l$, two constants $\varepsilon > 0$ and $\mu > 0$ and the relaxation parameters $\alpha_i > 0, i = 1, 2, \cdots, s$, choose the initial matrices $X_p^{(i)}(0)$ $(i = 1, 2, \cdots, s, p = 1, 2, \cdots, l)$, and set $k = 0$;

Step 2: If $\delta_k = \dfrac{\sqrt{\sum_{i=1}^{s}\left\|M_i - \sum_{j=1}^{l}\Gamma_{ij}(k)\right\|^2}}{\sqrt{\sum_{i=1}^{s}\|M_i\|^2}} < \varepsilon$, then stop; otherwise, go to Step 3;

Algorithm 1: *Cont.*

Step 3: Compute $X_p(k+1)$ using

$$\begin{aligned} X_p^{(1)}(k+1) &= X_p^{(1)}(k) + \frac{\mu}{4}\Bigg\{ A_{1p}^H \left[M_1 - \sum_{j=1}^{l} \Gamma_{1j}(k) \right] B_{1p}^H + C_{1p}^T \overline{\left[M_1 - \sum_{j=1}^{l} \Gamma_{1j}(k) \right]} D_{1p}^T \\ &\quad + \overline{F_{1p}} \left[M_1 - \sum_{j=1}^{l} \Gamma_{1j}(k) \right]^T \overline{E_{1p}} + H_{1p} \left[M_1 - \sum_{j=1}^{l} \Gamma_{1j}(k) \right]^H G_{1p} \Bigg\} \end{aligned}$$

$$\cdots\cdots\cdots$$

$$\begin{aligned} X_p^{(s)}(k+1) &= X_p^{(s)}(k) + \frac{\mu}{4}\Bigg\{ A_{sp}^H \left[M_s - \sum_{j=1}^{l} \Gamma_{sj}(k) \right] B_{sp}^H + C_{sp}^T \overline{\left[M_s - \sum_{j=1}^{l} \Gamma_{sj}(k) \right]} D_{sp}^T \\ &\quad + \overline{F_{sp}} \left[M_s - \sum_{j=1}^{l} \Gamma_{sj}(k) \right]^T \overline{E_{sp}} + H_{sp} \left[M_s - \sum_{j=1}^{l} \Gamma_{sj}(k) \right]^H G_{sp} \Bigg\} \end{aligned}$$

$$X_p(k+1) = \alpha_1 X_p^{(1)}(k+1) + \alpha_2 X_p^{(2)}(k+1) + \cdots + \alpha_s X_p^{(s)}(k+1), p = 1, \cdots, l.$$

Step 4: Set $k = k+1$ and return to Step 2.

In the following, we will discuss the convergence properties of the WRGI algorithm by applying the properties of the real representation of a complex matrix and the Kronecker product of the matrices. For convenience, we introduce the following notations. Let

$$V = V_1 + V_2 \tag{30}$$

and

$$R = R_1 + R_2, \tag{31}$$

where

$$V_1 = \begin{bmatrix} B_{11}^{\nabla} \otimes (A_{11}^{\nabla})^T + Q_{s_1} D_{11}^{\nabla} \otimes Q_{r_1} (C_{11}^{\nabla})^T & \cdots & B_{s1}^{\nabla} \otimes (A_{s1}^{\nabla})^T + Q_{s_1} D_{s1}^{\nabla} \otimes Q_{r_1} (C_{s1}^{\nabla})^T \\ B_{12}^{\nabla} \otimes (A_{12}^{\nabla})^T + Q_{s_2} D_{12}^{\nabla} \otimes Q_{r_2} (C_{12}^{\nabla})^T & \cdots & B_{s2}^{\nabla} \otimes (A_{s2}^{\nabla})^T + Q_{s_2} D_{s2}^{\nabla} \otimes Q_{r_2} (C_{s2}^{\nabla})^T \\ \vdots & & \vdots \\ B_{1l}^{\nabla} \otimes (A_{1l}^{\nabla})^T + Q_{s_l} D_{1l}^{\nabla} \otimes Q_{r_l} (C_{1l}^{\nabla})^T & \cdots & B_{sl}^{\nabla} \otimes (A_{sl}^{\nabla})^T + Q_{s_l} D_{sl}^{\nabla} \otimes Q_{r_l} (C_{sl}^{\nabla})^T \end{bmatrix},$$

$$R_1 = \begin{bmatrix} (B_{11}^{\nabla})^T \otimes A_{11}^{\nabla} + (D_{11}^{\nabla})^T Q_{s_1} \otimes C_{11}^{\nabla} Q_{r_1} & \cdots & (B_{1l}^{\nabla})^T \otimes A_{1l}^{\nabla} + (D_{1l}^{\nabla})^T Q_{s_l} \otimes C_{1l}^{\nabla} Q_{r_l} \\ (B_{21}^{\nabla})^T \otimes A_{21}^{\nabla} + (D_{21}^{\nabla})^T Q_{s_1} \otimes C_{21}^{\nabla} Q_{r_1} & \cdots & (B_{2l}^{\nabla})^T \otimes A_{2l}^{\nabla} + (D_{2l}^{\nabla})^T Q_{s_l} \otimes C_{2l}^{\nabla} Q_{r_l} \\ \vdots & & \vdots \\ (B_{s1}^{\nabla})^T \otimes A_{s1}^{\nabla} + (D_{s1}^{\nabla})^T Q_{s_1} \otimes C_{s1}^{\nabla} Q_{r_1} & \cdots & (B_{sl}^{\nabla})^T \otimes A_{sl}^{\nabla} + (D_{sl}^{\nabla})^T Q_{s_l} \otimes C_{sl}^{\nabla} Q_{r_l} \end{bmatrix},$$

$$V_2 = \begin{bmatrix} \left(Q_{s_1} (E_{11}^{\nabla})^T \otimes Q_{r_1} F_{11}^{\nabla} + (G_{11}^{\nabla})^T \otimes H_{11}^{\nabla}\right) P(2m_1, 2n_1) & \cdots & \left(Q_{s_1} (E_{s1}^{\nabla})^T \otimes Q_{r_1} F_{s1}^{\nabla} + (G_{s1}^{\nabla})^T \otimes H_{s1}^{\nabla}\right) P(2m_s, 2n_s) \\ \left(Q_{s_2} (E_{12}^{\nabla})^T \otimes Q_{r_2} F_{12}^{\nabla} + (G_{12}^{\nabla})^T \otimes H_{12}^{\nabla}\right) P(2m_1, 2n_1) & \cdots & \left(Q_{s_2} (E_{s2}^{\nabla})^T \otimes Q_{r_2} F_{s2}^{\nabla} + (G_{s2}^{\nabla})^T \otimes H_{s2}^{\nabla}\right) P(2m_s, 2n_s) \\ \vdots & & \vdots \\ \left(Q_{s_l} (E_{1l}^{\nabla})^T \otimes Q_{r_l} F_{1l}^{\nabla} + (G_{1l}^{\nabla})^T \otimes H_{1l}^{\nabla}\right) P(2m_1, 2n_1) & \cdots & \left(Q_{s_l} (E_{sl}^{\nabla})^T \otimes Q_{r_l} F_{sl}^{\nabla} + (G_{sl}^{\nabla})^T \otimes H_{sl}^{\nabla}\right) P(2m_s, 2n_s) \end{bmatrix},$$

$$R_2 = \begin{bmatrix} \left((F_{11}^{\nabla})^T Q_{r_1} \otimes E_{11}^{\nabla} Q_{s_1} + (H_{11}^{\nabla})^T \otimes G_{11}^{\nabla}\right) P(2r_1, 2s_1) & \cdots & \left((F_{1l}^{\nabla})^T Q_{r_l} \otimes E_{1l}^{\nabla} Q_{s_l} + (H_{1l}^{\nabla})^T \otimes G_{1l}^{\nabla}\right) P(2r_l, 2s_l) \\ \left((F_{21}^{\nabla})^T Q_{r_1} \otimes E_{21}^{\nabla} Q_{s_1} + (H_{21}^{\nabla})^T \otimes G_{21}^{\nabla}\right) P(2r_1, 2s_1) & \cdots & \left((F_{2l}^{\nabla})^T Q_{r_l} \otimes E_{2l}^{\nabla} Q_{s_l} + (H_{2l}^{\nabla})^T \otimes G_{2l}^{\nabla}\right) P(2r_l, 2s_l) \\ \vdots & & \vdots \\ \left((F_{s1}^{\nabla})^T Q_{r_1} \otimes E_{s1}^{\nabla} Q_{s_1} + (H_{s1}^{\nabla})^T \otimes G_{s1}^{\nabla}\right) P(2r_1, 2s_1) & \cdots & \left((F_{sl}^{\nabla})^T Q_{r_l} \otimes E_{sl}^{\nabla} Q_{s_l} + (H_{sl}^{\nabla})^T \otimes G_{sl}^{\nabla}\right) P(2r_l, 2s_l) \end{bmatrix}.$$

Theorem 1. *The matrix in Equation (5) has a unique solution if and only if* $\operatorname{rank}(R) = \sum_{p=1}^{l} 4s_p r_p$ *(that is, R is of a full-column rank) and* $\operatorname{rank}(R) = \operatorname{rank}((R, f))$*. The unique solution is given by*

$$x = \left(R^T R\right)^{-1} R^T f, \tag{32}$$

and the corresponding homogeneous matrix equations

$$\sum_{j=1}^{l} \left(A_{ij} X_j B_{ij} + C_{ij} \overline{X_j} D_{ij} + E_{ij} X_j^T F_{ij} + G_{ij} X_j^H H_{ij}\right) = 0, i = 1, 2, \cdots, s$$

have a unique solution $X^* = (X_1^*, X_2^*, \cdots, X_l^*) = (0, 0, \cdots, 0)$*, where*

$$x = \begin{bmatrix} \operatorname{vec}(X_1^{\nabla}) \\ \operatorname{vec}(X_2^{\nabla}) \\ \vdots \\ \operatorname{vec}(X_l^{\nabla}) \end{bmatrix}, f = \begin{bmatrix} \operatorname{vec}(M_1^{\nabla}) \\ \operatorname{vec}(M_2^{\nabla}) \\ \vdots \\ \operatorname{vec}(M_s^{\nabla}) \end{bmatrix}$$

and R is defined by Equation (31).

Proof. Applying the real representation of the complex matrix to Equation (5) leads to

$$\sum_{j=1}^{l} \left[A_{ij}^{\nabla} X_j^{\nabla} B_{ij}^{\nabla} + C_{ij}^{\nabla} Q_{r_j} X_j^{\nabla} Q_{s_j} D_{ij}^{\nabla} + E_{ij}^{\nabla} Q_{s_j} \left(X_j^{\nabla}\right)^T Q_{r_j} F_{ij}^{\nabla} + G_{ij}^{\nabla} \left(X_j^{\nabla}\right)^T H_{ij}^{\nabla}\right] = M_i^{\nabla}, i = 1, 2, \cdots, s.$$

By using the Kronecker products of the matrices and the vector-stretching operator in the above equations, we have

$$\sum_{j=1}^{l} \left[\left(B_{ij}^{\nabla}\right)^T \otimes A_{ij}^{\nabla} + \left(D_{ij}^{\nabla}\right)^T Q_{s_j} \otimes C_{ij}^{\nabla} Q_{r_j} + \left(\left(F_{ij}^{\nabla}\right)^T Q_{r_j} \otimes E_{ij}^{\nabla} Q_{s_j} + \left(H_{ij}^{\nabla}\right)^T \otimes G_{ij}^{\nabla}\right) P(2r_j, 2s_j)\right] \operatorname{vec}\left(X_j^{\nabla}\right) = \operatorname{vec}\left(M_i^{\nabla}\right),$$

where $i = 1, 2, \cdots, s$, which can be equivalently transformed into the following linear system:

$$Rx = f. \tag{33}$$

Therefore, Equation (33) has a unique solution if and only if R has a full-column rank and $\operatorname{rank}(R) = \operatorname{rank}((R, f))$. In this case, the unique solution in Equation (32) for the matrix in Equation (5) can be obtained. The conclusions follow immediately. □

Based on the properties of the matrix norms and Algorithm 1, we establish the sufficient condition for the convergence of the proposed WRGI algorithm in the following theorem:

Theorem 2. *Suppose that the matrix in Equation (5) has a unique solution* $X^* = \left(X_1^*, X_2^*, \cdots, X_l^*\right)$. *Then, the iterative sequences* $\{X_p(k)\}$ $(p = 1, \cdots, l)$ *generated by Algorithm 1 converge to* X^* *for any initial matrices* $X_p(0)$ $(p = 1, \cdots, l)$ *if* μ *satisfies*

$$0 < \mu < \min_{1 \le i \le s} \frac{8}{\alpha_i \|V\|_2^2}.$$

Proof. Define the error matrices

$$\tilde{X}_p(k) = X_p(k) - X_p^*, \tilde{X}_p^{(i)}(k) = X_p^{(i)}(k) - X_p^*, i = 1, \cdots, s, p = 1, \cdots, l,$$

and

$$\begin{aligned} \tilde{\Gamma}_{ij}(k) &= A_{ij}\tilde{X}_j(k)B_{ij} + C_{ij}\overline{\tilde{X}_j(k)}D_{ij} + E_{ij}\tilde{X}_j^T(k)F_{ij} + G_{ij}\tilde{X}_j^H(k)H_{ij}, \\ Z_i(k) &= \sum_{j=1}^{l}\left(A_{ij}\tilde{X}_j(k)B_{ij} + C_{ij}\overline{\tilde{X}_j(k)}D_{ij} + E_{ij}\tilde{X}_j^T(k)F_{ij} + G_{ij}\tilde{X}_j^H(k)H_{ij}\right), \\ (Z_i(k))^{\nabla} &= Z_i^{\nabla}(k), i = 1, 2, \cdots, s. \end{aligned} \tag{34}$$

It follows from Algorithm 1 that

$$\begin{aligned} &\tilde{X}_p^{(i)}(k+1) \\ &= X_p^{(i)}(k+1) - X_p^* \\ &= X_p^{(i)}(k) - X_p^* + \frac{\mu}{4}\left\{A_{ip}^H\left[M_i - \sum_{j=1}^{l}\Gamma_{ij}(k)\right]B_{ip}^H + C_{ip}^T\overline{\left[M_i - \sum_{j=1}^{l}\Gamma_{ij}(k)\right]}D_{ip}^T\right. \\ &\qquad \left. + \overline{F_{ip}}\left[M_i - \sum_{j=1}^{l}\Gamma_{ij}(k)\right]^T\overline{E_{ip}} + H_{ip}\left[M_i - \sum_{j=1}^{l}\Gamma_{ij}(k)\right]^H G_{ip}\right\} \\ &= \tilde{X}_p^{(i)}(k) - \frac{\mu}{4}\left\{A_{ip}^H\sum_{j=1}^{l}\tilde{\Gamma}_{ij}(k)B_{ip}^H + C_{ip}^T\overline{\sum_{j=1}^{l}\tilde{\Gamma}_{ij}(k)}D_{ip}^T\right. \\ &\qquad \left. + \overline{F_{ip}}\left[\sum_{j=1}^{l}\tilde{\Gamma}_{ij}(k)\right]^T\overline{E_{ip}} + H_{ip}\left[\sum_{j=1}^{l}\tilde{\Gamma}_{ij}(k)\right]^H G_{ip}\right\} \\ &= \tilde{X}_p^{(i)}(k) - \frac{\mu}{4}\left[A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right], \end{aligned} \tag{35}$$

and

$$\begin{aligned} \tilde{X}_p(k+1) &= X_p(k+1) - X_p^* \\ &= \alpha_1 X_p^{(1)}(k+1) + \alpha_2 X_p^{(2)}(k+1) + \cdots + \alpha_s X_p^{(s)}(k+1) - X_p^* \\ &= \sum_{i=1}^{s}\alpha_i\tilde{X}_p^{(i)}(k+1) \\ &= \sum_{i=1}^{s}\alpha_i\tilde{X}_p^{(i)}(k) - \frac{\mu}{4}\sum_{i=1}^{s}\alpha_i\left[A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right] \\ &= \tilde{X}_p(k) - \frac{\mu}{4}\sum_{i=1}^{s}\alpha_i\left[A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right]. \end{aligned} \tag{36}$$

Taking the Frobenius norm in Equation (36) and using the properties of the norm leads to

$$
\begin{aligned}
&\|\tilde{X}_p(k+1)\|^2 = \mathrm{tr}\left[\tilde{X}_p^H(k+1)\tilde{X}_p(k+1)\right] \\
&= \|\tilde{X}_p(k)\|^2 - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\tilde{X}_p^H(k)\left(A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right)\right] \\
&- \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\left(B_{ip}Z_i^H(k)A_{ip} + \overline{D_{ip}}Z_i^T(k)\overline{C_{ip}} + E_{ip}^T\overline{Z_i(k)}F_{ip}^T + G_{ip}^H Z_i(k)H_{ip}^H\right)\tilde{X}_p(k)\right] \\
&+ \frac{\mu^2}{16}\left\|\sum_{i=1}^{s}\alpha_i\left(A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right)\right\|^2 \\
&= \|\tilde{X}_p(k)\|^2 - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i B_{ip}^H\tilde{X}_p^H(k)A_{ip}^H Z_i(k)\right] - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i D_{ip}^T\tilde{X}_p^H(k)C_{ip}^T\overline{Z_i(k)}\right] - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\overline{E_{ip}}\tilde{X}_p^H(k)\overline{F_{ip}}Z_i^T(k)\right] \\
&- \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i G_{ip}\tilde{X}_p^H(k)H_{ip}Z_i^H(k)\right] - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i Z_i^H(k)A_{ip}\tilde{X}_p(k)B_{ip}\right] - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i Z_i^T(k)\overline{C_{ip}}\tilde{X}_p(k)\overline{D_{ip}}\right] \\
&- \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\overline{Z_i(k)}F_{ip}^T\tilde{X}_p(k)E_{ip}^T\right] - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i Z_i(k)H_{ip}^H\tilde{X}_p(k)G_{ip}^H\right] \\
&+ \frac{\mu^2}{16}\left\|\sum_{i=1}^{s}\alpha_i\left(A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right)\right\|^2.
\end{aligned}
\tag{37}
$$

It is not difficult to verify that

$$
\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i D_{ip}^T\tilde{X}_p^H(k)C_{ip}^T\overline{Z_i(k)}\right] + \mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i Z_i^T(k)\overline{C_{ip}}\tilde{X}_p(k)\overline{D_{ip}}\right]
$$

and

$$
\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\overline{E_{ip}}\tilde{X}_p^H(k)\overline{F_{ip}}Z_i^T(k)\right] + \mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\overline{Z_i(k)}F_{ip}^T\tilde{X}_p(k)E_{ip}^T\right]
$$

are real. Then, under Lemma 3, it holds that

$$
\begin{aligned}
&\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i D_{ip}^T\tilde{X}_p^H(k)C_{ip}^T\overline{Z_i(k)}\right] + \mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i Z_i^T(k)\overline{C_{ip}}\tilde{X}_p(k)\overline{D_{ip}}\right] \\
&= \mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i D_{ip}^H\tilde{X}_p^T(k)C_{ip}^H Z_i(k)\right] + \mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i Z_i^H(k)C_{ip}\overline{\tilde{X}_p(k)}D_{ip}\right],
\end{aligned}
\tag{38}
$$

and

$$
\begin{aligned}
&\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\overline{E_{ip}}\tilde{X}_p^H(k)\overline{F_{ip}}Z_i^T(k)\right] + \mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\overline{Z_i(k)}F_{ip}^T\tilde{X}_p(k)E_{ip}^T\right] \\
&= \mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i E_{ip}\tilde{X}_p^T(k)F_{ip}Z_i^H(k)\right] + \mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i Z_i(k)F_{ip}^H\overline{\tilde{X}_p(k)}E_{ip}^H\right].
\end{aligned}
\tag{39}
$$

With the relations in Equations (38) and (39), it follows from Equation (37) that

$$
\begin{aligned}
&\left\|\tilde{X}_p(k+1)\right\|^2 \\
&= \left\|\tilde{X}_p(k)\right\|^2 - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\left(B_{ip}^H\tilde{X}_p^H(k)A_{ip}^H + D_{ip}^H\tilde{X}_p^T(k)C_{ip}^H + F_{ip}^H\overline{\tilde{X}_p(k)}E_{ip}^H + H_{ip}^H\tilde{X}_p(k)G_{ip}^H\right)Z_i(k)\right] \\
&\quad - \frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i Z_i^H(k)\left(A_{ip}\tilde{X}_p(k)B_{ip} + C_{ip}\overline{\tilde{X}_p(k)}D_{ip} + E_{ip}\tilde{X}_p^T(k)F_{ip} + G_{ip}\tilde{X}_p^H(k)H_{ip}\right)\right] \\
&\quad + \frac{\mu^2}{16}\left\|\sum_{i=1}^{s}\alpha_i\left(A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right)\right\|^2.
\end{aligned} \tag{40}
$$

By making use of the relation in Equation (12), as well as Lemmas 2 and 4, we have

$$
\begin{aligned}
&\sum_{p=1}^{l}\left\|\sum_{i=1}^{s}\alpha_i\left(A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right)\right\|^2 \\
&= \frac{1}{2}\sum_{p=1}^{l}\left\|\sum_{i=1}^{s}\alpha_i\left(A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right)^{\nabla}\right\|^2 \\
&= \frac{1}{2}\sum_{p=1}^{l}\left\|\sum_{i=1}^{s}\alpha_i\left[\left(A_{ip}^H\right)^{\nabla}Z_i^{\nabla}(k)\left(B_{ip}^H\right)^{\nabla} + \left(C_{ip}^T\right)^{\nabla}\left(\overline{Z_i(k)}\right)^{\nabla}\left(D_{ip}^T\right)^{\nabla} + (\overline{F_{ip}})^{\nabla}\left(Z_i^T(k)\right)^{\nabla}(\overline{E_{ip}})^{\nabla} + H_{ip}^{\nabla}\left(Z_i^H(k)\right)^{\nabla}G_{ip}^{\nabla}\right]\right\|^2 \\
&= \frac{1}{2}\sum_{p=1}^{l}\left\|\sum_{i=1}^{s}\alpha_i\left[\left(A_{ip}^{\nabla}\right)^T Z_i^{\nabla}(k)\left(B_{ip}^{\nabla}\right)^T + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T Q_{m_i}Q_{m_i}Z_i^{\nabla}(k)Q_{n_i}Q_{n_i}\left(D_{ip}^{\nabla}\right)^T Q_{s_p}\right.\right. \\
&\qquad \left.\left. + Q_{r_p}F_{ip}^{\nabla}Q_{n_i}Q_{n_i}\left(Z_i^{\nabla}(k)\right)^T Q_{m_i}Q_{m_i}E_{ip}^{\nabla}Q_{s_p} + H_{ip}^{\nabla}\left(Z_i^{\nabla}(k)\right)^T G_{ip}^{\nabla}\right]\right\|^2 \\
&= \frac{1}{2}\sum_{p=1}^{l}\left\|\sum_{i=1}^{s}\alpha_i\left[\left(A_{ip}^{\nabla}\right)^T Z_i^{\nabla}(k)\left(B_{ip}^{\nabla}\right)^T + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T Z_i^{\nabla}(k)\left(D_{ip}^{\nabla}\right)^T Q_{s_p} + Q_{r_p}F_{ip}^{\nabla}\left(Z_i^{\nabla}(k)\right)^T E_{ip}^{\nabla}Q_{s_p} + H_{ip}^{\nabla}\left(Z_i^{\nabla}(k)\right)^T G_{ip}^{\nabla}\right]\right\|^2 \\
&= \frac{1}{2}\sum_{p=1}^{l}\left\|\sum_{i=1}^{s}\left\{B_{ip}^{\nabla}\otimes\left(A_{ip}^{\nabla}\right)^T + Q_{s_p}D_{ip}^{\nabla}\otimes Q_{r_p}\left(C_{ip}^{\nabla}\right)^T + \left[Q_{s_p}\left(E_{ip}^{\nabla}\right)^T\otimes Q_{r_p}F_{ip}^{\nabla} + \left(G_{ip}^{\nabla}\right)^T\otimes H_{ip}^{\nabla}\right]P(2m_i,2n_i)\right\}\mathrm{vec}\left[(\alpha_i Z_i(k))^{\nabla}\right]\right\|_2^2 \\
&= \frac{1}{2}\left\|\begin{array}{c}
\sum\limits_{i=1}^{s}\left\{B_{i1}^{\nabla}\otimes(A_{i1}^{\nabla})^T + Q_{s_1}D_{i1}^{\nabla}\otimes Q_{r_1}(C_{i1}^{\nabla})^T + \left[Q_{s_1}(E_{i1}^{\nabla})^T\otimes Q_{r_1}F_{i1}^{\nabla} + (G_{i1}^{\nabla})^T\otimes H_{i1}^{\nabla}\right]P(2m_i,2n_i)\right\}\mathrm{vec}\left[(\alpha_i Z_i(k))^{\nabla}\right] \\
\sum\limits_{i=1}^{s}\left\{B_{i2}^{\nabla}\otimes(A_{i2}^{\nabla})^T + Q_{s_2}D_{i2}^{\nabla}\otimes Q_{r_2}(C_{i2}^{\nabla})^T + \left[Q_{s_2}(E_{i2}^{\nabla})^T\otimes Q_{r_2}F_{i2}^{\nabla} + (G_{i2}^{\nabla})^T\otimes H_{i2}^{\nabla}\right]P(2m_i,2n_i)\right\}\mathrm{vec}\left[(\alpha_i Z_i(k))^{\nabla}\right] \\
\vdots \\
\sum\limits_{i=1}^{s}\left\{B_{il}^{\nabla}\otimes(A_{il}^{\nabla})^T + Q_{s_l}D_{il}^{\nabla}\otimes Q_{r_l}(C_{il}^{\nabla})^T + \left[Q_{s_l}(E_{il}^{\nabla})^T\otimes Q_{r_l}F_{il}^{\nabla} + (G_{il}^{\nabla})^T\otimes H_{il}^{\nabla}\right]P(2m_i,2n_i)\right\}\mathrm{vec}\left[(\alpha_i Z_i(k))^{\nabla}\right]
\end{array}\right\|_2^2 \\
&= \frac{1}{2}\|Vb\|_2^2,
\end{aligned} \tag{41}
$$

where V is defined by Equation (30) and

$$
b = \left[\left[\mathrm{vec}\left((\alpha_1 Z_1(k))^{\nabla}\right)\right]^T, \left[\mathrm{vec}\left((\alpha_2 Z_2(k))^{\nabla}\right)\right]^T, \cdots, \left[\mathrm{vec}\left((\alpha_s Z_s(k))^{\nabla}\right)\right]^T\right]^T.
$$

By applying the properties of the 2-norms of the matrices and vectors, we obtain that

$$
\sum_{p=1}^{l}\left\|\sum_{i=1}^{s}\alpha_i\left(A_{ip}^H Z_i(k)B_{ip}^H + C_{ip}^T\overline{Z_i(k)}D_{ip}^T + \overline{F_{ip}}Z_i^T(k)\overline{E_{ip}} + H_{ip}Z_i^H(k)G_{ip}\right)\right\|^2 = \frac{1}{2}\|Vb\|_2^2 \le \frac{1}{2}\|V\|_2^2\|b\|_2^2. \tag{42}
$$

That aside, in light of the relation in Equation (12), we deduce that

$$
\|b\|_2^2 = \sum_{i=1}^{s}\left\|(\alpha_i Z_i(k))^{\nabla}\right\|^2 = 2\sum_{i=1}^{s}\alpha_i^2\|Z_i(k)\|^2. \tag{43}
$$

Substituting Equation (43) into Equation (42) leads to

$$\sum_{p=1}^{l}\left\|\sum_{i=1}^{s}\alpha_i\left(A_{ip}^{H}Z_i(k)B_{ip}^{H}+C_{ip}^{T}\overline{Z_i(k)}D_{ip}^{T}+\overline{F_{ip}}Z_i^{T}(k)\overline{E_{ip}}+H_{ip}Z_i^{H}(k)G_{ip}\right)\right\|^2\leq\|V\|_2^2\sum_{i=1}^{s}\alpha_i^2\|Z_i(k)\|^2. \tag{44}$$

By combining Equations (40) and (44) and changing the order of addition, we derive

$$\begin{aligned}
&\sum_{p=1}^{l}\|\tilde{X}_p(k+1)\|^2\\
&\leq\sum_{p=1}^{l}\|\tilde{X}_p(k)\|^2-\frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_i\left(\sum_{p=1}^{l}\left(B_{ip}^{H}\tilde{X}_p^{H}(k)A_{ip}^{H}+D_{ip}^{H}\tilde{X}_p^{T}(k)C_{ip}^{H}+F_{ip}^{H}\overline{\tilde{X}_p(k)}E_{ip}^{H}+H_{ip}^{H}\tilde{X}_p(k)G_{ip}^{H}\right)\right)Z_i(k)\right]\\
&\quad-\frac{\mu}{4}\mathrm{tr}\left[Z_i^{H}(k)\sum_{i=1}^{s}\alpha_i\left(\sum_{p=1}^{l}\left(A_{ip}\tilde{X}_p(k)B_{ip}+C_{ip}\overline{\tilde{X}_p(k)}D_{ip}+E_{ip}\tilde{X}_p^{T}(k)F_{ip}+G_{ip}\tilde{X}_p^{H}(k)H_{ip}\right)\right)\right]+\frac{\mu^2}{16}\|V\|_2^2\sum_{i=1}^{s}\alpha_i^2\|Z_i(k)\|^2\\
&=\sum_{p=1}^{l}\|\tilde{X}_p(k)\|^2-\frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_iZ_i^{H}(k)Z_i(k)\right]-\frac{\mu}{4}\mathrm{tr}\left[\sum_{i=1}^{s}\alpha_iZ_i^{H}(k)Z_i(k)\right]+\frac{\mu^2}{16}\|V\|_2^2\sum_{i=1}^{s}\alpha_i^2\|Z_i(k)\|^2\\
&=\sum_{p=1}^{l}\|\tilde{X}_p(k)\|^2-\frac{\mu}{4}\sum_{i=1}^{s}\alpha_i\mathrm{tr}\left[Z_i^{H}(k)Z_i(k)\right]-\frac{\mu}{4}\sum_{i=1}^{s}\alpha_i\mathrm{tr}\left[Z_i^{H}(k)Z_i(k)\right]+\frac{\mu^2}{16}\|V\|_2^2\sum_{i=1}^{s}\alpha_i^2\|Z_i(k)\|^2\\
&=\sum_{p=1}^{l}\|\tilde{X}_p(k)\|^2-\frac{\mu}{2}\sum_{i=1}^{s}\alpha_i\|Z_i(k)\|^2+\frac{\mu^2}{16}\|V\|_2^2\sum_{i=1}^{s}\alpha_i^2\|Z_i(k)\|^2\\
&=\sum_{p=1}^{l}\|\tilde{X}_p(k)\|^2-\frac{\mu}{2}\sum_{i=1}^{s}\alpha_i\left(1-\frac{\mu}{8}\alpha_i\|V\|_2^2\right)\|Z_i(k)\|^2\\
&\leq\sum_{p=1}^{l}\|\tilde{X}_p(k-1)\|^2-\frac{\mu}{2}\sum_{i=1}^{s}\alpha_i\left(1-\frac{\mu}{8}\alpha_i\|V\|_2^2\right)\left(\|Z_i(k)\|^2+\|Z_i(k-1)\|^2\right)\\
&\leq\sum_{p=1}^{l}\|\tilde{X}_p(0)\|^2-\frac{\mu}{2}\sum_{i=1}^{s}\alpha_i\left(1-\frac{\mu}{8}\alpha_i\|V\|_2^2\right)\sum_{\omega=0}^{k}\|Z_i(\omega)\|^2.
\end{aligned} \tag{45}$$

If the parameter μ is chosen to satisfy

$$0<\mu<\min_{1\leq i\leq s}\frac{8}{\alpha_i\|V\|_2^2},$$

then

$$0<\frac{\mu}{2}\sum_{i=1}^{s}\alpha_i\left(1-\frac{\mu}{8}\alpha_i\|V\|_2^2\right)\sum_{\omega=0}^{\infty}\|Z_i(\omega)\|^2\leq\sum_{p=1}^{l}\|\tilde{X}_p(0)\|^2,$$

and therefore

$$\sum_{\omega=0}^{\infty}\|Z_i(\omega)\|^2<+\infty,\ i=1,2,\cdots,s,$$

According to the convergence theorem of series, we have $\lim_{\omega\to+\infty}Z_i(\omega)=0\ (i=1,2,\cdots,s)$. Having in mind that the matrix in Equation (5) has a unique solution, then it follows from the definition of $Z_i(k)$ in Equation (34) that

$$\lim_{k\to+\infty}X_j(k)=X_j^*,\ j=1,2,\cdots,l.$$

The proof is completed. □

In the sequel, by applying the properties of the real representation of a complex matrix and the vector-stretching operator, we study the necessary and sufficient condition for the convergence of the WRGI algorithm. This result can be stated as in Theorem 3:

Theorem 3. *Suppose that the matrix in Equation (5) has a unique solution $X^* = \left(X_1^*, X_2^*, \cdots, X_l^*\right)$. Then, the iterative sequences $\{X_p(k)\}$ $(p = 1, \cdots, l)$ generated by Algorithm 1 converge to X^* for any initial matrices $X_p(0)$ $(p = 1, \cdots, l)$ if and only if*

$$0 < \mu < \frac{8}{\left\| T^{\frac{1}{2}} R \right\|_2^2}. \tag{46}$$

Proof. For convenience, we define

$$\left(\tilde{X}_p(k)\right)^{\nabla} = \tilde{X}_p^{\nabla}(k), p = 1, 2, \cdots, l.$$

From Equation (36) and the definition in Equation (34) for $Z_i(k)$ in Theorem 2, it can be obtained that

$$\begin{aligned}
&\tilde{X}_p(k+1) \\
&= \tilde{X}_p(k) - \frac{\mu}{4} \sum_{i=1}^{s} \alpha_i \Bigg\{ A_{ip}^H \left[\sum_{j=1}^{l} \left(A_{ij} \tilde{X}_j(k) B_{ij} + C_{ij} \overline{\tilde{X}_j(k)} D_{ij} + E_{ij} \tilde{X}_j^T(k) F_{ij} + G_{ij} \tilde{X}_j^H(k) H_{ij} \right) \right] B_{ip}^H \\
&\quad + C_{ip}^T \overline{\left[\sum_{j=1}^{l} \left(A_{ij} \tilde{X}_j(k) B_{ij} + C_{ij} \overline{\tilde{X}_j(k)} D_{ij} + E_{ij} \tilde{X}_j^T(k) F_{ij} + G_{ij} \tilde{X}_j^H(k) H_{ij} \right) \right]} D_{ip}^T \\
&\quad + \overline{F_{ip}} \left[\sum_{j=1}^{l} \left(A_{ij} \tilde{X}_j(k) B_{ij} + C_{ij} \overline{\tilde{X}_j(k)} D_{ij} + E_{ij} \tilde{X}_j^T(k) F_{ij} + G_{ij} \tilde{X}_j^H(k) H_{ij} \right) \right]^T \overline{E_{ip}} \\
&\quad + H_{ip} \left[\sum_{j=1}^{l} \left(A_{ij} \tilde{X}_j(k) B_{ij} + C_{ij} \overline{\tilde{X}_j(k)} D_{ij} + E_{ij} \tilde{X}_j^T(k) F_{ij} + G_{ij} \tilde{X}_j^H(k) H_{ij} \right) \right]^H G_{ip} \Bigg\} \\
&= \tilde{X}_p(k) - \frac{\mu}{4} \sum_{i=1}^{s} \alpha_i \Bigg\{ A_{ip}^H \left[\sum_{j=1}^{l} \left(A_{ij} \tilde{X}_j(k) B_{ij} + C_{ij} \overline{\tilde{X}_j(k)} D_{ij} + E_{ij} \tilde{X}_j^T(k) F_{ij} + G_{ij} \tilde{X}_j^H(k) H_{ij} \right) \right] B_{ip}^H \\
&\quad + C_{ip}^T \left[\sum_{j=1}^{l} \left(\overline{A_{ij}}\, \overline{\tilde{X}_j(k)}\, \overline{B_{ij}} + \overline{C_{ij}} \tilde{X}_j(k) \overline{D_{ij}} + \overline{E_{ij}} \tilde{X}_j^H(k) \overline{F_{ij}} + \overline{G_{ij}} \tilde{X}_j^T(k) \overline{H_{ij}} \right) \right] D_{ip}^T \\
&\quad + \overline{F_{ip}} \left[\sum_{j=1}^{l} \left(B_{ij}^T \tilde{X}_j^T(k) A_{ij}^T + D_{ij}^T \tilde{X}_j^H(k) C_{ij}^T + F_{ij}^T \tilde{X}_j(k) E_{ij}^T + H_{ij}^T \overline{\tilde{X}_j(k)} G_{ij}^T \right) \right] \overline{E_{ip}} \\
&\quad + H_{ip} \left[\sum_{j=1}^{l} \left(B_{ij}^H \tilde{X}_j^H(k) A_{ij}^H + D_{ij}^H \tilde{X}_j^T(k) C_{ij}^H + F_{ij}^H \overline{\tilde{X}_j(k)} E_{ij}^H + H_{ij}^H \tilde{X}_j(k) G_{ij}^H \right) \right] G_{ip} \Bigg\} \\
&= \tilde{X}_p(k) - \frac{\mu}{4} \sum_{i=1}^{s} \sum_{j=1}^{l} \alpha_i \Big[A_{ip}^H A_{ij} \tilde{X}_j(k) B_{ij} B_{ip}^H + A_{ip}^H C_{ij} \overline{\tilde{X}_j(k)} D_{ij} B_{ip}^H + A_{ip}^H E_{ij} \tilde{X}_j^T(k) F_{ij} B_{ip}^H + A_{ip}^H G_{ij} \tilde{X}_j^H(k) H_{ij} B_{ip}^H \\
&\quad + C_{ip}^T \overline{A_{ij}}\, \overline{\tilde{X}_j(k)}\, \overline{B_{ij}} D_{ip}^T + C_{ip}^T \overline{C_{ij}} \tilde{X}_j(k) \overline{D_{ij}} D_{ip}^T + C_{ip}^T \overline{E_{ij}} \tilde{X}_j^H(k) \overline{F_{ij}} D_{ip}^T + C_{ip}^T \overline{G_{ij}} \tilde{X}_j^T(k) \overline{H_{ij}} D_{ip}^T \\
&\quad + \overline{F_{ip}} B_{ij}^T \tilde{X}_j^T(k) A_{ij}^T \overline{E_{ip}} + \overline{F_{ip}} D_{ij}^T \tilde{X}_j^H(k) C_{ij}^T \overline{E_{ip}} + \overline{F_{ip}} F_{ij}^T \tilde{X}_j(k) E_{ij}^T \overline{E_{ip}} + \overline{F_{ip}} H_{ij}^T \overline{\tilde{X}_j(k)} G_{ij}^T \overline{E_{ip}} \\
&\quad + H_{ip} B_{ij}^H \tilde{X}_j^H(k) A_{ij}^H G_{ip} + H_{ip} D_{ij}^H \tilde{X}_j^T(k) C_{ij}^H G_{ip} + H_{ip} F_{ij}^H \overline{\tilde{X}_j(k)} E_{ij}^H G_{ip} + H_{ip} H_{ij}^H \tilde{X}_j(k) G_{ij}^H G_{ip} \Big].
\end{aligned} \tag{47}$$

Then, by combining the real representation with Equation (47) and applying Lemma 4, we obtain

$$
\begin{aligned}
&\tilde{X}_p^{\nabla}(k+1)\\
&= \tilde{X}_p^{\nabla}(k) - \frac{\mu}{4}\sum_{i=1}^{s}\sum_{j=1}^{l}\alpha_i\Big[\left(A_{ip}^{\nabla}\right)^T A_{ij}^{\nabla}\tilde{X}_j^{\nabla}(k)B_{ij}^{\nabla}\left(B_{ip}^{\nabla}\right)^T + \left(A_{ip}^{\nabla}\right)^T C_{ij}^{\nabla}Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}D_{ij}^{\nabla}\left(B_{ip}^{\nabla}\right)^T\\
&\quad + \left(A_{ip}^{\nabla}\right)^T E_{ij}^{\nabla}Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}F_{ij}^{\nabla}\left(B_{ip}^{\nabla}\right)^T + \left(A_{ip}^{\nabla}\right)^T G_{ij}^{\nabla}\left(\tilde{X}_j^{\nabla}(k)\right)^T H_{ij}^{\nabla}\left(B_{ip}^{\nabla}\right)^T\\
&\quad + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T Q_{m_i}Q_{m_i}A_{ij}^{\nabla}Q_{r_j}Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}Q_{s_j}B_{ij}^{\nabla}Q_{n_i}Q_{n_i}\left(D_{ip}^{\nabla}\right)^T Q_{s_p}\\
&\quad + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T Q_{m_i}Q_{m_i}C_{ij}^{\nabla}Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}D_{ij}^{\nabla}Q_{n_i}Q_{n_i}\left(D_{ip}^{\nabla}\right)^T Q_{s_p}\\
&\quad + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T Q_{m_i}Q_{m_i}E_{ij}^{\nabla}Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}F_{ij}^{\nabla}Q_{n_i}Q_{n_i}\left(D_{ip}^{\nabla}\right)^T Q_{s_p}\\
&\quad + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T Q_{m_i}Q_{m_i}G_{ij}^{\nabla}Q_{s_j}Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}Q_{r_j}H_{ij}^{\nabla}Q_{n_i}Q_{n_i}\left(D_{ip}^{\nabla}\right)^T Q_{s_p}\\
&\quad + Q_{r_p}F_{ip}^{\nabla}Q_{n_i}Q_{n_i}\left(B_{ij}^{\nabla}\right)^T Q_{s_j}Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}Q_{r_j}\left(A_{ij}^{\nabla}\right)^T Q_{m_i}Q_{m_i}E_{ip}^{\nabla}Q_{s_p}\\
&\quad + Q_{r_p}F_{ip}^{\nabla}Q_{n_i}Q_{n_i}\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}\left(C_{ij}^{\nabla}\right)^T Q_{m_i}Q_{m_i}E_{ip}^{\nabla}Q_{s_p}\\
&\quad + Q_{r_p}F_{ip}^{\nabla}Q_{n_i}Q_{n_i}\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}\left(E_{ij}^{\nabla}\right)^T Q_{m_i}Q_{m_i}E_{ip}^{\nabla}Q_{s_p}\\
&\quad + Q_{r_p}F_{ip}^{\nabla}Q_{n_i}Q_{n_i}\left(H_{ij}^{\nabla}\right)^T Q_{r_j}Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}Q_{s_j}\left(G_{ij}^{\nabla}\right)^T Q_{m_i}Q_{m_i}E_{ip}^{\nabla}Q_{s_p}\\
&\quad + H_{ip}^{\nabla}\left(B_{ij}^{\nabla}\right)^T\left(\tilde{X}_j^{\nabla}(k)\right)^T\left(A_{ij}^{\nabla}\right)^T G_{ip}^{\nabla} + H_{ip}^{\nabla}\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}\left(C_{ij}^{\nabla}\right)^T G_{ip}^{\nabla}\\
&\quad + H_{ip}^{\nabla}\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}\left(E_{ij}^{\nabla}\right)^T G_{ip}^{\nabla} + H_{ip}^{\nabla}\left(H_{ij}^{\nabla}\right)^T\tilde{X}_j^{\nabla}(k)\left(G_{ij}^{\nabla}\right)^T G_{ip}^{\nabla}\Big]\\
&= \tilde{X}_p^{\nabla}(k) - \frac{\mu}{4}\sum_{i=1}^{s}\sum_{j=1}^{l}\alpha_i\Big[\left(A_{ip}^{\nabla}\right)^T A_{ij}^{\nabla}\tilde{X}_j^{\nabla}(k)B_{ij}^{\nabla}\left(B_{ip}^{\nabla}\right)^T + \left(A_{ip}^{\nabla}\right)^T C_{ij}^{\nabla}Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}D_{ij}^{\nabla}\left(B_{ip}^{\nabla}\right)^T\\
&\quad + \left(A_{ip}^{\nabla}\right)^T E_{ij}^{\nabla}Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}F_{ij}^{\nabla}\left(B_{ip}^{\nabla}\right)^T + \left(A_{ip}^{\nabla}\right)^T G_{ij}^{\nabla}\left(\tilde{X}_j^{\nabla}(k)\right)^T H_{ij}^{\nabla}\left(B_{ip}^{\nabla}\right)^T\\
&\quad + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T A_{ij}^{\nabla}\tilde{X}_j^{\nabla}(k)B_{ij}^{\nabla}\left(D_{ip}^{\nabla}\right)^T Q_{s_p} + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T C_{ij}^{\nabla}Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}D_{ij}^{\nabla}\left(D_{ip}^{\nabla}\right)^T Q_{s_p}\\
&\quad + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T E_{ij}^{\nabla}Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}F_{ij}^{\nabla}\left(D_{ip}^{\nabla}\right)^T Q_{s_p} + Q_{r_p}\left(C_{ip}^{\nabla}\right)^T G_{ij}^{\nabla}\left(\tilde{X}_j^{\nabla}(k)\right)^T H_{ij}^{\nabla}\left(D_{ip}^{\nabla}\right)^T Q_{s_p}\\
&\quad + Q_{r_p}F_{ip}^{\nabla}\left(B_{ij}^{\nabla}\right)^T\left(\tilde{X}_j^{\nabla}(k)\right)^T\left(A_{ij}^{\nabla}\right)^T E_{ip}^{\nabla}Q_{s_p} + Q_{r_p}F_{ip}^{\nabla}\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}\left(C_{ij}^{\nabla}\right)^T E_{ip}^{\nabla}Q_{r_p}\\
&\quad + Q_{r_p}F_{ip}^{\nabla}\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}\left(E_{ij}^{\nabla}\right)^T E_{ip}^{\nabla}Q_{s_p} + Q_{r_p}F_{ip}^{\nabla}\left(H_{ij}^{\nabla}\right)^T\tilde{X}_j^{\nabla}(k)\left(G_{ij}^{\nabla}\right)^T E_{ip}^{\nabla}Q_{s_p}\\
&\quad + H_{ip}^{\nabla}\left(B_{ij}^{\nabla}\right)^T\left(\tilde{X}_j^{\nabla}(k)\right)^T\left(A_{ij}^{\nabla}\right)^T G_{ip}^{\nabla} + H_{ip}^{\nabla}\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\left(\tilde{X}_j^{\nabla}(k)\right)^T Q_{r_j}\left(C_{ij}^{\nabla}\right)^T G_{ip}^{\nabla}\\
&\quad + H_{ip}^{\nabla}\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\tilde{X}_j^{\nabla}(k)Q_{s_j}\left(E_{ij}^{\nabla}\right)^T G_{ip}^{\nabla} + H_{ip}^{\nabla}\left(H_{ij}^{\nabla}\right)^T\tilde{X}_j^{\nabla}(k)\left(G_{ij}^{\nabla}\right)^T G_{ip}^{\nabla}\Big].
\end{aligned}
\tag{48}
$$

By taking the vector-stretching operator on both sides of Equation (48) and using Lemma 2, we have

$$
\begin{aligned}
&\mathrm{vec}\left[\tilde{X}_p^{\nabla}(k+1)\right]\\
&= \mathrm{vec}\left[\tilde{X}_p^{\nabla}(k)\right] - \frac{\mu}{4}\sum_{i=1}^{s}\alpha_i\left[B_{ip}^{\nabla}\otimes\left(A_{ip}^{\nabla}\right)^T + Q_{s_p}D_{ip}^{\nabla}\otimes Q_{r_p}\left(C_{ip}^{\nabla}\right)^T + P(2s_p,2r_p)\left(Q_{r_p}F_{ip}^{\nabla}\otimes Q_{s_p}\left(E_{ip}^{\nabla}\right)^T + H_{ip}^{\nabla}\otimes\left(G_{ip}^{\nabla}\right)^T\right)\right]\\
&\sum_{j=1}^{l}\left[\left(B_{ij}^{\nabla}\right)^T\otimes A_{ij}^{\nabla} + \left(D_{ij}^{\nabla}\right)^T Q_{s_j}\otimes C_{ij}^{\nabla}Q_{r_j} + \left(\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\otimes E_{ij}^{\nabla}Q_{s_j} + \left(H_{ij}^{\nabla}\right)^T\otimes G_{ij}^{\nabla}\right)P(2r_j,2s_j)\right]\mathrm{vec}\left[\tilde{X}_j^{\nabla}(k)\right],
\end{aligned}
\tag{49}
$$

for $p = 1, 2, \cdots, l$. This gives the following result:

$$\begin{aligned} \mathrm{vec}\left[\tilde{X}^{\nabla}(k+1)\right] &= \mathrm{vec}\left[\tilde{X}^{\nabla}(k)\right] - \frac{\mu}{4}R^T T R \mathrm{vec}\left[\tilde{X}^{\nabla}(k)\right] \\ &= \left(I - \frac{\mu}{4}R^T T R\right)\mathrm{vec}\left[\tilde{X}^{\nabla}(k)\right] \\ &= \left(I - \frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\mathrm{vec}\left[\tilde{X}^{\nabla}(k)\right], \end{aligned} \tag{50}$$

where

$$\mathrm{vec}\left[\tilde{X}^{\nabla}(k)\right] = \left[\left[\mathrm{vec}\left(\tilde{X}_1^{\nabla}(k)\right)\right]^T, \left[\mathrm{vec}\left(\tilde{X}_2^{\nabla}(k)\right)\right]^T, \cdots, \left[\mathrm{vec}\left(\tilde{X}_l^{\nabla}(k)\right)\right]^T\right]^T,$$

$$T = \begin{bmatrix} \alpha_1 I_{4m_1n_1} & 0 & \cdots & 0 \\ 0 & \alpha_2 I_{4m_2n_2} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & \alpha_s I_{4m_sn_s} \end{bmatrix}$$

and R is defined by Equation (31).

Equation (50) implies that the sufficient and necessary condition for the convergent of Algorithm 1 (the WRGI algorithm) is

$$\rho\left(I - \frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right) < 1.$$

Since $R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R$ is a symmetric matrix, it holds that

$$\begin{aligned} \lambda\left(I - \frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right) &= \left\{1 - \frac{\mu}{4}\lambda_q\left(R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right), q = 1, 2, \cdots, \eta\right\} \\ &= \left\{1 - \frac{\mu}{4}\sigma_q^2\left(T^{\frac{1}{2}} R\right), q = 1, 2, \cdots, \eta\right\}, \end{aligned} \tag{51}$$

where $\eta = \mathrm{rank}(R^T R) = \mathrm{rank}(R) = \mathrm{rank}\left(T^{\frac{1}{2}} R\right)$. Then, $\rho\left(I - \frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right) < 1$ is equivalent to

$$\left|1 - \frac{\mu}{4}\sigma_q^2\left(T^{\frac{1}{2}} R\right)\right| < 1, q = 1, 2, \cdots, \eta. \tag{52}$$

It follows from Equation (52) that

$$-1 < 1 - \frac{\mu}{4}\sigma_q^2\left(T^{\frac{1}{2}} R\right) < 1, q = 1, 2, \cdots, \eta.$$

After simple computations, we derive

$$0 < \mu < \frac{8}{\left\|T^{\frac{1}{2}} R\right\|_2^2},$$

which completes the proof. □

Remark 2. *When the relaxation parameters α_i of the WRGI algorithm are $\alpha_i = \frac{1}{s}, i = 1, 2, \cdots, s$, then the condition in Equation (46) reduces to the following condition:*

$$0 < \mu < \frac{8s}{\|R\|_2^2}.$$

Based on Theorem 3, we will study the optimal step size μ and the corresponding optimal convergence factor of the WRGI algorithm in the following theorem:

Theorem 4. *Assume that the conditions of Theorem 2 are valid. Then, it holds that*

$$\|X(k)-X^*\| \le \rho^k\left(I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\|X(0)-X^*\|, \tag{53}$$

where

$$X(k)=(X_1(k),X_2(k),\cdots,X_l(k)).$$

In addition, the optimal convergence factor μ_{opt} is

$$\mu_{opt}=\frac{8}{\sigma^2_{\max}\left(T^{\frac{1}{2}}R\right)+\sigma^2_{\min}\left(T^{\frac{1}{2}}R\right)}. \tag{54}$$

Under this situation, the convergence rate is maximized, and we have

$$\|X(k)-X^*\| \le \left(\frac{\mathrm{cond}^2\left(T^{\frac{1}{2}}R\right)-1}{\mathrm{cond}^2\left(T^{\frac{1}{2}}R\right)+1}\right)^k\|X(0)-X^*\|. \tag{55}$$

Proof. Due to the fact that $I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R$ is symmetric, it follows that

$$\left\|I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right\|_2=\rho\left(I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right). \tag{56}$$

In light of Equations (50) and (56), one has

$$\begin{aligned}\left\|\tilde{X}^\nabla(k+1)\right\|=\left\|\mathrm{vec}\left[\tilde{X}^\nabla(k+1)\right]\right\| &= \left\|\mathrm{vec}\left[\tilde{X}^\nabla(k+1)\right]\right\|_2\\ &= \left\|\left(I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\mathrm{vec}\left[\tilde{X}^\nabla(k)\right]\right\|_2\\ &\le \left\|I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right\|_2\left\|\mathrm{vec}\left[\tilde{X}^\nabla(k)\right]\right\|_2\\ &= \rho\left(I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\left\|\tilde{X}^\nabla(k)\right\|,\end{aligned} \tag{57}$$

where

$$\tilde{X}^\nabla(k)=\left[\tilde{X}_1^\nabla(k),\tilde{X}_2^\nabla(k),\cdots,\tilde{X}_l^\nabla(k)\right].$$

Combining the relations in Equations (12) and (57) results in

$$\|\tilde{X}(k+1)\| \quad \le \quad \rho\left(I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\|\tilde{X}(k)\|, \tag{58}$$

where

$$\tilde{X}(k)=[\tilde{X}_1(k),\tilde{X}_2(k),\cdots,\tilde{X}_l(k)].$$

Having in mind that $\tilde{X}_p(k)=X_p(k)-X_p^*$ $(p=1,2,\cdots,l)$, then it follows from Equation (58) that

$$\begin{aligned}\|X(k)-X^*\| &\le \rho\left(I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\|X(k-1)-X^*\|\\ &\le \rho^2\left(I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\|X(k-2)-X^*\|\\ &\le \rho^k\left(I-\frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\|X(0)-X^*\|.\end{aligned} \tag{59}$$

It can be seen from Equation (59) that the smaller the $\rho\left(I - \frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)$, the faster the convergence rate of the WRGI algorithm. Direct calculations give that $\rho\left(I - \frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)$ is minimized if and only if

$$1 - \frac{\mu}{4}\sigma_{\min}^2\left(T^{\frac{1}{2}} R\right) = \frac{\mu}{4}\sigma_{\max}^2\left(T^{\frac{1}{2}} R\right) - 1, \tag{60}$$

from which one can deduce that

$$\mu_{opt} = \frac{8}{\sigma_{\max}^2\left(T^{\frac{1}{2}} R\right) + \sigma_{\min}^2\left(T^{\frac{1}{2}} R\right)}.$$

If $\mu = \mu_{opt}$, then we have

$$\begin{aligned} \rho\left(I - \frac{\mu}{4}R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right) &= \max_i \left\{1 - \frac{2}{\sigma_{\max}^2\left(T^{\frac{1}{2}} R\right) + \sigma_{\min}^2\left(T^{\frac{1}{2}} R\right)} \lambda_i\left(R^T T^{\frac{1}{2}} T^{\frac{1}{2}} R\right)\right\} \\ &= \max_i \left\{1 - \frac{2}{\sigma_{\max}^2\left(T^{\frac{1}{2}} R\right) + \sigma_{\min}^2\left(T^{\frac{1}{2}} R\right)} \sigma_i^2\left(T^{\frac{1}{2}} R\right)\right\} \\ &= 1 - \frac{2\sigma_{\min}^2\left(T^{\frac{1}{2}} R\right)}{\sigma_{\max}^2\left(T^{\frac{1}{2}} R\right) + \sigma_{\min}^2\left(T^{\frac{1}{2}} R\right)} \\ &= \frac{\sigma_{\max}^2\left(T^{\frac{1}{2}} R\right) - \sigma_{\min}^2\left(T^{\frac{1}{2}} R\right)}{\sigma_{\max}^2\left(T^{\frac{1}{2}} R\right) + \sigma_{\min}^2\left(T^{\frac{1}{2}} R\right)} \\ &= \frac{\mathrm{cond}^2\left(T^{\frac{1}{2}} R\right) - 1}{\mathrm{cond}^2\left(T^{\frac{1}{2}} R\right) + 1}. \end{aligned} \tag{61}$$

Substituting Equation (61) into Equation (59) yields Equation (55). The proof is completed. □

Although the convergence conditions of the WRGI algorithm are given in Theorems 2 and 3, the Kronecker product and the real representations of the system matrices are involved in computing $\|V\|_2$ and $\left\|T^{\frac{1}{2}} R\right\|_2$. This may bring about a high dimensionality problem and lead to high computing consumption. To overcome this drawback, we will give a sufficient condition for the convergence of the proposed WRGI algorithm below.

Corollary 1. *Suppose that the matrix in Equation (5) has a unique solution $X^* = (X_1^*, X_2^*, \cdots, X_l^*)$. Then, the iterative sequences $\{X_p(k)\}$ $(p = 1, \cdots, l)$ generated by Algorithm 1 converge to X^* for the arbitrary initial matrices $X_p(0)$ $(p = 1, \cdots, l)$ if*

$$0 < \mu < \frac{2}{\sum_{i=1}^{s}\sum_{j=1}^{l} \alpha_i\left(\|B_{ij}\|_2^2\|A_{ij}\|_2^2 + \|D_{ij}\|_2^2\|C_{ij}\|_2^2 + \|F_{ij}\|_2^2\|E_{ij}\|_2^2 + \|H_{ij}\|_2^2\|G_{ij}\|_2^2\right)}.$$

Proof. According to Lemmas 4–6 and the fact that $\|E \otimes F\|_2 = \|E\|_2\|F\|_2$, we have

$$
\begin{aligned}
\left\|T^{\frac{1}{2}}R\right\|_2^2 \;\leq\;& \sum_{i=1}^{s}\sum_{j=1}^{l}\left\|\sqrt{\alpha_i}\left[\left(B_{ij}^{\nabla}\right)^T\otimes A_{ij}^{\nabla}+\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\otimes C_{ij}^{\nabla}Q_{r_j}\right]+\sqrt{\alpha_i}\left[\left(\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\otimes E_{ij}^{\nabla}Q_{s_j}+\left(H_{ij}^{\nabla}\right)^T\otimes G_{ij}^{\nabla}\right)P(2r_j,2s_j)\right]\right\|_2^2\\
\leq\;& \sum_{i=1}^{s}\sum_{j=1}^{l}\alpha_i\left(\left\|\left(B_{ij}^{\nabla}\right)^T\otimes A_{ij}^{\nabla}+\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\otimes C_{ij}^{\nabla}Q_{r_j}\right\|_2+\left\|\left(\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\otimes E_{ij}^{\nabla}Q_{s_j}+\left(H_{ij}^{\nabla}\right)^T\otimes G_{ij}^{\nabla}\right)P(2r_j,2s_j)\right\|_2\right)^2\\
\leq\;& \sum_{i=1}^{s}\sum_{j=1}^{l}2\alpha_i\left(\left\|\left(B_{ij}^{\nabla}\right)^T\otimes A_{ij}^{\nabla}+\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\otimes C_{ij}^{\nabla}Q_{r_j}\right\|_2^2+\left\|\left(\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\otimes E_{ij}^{\nabla}Q_{s_j}+\left(H_{ij}^{\nabla}\right)^T\otimes G_{ij}^{\nabla}\right)P(2r_j,2s_j)\right\|_2^2\right)\\
\leq\;& \sum_{i=1}^{s}\sum_{j=1}^{l}2\alpha_i\left[\left(\left\|\left(B_{ij}^{\nabla}\right)^T\otimes A_{ij}^{\nabla}\right\|_2+\left\|\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\otimes C_{ij}^{\nabla}Q_{r_j}\right\|_2\right)^2\right.\\
+\;& \left.\left(\left\|\left(\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\otimes E_{ij}^{\nabla}Q_{s_j}\right)P(2r_j,2s_j)\right\|_2+\left\|\left(\left(H_{ij}^{\nabla}\right)^T\otimes G_{ij}^{\nabla}\right)P(2r_j,2s_j)\right\|_2\right)^2\right] \qquad (62)\\
\leq\;& \sum_{i=1}^{s}\sum_{j=1}^{l}4\alpha_i\left[\left\|\left(B_{ij}^{\nabla}\right)^T\otimes A_{ij}^{\nabla}\right\|_2^2+\left\|\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\otimes C_{ij}^{\nabla}Q_{r_j}\right\|_2^2\right.\\
+\;& \left.\left\|\left(\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\otimes E_{ij}^{\nabla}Q_{s_j}\right)P(2r_j,2s_j)\right\|_2^2+\left\|\left(\left(H_{ij}^{\nabla}\right)^T\otimes G_{ij}^{\nabla}\right)P(2r_j,2s_j)\right\|_2^2\right]\\
=\;& \sum_{i=1}^{s}\sum_{j=1}^{l}4\alpha_i\left[\left\|\left(B_{ij}^{\nabla}\right)^T\otimes A_{ij}^{\nabla}\right\|_2^2+\left\|\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\otimes C_{ij}^{\nabla}Q_{r_j}\right\|_2^2+\left\|\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\otimes E_{ij}^{\nabla}Q_{s_j}\right\|_2^2+\left\|\left(H_{ij}^{\nabla}\right)^T\otimes G_{ij}^{\nabla}\right\|_2^2\right]\\
=\;& \sum_{i=1}^{s}\sum_{j=1}^{l}4\alpha_i\left[\left\|\left(B_{ij}^{\nabla}\right)^T\right\|_2^2\left\|A_{ij}^{\nabla}\right\|_2^2+\left\|\left(D_{ij}^{\nabla}\right)^T Q_{s_j}\right\|_2^2\left\|C_{ij}^{\nabla}Q_{r_j}\right\|_2^2+\left\|\left(F_{ij}^{\nabla}\right)^T Q_{r_j}\right\|_2^2\left\|E_{ij}^{\nabla}Q_{s_j}\right\|_2^2+\left\|\left(H_{ij}^{\nabla}\right)^T\right\|\left\|G_{ij}^{\nabla}\right\|_2^2\right]\\
=\;& \sum_{i=1}^{s}\sum_{j=1}^{l}4\alpha_i\left[\|B_{ij}\|_2^2\|A_{ij}\|_2^2+\|D_{ij}\|_2^2\|C_{ij}\|_2^2+\|F_{ij}\|_2^2\|E_{ij}\|_2^2+\|H_{ij}\|_2^2\|G_{ij}\|_2^2\right].
\end{aligned}
$$

Substituting Equation (62) into the relation in Equation (46) gives the conclusion of this corollary. □

4. Numerical Experiments

In this section, two numerical examples are provided to verify the effectiveness and advantage of the proposed WRGI algorithm for solving the matrix in Equation (5) in terms of the number of iterations (IT) and the computing times in seconds (CPU). All the computations were performed in MATLAB R2018b on a personal computer with an AMD Ryzen 7 5800H, where the CPU was 3.20 GHz and the memory was 16.0 GB.

Example 1 ([32]). *Consider the generalized coupled conjugate and transpose Sylvester matrix equations*

$$
\begin{cases}
A_{11}X_1B_{11}+E_{12}X_2^TF_{12}+C_{13}\overline{X_3}D_{13}+G_{14}X_4^HH_{14}=M_1,\\
A_{22}X_2B_{22}+E_{23}X_3^TF_{23}+C_{24}\overline{X_4}D_{24}+G_{21}X_1^HH_{21}=M_2,\\
A_{33}X_3B_{33}+E_{34}X_4^TF_{34}+C_{31}\overline{X_1}D_{31}+G_{32}X_2^HH_{32}=M_3,\\
A_{44}X_4B_{44}+E_{41}X_1^TF_{41}+C_{42}\overline{X_2}D_{42}+G_{43}X_3^HH_{43}=M_4,
\end{cases}
$$

with the following coefficient matrices:

$$
A_{11}=\begin{bmatrix}-12-7i & 10-11i & -9+10i\\ 2-32i & 27-3i & 1-3i\\ 10+11i & 3-7i & -14-4i\end{bmatrix}, B_{11}=\begin{bmatrix}-17-7i & -8-25i & 13+1i\\ 7+4i & -2-9i & 0+6i\\ 7-11i & -4-2i & 7+6i\end{bmatrix},
$$

$$
A_{22}=\begin{bmatrix}11-9i & -8-7i & -18-2i\\ 33-25i & -3+6i & 5-23i\\ 7-7i & -3+11i & -12-15i\end{bmatrix}, B_{22}=\begin{bmatrix}-4+13i & 7-14i & -10+2i\\ 12+5i & -4+3i & 8-16i\\ 1-5i & 19+7i & 7-7i\end{bmatrix},
$$

$$
A_{33}=\begin{bmatrix}7+6i & -5+11i & 4+8i\\ -24-1i & 11-3i & 0+22i\\ 0-7i & -2-9i & 0-6i\end{bmatrix}, B_{33}=\begin{bmatrix}4-9i & -20+15i & -23+20i\\ -25+4i & -4-2i & 11+1i\\ -14-4i & -4+8i & 10-1i\end{bmatrix},
$$

$$A_{44} = \begin{bmatrix} -12+12i & -2+5i & 1-9i \\ 2-2i & 10-6i & -26+3i \\ -19+13i & -3+15i & -7-17i \end{bmatrix}, B_{44} = \begin{bmatrix} -1-5i & 2+9i & -3+4i \\ 16-7i & 1+4i & -9-5i \\ -8-6i & -6+4i & 1 \end{bmatrix},$$

$$C_{13} = \begin{bmatrix} 5+6i & 11+7i & -12+4i \\ -15+2i & -5+7i & 0-14i \\ 11+4i & -9-17i & 2+21i \end{bmatrix}, D_{13} = \begin{bmatrix} 6+3i & -22+9i & 10-4i \\ 16+17i & 6+2i & 0+2i \\ 14+3i & -12-2i & -7-2i \end{bmatrix},$$

$$C_{24} = \begin{bmatrix} -16+1i & 0-3i & -7-10i \\ -3 & -4+6i & -9+7i \\ -9-2i & 19-9i & 6-3i \end{bmatrix}, D_{24} = \begin{bmatrix} -6-14i & -7+20i & 4 \\ 0-19i & -6-8i & -5+8i \\ -12-6i & 0 & -7-17i \end{bmatrix},$$

$$C_{31} = \begin{bmatrix} -13-2i & 15+16i & 12-23i \\ -10-4i & -10-3i & -15+12i \\ -3-9i & 5+20i & -5+5i \end{bmatrix}, D_{31} = \begin{bmatrix} -6+6i & -9+6i & -20-19i \\ 9+14i & 1+21i & -8-12i \\ -12-12i & -17-4i & 3+8i \end{bmatrix},$$

$$C_{42} = \begin{bmatrix} -1+5i & -5+25i & -2-4i \\ 19-11i & -2-6i & -2+9i \\ -10+2i & -5-9i & 13+10i \end{bmatrix}, D_{42} = \begin{bmatrix} -5 & 0+1i & 10+16i \\ 2-1i & -24+2i & -6+1i \\ -1-12i & -6+14i & 4-13i \end{bmatrix},$$

$$E_{12} = \begin{bmatrix} -14+2i & 2+5i & 4-3i \\ 9+5i & -3+10i & 8+6i \\ -8-1i & 5+9i & -21-1i \end{bmatrix}, F_{12} = \begin{bmatrix} -13+1i & -21+6i & -11-3i \\ 2+7i & -10+4i & -10+10i \\ -3+5i & 9-11i & -13+13i \end{bmatrix},$$

$$E_{23} = \begin{bmatrix} 8-2i & -8+2i & 0-5i \\ 3+6i & 8-2i & -3-1i \\ 10-3i & -13+3i & -14+7i \end{bmatrix}, F_{23} = \begin{bmatrix} -16-27i & -26+2i & -12-10i \\ 6-8i & 2+33i & 1-9i \\ -5-12i & 12-18i & 29+7i \end{bmatrix},$$

$$E_{34} = \begin{bmatrix} 14+21i & -9+8i & 4-14i \\ 7+7i & 19-2i & 0+5i \\ 7+10i & -5+8i & 7-18i \end{bmatrix}, F_{34} = \begin{bmatrix} 0+8i & 16-12i & -8+17i \\ -5+18i & 4+12i & 10+9i \\ 10-1i & 12-14i & 17-5i \end{bmatrix},$$

$$E_{41} = \begin{bmatrix} 3+8i & -18+10i & -9+14i \\ -17+6i & 3+6i & 14+7i \\ 5 & -5-17i & -6+5i \end{bmatrix}, F_{41} = \begin{bmatrix} -5-1i & 16+3i & -5-13i \\ -19+6i & -12 & -9-3i \\ -12+4i & 6-12i & 6+3i \end{bmatrix},$$

$$G_{14} = \begin{bmatrix} 2-29i & -6-3i & 8+8i \\ -3+3i & -6-3i & 1-14i \\ 1 & -3-11i & -1+19i \end{bmatrix}, H_{14} = \begin{bmatrix} 8-8i & -10-2i & -5+6i \\ 1 & -3+6i & 16+6i \\ -9+8i & 9 & -3-2i \end{bmatrix},$$

$$G_{21} = \begin{bmatrix} 15+15i & -5+19i & -16-4i \\ -7-14i & 5+12i & -11-16i \\ -1+5i & 5-3i & -1-3i \end{bmatrix}, H_{21} = \begin{bmatrix} -4+15i & -18+3i & 6-10i \\ 6-1i & 4-3i & 9-16i \\ 10+2i & -17+12i & -2-8i \end{bmatrix},$$

$$G_{32} = \begin{bmatrix} 14-4i & 25-3i & 5-2i \\ 4-5i & 2-16i & 3-2i \\ 7-22i & 2+5i & 7-18i \end{bmatrix}, H_{32} = \begin{bmatrix} 10+5i & 8+1i & 12-9i \\ 12-3i & 6-7i & -3+13i \\ 4-3i & -7-4i & -6-10i \end{bmatrix},$$

$$G_{43} = \begin{bmatrix} 5+21i & 1+20i & -4+12i \\ -7-8i & 6-6i & -2+15i \\ -13+5i & -22+4i & -2+5i \end{bmatrix}, H_{43} = \begin{bmatrix} -13-18i & 3-26i & -3-8i \\ -2+16i & 9-4i & 11-5i \\ 9+7i & 5+4i & 2-14i \end{bmatrix},$$

$$M_1 = \begin{bmatrix} -2418+3322i & 10353-966i & 5238+1933i \\ -11927-7210i & -7206-12568i & 4614+7638i \\ 1619-9753i & 16692+11938i & -3798-2865i \end{bmatrix},$$

$$M_2 = \begin{bmatrix} -4750+14828i & 10137-3634i & -18315+2472i \\ -11651+15269i & -11063-9783i & -16515+18210i \\ -2388+17370i & 2934-1222i & -3823+2947i \end{bmatrix},$$

$$M_3 = \begin{bmatrix} 22162 - 7358i & 22122 - 18790i & -3570 - 2827i \\ 7263 + 4634i & 4142 + 7164i & 6578 + 26838i \\ 12844 - 6822i & -11828 - 13695i & 9789 - 20700i \end{bmatrix},$$

$$M_4 = \begin{bmatrix} -5068 - 9357i & 6306 - 13376i & -3738 - 6683i \\ 9215 + 5146i & 8946 - 8825i & 14012 - 7731i \\ -2927 - 1660i & -5342 + 4327i & 6279 - 2721i \end{bmatrix}.$$

It can be verified that the unique solution of this matrix equation is

$$X_1^* = \begin{bmatrix} -9 + 8i & -12 - 5i & 9 + 7i \\ 6 + 7i & -10 - 4i & 14 + 2i \\ 6 - 10i & 10 + 11i & -1 \end{bmatrix}, X_2^* = \begin{bmatrix} 13 + 6i & -19 + 6i & -11 - 11i \\ 16 - 15i & 16 - 16i & 9 + 11i \\ -7 - 10i & 5 + 6i & -5 - 2i \end{bmatrix},$$

$$X_3^* = \begin{bmatrix} -13 + 7i & 14 + 3i & 2 + 1i \\ -3 + 9i & 8 + 3i & -1 + 6i \\ -15 - 1i & -6 - 4i & 5 + 4i \end{bmatrix}, X_4^* = \begin{bmatrix} -2 - 14i & 9 + 11i & -9 - 17i \\ -9 - 13i & 2 + 13i & 9 + 8i \\ 10 - 13i & -2 - 6i & 7 - 14i \end{bmatrix}.$$

We take the initial iterative matrices

$$X_p^{(i)}(0) = 10^{-6} \times I_3, X_p(0) = 10^{-6} \times I_3, i \in [1,4], p \in [1,4]$$

and set

$$RES = \sqrt{\frac{M_1(k) + M_2(k) + M_3(k) + M_4(k)}{M_1(0) + M_2(0) + M_3(0) + M_4(0)}} \leq \xi$$

to be the termination condition with a constant $\xi > 0$, where

$$M_1(k) = \left\| M_1 - A_{11}X_1(k)B_{11} - E_{12}X_2^T(k)F_{12} - C_{13}\overline{X_3(k)}D_{13} - G_{14}X_4^H(k)H_{14} \right\|^2,$$

$$M_2(k) = \left\| M_2 - A_{22}X_2(k)B_{22} - E_{23}X_3^T(k)F_{23} - C_{24}\overline{X_4(k)}D_{24} - G_{21}X_1^H(k)H_{21} \right\|^2,$$

$$M_3(k) = \left\| M_3 - A_{33}X_3(k)B_{33} - E_{34}X_4^T(k)F_{34} - C_{31}\overline{X_1(k)}D_{31} - G_{32}X_2^H(k)H_{32} \right\|^2,$$

$$M_4(k) = \left\| M_4 - A_{44}X_4(k)B_{44} - E_{41}X_1^T(k)F_{41} - C_{42}\overline{X_2(k)}D_{42} - G_{43}X_3^H(k)H_{43} \right\|^2,$$

Here, the prescribed maximum iterative number is k_{max} = 20,000, and $X_p(k)$ $(p = 1,2,3,4)$ represents the kth iteration solution.

It follows from Remark 2 that the WRGI algorithm reduces to the GI one [25] as $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.25$. In this case, according to Theorem 4, the optimal convergence factor is $\mu_{opt} = 4.5603 \times 10^{-6}$. However, the optimal convergence factor μ_{opt} calculated by Theorem 4 may not minimize the IT of the GI algorithm, and the reason for this is that there are errors in the calculation process. Thus, the parameter μ adopted in the GI algorithm is the experimentally found optimal one, which minimizes the IT of the GI algorithm. After experimental debugging, the experimental optimal parameters of the GI algorithm are $\mu_{GI1} = 4.2 \times 10^{-6}, \mu_{GI2} = 4.53 \times 10^{-6}, \mu_{GI3} = 4.556 \times 10^{-6}$ and $\mu_{GI4} = 4.558 \times 10^{-6}$ when $\xi = 0.1, \xi = 0.01, \xi = 0.001$ and $\xi = 0.0001$, respectively. Aside from that, we can obtain that the optimal parameter of the WRGI algorithm is $\mu = 4.6493 \times 10^{-6}$ in terms of Theorem 4. Owing to the existence of computation errors, the parameter μ in the WRGI algorithm is adopted to be the experimentally found one as demonstrated above. Through experimental debugging, the experimental optimal parameters of the WRGI algorithm are $\mu_{WRGI1} = 4.2 \times 10^{-6}$, $\mu_{WRGI2} = 4.61 \times 10^{-6}, \mu_{WRGI3} = 4.645 \times 10^{-6}$ and $\mu_{WRGI4} = 4.647 \times 10^{-6}$ when $\xi = 0.1$, $\xi = 0.01$, $\xi = 0.001$ and $\xi = 0.0001$, respectively. Under these circumstances, the relaxation factors are taken to be $\alpha_1 = 0.26, \alpha_2 = 0.26$, $\alpha_3 = 0.24$ and $\alpha_4 = 0.24$.

Table 1 lists the numerical results of the GI and the WRGI algorithms. As observed earlier in the comparisons of the GI and the WRGI algorithms in Table 1, the two tested algorithms were convergent for all cases, and the IT of them increased gradually with the decreases in the value of ξ. Additionally, it can be seen from Table 1 that the proposed WRGI algorithm performed better than the GI one with respect to the IT and CPU times. This implies that by applying the weighted relaxation technique, the WRGI algorithm can own better numerical behaviors than the GI one.

To better validate the advantage of the WRGI algorithm, Figures 1–4 describe the convergence curves of the GI and WRGI algorithms for four different values of ξ. These curves reflect that the WRGI algorithm is convergent and its convergence speed is faster than the GI one, except for the case of $\xi = 0.1$. Additionally, the advantage of the WRGI algorithm becomes more pronounced as the value of ξ decreases. This is consistent with the results in Table 1.

Table 1. Numerical results of the tested algorithms for Example 1 with different values of ξ.

Algorithm		ξ			
		0.1	**0.01**	**0.001**	**0.0001**
GI	IT	17	496	5312	12347
	CPU	0.0450	0.1784	1.9372	3.6098
	RES	9.93×10^{-2}	1.00×10^{-2}	9.9993×10^{-4}	9.9969×10^{-5}
WRGI	IT	17	492	5228	12128
	CPU	0.0381	0.1270	1.7313	3.1916
	RES	9.9×10^{-2}	1.00×10^{-2}	9.9998×10^{-4}	9.9999×10^{-5}

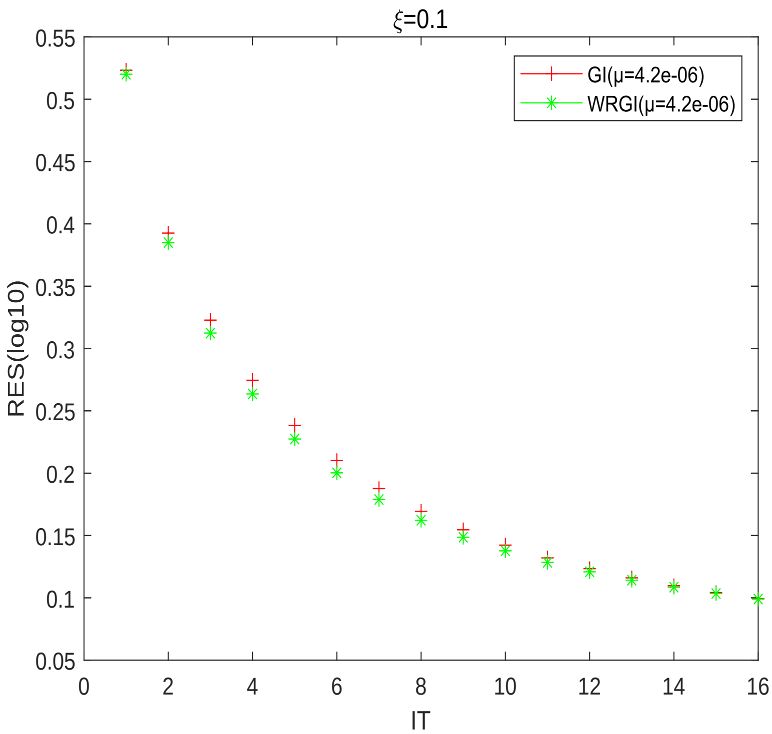

Figure 1. Convergence curves of the tested algorithms for Example 1 with $\xi = 0.1$.

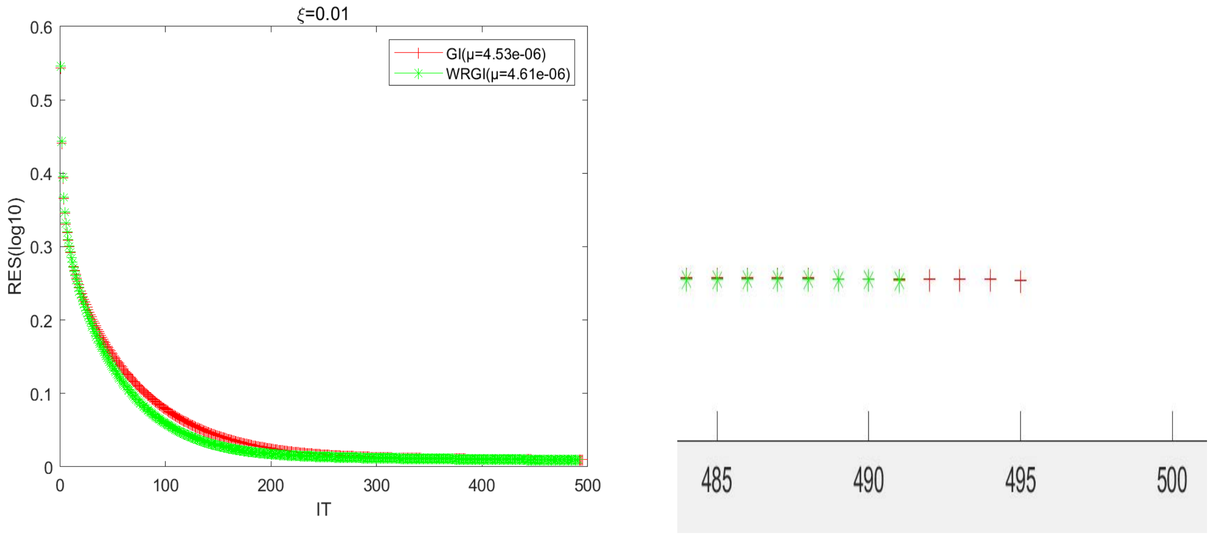

Figure 2. Convergence curves of the tested algorithms for Example 1 with $\xi = 0.01$.

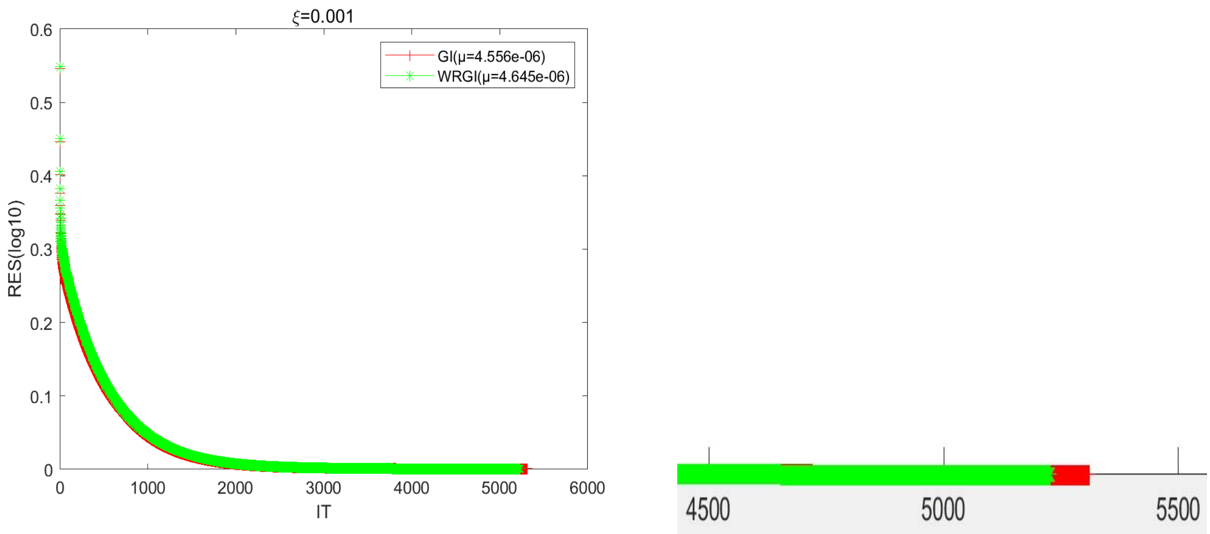

Figure 3. Convergence curves of the tested algorithms for Example 1 with $\xi = 0.001$.

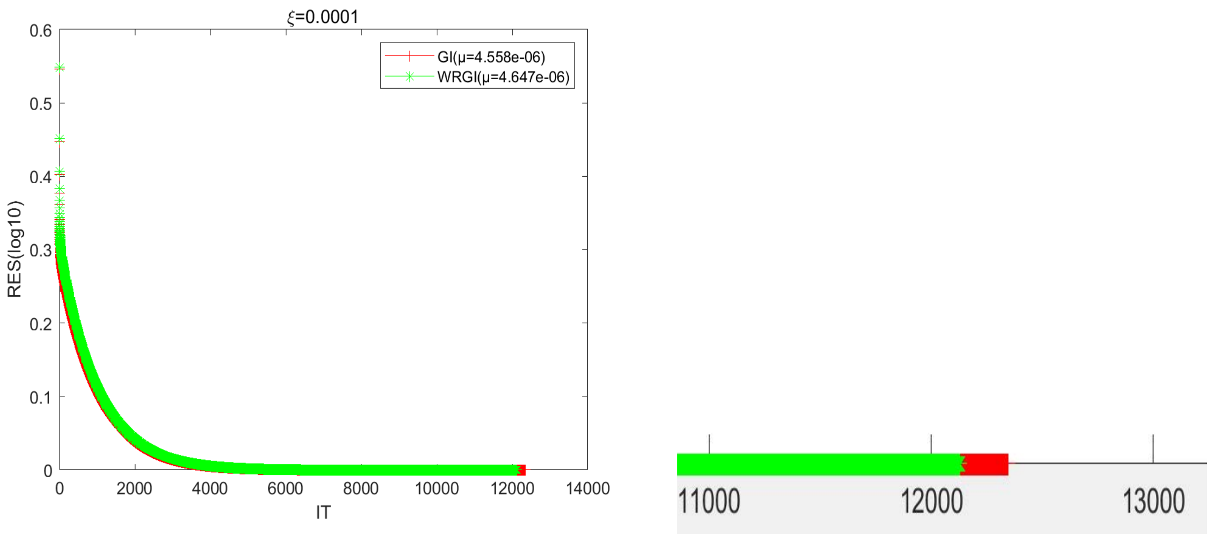

Figure 4. Convergence curves of the tested algorithms for Example 1 with $\xi = 0.0001$.

Example 2. *We consider the generalized coupled conjugate and transpose Sylvester matrix equations*

$$\begin{cases} A_{11}X_1B_{11} + C_{11}\overline{X_1}D_{11} + E_{12}X_2^TF_{12} + G_{12}X_2^HH_{12} = M_1, \\ A_{21}X_1B_{21} + C_{21}\overline{X_1}D_{21} + E_{22}X_2^TF_{22} + G_{22}X_2^HH_{22} = M_2, \end{cases}$$

with the coefficient matrices

$$A_{11} = \begin{bmatrix} 12 & 5 & 3 \\ 19 & 8 & 2 \\ -9 & -3 & 21 \end{bmatrix}, B_{11} = \begin{bmatrix} -11 & 23 & 30 \\ 4 & -9 & 3 \\ 1 & -35 & 20 \end{bmatrix}, C_{11} = \begin{bmatrix} 8+1i & 1+21i & 7 \\ -1+8i & 1 & 26 \\ 1i & 8 & 41i \end{bmatrix},$$

$$D_{11} = \begin{bmatrix} 1 & -1i & 1 \\ 1i & 1 & 1 \\ 2 & 1i & 1 \end{bmatrix}, E_{12} = \begin{bmatrix} 1 & 1i & 1 \\ 2i & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, F_{12} = \begin{bmatrix} 1 & 1 & 6 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

$$G_{12} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 4 & 1 \\ -2 & 0 & -2 \end{bmatrix}, H_{12} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1i & 0 \\ -1 & 0 & 1 \end{bmatrix}, A_{21} = \begin{bmatrix} 1 & 0 & 5i \\ 35 & -1 & 45 \\ 1 & -1 & 1 \end{bmatrix},$$

$$B_{21} = \begin{bmatrix} -1+4i & 28 & 1 \\ 1 & -42 & 1 \\ 1 & 33 & -3 \end{bmatrix}, C_{21} = \begin{bmatrix} 1i & 0 & 1 \\ -1 & 1 & -2 \\ 0 & 2 & 1 \end{bmatrix}, D_{21} = \begin{bmatrix} 1 & -1i & 3 \\ 1 & 1 & 1 \\ 1i & 1i & 1 \end{bmatrix},$$

$$E_{22} = \begin{bmatrix} 1 & -1i & 1 \\ 1i & 3 & 4-2i \\ -1 & 2 & 1 \end{bmatrix}, F_{22} = \begin{bmatrix} 1 & 8 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & -3 \end{bmatrix}, G_{22} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ -2 & 0 & -2 \end{bmatrix}$$

$$H_{22} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & -1i & 1 \\ -1 & 2 & 1 \end{bmatrix}, M_1 = \begin{bmatrix} -136-344 & 1780-688i & 5777+2083i \\ -1508-670i & 1819-1450i & 8360+3079i \\ 199-1526i & 4844+2846i & -4882+2714i \end{bmatrix},$$

$$M_2 = \begin{bmatrix} -19-76i & -1004-949i & 3+166i \\ -1222+1753i & -31239+8414i & 2088+432i \\ -28+10i & -1133-147i & 69+5i \end{bmatrix}.$$

The unique solution for these matrix equations is

$$X_1^* = \begin{bmatrix} 12+2i & 23-2i & 1 \\ 1 & -8 & 10i \\ 2+6i & 2+2i & -6 \end{bmatrix}, X_2^* = \begin{bmatrix} 1+2i & 2-3i & 3+4i \\ 2-4i & -6+3i & -3+2i \\ 1+1i & 2+2i & -1-2i \end{bmatrix}.$$

In this example, we chose

$$X_p^{(i)}(0) = 10^{-6} \times I_3, X_p(0) = 10^{-6} \times I_3, i \in [1,2], p \in [1,2],$$

as the initial matrices and took

$$RES = \sqrt{\frac{M_1(k) + M_2(k)}{M_1(0) + M_2(0)}} \leq \xi$$

as the termination condition with a positive number ξ, where

$$M_1(k) = \left\| M_1 - A_{11}X_1(k)B_{11} - C_{11}\overline{X_1(k)}D_{11} - E_{12}X_2^T(k)F_{12} - G_{12}X_2^H(k)H_{12} \right\|^2$$

and

$$M_2(k) = \left\| M_2 - A_{21}X_1(k)B_{21} - C_{21}\overline{X_1(k)}D_{21} - E_{22}X_2^T(k)F_{22} - G_{22}X_2^H(k)H_{22} \right\|^2,$$

and the prescribed maximum iterative number was k_{max} = 20,000, while $X_p(k)$ $(p = 1, 2)$ denotes the kth iteration solution.

When $\alpha_1 = \alpha_2 = 0.5$, the WRGI algorithm degenerates into the GI one, and the optimal convergence factor calculated by Theorem 4 is $\mu = 1.3151 \times 10^{-6}$. When $\alpha_1 = 0.6$ or $\alpha_2 = 0.4$, the optimal parameter of the WRGI algorithm is $\mu = 1.6269 \times 10^{-6}$ in light of Theorem 4. Due to the existence of computation errors, as in Example 1, the parameters used in the GI and WRGI algorithms for Example 2 are the experimentally found optimal ones which minimize the IT of the tested algorithms. Through experimental debugging, the experimental optimal parameters of the GI algorithm were $\mu_{GI1} = 1.21 \times 10^{-6}$ and $\mu_{GI2} = 1.299 \times 10^{-6}$ for $\xi = 0.1$ and $\xi = 0.01$, respectively. And when $\alpha_1 = 0.6$ and $\alpha_2 = 0.4$, the experimental optimal parameters of the WRGI algorithm were $\mu_{WRGI1} = 1.43 \times 10^{-6}$ and $\mu_{WRGI2} = 1.625 \times 10^{-6}$ for $\xi = 0.1$ and $\xi = 0.01$, respectively.

In Table 2, we compare the numerical results of the GI and WRGI algorithms to solve the generalized coupled conjugate and transpose Sylvester matrix equations in Example 2 with respect to two different values of ξ. Since the IT of the algorithms exceeded the maximum number of iterations when $\xi = 0.001$ and $\xi = 0.0001$, we did not list the corresponding numerical results here. From the numerical results listed in Table 2, we can conclude that when the value of ξ decreases, the IT and CPU times of the tested algorithms increase. The changing scope of the IT of the proposed WRGI algorithm was smaller than that for the GI one. In addition, the WRGI algorithm outperformed the GI one from the point of view of computing efficiency, and the advantage of the WRGI algorithm became more pronounced as ξ became small.

Figure 5 describes the convergence curves of the GI and WRGI algorithms for Example 2 with $\xi = 0.1$ and $\xi = 0.01$, respectively. It follows from Figure 5 that the error curves of the WRGI algorithm decreased faster than those for the GI one, which means that the proposed WRGI algorithm is superior to the GI one in terms of IT.

Table 2. Numerical results of the tested algorithms for Example 2 with different values of ξ.

Algorithm		ξ	
		0.1	**0.01**
GI	IT	21	2495
	CPU	0.0482	0.4058
	RES	9.82×10^{-2}	1.00×10^{-2}
WRGI	IT	15	2014
	CPU	0.0179	0.2831
	RES	9.997×10^{-2}	1.00×10^{-2}

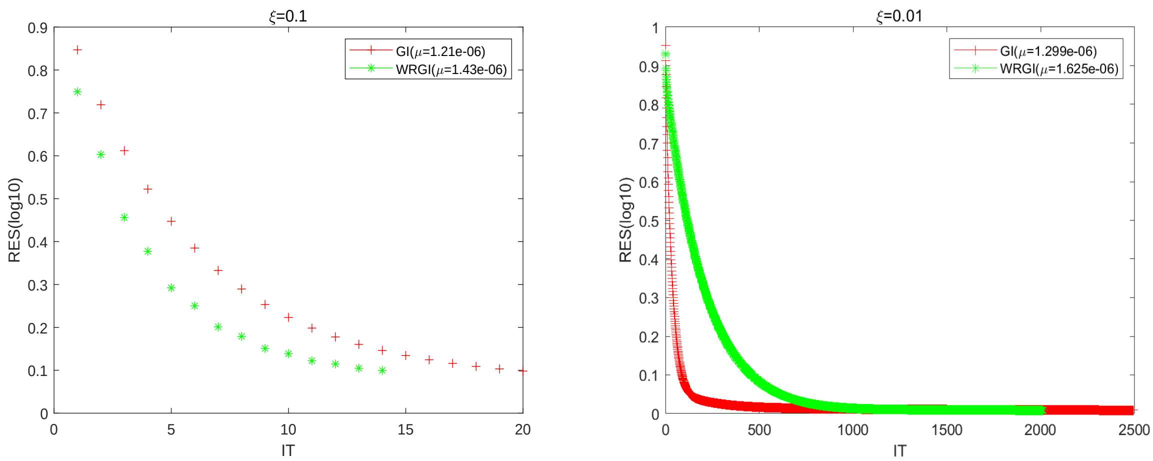

Figure 5. Convergence curves of the tested algorithms for Example 2 with $\xi = 0.1$ (**left**) and $\xi = 0.01$ (**right**).

5. Concluding Remarks

In this work, by applying the weighted technique and introducing several relaxation factors into the GI algorithm, the weighted, relaxed gradient-based iterative (WRGI) algorithm was constructed for the generalized coupled conjugate and transpose Sylvester matrix equations. By using the real representation of a complex matrix and the vector-stretching operator, the necessary and sufficient conditions for the convergence of the WRGI algorithm were given, and its optimal convergence factor was settled theoretically. Finally, two numerical examples were offered to show the effectiveness and superiority of the proposed WRGI algorithm.

Note that only one step size factor μ is used in the WRGI algorithm. We will consider to adopt different step size factors in the WRGI algorithm and investigate its algebraic properties. Aside from that, establishing different versions of the WRGI algorithm and their convergent conditions when the system matrix is of a full-row rank, full-column rank or reduced rank deserve to be studied in our future work.

Author Contributions: X.W. wrote the main manuscript text, Z.H. supervised and revised the manuscript, J.C. performed the numerical experiments, and Y.L. examined the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Science Foundations of China (No. 12361078 and 11901123), the Guangxi Natural Science Foundations (No. Guike AD20159056, 2019GXNSFBA185014, 2021GXNSFBA196064 and Guike AD21220129).

Data Availability Statement: Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Acknowledgments: We would like to express our sincere thanks to the editor and the anonymous reviewers for their valuable suggestions and constructive comments which greatly improved the presentation of this paper.

Conflicts of Interest: The authors declare that they have no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Fletcher, L.R.; Kuatslcy, J.; Nichols, N.K. Eigenstructure assignment in descriptor systems. *IEEE Trans. Autom. Control.* **1986**, *31*, 1138–1141. [CrossRef]
2. Zhou, B.; Zheng, W.-X.; Duan, G.-R. Stability and stabilization of discrete-time periodic linear systems with actuator saturation. *Automatica* **2011**, *47*, 1813–1820. [CrossRef]
3. Zhou, B.; Duan, G.-R. Periodic Lyapunov equation based approaches to the stabilization of continuous-time periodic linear systems. *IEEE Trans. Autom. Control* **2012**, *57*, 2139–2146. [CrossRef]
4. Ding, F.; Wang, F.-F.; Xu, L.; Wu, M.-H. Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering. *J. Frankl. Inst.* **2017**, *354*, 1321–1339. [CrossRef]
5. Lv, L.-L.; Chen, J.-B.; Zhang, L.; Zhang, F.-R. Gradient-based neural networks for solving periodic Sylvester matrix equations. *J. Frankl. Inst.* **2022**, *359*, 10849–10866. [CrossRef]
6. Bai, Z.-Z.; Guo, X.-X.; Yin, J.-F. On two iteration methods for the quadratic matrix equations. *Int. J. Numer. Anal. Model.* **2005**, *2*, 114–122.
7. Bai, Z.-Z.; Guo, X.-X.; Xu, S.-F. Alternately linearized implicit iteration methods for the minimal nonnegative solutions of the nonsymmetric algebraic Riccati equations. *Numer. Linear Algebra Appl.* **2006**, *13*, 655–674. [CrossRef]
8. Bai, Z.-Z.; Huang, Y.-M.; Ng, M.K. On preconditioned iterative methods for Burgers equations. *SIAM J. Sci. Comput.* **2007**, *29*, 415–439. [CrossRef]
9. Dehghan, M.; Hajarian, M. An iterative algorithm for the reflexive solutions of the generalized coupled Sylvester matrix equations and its optimal approximation. *Appl. Math. Comput.* **2008**, *202*, 571–588.
10. Deng, Y.-B.; Bai, Z.-Z.; Gao, Y.-H. Iterative orthogonal direction methods for Hermitian minimum norm solutions of two consistent matrix equations. *Numer. Linear Algebra Appl.* **2006**, *13*, 801–823. [CrossRef]
11. Guo, K.-H.; Hu, X.-Y.; Zhang, L. A new iteration method for the matrix equation $AX = B$. *Appl. Math. Comput.* **2007**, *187*, 1434–1441.
12. Li, F.-L.; Gong, L.-S.; Hu, X.-Y.; Zhang, L. Successive projection iterative method for solving matrix $AX = B$. *J. Comput. Appl. Math.* **2010**, *234*, 2405–2410. [CrossRef]

13. Ding, F.; Chen, T.-W. On iterative solutions of general coupled matrix equations. *Siam J. Control. Optim.* **2006**, *44*, 2269–2284. [CrossRef]
14. Chen, Z.-B.; Chen, X.-S. Modification on the convergence results of the Sylvester matrix equation $AX + XB = C$. *J. Frankl. Inst.* **2022**, *359*, 3126–3147. [CrossRef]
15. Ramadan, M.A.; Bayoumi, A.M.E. A modified gradient-based algorithm for solving extended Sylvester-conjufate matrix equations. *Asian J. Control.* **2018**, *20*, 228–235. [CrossRef]
16. Zhang, J.-J. A note on the iterative solutions of general coupled matrix equation. *Appl. Math. Comput.* **2011**, *217*, 9380–9386. [CrossRef]
17. Zhang, H.-M. Reduced-rank gradient-based algorithms for generalized coupled Sylvester matrix equations and its applications. *Comput. Math. Appl.* **2015**, *70*, 2049–2062. [CrossRef]
18. Zhang, H.-M.; Yin, H.-C. Conjugate gradient least squares algorithm for solving the generalized coupled Sylvester matrix equations. *Comput. Math. Appl.* **2017**, *73*, 2529–2547. [CrossRef]
19. Wu, A.-G.; Feng, G.; Duan, G.-R.; Wu, W.-J. Iterative solutions to coupled Sylvester-conjugate matrix equations. *Comput. Math. Appl.* **2010**, *60*, 54–66. [CrossRef]
20. Huang, B.-H.; Ma, C.-F. Gradient-based iterative algorithms for generalized coupled Sylvester-conjugate matrix equations. *Comput. Math. Appl.* **2018**, *75*, 2295–2310. [CrossRef]
21. Huang, B.-H.; Ma, C.-F. The relaxed gradient-based iterative algorithms for a class of generalized coupled Sylvester-conjugate matrix equations. *J. Frankl. Inst.* **2018**, *355*, 3168–3195. [CrossRef]
22. Wang, W.-L.; Qu, G.-R.; Song, C.-Q. Cyclic gradient based iterative algorithm for a class of generalized coupled Sylvester-conjugate matrix equations. *J. Frankl. Inst.* **2023**, *360*, 7206–7229. [CrossRef]
23. Song, C.-Q.; Chen, G.-L.; Zhao, L.-L. Iterative solution to coupled Sylvester-transpose matrix equations. *Appl. Math. Model.* **2011**, *35*, 4675–4683. [CrossRef]
24. Huang, B.-H.; Ma, C.-F. On the relaxed gradient-based iterative methods for the generalized coupled Sylvester-transpose matrix equations. *J. Frankl. Inst.* **2022**, *359*, 10688–10725. [CrossRef]
25. Beik, F.P.A.; Salkuyeh, D.K.; Moghadam, M.M. Gradient-based iterative algorithm for solving the generalized coupled Sylvester-transpose and conjugate matrix equations over reflexive (anti-reflexive) matrices. *Trans. Inst. Meas. Control.* **2014**, *36*, 99–110. [CrossRef]
26. Wang, W.-L.; Song, C.-Q.; Ji, S.-P. Iterative solution to a class of complex matrix equations and its application in time-varying linear system. *J. Appl. Math. Comput.* **2021**, *67*, 317–341. [CrossRef]
27. Zhang, H.-M.; Yin, H.-C. New proof of the gradient-based iterative algorithm for a complex conjugate and transpose matrix equation. *J. Frankl. Inst.* **2017**, *354*, 7585–7603. [CrossRef]
28. Niu, Q.; Wang, X.; Lu, L.-Z. A relaxed gradient based algorithm for solving Sylvester equations. *Asian J. Control.* **2011**, *13*, 461–464. [CrossRef]
29. Wang, W.-L.; Song, C.-Q. Iterative algorithms for discrete-time periodic Sylvester matrix equations and its application in antilinear periodic system. *Appl. Numer. Math.* **2021**, *168*, 251–273. [CrossRef]
30. Wu, A.-G.; Zhang, Y.; Qian, Y.-Y. *Complex Conjugate Matrix Equations*; Science Press: Beijing, China, 2017.
31. Khalil, I.S.; Doyle, J.C.; Glover, K. *Robust and Optimal Control*; Prentice Hall: Hoboken, NJ, USA, 1996.
32. Zhang, H.-M. A finite iterative algorithm for solving the complex generalized coupled Sylvester matrix equations by using the linear operators. *J. Frankl. Inst.* **2017**, *354*, 1856–1874. [CrossRef]