

Research Article

Iterative Solutions of a Set of Matrix Equations by Using the Hierarchical Identification Principle

Huamin Zhang

Department of Mathematics and Physics, Bengbu College, Bengbu 233030, China

Correspondence should be addressed to Huamin Zhang; zhangeasymail@126.com

Received 13 January 2014; Revised 29 March 2014; Accepted 30 March 2014; Published 4 May 2014

Academic Editor: Weinian Zhang

Copyright © 2014 Huamin Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper is concerned with iterative solution to a class of the real coupled matrix equations. By using the hierarchical identification principle, a gradient-based iterative algorithm is constructed to solve the real coupled matrix equations $A_1XB_1 + A_2XB_2 = F_1$ and $C_1XD_1 + C_2XD_2 = F_2$. The range of the convergence factor is derived to guarantee that the iterative algorithm is convergent for any initial value. The analysis indicates that if the coupled matrix equations have a unique solution, then the iterative solution converges fast to the exact one for any initial value under proper conditions. A numerical example is provided to illustrate the effectiveness of the proposed algorithm.

1. Introduction

For systems with certain parameters, the controllability and stability are the important topics which are worth studying [1–3]. If the parameters of systems are uncertain, then how to identify these parameters is to put on the agenda. To identify the parameters of the large-scale systems, the hierarchical identification principle was proposed in [4–6]. The hierarchical gradient-based and least squares based identifications were presented for multivariable system [7]. The fruits of these effective strategies include identifications and adaptive control for dual-rate systems [8] and Hammerstein nonlinear systems [9].

Many publications have studied the solutions to matrix equations from the different points of view [10–14]. Zhou et al. proposed the positive definite solutions of the nonlinear matrix equation $X + A^H \bar{X}^{-1} A = I$ [15]; Li et al. discussed a class of iterative methods for the generalized Sylvester equation [16]; the Riccati equation and a class of the coupled transpose matrix equations are investigated in [17, 18].

Unlike the above methods and just like the Jacobi and Gauss-Seidel iterations, Ding and Chen proposed the gradient-based and least squares based iterations for solving $Ax = b$ and $AXB = F$ [19, 20] and a large family of iterations. These iterations include the gradient iteration,

the least squares iteration, and some classical iterations as their special cases [21, 22]. By using the hierarchical identification principle, the gradient-based iterative algorithms were derived for solving different real matrix equations, such as the generalized Sylvester matrix equations and general linear matrix equations [23–25].

Ding's strategy has received much attention. With the real representations of complex matrices as tools, Wu et al. applied the Ding's strategy to solve the extended Sylvester-conjugate matrix equations $AXB + C\bar{X}D = F$ [26], the complex conjugate and transpose matrix equations [27], and the extended coupled Sylvester conjugate matrix equations [28]; Song and Chen presented the gradient based iterative algorithm for solving the extended Sylvester-conjugate transpose matrix equations [29].

Different from the various single real matrix equations in [21–25] and the complex matrix equations in [26–29], the iterative algorithm for the real coupled matrix equations $A_1XB_1 + A_2XB_2 = F_1$ and $C_1XD_1 + C_2XD_2 = F_2$ is discussed by using the hierarchical identification principle, and the gradient-based iterative algorithm is proposed. Moreover, the gradient-based iteration is reported for solving the general real coupled matrix equations. We prove that the iterative solution always converges to the exact one for any initial

value, if there exists a unique solution of these kinds of real coupled matrix equations.

The iterative methods can be applied to nonlinear system identification [30, 31]. The proposed methods of this paper can combine the iterative identification methods [32–35], the auxiliary model identification methods [36–39], the multi-innovation identification methods [40–44], and the two-stage or multistage identification methods [45] to study identification problems for other linear systems [46–50] or nonlinear systems [51–54] and other systems with colored noises [55–58].

This paper is organized as follows. Section 2 offers some notation and basic lemmas. Section 3 derives the gradient-based iterative algorithm for solving the matrix equations $\mathbf{A}_1\mathbf{X}\mathbf{B}_1 + \mathbf{A}_2\mathbf{X}\mathbf{B}_2 = \mathbf{F}_1$ and $\mathbf{C}_1\mathbf{X}\mathbf{D}_1 + \mathbf{C}_2\mathbf{X}\mathbf{D}_2 = \mathbf{F}_2$. The iteration of real general coupled matrix equations is discussed in Section 4. Section 5 presents a numerical example to illustrate the effectiveness of the proposed algorithm. Finally, we offer some concluding remarks in Section 6.

2. Notations and Basic Lemmas

Some notation and lemmas are introduced first. The symbol \mathbf{A}^T represents the transpose of \mathbf{A} . For a square matrix \mathbf{A} , $\lambda_{\max}[\mathbf{A}]$ indicates the maximum eigenvalue of \mathbf{A} . $\|\mathbf{A}\|$ denotes the norm of \mathbf{A} and is defined as $\|\mathbf{A}\|^2 := \text{tr}[\mathbf{A}^T \mathbf{A}]$. For an $m \times n$ matrix

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}, \quad \mathbf{x}_i \in \mathbb{R}^m \quad (1)$$

$\text{col}[\mathbf{X}]$ is defined as

$$\text{col}[\mathbf{X}] = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_n^T]^T \in \mathbb{R}^{mn}. \quad (2)$$

If $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{p \times q}$, then their Kronecker product is defined as

$$\mathbf{A} \otimes \mathbf{B} = [a_{ij}\mathbf{B}] = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix} \in \mathbb{R}^{(mp) \times (nq)}. \quad (3)$$

The relationship between the vec-operator col and the Kronecker product can be expressed as the following lemma [59].

Lemma 1. *If $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{p \times q}$, and $\mathbf{X} \in \mathbb{R}^{n \times p}$, then we have*

$$\text{col}[\mathbf{AXB}] = (\mathbf{B}^T \otimes \mathbf{A}) \text{col}[\mathbf{X}]. \quad (4)$$

The gradient-based iterative algorithm for solving matrix equation $\mathbf{AXB} = \mathbf{F}$ is listed as follows [22].

Lemma 2. *For $\mathbf{AXB} = \mathbf{F}$, if \mathbf{A} is a full-column rank matrix and \mathbf{B} is a full-row rank matrix, then the iterative solution $\mathbf{X}(k)$ given by the following gradient-based iterative algorithm*

converges to the exact solution \mathbf{X} (i.e., $\lim_{k \rightarrow \infty} \mathbf{X}(k) = \mathbf{X}$) for any initial values $\mathbf{X}(0)$ [22]:

$$\begin{aligned} \mathbf{X}(k) &= \mathbf{X}(k-1) + \mu \mathbf{A}^T [\mathbf{F} - \mathbf{AX}(k-1)\mathbf{B}] \mathbf{B}^T, \\ 0 < \mu < \frac{2}{\lambda_{\max}[\mathbf{AA}^T] \lambda_{\max}[\mathbf{B}^T\mathbf{B}]} \quad \text{or} \quad 0 < \mu \leq \frac{2}{\|\mathbf{A}\|^2 \|\mathbf{B}\|^2}. \end{aligned} \quad (5)$$

3. The Coupled Matrix Equations

In this section, we consider the following real coupled matrix equations:

$$\begin{aligned} \mathbf{A}_1\mathbf{X}\mathbf{B}_1 + \mathbf{A}_2\mathbf{X}\mathbf{B}_2 &= \mathbf{F}_1, \\ \mathbf{C}_1\mathbf{X}\mathbf{D}_1 + \mathbf{C}_2\mathbf{X}\mathbf{D}_2 &= \mathbf{F}_2, \end{aligned} \quad (6)$$

where $\mathbf{A}_i, \mathbf{C}_i \in \mathbb{R}^{m \times m}$, $\mathbf{B}_i, \mathbf{D}_i \in \mathbb{R}^{n \times n}$, and $\mathbf{F}_i \in \mathbb{R}^{m \times n}$, $i = 1, 2$, are the given matrices, and $\mathbf{X} \in \mathbb{R}^{m \times n}$ is the unknown matrix to be determined.

3.1. The Exact Solution. According to Lemma 1, we rewrite (6) as

$$\begin{bmatrix} \mathbf{B}_1^T \otimes \mathbf{A}_1 + \mathbf{B}_2^T \otimes \mathbf{A}_2 \\ \mathbf{D}_1^T \otimes \mathbf{C}_1 + \mathbf{D}_2^T \otimes \mathbf{C}_2 \end{bmatrix} \text{col}[\mathbf{X}] = \text{col}[\mathbf{F}_1, \mathbf{F}_2], \quad (7)$$

where $[\mathbf{F}_1, \mathbf{F}_2]$ is a block matrix. The exact solution of (6) can be given by the following theorem.

Theorem 3. *Equation (6) has a unique solution if and only if \mathbf{S}_1 is a full-column rank matrix; in this case, the unique solution is*

$$\text{col}[\mathbf{X}] = (\mathbf{S}_1^T \mathbf{S}_1)^{-1} \mathbf{S}_1^T \text{col}[\mathbf{F}_1, \mathbf{F}_2], \quad (8)$$

where

$$\mathbf{S}_1 := \begin{bmatrix} \mathbf{B}_1^T \otimes \mathbf{A}_1 + \mathbf{B}_2^T \otimes \mathbf{A}_2 \\ \mathbf{D}_1^T \otimes \mathbf{C}_1 + \mathbf{D}_2^T \otimes \mathbf{C}_2 \end{bmatrix} \in \mathbb{R}^{(2mn) \times (mn)}. \quad (9)$$

If $\mathbf{F}_1 = \mathbf{F}_2 = \mathbf{0}$, then (6) has a unique solution $\mathbf{X} = \mathbf{0}$.

3.2. The Gradient Iterative Algorithm. The hierarchical identification principle implies that the related system can be decomposed into several subsystems, so we introduce four intermediate matrices:

$$\begin{aligned} \mathbf{M}_1 &:= \mathbf{F}_1 - \mathbf{A}_2\mathbf{X}\mathbf{B}_2, \\ \mathbf{M}_2 &:= \mathbf{F}_1 - \mathbf{A}_1\mathbf{X}\mathbf{B}_1, \\ \mathbf{N}_1 &:= \mathbf{F}_2 - \mathbf{C}_2\mathbf{X}\mathbf{D}_2, \\ \mathbf{N}_2 &:= \mathbf{F}_2 - \mathbf{C}_1\mathbf{X}\mathbf{D}_1. \end{aligned} \quad (10)$$

Using these four intermediate matrices, we decompose (6) into four subsystems

$$\begin{aligned} \mathbf{S}_1: \mathbf{A}_1\mathbf{X}\mathbf{B}_1 &= \mathbf{M}_1, \\ \mathbf{S}_2: \mathbf{A}_2\mathbf{X}\mathbf{B}_2 &= \mathbf{M}_2, \\ \mathbf{S}_3: \mathbf{C}_1\mathbf{X}\mathbf{D}_1 &= \mathbf{N}_1, \\ \mathbf{S}_4: \mathbf{C}_2\mathbf{X}\mathbf{D}_2 &= \mathbf{N}_2. \end{aligned} \quad (11)$$

According to Lemma 2, it is not hard to get the iterative solutions $\mathbf{X}_1(k)$, $\mathbf{X}_2(k)$, $\mathbf{X}_3(k)$, and $\mathbf{X}_4(k)$ of these four subsystems S_1 , S_2 , S_3 , and S_4 as follows:

$$\begin{aligned} \mathbf{X}_1(k) &= \mathbf{X}_1(k-1) + \mu \mathbf{A}_1^T [\mathbf{M}_1 - \mathbf{A}_1 \mathbf{X}_1(k-1) \mathbf{B}_1] \mathbf{B}_1^T, \\ \mathbf{X}_2(k) &= \mathbf{X}_2(k-1) + \mu \mathbf{A}_2^T [\mathbf{M}_2 - \mathbf{A}_2 \mathbf{X}_2(k-1) \mathbf{B}_2] \mathbf{B}_2^T, \\ \mathbf{X}_3(k) &= \mathbf{X}_3(k-1) + \mu \mathbf{C}_1^T [\mathbf{N}_1 - \mathbf{C}_1 \mathbf{X}_3(k-1) \mathbf{D}_1] \mathbf{D}_1^T, \\ \mathbf{X}_4(k) &= \mathbf{X}_4(k-1) + \mu \mathbf{C}_2^T [\mathbf{N}_2 - \mathbf{C}_2 \mathbf{X}_4(k-1) \mathbf{D}_2] \mathbf{D}_2^T. \end{aligned} \tag{12}$$

We will determine the convergence factor μ later. Using (10), we have

$$\begin{aligned} \mathbf{X}_1(k) &= \mathbf{X}_1(k-1) \\ &\quad + \mu \mathbf{A}_1^T [\mathbf{F}_1 - \mathbf{A}_2 \mathbf{X} \mathbf{B}_2 - \mathbf{A}_1 \mathbf{X}_1(k-1) \mathbf{B}_1] \mathbf{B}_1^T, \\ \mathbf{X}_2(k) &= \mathbf{X}_2(k-1) \\ &\quad + \mu \mathbf{A}_2^T [\mathbf{F}_1 - \mathbf{A}_1 \mathbf{X} \mathbf{B}_1 - \mathbf{A}_2 \mathbf{X}_2(k-1) \mathbf{B}_2] \mathbf{B}_2^T, \\ \mathbf{X}_3(k) &= \mathbf{X}_3(k-1) \\ &\quad + \mu \mathbf{C}_1^T [\mathbf{F}_2 - \mathbf{C}_2 \mathbf{X} \mathbf{D}_2 - \mathbf{C}_1 \mathbf{X}_3(k-1) \mathbf{D}_1] \mathbf{D}_1^T, \\ \mathbf{X}_4(k) &= \mathbf{X}_4(k-1) \\ &\quad + \mu \mathbf{C}_2^T [\mathbf{F}_2 - \mathbf{C}_1 \mathbf{X} \mathbf{D}_1 - \mathbf{C}_2 \mathbf{X}_4(k-1) \mathbf{D}_2] \mathbf{D}_2^T. \end{aligned} \tag{13}$$

Because the unknown matrix \mathbf{X} appears in the right-hand side, this algorithm is impossible to realize. By using the hierarchical identification principle, we replace the unknown matrix \mathbf{X} with its estimate at iteration $k-1$. Hence, we obtain

$$\begin{aligned} \mathbf{X}_1(k) &= \mathbf{X}_1(k-1) + \mu \mathbf{A}_1^T \\ &\quad \times [\mathbf{F}_1 - \mathbf{A}_2 \mathbf{X}_1(k-1) \mathbf{B}_2 - \mathbf{A}_1 \mathbf{X}_1(k-1) \mathbf{B}_1] \mathbf{B}_1^T, \\ \mathbf{X}_2(k) &= \mathbf{X}_2(k-1) + \mu \mathbf{A}_2^T \\ &\quad \times [\mathbf{F}_1 - \mathbf{A}_1 \mathbf{X}_2(k-1) \mathbf{B}_1 - \mathbf{A}_2 \mathbf{X}_2(k-1) \mathbf{B}_2] \mathbf{B}_2^T, \\ \mathbf{X}_3(k) &= \mathbf{X}_3(k-1) + \mu \mathbf{C}_1^T \\ &\quad \times [\mathbf{F}_2 - \mathbf{C}_2 \mathbf{X}_3(k-1) \mathbf{D}_2 - \mathbf{C}_1 \mathbf{X}_3(k-1) \mathbf{D}_1] \mathbf{D}_1^T, \\ \mathbf{X}_4(k) &= \mathbf{X}_4(k-1) + \mu \mathbf{C}_2^T \\ &\quad \times [\mathbf{F}_2 - \mathbf{C}_1 \mathbf{X}_4(k-1) \mathbf{D}_1 - \mathbf{C}_2 \mathbf{X}_4(k-1) \mathbf{D}_2] \mathbf{D}_2^T. \end{aligned} \tag{14}$$

In fact, only an iterative solution $\mathbf{X}(k)$ is needed in this algorithm; we take the average of $\mathbf{X}_1(k)$, $\mathbf{X}_2(k)$, $\mathbf{X}_3(k)$,

and $\mathbf{X}_4(k)$ as $\mathbf{X}(k)$ and obtain the gradient-based iterative algorithm:

$$\mathbf{X}(k) = \frac{\mathbf{X}_1(k) + \mathbf{X}_2(k) + \mathbf{X}_3(k) + \mathbf{X}_4(k)}{4}, \tag{15}$$

$$\begin{aligned} \mathbf{X}_1(k) &= \mathbf{X}(k-1) + \mu \mathbf{A}_1^T \\ &\quad \times [\mathbf{F}_1 - \mathbf{A}_2 \mathbf{X}(k-1) \mathbf{B}_2 - \mathbf{A}_1 \mathbf{X}(k-1) \mathbf{B}_1] \mathbf{B}_1^T, \end{aligned} \tag{16}$$

$$\begin{aligned} \mathbf{X}_2(k) &= \mathbf{X}(k-1) + \mu \mathbf{A}_2^T \\ &\quad \times [\mathbf{F}_1 - \mathbf{A}_1 \mathbf{X}(k-1) \mathbf{B}_1 - \mathbf{A}_2 \mathbf{X}(k-1) \mathbf{B}_2] \mathbf{B}_2^T, \end{aligned} \tag{17}$$

$$\begin{aligned} \mathbf{X}_3(k) &= \mathbf{X}(k-1) + \mu \mathbf{C}_1^T \\ &\quad \times [\mathbf{F}_2 - \mathbf{C}_2 \mathbf{X}(k-1) \mathbf{D}_2 - \mathbf{C}_1 \mathbf{X}(k-1) \mathbf{D}_1] \mathbf{D}_1^T, \end{aligned} \tag{18}$$

$$\begin{aligned} \mathbf{X}_4(k) &= \mathbf{X}(k-1) + \mu \mathbf{C}_2^T \\ &\quad \times [\mathbf{F}_2 - \mathbf{C}_1 \mathbf{X}(k-1) \mathbf{D}_1 - \mathbf{C}_2 \mathbf{X}(k-1) \mathbf{D}_2] \mathbf{D}_2^T, \end{aligned} \tag{19}$$

$$0 < \mu < \frac{2}{\mu_0}, \quad \mu_0 := \max \left\{ \sum_{i=1}^2 \|\mathbf{A}_i\|^2 \|\mathbf{B}_i\|^2, \sum_{i=1}^2 \|\mathbf{C}_i\|^2 \|\mathbf{D}_i\|^2 \right\}. \tag{20}$$

Theorem 4. If (6) has a unique solution \mathbf{X} , then the iterative solution $\mathbf{X}(k)$ given by (15)–(20) converges to \mathbf{X} , that is, $\lim_{k \rightarrow \infty} \mathbf{X}(k) = \mathbf{X}$, or the error matrix $\mathbf{X}(k) - \mathbf{X}$ converges to zero for any initial value $\mathbf{X}(0)$.

Proof. We define estimation error matrices:

$$\begin{aligned} \widetilde{\mathbf{X}}_i(k) &:= \mathbf{X}_i(k) - \mathbf{X}, \quad i = 1, 2, 3, 4, \\ \widetilde{\mathbf{X}}(k) &:= \mathbf{X}(k) - \mathbf{X}. \end{aligned} \tag{21}$$

From this definition and (15)–(19), we have the following error formulas:

$$\begin{aligned} \widetilde{\mathbf{X}}_1(k) &= \widetilde{\mathbf{X}}(k-1) \\ &\quad - \mu \mathbf{A}_1^T [\mathbf{A}_1 \widetilde{\mathbf{X}}(k-1) \mathbf{B}_1 + \mathbf{A}_2 \widetilde{\mathbf{X}}(k-1) \mathbf{B}_2] \mathbf{B}_1^T, \\ \widetilde{\mathbf{X}}_2(k) &= \widetilde{\mathbf{X}}(k-1) \\ &\quad - \mu \mathbf{A}_2^T [\mathbf{A}_1 \widetilde{\mathbf{X}}(k-1) \mathbf{B}_1 + \mathbf{A}_2 \widetilde{\mathbf{X}}(k-1) \mathbf{B}_2] \mathbf{B}_2^T, \\ \widetilde{\mathbf{X}}_3(k) &= \widetilde{\mathbf{X}}(k-1) \\ &\quad - \mu \mathbf{C}_1^T [\mathbf{C}_1 \widetilde{\mathbf{X}}(k-1) \mathbf{D}_1 + \mathbf{C}_2 \widetilde{\mathbf{X}}(k-1) \mathbf{D}_2] \mathbf{D}_1^T, \\ \widetilde{\mathbf{X}}_4(k) &= \widetilde{\mathbf{X}}(k-1) \\ &\quad - \mu \mathbf{C}_2^T [\mathbf{C}_1 \widetilde{\mathbf{X}}(k-1) \mathbf{D}_1 + \mathbf{C}_2 \widetilde{\mathbf{X}}(k-1) \mathbf{D}_2] \mathbf{D}_2^T, \end{aligned} \tag{22}$$

$$\widetilde{\mathbf{X}}(k) = \frac{\widetilde{\mathbf{X}}_1(k) + \widetilde{\mathbf{X}}_2(k) + \widetilde{\mathbf{X}}_3(k) + \widetilde{\mathbf{X}}_4(k)}{4}. \tag{23}$$

Set

$$\begin{aligned}\tilde{\xi}(k-1) &:= \mathbf{A}_1 \tilde{\mathbf{X}}(k-1) \mathbf{B}_1 + \mathbf{A}_2 \tilde{\mathbf{X}}(k-1) \mathbf{B}_2, \\ \tilde{\eta}(k-1) &:= \mathbf{C}_1 \tilde{\mathbf{X}}(k-1) \mathbf{D}_1 + \mathbf{C}_2 \tilde{\mathbf{X}}(k-1) \mathbf{D}_2.\end{aligned}\quad (24)$$

By the trace formula $\|\mathbf{A}\|^2 = \text{tr}[\mathbf{A}^T \mathbf{A}]$ and from (22), we obtain

$$\begin{aligned}\|\tilde{\mathbf{X}}_1(k)\|^2 &= \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr} \left[\tilde{\mathbf{X}}^T(k-1) \mathbf{A}_1^T \tilde{\xi}(k-1) \mathbf{B}_1^T \right] \\ &\quad + \mu^2 \|\mathbf{A}_1^T \tilde{\xi}(k-1) \mathbf{B}_1^T\|^2,\end{aligned}\quad (25)$$

$$\begin{aligned}\|\tilde{\mathbf{X}}_2(k)\|^2 &= \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr} \left[\tilde{\mathbf{X}}^T(k-1) \mathbf{A}_2^T \tilde{\xi}(k-1) \mathbf{B}_2^T \right] \\ &\quad + \mu^2 \|\mathbf{A}_2^T \tilde{\xi}(k-1) \mathbf{B}_2^T\|^2,\end{aligned}\quad (26)$$

$$\begin{aligned}\|\tilde{\mathbf{X}}_3(k)\|^2 &= \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr} \left[\tilde{\mathbf{X}}^T(k-1) \mathbf{C}_1^T \tilde{\eta}(k-1) \mathbf{D}_1^T \right] \\ &\quad + \mu^2 \|\mathbf{C}_1^T \tilde{\eta}(k-1) \mathbf{D}_1^T\|^2,\end{aligned}\quad (27)$$

$$\begin{aligned}\|\tilde{\mathbf{X}}_4(k)\|^2 &= \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr} \left[\tilde{\mathbf{X}}^T(k-1) \mathbf{C}_2^T \tilde{\eta}(k-1) \mathbf{D}_2^T \right] \\ &\quad + \mu^2 \|\mathbf{C}_2^T \tilde{\eta}(k-1) \mathbf{D}_2^T\|^2.\end{aligned}\quad (28)$$

Adding (25) to (26) gives

$$\begin{aligned}\|\tilde{\mathbf{X}}_1(k)\|^2 + \|\tilde{\mathbf{X}}_2(k)\|^2 &= 2\|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr} \left[\tilde{\mathbf{X}}^T(k-1) \mathbf{A}_1^T \tilde{\xi}(k-1) \mathbf{B}_1^T \right] \\ &\quad - 2\mu \text{tr} \left[\tilde{\mathbf{X}}^T(k-1) \mathbf{A}_2^T \tilde{\xi}(k-1) \mathbf{B}_2^T \right] \\ &\quad + \mu^2 \|\mathbf{A}_1^T \tilde{\xi}(k-1) \mathbf{B}_1^T\|^2 \\ &\quad + \mu^2 \|\mathbf{A}_2^T \tilde{\xi}(k-1) \mathbf{B}_2^T\|^2 \\ &= 2\|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr} \left\{ \left[\mathbf{B}_1^T \tilde{\mathbf{X}}^T(k-1) \mathbf{A}_1^T \right. \right. \\ &\quad \quad \left. \left. + \mathbf{B}_2^T \tilde{\mathbf{X}}^T(k-1) \mathbf{A}_2^T \right] \right. \\ &\quad \quad \left. \times \tilde{\xi}(k-1) \right\} \\ &\quad + \mu^2 \|\mathbf{A}_1^T \tilde{\xi}(k-1) \mathbf{B}_1^T\|^2 \\ &\quad + \mu^2 \|\mathbf{A}_2^T \tilde{\xi}(k-1) \mathbf{B}_2^T\|^2 \\ &= 2\|\tilde{\mathbf{X}}(k-1)\|^2\end{aligned}$$

$$\begin{aligned}&\quad - 2\mu \text{tr} \left\{ \left[\mathbf{A}_1 \tilde{\mathbf{X}}(k-1) \mathbf{B}_1 \right. \right. \\ &\quad \quad \left. \left. + \mathbf{A}_2 \tilde{\mathbf{X}}(k-1) \mathbf{B}_2 \right]^T \right. \\ &\quad \quad \left. \times \tilde{\xi}(k-1) \right\} \\ &\quad + \mu^2 \|\mathbf{A}_1^T \tilde{\xi}(k-1) \mathbf{B}_1^T\|^2 \\ &\quad + \mu^2 \|\mathbf{A}_2^T \tilde{\xi}(k-1) \mathbf{B}_2^T\|^2 \\ &= 2\|\tilde{\mathbf{X}}(k-1)\|^2 - 2\mu \|\tilde{\xi}(k-1)\|^2 \\ &\quad + \mu^2 \|\mathbf{A}_1^T \tilde{\xi}(k-1) \mathbf{B}_1^T\|^2 \\ &\quad + \mu^2 \|\mathbf{A}_2^T \tilde{\xi}(k-1) \mathbf{B}_2^T\|^2 \\ &\leq 2\|\tilde{\mathbf{X}}(k-1)\|^2 - 2\mu \|\tilde{\xi}(k-1)\|^2 \\ &\quad + \mu^2 (\|\mathbf{A}_1\|^2 \|\mathbf{B}_1\|^2 + \|\mathbf{A}_2\|^2 \|\mathbf{B}_2\|^2) \\ &\quad \times \|\tilde{\xi}(k-1)\|^2 \\ &= 2\|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - \mu \left[2 - \mu (\|\mathbf{A}_1\|^2 \|\mathbf{B}_1\|^2 \right. \\ &\quad \quad \left. + \|\mathbf{A}_2\|^2 \|\mathbf{B}_2\|^2) \right] \\ &\quad \times \|\tilde{\xi}(k-1)\|^2.\end{aligned}\quad (29)$$

Similarly, adding (27) to (28) yields

$$\begin{aligned}\|\tilde{\mathbf{X}}_3(k)\|^2 + \|\tilde{\mathbf{X}}_4(k)\|^2 &= 2\|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr} \left[\tilde{\mathbf{X}}^T(k-1) \mathbf{C}_1^T \tilde{\eta}(k-1) \mathbf{D}_1^T \right] \\ &\quad - 2\mu \text{tr} \left[\tilde{\mathbf{X}}^T(k-1) \mathbf{C}_2^T \tilde{\eta}(k-1) \mathbf{D}_2^T \right] \\ &\quad + \mu^2 \|\mathbf{C}_1^T \tilde{\eta}(k-1) \mathbf{D}_1^T\|^2 \\ &\quad + \mu^2 \|\mathbf{C}_2^T \tilde{\eta}(k-1) \mathbf{D}_2^T\|^2 \\ &= 2\|\tilde{\mathbf{X}}(k-1)\|^2 - 2\mu \|\tilde{\eta}(k-1)\|^2 \\ &\quad + \mu^2 \|\mathbf{C}_1^T \tilde{\eta}(k-1) \mathbf{D}_1^T\|^2 \\ &\quad + \mu^2 \|\mathbf{C}_2^T \tilde{\eta}(k-1) \mathbf{D}_2^T\|^2 \\ &\leq 2\|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - \mu \left[2 - \mu (\|\mathbf{C}_1\|^2 \|\mathbf{D}_1\|^2 \right. \\ &\quad \quad \left. + \|\mathbf{C}_2\|^2 \|\mathbf{D}_2\|^2) \right] \\ &\quad \times \|\tilde{\eta}(k-1)\|^2.\end{aligned}\quad (30)$$

Using (29) and (30), we have

$$\begin{aligned} & \|\tilde{\mathbf{X}}_1(k)\|^2 + \|\tilde{\mathbf{X}}_2(k)\|^2 + \|\tilde{\mathbf{X}}_3(k)\|^2 + \|\tilde{\mathbf{X}}_4(k)\|^2 \\ & \leq 4\|\tilde{\mathbf{X}}(k-1)\|^2 \\ & \quad - \mu \left[2 - \mu (\|\mathbf{A}_1\|^2\|\mathbf{B}_1\|^2 + \|\mathbf{A}_2\|^2\|\mathbf{B}_2\|^2) \right] \|\tilde{\boldsymbol{\xi}}(k-1)\|^2 \\ & \quad - \mu \left[2 - \mu (\|\mathbf{C}_1\|^2\|\mathbf{D}_1\|^2 + \|\mathbf{C}_2\|^2\|\mathbf{D}_2\|^2) \right] \|\tilde{\boldsymbol{\eta}}(k-1)\|^2. \end{aligned} \tag{31}$$

Taking the norm of both sides of (23) gives

$$\begin{aligned} \|\tilde{\mathbf{X}}(k)\|^2 &= \frac{\|\tilde{\mathbf{X}}_1(k) + \tilde{\mathbf{X}}_2(k) + \tilde{\mathbf{X}}_3(k) + \tilde{\mathbf{X}}_4(k)\|^2}{16} \\ &\leq \frac{\|\tilde{\mathbf{X}}_1(k)\|^2 + \|\tilde{\mathbf{X}}_2(k)\|^2 + \|\tilde{\mathbf{X}}_3(k)\|^2 + \|\tilde{\mathbf{X}}_4(k)\|^2}{4}. \end{aligned} \tag{32}$$

Substituting (31) into (32) yields

$$\begin{aligned} \|\tilde{\mathbf{X}}(k)\|^2 &\leq \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - \frac{\mu}{4} \left[2 - \mu (\|\mathbf{A}_1\|^2\|\mathbf{B}_1\|^2 + \|\mathbf{A}_2\|^2\|\mathbf{B}_2\|^2) \right] \\ &\quad \times \|\tilde{\boldsymbol{\xi}}(k-1)\|^2 \\ &\quad - \frac{\mu}{4} \left[2 - \mu (\|\mathbf{C}_1\|^2\|\mathbf{D}_1\|^2 + \|\mathbf{C}_2\|^2\|\mathbf{D}_2\|^2) \right] \\ &\quad \times \|\tilde{\boldsymbol{\eta}}(k-1)\|^2 \\ &\leq \|\tilde{\mathbf{X}}(0)\|^2 - \frac{\mu}{4} \left[2 - \mu (\|\mathbf{A}_1\|^2\|\mathbf{B}_1\|^2 + \|\mathbf{A}_2\|^2\|\mathbf{B}_2\|^2) \right] \\ &\quad \times \sum_{i=0}^{k-1} \|\tilde{\boldsymbol{\xi}}(i)\|^2 \\ &\quad - \frac{\mu}{4} \left[2 - \mu (\|\mathbf{C}_1\|^2\|\mathbf{D}_1\|^2 + \|\mathbf{C}_2\|^2\|\mathbf{D}_2\|^2) \right] \\ &\quad \times \sum_{i=0}^{k-1} \|\tilde{\boldsymbol{\eta}}(i)\|^2. \end{aligned} \tag{33}$$

Thus, we have

$$\begin{aligned} & \frac{\mu}{4} \left(2 - \mu \sum_{i=1}^2 \|\mathbf{A}_i\|^2\|\mathbf{B}_i\|^2 \right) \sum_{i=0}^{k-1} \|\tilde{\boldsymbol{\xi}}(i)\|^2 \\ & + \frac{\mu}{4} \left(2 - \mu \sum_{i=1}^2 \|\mathbf{C}_i\|^2\|\mathbf{D}_i\|^2 \right) \sum_{i=0}^{k-1} \|\tilde{\boldsymbol{\eta}}(i)\|^2 \leq \|\tilde{\mathbf{X}}(0)\|^2 < \infty. \end{aligned} \tag{34}$$

Since

$$\mu_0 = \max \left\{ \sum_{i=1}^2 \|\mathbf{A}_i\|^2\|\mathbf{B}_i\|^2, \sum_{i=1}^2 \|\mathbf{C}_i\|^2\|\mathbf{D}_i\|^2 \right\}, \tag{35}$$

that is,

$$\frac{\mu}{4} (2 - \mu\mu_0) \left[\sum_{i=0}^{k-1} (\|\tilde{\boldsymbol{\xi}}(i)\|^2 + \|\tilde{\boldsymbol{\eta}}(i)\|^2) \right] \leq \|\tilde{\mathbf{X}}(0)\|^2, \tag{36}$$

if the convergence factor μ is chosen to satisfy $0 < \mu < 2/\mu_0$, then we have

$$\sum_{k=1}^{\infty} (\|\tilde{\boldsymbol{\xi}}(k)\|^2 + \|\tilde{\boldsymbol{\eta}}(k)\|^2) \leq \|\tilde{\mathbf{X}}(0)\|^2. \tag{37}$$

It follows that $\tilde{\boldsymbol{\xi}}(k) \rightarrow \mathbf{0}$ and $\tilde{\boldsymbol{\eta}}(k) \rightarrow \mathbf{0}$, as $i \rightarrow \infty$ or as $k \rightarrow \infty$, we obtain

$$\begin{aligned} \mathbf{A}_1\tilde{\mathbf{X}}(k)\mathbf{B}_1 + \mathbf{A}_2\tilde{\mathbf{X}}(k)\mathbf{B}_2 &\rightarrow \mathbf{0}, \\ \mathbf{C}_1\tilde{\mathbf{X}}(k)\mathbf{D}_1 + \mathbf{C}_2\tilde{\mathbf{X}}(k)\mathbf{D}_2 &\rightarrow \mathbf{0}. \end{aligned} \tag{38}$$

From Theorem 3, we have $\tilde{\mathbf{X}}(k) \rightarrow \mathbf{0}$, as $k \rightarrow \infty$. This completes the proof of Theorem 4. \square

4. General Real Coupled Matrix Equations

In this section, we study the general real coupled matrix equations of form

$$\begin{aligned} \mathbf{A}_1\mathbf{X}\mathbf{B}_1 + \mathbf{A}_2\mathbf{X}\mathbf{B}_2 + \dots + \mathbf{A}_p\mathbf{X}\mathbf{B}_p &= \mathbf{F}_1, \\ \mathbf{C}_1\mathbf{X}\mathbf{D}_1 + \mathbf{C}_2\mathbf{X}\mathbf{D}_2 + \dots + \mathbf{C}_q\mathbf{X}\mathbf{D}_q &= \mathbf{F}_2, \end{aligned} \tag{39}$$

where $\mathbf{A}_i, \mathbf{C}_j \in \mathbb{R}^{m \times m}$, $\mathbf{B}_i, \mathbf{D}_j \in \mathbb{R}^{n \times n}$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, q$, and $\mathbf{F}_1, \mathbf{F}_2 \in \mathbb{R}^{m \times n}$ are the given matrices, and $\mathbf{X} \in \mathbb{R}^{m \times n}$ is the unknown matrix to be determined.

4.1. The Exact Solution. According to Lemma 1, (39) can be rewritten as

$$\begin{bmatrix} \sum_{i=1}^p \mathbf{B}_i^T \otimes \mathbf{A}_i \\ \sum_{i=1}^q \mathbf{D}_i^T \otimes \mathbf{C}_i \end{bmatrix} \text{col}[\mathbf{X}] = \text{col}[\mathbf{F}_1, \mathbf{F}_2], \tag{40}$$

where $[\mathbf{F}_1, \mathbf{F}_2]$ is a block matrix. Set

$$\mathbf{S}_2 := \begin{bmatrix} \sum_{i=1}^p \mathbf{B}_i^T \otimes \mathbf{A}_i \\ \sum_{i=1}^q \mathbf{D}_i^T \otimes \mathbf{C}_i \end{bmatrix} \in \mathbb{R}^{(2mn) \times (mn)}. \tag{41}$$

The exact solution of (39) can be given by the following theorem.

Theorem 5. Equation (39) has a unique solution if and only if \mathbf{S}_2 is a full-column rank matrix; in this case, the unique solution is

$$\text{col}[\mathbf{X}] = (\mathbf{S}_2^T \mathbf{S}_2)^{-1} \mathbf{S}_2^T \text{col}[\mathbf{F}_1, \mathbf{F}_2]. \tag{42}$$

If $\mathbf{F}_1 = \mathbf{F}_2 = \mathbf{0}$, then (39) has a unique solution $\mathbf{X} = \mathbf{0}$.

4.2. *The Gradient-Based Iterative Algorithm.* We define the intermediate matrices:

$$\begin{aligned} \mathbf{G}_i &:= \mathbf{F}_1 - \sum_{r=1, r \neq i}^p \mathbf{A}_r \mathbf{X} \mathbf{B}_r, \quad i = 1, 2, \dots, p, \\ \mathbf{H}_j &:= \mathbf{F}_2 - \sum_{s=1, s \neq j}^q \mathbf{C}_s \mathbf{X} \mathbf{D}_s, \quad j = 1, 2, \dots, q. \end{aligned} \quad (43)$$

By using the hierarchical identification principle, we decompose (39) into

$$\begin{aligned} \mathbf{G}_i &= \mathbf{A}_i \mathbf{X} \mathbf{B}_i, \quad i = 1, 2, \dots, p, \\ \mathbf{H}_j &= \mathbf{C}_j \mathbf{X} \mathbf{D}_j, \quad j = 1, 2, \dots, q. \end{aligned} \quad (44)$$

According to Lemma 2, we have

$$\mathbf{Y}_i(k) = \mathbf{Y}_i(k-1) + \mu \mathbf{A}_i^T [\mathbf{G}_i - \mathbf{A}_i \mathbf{Y}_i(k-1) \mathbf{B}_i] \mathbf{B}_i^T, \quad (45)$$

$$\mathbf{Z}_j(k) = \mathbf{Z}_j(k-1) + \mu \mathbf{C}_j^T [\mathbf{H}_j - \mathbf{C}_j \mathbf{Z}_j(k-1) \mathbf{D}_j] \mathbf{D}_j^T, \quad (46)$$

where $\mathbf{Y}_i(k)$ and $\mathbf{Z}_j(k)$ denote the iterative solutions of equations $\mathbf{G}_i = \mathbf{A}_i \mathbf{X} \mathbf{B}_i$ and $\mathbf{H}_j = \mathbf{C}_j \mathbf{X} \mathbf{D}_j$ at iteration k , respectively. Substituting \mathbf{G}_i into (45) and \mathbf{H}_j into (46) gives

$$\begin{aligned} \mathbf{Y}_i(k) &= \mathbf{Y}_i(k-1) \\ &+ \mu \mathbf{A}_i^T \left[\mathbf{F}_1 - \sum_{r=1, r \neq i}^p \mathbf{A}_r \mathbf{X} \mathbf{B}_r - \mathbf{A}_i \mathbf{Y}_i(k-1) \mathbf{B}_i \right] \mathbf{B}_i^T, \\ \mathbf{Z}_j(k) &= \mathbf{Z}_j(k-1) \\ &+ \mu \mathbf{C}_j^T \left[\mathbf{F}_2 - \sum_{s=1, s \neq j}^q \mathbf{C}_s \mathbf{X} \mathbf{D}_s - \mathbf{C}_j \mathbf{Z}_j(k-1) \mathbf{D}_j \right] \mathbf{D}_j^T. \end{aligned} \quad (47)$$

To realize the above algorithm, we replace the unknown matrix \mathbf{X} in the right-hand side of the above two equations with $\mathbf{Y}_i(k-1)$ and $\mathbf{Z}_j(k-1)$, respectively, and obtain

$$\begin{aligned} \mathbf{Y}_i(k) &= \mathbf{Y}_i(k-1) + \mu \mathbf{A}_i^T \left[\mathbf{F}_1 - \sum_{r=1}^p \mathbf{A}_r \mathbf{Y}_i(k-1) \mathbf{B}_r \right] \mathbf{B}_i^T, \\ & \quad i = 1, 2, \dots, p, \\ \mathbf{Z}_j(k) &= \mathbf{Z}_j(k-1) + \mu \mathbf{C}_j^T \left[\mathbf{F}_2 - \sum_{s=1}^q \mathbf{C}_s \mathbf{Z}_j(k-1) \mathbf{D}_s \right] \mathbf{D}_j^T, \\ & \quad j = 1, 2, \dots, q. \end{aligned} \quad (48)$$

In fact, only an iterative solution $\mathbf{X}(k)$ is needed; taking the average of $\mathbf{Y}_i(k)$ and $\mathbf{Z}_j(k)$ as $\mathbf{X}(k)$, we obtain the gradient iterative algorithm for solving (39) as follows:

$$\mathbf{X}(k) = \frac{\sum_{i=1}^p \mathbf{Y}_i(k) + \sum_{j=1}^q \mathbf{Z}_j(k)}{p+q}, \quad (49)$$

$$\mathbf{Y}_i(k) = \mathbf{X}(k-1) + \mu \mathbf{A}_i^T \left[\mathbf{F}_1 - \sum_{r=1}^p \mathbf{A}_r \mathbf{X}(k-1) \mathbf{B}_r \right] \mathbf{B}_i^T, \quad (50)$$

$$\mathbf{Z}_j(k) = \mathbf{X}(k-1) + \mu \mathbf{C}_j^T \left[\mathbf{F}_2 - \sum_{s=1}^q \mathbf{C}_s \mathbf{X}(k-1) \mathbf{D}_s \right] \mathbf{D}_j^T, \quad (51)$$

$$0 < \mu < \frac{2}{\mu_0}, \quad \mu_0 := \max \left\{ \sum_{i=1}^p \|\mathbf{A}_i\|^2 \|\mathbf{B}_i\|^2, \sum_{j=1}^q \|\mathbf{C}_j\|^2 \|\mathbf{D}_j\|^2 \right\}. \quad (52)$$

Theorem 6. *If (39) has a unique solution \mathbf{X} , then the iterative solution $\mathbf{X}(k)$ given by (49)–(52) converges to \mathbf{X} , that is, $\lim_{k \rightarrow \infty} \mathbf{X}(k) = \mathbf{X}$, or the error matrix $\mathbf{X}(k) - \mathbf{X}$ converges to zero for any initial value $\mathbf{X}(0)$.*

Proof. The estimation error matrices are defined as

$$\begin{aligned} \tilde{\mathbf{Y}}_i(k) &:= \mathbf{Y}_i(k) - \mathbf{X}, \quad i = 1, 2, \dots, p, \\ \tilde{\mathbf{Z}}_j(k) &:= \mathbf{Z}_j(k) - \mathbf{X}, \quad j = 1, 2, \dots, q, \\ \tilde{\mathbf{X}}(k) &:= \mathbf{X}(k) - \mathbf{X}. \end{aligned} \quad (53)$$

By using (49)–(51), we have

$$\begin{aligned} \tilde{\mathbf{Y}}_i(k) &= \tilde{\mathbf{X}}(k-1) - \mu \mathbf{A}_i^T \left[\sum_{r=1}^p \mathbf{A}_r \tilde{\mathbf{X}}(k-1) \mathbf{B}_r \right] \mathbf{B}_i^T, \\ & \quad i = 1, 2, \dots, p, \\ \tilde{\mathbf{Z}}_j(k) &= \tilde{\mathbf{X}}(k-1) - \mu \mathbf{C}_j^T \left[\sum_{s=1}^q \mathbf{C}_s \tilde{\mathbf{X}}(k-1) \mathbf{D}_s \right] \mathbf{D}_j^T, \\ & \quad j = 1, 2, \dots, q, \end{aligned} \quad (54)$$

$$\tilde{\mathbf{X}}(k) = \frac{\sum_{i=1}^p \tilde{\mathbf{Y}}_i(k) + \sum_{j=1}^q \tilde{\mathbf{Z}}_j(k)}{p+q}. \quad (55)$$

Set

$$\begin{aligned} \tilde{\xi}(k-1) &:= \sum_{r=1}^p \mathbf{A}_r \tilde{\mathbf{X}}(k-1) \mathbf{B}_r, \\ \tilde{\eta}(k-1) &:= \sum_{s=1}^q \mathbf{C}_s \tilde{\mathbf{X}}(k-1) \mathbf{D}_s. \end{aligned} \quad (56)$$

According to the trace formula $\|\mathbf{X}\|^2 = \text{tr}[\mathbf{X}^T \mathbf{X}]$ and using (54), for $i = 1, 2, \dots, p$ and $j = 1, 2, \dots, q$, we obtain

$$\begin{aligned} \|\tilde{\mathbf{Y}}_i(k)\|^2 &= \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr}[\tilde{\mathbf{X}}(k-1)^T \mathbf{A}_i^T \tilde{\boldsymbol{\xi}}(k-1) \mathbf{B}_i^T] \\ &\quad + \mu^2 \|\mathbf{A}_i^T \tilde{\boldsymbol{\xi}}(k-1) \mathbf{B}_i^T\|^2, \end{aligned} \quad (57)$$

$$\begin{aligned} \|\tilde{\mathbf{Z}}_j(k)\|^2 &= \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \text{tr}[\tilde{\mathbf{X}}(k-1)^T \mathbf{C}_j^T \tilde{\boldsymbol{\eta}}(k-1) \mathbf{D}_j^T] \\ &\quad + \mu^2 \|\mathbf{C}_j^T \tilde{\boldsymbol{\eta}}(k-1) \mathbf{D}_j^T\|^2. \end{aligned} \quad (58)$$

For $i = 1, 2, \dots, p$ and $j = 1, 2, \dots, q$, adding (57) to (58) gives

$$\begin{aligned} &\sum_{i=1}^p \|\tilde{\mathbf{Y}}_i(k)\|^2 + \sum_{j=1}^q \|\tilde{\mathbf{Z}}_j(k)\|^2 \\ &= (p+q) \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \left(\|\tilde{\boldsymbol{\xi}}(k-1)\|^2 + \|\tilde{\boldsymbol{\eta}}(k-1)\|^2 \right) \\ &\quad + \mu^2 \sum_{i=1}^p \|\mathbf{A}_i^T \tilde{\boldsymbol{\xi}}(k-1) \mathbf{B}_i^T\|^2 \\ &\quad + \mu^2 \sum_{j=1}^q \|\mathbf{C}_j^T \tilde{\boldsymbol{\eta}}(k-1) \mathbf{D}_j^T\|^2 \\ &\leq (p+q) \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - 2\mu \left(\|\tilde{\boldsymbol{\xi}}(k-1)\|^2 + \|\tilde{\boldsymbol{\eta}}(k-1)\|^2 \right) \\ &\quad + \mu^2 \|\tilde{\boldsymbol{\xi}}(k-1)\|^2 \sum_{i=1}^p \|\mathbf{A}_i\|^2 \|\mathbf{B}_i\|^2 \\ &\quad + \mu^2 \|\tilde{\boldsymbol{\eta}}(k-1)\|^2 \sum_{j=1}^q \|\mathbf{C}_j\|^2 \|\mathbf{D}_j\|^2 \\ &= (p+q) \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - \mu \left(2 - \mu \sum_{i=1}^p \|\mathbf{A}_i\|^2 \|\mathbf{B}_i\|^2 \right) \|\tilde{\boldsymbol{\xi}}(k-1)\|^2 \\ &\quad - \mu \left(2 - \mu \sum_{j=1}^q \|\mathbf{C}_j\|^2 \|\mathbf{D}_j\|^2 \right) \|\tilde{\boldsymbol{\eta}}(k-1)\|^2. \end{aligned} \quad (59)$$

Taking the norm of both sides of (55) and using the norm inequality $\|\sum_{i=1}^m \mathbf{A}_i\|^2 \leq m \sum_{i=1}^m \|\mathbf{A}_i\|^2$ gives

$$\begin{aligned} \|\tilde{\mathbf{X}}(k)\|^2 &= \frac{\|\sum_{i=1}^p \tilde{\mathbf{Y}}_i(k) + \sum_{j=1}^q \tilde{\mathbf{Z}}_j(k)\|^2}{(p+q)^2} \\ &\leq \frac{\sum_{i=1}^p \|\tilde{\mathbf{Y}}_i(k)\|^2 + \sum_{j=1}^q \|\tilde{\mathbf{Z}}_j(k)\|^2}{p+q}. \end{aligned} \quad (60)$$

Dividing (59) by $p+q$ and using the inequality (60) gives

$$\begin{aligned} \|\tilde{\mathbf{X}}(k)\|^2 &\leq \|\tilde{\mathbf{X}}(k-1)\|^2 \\ &\quad - \frac{\mu}{p+q} \left(2 - \mu \sum_{i=1}^p \|\mathbf{A}_i\|^2 \|\mathbf{B}_i\|^2 \right) \|\tilde{\boldsymbol{\xi}}(k-1)\|^2 \\ &\quad - \frac{\mu}{p+q} \left(2 - \mu \sum_{j=1}^q \|\mathbf{C}_j\|^2 \|\mathbf{D}_j\|^2 \right) \|\tilde{\boldsymbol{\eta}}(k-1)\|^2 \\ &\leq \|\tilde{\mathbf{X}}(k-1)\|^2 - \frac{\mu}{p+q} (2 - \mu\mu_0) \\ &\quad \times \left(\|\tilde{\boldsymbol{\xi}}(k-1)\|^2 + \|\tilde{\boldsymbol{\eta}}(k-1)\|^2 \right) \\ &\leq \|\tilde{\mathbf{X}}(0)\|^2 \\ &\quad - \frac{\mu}{p+q} (2 - \mu\mu_0) \sum_{t=0}^{k-1} \left(\|\tilde{\boldsymbol{\xi}}(t)\|^2 + \|\tilde{\boldsymbol{\eta}}(t)\|^2 \right). \end{aligned} \quad (61)$$

In the second “ \leq ” the assumption that $\mu_0 = \max\{\sum_{i=1}^p \|\mathbf{A}_i\|^2 \|\mathbf{B}_i\|^2, \sum_{j=1}^q \|\mathbf{C}_j\|^2 \|\mathbf{D}_j\|^2\}$ is used. Since $\|\tilde{\mathbf{X}}(k)\|^2 \geq 0$, we have

$$\frac{\mu}{p+q} (2 - \mu\mu_0) \sum_{t=0}^{k-1} \left(\|\tilde{\boldsymbol{\xi}}(t)\|^2 + \|\tilde{\boldsymbol{\eta}}(t)\|^2 \right) \leq \|\tilde{\mathbf{X}}(0)\|^2 < \infty. \quad (62)$$

If the convergence factor μ is chosen to satisfy $0 < \mu < 2/\mu_0$, then we have

$$\sum_{k=0}^{\infty} \left(\|\tilde{\boldsymbol{\xi}}(k)\|^2 + \|\tilde{\boldsymbol{\eta}}(k)\|^2 \right) \leq \|\tilde{\mathbf{X}}(0)\|^2. \quad (63)$$

It follows that $\|\tilde{\boldsymbol{\eta}}(k)\|^2 + \|\tilde{\boldsymbol{\xi}}(k)\|^2 \rightarrow 0$, as $k \rightarrow \infty$ or as $k \rightarrow \infty$, we obtain

$$\begin{aligned} &\mathbf{A}_1 \tilde{\mathbf{X}}(k-1) \mathbf{B}_1 + \mathbf{A}_2 \tilde{\mathbf{X}}(k-1) \mathbf{B}_2 \\ &\quad + \dots + \mathbf{A}_p \tilde{\mathbf{X}}(k-1) \mathbf{B}_p \rightarrow \mathbf{0}, \\ &\mathbf{C}_1 \tilde{\mathbf{X}}(k-1) \mathbf{D}_1 + \mathbf{C}_2 \tilde{\mathbf{X}}(k-1) \mathbf{D}_2 \\ &\quad + \dots + \mathbf{C}_q \tilde{\mathbf{X}}(k-1) \mathbf{D}_q \rightarrow \mathbf{0}. \end{aligned} \quad (64)$$

According to Theorem 5, we have $\tilde{\mathbf{X}}(k) \rightarrow \mathbf{0}$, as $k \rightarrow \infty$. This completes the proof of Theorem 6. \square

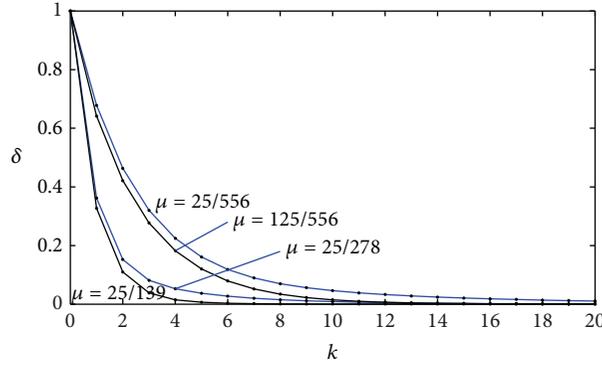


FIGURE 1: The relative error δ versus k (dots: $\mu_0 = \max\{\sum_{i=1}^2 \|A_i\|^2 \|B_i\|^2, \sum_{j=1}^2 \|C_j\|^2 \|D_j\|^2\} = 11.12$).

5. A Numerical Example

This section offers a numerical example to illustrate the performance of the proposed algorithm. Consider (6) with

$$\begin{aligned}
 A_1 &= \begin{bmatrix} -0.8 & -0.3 \\ -0.2 & -2.1 \end{bmatrix}, & B_1 &= \begin{bmatrix} -1.7 & 0.4 \\ -0.8 & -1.5 \end{bmatrix}, \\
 A_2 &= \begin{bmatrix} 0.9 & -1.5 \\ 0.6 & -0.2 \end{bmatrix}, & B_2 &= \begin{bmatrix} 0.3 & 1.4 \\ 1.1 & 0.9 \end{bmatrix}, \\
 C_1 &= \begin{bmatrix} -2.4 & 1.3 \\ 0.4 & -0.2 \end{bmatrix}, & D_1 &= \begin{bmatrix} 0.6 & -0.3 \\ 0.6 & 1 \end{bmatrix}, \\
 C_2 &= \begin{bmatrix} 0.9 & 0.6 \\ -0.4 & -0.1 \end{bmatrix}, & D_2 &= \begin{bmatrix} 0 & 1.2 \\ 0.1 & -1.5 \end{bmatrix}, \\
 F_1 &= \begin{bmatrix} -24.34 & -38.9 \\ 25.73 & 1.74 \end{bmatrix}, & F_2 &= \begin{bmatrix} 24.18 & 31.11 \\ -3.7 & -7.52 \end{bmatrix}.
 \end{aligned} \tag{65}$$

From Theorem 3, the exact solution is

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} -2 & -8 \\ 7 & 6 \end{bmatrix}. \tag{66}$$

Taking $X(0) = 10^{-6} \mathbf{1}_{2 \times 2}$ as initial iterative value, we apply the gradient-based iterative algorithm in (15)–(19) to compute $X(k)$. The iterative solutions $X(k)$ are shown in Table 1, where $\delta = \|X(k) - X\|/\|X\|$ is the relative error and the convergence factor μ is 25/139. The relation of the relative error δ with different convergence factors is shown in Figure 1. From Table 1 and Figure 1, we can find that δ becomes smaller and smaller and tends to zero as the iterative times increase. This demonstrates that the algorithm proposed in this paper is effective.

A simple calculation indicates that the range of the convergence factor is conservative. Based on the deep analysis of Figure 1, we find that the rate of convergence increases when μ enlarges from 25/556 to 25/278. However, if we keep enlarging μ from 25/139 to 125/556, the rate of convergence will drop. This shows that the best convergence factor is uncovered in this algorithm. How to find the best convergence factor subtly is our work in the future.

TABLE 1: The gradient iterative solution ($\mu = 25/139$).

k	x_{11}	x_{12}	x_{21}	x_{22}	δ (%)
2	-1.60864	-7.04633	6.47331	5.28946	10.98102
4	-1.89505	-7.89255	6.97696	5.89765	1.48121
6	-1.96855	-7.98772	7.00523	5.98269	0.30968
8	-1.99026	-7.99846	7.00259	5.99669	0.08667
10	-1.99696	-7.99976	7.00084	5.99931	0.02617
12	-1.99905	-7.99995	7.00025	5.99985	0.00805
14	-1.99970	-7.99999	7.00007	5.99996	0.00249
16	-1.99991	-8.00000	7.00002	5.99999	0.00077
18	-1.99997	-8.00000	7.00001	6.00000	0.00024
20	-1.99999	-8.00000	7.00000	6.00000	0.00007
Solution	-2.00000	-8.00000	7.00000	6.00000	

6. Conclusions

This paper has proposed a gradient-based iterative algorithm for solving a class of the real coupled matrix equations. By using the hierarchical identification principle, we prove that the iterative solution is convergent if the unique solution exists. An example demonstrates that the algorithm is effective and indicates that the best convergence factor is existent. We will find the best convergence factor of this algorithm in the future.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China.

References

[1] Y. Shi and B. Yu, “Output feedback stabilization of networked control systems with random delays modeled by Markov

- chains," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1668–1674, 2009.
- [2] H. Li and Y. Shi, "State-feedback H_∞ control for stochastic time-delay nonlinear systems with state and disturbance-dependent noise," *International Journal of Control*, vol. 85, no. 10, pp. 1515–1531, 2012.
- [3] Y. Shi and B. Yu, "Robust mixed H_2/H_∞ control of networked control systems with random time delays in both forward and backward communication links," *Automatica*, vol. 47, no. 4, pp. 754–760, 2011.
- [4] F. Ding and T. Chen, "Hierarchical least squares identification methods for multivariable systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 397–402, 2005.
- [5] F. Ding, L. Qiu, and T. Chen, "Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems," *Automatica*, vol. 45, no. 2, pp. 324–332, 2009.
- [6] Y. Liu, F. Ding, and Y. Shi, "An efficient hierarchical identification method for general dual-rate sampled-data systems," *Automatica*, vol. 50, no. 3, pp. 962–973, 2014.
- [7] F. Ding and T. Chen, "Hierarchical gradient-based identification of multivariable discrete-time systems," *Automatica*, vol. 41, no. 2, pp. 315–325, 2005.
- [8] F. Ding and T. Chen, "Hierarchical identification of lifted state-space models for general dual-rate systems," *IEEE Transactions on Circuits and Systems. I. Regular Papers*, vol. 52, no. 6, pp. 1179–1187, 2005.
- [9] F. Ding, T. Chen, and Z. Iwai, "Adaptive digital control of Hammerstein nonlinear systems with limited output sampling," *SIAM Journal on Control and Optimization*, vol. 45, no. 6, pp. 2257–2276, 2007.
- [10] B. Zhou, J. Lam, and G.-R. Duan, "Toward solution of matrix equation $X = Af(X)B + C$," *Linear Algebra and Its Applications*, vol. 435, no. 6, pp. 1370–1398, 2011.
- [11] H. Zhang and F. Ding, "A property of the eigenvalues of the symmetric positive definite matrix and the iterative algorithm for coupled Sylvester matrix equations," *Journal of the Franklin Institute*, vol. 351, no. 1, pp. 340–357, 2014.
- [12] X.-F. Duan, Q.-W. Wang, and J.-F. Li, "On the low-rank approximation arising in the generalized Karhunen-Loeve transform," *Abstract and Applied Analysis*, vol. 2013, Article ID 528281, 8 pages, 2013.
- [13] J.-F. Li, X.-Y. Hu, and L. Zhang, "New symmetry preserving method for optimal correction of damping and stiffness matrices using measured modes," *Journal of Computational and Applied Mathematics*, vol. 234, no. 5, pp. 1572–1585, 2010.
- [14] J. Liu, Z. Huang, L. Zhu, and Z. Huang, "Theorems on Schur complement of block diagonally dominant matrices and their application in reducing the order for the solution of large scale linear systems," *Linear Algebra and Its Applications*, vol. 435, no. 12, pp. 3085–3100, 2011.
- [15] B. Zhou, G.-B. Cai, and J. Lam, "Positive definite solutions of the nonlinear matrix equation $X + A^H X^{-1} A = I$," *Applied Mathematics and Computation*, vol. 219, no. 14, pp. 7377–7391, 2013.
- [16] J.-F. Li, X.-Y. Hu, and X.-F. Duan, "A symmetric preserving iterative method for generalized Sylvester equation," *Asian Journal of Control*, vol. 13, no. 3, pp. 408–417, 2011.
- [17] J. Liu, J. Zhang, and Y. Liu, "New solution bounds for the continuous algebraic Riccati equation," *Journal of the Franklin Institute*, vol. 348, no. 8, pp. 2128–2141, 2011.
- [18] K. Liang and J. Liu, "Iterative algorithms for the minimum-norm solution and the least-squares solution of the linear matrix equations $A_1 X B_1 + C_1 X^T D_1 = M_1$, $A_2 X B_2 + C_2 X^T D_2 = M_2$," *Applied Mathematics and Computation*, vol. 218, no. 7, pp. 3166–3175, 2011.
- [19] F. Ding and T. Chen, "Gradient based iterative algorithms for solving a class of matrix equations," *IEEE Transactions on Automatic Control*, vol. 50, no. 8, pp. 1216–1221, 2005.
- [20] F. Ding and T. Chen, "Iterative least-squares solutions of coupled Sylvester matrix equations," *Systems & Control Letters*, vol. 54, no. 2, pp. 95–107, 2005.
- [21] F. Ding and T. Chen, "On iterative solutions of general coupled matrix equations," *SIAM Journal on Control and Optimization*, vol. 44, no. 6, pp. 2269–2284, 2006.
- [22] F. Ding, P. X. Liu, and J. Ding, "Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle," *Applied Mathematics and Computation*, vol. 197, no. 1, pp. 41–50, 2008.
- [23] L. Xie, J. Ding, and F. Ding, "Gradient based iterative solutions for general linear matrix equations," *Computers & Mathematics with Applications*, vol. 58, no. 7, pp. 1441–1448, 2009.
- [24] L. Xie, Y. Liu, and H. Yang, "Gradient based and least squares based iterative algorithms for matrix equations $AXB + CX^T D = F$," *Applied Mathematics and Computation*, vol. 217, no. 5, pp. 2191–2199, 2010.
- [25] J. Ding, Y. Liu, and F. Ding, "Iterative solutions to matrix equations of the form $A_i X B_i = F_i$," *Computers & Mathematics with Applications*, vol. 59, no. 11, pp. 3500–3507, 2010.
- [26] A.-G. Wu, X. Zeng, G.-R. Duan, and W.-J. Wu, "Iterative solutions to the extended Sylvester-conjugate matrix equations," *Applied Mathematics and Computation*, vol. 217, no. 1, pp. 130–142, 2010.
- [27] A.-G. Wu, L. Lv, and G.-R. Duan, "Iterative algorithms for solving a class of complex conjugate and transpose matrix equations," *Applied Mathematics and Computation*, vol. 217, no. 21, pp. 8343–8353, 2011.
- [28] A.-G. Wu, G. Feng, G.-R. Duan, and W.-J. Wu, "Iterative solutions to coupled Sylvester-conjugate matrix equations," *Computers & Mathematics with Applications*, vol. 60, no. 1, pp. 54–66, 2010.
- [29] C. Song and G. Chen, "An efficient algorithm for solving extended Sylvester-conjugate transpose matrix equations," *Arab Journal of Mathematical Sciences*, vol. 17, no. 2, pp. 115–134, 2011.
- [30] F. Ding, X. P. Liu, and G. Liu, "Identification methods for Hammerstein nonlinear systems," *Digital Signal Processing*, vol. 21, no. 2, pp. 215–238, 2011.
- [31] F. Ding, "Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling," *Applied Mathematical Modelling*, vol. 37, no. 4, pp. 1694–1704, 2013.
- [32] F. Ding, Y. Liu, and B. Bao, "Gradient-based and least-squares-based iterative estimation algorithms for multi-input multi-output systems," *Proceedings of the Institution of Mechanical Engineers I: Journal of Systems and Control Engineering*, vol. 226, no. 1, pp. 43–55, 2012.
- [33] F. Ding, P. X. Liu, and G. Liu, "Gradient based and least-squares based iterative identification methods for OE and OEMA systems," *Digital Signal Processing*, vol. 20, no. 3, pp. 664–677, 2010.
- [34] F. Ding, "Decomposition based fast least squares algorithm for output error systems," *Signal Processing*, vol. 93, no. 5, pp. 1235–1242, 2013.

- [35] D. Q. Wang, "Least squares-based recursive and iterative estimation for output error moving average systems using data filtering," *IET Control Theory & Applications*, vol. 5, no. 14, pp. 1648–1657, 2011.
- [36] F. Ding, Y. Shi, and T. Chen, "Auxiliary model-based least-squares identification methods for Hammerstein output-error systems," *Systems & Control Letters*, vol. 56, no. 5, pp. 373–380, 2007.
- [37] F. Ding, P. X. Liu, and G. Liu, "Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises," *Signal Processing*, vol. 89, no. 10, pp. 1883–1890, 2009.
- [38] F. Ding and Y. Gu, "Performance analysis of the auxiliary model-based least-squares identification algorithm for one-step state-delay systems," *International Journal of Computer Mathematics*, vol. 89, no. 15, pp. 2019–2028, 2012.
- [39] F. Ding and Y. Gu, "Performance analysis of the auxiliary model-based stochastic gradient parameter estimation algorithm for state-space systems with one-step state delay," *Circuits, Systems, and Signal Processing*, vol. 32, no. 2, pp. 585–599, 2013.
- [40] F. Ding and T. Chen, "Performance analysis of multi-innovation gradient type identification methods," *Automatica*, vol. 43, no. 1, pp. 1–14, 2007.
- [41] F. Ding, "Several multi-innovation identification methods," *Digital Signal Processing*, vol. 20, no. 4, pp. 1027–1039, 2010.
- [42] L. Han and F. Ding, "Multi-innovation stochastic gradient algorithms for multi-input multi-output systems," *Digital Signal Processing*, vol. 19, no. 4, pp. 545–554, 2009.
- [43] D. Wang and F. Ding, "Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems," *Digital Signal Processing*, vol. 20, no. 3, pp. 750–762, 2010.
- [44] Y. Liu, L. Yu, and F. Ding, "Multi-innovation extended stochastic gradient algorithm and its performance analysis," *Circuits, Systems, and Signal Processing*, vol. 29, no. 4, pp. 649–667, 2010.
- [45] F. Ding, "Two-stage least squares based iterative estimation algorithm for CARARMA system modeling," *Applied Mathematical Modelling*, vol. 37, no. 7, pp. 4798–4808, 2013.
- [46] Y. Liu and R. Ding, "Consistency of the extended gradient identification algorithm for multi-input multi-output systems with moving average noises," *International Journal of Computer Mathematics*, vol. 90, no. 9, pp. 1840–1852, 2013.
- [47] Y. Zhang, "Unbiased identification of a class of multi-input single-output systems with correlated disturbances using bias compensation methods," *Mathematical and Computer Modelling*, vol. 53, no. 9–10, pp. 1810–1819, 2011.
- [48] Y. Zhang and G. Cui, "Bias compensation methods for stochastic systems with colored noise," *Applied Mathematical Modelling*, vol. 35, no. 4, pp. 1709–1716, 2011.
- [49] F. Ding, "Combined state and least squares parameter estimation algorithms for dynamic systems," *Applied Mathematical Modelling*, vol. 38, no. 1, pp. 403–412, 2014.
- [50] F. Ding, "Coupled-least-squares identification for multivariable systems," *IET Control Theory & Applications*, vol. 7, no. 1, pp. 68–79, 2013.
- [51] D. Wang and F. Ding, "Least squares based and gradient based iterative identification for Wiener nonlinear systems," *Signal Processing*, vol. 91, no. 5, pp. 1182–1189, 2011.
- [52] D. Q. Wang and F. Ding, "Hierarchical least squares estimation algorithm for Hammerstein-Wiener systems," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 825–828, 2012.
- [53] D. Wang, F. Ding, and Y. Chu, "Data filtering based recursive least squares algorithm for Hammerstein systems using the key-term separation principle," *Information Sciences*, vol. 222, pp. 203–212, 2013.
- [54] F. Ding, X. Liu, and J. Chu, "Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle," *IET Control Theory & Applications*, vol. 7, no. 2, pp. 176–184, 2013.
- [55] F. Ding, Y. Shi, and T. Chen, "Performance analysis of estimation algorithms of nonstationary ARMA processes," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1041–1053, 2006.
- [56] F. Ding, T. Chen, and L. Qiu, "Bias compensation based recursive least-squares identification algorithm for MISO systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 5, pp. 349–353, 2006.
- [57] X. Luan, P. Shi, and F. Liu, "Stabilization of networked control systems with random delays," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4323–4330, 2011.
- [58] X. Luan, S. Zhao, and F. Liu, " H_∞ control for discrete-time Markov jump systems with uncertain transition probabilities," *IEEE Transactions on Automatic Control*, vol. 58, no. 6, pp. 1566–1572, 2013.
- [59] H. Zhang and F. Ding, "On the Kronecker products and their applications," *Journal of Applied Mathematics*, vol. 2013, Article ID 296185, 8 pages, 2013.