Arab Journal of Mathematical Sciences (2011) **17**, 115–134

King Saud University
Arab Journal of Mathematical Sciences
www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

An efficient algorithm for solving extended Sylvester-conjugate transpose matrix equations ☆

Caiqin Song, Guoliang Chen *

Department of Mathematics, East China Normal University, Shanghai 200241, PR China

Received 26 February 2011; revised 21 March 2011; accepted 21 March 2011
Available online 8 April 2011

KEYWORDS
Extended Sylvester-conjugate transpose matrix equation;
Iterative algorithm;
Real representation;
Convergence;
Spectral norm

Abstract This note studies the iterative solutions to the extended Sylvester-conjugate transpose matrix equations with a unique solution. By using the hierarchical identification principle, an iterative algorithm is presented for solving this class of extended matrix equations. It is proved that the iterative solution consistently converges to the exact solution for any initial values. Meanwhile, by means of a real representation of a complex matrix, sufficient conditions are derived to guarantee that the iterative solutions given by the proposed algorithm converge to the exact solution for any initial matrices. Finally, a numerical example is given to illustrate the efficiency of the proposed approach.

☆ This project is granted financial support from NSFC (No. 11071079), NSFC(No. 10901056), the Fundamental Research Funds for the Central Universities.
* Corresponding author. Tel.: +86 13601659126.
E-mail address: glchen@math.ecnu.edu.cn (G. Chen).

Peer review under responsibility of King Saud University.
doi:10.1016/j.ajmsc.2011.03.003

Production and hosting by Elsevier

1. Introduction

Matrix equations are often encountered in systems and control, such as Lyapunov matrix equations, Sylvester matrix equations and so on. Traditional methods convert such matrix equations into their equivalent forms by using the Kronecker product, however, which involve the inversion of the associated large matrix and result in increasing computation and excessive computer memory. In the matrix algebra field, some complex matrix equations have attached much attention from many researchers since it was shown in Bevis et al. (1987) that the consistence of the matrix equation $AX - \overline{X}B = C$ was related to the consimilarity (Horn and John, 1990; Huang, 2001; Jiang et al., 2006) of two partitioned matrices associated with the matrices A, B and C. In the preceding matrix equations, $\overline{X}$ denotes the matrix obtained by taking the complex conjugate of each element of X. By consimilarity decomposition, explicit solutions can were established in Bevis et al. (1987) and Bevis et al. (1988). Recently, in Wu et al. (2006) some explicit expressions of the solution to the matrix equation $AX - \overline{X}B = C$ were established by means of real representation of a complex matrix, and it shown that there exists a unique solution if and only if $A\overline{A}$ and $B\overline{B}$ have no common eigenvalues. The explicit solution of the matrix equation $X - A\overline{X}B = C$ was proposed in Jiang and Wei (2003) with matrix polynomial as a tool. In Wu et al. (2010), the Homogeneous Sylvester-conjugate matrix equations $A\overline{X} + BY = XF$ and Nonhomogeneous Sylvester-conjugate matrix equations $A\overline{X} + BY = XF + R$ are investigated. Some explicit closed-form solutions of the above two matrix equations are provided. Very recently, in Wu et al. (2011) proposed a new operator of conjugate product for complex polynomial matrices. It is shown that an arbitrary complex polynomial matrix can be converted into the so-called Smith normal form by elementary transformations in the framework of conjugate product. Meanwhile, the conjugate product and the Sylvester-conjugate sum are also proposed by in Wu et al. (2011). Based on the important properties of the above new operators, a unified approach to solve a general class of Sylvester-polynomial-conjugate matrix equations is given. The complete solution of the Sylvester-polynomial-conjugate matrix equation is obtained.

Iterative approaches for solving matrix equations and recursive identifications have attached much attention from many researches since Huang et al. (2008) proposed an iterative method for solving the linear matrix equation $AXB = F$ over skew-symmetric matrix X. By extending the well-known Jacobi and Gauss Seidel iterations for $Ax = b$, Ding et al. (2008) derived iterative solutions of matrix equation $AXB = F$ and the generalized Sylvester matrix equation $AXB + CXD = F$. The gradient based iterative algorithm (Ding and Chen, 2006, Ding and Chen, 2005, Wu et al., 2010) and least squares based iterative algorithm (Ding and Chen, 2006) for solving (coupled) matrix equations are a novel and efficiently numerical algorithms were presented based on the hierarchical identification principle (Ding and Chen, 2005; Ding and Chen, 2005) which regards the unknown matrix as the system parameter matrix to be identified. In Wu et al. (2011), the matrix equation

$X - A\overline{X}B = C$ is considered. Some Smith-type iterative algorithms were established and the corresponding convergence analysis were also given in Wu et al. (2010, 2011) investigated the following so-called coupled Sylvester-conjugate matrix equations:

$$\sum_{\eta=1}^{p}(A_{i\eta}X_{\eta}B_{i\eta} + C_{i\eta}\overline{X_{\eta}}D_{i\eta}) = F_i, \quad i \in I[1.N]. \tag{1.1}$$

In Wu et al. (2010), by applying a hierarchical identification principle, an iterative algorithm is established and the iterative solutions with a unique solution is given. In Wu et al. (2011), by using a real inner product in complex matrix spaces, a solution can be obtained within finite iterative steps for any initial values in the absence of roundoff errors. In Li et al. (2010) and Xie et al. (2009), the following linear equations

$$\sum_{i=1}^{r} A_iXB_i + \sum_{j=1}^{s} C_jX^TD_j = E, \tag{1.2}$$

where $A_i, B_i, C_j, D_j, i = 1, \ldots, r; j = 1, \ldots, s.$ and E are some known constant matrices of appropriate dimensions and X is a matrix to be determined, was considered. In Wang et al. (2007), the special case of (1) $AXB + CX^TD = E$ was considered by the iterative algorithm. A more special case of (1), namely the matrix equation $AX + X^TB = C$, was investigated by Piao et al. (2007). The Moore-Penrose generalized inverse was used in Piao et al. (2007) to find explicit solutions to this matrix equation. Now in this paper we consider the following matrix equation

$$\sum_{i=1}^{r} A_iXB_i + \sum_{j=1}^{s} C_jX^HD_j = E, \tag{1.3}$$

where $A_i, B_i, C_j, D_j, i = 1, \ldots, r;\ j = 1, \ldots, s.$ and E are some known constant matrices of appropriate dimensions and X is a matrix to be determined, was considered. The current paper uses the iterative approach to study the iterative solutions of complex Sylvester matrix equations. The method in this paper differs from ones in Ding and Chen (2006), Ding and Chen (2005), Wu et al. (2010), Ding and Chen (2005) and Liang et al. (2007) because the proposed methods in this paper is for a complex matrixs using the hierarchical identification principle (Wu et al., 2010,) and some properties of the real representation of a complex matrix. These results, however, are difficult to be extended to the more general case (1).

The remainder of this paper is organized as follows. In Section 3, we introduce an iterative method for solving the iterative solution to the coupled Sylvester-conjugate transpose matrix equation, and give the convergence properties of this iterative algorithm. We also show that this iterative method can be used to a more general coupled Sylvester-conjugate transpose matrix equations in Section 4. In Section 5, we present a numerical example to verify our results. Conclusions will be put in Section 6.

Throughout this paper, we use $A^T, \overline{A}, A^H$, $\|A\|$ and $\|A\|_2$ to denote transpose, conjugate, conjugate transpose, the Frobenius norm and the spectral norm of A, respectively. $A \otimes B = (a_{ij}B)$ denotes the Kronecker product of two matrices A and B. For two integers $m < n$, the symbol $I[m, n]$ is used to denote the set $\{m, m+1, \ldots, n\}$. Let R denote the real number field, C the complex number field. For a matrix $X = [\, x_1 \quad x_2 \quad \ldots \quad x_n \,] \in C^{m\times n}$, $vec(X)$ is the column stretching operation of X, and defined as $vec(X) = \left[\, x_1^T \quad x_2^T \quad \ldots \quad x_n^T \,\right]^T$. In addition, it is obvious that $\|A\| = \|vec(A)\|$ for any matrix A.

2. Preliminaries

In this section, we provide some useful results which will play vital roles in the sequel section.

Referring to work Al Zhour and Kilicman (2007), let $P(m, n) \in R^{mn\times mn}$ be a square $mn \times mn$ matrix partitioned into $m \times n$ sub-matrices such that jith position and zeros elsewhere, i.e.,

$$P(m, n) = \sum_{i=1}^{m} \sum_{j=1}^{n} E_{ij} \otimes E_{ij}^T,$$

where $E_{ij} = e_i e_j^T$ called an elementary matrix of order $m \times 1(n \times 1)$. Using this definition, we have

$$\begin{aligned} vec(X^T) &= P(m, n)vec(X), P(m, n)P(n, m) = I_{mn}, P(m, n)^T = P(m, n)^{-1} \\ &= P(m, n). \end{aligned}$$

Next we give a real representation of complex matrix. This concept is firstly proposed in Jiang and Wei (2003). Let $A \in C^{m\times n}$, then A can be uniquely written as $A = A_1 + A_2 k$ with $A_1, A_2 \in R^{m\times n}, k = \sqrt{-1}$. Define real representation σ as

$$A_\sigma = \begin{bmatrix} A_1 & A_2 \\ A_2 & -A_1 \end{bmatrix} \in R^{2m\times 2n},$$

A_σ is called the real representation of the matrix A.

For an $n \times n$ complex matrix A, Define $A_\sigma^i = (A_\sigma)^i$, and

$$R_j = \begin{bmatrix} I_j & 0 \\ 0 & -I_j \end{bmatrix}, \quad Q_j = \begin{bmatrix} 0 & I_j \\ -I_j & 0 \end{bmatrix},$$

where I_j is the $j \times j$ identity matrix. The real representation possesses the following properties.

Lemma 1 Jiang and Wei, 2003.

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

$$(A + B)_\sigma = A_\sigma + B_\sigma, (aA)_\sigma = aA_\sigma, R_m A_\sigma R_n = \overline{A}_\sigma.$$

(2) If $A \in C^{m\times n},\ B \in C^{n\times r},\ C \in C^{r\times p}, a \in R$, *then*

$$(AB)_\sigma = A_\sigma R_n B_\sigma = A_\sigma \overline{B}_\sigma R_r, (ABC)_\sigma = A_\sigma \overline{B}_\sigma C_\sigma.$$

(3) If $A \in C^{m\times n}$, *then* $Q_m A_\sigma Q_n = A_\sigma$.
(4) If $A \in C^{m\times n}$, *then* $A_\sigma^H = R_n (A_\sigma)^T R_m$ *and* $A_\sigma^T = (A_\sigma)^T$.

In Lemma 1, Items (1)–(3) can be found in Jiang and Wei (2003). Item 4 can be obtained by simple computation.

Lemma 2 Wu et al., 2010. *Given a complex matrix,the following relations hold.*

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

The proof can be found in Wu et al., 2010.

Lemma 3 Wu et al., 2010. *Consider the following matrix equation*

$$AXB = F$$

where $A \in C^{m\times r},\ B \in C^{s\times n}$ *and* $F \in C^{m\times n}$ *are known matrices, and* $X \in C^{r\times s}$ *is the matrix to be determined. For this matrix equation, an iterative algorithm is constructed as*

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

with

$$0 < \mu < \frac{2}{\|A\|_2^2 \|B\|_2^2}.$$

If this matrix equation has a unique solution X_*, *then the iterative solution* $X(k)$ *converges to the unique solution* X_*, *that is* $\lim_{k\to\infty} X(k) = X_*$.

3. The matrix equation $AXB + CX^H D = F$

In this section, we consider the following extended Sylvester-conjugate transpose matrix equation

$$AXB + CX^H D = F, \tag{3.1}$$

where $A \in C^{m\times r}, B \in C^{s\times n}, C \in C^{m\times s}, D \in C^{r\times n},\ F \in C^{m\times n}$, are the given known matrices, and $X \in C^{r\times s}$ is the matrix to be determined. The hierarchical identification principle (Wu et al., 2010; Ding and Chen, 2005) implies that by defining two matrices

$$F_1 = F - CX^H D, \tag{3.2}$$

$$F_2 = F^H - (AXB)^H. \tag{3.3}$$

With the preceding definitions, the matrix Eq. (3.1) can be decomposed into the following matrix equations

$$AXB = F_1, \tag{3.4}$$

$$D^H X C^H = F_2. \tag{3.5}$$

According to Lemma 4, for these matrix equation one can construct the following respective iterative forms

$$X_1(k+1) = X_1(k) + \mu A^H (F_1 - AX_1(k)B)B^H, \tag{3.6}$$

$$X_2(k+1) = X_2(k) + \mu D(F_2 - D^H X_2(k)C^H)D. \tag{3.7}$$

Substituting (3.4) and (3.5) into (3.6) and (3.7), respectively, gives

$$X_1(k+1) = X_1(k) + \mu A^H [F - CX^H D - AX_1(k)B]B^H, \tag{3.8}$$

$$X_2(k+1) = X_2(k) + \mu D[F - AXB - CX_2^H(k)D]^H C. \tag{3.9}$$

The right-hand sides of these equations contain the unknown matrices matrices X, so it is impossible to realize these algorithms. In order to make the algorithm in (3.8) and (3.9) work, the unknown variable matrices X in (3.8) and (3.9) are respectively replaced with their estimates $X(k)$ by applying the hierarchical identification principle (Wu et al., 2010; Ding and Chen, 2005). Hence, one obtains the following recursive forms:

$$X_1(k+1) = X_1(k) + \mu A^H [F - CX_1^H(k)D - AX_1(k)B]B^H, \tag{3.10}$$

$$X_2(k+1) = X_2(k) + \mu D[F - AX_2(k)B - CX_2^H(k)D]^H C, \tag{3.11}$$

Taking the average of $X_1(k)$ and $X_2(k)$, we give the following iterative algorithm

$$X_1(k+1) = X(k) + \mu A^H [F - CX^H(k)D - AX(k)B]B^H, \tag{3.12}$$

$$X_2(k+1) = X(k) + \mu D[F - AX(k)B - CX^H(k)D]^H C, \tag{3.13}$$

$$X(k) = \frac{X_1(k) + X_2(k)}{2}. \tag{3.14}$$

This algorithm can be equivalently rewritten as

$$X(k+1) = X(k) + \frac{\mu}{2} A^H [F - AX(k)B - CX^H(k)D]B^H + \frac{\mu}{2} D[F - AX(k)B - CX^H(k)D]^H C. \tag{3.15}$$

In the following, we consider convergence properties of the proposed algorithm (3.15). During the proof of the convergence properties, it adopts the line of the one in Wu et al. (2010).

Theorem 1. *If the extended Sylvester-conjugate transpose matrix Eq. (3.1) has a unique solution X_*, then the iterative solution $X(k)$ given by the algorithm in 3.12,* *3.13, 3.14, or equivalently, algorithm (3.15), converges to X_* for arbitrary initial $X(0)$ if*

$$0 < \mu < \frac{4N}{\|[(R_n(B^H)_\sigma)^T \otimes (A^H)_\sigma R_m + ((C_\sigma)^T \otimes D_\sigma)P(m,n)]\|_2^2}. \tag{3.16}$$

Proof. Define error matrices

$$\widetilde{X}_i(k) = X_i(k) - X_*$$

Thus one has

$$\widetilde{X}_i(k) = X(k) - X_* = \frac{\widetilde{X}_1(k) + \widetilde{X}_2(k)}{2}$$

By using the algorithm (3.12) and (3.13), one has

$$\begin{aligned}\widetilde{X}_1(k+1) &= \widetilde{X}(k) + \mu A^H[F - AX(k)B - CX^H(k)D]B^H \\ &= \widetilde{X}(k) + \mu A^H[AX_*B + CX_*^H(k)D - AX(k)B - CX^H(k)D]B^H \\ &= \widetilde{X}(k) - \mu A^H[A\widetilde{X}(k)B + C\widetilde{X}^H(k)D]B^H\end{aligned}$$

and

$$\begin{aligned}\widetilde{X}_2(k+1) &= \widetilde{X}(k) + \mu D[F - AX(k)B - CX^H(k)D]^H C \\ &= \widetilde{X}(k) - \mu 1 D[A\widetilde{X}(k)B + C\widetilde{X}^H(k)D]^H C\end{aligned}$$

Denote

$$Z(k) = A\widetilde{X}(k)B + C\widetilde{X}^H(k)D. \tag{3.17}$$

Then, combining this relation with the preceding expression, one has

$$\widetilde{X}(k+1) = \widetilde{X}(k) - \frac{1}{2}\mu A^H Z(k)B^H - \frac{1}{2}\mu D Z^H(k)C$$

Recall the well known fact that $tr(AB) = tr(BA),\ tr(A+B) = tr(A) + tr(B)$. Then by simple computations it follows

$$\begin{aligned}&\|\widetilde{X}(k+1)\|^2 \\ &= tr[\widetilde{X}^H(k+1)\widetilde{X}(k+1)] = \|\widetilde{X}(k)\|^2 - \frac{1}{2}\mu tr[\widetilde{X}^H(k)A^H Z(k)B^H + DZ^H(k)C] \\ &\quad - \frac{1}{2}\mu tr[(BZ^H(k)A + C^H Z(k)D^H)\widetilde{X}(k)] + \frac{1}{4}\mu^2\|A^H Z(k)B^H + DZ^H(k)C\|^2 \\ &= \|\widetilde{X}(k)\|^2 - \frac{1}{2}\mu tr[B^H\widetilde{X}^H(k)A^H Z(k) - \frac{1}{2}\mu tr[Z^H(k)A\widetilde{X}(k)B] \\ &\quad - \frac{1}{2}\mu tr[Z(k)D^H\widetilde{X}(k)C^H] - \frac{1}{2}\mu tr[C\widetilde{X}^H(k)DZ^H(k)] - \frac{1}{4}\mu^2\|A^H Z(k)B^H + DZ^H(k)C\|^2\end{aligned}$$

It is obvious that $tr[C\widetilde{X}^H(k)DZ^H(k)] + tr[Z(k)D^H\widetilde{X}(k)C^H]$ is real. Then, by using Lemma 4 one has

$$\begin{aligned} & tr[C\widetilde{X}^H(k)DZ^H(k)] + tr[Z(k)D^H\widetilde{X}(k)C^H] \\ & \quad = tr[D^H\widetilde{X}(k)C^H Z(k)] + tr[Z^H(k)C\widetilde{X}^H(k)D] \end{aligned}$$

Combining this expression with (3.17), it is easily obtained that

$$\begin{aligned} & tr[B^H\widetilde{X}^H(k)A^H Z(k)] + tr[Z^H(k)A\widetilde{X}(k)B] + tr[C\widetilde{X}^H(k)DZ^H(k)] \\ & \quad + tr[Z(k)D^H\widetilde{X}(k)C^H] = tr[(B^H\widetilde{X}^H(k)A^H + D^H\widetilde{X}(k)C^H)Z(k)] \\ & \quad + tr[Z^H(k)(A\widetilde{X}(k)B + C\widetilde{X}^H(k)D)] = 2\|Z(k)\|^2. \end{aligned} \tag{3.19}$$

In addition, by using Lemma 3 and Lemma 2 one has

$$\begin{aligned} & \|A^H Z(k)B^H + DZ^H(k)C\|^2 \\ & \quad = \frac{1}{2}\|(A^H Z(k)B^H + DZ^H(k)C)_\sigma\|^2 \\ & \quad = \frac{1}{2}\|((A^H)_\sigma R_m (Z(k))_\sigma R_n (B^H)_\sigma + (D)_\sigma(((Z(k))_\sigma)^T C_\sigma)\|^2 \\ & \quad = \frac{1}{2}\|[(R_n(B^H)_\sigma)^T \otimes (A^H)_\sigma R_m + ((C_\sigma)^T \otimes D_\sigma)P(m,n)]vec((Z(k))_\sigma)\|^2 \\ & \quad \leqslant \frac{1}{2}\|[(R_n(B^H)_\sigma)^T \otimes (A^H)_\sigma R_m + ((C_\sigma)^T \otimes D_\sigma)P(m,n)]\|_2^2 \|vec((Z(k))_\sigma)\|^2 \\ & \quad = \|[(R_n(B^H)_\sigma)^T \otimes (A^H)_\sigma R_m + ((C_\sigma)^T \otimes D_\sigma)P(m,n)]\|_2^2 \|Z(k)\|^2. \end{aligned} \tag{3.20}$$

Denote

$$T = \|[(R_n(B^H)_\sigma)^T \otimes (A^H)_\sigma R_m + ((C_\sigma)^T \otimes D_\sigma)P(m,n)]\|_2^2, \tag{3.21}$$

Then combining (3.19), (3.20), (3.21) with (3.18), yields

$$\begin{aligned} \|\widetilde{X}(k+1)\|^2 & \leqslant \|\widetilde{X}(k)\|^2 - \mu\left(1 - \frac{1}{4}\mu T\right)(\|Z(k-1)\|^2 + \|Z(k)\|^2) \\ & \leqslant \|\widetilde{X}(0)\|^2 - \mu\left(1 - \frac{1}{4}\mu T\right)\sum_{i=0}^{k}\|Z(i)\|. \end{aligned} \tag{3.22}$$

If the parameter μ is chosen in (3.16), then one has

$$0 < \mu\left(1 - \frac{1}{4}\mu T\right)\sum_{i=0}^{k}\|Z(i)\| \leqslant \|\widetilde{X}(0)\|^2.$$

Then we have

$$0 < \mu\left(1 - \frac{1}{4}\mu T\right)\sum_{i=0}^{\infty}\|Z(i)\| \leqslant \|\widetilde{X}(0)\|^2.$$

It follows the convergence theorem of series that

$$\lim_{i\to\infty}\|Z(i)\| = 0,$$

Since the matrix Eq. (3.1) has a unique solution, it follows from the definition Eq. (3.17) of $Z(k)$ that

$$\lim_{i\to\infty}\widetilde{X}(i) = 0.$$

Thus we complete the proof. □

Remark 1. In the proof of Theorem 1, vector $vec((Z^H(k))_\sigma)$ has a special structure. This implies that the estimation (3.16) is a bit conservative. Therefore, the range of the parameter μ given in Theorem 1 may be a bit conservative. In other words, the algorithm (3.15) might be still convergent even if the parameter μ does not satisfy (3.16). It is our future work to reduce this conservatism.

In view of the expression of the algorithm (3.15) and the condition (3.16), we give the following corollary.

Corollary 1. If the extended Sylvester-conjugate transpose matrix Eq. (3.1) has a unique solution X_*, then the iterative solution $X(k)$ given by the algorithm (3.15) converge to X_* for arbitrary initial values $X(0)$ if

$$0 < \mu < \frac{1}{[\|A\|_2^2\|B\|_2^2 + \|C\|_2^2\|D\|_2^2]}. \tag{3.23}$$

Proof. Recall the fact that $\|(A\otimes B)P_{(m,n)}\|_2 = \|A\otimes B\|_2 = \|A\|_2\|B\|_2$, one has

$$\|A\otimes B\|_2 = \|A\|_2\|B\|_2.$$

Applying this fact and Lemma 3, one has

$$\begin{aligned}
&\|[(R_n(B^H)_\sigma)^T\otimes (A^H)_\sigma R_m + ((C_\sigma)^T\otimes D_\sigma)P(m,n)]\|_2^2\\
&\quad\leqslant (\|(R_n(B^H)_\sigma)^T\otimes (A^H)_\sigma R_m\|_2 + \|((C_\sigma)^T\otimes (D)_\sigma)P(m,n)\|_2)^2\\
&\quad= (\|(B^H)_\sigma\|_2\|A^H\|_2 + \|(C_\sigma)^T\|_2\|D_\sigma P(m,n)\|_2)^2.\\
&\quad= (\|A\|_2\|B\|_2 + \|C\|_2\|D\|_2)^2. \leqslant 2(\|A\|_2^2\|B\|_2^2 + \|C\|_2^2\|D\|_2^2).
\end{aligned}$$

Combining this relation with (3.15), gives the conclusion. □

Remark 2. Compared with Theorem 1, the range of the parameter μ guaranteeing the convergence of the algorithm (3.15) is given in terms of the original coefficient matrices instead of their real representation matrices. It is easy to compute than that of (3.16). Nevertheless, it is obvious that the result of Corollary 2is more conservative than the one of Theorem 1.

4. A more general extended Sylvester-conjugate transpose matrix equation

In this section, we consider a class of more general extended Sylvester-conjugate transpose matrix equations which include the extended Sylvester- conjugate transpose matrix Eq. (3.1) as a special case. Such a class of matrix equation is in the form of

$$\sum_{i=1}^{p} A_i X B_i + \sum_{j=1}^{q} C_j X^H D_j = F, \tag{4.1}$$

where $A_i \in C^{m\times r}, B_i \in C^{s\times n}, C_j \in C^{m\times s}, D_j \in C^{r\times n} i \in I[1,p], j \in I[1,q]$, $F \in C^{m\times n}$, are the given known matrices, and $X \in C^{r\times s}$ is the matrix to be determined. Now we define the following matrices (4.2) and (4.3).

$$F_i = F - \sum_{l=1}^{p} A_l X B_l - \sum_{j=1}^{q} C_{j\eta} X^H D_j + A_i X B_i, \quad i \in I[1,N], \tag{4.2}$$

$$F_{p+j} = \left[F - \sum_{i=1}^{p} A_i X B_i - \sum_{l=1}^{q} C_l X^H D_l\right]^H + D_j^H X C_j^H, \quad j \in I[1,q]. \tag{4.3}$$

According to the preceding definitions, the matrix Eq. (4.1) can be decomposed into the following matrix equations

$$A_i X B_i = F_i, \quad i \in I[1,p],$$
$$D_j^H X C_j^H = F_{p+j}, \quad j \in I[1,p].$$

According to Lemma 1, for these matrix equations one can construct the following respective iterative forms

$$X_i(k+1) = X_i(k) + \mu A_i^H (F_i - A_i X_i(k) B_i) B_i^H, \quad i \in I[1,p], \tag{4.4}$$

$$X_{p+j}(k+1) = X_{p+j}(k) + \mu D_j (F_{p+j} - D_j^H X_{p+j}(k) C_j^H) C_j, \quad j \in I[1,q]. \tag{4.5}$$

Substituting (4.2) and (4.3) into (4.4) and (4.5), respectively, gives

$$X_i(k+1) = X_i(k) + \mu A_i^H \left[F - \sum_{l=1}^{p} A_l X B_l - \sum_{l=1}^{q} C_l X^H D_l + A_i X B_i - A_i X_i(k) B_i\right] B_i^H,$$
$$i \in I[1,p], \tag{4.6}$$

$$X_{p+j}(k+1)$$
$$= X_{p+j}(k) + \mu D_j \left[F - \sum_{l=1}^{p} A_l X B_l - \sum_{l=1}^{q} C_l X^H D_l + C_j X^H D_j - C_j X_{p+j}^H(k) D_j\right]^H C_j,$$
$$j \in I[1,q]. \tag{4.7}$$

The right-hand sides of these equations contain the unknown matrices X, so it is impossible to realize these algorithms in (4.6) and (4.7). Similar to Section 3, by applying the hierarchical identification principle (Wu et al., 2010; Ding and Chen, 2005), the unknown variable matrices X in these two expressions is replaced with its estimate $X(k)$ at time k. Hence, one obtains the following recursive forms:

$$X_i(k+1)=X_i(k)+\mu A_i^H\left[F-\sum_{l=1}^{p}A_lX_i(k)B_l-\sum_{l=1}^{q}C_lX_i^H(k)D_l\right]B_i^H,\quad i\in I[1,p],$$

$$X_{p+j}(k+1)=X_{p+j}(k)+\mu D_j\left[F-\sum_{l=1}^{p}A_lX_{p+j}(k)B_l-\sum_{l=1}^{q}C_lX_{p+j}^H(k)D_l\right]^H C_j,\quad j\in I[1,q].$$

Taking the average of $X_i(k), i\in I[1,p+q]$, one obtain an iterative algorithm

$$X_i(k+1)=X_i(k)+\mu A_i^H\left[F-\sum_{l=1}^{p}A_lX_i(k)B_l-\sum_{l=1}^{q}C_lX_i^H(k)D_l\right]B_i^H,\quad i\in I[1,p], \tag{4.8}$$

$$X_{p+j}(k+1) = X_{p+j}(k) + \mu D_j\left[F-\sum_{l=1}^{p}A_lX_{p+j}(k)B_l-\sum_{l=1}^{q}C_lX_{p+j}^H(k)D_l\right]^H C_j,\quad j \in I[1,q], \tag{4.9}$$

$$X(k) = \frac{1}{p+q}\sum_{i=1}^{p+q}X_i(k). \tag{4.10}$$

This algorithm can be equivalently rewritten as

$$\begin{aligned}X(k+1) &= X(k) + \frac{\mu}{p+q}\\ &\quad\times\sum_{i=1}^{p}A_i^H\left[F-\sum_{l=1}^{p}A_lX(k)B_l-\sum_{l=1}^{q}C_lX^H(k)D_l\right]B_i^H\\ &\quad+\frac{\mu}{p+q}\sum_{j=1}^{q}D_j\left[F-\sum_{l=1}^{p}A_lX(k)B_l-\sum_{l=1}^{q}C_lX^H(k)D_l\right]^H C_j.\end{aligned} \tag{4.11}$$

Theorem 2. *If the general Sylvester-transpose matrix Eq. (4.1) has a unique solution X_*, then the iterative solution $X(k)$ given by the algorithm in (4.11) convergence to X_* for any initial matrix $X(0)$ if*

$$0<\mu<\frac{2(p+q)}{\left\|\sum_{i=1}^{p}(R_n(B_i^H)_\sigma)^T\otimes(A_i^H)_\sigma R_m+\sum_{j=1}^{q}(((C_j)_\sigma)^T\otimes(D_j)_\sigma)P(m,n)\right\|_2^2}. \tag{4.12}$$

Proof. Define error matrices

$$\widetilde{X}_i(k) = X_i(k) - X_{i*}, \quad i \in I[1, p+q].$$

Thus, one has

$$\widetilde{X}(k) = X(k) - X_* = \frac{\sum_{i=1}^{p+q} \widetilde{X}_i(k)}{p+q}$$

Using the algorithm (4.8) and (4.9), one has for $i \in I[1, p]$

$$\begin{aligned}
\widetilde{X}_i(k+1) &= \widetilde{X}(k) + \mu A_i^H \left[F - \sum_{l=1}^{p} A_l X(k) B_l - \sum_{l=1}^{q} C_l X^H(k) D_l \right] B_i^H \\
&= \widetilde{X}(k) + \mu A_i^H \left[\sum_{l=1}^{p} A_l X_* B_l + \sum_{l=1}^{q} C_l X_*^H D_l - \sum_{l=1}^{p} A_l X(k) B_l - \sum_{l=1}^{q} C_l X^H(k) D_l \right] B_i^H \\
&= \widetilde{X}(k) - \mu A_i^H \left[\sum_{l=1}^{p} A_l \widetilde{X}(k) B_l + \sum_{l=1}^{q} C_l \widetilde{X}^H(k) D_l \right] B_i^H.
\end{aligned}$$

and for $j \in I[1, q]$

$$\begin{aligned}
\widetilde{X}_{p+j}(k+1) &= \widetilde{X}(k) + \mu D_j \left[F - \sum_{l=1}^{p} A_l X(k) B_l - \sum_{l=1}^{q} C_l X^H(k) D_l \right]^H C_j \\
&= \widetilde{X}(k) - \mu D_j \left[\sum_{l=1}^{p} A_l \widetilde{X}(k) B_l + \sum_{l=1}^{q} C_l \widetilde{X}^H(k) D_l \right]^H C_j
\end{aligned}$$

Denote

$$Z(k) = \sum_{l=1}^{p} A_l \widetilde{X}(k) B_l + \sum_{l=1}^{q} C_l \widetilde{X}^H(k) D_l, \tag{4.13}$$

Then, combining this relation with the preceding expression, one has

$$\begin{aligned}
\|\widetilde{X}(k+1)\|^2 &= tr[\widetilde{X}^H(k+1) \widetilde{X}(k+1)] \\
&= \|\widetilde{X}(k)\|^2 - \frac{\mu}{p+q} tr\left[\widetilde{X}^H(k) \left(\sum_{i=1}^{p} A_i^H Z(k) B_i^H + \sum_{j=1}^{q} D_j Z^H(k) C_j \right) \right]
\end{aligned}$$

$$
\begin{aligned}
&-\frac{\mu}{p+q}tr\left[\left(\sum_{i=1}^{p}B_iZ^H(k)A_i+\sum_{j=1}^{q}C_j^HZ(k)D_j^H\right)\widetilde{X}(k)\right]\\
&+\frac{\mu^2}{(p+q)^2}\left\|\sum_{i=1}^{p}A_i^HZ(k)B_i^H+\sum_{j=1}^{q}D_jZ^H(k)C_j\right\|^2.\\
=&\|\widetilde{X}(k)\|^2-\frac{\mu}{p+q}tr\left[\sum_{i=1}^{p}B_i^H\widetilde{X}^H(k)A_i^HZ(k)\right]\\
&-\frac{\mu}{p+q}tr\left[\sum_{i=1}^{p}Z^H(k)A_i\widetilde{X}(k)B_i\right]-\frac{\mu}{p+q}tr\left[\sum_{j=1}^{q}Z^H(k)C_j\widetilde{X}^H(k)D_j\right]\\
&-\frac{\mu}{p+q}tr\left[\sum_{j=1}^{q}D_j^H\widetilde{X}(k)C_j^HZ(k)\right]+\frac{\mu^2}{(p+q)^2}\left\|\sum_{i=1}^{p}A_i^HZ(k)B_i^H\right.\\
&\left.+\sum_{j=1}^{q}D_jZ^H(k)C_j\right\|^2\\
=&\|\widetilde{X}(k)\|^2-\frac{\mu}{p+q}tr\left[\left(\sum_{i=1}^{p}B_i^H\widetilde{X}^H(k)A_i^H+\sum_{j=1}^{q}D_j^H\widetilde{X}(k)C_j^H\right)Z(k)\right]\\
&-\frac{\mu}{p+q}tr\left[Z^H(k)\left(\sum_{i=1}^{p}A_i\widetilde{X}(k)B_i+\sum_{j=1}^{q}C_j\widetilde{X}^H(k)D_j\right)\right]\\
&+\frac{\mu^2}{(p+q)^2}\left\|\sum_{i=1}^{p}A_i^HZ(k)B_i^H+\sum_{j=1}^{q}D_jZ^H(k)C_j\right\|^2\\
=&\|\widetilde{X}(k)\|^2-\frac{\mu}{p+q}tr[Z^H(k)Z(k)]-\frac{\mu}{p+q}tr[Z(k)Z^H(k)]\\
&+\frac{\mu^2}{(p+q)^2}\left\|\sum_{i=1}^{p}A_i^HZ(k)B_i^H+\sum_{j=1}^{q}D_jZ^H(k)C_j\right\|^2\\
=&\|\widetilde{X}(k)\|^2-\frac{2\mu}{p+q}\|Z(k)\|^2+\frac{\mu^2}{(p+q)^2}\left\|\sum_{i=1}^{p}A_i^HZ(k)B_i^H\right.\\
&\left.+\sum_{j=1}^{q}D_jZ^H(k)C_j\right\|^2
\end{aligned}
$$

In addition, by using Lemma 2 and 3 one has

$$\begin{aligned}
&\left\|\sum_{i=1}^{p}A_i^H Z(k)B_i^H+\sum_{j=1}^{q}D_j Z^H(k)C_j\right\|^2\\
&=\frac{1}{2}\left\|\left(\sum_{i=1}^{p}A_i^H Z(k)B_i^H+\sum_{j=1}^{q}D_j Z^H(k)C_j\right)_\sigma\right\|^2\\
&=\frac{1}{2}\left\|\sum_{i=1}^{p}(A_i^H)_\sigma R_m(Z(k))_\sigma R_n(B_i^H)_\sigma+\sum_{j=1}^{q}(D_j)_\sigma((Z(k))_\sigma)^T(C_j)_\sigma\right\|^2\\
&=\frac{1}{2}\left\|\sum_{i=1}^{p}(R_n(B_i^H)_\sigma)^T\otimes(A_i^H)_\sigma R_m+\sum_{j=1}^{q}(((C_j)_\sigma)^T\otimes(D_j)_\sigma)P(m,n)\|^2\|vec(Z(k))_\sigma\right\|^2\\
&\leqslant\frac{1}{2}\left\|\sum_{i=1}^{p}(R_n(B_i^H)_\sigma)^T\otimes(A_i^H)_\sigma R_m+\sum_{j=1}^{q}(((C_j)_\sigma)^T\otimes(D_j)_\sigma)P(m,n)\right\|_2^2\|vec(Z(k))_\sigma\|^2\\
&=\left\|\sum_{i=1}^{p}(R_n(B_i^H)_\sigma)^T\otimes(A_i^H)_\sigma R_m+\sum_{j=1}^{q}(((C_j)_\sigma)^T\otimes(D_j)_\sigma)P(m,n)\right\|_2^2\|Z(k)\|^2
\end{aligned}$$

Denote

$$T=\left\|\sum_{i=1}^{p}(R_n(B_i^H)_\sigma)^T\otimes(A_i^H)_\sigma R_m+\sum_{j=1}^{q}(((C_j)_\sigma)^T\otimes(D_j)_\sigma)P(m,n)\right\|_2^2$$

Then, combining the preceding three relations, yields

$$\begin{aligned}
\|\widetilde{X}(k+1)\|^2&\leqslant\|\widetilde{X}(k)\|^2-\frac{\mu}{p+q}(2-\frac{\mu}{p+q}T)\|Z(k)\|^2\\
&\leqslant\|\widetilde{X}(k-1)\|^2-\frac{\mu}{p+q}\left(2-\frac{\mu}{p+q}T\right)(\|Z(k-1)\|^2+\|Z(k)\|^2)\\
&\leqslant\|\widetilde{X}(0)\|^2-\frac{\mu}{p+q}\left(2-\frac{\mu}{p+q}T\right)\sum_{i=0}^{k}(\|Z(i)\|^2).
\end{aligned}$$

If the parameter μ is chosen in (4.12), then one has

$$0<\frac{\mu}{p+q}\left(2-\frac{\mu}{p+q}\right)\sum_{i=0}^{k}\|Z(i)\|\leqslant\|\widetilde{X}(0)\|^2.$$

Then we have

$$0<\frac{\mu}{p+q}\left(2-\frac{\mu}{p+q}\right)\sum_{i=0}^{\infty}\|Z(i)\|\leqslant\|\widetilde{X}(0)\|^2.$$

It follows the convergence theorem of series that

$$\lim_{i\to\infty} Z(i) = 0.$$

Since the matrix (4.1) has a unique solution, it follows from the definition (4.13) of $Z(k)$ that

$$\lim_{i\to\infty} X(i) = 0.$$

Thus we complete the proof. □

In view of the expression of the algorithm (4.11) and the condition (4.12), we give the following corollary.

Corollary 2. *If the general Sylvester-conjugate transpose matrix Eq. (4.1) has a unique solution X_*, then the iterative solution $X(k)$ given by*

$$\begin{aligned} X(k+1) &= X_l(k) \\ &\quad + \mu \sum_{i=1}^{p} A_i^H \left[F - \sum_{l=1}^{p} (A_l X(k) B_l - \sum_{l=1}^{q} C_l X^H(k) D_l) \right] B_i^H \\ &\quad + \mu \sum_{j=1}^{q} D_j \left[F - \sum_{l=1}^{p} (A_l X(k) B_l - \sum_{l=1}^{q} C_l X^H(k) D_l) \right]^H C_j, \end{aligned} \tag{4.15}$$

converge to the exact solution X_* for arbitrary initial matrices $X(0)$ if

$$\begin{aligned} 0 &< \mu \\ &< \frac{2}{\left\| \sum_{i=1}^{p} (R_n (B_i^H)_\sigma)^T \otimes (A_i^H)_\sigma R_m + \sum_{j=1}^{q} (((C_j)_\sigma)^T \otimes (D_j)_\sigma) P(m,n) \right\|_2^2}. \end{aligned} \tag{4.16}$$

Similar to the Section 1, at the end of this section, we also provide a sufficient condition that is easier to compute. The given condition does not invoke Kronecker product and the real representations of the coefficient matrices.

Corollary 3. *If the general Sylvester- conjugate transpose matrix Eq. (4.1) has a unique solution X_*, then the iterative solution $X(k)$ given by the algorithm (4.11) converge to X_* for arbitrary initial values $X(0)$ if*

$$0 < \mu < \frac{2}{\sum_{i=1}^{p} \|A_i\|_2^2 \|B_i\|_2^2 + \sum_{j=1}^{q} \|C_j\|_2^2 \|D_j\|_2^2}. \tag{4.17}$$

Proof. By applying Lemma 2 and Lemma 3, one has

$$\begin{aligned}
&\left\|\sum_{i=1}^{p}(R_n(B_i^H)_\sigma)^T\otimes(A_i^H)_\sigma R_m+\sum_{j=1}^{q}((C_j^T)_\sigma\otimes(D_j)_\sigma)P(m,n)\right\|_2^2\\
&\leqslant\left(\sum_{i=1}^{p}\|(R_n(B_i^H)_\sigma)^T\otimes(A_i^H)_\sigma R_m\|_2^2+\sum_{j=1}^{q}\|(C_j^T)_\sigma\otimes(D_j)_\sigma\|_2^2\right)^2\\
&=\left(\sum_{i=1}^{p}\|A_i\|_2\|B_i\|_2+\sum_{j=1}^{q}\|C_j\|_2\|D_j\|_2\right)^2\\
&\leqslant(p+q)(\sum_{i=1}^{p}\|A_i\|_2^2\|B_i\|_2^2+\sum_{j=1}^{q}\|C_j\|_2^2\|D_j\|_2^2)
\end{aligned}$$

Combining this relation with (4.17), gives the conclusion. □

Corollary 4. If the general Sylvester-conjugate transpose matrix Eq. (4.1) has a unique solution X_*, then the iterative solution $X(k)$ given by the algorithm (4.11) converge to X_* for arbitrary initial values $X(0)$ if

$$0<\mu<\frac{2}{p\sum_{i=1}^{p}\|A_i\|_2^2\|B_i\|_2^2+q\sum_{j=1}^{q}\|C_j\|_2^2\|D_j\|_2^2}.$$

Proof. It follows from the proof of the above corollary that

$$\begin{aligned}
&\left\|\sum_{i=1}^{p}(R_n(B_i^H)_\sigma)^T\otimes(A_i^H)_\sigma R_m+\sum_{j=1}^{q}((C_j^T)_\sigma\otimes(D_j)_\sigma)P(m,n)\right\|_2^2\\
&\leqslant\left(\sum_{i=1}^{p}\|A_i\|_2\|B_i\|_2+\sum_{j=1}^{q}\|C_j\|_2\|D_j\|_2\right)^2\\
&\leqslant 2\left(\sum_{i=1}^{p}\|A_i\|_2\|B_i\|_2\right)^2+2\left(\sum_{j=1}^{q}\|C_j\|_2\|D_j\|_2\right)^2\\
&\leqslant 2\left(p\sum_{i=1}^{p}\|A_i\|_2^2\|B_i\|_2^2\right)+2\left(q\sum_{j=1}^{q}\|C_j\|_2^2\|D_j\|_2^2\right).
\end{aligned}$$

□

5. Numerical example

In this section, we report two numerical examples to test the proposed iterative method.

Example 1. Consider the following extended Sylvester-conjugate transpose matrix equation

$$AXB + MXN + CX^H D = F$$

with the following parameter matrices:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & i \\ 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 0 & 1-i \end{bmatrix}, \quad D = \begin{bmatrix} 1-i & 1 \\ 2 & 4 \end{bmatrix},$$

$$M = \begin{bmatrix} 1 & 0 \\ 2 & i \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 0 \\ 2 & 4 \end{bmatrix}, \quad F = \begin{bmatrix} 5-i & 7+11i \\ 4+5i & -1+16i \end{bmatrix},$$

It can be verified that the above coupled matrix equation has a unique solution

$$X = \begin{bmatrix} 1 & i \\ 1-i & 0 \end{bmatrix}.$$

We apply algorithm (21) to compute the above extended Sylvester-conjugate transpose matrix equation. The initial matrices are chosen as $X(0) = 10^{-6} \times I_2$. According to Theorem 2, the algorithm (21) is convergent for $0 < \mu < 1.03 \times 10^{-2}$. According to Corollary 2, the algorithm (21) is convergent for $0 < \mu < 6.2 \times 10^{-3}$. Define the relative iterative error as

$$f(k) = \sqrt{\frac{\|X(k) - X\|^2 + \|Y(k) - Y\|^2}{\|X\|^2 + \|Y\|^2}}.$$

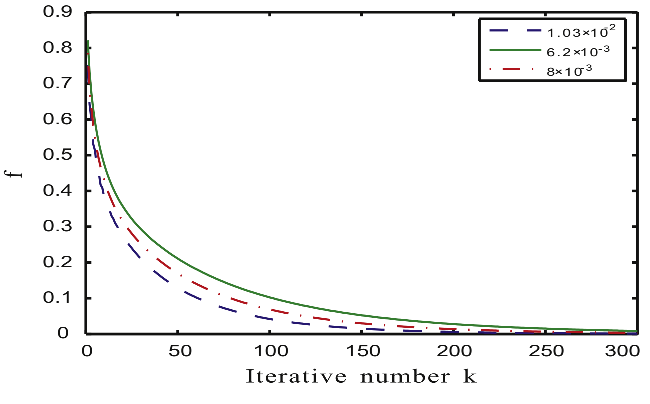

Figure 1 The convergence performance of algorithm (21).

Table 1 The iterative solution of Example for $\mu = 1.03 \times 10^{-2}$.

k	x_{11}	x_{12}	x_{21}	x_{22}	f
20	0.9680 − 0.1242i	−0.1960 + 0.9649i	0.7732 − 0.7217i	0.2714 − 0.1688i	0.2752
40	1.0025 − 0.1209i	−0.1310 + 0.9711i	0.8931 − 0.8462i	0.1515 − 0.1095i	0.1461
60	1.0002 − 0.0847i	−0.0834 + 0.9814i	0.9487 − 0.9014i	0.0891 − 0.0704i	0.1022
80	0.9985 − 0.0561i	−0.0527 + 0.9883i	0.9769 − 0.9337i	0.0531 − 0.0456i	0.0649
100	0.9981 − 0.0367i	−0.0335 + 0.9926i	0.9910 − 0.9547i	0.0319 − 0.0296i	0.0419
120	0.9983 − 0.0241i	−0.0214 + 0.9952i	0.9979 − 0.9689i	0.0192 − 0.0194i	0.0275
140	0.9985 − 0.0159i	−0.0137 + 0.9969i	1.0009 − 0.9786i	0.0117 − 0.0127i	0.0182
160	0.9988 − 0.0106i	−0.0089 + 0.9980i	1.0020 − 0.9852i	0.0071 − 0.0084i	0.0122
180	0.9991 − 0.0071i	−0.0058 + 0.9987i	1.0022 − 0.9898i	0.0044 − 0.0056i	0.0082
200	0.9993 − 0.0047i	−0.0038 + 0.9992i	1.0020 − 0.9930i	0.0027 − 0.0038i	0.0056
220	0.9995 − 0.0032i	−0.0025 + 0.9994i	1.0017 − 0.9951i	0.0017 − 0.0025i	0.0038
240	0.9996 − 0.0022i	−0.0017 + 0.9996i	1.0013 − 0.9966i	0.0010 − 0.0017i	0.0026
260	0.9997 − 0.0014i	−0.0011 + 0.9998i	1.0010 − 0.9977i	0.0006 − 0.0011i	0.0018
280	0.9998 − 0.0009i	−0.0007 + 0.9998i	1.0007 − 0.9985i	0.0004 − 0.0007i	0.0012
300	0.9999 − 0.0007i	−0.0005 + 0.9999i	1.0006 − 0.9989i	0.0003 − 0.0005i	0.0009
solution	1	i	1 − i	0	0.0009

From Fig 1, it is clear that the error f become smaller and go to zero as k increases. The effect of changing the convergence factor μ is illustrated in Fig 1. We can see that for $\mu = 1.02 \times 10^{-2}, 6.2 \times 10^{-3}, 8 \times 10^{-3}$, the larger the convergence factor μ, the faster the convergence rate. However, if we keep enlarging μ, the algorithm will diverge. How to choose a best convergence factor is still a project to be studied.

Algorithm (3.15) can also be constructed to the complex number equation. In the following, we give a numerical example (see Table 1).

6. Conclusions

The gradient based iterative algorithms for solving the coupled Sylvester-transpose matrix equation are studied by using the hierarchical identification principle. The basic idea of a hierarchical identification principle is to regard the unknown matrix as the system parameter matrix to be identified (Ding and Chen, 2005; Wu et al., 2010; Xie et al., 2009; Ding and Chen, 2005; Wu et al., 2010). We prove that the iterative solutions given by the proposed algorithms converge fast to their true solutions for any initial values and in the roundoff errors. We test the proposed algorithm using MATLAB and the results verify our theoretical findings. Sufficient conditions that guarantee the convergence of the proposed algorithm are given. The analysis indicates that the proposed convergence conditions may be conservative. Such statement is also confirmed by the given numerical example. It is our future work to establish a sufficient and necessary condition guaranteeing that the proposed iterative algorithm converges to the exact solution for any initial matrices. Furthermore, extending the adopted idea to study iterative solutions for nonlinear matrix equations and some more complicated linear matrix equations requires further research.

References

Al Zhour Z, Kilicman A. Some new connections between matrix products for partitioned and non-partitioned matrices. Comput Math Appl 2007;54(6):763–84.

Bevis JH, Hall FJ, Hartwing RE. Consimilarily and the matrix equation $A\overline{X} - XB = C$, in: current trends in matrix theory. New York: North-Holland; 1987.

Bevis JH, Hall FJ, Hartwig RE. The matrix equation $A\overline{X} - XB = C$ and its special cases. SIAM J Matrix Anal Appl 1988;9(3):348–59.

Ding F, Chen T. Iterative least squares solutions of coupled matrix equations. System Control Lett 2005;54:95–107.

Ding F, Chen T. Hierarchical gradient-based identification of multivariable discrete-time systems. Automatica 2005;41:315–25.

Ding F, Chen T. Hierarchical least squares identification methods for multivariable systems. IEEE Trans Autom Control 2005;50:397–402.

Ding F, Chen T. Gradient based iterative algorithms for solving a class of matrix equations. IEEE Trans Autom Control 2005;50:1216–21.

Ding F, Chen T. Hierarchical gradient-based identification of multivariable discrete-time systems. Automatica 2005;41:2269–84.

Ding F, Chen T. On iterative solutions of general coupled matrix equations. SIAM J Control Optim 2006;44:2269–84.

Ding F, Liu PX, Ding J. Iterative solutions of the generalized Sylvester matrix equation by using the hierarchical identification principle. Appl Math Comput 2008;197:41–50.

Horn RA, John CR. Matrix analysis. Cambridge: Cambridge University Press; 1990.

Huang LP. Consimilarity of quaternion matrices and complex matrices. Linear Algebra Appl 2001;331:21–30.

Huang GX, Yin F, Guo K. An iterative method for the skew-symmetric solution and the optimal approximate solution of the matrix equation $AXB = C$. J Comput Appl Math 2008;212:231–44.

Jiang TS, Wei MS. On solutions of $X - AXB = C$, and $X - A\overline{X}B = C$. Linear Algebra Appl 2003;367:225–33.

Jiang TS, Cheng XH, Chen L. An algebraic relation between consimilarity and similarity of complex matrices and its applications. J Phys A Math General 2006;39:9215–22.

Liang ML, You CH, Dai LF. An efficient algorithm for the generalized centro-symmetric solution of matrix equation $AXB = C$. Numer Algorithm 2007;4:173–84.

Li ZY, Wang Y, Zhou B, Duan GR. Least squares solution with the minimum-norm to general matrix equations via iteration. Appl Math Comput 2010;215:3547–62.

Piao F, Zhang Q, Wang Z. The solution to matrix equation $AX + X^TC = B$. J Franklin Inst 2007;344:1056–62.

Wang MH, Cheng XH, Wei MS. Iterative algorithms for solving the matrix equation $AXB + CX^TD = E$. Appl Math Comput 2007;187:622–9.

Wu AG, Duan GR, Yu HH. On solutions of $XF - AX = C$ and $XF - A\overline{X} = C$. Appl Math Comput 2006;182(2):932–41.

Wu AG, Feng G, Duan GR, Wu WJ. Iterative solutions of coupled Sylvester-conjugte matrix equations. Comput Math Appl 2010;60(1):54–66.

Wu AG, Zeng XL, Duan GR, Wu WJ. Iterative solutions to the extended Sylvester-conjugte matrix equations. Appl Math Comput 2010;217:130–42.

Wu AG, Feng G, Duan GR, Wu WJ. Iterative solutions to coupled Sylvester-conjugate matrix equations. Comput Math Appl 2010;60(1):54–66.

Wu AG, Feng G, Duan GR, Wu WJ. Closed-form solutions to Sylvester-conjugate matrix equations. Comput Math Appl 2010;60(1):95–111.

Wu AG, Liu WQ, Duan GR. On the conjugate product of matrix polynomial matrices. Math Comput Modell 2011;53:2031–43.

Wu AG, Feng G, Liu WQ, Duan GR. The complete solution to the Sylvester-polynomial-conjugate matrix equations. Math Comput Modell 2011;53:2044–56.

Wu AG, Feng G, Duan GR, Liu WQ. Iterative solutions to the Kalman–Yakubovich-conjugate matrix equation. Appl Math Comput 2011;217:4427–38.

Wu AG, Li B, Zhang Y, Duan GR. Finite iterative solutions to coupled Sylvester-conjugate matrix equations. Appl Math Modell 2011;35(3):1065–80.

Xie L, Ding J, Ding F. Gradient based iterative solutions for general linear matrix equations. Comput Math Appl 2009;58:1441–8.