MINIMUM NORM LEAST-SQUARES SOLUTION TO GENERAL COMPLEX COUPLED LINEAR MATRIX EQUATIONS VIA ITERATION

Davod Khojasteh Salkuyeh[*,§] **and Fatemeh Panjeh Ali Beik**[‡]

[§]*Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran*
e-mail: khojasteh@guilan.ac.ir; salkuyeh@gmail.com

[‡]*Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran*
e-mail: f.beik@vru.ac.ir; beik.fatemeh@gmail.com

Abstract. This paper deals with the problem of finding the minimum norm least-squares solution of a quite general class of coupled linear matrix equations defined over over field of complex numbers. To this end, we examine a gradient-based approach and present the convergence properties of the algorithm. The highlight of the elaborated results in the current work is using a new sight of view for construction of the gradient-based algorithm which turns out that we can ignore some of the limitations assumed by the authors in the recently published works for the application of the algorithm to obtain the solution of the referred problems. To the best of our knowledge, so far, computing the optimal convergence factor of the algorithm to determine the (least-squares) solution of general complex linear matrix equations has left as a project to be investigated. In the current work, we determine the optimal convergence factor of the algorithm. Some numerical experiments are reported to illustrate the validity of the presented results.

Keywords: Gradient-based, Linear matrix equation; Iterative method; Convergence; Least-squares solution.

AMS Subject Classification: Primely: 65A24 Secondary: 65F10.

1. Introduction and preliminaries

Throughout this paper, we use $\mathrm{tr}(A), A^T, \overline{A}, A^H$, $\mathcal{N}(A)$ to represent the trace, the transpose, the conjugate, the conjugate transpose and the null space of the matrix A, respectively. Furthermore, $\mathbb{C}^{m\times n}$ denotes the set of all $m\times n$ complex matrices. For an arbitrary complex number z, the real and imaginary parts of z are indicated by $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$, respectively. For an arbitrary $n\times p$ complex matrix $A=[a_{ij}]$, the matrices $\mathrm{Re}(A)$ and $\mathrm{Im}(A)$ are $n\times p$ real matrices specified by $\mathrm{Re}(A)=[\mathrm{Re}(a_{ij})]$ and $\mathrm{Im}(A)=[\mathrm{Im}(a_{ij})]$. For a given matrix $X\in\mathbb{C}^{n\times p}$, the notation $\mathrm{vec}(X)$ stands for a vector of dimension np obtained by stacking the columns of the matrix X. If $\mathcal{X}=\mathrm{vec}(X)$, then we define $\mathrm{unvec}(\mathcal{X})$ so that $X=\mathrm{unvec}(\mathcal{X})$. For an arbitrary square matrix Z, the symbols $\rho(Z)$ and $\sigma(Z)$ represent the spectral radius and the spectrum of the matrix of Z, respectively. For two given matrices $X\in\mathbb{C}^{n\times p}$ and $Y\in\mathbb{C}^{q\times l}$, the Kronecker product $X\otimes Y$ is the

[*] Corresponding author.

$nq \times pl$ matrix determined by $X \otimes Y = [X_{i,j}Y]$. In this paper, the following relation is utilized (See [5])

$$\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X),$$

in which A, B and X are given matrices with proper dimensions defined over field of complex (real) numbers. For a given matrices $Y \in \mathbb{C}^{n\times p}$, the well-known Frobenius norm is given by $\|Y\|^2 = \mathrm{Re}(\mathrm{tr}(Y^H Y))$. As a natural extension, the norm of $X = (X_1, X_2, \ldots, X_p)$ where $X_i \in \mathbb{C}^{n_i \times m_i}$ for $i = 1, 2, \ldots, p$ is defined by

$$\|X\|^2 = \|X_1\|^2 + \|X_2\|^2 + \ldots + \|X_p\|^2 .$$

The following result can be deduced by referring to Al-Zhour and Kilicman's work [1]; for more details see [22].

Lemma 1.1. *Let $X \in \mathbb{R}^{m\times n}$ be an arbitrary matrix. Then*

$$vec(X^T) = P(m,n)vec(X),$$

where $P(m,n) = (E_{ij}^T) \in \mathbb{R}^{(mn)\times(mn)}$ in which E_{ij} for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ is an $m \times n$ matrix with the element at position (i,j) being one and the others being zero.

The proof of the next theorem can be found in [6].

Theorem 1.2. *Assume that positive integers m, n, p and q are given and let $P(p,m)$ and $P(n,q)$ be defined as before. Then, $B \otimes A = P(m,p)^T (A \otimes B)P(n,q)$ for all $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$. Moreover, $P(m,n)$ is unitary and $P(m,n) = P(n,m)^T$.*

Consider the following coupled linear matrix equations

$$\sum_{i=1}^{p} \left(A_{\ell i} X_i B_{\ell i} + C_{\ell i} X_i^T D_{\ell i} + M_{\ell i} \overline{X}_i N_{\ell i} + H_{\ell i} X_i^H G_{\ell i} \right) = F_\ell, \tag{1.1}$$

in which $A_{\ell i} \in \mathbb{C}^{r_\ell \times n_i}, B_{\ell i} \in \mathbb{C}^{m_i \times k_\ell}$, $C_{\ell i} \in \mathbb{C}^{r_\ell \times m_i}, D_{\ell i} \in \mathbb{C}^{n_i \times k_\ell}, M_{\ell i} \in \mathbb{C}^{r_\ell \times n_i}$, $N_{\ell i} \in \mathbb{C}^{m_i \times k_\ell}, H_{\ell i} \in \mathbb{C}^{r_\ell \times m_i}, G_{\ell i} \in \mathbb{C}^{n_i \times k_\ell}$ and $F_\ell \in \mathbb{C}^{r_\ell \times k_\ell}$ $(\ell = 1, 2, \ldots, N)$ are given matrices.

Linear matrix equations play a cardinal role in control theory, signal processing, model reduction, image restoration, filtering theory for continuous or discrete-time large-scale dynamical systems, decoupling techniques for ordinary and partial differential equations, implementation of implicit numerical methods for ordinary differential equations, and block-diagonalization of matrices; for further details see [8, 9] and the references therein.

In the literature, the performance of several iterative methods to find (least-squares) solution of the (in)consistent linear matrix equations have been examined widely; for instance see [2, 3, 7, 8, 9, 10, 11, 12, 13, 14, 18, 16, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31] and the references therein. Most of the earlier cited works have been utilized the gradient-based approaches to resolve their mentioned problems. However, there is two kinds of restrictions in these works. Some of the refereed papers apply the gradient-based algorithms to obtain the solution of consistent (coupled) linear matrix equations as their main problem, the restriction in these works is the assumption that the mentioned problem has a unique solution. In other cited works, the gradient-based algorithms are exploited to find the solution of the least-squares problems corresponding to inconsistent (coupled) linear matrix equations. In these works, the authors first consider an equivalent least-squares problem associated with a linear system by using the "vec(.)" operator. The reduction in hypothesis of these works is that the coefficient matrix of the obtained linear system is supposed to be a full row (column) rank matrix. The limitations mentioned in these works motivate us to study the convergence of the gradient-based iterative algorithm in a such way that these curtailments can be ignored. To this end, we focus on the coupled linear matrix equations (1.1) which are entirely general and incorporate many of the recently mentioned (coupled) linear matrix equations. In addition, the optimal convergence factor of the gradient-based iterative algorithm for solving the mentioned problems is derived which has not been investigated so far.

In the sequel, we first reformulate (1.1) for more clarity. Afterwards, we briefly review some of the recently published research papers which their subjects are relevant to the current paper. In addition, we momentarily state our main contribution and shortly describe the outline of the rest of the paper. At the end of this section, the main problems are formulated.

For simplicity, we consider the linear operator $\mathcal{A}(X)$ which specified as follows:

$$\begin{aligned} \mathcal{A} : \mathbb{C}^{n_1\times m_1} \times \mathbb{C}^{n_2\times m_2} \times \ldots \times \mathbb{C}^{n_p\times m_p} &\rightarrow \mathbb{C}^{r_1\times k_1} \times \mathbb{C}^{r_2\times k_2} \times \ldots \times \mathbb{C}^{r_N\times k_N}, \\ X = (X_1, X_2, \ldots, X_p) &\mapsto \mathcal{A}(X) := (\mathcal{A}_1(X), \mathcal{A}_2(X), \ldots, \mathcal{A}_N(X)), \end{aligned}$$

where

$$\mathcal{A}_\ell(X) = \sum_{i=1}^{p} \left(A_{\ell i} X_i B_{\ell i} + C_{\ell i} X_i^T D_{\ell i} + M_{\ell i} \overline{X}_i N_{\ell i} + H_{\ell i} X_i^H G_{\ell i} \right), \qquad \ell = 1, 2, \ldots, N.$$

Thence, the coupled linear matrix equations (1.1) can be rewritten by

$$\mathcal{A}(X) = F, \tag{1.2}$$

in which $F = (F_1, F_2, \ldots, F_N)$.

We would like to comment here that Eq. (1.1) is very general and contains many of the recently investigated (coupled) linear matrix equations. For instance, Wu et al. [24] have focused on the following coupled Sylvester-conjugate matrix equations

$$\sum_{j=1}^{q} (A_{ij} X_j B_{ij} + C_{ij} \overline{X}_j D_{ij}) = F_i, \quad i = 1, 2, \ldots, p.$$

In the case that the above coupled linear matrix equations have a unique solution the gradient-based method is applied to solve it. In [22], the authors have offered the gradient-based method to solve the following coupled Sylvester-transpose matrix equations

$$\sum_{j=1}^{q} (A_{ij} X_j B_{ij} + C_{ij} X_j^T D_{ij}) = F_i, \quad i = 1, 2, \ldots, p,$$

under the assumption that the mentioned coupled linear matrix equations have a unique solution.

In [7], Dehghan and Hajarian have focused on the following coupled linear matrix equations

$$\begin{cases} A_1XB_1 + C_1X^TD_1 = M_1 \\ A_2XB_2 + C_2X^TD_2 = M_2 \end{cases}.$$

Under the hypothesis that the above coupled linear matrix equations have unique (anti-) reflexive solution, the authors have propounded two gradient-based iterative algorithms for solving the referred coupled matrix equations over reflexive and anti-reflexive matrices, respectively.

Presume that the subsequent matrix equation

$$\sum_{i=1}^{p} A_iXB_i + \sum_{i=1}^{q} C_iX^TD_i = F, \tag{1.3}$$

where $A_i \in \mathbb{R}^{r\times m}, B_i \in \mathbb{R}^{n\times s}, C_i \in \mathbb{R}^{r\times n}, D_i \in \mathbb{R}^{m\times s}, F \in \mathbb{R}^{r\times s}$. Under the assumption that (1.3) has unique solution, Wang and Liao [23] derived the optimal convergence factor of the gradient-based iterative algorithm to solve (1.3).

Lately, Hajarian and Dehghan [13] have considered the ensuing coupled linear matrix equations and examined a gradient-based algorithm

$$\sum_{t=1}^{\ell} E_{st}Y_tF_{st} = G_s, \quad s = 1, 2, \ldots, \ell,$$

to find the unique generalized reflexive matrix group $(Y_1, Y_2, \ldots, Y_\ell)$.

More recently, Hajarian [14] has developed a gradient-based iterative algorithm to solve the next coupled linear matrix equations

$$\begin{cases} A_1X_1B_1 + C_1X_2D_1 = E_1, \\ A_2X_1B_2 + C_2X_2D_2 = E_2, \end{cases} \tag{1.4}$$

to obtain the unique solution (X_1, X_2) where X_1 and X_2 are the generalized centro-symmetric matrices.

In [29, 30, 31, 32, 33], Zhou et al. have offered the gradient-based algorithms to resolve different kinds of (coupled) matrix equations. As known, the handled algorithms depend on a fixed parameter denoted by μ. In each of the earlier referred works, Zhou et al. have assumed that the considered problem has a unique solution. Afterward, a necessary and sufficient condition for the parameter μ has been given under which the proposed algorithm is convergent to the unique solution of the mentioned problem. Furthermore, the optimum value for the fixed parameter μ has been derived in these works. The presented results of this paper turn out that the restriction of the existence of the unique solution can be relaxed when a gradient-based algorithm is applied for solving (coupled) matrix equations. In addition, it reveals the formula for obtaining the optimum value of the μ is changed in general situations.

In [16], for computing a minimum norm least squares solution to

$$\sum_{i=1}^{r} A_iXB_i = C,$$

the authors have focused on the following problem

$$\alpha = \min_{X\in\mathbb{R}^{m\times n}} \left\{ \left\| \sum_{i=1}^{r} A_i X B_i - C \right\|_F \right\}. \tag{1.5}$$

Exploiting the "vec(.)" operator, an identical problem is mentioned which aims to determine X^* such that

$$f(X^*) = \min_{X\in\mathbb{R}^{m\times n}} \left\| \Upsilon \mathrm{vec}(X) - \mathrm{vec}(C) \right\|_2,$$

in which $\Upsilon = \sum_{i=1}^{r} (B_i^T \otimes A_i)$. In the case that (1.5) has a unique solution, a gradient-based algorithm has been examined. The convergence of the algorithm has been studied for the circumstance that Υ is a full row (column) rank matrix. The case that Υ is neither of full column rank nor of full row rank has been left as a project to be investigated. In [28, 29, 31, 30], the proposed gradient-based algorithms have been offered under the duplicate hypothesis mentioned by Li and Wang [16], i.e., the coefficient matrix appears after using the "vec(.)" operator is assumed to be full rank.

In [17], Li et al. have focused on finding the least norm solution of the following problem

$$\min_{X\in\mathbb{R}^{m\times n}} \left\| \sum_{i=1}^{r} A_i X B_i + \sum_{j=1}^{s} C_j X^T D_j - E \right\|_F, \tag{1.6}$$

where $E \in \mathbb{R}^{p\times q}$, $A_i \in \mathbb{R}^{p\times m}$, $B_i \in \mathbb{R}^{n\times q}$, $C_j \in \mathbb{R}^{p\times n}$ and $D_j \in \mathbb{R}^{m\times q}$ for $i = 1, 2, \ldots, r$ and $j = 1, 2, \ldots, s$ are known matrices and $X \in \mathbb{R}^{m\times n}$ is a matrix to be determined. In order to solve the mentioned problem, the authors have offered a gradient-based iterative algorithm. For studying the convergence properties of the proposed algorithm, (1.6) has been transformed to the next problem via the "vec(.)" operator

$$\min_{x\in\mathbb{R}^{mn}} \left\| \Upsilon x - \mathrm{vec}(E) \right\|_2,$$

where

$$\Upsilon = \sum_{i=1}^{r} \left(B_i^T \otimes A_i\right) + \sum_{j=1}^{s} \left(D_j^T \otimes C_j\right) P(m,n) \in \mathbb{R}^{pq\times mn},$$

in which $P(m,n)$ is a symmetric and unitary matrix satisfying $\mathrm{vec}(X^T) = P(m,n)\mathrm{vec}(X)$. Nevertheless, the convergence of the algorithm is established under the condition that $\mathrm{rank}(\Upsilon) = mn$, i.e., Υ has full column rank which is equivalent to say that $\Upsilon^T\Upsilon$ is nonsingular. The handled gradient-based algorithm relies on a fixed parameter, the optimum value of this parameter has been also derived under the assumption that Υ has full column rank. Here we would like to clarify that beside the fact that our mentioned problem incorporates the problem considered by Li et al. [17], we relax the restriction of the invertibility of $\Upsilon^T\Upsilon$ and investigate the convergence properties of the gradient-based algorithm for solving our problem. In addition, it reveals that the optimum value of the fixed parameter of the gradient-based algorithm is derived with a slightly different formula when Υ has not full column rank.

In the present work, we demonstrate that the gradient-based algorithm can be constructed in an alternative way. With the assistance of this new point of view, we capable to study the semi-convergence of the algorithm. We emphasize that, so far, the gradient-based iterative algorithm for solving (1.1) (and its special cases) has been presented under the restriction that the mentioned problem has unique solution (or the coefficient matrix of the linear system corresponding to theses matrix equations obtained after using "vec(.)" operator is full rank); for further details see [10, 8, 9, 12, 11, 28, 29, 31, 30, 32, 33] and the reference therein. We show that under a mild condition the assumed restrictions in the earlier cited works can be disregarded.

The rest of this paper is organized as follows. Before ending this section, we state the main problems. In Section 2, it has been discussed that how the Richardson iterative method can be applied to obtain the minimum norm (lease-squares) solution of (in)consistent linear system of equations. As the presented results in the second section is a complex version of the results which have been recently elaborated by Salkuyeh and Beik [21], we omit the details and believe that the required generalizations are straightforward. In Section 3, we demonstrate that the described results in the second section can be exploited to solve our mentioned problems. Numerical results are given in Section 4 which reveal the validity of the presented results. Finally, the paper is ended with a brief conclusion in Section 5.

1.1. **Problem reformulation.**

The current paper deals with solution of the following two problems.

Problem I. Suppose that the coupled linear matrix equations (1.2) are consistent. Find the solution $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_p)$ of (1.2) such that

$$\|\tilde{X}\| = \min\{\ \|X\| \mid \mathcal{A}(X) = F\ \}.$$

Problem II. Suppose that the coupled linear matrix equations (1.2) are inconsistent. Find $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_p)$ such that

$$\|\tilde{X}\| = \min\{\ \|\hat{X}\| \mid \hat{X} = \text{argmin}\ \|F - \mathcal{A}(X)\|\ \}.$$

At the end of this section, we would like to comment that our main contributions are studding the convergence of the gradient-based algorithm and deriving its best convergence factor to solve Problems I and II for more general cases which have not been studied so far.

2. Richardson method for normal equations

In this section we give a brief survey on the required theorems and properties. More precisely, we present an overview of the recently established results by Salkuyeh and Beik [21] which demonstrate the convergence properties of the Richardson method for solving the normal equations.

Consider the following linear system

$$Ax = b, \tag{2.1}$$

where $A \in \mathbb{C}^{m\times n}$ and $b \in \mathbb{C}^m$ are given and $x \in \mathbb{C}^n$ is the unknown. Here, we would like to comment that the coefficient matrix is not necessarily of full column (row) rank.

Theorem 2.1. [15, Chapter 8] *Suppose that $A \in \mathbb{C}^{m\times n}$, $b \in \mathbb{C}^m$ and*

$$\mathcal{X} = \{\ x \in \mathbb{C}^n \mid x = \underset{y\in\mathbb{C}^n}{\operatorname{argmin}} \|Ay - b\|_2\ \}.$$

Then $x \in \mathcal{X}$ if and only if $A^HAx = A^Hb$. Moreover, $x^* = A^+b$ is the unique solution of the problem

$$\min_{x\in\mathcal{X}} \|x\|_2,$$

where A^+ is the pseudoinverse of A.

The linear system $A^HAx = A^Hb$ is known as the normal equations. The vector $x^* = A^+b$ in Theorem 2.1 is called the minimum norm solution.

Remark 2.2. *Presume that $A \in \mathbb{C}^{m\times n}$ and $b \in \mathbb{C}^m$. It is known that if $x = A^+b + y$ where $y \in \mathcal{N}(A)$ then the following two statements hold.*

- *For the consistent linear system $Ax = b$, the vector x is a solution of the linear system $Ax = b$.*
- *In the case that the linear system $Ax = b$ is not consistent, the vector x is a solution of the least-squares problem $\min_{x\in\mathbb{C}^m} \|b - Ax\|_2$.*

Thence, x^* is the unique minimum norm solution if and only if $x^* \in \mathrm{Range}(A^+)$. Invoking the fact that $\mathrm{Range}\ (A^+) = \mathrm{Range}\ (A^H)$, we conclude that x^* is the unique minimum norm solution iff $x^* \in \mathrm{Range}\ (A^H)$; for more details see [15].

Let us apply the Richardson iterative method [20] to solve the normal linear system $A^HAx = A^Hb$ as follows:

$$x(k+1) = x(k) + \mu A^H(b - Ax(k)) = Hx(k) + \mu A^Hb, \tag{2.2}$$

where $H = I - \mu A^HA$ and μ is a positive real number.

Now, we state the ensuing theorem concerning the convergence of the iterative method (2.2). As the theorem can be established with an analogous strategy applied in [21], we omit its proof.

Theorem 2.3. *Assume that*

$$0 < \mu < \frac{2}{\sigma_{\max}^2(A)}, \tag{2.3}$$

where $\sigma_{\max}$ is the largest singular value of A. Then, the iterative method (2.2) converges to a solution of the normal equations $A^HAx = A^Hb$ for any initial guess $x(0)$. In the case that $x(0) \in \mathrm{Range}(A^H)$, the iterative method (2.2) converges to $x^* = A^+b$. Furthermore, the optimal value of μ is given by

$$\mu_{opt} = \frac{2}{\bar{\sigma}_{\min}^2(A) + \sigma_{\max}^2(A)},$$

where $\bar{\sigma}_{\min}$ is the smallest nonzero singular value of A.

Remark 2.4. *We would like to point here that if the matrix A in the assumption of Theorem 2.3 has full column rank then $\bar{\sigma}_{\min} = \sigma_{\min}(A)$. In this case the optimum value of μ is determined by*

$$\mu_{opt} = \frac{2}{\sigma_{\min}^2(A) + \sigma_{\max}^2(A)},$$

which has been originally derived by Zhou et al. [29].

Remark 2.5. *By Theorems 2.1 and 2.3, we may conclude the following two statements:*

- *If the linear system (2.1) is consistent, i.e., $b \in \mathrm{Range}(A)$, the iterative method (2.2) converges to a solution of (2.1).*
- *If the linear system (2.1) is inconsistent, i.e., $b \notin \mathrm{Range}(A)$, the iterative method (2.2) converges to a solution of the least-squares problem*

$$\min_{x\in\mathbb{R}^n} \|b - Ax\|_2 .$$

Let us assume that $x(j) \in \mathrm{Range}(A^H)$, i.e., $x(j) = A^H w$ for some $w \in \mathbb{C}^m$. Therefore, for the $(j+1)$th approximate solution obtained by the the iterative method (2.2), we have

$$\begin{aligned} x(j+1) &= (I - \mu A^H A)x(j) + \mu A^H b \\ &= (I - \mu A^H A)A^H w + \mu A^H b \\ &= A^H((I - \mu A A^H)w + \mu b) \in \mathrm{Range}(A^H). \end{aligned}$$

By Remark 2.2, it can be deduced that the minimum norm solution is obtained by proper selection of the initial guess $x(0)$ so that $x(0) \in \mathrm{Range}(A^H)$ for the two mentioned different cases in Remark 2.5. Without loss of generality, we may set $x(0) = 0$.

3. An iterative algorithm for the solution of Problem I (II)

In continuation, we concentrate on the solution of Problems I and II. Note that using the "vec(.)" operator, we may equivalently rewrite (1.2) into the following linear system

$$\mathcal{M}_1\mathcal{X} + \mathcal{M}_2\mathcal{X} + \mathcal{M}_3\overline{\mathcal{X}} + \mathcal{M}_4\overline{\mathcal{X}} = \mathcal{F}, \tag{3.1}$$

where

$$\mathcal{X} = \begin{pmatrix} \mathrm{vec}(X_1) \\ \vdots \\ \mathrm{vec}(X_p) \end{pmatrix}, \qquad \mathcal{F} = \begin{pmatrix} \mathrm{vec}(F_1) \\ \vdots \\ \mathrm{vec}(F_N) \end{pmatrix},$$

$$\mathcal{M}_1 = \begin{pmatrix} B_{11}^T \otimes A_{11} & \cdots & B_{1p}^T \otimes A_{1p} \\ B_{21}^T \otimes A_{21} & \cdots & B_{2p}^T \otimes A_{2p} \\ \vdots & \ddots & \vdots \\ B_{N1}^T \otimes A_{N1} & \cdots & B_{Np}^T \otimes A_{Np} \end{pmatrix},$$

$$\mathcal{M}_3 = \begin{pmatrix} N_{11}^T \otimes M_{11} & \cdots & N_{1p}^T \otimes M_{1p} \\ N_{21}^T \otimes M_{21} & \cdots & N_{2p}^T \otimes M_{2p} \\ \vdots & \ddots & \vdots \\ N_{N1}^T \otimes M_{N1} & \cdots & N_{Np}^T \otimes M_{Np} \end{pmatrix},$$

$$\mathcal{M}_2 = \begin{pmatrix} (D_{11}^T \otimes C_{11})P(n_1,m_1) & \cdots & (D_{1p}^T \otimes C_{1p})P(n_p,m_p) \\ (D_{21}^T \otimes C_{21})P(n_1,m_1) & \cdots & (D_{2p}^T \otimes C_{2p})P(n_p,m_p) \\ \vdots & \ddots & \vdots \\ (D_{N1}^T \otimes C_{N1})P(n_1,m_1) & \cdots & (D_{Np}^T \otimes C_{Np})P(n_p,m_p) \end{pmatrix},$$

and

$$\mathcal{M}_4 = \begin{pmatrix} (G_{11}^T \otimes H_{11})P(n_1,m_1) & \cdots & (G_{1p}^T \otimes H_{1p})P(n_p,m_p) \\ (G_{21}^T \otimes H_{21})P(n_1,m_1) & \cdots & (G_{2p}^T \otimes H_{2p})P(n_p,m_p) \\ \vdots & \ddots & \vdots \\ (G_{N1}^T \otimes H_{N1})P(n_1,m_1) & \cdots & (G_{Np}^T \otimes H_{Np})P(n_p,m_p) \end{pmatrix}.$$

Consequently, we may consider the next two equivalent problems instead of our mentioned two main problems. That is, we focus on Problems III and IV instead of Problems I and II, respectively.

Problem III. Suppose that the linear system (3.1) is consistent. Find the solution $\tilde{\mathcal{X}}$ of (3.1) such that

$$\|\tilde{\mathcal{X}}\| = \min\{\ \|\mathcal{X}\| \mid \mathcal{M}_1\mathcal{X} + \mathcal{M}_2\mathcal{X} + \mathcal{M}_3\overline{\mathcal{X}} + \mathcal{M}_4\overline{\mathcal{X}} = \mathcal{F}\ \}.$$

Problem IV. Suppose that the linear system (3.1) is inconsistent. Find $\tilde{\mathcal{X}}$ such that

$$\|\tilde{\mathcal{X}}\|_2 = \min\{\ \|\hat{\mathcal{X}}\| \mid \hat{\mathcal{X}} = \text{argmin}\ \|\mathcal{F} - \left(\mathcal{M}_1\mathcal{X} + \mathcal{M}_2\mathcal{X} + \mathcal{M}_3\overline{\mathcal{X}} + \mathcal{M}_4\overline{\mathcal{X}}\right)\|\}.$$

Hence, we first present an iterative method based on the Richardson method for computing the solution of Problem III (IV). Then the derived method is converted to an identical iterative method for solving Problem I (II).

The (i,j)th block of $\mathcal{M}_\mu$ is denoted by $(\mathcal{M}_\mu)_{ij}$ ($\mu = 1,2,3,4$) and given as follows:

$$(\mathcal{M}_1)_{ij} = B_{ij}^T \otimes A_{ij}, \quad (\mathcal{M}_2)_{ij} = (D_{ij}^T \otimes C_{ij})P(n_j,m_j),$$

$$(\mathcal{M}_3)_{ij} = N_{ij}^T \otimes M_{ij}, \quad \text{and} \quad (\mathcal{M}_4)_{ij} = (G_{ij}^T \otimes H_{ij})P(n_j,m_j).$$

Using the properties of the Kronecker product and Theorem 1.2, we get

$$\begin{aligned} (\mathcal{M}_2)_{ij} &= (D_{ij}^T \otimes C_{ij})P(n_j,m_j) \\ &= P(r_i,k_i)^T(C_{ij} \otimes D_{ij}^T)P(m_j,n_j)P(n_j,m_j) \\ &= P(r_i,k_i)^T(C_{ij} \otimes D_{ij}^T), \end{aligned}$$

and similarly $(\mathcal{M}_4)_{ij} = P(r_i,k_i)^T(H_{ij} \otimes G_{ij}^T)$.

For a given arbitrary matrix group $W = [W_1, W_2, \ldots, W_N]$, we set

$$\mathcal{W} = \left(\text{vec}(W_1)^T, \text{vec}(W_2)^T, \ldots, \text{vec}(W_N)^T\right)^T.$$

It can be verified that

$$(\mathcal{M}_1^H\mathcal{W})_i = \mathrm{vec}\left(\sum_{\ell=1}^{N} A_{\ell i}^H W_\ell B_{\ell i}^H\right), \tag{3.2}$$

$$(\mathcal{M}_2^H\mathcal{W})_i = \mathrm{vec}\left(\sum_{\ell=1}^{N} \overline{D}_{\ell i} W_\ell^T \overline{C}_{\ell i}\right), \tag{3.3}$$

$$(\mathcal{M}_3^H\mathcal{W})_i = \overline{\mathrm{vec}\left(\sum_{\ell=1}^{N} M_{\ell i}^T \overline{W}_\ell N_{\ell i}^T\right)}, \tag{3.4}$$

$$(\mathcal{M}_4^H\mathcal{W})_i = \overline{\mathrm{vec}\left(\sum_{\ell=1}^{N} G_{\ell i} W_\ell^H H_{\ell i}\right)}. \tag{3.5}$$

Let us rewrite (3.1) in the real representation. As a matter of fact, the complex linear system (3.1) is equivalent to the following linear system defined over real number field

$$\mathcal{U}\begin{pmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \end{pmatrix} = \begin{pmatrix} \mathcal{F}_1 \\ \mathcal{F}_2 \end{pmatrix}, \tag{3.6}$$

where

$$\mathcal{U} = \begin{pmatrix} \mathrm{Re}(\mathcal{M}_1+\mathcal{M}_2)+\mathrm{Re}(\mathcal{M}_3+\mathcal{M}_4) & -\mathrm{Im}(\mathcal{M}_1+\mathcal{M}_2)+\mathrm{Im}(\mathcal{M}_3+\mathcal{M}_4) \\ \mathrm{Im}(\mathcal{M}_1+\mathcal{M}_2)+\mathrm{Im}(\mathcal{M}_3+\mathcal{M}_4) & \mathrm{Re}(\mathcal{M}_1+\mathcal{M}_2)-\mathrm{Re}(\mathcal{M}_3+\mathcal{M}_4) \end{pmatrix},$$

and $\mathcal{X}_1 = \mathrm{Re}(\mathcal{X})$, $\mathcal{X}_2 = \mathrm{Im}(\mathcal{X})$, $\mathcal{F}_1 = \mathrm{Re}(\mathcal{F})$ and $\mathcal{F}_2 = \mathrm{Im}(\mathcal{F})$.

It is not difficult to see that

$$\mathcal{U}^T = \begin{pmatrix} \mathrm{Re}(\mathcal{M}_1^H+\mathcal{M}_2^H)+\mathrm{Re}(\overline{\mathcal{M}_3^H+\mathcal{M}_4^H}) & -\mathrm{Im}(\mathcal{M}_1^H+\mathcal{M}_2^H)+\mathrm{Im}(\overline{\mathcal{M}_3^H+\mathcal{M}_4^H}) \\ \\ \mathrm{Im}(\mathcal{M}_1^H+\mathcal{M}_2^H)+\mathrm{Im}(\overline{\mathcal{M}_3^H+\mathcal{M}_4^H}) & \mathrm{Re}(\mathcal{M}_1^H+\mathcal{M}_2^H)-\mathrm{Re}(\overline{\mathcal{M}_3^H+\mathcal{M}_4^H}) \end{pmatrix}.$$

For an arbitrary real vector of the form

$$\mathcal{R} = \begin{pmatrix} \mathcal{R}_1 \\ \mathcal{R}_2 \end{pmatrix},$$

such that $\mathcal{R}_i = \left((\mathcal{R}_i^{(1)})^T, (\mathcal{R}_i^{(2)})^T, \ldots, (\mathcal{R}_i^{(N)})^T\right)^T$ and $\mathcal{R}_i^{(j)} \in \mathbb{R}^{r_j k_j}$ for $i = 1, 2$ and $j = 1, 2, \ldots, N$. Our object is to determine $\mathcal{U}^T\mathcal{R}$, i.e.,

$$\begin{pmatrix} \mathcal{Y}_1 \\ \mathcal{Y}_2 \end{pmatrix} = \mathcal{U}^T\mathcal{R}.$$

In the sequel, assume that $R_i = \left(R_i^{(1)}, R_i^{(2)}, \ldots, R_i^{(N)}\right)$ such that $R_i^{(j)} := \text{unvec }(\mathcal{R}_i^{(j)})$ for $i = 1, 2$ and $j = 1, 2, \ldots, N$. By Eqs (3.2)-(3.5), we may derive $\mathcal{Y}_1$ and $\mathcal{Y}_2$ as follows:

$$\begin{aligned}
\mathcal{Y}_1 = {} & \text{Re}(\text{vec}\left(\sum_{\ell=1}^{N} A_{\ell i}^H R_1^{(\ell)} B_{\ell i}^H\right)) + \text{Re}(\text{vec}\left(\sum_{\ell=1}^{N} \overline{D}_{\ell i}(R_1^{(\ell)})^T \overline{C}_{\ell i}\right)) \\
& + \text{Re}(\text{vec}\left(\sum_{\ell=1}^{N} M_{\ell i}^T R_1^{(\ell)} N_{\ell i}^T\right)) + \text{Re}(\text{vec}\left(\sum_{\ell=1}^{N} G_{\ell i}(R_1^{(\ell)})^T H_{\ell i}\right)) \\
& - \text{Im}(\text{vec}\left(\sum_{\ell=1}^{N} A_{\ell i}^H R_2^{(\ell)} B_{\ell i}^H\right)) - \text{Im}(\text{vec}\left(\sum_{\ell=1}^{N} \overline{D}_{\ell i}(R_2^{(\ell)})^T \overline{C}_{\ell i}\right)) \\
& + \text{Im}(\text{vec}\left(\sum_{\ell=1}^{N} M_{\ell i}^T R_2^{(\ell)} N_{\ell i}^T\right)) + \text{Im}(\text{vec}\left(\sum_{\ell=1}^{N} G_{\ell i}(R_2^{(\ell)})^T H_{\ell i}\right)),
\end{aligned} \tag{3.7}$$

and

$$\begin{aligned}
\mathcal{Y}_2 = {} & \text{Im}(\text{vec}\left(\sum_{\ell=1}^{N} A_{\ell i}^H R_1^{(\ell)} B_{\ell i}^H\right)) + \text{Im}(\text{vec}\left(\sum_{\ell=1}^{N} \overline{D}_{\ell i}(R_1^{(\ell)})^T \overline{C}_{\ell i}\right)) \\
& + \text{Im}(\text{vec}\left(\sum_{\ell=1}^{N} M_{\ell i}^T R_1^{(\ell)} N_{\ell i}^T\right)) + \text{Im}(\text{vec}\left(\sum_{\ell=1}^{N} G_{\ell i}(R_1^{(\ell)})^T H_{\ell i}\right)) \\
& + \text{Re}(\text{vec}\left(\sum_{\ell=1}^{N} A_{\ell i}^H R_2^{(\ell)} B_{\ell i}^H\right)) + \text{Re}(\text{vec}\left(\sum_{\ell=1}^{N} \overline{D}_{\ell i}(R_2^{(\ell)})^T \overline{C}_{\ell i}\right)) \\
& - \text{Re}(\text{vec}\left(\sum_{\ell=1}^{N} M_{\ell i}^T R_2^{(\ell)} N_{\ell i}^T\right)) - \text{Re}(\text{vec}\left(\sum_{\ell=1}^{N} G_{\ell i}(R_2^{(\ell)})^T H_{\ell i}\right)).
\end{aligned} \tag{3.8}$$

Now Eqs (3.7) and (3.8) implies that

$$\begin{aligned}
\mathcal{Y} := \mathcal{Y}_1 + i\mathcal{Y}_2 = {} & \text{vec}\left(\sum_{\ell=1}^{N} A_{\ell i}^H (R_1^{(\ell)} + iR_2^{(\ell)}) B_{\ell i}^H\right) + \text{vec}\left(\sum_{\ell=1}^{N} \overline{D}_{\ell i}(R_1^{(\ell)} + iR_2^{(\ell)})^T \overline{C}_{\ell i}\right) \\
& + \text{vec}\left(\sum_{\ell=1}^{N} M_{\ell i}^T \overline{(R_1^{(\ell)} + iR_2^{(\ell)})} N_{\ell i}^T\right) + \text{vec}\left(\sum_{\ell=1}^{N} G_{\ell i}\left(\overline{(R_1^{(\ell)} + iR_2^{(\ell)})}\right)^T H_{\ell i}\right).
\end{aligned} \tag{3.9}$$

For solving Problem III and IV, we apply the Richardson iterative method to resolve the normal equations associated with (3.6) as follows:

$$\begin{pmatrix} \mathcal{X}_1(k+1) \\ \mathcal{X}_2(k+1) \end{pmatrix} = \begin{pmatrix} \mathcal{X}_1(k) \\ \mathcal{X}_2(k) \end{pmatrix} + \mu\, \mathcal{U}^T \left(\begin{pmatrix} \mathcal{F}_1 \\ \mathcal{F}_2 \end{pmatrix} - \mathcal{U} \begin{pmatrix} \mathcal{X}_1(k) \\ \mathcal{X}_2(k) \end{pmatrix} \right), \quad k = 0, 1, 2, \ldots\ .$$

Or equivalently, we may rewrite the recursive formula in the ensuing identical form

$$\begin{pmatrix} \mathcal{X}_1(k+1) \\ \mathcal{X}_2(k+1) \end{pmatrix} = \begin{pmatrix} \mathcal{X}_1(k) \\ \mathcal{X}_2(k) \end{pmatrix} + \mu\, \mathcal{U}^T \begin{pmatrix} \mathcal{R}_1(k) \\ \mathcal{R}_2(k) \end{pmatrix}, \quad k = 0, 1, 2, \ldots, \tag{3.10}$$

where straightforward computations reveal that

$$\mathcal{R}_1(k) = \mathrm{Re}(\mathcal{F} - \left(\mathcal{M}_1\mathcal{X}(k) + \mathcal{M}_2\mathcal{X}(k) + \mathcal{M}_3\overline{\mathcal{X}}(k) + \mathcal{M}_4\overline{\mathcal{X}}(k)\right)),$$

and

$$\mathcal{R}_2(k) = \mathrm{Im}(\mathcal{F} - \left(\mathcal{M}_1\mathcal{X}(k) + \mathcal{M}_2\mathcal{X}(k) + \mathcal{M}_3\overline{\mathcal{X}}(k) + \mathcal{M}_4\overline{\mathcal{X}}(k)\right)).$$

By (3.9) and the recursive formula (3.10), we find the following recursive formula for solving Problems I and II,

$$\begin{aligned} X_\ell(k+1) &= X_\ell(k) + \mu \left(\sum_{\ell=1}^{N} A_{\ell i}^H R_\ell(k) B_{\ell i}^H + \overline{D}_{\ell i} \left(R_\ell(k)\right)^T \overline{C}_{\ell i} \right. \\ &\qquad \left. + M_{\ell i}^T \overline{R}_\ell(k) N_{\ell i}^T + G_{\ell i} \left(R_\ell(k)\right)^H H_{\ell i} \right), \quad \ell = 1,2,\ldots,N, \end{aligned} \tag{3.11}$$

for $k = 0,1,2,\ldots$ where $\mathcal{X}_1(0), \mathcal{X}_2(0)$ are given arbitrary real vectors,

$$X(k) = \mathrm{unvec}\left(\mathcal{X}_1(k) + i\mathcal{X}_2(k)\right), \quad k = 0,1,2,\ldots,$$

and

$$R_\ell(k) = F_\ell - \left(\sum_{i=1}^{p} A_{\ell i} X_i(k) B_{\ell i} + C_{\ell i} X_i(k)^T D_{\ell i} + M_{\ell i} \overline{X}_i(k) N_{\ell i} + H_{\ell i} X_i(k)^H G_{\ell i} \right).$$

From Remark 2.2, for finding the minimum norm solution $\tilde{\mathcal{X}}$ (solution of Problems III and IV), the initial iterate $\mathcal{X}(0) = \mathcal{X}_1(0) + i\mathcal{X}_2(0)$ should be chosen such that

$$\begin{pmatrix} \mathcal{X}_1(0) \\ \mathcal{X}_2(0) \end{pmatrix} \in \mathrm{Range}\ (\mathcal{U}^T).$$

From (3.9), we may conclude that the initial guess $X(0)$ in (3.11) should be chosen such that

$$X(0) = \mathrm{unvec}(\mathcal{X}(0)) = \left(\sum_{\ell=1}^{N} A_{\ell i}^H W_\ell B_{\ell i}^H + \overline{D}_{\ell i} W_\ell^T \overline{C}_{\ell i} + M_{\ell i}^T \overline{W}_\ell N_{\ell i}^T + G_{\ell i} W_\ell^H H_{\ell i} \right), \tag{3.12}$$

for a given arbitrary matrix group $W = [W_1, W_2, \ldots, W_N]$.

In view of Theorem 2.3, we may instantly conclude the following two theorems.

Theorem 3.1. *Suppose that (1.2) is consistent. Presume that*

$$0 < \mu < \frac{2}{\sigma_{\max}^2(\mathcal{U})}, \tag{3.13}$$

where $\sigma_{\max}$ is the largest singular value of $\mathcal{U}$. Then, the sequence of approximate solutions produced by recursive formula (3.11) converges to a solution of (1.2) for any initial guess $X(0)$. In the case that $X(0)$ is chosen of the form (3.12), then the iterative method defined by (3.11) converges to the solution of Problem I. Furthermore, the optimal value of μ is given by

$$\mu_{opt} = \frac{2}{\bar{\sigma}_{\min}^2(\mathcal{U}) + \sigma_{\max}^2(\mathcal{U})}, \tag{3.14}$$

where $\bar{\sigma}_{\min}$ is the smallest nonzero singular value of $\mathcal{U}$.

Theorem 3.2. *Suppose that (1.2) is not consistent. Presume that μ satisfies (3.13). Then, the sequence of approximate solutions produced by recursive formula (3.11) converges to a solution of the subsequent least-squares problem*

$$\min \|F - \mathcal{A}(X)\|,$$

for any initial guess $X(0)$. Moreover assume that $X(0)$ is chosen of the form (3.12), then the iterative method defined by (3.11) converges to the solution of Problem II. Furthermore, the optimal value of μ is obtained by (3.14).

4. Numerical experiments

In this section, we examine some numerical experiments to illustrate the effectiveness of the proposed algorithm and the presented theoretical results. All the reported numerical experiments in this section were computed in double precision with some Matlab codes.

Example 4.1. In this example we mention the matrix equation

$$A_{11}X_1B_{11} + C_{11}X_1^TD_{11} + M_{11}\overline{X_1}N_{11} + H_{11}X_2^HC_{11} = F_1, \tag{4.1}$$

where

$$A_{11} = \begin{pmatrix} -2-2i & 2+2i \\ 1-i & -2-1i \end{pmatrix}, \quad B_{11} = \begin{pmatrix} 0 & 1+3i \\ 0 & 5-10i \end{pmatrix}, \quad C_{11} = \begin{pmatrix} 5+3i & 2-i \\ 4-2i & 1+2i \end{pmatrix},$$
$$D_{11} = \begin{pmatrix} 0 & 2-5i \\ 0 & 4i \end{pmatrix}, \quad M_{11} = \begin{pmatrix} 2-3i & 3+5i \\ 0 & 0 \end{pmatrix}, \quad N_{11} = \begin{pmatrix} 6i & 1+3i \\ -1+i & -1+4i \end{pmatrix},$$
$$G_{11} = \begin{pmatrix} 5i & i \\ 3+4i & 2+i \end{pmatrix}.$$

We consider following three cases for the matrices H_{11} and F_1:

Case 1.

$$H_{11} = \begin{pmatrix} 2+3i & 3i \\ 0 & 10i \end{pmatrix} \quad \text{and} \quad F_1 = \begin{pmatrix} -52+248i & -72+240i \\ -10+70i & -59-115i \end{pmatrix};$$

Case 2.

$$H_{11} = \begin{pmatrix} 2+3i & 3i \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad F_1 = \begin{pmatrix} -52+248i & -72+240i \\ 0 & -99-135i \end{pmatrix};$$

Case 3.

$$H_{11} = \begin{pmatrix} 2+3i & 3i \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad F_1 = \begin{pmatrix} -1-5i & -11-6i \\ 10-i & -12-3i \end{pmatrix}.$$

It is easy to verify that the matrix

$$X_1^* = \begin{pmatrix} 2-2i & 2-i \\ 2+2i & 3i \end{pmatrix}, \tag{4.2}$$

is a solution of Eq. (4.1) for both of the Cases 1 and 2. We apply the iterative method given by (3.11) to solve Eq. (4.1). We use a null matrix as an initial guess and

$$\delta_k = \frac{\|R_1(k)\|}{\|R_1(0)\|} < 10^{-7},$$

as the stooping criterion for the Cases 1 and 2, where $R_1(k)$ is the residual matrix at kth iteration.

In Case 1, it can be seen that the corresponding $\mathcal{U}$ appeared in Eq. (3.6) is a nonsingular 8-by-8 matrix. Therefore, system (3.6) and as a result system (4.1) has a unique solution given by (4.2). According to Theorem 3.1, the iterative method (3.11) converges to the exact solution X_1^* if

$$0 < \mu < 1.9328 \times 10^{-4}.$$

Moreover the optimum value of μ is given by

$$\mu_{opt} = 1.7378 \times 10^{-4}.$$

In Figure 1, the convergence history of the method for three values μ_{opt}, $\mu = 1.0 \times 10^{-4}$ and $\mu = 1.9 \times 10^{-4}$ are depicted where the method converges, respectively, in 71, 119 and 463 iterations. As seen, the best convergence curve corresponds to μ_{opt} among the chosen values for μ.

Now, we consider the second case. It is straightforward to see that the corresponding $\mathcal{U}$ is an 8-by-8 matrix with $\text{rank}(\mathcal{U}) = 6$. This means that Eq. (3.6), and as a result system (4.1), has infinitely number of solutions. Hence, if the assumptions of Theorem 3.1 hold, then the proposed iterative method converges to a solution of the given matrix equation for any initial guess. According to Theorem 3.1, the convergence interval for μ is

$$0 < \mu < 1.9280 \times 10^{-4}.$$

Moreover the optimum value of μ is given by

$$\mu_{opt} = 1.6845 \times 10^{-4}.$$

In Figure 2, the convergence history of the method for three values μ_{opt}, $\mu = 1.0 \times 10^{-4}$ and $\mu = 1.9 \times 10^{-4}$ are depicted where the method converges, respectively, in 55, 92 and 542 iterations. As observed, the least number of iterations is due to μ_{opt} between the chosen values for μ. It is necessary to mention that the method with the optimum value converges to the solution

$$X(55) = \begin{pmatrix} 1.7447 - 2.1364i & 1.8825 - 0.5580i \\ 1.9696 + 1.8761i & -0.0606 + 3.2778i \end{pmatrix},$$

which is different from X_1^*.

Finally, we consider Case 3. Since the left-hand side of (4.1) for both Cases 2 and 3 are the same, we see that the matrix $\mathcal{U}$ is not of full rank. It is not difficult to check that the system (4.1) (as a result Eq. (3.6)) is inconsistent. Therefore, we look for the solution of Problem II. We apply the proposed iterative method with a null matrix as an initial guess and

$$\delta_k = \|X(k) - X(k-1)\|_F < 10^{-7},$$

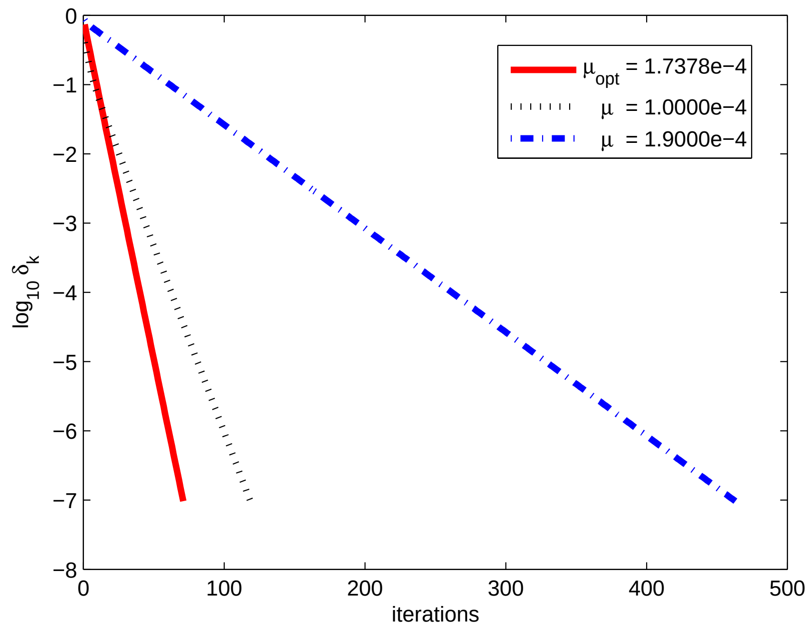

FIGURE 1. $\log_{10}\delta_k$ for Example 4.1 in Case 1 (system (3.6) is consistent and $\mathcal{U}$ is of full rank).

as the stopping criterion. Obviously, the range of the parameter μ and its optimal value are as Case 2. In Figure 3, the convergence history of the method for three values of μ, i.e., $\mu=\mu_{opt}$, $\mu=1.5\times10^{-4}$ and $\mu=1.8\times10^{-4}$ has been displayed. For these values of μ, the method converges in 48, 54 and 90 iterations, respectively. With $\mu=\mu_{opt}$ the obtained solution is

$$X(48)=\begin{pmatrix} -0.0645-0.3148i & 0.0808-0.1287i \\ -0.1723+0.0554i & 0.0253+0.0365i \end{pmatrix}.$$

This is the solution of Problem II corresponding to (4.1) (least squares solution of (3.6)) which confirms the presented results in Theorem 3.2 .

Example 4.2. In this example, we consider the coupled linear matrix equations

$$\begin{cases} A_{11}X_1B_{11}+C_{11}X_1^TD_{11}+M_{12}\overline{X_2}N_{12}+H_{12}X_2^HG_{12}=F_1, \\ A_{21}X_1B_{21}+C_{21}X_1^TD_{21}+M_{22}\overline{X_2}N_{22}+H_{22}X_2^TG_{22}=F_2, \end{cases} \tag{4.3}$$

where

$$A_{11}=\begin{pmatrix} -1-2i & 2+2i \\ -6i & -2-3i \end{pmatrix},\quad B_{11}=\begin{pmatrix} 9i & 1+3i \\ 0 & 5-i \end{pmatrix},\quad C_{11}=\begin{pmatrix} 8i9i & 2-i \\ 1-4i & 3+i \end{pmatrix},$$

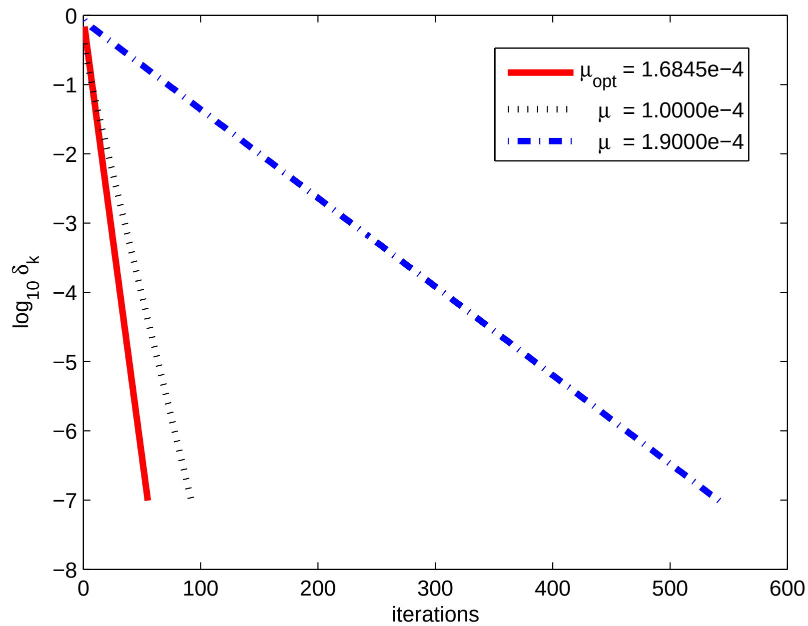

FIGURE 2. $\log_{10}\delta_k$ for Example 4.1 in Case 2 (system (3.6) is consistent, but $\mathcal{U}$ is not of full rank).

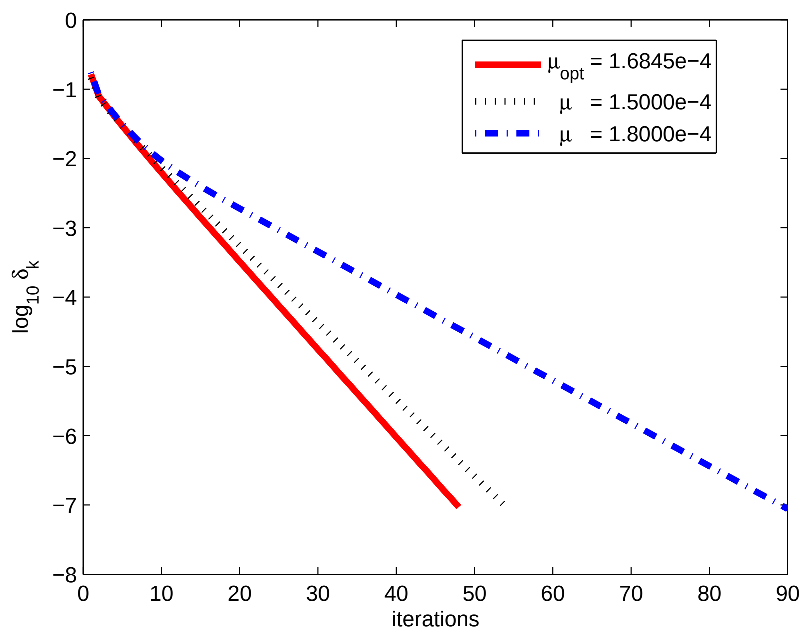

FIGURE 3. $\log_{10}\delta_k$ for Example 4.1 in Case 3 (system (3.6) is not consistent).

$$D_{11} = \begin{pmatrix} 1 & 2-5i \\ -9 & i \end{pmatrix}, \quad M_{12} = \begin{pmatrix} 0 & 3+i \\ 0 & 1-i \end{pmatrix}, \quad N_{12} = \begin{pmatrix} 2-1i & 1+3i \\ 10+i & -2-i \end{pmatrix},$$
$$H_{12} = \begin{pmatrix} 5i & 0 \\ 0 & 0 \end{pmatrix}, \quad G_{12} = \begin{pmatrix} 1-i & i \\ 3-i & 2+i \end{pmatrix}, \quad A_{21} = \begin{pmatrix} -2-6i & 4+2i \\ 1-i & -2-2i \end{pmatrix},$$
$$B_{21} = \begin{pmatrix} 2i & 1-3i \\ 1-i & 1+i \end{pmatrix}, \quad C_{21} = \begin{pmatrix} -2i & 2+i \\ 0 & -1+2i \end{pmatrix}, \quad D_{21} = \begin{pmatrix} 2-i & 2+5i \\ -7i & 1+i \end{pmatrix},$$
$$N_{22} = \begin{pmatrix} 5-6i & 4+3i \\ 3i & -9i \end{pmatrix}, \quad H_{22} = \begin{pmatrix} -5-9i & 0 \\ 0 & 0 \end{pmatrix}, \quad G_{22} = \begin{pmatrix} 0 & 5i \\ 3-i & 4-2i \end{pmatrix}.$$

We set following three cases for M_{22}, F_1 and F_2.

Case 1.

$$M_{22} = \begin{pmatrix} -1+9i & 1-2i \\ 3 & 5-4i \end{pmatrix},$$
$$F_1 = \begin{pmatrix} 135-192i & 32-11i \\ 51-87i & 144-57i \end{pmatrix} \quad \text{and} \quad F_2 = \begin{pmatrix} -13-183i & 98-112i \\ 122-86i & -193-57i \end{pmatrix};$$

Case 2.

$$M_{22} = \begin{pmatrix} 0 & 3+i \\ 0 & 1-6i \end{pmatrix},$$
$$F_1 = \begin{pmatrix} 135-192i & 32-11i \\ 51-87i & 144-57i \end{pmatrix} \quad \text{and} \quad F_2 = \begin{pmatrix} 46-82i & -126-162i \\ 29-135i & -103+85i \end{pmatrix};$$

Case 3.

$$M_{22} = \begin{pmatrix} 0 & 3+i \\ 0 & 1-6i \end{pmatrix},$$
$$F_1 = \begin{pmatrix} -5-i & 3-2i \\ -6 & -1-2i \end{pmatrix} \quad \text{and} \quad F_2 = \begin{pmatrix} 2i & i \\ 5i & 3-3i \end{pmatrix};$$

It is not onerous to check that in both of the Cases 1 and 2, (X_1^*, X_2^*) is the solution of (4.3) where

$$X_1^* = \begin{pmatrix} 1-i & 1+3i \\ 2+i & 1+i \end{pmatrix} \quad \text{and} \quad X_2^* = \begin{pmatrix} i & 2+i \\ 2-i & 2+3i \end{pmatrix}. \tag{4.4}$$

This shows that in both of these cases the system (4.3) is consistent. We use null matrices as the initial guess and

$$\delta_k = \max\{\frac{\|R_1(k)\|}{\|R_1(0)\|}, \frac{\|R_2(k)\|}{\|R_2(0)\|}\} < 10^{-7},$$

as the stoping criterion for the Cases 1 and 2 where $R_1(k)$ and $R_2(k)$ are the residual matrices at kth iteration.

We first consider Case 1. It is easy to see that the corresponding $\mathcal{U}$ appeared in Eq. (3.6) is a nonsingular 16-by-16 matrix. Therefore both of the systems (3.6) and (4.3) have a unique solution given by (4.4). According to Theorem 3.1, the iterative method (3.11) converges to the exact solution (X_1^*, X_2^*) if

$$0 < \mu < 1.6332 \times 10^{-4}.$$

In addition, the optimum value of μ is given by

$$\mu_{opt} = 1.5403 \times 10^{-4}.$$

In Figure 4, the convergence history of the method for three values μ_{opt}, $\mu = 1.58 \times 10^{-4}$ and $\mu = 1.35 \times 10^{-4}$ are represented where the method converges in 114, 199 and 128 iterations , respectively. As seen, the best convergence curve is occurred for μ_{opt} among the different chosen values for μ.

Now, we mention Case 2. In this case, it is easy to see that the corresponding $\mathcal{U}$ is an 16-by-16 matrix with $\mathrm{rank}(\mathcal{U}) = 14$. This means that Eq. (3.6), and as a result system (4.3), has infinitely number of solutions. Hence, if the assumptions of Theorem 3.1 hold, then the proposed iterative method converges to a solution of the given matrix equation for any initial guess. According to Theorem 3.1, the convergence interval for μ is

$$0 < \mu < 2.1538 \times 10^{-4}.$$

Moreover the optimum value of μ is given by

$$\mu_{opt} = 1.9619 \times 10^{-4}.$$

In Figure 5, the convergence history of the method for three values μ_{opt}, $\mu = 1.6 \times 10^{-4}$ and $\mu = 2.05 \times 10^{-4}$ are illustrated where the method converges in 77, 93 and 146 iterations, respectively. As observed, the least number of iterations belongs to μ_{opt} in comparison with different choices of μ. It is necessary to consider that the method with the optimum value converges to the solution

$$\hat{X}_1 = \begin{pmatrix} 1-i & 1+3i \\ 2+i & 1+i \end{pmatrix} \quad \text{and} \quad \hat{X}_2 = \begin{pmatrix} i & 0 \\ 2-i & 2+3i \end{pmatrix},$$

which is different from (X_1^*, X_2^*). In other words, the method has converged to another solution of (4.3).

At last, we consider Case 3. Since the left-hand side of the system (4.3) in the Cases 1 and 2 are the same, we observe that the corresponding $\mathcal{U}$ is an 16-by-16 matrix with $\mathrm{rank}(\mathcal{U}) = 14$. However, it is not difficult to verify that with the chosen right-hand side (F_1, F_2), system (4.3) is not consistent. Therefore, we look for the solution of Problem II. We apply the proposed iterative method with null matrices as the initial guess and

$$\delta_k = \max\{\|X_1(k) - X_1(k-1)\|, \|X_2(k) - X_2(k-1)\|\} < 10^{-7},$$

as the stopping criterion. Evidently, the range of the parameter μ and its optimal value are as Case 2. In Figure 3 the convergence history of the method for three values of μ, i.e., $\mu = \mu_{opt}$, $\mu = 1.7 \times 10^{-4}$ and $\mu = 2.05 \times 10^{-4}$ has been depicted. For these values of μ, the method converges, respectively, in 65, 75 and 113 iterations. With $\mu = \mu_{opt}$ the computed solution is given by $(X_1(65), X_2(65))$ where

$$X_1(65) = \begin{pmatrix} -0.0329 - 0.0205i & 0.0183 + 0.0247i \\ -0.0124 + 0.0068i & 0.0928 + 0.0098i \end{pmatrix},$$

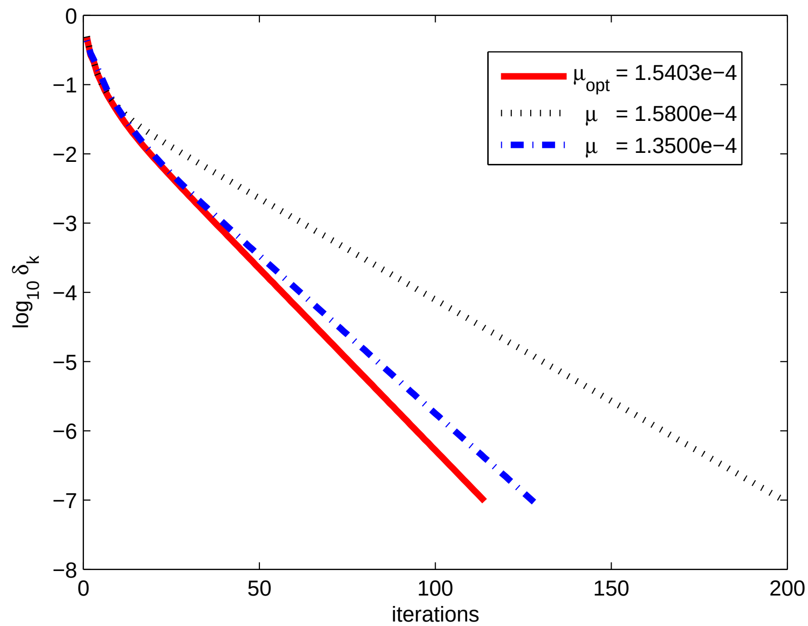

FIGURE 4. $\log_{10} \delta_k$ for Example 4.2 in Case 1 (system (3.6) is consistent and $\mathcal{U}$ is of full rank).

and

$$X_2(65) = \begin{pmatrix} -0.0017 + 0.1390i & 0 \\ -0.0908 + 0.0379i & -0.1001 - 0.0704i \end{pmatrix},$$

which is the solution of Problem II corresponding to (4.1) (least squares solution of (3.6)).

5. Conclusion

The convergence of the gradient-based algorithm together with determining its best convergence factor have been studied to compute the minimum norm (least-squares) solution of the general complex of the (in)consistent coupled matrix equations. Our main inspiration to discuss the convergence of the algorithm for the mentioned problems was the restrictions in the assumptions of the preceding research works published in the discipline of the application of the algorithm to solve the considered problems and their special cases. The optimal convergence factor of the algorithm has been also derived which was not obtained without setting the refereed restrictions so far. As a matter of fact, the presented results in this work have elaborated for more general situations which left as a subject to be investigated in the previously published research papers in the literature.

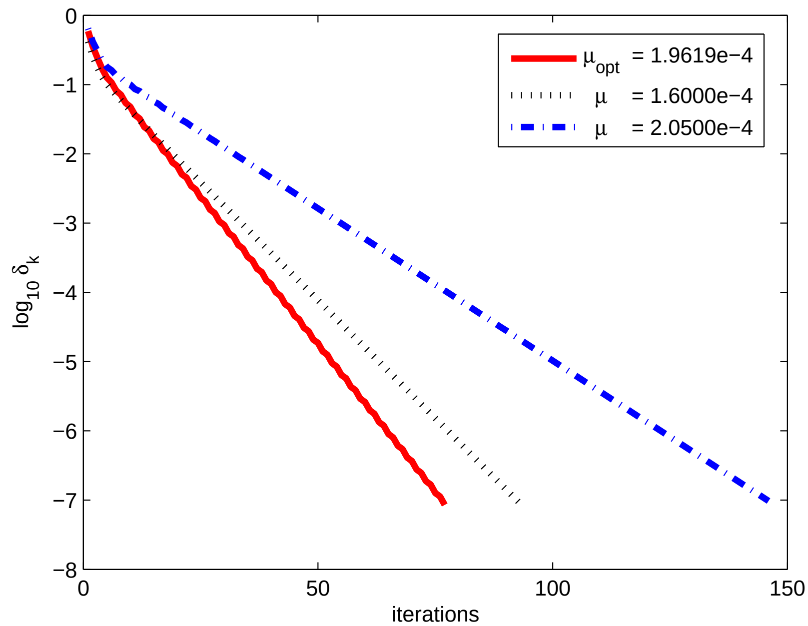

FIGURE 5. $\log_{10}\delta_k$ for Example 4.2 in Case 2 (system (3.6) is consistent, but $\mathcal{U}$ is not of full rank).

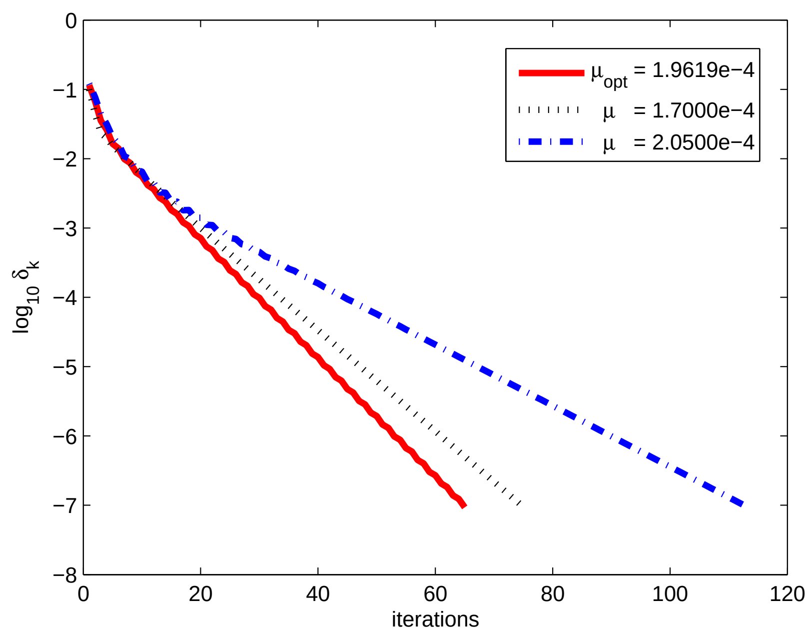

FIGURE 6. $\log_{10}\delta_k$ for Example 4.2 in Case 2 (system (3.6) is not consistent).

Acknowledgments

The authors would like to express their heartfelt gratitude to the anonymous referee for her/his valuable recommendations and useful comments which have improved the quality of the paper.

References

[1] Z. Al-Zhour and A. Kilicman, Some new connections between matrix products for partitioned and non-partitioned matrices, *Comput. Appl. Math.* 54 (2008), no. 6, 763–784.

[2] F. P. A. Beik and D. K. Salkuyeh, On the global Krylov subspace methods for solving general coupled matrix equation, *Comput. Math. Appl.* 62 (2011), no. 11, 4605–4613.

[3] F. P. A. Beik, D. K. Salkuyeh and M. M. Moghadam, Gradient-based iterative algorithm for solving the generalized coupled Sylvester-transpose and conjugate matrix equations over reflexive (anti-reflexive) matrices, *Transactions of the Institute of Measurement and Control,* 36 (2014), no. 1, 99–110.

[4] A. Berman and R. J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York, third edition, Reprinted by SIAM, Philadelphia, 1994.

[5] D. S. Bernstein, Matrix mathematics: theory, facts, and formulas. Princeton University Press, New Jersey, 2009.

[6] B. J. Broxson, *The Kronecker Product*, University of North Florida, UNF Theses and Dissertations, 2006.

[7] M. Dehghan, M. Hajarian, Two iterative algorithms for solving coupled matrix equations over reflexive and anti-reflexive matrices, *Comput. Appl. Math.* 31 (2012), no. 2, 353–371.

[8] F. Ding and T. Chen, Iterative least-squares solutions of coupled Sylvester matrix equations, *Systems Control Lett.* 54 (2005), no. 2, 95–107.

[9] F. Ding and T. Chen, On iterative solutions of general coupled matrix equations, *SIAM J. Control Optim.* 44 (2006), no. 6, 2269-2284.

[10] F. Ding, Coupled-least-squares identification for multivariable systems, *IET Control Theory and Applications,* 7 (2013), no. 1, 68–79.

[11] F. Ding, P. X. Liu and J. Ding, Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle, *Appl. Math. Comput.* 197 (2008), no. 1, 41–50.

[12] J. Ding, Y. Liu and F. Ding, Iterative solutions to matrix equations of the form $A_iXB_i = F_i$, *Comput. Math. Appl.* 59 (2010), no. 11, 3500–3507.

[13] M. Hajarian and M. Dehghan, Efficient iterative solutions to general coupled matrix equations, *International Journal of Automation and Computing,* In press.

[14] M. Hajarian, A gradient-based iterative algorithm for generalized coupled Sylvester matrix equations over generalized centro-symmetric matrices, *Transactions of the Institute of Measurement and Control,* DOI: 10.1177/0142331213497499.

[15] A. J. Laub, Matrix Analysis for Scientists & Engineers, SIAM, Philadelphia, 2005.

[16] Z. Y. Li and Y. Wang, Iterative algorithm for minimal norm least squares solution to general linear matrix equations, *Int. J. Comput. Math.* 87 (2010), no. 11, 2552–2567.

[17] Z. Y. Li, Y. Wang, B. Zhou and G. R. Duan, Least squares solution with the minimum-norm to general matrix equations via iteration, *Appl. Math. Comput.* 215 (2010), no. 10, 3547–3562.

[18] Z. Y. Peng and Y. X. Peng, An efficient iterative method for solving the matrix equation $AXB + CYD = E$, *Numer. Linear Algebra Appl.* 13 (2006), no. 6, 473–485.

[19] M. A. Ramadan, T. S. El-Danaf and A. M. E. Bayoumi, Two iterative algorithms for the reflexive and Hermitian reflexive solutions of the generalized Sylvester matrix equation, *Journal of Vibration and Control,* DOI: 10.1177/1077546312474859.

[20] Y. Saad, Iterative Methods for Sparse Linear Systems, PWS press, New York, 1996.

[21] D. K. Salkuyeh and F. P. A. Beik, On the gradient based algorithm for solving the general coupled matrix equations, *Transactions of the Institute of Measurement and Control,* DOI: 10.1177/0142331213504429.

[22] C. Song, G. Chen and L. Zhao, Iterative solutions to coupled Sylvester-transpose matrix equations, *Applied Mathematical Modelling*, 35 (2011), no. 10, 4675–4683.

[23] X. Wang and D. Liao, The optimal convergence factor of the gradient based iterative algorithm for linear matrix equations, *Filomat*, 26 (2012), no. 3, 607–613.

[24] A. Wu, G. Feng, G. Duan and W. Wu, Iterative solutions to coupled Sylvester-conjugate matrix equations, *Comput. Math. Appl.* 60 (2010), no. 1, 54–66.

[25] L. Xie, J. Ding and F. Ding, Gradient based iterative solutions for general linear matrix equations, *Comput. Math. Appl.* 58 (2009), no. 7, 1441–1448.

[26] L. Xie, Y. J. Liu and H. Z. Yang, Gradient based and least squares based iterative algorithms for matrix equations $AXB + CX^T D = F$, *Appl. Math. Comput.* 217 (2010), no. 5, 2191–2199.

[27] J. J. Zhang, A note on the iterative solutions of general coupled matrix equation. *Appl. Math. Comput.* 217 (2011), no. 22, 9380–9386.

[28] B. Zhou and G. R. Duan, On the generalized Sylvester mapping and matrix equations. *Systems Control Lett.* 57 (2008), no. 3, 200–208.

[29] B. Zhou, G. R. Duan and Z. Y. Li, Gradient based iterative algorithm for solving coupled matrix equations. *Systems Control Lett.* 58 (2009), no. 5, 327–333.

[30] B. Zhou, Z. Y. Li, G. R. Duan and Y. Wang, Weighted least squares solutions to general coupled Sylvester matrix equations, *J. Comput. Appl. Math.* 224 (2009), no. 2, 759–776.

[31] B. Zhou, J. Lam and G. R. Duan, Gradient-based maximal convergence rate iterative method for solving linear matrix equations, *Int. J. Comput. Math.* 87 (2010), no. 3, 515–527.

[32] B. Zhou, J. Lam and G.R. Duan, Convergence of gradient-based iterative solution of coupled Markovian jump Lyapunov equations, *Comput. Math. Appl.* 56 (2008), no. 2, 3070-3078.

[33] B. Zhou, Z. Y. Li, G. R. Duan and Y. Wang, Weighted least squares solutions to general coupled Sylvester matrix equations, *J. Comput. Appl. Math.* 224 (2009), no. 2, 759–776.