# On Iterative Solutions of a Class of Matrix Equations in Systems and Control

Feng Ding Tongwen Chen
Department of Electrical and Computer Engineering
University of Alberta, Edmonton
Alberta, Canada T6G 2V4

Email: fding@ece.ualberta.ca; tchen@ece.ualberta.ca.

*Abstract*— **In this paper, we present a general family of iterative methods to solve linear equations, which includes the well-known Jacobi and Gauss-Seidel iterations as its special cases. We give the necessary and sufficient conditions for convergence of the iterative solutions. Furthermore, the methods are extended to solve coupled Sylvester matrix equations. In our approach, we regard the unknown matrices to be solved as the system parameters to be identified, and propose a least squares iterative algorithm by applying a hierarchical identification principle. We prove that the iterative solution consistently converges to the exact solution for any initial value. The algorithms proposed require less storage capacity than the existing numerical ones. Finally, the algorithms are tested on computer and the results verify the theoretical findings.**

**Keywords: Sylvester matrix equation, Lyapunov matrix equation, identification, estimation, least squares, Jacobi iteration, Gauss-Seidel iteration, Hadamard product, star product, hierarchical identification principle.**

## I. Introduction

Lyapunov and Sylvester matrix equations play important roles in system theory [1], [2], [3], [4], [5]. Although exact solutions, which can be computed by using the Kronecker product, are important, the computational efforts rapidly increase with the dimensions of the matrices to be solved. For some applications such as stability analysis, it is often not necessary to compute exact solutions; approximate solutions or bounds of solutions are sufficient. Also, if the parameters in system matrices are uncertain, it is not possible to obtain exact solutions for robust stability results [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20].

Alternative ways exist which transform the matrix equations into forms for which solutions may be readily computed, for example, the Jordan canonical form [21], companion-type form [22], [23], Hessenberg-Schur form [24], [25]. In this area, Chu gave a numerical algorithm for solving the coupled Sylvester equations [26]; and Borno presented a parallel algorithm for solving the coupled Lyapunov equations [27]. But, these algorithms require computing some additional matrix transformation/decomposition; moreover, they are not suitable for more general coupled matrix equations of the form:

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

which includes the coupled Lyapunov and Sylvester equations as its special cases. In (1), $X_i \in \mathbb{R}^{m\times n}$ are the unknown matrices to be solved; $A_{ij}$, $B_{ij}$, and $C_{ij}$ represent constant (coefficient) matrices of appropriate dimensions. For such coupled matrix equations, the conventional methods require dealing with matrices whose dimensions are $mnp \times mnp$. Such a dimensionality problem leads to computational difficulty in that excessive computer memory is required for computation and inversion of large matrices of size $mnp \times mnp$. For instance, if $m = n = p = 100$, then $mnp \times mnp = 10^6 \times 10^6$.

In the field of matrix algebra and system identification, iterative algorithms have received much attention [28], [29], [30], [6], [7]. For example, Starke presented an iterative method for solutions of the Sylvester equations by using the SOR technique [31]; Jonsson and Kägström proposed recursive block algorithms for solving the coupled Sylvester matrix equations [32], [33]; Kägström derived an approximate solution of the coupled Sylvester equation [34]. To our best knowledge, numerical algorithms for general matrix equations have not been fully investigated, especially the iterative solutions of the coupled Sylvester matrix equations, as well as the general coupled matrix equations in (1), and the convergence of the iterative solutions involved, which are the focus of this work.

In this paper, the problem will be tackled in a new way – we regard the unknown matrices $X_j$ to be solved as the parameters (parameter matrices) of the system to be identified, and apply the so-called *hierarchical identification principle* to decompose the system into some subsystems, and derive iterative algorithms of the matrix equations involved. Our methods will generate solutions to the matrix equations which are arbitrarily close to the exact solutions.

The paper is organized as follows. In Section II, we extend the well-known Jacobi and Gauss-Seidel iterations and present a large family of iterative methods. In Sections III and IV, we define the block-matrix inner product (the star product for short) and derive iterative algorithms for the coupled Sylvester matrix equations and general coupled matrix equations, respectively, and study the convergence properties of the algorithms. In Section V we give an example for illustrating the effectiveness of the algorithms proposed in the paper. Finally, we offer some concluding remarks in Section VI.

0-7803-8335-4/04/$17.00 ©2004 AACC

## II. Extension of the Jacobi and Guass-Seidel Iterations

Consider the following linear equation:

$$Ax = b. \tag{2}$$

Here, $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ is a given full-rank square matrix with non-zero diagonal elements, $b \in \mathbb{R}^n$ is a constant vector, and $x \in \mathbb{R}^n$ an unknown vector to be solved. Let $D$ be the diagonal part of $A$, $L$ and $U$ be the strictly lower and upper triangular parts of $A$:

$$\begin{aligned} D &= \operatorname{diag}[a_{11},\ a_{22},\ \cdots,\ a_{nn}] \in \mathbb{R}^{n\times n}, \\ L &= \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ a_{21} & 0 & \ddots & & \vdots \\ a_{31} & a_{32} & 0 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}, \\ U &= \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & & a_{2n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}, \end{aligned}$$

which satisfy $L + D + U = A$. Then both the Jacobi and Gauss-Seidel iterations can be expressed as [28], [29]

$$Mx(k) = Nx(k-1) + b, \quad k = 1, 2, 3, \cdots,$$

where $x(k)$ is the iterative solution of $x$. For the Jacobi method, $M = D$ and $N = -(L+U)$; for the Gauss-Seidel method, $M = L + D$ and $N = -U$.

Unfortunately, the Jacobi and Gauss-Seidel iterations cannot guarantee that $x(k)$ converges to the exact solution $x = A^{-1}b$, and are not suitable for solving the non-square system: $Hx = g$ with $H \in \mathbb{R}^{m\times n}$. This motivates us to study new iterative methods.

Let $G \in \mathbb{R}^{n\times n}$ be a full-rank matrix to be determined and $\mu > 0$ be the step-size or convergence factor. We present a large family of iterative methods as follows:

$$x(k) = x(k-1) + \mu G\,[b - Ax(k-1)], \quad k = 1, 2, 3, \cdots, \tag{3}$$

which includes the Jacobi and Gauss-Seidel iterations as special cases. For example, when $G = D^{-1}$ and $\mu = 1$, we get the Jacobi method; when $G = (L+D)^{-1}$ and $\mu = 1$, we obtain the Gauss-Seidel method.

*Theorem 1:* For the iterative algorithm in (3), assume the system in (2) has a unique solution. Then the iterative solution $x(k)$ given by the algorithm in (3) converges to the exact solution $x$ (i.e., $\lim\limits_{k\to\infty} x(k) = x = A^{-1}b$) for any finite initial values $x(0)$ if and only if

$$\mu (GA)^{\mathrm{T}}(GA) < (GA)^{\mathrm{T}} + (GA). \tag{4}$$

In fact, if $(GA)^{\mathrm{T}} + (GA)$ is positive-definite, we can take

$$0 < \mu < \frac{\lambda_{\min}[(GA)^{\mathrm{T}} + (GA)]}{\lambda_{\max}[(GA)^{\mathrm{T}}(GA)]}.$$

where $\lambda_{\max}$ ($\lambda_{\min}$) denotes the maximum (minimum) eigenvalue.

□

The proof of Theorem 1 is straightforward and hence omitted here. We may draw the following corollaries from Theorem 1.

*Corollary 1:* For the Jacobi iteration, if $A^{\mathrm{T}}D^{-1} + D^{-1}A$ is a positive-definite matrix, and

$$0 < \mu < \frac{\lambda_{\min}[A^{\mathrm{T}}D^{-1} + D^{-1}A]}{\lambda_{\max}[A^{\mathrm{T}}D^{-2}A]},$$

then $\lim\limits_{k\to\infty} x(k) = x = A^{-1}b$.

□

*Corollary 2:* For the Gauss-Seidel iteration, if $A^{\mathrm{T}}(L+D)^{-\mathrm{T}} + (L+D)^{-1}A > 0$, and

$$0 < \mu < \frac{\lambda_{\min}[A^{\mathrm{T}}(L+D)^{-\mathrm{T}} + (L+D)^{-1}A]}{\lambda_{\max}[A^{\mathrm{T}}(L+D)^{-\mathrm{T}}(L+D)A]},$$

then $\lim\limits_{k\to\infty} x(k) = x$.

□

*Corollary 3:* If $A^{\mathrm{T}} + A > 0$, take $G = I$ (an identity matrix) to get a simple iteration,

$$\left\{\begin{array}{l} x(k) = x(k-1) + \mu[b - Ax(k-1)], \\ 0 < \mu < \frac{\lambda_{\min}[A^{\mathrm{T}}+A]}{\lambda_{\max}[A^{\mathrm{T}}A]}. \end{array}\right. \tag{5}$$

Or, if $A^{\mathrm{T}} + A < 0$, taking $G = -I$, we have

$$\left\{\begin{array}{l} x(k) = x(k-1) - \mu[b - Ax(k-1)], \\ 0 < \mu < \frac{\lambda_{\min}[-A^{\mathrm{T}}-A]}{\lambda_{\max}[A^{\mathrm{T}}A]}. \end{array}\right. \tag{6}$$

Both cases yield $\lim\limits_{k\to\infty} x(k) = x$.

□

*Corollary 4:* If we take $G = A^{\mathrm{T}}$, then the iterative algorithm,

$$\left\{\begin{array}{l} x(k) = x(k-1) + \mu A^{\mathrm{T}}[b - Ax(k-1)], \\ 0 < \mu < \frac{2}{\lambda_{\max}[A^{\mathrm{T}}A]}, \quad or \quad 0 < \mu < \frac{2}{\|A\|^2}, \end{array}\right. \tag{7}$$

yields $\lim\limits_{k\to\infty} x(k) = x$. Here, $\|X\|^2 = \operatorname{tr}[XX^{\mathrm{T}}]$.

□

*Corollary 5:* If we take $G = A^{-1}$, then the following iterative algorithm converges to $x$:

$$\left\{\begin{array}{l} x(k) = x(k-1) + \mu A^{-1}[b - Ax(k-1)], \\ 0 < \mu < 2. \end{array}\right. \tag{8}$$

If $A$ is a non-square $m\times n$ full column-rank matrix, then we have $\lim\limits_{k\to\infty} x(k) = x$ in the following:

$$\left\{\begin{array}{l} x(k) = x(k-1) + \mu(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}[b - Ax(k-1)], \\ 0 < \mu < 2. \end{array}\right. \tag{9}$$

□

When $\mu = 1$, $x(1) = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b$ is the least squares solution, so (9) is also called the least squares iterative algorithm or the iterative least squares algorithm.

From Corollaries 1 to 3, we can see that the Jacobi iteration, Gauss-Seidel iteration and the iterative algorithm in (5) or (6) all require doing matrix eigenvalue analysis and additional computation because it is more complicated to compute eigenvalues than the trace of a matrix; the iterative algorithms in (7) and (9) are also suitable for solving non-square systems and are very useful for finding the iterative solutions of general matrix equations to be studied later; the convergence factors $\mu$ in (8) and (9) do not rely on the matrix $A$ and is easy to choose, although the algorithms in (8) and (9) require computing matrix inversion.

## III. Coupled Sylvester matrix equations

In this section, we study iterative algorithms to solve the coupled Sylvester matrix equation:

$$\left\{\begin{array}{l} AX + YB = C, \\ DX + YE = F. \end{array}\right. \tag{10}$$

Here, $A, D \in \mathbb{R}^{m\times m}$, $B, E \in \mathbb{R}^{n\times n}$ and $C, F \in \mathbb{R}^{m\times n}$ are given constant matrices, $X, Y \in \mathbb{R}^{m\times n}$ are the unknown matrices to be solved.

First, let us introduce some notation. The notation $I_n$ is the identity matrix of $n\times n$. For two matrices $M$ and $N$, $M\otimes N$ is their Kronecker product. For two $m\times n$ matrices $X$ and $Y$ with

$$X = \left[\begin{array}{cccc} x_1, & x_2, & \cdots, & x_n \end{array}\right] \in \mathbb{R}^{m\times n},$$

$\mathrm{col}[X]$ is an $mn$-dimensional vector formed by columns of $X$:

$$\mathrm{col}[X] = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right] \in \mathbb{R}^{(mn)}$$

and

$$\mathrm{col}[X, Y] = \left[\begin{array}{c} \mathrm{col}[X] \\ \mathrm{col}[Y] \end{array}\right] \in \mathbb{R}^{(2mn)}.$$

The following result is well-known.

*Lemma 1:* Equation (10) has a unique solution if and only if the matrix

$$S_2 := \left[\begin{array}{cc} I_n\otimes A & B^{\mathrm{T}}\otimes I_m \\ I_n\otimes D & E^{\mathrm{T}}\otimes I_m \end{array}\right] \in \mathbb{R}^{(2mn)\times(2mn)}$$

is nonsingular; in this case, the unique solution is given by

$$\mathrm{col}[X, Y] = S_2^{-1}\mathrm{col}[C, F], \tag{11}$$

and the corresponding homogeneous matrix equation $AX + YB = \mathbf{0}$, $DX + YE = \mathbf{0}$ has a unique solution: $X = Y = \mathbf{0}$.

□

In order to derive the iterative solution to (10), we need to introduce the intermediate matrices $b_1$ and $b_2$ as follows:

$$b_1 := \left[\begin{array}{c} C - YB \\ F - YE \end{array}\right], \tag{12}$$

$$b_2 := [C - AX,\ F - DX]. \tag{13}$$

Then from (10), we obtain two fictitious subsystems

$$\begin{array}{ll} S_1: & G_1X = b_1, \\ S_2: & YH_1 = b_2. \end{array}$$

Here, $G_1 := \left[\begin{array}{c} A \\ D \end{array}\right]$ and $H_1 := [B,\ E]$.

Let $X(k)$ and $Y(k)$ be the iterative solutions of $X$ and $Y$. Referring to Corollary 5, it is not difficult to get the iterative solutions to $S_1$ and $S_2$ as follows:

$$\begin{aligned} X(k) &= X(k-1) + \mu(G_1^{\mathrm{T}}G_1)^{-1} \\ &\quad \left[\begin{array}{c} A \\ D \end{array}\right]^{\mathrm{T}} \left\{b_1 - \left[\begin{array}{c} A \\ D \end{array}\right] X(k-1)\right\}, \quad (14) \\ Y(k) &= Y(k-1) + \mu\{b_2 - Y(k-1)[B,\ E]\} \\ &\quad [B,\ E]^{\mathrm{T}}(H_1H_1^{\mathrm{T}})^{-1}. \quad (15) \end{aligned}$$

Substituting (12) into (14) and (13) into (15) gives

$$\begin{aligned} X(k) &= X(k-1) + \mu(G_1^{\mathrm{T}}G_1)^{-1} \\ &\quad \left[\begin{array}{c} A \\ D \end{array}\right]^{\mathrm{T}} \left\{\left[\begin{array}{c} C - YB \\ F - YE \end{array}\right] - \left[\begin{array}{c} A \\ D \end{array}\right] X(k-1)\right\} \\ &= X(k-1) + \mu(G_1^{\mathrm{T}}G_1)^{-1} \\ &\quad \left[\begin{array}{c} A \\ D \end{array}\right]^{\mathrm{T}} \left[\begin{array}{c} C - YB - AX(k-1) \\ F - YE - DX(k-1) \end{array}\right], \quad (16) \\ Y(k) &= Y(k-1) + \mu\{[C - AX,\ F - DX] \\ &\quad -Y(k-1)[B,\ E]\}[B,\ E]^{\mathrm{T}}(H_1H_1^{\mathrm{T}})^{-1} \\ &= Y(k-1) + \mu[C - AX - Y(k-1)B,\ F \\ &\quad -DX - Y(k-1)E][B,\ E]^{\mathrm{T}}(H_1H_1^{\mathrm{T}})^{-1}. \quad (17) \end{aligned}$$

Here, a difficulty arises in that the expressions on the right-hand sides of (16) and (17) contain the unknown parameter matrix $Y$ and $X$, respectively, so it is impossible to realize the algorithm in (16) and (17). Our solution is based on the hierarchical identification principle: The unknown variables $Y$ in (16) and $X$ in (17) are replaced by their estimates $Y(k-1)$ and $X(k-1)$. Thus, we obtain the iterative solutions $X(k)$ and $Y(k)$ of the coupled Sylvester equation in (10):

$$\begin{aligned} X(k) &= X(k-1) + \mu(G_1^{\mathrm{T}}G_1)^{-1} \\ &\quad \left[\begin{array}{c} A \\ D \end{array}\right]^{\mathrm{T}} \left[\begin{array}{c} C - AX(k-1) - Y(k-1)B \\ F - DX(k-1) - Y(k-1)E \end{array}\right], \end{aligned} \tag{18}$$

$$\begin{aligned} Y(k) &= Y(k-1) + \mu[C - AX(k-1) \\ &\quad -Y(k-1)B,\ F - DX(k-1) - Y(k-1)E] \\ &\quad [B,\ E]^{\mathrm{T}}(H_1H_1^{\mathrm{T}})^{-1}, \end{aligned} \tag{19}$$

$$\mu = \frac{1}{m+n}, \tag{20}$$

or

$$\mu = \frac{1}{\lambda_{\max}[G_1(G_1^{\mathrm{T}}G_1)^{-1}G_1^{\mathrm{T}}] + \lambda_{\max}[H_1^{\mathrm{T}}(H_1H_1^{\mathrm{T}})^{-1}H_1]},$$

To initialize the algorithm, we take $X(0) = Y(0) = \mathbf{0}$ or some small real matrix, e.g., $X(0) = Y(0) = 10^{-6}\mathbf{1}_{m\times n}$ with $\mathbf{1}_{m\times n}$ being an $m \times n$ matrix whose elements are 1.

*Theorem 2:* If the coupled Sylvester equation in (10) has a unique solution $X$ and $Y$, then the iterative solution $X(k)$ and $Y(k)$ given by the algorithm in (18)-(20) converges to $X$ and $Y$ for any finite initial values $X(0)$ and $Y(0)$, i.e.,

$$\lim_{k\to\infty} X(k) = X, \quad and \quad \lim_{k\to\infty} Y(k) = Y.$$

□

Due to the limit of pages, the proofs of this theorem and following theorem are omitted, but can be obtained from the authors.

The convergence factor in (20) may not be the best and may be conservative. In fact, there exists a best $\mu$ such that the fast convergence rate of $X(k)$ to $X$ and $Y(k)$ to $Y$ can be obtained - see the example to be studied later.

## IV. General coupled matrix equations

In this section, we will extend the iterative method to solve more general coupled matrix equations of the form:

$$\left\{\begin{array}{l} A_{11}X_1B_{11} + A_{12}X_2B_{12} + \cdots + A_{1p}X_pB_{1p} = C_1, \\ A_{21}X_1B_{21} + A_{22}X_2B_{22} + \cdots + A_{2p}X_pB_{2p} = C_2, \\ \quad \cdots \\ A_{p1}X_1B_{p1} + A_{p2}X_2B_{p2} + \cdots + A_{pp}X_pB_{pp} = C_p. \end{array}\right. \tag{21}$$

Here, $A_{ij} \in \mathbb{R}^{m\times m},\ B_{ij} \in \mathbb{R}^{n\times n}$ and $C_i \in \mathbb{R}^{m\times n}$ are given constant matrices, $X_j \in \mathbb{R}^{m\times n}$ are the unknown matrix to be solved.

In order to more succinctly express the iterative algorithm to be presented later, we introduce the block-matrix inner product - the star ($\star$) product for short, denoted by notation $\star$, which differs from Hadamard (inner) product [35], [36], [37], [38] and general matrix multiplication. Let

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} \in \mathbb{R}^{(mp)\times n}, \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} \in \mathbb{R}^{(np)\times m}, \tag{22}$$

Then the block-matrix star product is defined as

$$X \star Y = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} \star \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} = \begin{bmatrix} X_1Y_1 \\ X_2Y_2 \\ \vdots \\ X_pY_p \end{bmatrix}.$$

In the above definition, we assume that the dimensions of multiplier matrix and multiplicand matrix are compatible. Taking into account the dimension compatibility, the star product is superior to matrix multiplication. Note that $AB \star C = A(B \star C) \neq (AB) \star C$.

For the Hadamard ($\circ$) product, we have $X \circ Y = Y \circ X$. For the star product, since the multiplier matrix and multiplicand matrix are not necessarily of the same size, in general, $A \star B \neq B \star A$, $A \star B \star C = (A \star B) \star C \neq A \star (B \star C)$.

*Lemma 2:* Equation (21) has a unique solution if and only if the matrix $S_p = [B_{ij}^{\mathrm{T}} \otimes A_{ij}] \in \mathbb{R}^{(mnp)\times(mnp)}$ is nonsingular; in this case, the solution is

$$\mathrm{col}[X_1, X_2, \cdots, X_p] = S_p^{-1}\mathrm{col}[C_1, C_2, \cdots, C_p];$$

and if $C_i = \mathbf{0}$ $(i = 1,\ 2,\ \cdots,\ p)$, then the matrix equation in (21) has unique solutions $X_i = \mathbf{0}$ $(i = 1,\ 2,\ \cdots,\ p)$.

In order to derive the iterative algorithm for solving the general coupled matrix equation in (21), we first consider the coupled Sylvester equation in (10) to a more general form:

$$\left\{\begin{array}{l} AXI_B + I_AYB = C, \\ DXI_E + I_DYE = F, \end{array}\right.$$

whose iterative solution can be expressed as

$$X(k) = X(k-1) + \mu(G_1^{\mathrm{T}}G_1)^{-1}\begin{bmatrix} A \\ D \end{bmatrix}^{\mathrm{T}}$$

$$\left\{\begin{bmatrix} C - AX(k-1)I_B - I_AY(k-1)B \\ F - DX(k-1)I_E - I_DY(k-1)E \end{bmatrix} \star [I_B,\ I_E]^{\mathrm{T}}\right\}, \tag{23}$$

$$Y(k) = Y(k-1) + \mu \begin{bmatrix} I_A \\ I_D \end{bmatrix}^{\mathrm{T}}$$

$$\begin{bmatrix} C - AX(k-1)I_E - I_DY(k-1)B \\ F - DX(k-1)I_E - I_DY(k-1)E \end{bmatrix}$$

$$\star \begin{bmatrix} B^{\mathrm{T}} \\ E^{\mathrm{T}} \end{bmatrix}(H_1H_1^{\mathrm{T}})^{-1}. \tag{24}$$

If $I_A$, $I_B$, $I_D$ and $I_E$ are identity matrices of appropriate dimensions, then the algorithm in (23) and (24) is equivalent to the one in (18) and (19).

Let $X_i(k)$ be the estimates or iterative solutions of $X_i$, and

$$A_i = \begin{bmatrix} A_{1i} \\ A_{2i} \\ \vdots \\ A_{pi} \end{bmatrix} \in \mathbb{R}^{(mp)\times m},$$

$$B_i = [B_{1i},\ B_{2i},\ \cdots,\ B_{pi}] \in \mathbb{R}^{n\times(np)}.$$

We present the least squares iterative algorithm of computing the solutions $X_i(k)$ $(i = 1,\ 2,\ \cdots,\ p)$ of the matrix equations in (21) as follows:

$$X_i(k) = X_i(k-1) + \mu(A_i^{\mathrm{T}}A_i)^{-1}A_i^{\mathrm{T}}$$

$$\begin{bmatrix} C_1 - \sum\limits_{j=1}^{p} A_{1j}X_j(k-1)B_{1j} \\ C_2 - \sum\limits_{j=1}^{p} A_{2j}X_j(k-1)B_{2j} \\ \vdots \\ C_p - \sum\limits_{j=1}^{p} A_{pj}X_j(k-1)B_{pj} \end{bmatrix} \star \begin{bmatrix} B_{1i}^{\mathrm{T}} \\ B_{2i}^{\mathrm{T}} \\ \vdots \\ B_{pi}^{\mathrm{T}} \end{bmatrix} (B_iB_i^{\mathrm{T}})^{-1}, \tag{25}$$

$$\mu = \frac{1}{\sum\limits_{i=1}^{p} \lambda_{\max}[A_i(A_i^{\mathrm{T}}A_i)^{-1}A_i^{\mathrm{T}}]\lambda_{\max}[B_i^{\mathrm{T}}(B_iB_i^{\mathrm{T}})^{-1}B_i]},$$

or

$$\mu \;=\; \frac{1}{mnp}. \tag{26}$$

Since (25) and (26) are established based on the least squares iterative idea of Corollary 5, the algorithm in (25) and (26) is known as the least squares iterative algorithm. In this algorithm, we only require computing the inversion of the $m\times m$ and $n\times n$ matrices instead of the $mnp\times mnp$ matrix, e.g., in Lemma 2.

*Theorem 3:* If the coupled matrix equation in (21) has unique solutions $X_i,\ i = 1,\ 2,\ \cdots,\ p$, then the iterative solutions $X_i(k)$ given by the algorithm in (25)-(26) converge to the solutions $X_i$ for any finite initial values $X_i(0)$, i.e.,

$$\lim_{k\to\infty} X_i(k) = X_i, \quad i = 1,\ 2,\ \cdots,\ p.$$

□

## V. Example

In this Section, we give an example to illustrate the performance of the proposed algorithms.

Suppose that the coupled matrix equations are $AX + YB = C,\ DX + YE = F$ with

$$A = \begin{bmatrix} 2.00 & 1.00 \\ -1.00 & 2.00 \end{bmatrix}, \quad B = \begin{bmatrix} 1.00 & -0.20 \\ 0.20 & 1.00 \end{bmatrix},$$
$$C = \begin{bmatrix} 13.20 & 10.60 \\ 0.60 & 8.40 \end{bmatrix}, \quad D = \begin{bmatrix} -2.00 & -0.50 \\ 0.50 & 2.00 \end{bmatrix},$$
$$E = \begin{bmatrix} -1.00 & -3.00 \\ 2.00 & -4.00 \end{bmatrix}, \quad F = \begin{bmatrix} -9.50 & -18.00 \\ 16.00 & 3.50 \end{bmatrix}.$$

Then the solutions of $X$ and $Y$ from (11) are

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 4.00 & 3.00 \\ 3.00 & 4.00 \end{bmatrix},$$
$$Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} 2.00 & 1.00 \\ -2.00 & 3.00 \end{bmatrix}.$$

Taking $X(0) = Y(0) = 10^{-6}\mathbf{1}_{2\times 2}$, we apply the algorithm in (18) and (19) to compute $X(k)$ and $Y(k)$. The iterative solutions $X(k)$ and $Y(k)$ is shown in Table I, where

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

is the relative error. The errors $\delta$ with different convergence factors are shown in Fig. 1. From Table I and Fig. 1, it is clear that $\delta$ are becoming smaller and smaller and goes to zero as $k$ increases. This indicates that the proposed algorithm is effective.

TABLE I

The iterative solutions ($\mu = 1/1.10$)

| $k$ | $x_{11}$ | $x_{12}$ | $x_{21}$ | $x_{22}$ |
|---|---|---|---|---|
| 5 | 3.61430 | 2.99005 | 2.94096 | 3.69706 |
| 10 | 3.58609 | 3.05453 | 2.90272 | 3.87639 |
| 15 | 3.82227 | 3.06025 | 2.95326 | 3.97523 |
| 20 | 3.89469 | 3.05144 | 2.97031 | 3.99632 |
| 25 | 3.94038 | 3.03387 | 2.98259 | 4.00113 |
| 30 | 3.96448 | 3.02170 | 2.98944 | 4.00170 |
| 35 | 3.97879 | 3.01341 | 2.99364 | 4.00132 |
| 40 | 3.98723 | 3.00821 | 2.99615 | 4.00089 |
| 45 | 3.99229 | 3.00500 | 2.99767 | 4.00056 |
| 50 | 3.99534 | 3.00303 | 2.99859 | 4.00035 |
| 55 | 3.99718 | 3.00184 | 2.99915 | 4.00021 |
| 60 | 3.99829 | 3.00111 | 2.99948 | 4.00013 |
| Solution | 4.00000 | 3.00000 | 3.00000 | 4.00000 |

TABLE II

The iterative solutions ($\mu = 1/1.10$)

| $k$ | $y_{11}$ | $y_{12}$ | $y_{21}$ | $y_{22}$ | $\delta$ (%) |
|---|---|---|---|---|---|
| 5 | 3.32282 | 0.38948 | -2.97539 | 3.27086 | 22.33259974 |
| 10 | 2.34456 | 0.78180 | -2.21107 | 3.09466 | 7.84857813 |
| 15 | 2.21169 | 0.83128 | -2.10876 | 3.07171 | 4.34305171 |
| 20 | 2.10743 | 0.90351 | -2.04993 | 3.04066 | 2.41409661 |
| 25 | 2.06247 | 0.93997 | -2.02722 | 3.02519 | 1.42914360 |
| 30 | 2.03639 | 0.96383 | -2.01531 | 3.01515 | 0.85256301 |
| 35 | 2.02173 | 0.97803 | -2.00897 | 3.00919 | 0.51331998 |
| 40 | 2.01304 | 0.98670 | -2.00533 | 3.00556 | 0.30979089 |
| 45 | 2.00787 | 0.99195 | -2.00320 | 3.00337 | 0.18728213 |
| 50 | 2.00475 | 0.99512 | -2.00193 | 3.00204 | 0.11329119 |
| 55 | 2.00287 | 0.99705 | -2.00117 | 3.00123 | 0.06855766 |
| 60 | 2.00174 | 0.99821 | -2.00071 | 3.00075 | 0.04149393 |
| Solution | 2.00000 | 1.00000 | -2.00000 | 3.00000 | |

The effect of changing the convergence factor $\mu$ is illustrated in Fig. 1. We see that the larger the convergence factor $\mu$ is, the faster the convergence the algorithm (or, the smaller the estimation error). However, if $\mu$ is too large, the algorithm may diverge. How to choose a best convergence factor is still a project to be studied.

## VI. Conclusions

A family of iterative methods for linear systems is presented and a least squares iterative solution to coupled matrix equations are studied by using the hierarchical identification principle. The analysis indicates that the algorithms proposed can achieve a good convergence property for any initial values. Although the algorithms are presented for linear coupled matrix equations, the idea adopted can be easily extended to study iterative solutions of more complex matrix equations and nonlinear matrix equations, e.g., the Riccati equation.

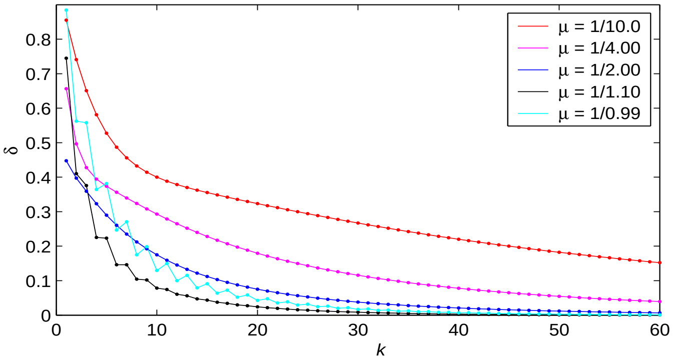

Fig. 1. The relative error $\delta$ of Example 1 versus $k$ (dots)

$$\mu = \frac{1}{m+n} = \frac{1}{4},$$

$$\mu = \frac{1}{\lambda_{\max}[G_1(G_1^{\mathrm{T}}G_1)^{-1}G_1^{\mathrm{T}}]+\lambda_{\max}[H_1^{\mathrm{T}}(H_1H_1^{\mathrm{T}})^{-1}H_1]} = \frac{1}{2}.$$

## VII. ACKNOWLEDGMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada and the National Natural Science Foundation of China (No. 60074029).

## VIII. REFERENCES

[1] T. Chen, B.A. Francis, Optimal Sampled-data Control Systems, London, Springer, 1995.

[2] T. Chen, L. Qiu, $\mathcal{H}_\infty$ design of general multirate sampled-data control systems, Automatica 30 (1994) 139-1152.

[3] L. Qiu, T. Chen, Contractive completion of block matrices and its application to $\mathcal{H}_\infty$ control of periodic systems, in: I. Gohberg, P. Lancaster, P.N. Shivakumar (Eds.), Recent Developments in Operator Theory and Its Applications, Birkhauser Verlag, Basel, Switzerland, 1996, pp. 263-281.

[4] L. Qiu, T. Chen, Multirate sampled-data systems: all $\mathcal{H}_\infty$ suboptimal controllers and the minimum entropy controller, IEEE Trans. Automat. Control 44 (1999) 537-550.

[5] L. Qiu, T. Chen, Unitary dilation approach to contractive matrix completion, Linear Algebra and its Applications, In Press, Corrected Proof, Available online 17 July 2003.

[6] Y. Fang, K.A. Loparo, X. Feng, New estimates for solutions of Lyapunov equations, IEEE Trans. Automat. Control 42 (1997) 408-411.

[7] H. Mukaidani, H. Xu, K. Mizukami, New iterative algorithm for algebraic Riccati equation related to $H_\infty$ control problem of singularly perturbed systems, IEEE Trans. Automat. Control 46 (2001) 1659-1666.

[8] T. Mori, A. Derese, A brief summary of the bounds on the solution of the algebraic matrix equations in control theory, Int. J. Control 39 (1984) 247-256.

[9] J. Garloff, Bounds for the eigenvalues of the solution of the discrete Riccati and Lyapunov equation and the continuous Lyapunov equation, Int. J. Control 43 (1986) 423-431.

[10] N. Komaroff, Simultaneous eigenvalue lower bounds for the Lyapunov matrix equation, IEEE Trans. Automat. Control 33 (1988) 126-128.

[11] A. Hmamed, Discrete Lyapunov equation: Simultaneous eigenvalue bounds, Int. J. Control 22 (1991) 1121-1126.

[12] M. Mrabti, A. Hmamed, Bounds for the solution of the Lyapunov matrix equation - A unified approach, Systems Control Lett. 18 (1992) 73-81.

[13] N. Komaroff, Lower bounds for the solution of the discrete algebraic Lyapunov equation, IEEE Trans. Automat. Control 37 (1992) 1017-1019.

[14] N. Komaroff, Upper summation and product bounds for solution eigenvalues of the Lyapunov matrix equation, IEEE Trans. Automat. Control 37 (1992) 1040-1042.

[15] M. Mrabti, and M. Benseddik, Unified type non-stationary Lyapunov matrix equation - Simultaneous eigenvalue bounds. Systems Control Lett. 24 (1995) 53-59.

[16] W.H. Kwon, Y.S. Moon, S.C. Ahn, Bounds in algebraic Riccati and Lyapunov equations: A survey and some new results, Int. J. Control 64 (1996) 377-389.

[17] C. H. Lee, Upper and lower matrix bounds of the solution for the discrete Lyapunov equation, IEEE Trans. Automat. Control 41 (1996) 1338-1341.

[18] C. H. Lee, On the matrix bounds for the solution matrix of the discrete algebraic Riccati equation, IEEE Trans. Circuits Syst. I, 43 (1996) 402-407.

[19] M. K. Tippert, D. Marchesin, Upper bounds for the solution of the discrete algebraic Lyapunov equation, Automatica 35 (1999) 1485-1489.

[20] T. Mori and H. Kokame, On Solution bounds for three types of Lyapunov matrix equations: Continuous, discrete and unified Equations, IEEE Trans. Automat. Control 47 (2002) 1767-1770.

[21] J. Heinen, A technique for solving the extended discrete Lyapunov matrix equation, IEEE Trans. Automat. Control 17 (1972) 156-157.

[22] R. Bitmead, H. Weiss, On the solution of the discrete-time Lyapunov matrix equation in controllable canonical form, IEEE Trans. Automat. Control 24 (1979) 481-482.

[23] R. Bitmead, Explicit solutions of the discrete-time Lyapunov matrix equation and Kalman-Yakubovich equations, IEEE Trans. Automat. Control 26 (1981) 1291-1294.

[24] A. Barraud, A numerical algorithm to solve $A^{\mathrm{T}}XA - X = Q$, IEEE Trans. Automat. Control 22 (1977) 883-885.

[25] G.H. Golub, S. Nash, C.F. Van Loan, A Hessenberg-Schur method for the matrix problem $AX + XB = C$, IEEE Trans. Automat. Control 24 (1979) 909-913.

[26] K.E. Chu, The solution of the matrix equations $AXB - CXD = E$ and $(YA - DZ, YC - BZ) = (E, F)$, Linear Algebra Appl. 93 (1987) 93-105.

[27] I. Borno. Parallel computation of the solutions of coupled algebraic Lyapunov equations, Automatica 31 (1995) 1345-1347.

[28] G. H. Golub, C.F. Van Loan, Matrix computations (Third Edition). Johns Hopkins University Press, Baltimore, MD, 1996.

[29] J.J. Climent, C. Perea, Convergence and comparison theorems for a generalized alternating iterative method, Applied Mathematics and Computation 143 (2003) 1-14.

[30] L. Ljung, System Identification: Theory for the User (2nded), Englewood Cliffs, NJ: Prentice-hall, 1999.

[31] G. Starke, W. Niethammer, SOR for $AX - XB = C$, Linear Algebra Appl. 154 (1991) 355-375.

[32] I. Jonsson, B. Kägström, Recursive blocked algorithms for solving triangular systems–Part I: One-sided and coupled Sylvester-type matrix equations, ACM Transactions on Mathematical Software 28 (2002) 392-415.

[33] I. Jonsson, B. Kägström, Recursive blocked algorithms for solving triangular systems–Part II: Two-sided and generalized Sylvester and Lyapunov matrix equations, ACM Transactions on Mathematical Software 28 (2002) 416-435.

[34] B. Kägström, A perturbation analysis of the generalized Sylvester equation $(AR - LB, DR - LE) = (C, F)$, SIAM Journal on Matrix analysis and applications 15 (1994) 1045-1060.

[35] C.R. Johnson, L. Elsner, The relationship between Hadamard and conventional multiplication for positive definite matrices, Linear Algebra Appl. 92 (1987) 231-240.

[36] G. Corach, D. Stojanoff, Index of Hadamard multiplication by positive matrices II, Linear Algebra Appl. 332-334 (2001) 503-517.

[37] S. Xiang, On an inequality for the Hadamard product of an M-matrix or an H-matrix and its inverse. Linear Algebra Appl. 367 (2003) 17-27.

[38] P. Fischer, J.D. Stegeman, Fractional Hadamard powers of positive semidefinite matrices, Linear Algebra Appl. (2003) (Article in Press)