

Iterative solutions for general coupled matrix equations with real coefficients

Li Xie, Huizhong Yang, Yanjun Liu, Feng Ding

Abstract—This paper applies the hierarchical identification principle and the gradient search method to study iterative solutions for a class of general coupled matrix equations with real coefficients. As long as the convergence factors are appropriately chosen, the proposed algorithms for any initial values can provide iterative solutions that are arbitrarily close to the unique solutions of the equations. Two numerical examples are given to demonstrate the effectiveness of the proposed algorithms.

Index Terms: Coupled matrix equations; hierarchical identification; gradient search; iterative algorithm; estimation.

I. INTRODUCTION

Throughout the paper, the transpose and the trace of matrix X are denoted by X^T and $\text{tr}[X]$, respectively, and the norm is defined as $\|X\|^2 = \text{tr}[X^T X]$. The symbol I_n stands for an identity matrix of size $n \times n$, $\mathbf{0}$ is a zero matrix with appropriate dimension. For two matrices A and B , $A \otimes B$ is their Kronecker product; and for matrix $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$, $x_i \in \mathbb{R}^m$, the vector operator is defined as $\text{col}[X] = [x_1^T, x_2^T, \dots, x_n^T]^T \in \mathbb{R}^{mn}$. Furthermore, we have $\text{col}[X^T] = P_{mn} \text{col}[X]$, where P_{mn} is a square $mn \times mn$ matrix partitioned into $m \times n$ submatrices such that the (i, j) th submatrix has a 1 in its (j, i) th position and zeros elsewhere.

Matrix equations have received much attention because of their important applications in control theory, signal processing, filtering and many other fields [1]–[4]. For example, the Sylvester matrix equations can be used for pole assignment, feedback design and fault detection [5], [6]; and the coupled Lyapunov matrix equations are encountered in stability analysis of linear jump systems with Markovian transitions [7]. There exist numerous methods for solving matrix equations, in which the iterative ones are especially computational efficient [8]–[10]. In this literature, Dehghan and Hajarian introduced three iterative algorithms to obtain the reflexive and anti-reflexive solutions for the matrix equations $A_1 X_1 B_1 + A_2 X_2 B_2 = C$ [11]; Zhou et al. developed a gradient based iterative algorithm to find the weighted least squares solutions for general coupled Sylvester matrix equations [12]; Wu et al. considered iterative solutions for a class of general complex matrix equations with the conjugate and transpose of the unknowns, and demonstrated that the

proposed algorithm can converge to the exact solutions within finite iteration steps [13].

Recently, by applying the hierarchical identification principle, Ding, Liu and Ding presented a gradient based and a least squares based iterative algorithms for solving the generalized Sylvester matrix equations, which regarded the unknown matrices as the parameter matrices to be identified [14]. Such methods are also developed to solve the extended Sylvester-conjugate matrix equations [15], general linear matrix equations $\sum_{i=1}^p A_i X B_i = F$ [16] and $\sum_{i=1}^p A_i X B_i + \sum_{i=1}^q C_i X^T D_i = F$ [17], and general coupled matrix equations $\sum_{j=1}^p A_{ij} X_j B_{ij} = F_i$, $i = 1, 2, \dots, p$ [18], [19]. Based on the work mentioned above, solutions for more general coupled matrix equations with real coefficients are considered in this paper, and the gradient based iterative algorithms are derived by using the hierarchical identification principle and the gradient search method.

The iterative algorithms are related to recursive estimation algorithms in system identification, e.g., the multi-innovation identification methods [20]–[32] and the iterative identification methods [33]–[37].

The rest of this paper is organized as follows. Section II presents a gradient based iterative algorithm to solve a simple form of coupled matrix equations with real coefficients; and Section III develops the algorithm for more general cases. Section IV provides two examples to demonstrate the effectiveness of the proposed algorithms. Finally, we end the paper with some concluding remarks in Section V.

II. SIMPLE COUPLED MATRIX EQUATIONS

This section concentrates on solutions for the coupled matrix equations

$$\begin{cases} A_1 X + X^T B_1 + C_1 Y + Y^T D_1 = F_1 \\ A_2 X + X^T B_2 + C_2 Y + Y^T D_2 = F_2 \end{cases}, \quad (1)$$

where $A_1, A_2, C_1, C_2 \in \mathbb{R}^{n \times m}$, $B_1, B_2, D_1, D_2 \in \mathbb{R}^{m \times n}$, and $F_1, F_2 \in \mathbb{R}^{n \times n}$ are given constant matrices, $X, Y \in \mathbb{R}^{m \times n}$ are two unknown matrices to be solved.

By using the vector operator, equation (1) can be converted into the following equivalent form,

$$S \text{col}[X, Y] = \text{col}[F_1, F_2],$$

$$S := \begin{bmatrix} I_n \otimes A_1 + (B_1^T \otimes I_n) P_{nm} & I_n \otimes C_1 + (D_1^T \otimes I_n) P_{nm} \\ I_n \otimes A_2 + (B_2^T \otimes I_n) P_{nm} & I_n \otimes C_2 + (D_2^T \otimes I_n) P_{nm} \end{bmatrix},$$

$$P_{nm} = P_{mn}^{-1} = P_{mn}^T,$$

$$\text{col}[X, Y] := \begin{bmatrix} \text{col}[X] \\ \text{col}[Y] \end{bmatrix}, \quad \text{col}[F_1, F_2] := \begin{bmatrix} \text{col}[F_1] \\ \text{col}[F_2] \end{bmatrix}.$$

This work was supported by the National Natural Science Foundation of China (60674092).

The authors are with the Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), and the Control Science and Engineering Research Center, Jiangnan University, Wuxi 214122, P.R. China. xiel2412@126.com (L. Xie); yhz@jiangnan.edu.cn (H.Z. Yang); yanjunliu.1983@126.com (Y.J. Liu); fding@jiangnan.edu.cn (F. Ding).

Then the exact solutions of equation (1) can be given by the following lemma.

Lemma 1: Equation (1) has unique solutions if and only if $\text{rank}\{S, \text{col}[F_1, F_2]\} = \text{rank}[S] = mn$ (i.e., S has a full column rank). In this case, the unique solutions can be given by

$$\text{col}[X, Y] = (S^T S)^{-1} S^T \text{col}[F_1, F_2], \quad (2)$$

and the corresponding homogeneous coupled matrix equation in (1) with $F_1 = \mathbf{0}$, $F_2 = \mathbf{0}$ has unique solutions $X = Y = \mathbf{0}$.

However, when the dimensions of X and Y become large, computing the solutions in (2) requires excessive computer memory and the resulted computational cost is high. This motivates us to study the iterative algorithm to solve (1).

Based on the hierarchical identification principle, regarding the unknown matrices X and Y in (1) as the parameter matrices to be identified, we define the following matrices,

$$Q_1 := \begin{bmatrix} F_1 - X^T B_1 - C_1 Y - Y^T D_1 \\ F_2 - X^T B_2 - C_2 Y - Y^T D_2 \end{bmatrix}, \quad (3)$$

$$Q_2 := [F_1 - A_1 X - C_1 Y - Y^T D_1, F_2 - A_2 X - C_2 Y - Y^T D_2], \quad (4)$$

$$Q_3 := \begin{bmatrix} F_1 - A_1 X - X^T B_1 - Y^T D_1 \\ F_2 - A_2 X - X^T B_2 - Y^T D_2 \end{bmatrix}, \quad (5)$$

$$Q_4 := [F_1 - A_1 X - X^T B_1 - C_1 Y, F_2 - A_2 X - X^T B_2 - C_2 Y]. \quad (6)$$

Then, from (1), we obtain four fictitious subsystems,

$$\text{Sub1: } \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} X = Q_1, \quad \text{Sub2: } X^T [B_1, B_2] = Q_2,$$

$$\text{Sub3: } \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} Y = Q_3, \quad \text{Sub4: } Y^T [D_1, D_2] = Q_4.$$

Applying the gradient search method [16], [19] to the above four subsystems leads to the following iterative equations:

$$X(k) = X(k-1) + \mu \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}^T \left\{ Q_1 - \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} X(k-1) \right\}, \quad (7)$$

$$X(k) = X(k-1) + \mu [B_1, B_2] \{ Q_2 - X^T(k-1) [B_1, B_2] \}^T, \quad (8)$$

$$Y(k) = Y(k-1) + \mu \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}^T \left\{ Q_3 - \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} Y(k-1) \right\}, \quad (9)$$

$$Y(k) = Y(k-1) + \mu [D_1, D_2] \{ Q_4 - Y^T(k-1) [D_1, D_2] \}^T, \quad (10)$$

where $\mu > 0$ is the iterative step size or convergence factor to be given later. Substituting (3)-(6) into (7)-(10), respectively, and replacing the unknown matrices X and Y with their estimates $X(k-1)$ and $Y(k-1)$ yield

$$X(k) = X(k-1) + \mu \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}^T \left\{ \begin{bmatrix} F_1 - X^T(k-1) B_1 - C_1 Y(k-1) \\ F_2 - X^T(k-1) B_2 - C_2 Y(k-1) \end{bmatrix} - \begin{bmatrix} Y^T(k-1) D_1 + A_1 X(k-1) \\ Y^T(k-1) D_2 + A_2 X(k-1) \end{bmatrix} \right\}, \quad (11)$$

$$X(k) = X(k-1) + \mu [B_1, B_2] \times \begin{bmatrix} F_1 - A_1 X(k-1) - C_1 Y(k-1) - Y^T(k-1) D_1 - X^T(k-1) B_1, \\ F_2 - A_2 X(k-1) - C_2 Y(k-1) - Y^T(k-1) D_2 - X^T(k-1) B_2 \end{bmatrix}^T, \quad (12)$$

$$Y(k) = Y(k-1) + \mu \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}^T \left\{ \begin{bmatrix} F_1 - A_1 X(k-1) - X^T(k-1) B_1 \\ F_2 - A_2 X(k-1) - X^T(k-1) B_2 \end{bmatrix} - \begin{bmatrix} Y^T(k-1) D_1 + C_1 Y(k-1) \\ Y^T(k-1) D_2 + C_2 Y(k-1) \end{bmatrix} \right\}, \quad (13)$$

$$Y(k) = Y(k-1) + \mu [D_1, D_2] \times \begin{bmatrix} F_1 - A_1 X(k-1) - X^T(k-1) B_1 - C_1 Y(k-1) - Y^T(k-1) D_1, \\ F_2 - A_2 X(k-1) - X^T(k-1) B_2 - C_2 Y(k-1) - Y^T(k-1) D_2 \end{bmatrix}^T. \quad (14)$$

Taking the average of (11) and (12) as the iterative solution $X(k)$, and the average of (13) and (14) as the iterative solution $Y(k)$, we obtain a gradient based iterative algorithm for the solutions of (1):

$$X(k) = \frac{X_a(k) + X_b(k)}{2}, \quad (15)$$

$$X_a(k) = X(k-1) + \mu \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}^T \begin{bmatrix} \Delta F_1(k) \\ \Delta F_2(k) \end{bmatrix}, \quad (16)$$

$$X_b(k) = X(k-1) + \mu [B_1, B_2] [\Delta F_1(k), \Delta F_2(k)]^T, \quad (17)$$

$$Y(k) = \frac{Y_a(k) + Y_b(k)}{2}, \quad (18)$$

$$Y_a(k) = Y(k-1) + \mu \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}^T \begin{bmatrix} \Delta F_1(k) \\ \Delta F_2(k) \end{bmatrix}, \quad (19)$$

$$Y_b(k) = Y(k-1) + \mu [D_1, D_2] [\Delta F_1(k), \Delta F_2(k)]^T, \quad (20)$$

$$\Delta F_i(k) = F_i - A_i X(k-1) - X^T(k-1) B_i - C_i Y(k-1) - Y^T(k-1) D_i, \quad i = 1, 2. \quad (21)$$

The convergence factor μ can be simply taken to satisfy

$$0 < \mu < 2 \{ \|A_1\|^2 + \|B_1\|^2 + \|C_1\|^2 + \|D_1\|^2 + \|A_2\|^2 + \|B_2\|^2 + \|C_2\|^2 + \|D_2\|^2 \}^{-1}. \quad (22)$$

To initialize the algorithm, we take $X(0)$ and $Y(0)$ as some small real matrices, 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 all 1.

Theorem 1: If the equation in (1) has unique solutions X and Y , then for any initial values, the iterative solutions $X(k)$ and $Y(k)$ given by the algorithm in (15)-(22) converge to the true solutions X and Y , i.e.,

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

III. GENERAL COUPLED MATRIX EQUATIONS

Consider the following more general coupled matrix equations with transpose of p unknown matrices,

$$\begin{cases} \sum_{j=1}^p [A_{1j} X_j B_{1j} + C_{1j} X_j^T D_{1j}] = F_1, \\ \sum_{j=1}^p [A_{2j} X_j B_{2j} + C_{2j} X_j^T D_{2j}] = F_2, \\ \vdots \\ \sum_{j=1}^p [A_{pj} X_j B_{pj} + C_{pj} X_j^T D_{pj}] = F_p, \end{cases} \quad (23)$$

where $A_{ij} \in \mathbb{R}^{r \times m}$, $B_{ij} \in \mathbb{R}^{n \times s}$, $C_{ij} \in \mathbb{R}^{r \times n}$, $D_{ij} \in \mathbb{R}^{m \times s}$, and $F_i \in \mathbb{R}^{r \times s}$ are given constant matrices, $X_j \in \mathbb{R}^{m \times n}$, $j = 1, 2, \dots, p$, are the unknown matrices to be determined.

In order to simplify the representation of the gradient based iterative algorithm to be proposed later, we use the block-matrix star product, denoted by notation \star in [18], [19].

Let

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}, Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix}, X_i, Y_i^T \in \mathbb{R}^{m \times n},$$

$$S_A = [A_{ij}]_{p \times p} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pp} \end{bmatrix},$$

$$S_{A^T} = [A_{ij}^T]_{p \times p} = \begin{bmatrix} A_{11}^T & A_{12}^T & \cdots & A_{1p}^T \\ A_{21}^T & A_{22}^T & \cdots & A_{2p}^T \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1}^T & A_{p2}^T & \cdots & A_{pp}^T \end{bmatrix},$$

$$S_B = [B_{ij}]_{p \times p}, S_{B^T} = [B_{ij}^T]_{p \times p},$$

$$S_C = [C_{ij}]_{p \times p}, S_{C^T} = [C_{ij}^T]_{p \times p},$$

$$S_D = [D_{ij}]_{p \times p}, S_{D^T} = [D_{ij}^T]_{p \times p},$$

$$S_p = [B_{ij}^T \otimes A_{ij} + (D_{ij}^T \otimes C_{ij})P_{nm}]_{p \times p}.$$

Then the block-matrix star product is defined as

$$X \star Y = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} \star \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} = \begin{bmatrix} X_1 Y_1 \\ X_2 Y_2 \\ \vdots \\ X_p Y_p \end{bmatrix},$$

$$S_A \star X = \begin{bmatrix} A_{11}X_1 & A_{12}X_2 & \cdots & A_{1p}X_p \\ A_{21}X_1 & A_{22}X_2 & \cdots & A_{2p}X_p \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1}X_1 & A_{p2}X_2 & \cdots & A_{pp}X_p \end{bmatrix},$$

$$S_A \star S_B = \begin{bmatrix} A_{11}B_{11} & A_{12}B_{12} & \cdots & A_{1p}B_{1p} \\ A_{21}B_{21} & A_{22}B_{22} & \cdots & A_{2p}B_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1}B_{p1} & A_{p2}B_{p2} & \cdots & A_{pp}B_{pp} \end{bmatrix}.$$

The star product is superior to matrix multiplication, thus $AB \star C = A(B \star C) \neq (AB) \star C$. Furthermore, the following properties exist:

$$\text{tr} \left\{ X_j^T \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{pj} \end{bmatrix}^T \begin{bmatrix} \tilde{F}_1 \\ \tilde{F}_2 \\ \vdots \\ \tilde{F}_p \end{bmatrix} \star \begin{bmatrix} B_{1j}^T \\ B_{2j}^T \\ \vdots \\ B_{pj}^T \end{bmatrix} \right\} = \text{tr} \left\{ \begin{bmatrix} A_{1j}X_jB_{1j} \\ A_{2j}X_jB_{2j} \\ \vdots \\ A_{pj}X_jB_{pj} \end{bmatrix}^T \begin{bmatrix} \tilde{F}_1 \\ \tilde{F}_2 \\ \vdots \\ \tilde{F}_p \end{bmatrix} \right\},$$

$$\left\| \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{pj} \end{bmatrix}^T \begin{bmatrix} \tilde{F}_1 \\ \tilde{F}_2 \\ \vdots \\ \tilde{F}_p \end{bmatrix} \star \begin{bmatrix} B_{1j}^T \\ B_{2j}^T \\ \vdots \\ B_{pj}^T \end{bmatrix} \right\|^2 \leq \sum_{i=1}^p \|A_{ij}\|^2 \|B_{ij}\|^2 \left\| \begin{bmatrix} \tilde{F}_1 \\ \tilde{F}_2 \\ \vdots \\ \tilde{F}_p \end{bmatrix} \right\|^2.$$

Lemma 2: Provided that the matrix S_p is full column rank, equation (23) has unique solutions and can be given by

$$\text{col}[X_1, X_2, \dots, X_p] = (S_p^T S_p)^{-1} \text{col}[F_1, F_2, \dots, F_p], \quad (24)$$

the corresponding homogeneous matrix equation in (23) has unique solutions $X_j = \mathbf{0}$, $j = 1, 2, \dots, p$.

Although (24) can be used to obtain the exact solutions of the coupled matrix equations in (23), it requires excessive computer memory because of computing the inversion of the large matrix $S_p^T S_p$ of size $(mnp) \times (mnp)$ as the dimension of X_j increases. Thus, we need to seek an alternative way to study the iterative solutions for (23), the details are as follows.

According to the hierarchical identification principle, regard the unknown matrices X_j , $j = 1, 2, \dots, p$, as the parameter matrices to be identified and decompose equation (23) into p subsystems,

$$\begin{cases} A_{1j}X_jB_{1j} = F_1 - \sum_{i=1, i \neq j}^p [A_{1i}X_iB_{1i} + C_{1i}X_i^T D_{1i}] - C_{1j}X_j^T D_{1j}, \\ A_{2j}X_jB_{2j} = F_2 - \sum_{i=1, i \neq j}^p [A_{2i}X_iB_{2i} + C_{2i}X_i^T D_{2i}] - C_{2j}X_j^T D_{2j}, \\ \vdots \\ A_{pj}X_jB_{pj} = F_p - \sum_{i=1, i \neq j}^p [A_{pi}X_iB_{pi} + C_{pi}X_i^T D_{pi}] - C_{pj}X_j^T D_{pj}. \end{cases} \quad (25)$$

Let

$$\Delta F_i(k) := F_i - \sum_{j=1}^p [A_{ij}X_j(k-1)B_{ij} + C_{ij}X_j^T(k-1)D_{ij}]. \quad (26)$$

Applying the gradient search method to solve (25) gives

$$X_j(k) = X_j(k-1) + \mu \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{pj} \end{bmatrix}^T \begin{bmatrix} \Delta F_1(k) \\ \Delta F_2(k) \\ \vdots \\ \Delta F_p(k) \end{bmatrix} \star [B_{1j}, B_{2j}, \dots, B_{pj}]^T, \quad j = 1, 2, \dots, p. \quad (27)$$

where $\mu > 0$ is the iterative step size or convergence factor to be given later.

Similarly, From (23) we have

$$\begin{cases} C_{1j}X_j^T D_{1j} = F_1 - \sum_{i=1, i \neq j}^p [A_{1i}X_iB_{1i} + C_{1i}X_i^T D_{1i}] - A_{1j}X_jB_{1j}, \\ C_{2j}X_j^T D_{2j} = F_2 - \sum_{i=1, i \neq j}^p [A_{2i}X_iB_{2i} + C_{2i}X_i^T D_{2i}] - A_{2j}X_jB_{2j}, \\ \vdots \\ C_{pj}X_j^T D_{pj} = F_p - \sum_{i=1, i \neq j}^p [A_{pi}X_iB_{pi} + C_{pi}X_i^T D_{pi}] - A_{pj}X_jB_{pj}. \end{cases}$$

Using the gradient search method gives

$$X_j(k) = X_j(k-1) + \mu \begin{bmatrix} D_{1j}^T \\ D_{2j}^T \\ \vdots \\ D_{pj}^T \end{bmatrix}^T \begin{bmatrix} \Delta F_1^T(k) \\ \Delta F_2^T(k) \\ \vdots \\ \Delta F_p^T(k) \end{bmatrix} \star [C_{1j}^T, C_{2j}^T, \dots, C_{pj}^T]^T, \quad j = 1, 2, \dots, p. \quad (28)$$

Taking the average of equations (27) and (28), we propose the following gradient based iterative algorithm to compute

the solutions $X_j(k)$ for (23):

$$X_{ja}(k) = X_j(k-1) + \mu \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{pj} \end{bmatrix}^T \begin{bmatrix} \Delta F_1(k) \\ \Delta F_2(k) \\ \vdots \\ \Delta F_p(k) \end{bmatrix} \star [B_{1j}, B_{2j}, \dots, B_{pj}]^T, \quad (29)$$

$$X_{jb}(k) = X_j(k-1) + \mu \begin{bmatrix} D_{1j}^T \\ D_{2j}^T \\ \vdots \\ D_{pj}^T \end{bmatrix} \begin{bmatrix} \Delta F_1^T(k) \\ \Delta F_2^T(k) \\ \vdots \\ \Delta F_p^T(k) \end{bmatrix} \star [C_{1j}^T, C_{2j}^T, \dots, C_{pj}^T]^T, \quad (30)$$

$$X_j(k) = \frac{X_{ja}(k) + X_{jb}(k)}{2}, \quad (31)$$

$$0 < \mu < 2 \left\{ \sum_{i=1}^p \sum_{j=1}^p \|A_{ij}\|^2 \|B_{ij}\|^2 + \|C_{ij}\|^2 \|D_{ij}\|^2 \right\}^{-1}. \quad (32)$$

Theorem 2: If the coupled matrix equation in (23) has unique solutions X_j , $j = 1, 2, \dots, p$, then the iterative solutions $X_j(k)$ given by the algorithm in (29)-(32) converge to the true solutions X_j for any initial values $X_j(0)$, i.e., $\lim_{k \rightarrow \infty} X_j(k) = X_j$; in other words, the error matrices $X_j(k) - X_j$ converge to zero when k is infinite.

Considering the space restrictions, the proofs of *Theorem 1* and *2* are omitted here, but they can be derived similarly as in [18], [19].

Let

$$X(k) := \begin{bmatrix} X_1(k) \\ X_2(k) \\ \vdots \\ X_p(k) \end{bmatrix}, \quad F := \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_p \end{bmatrix}, \quad \Delta F(k) := \begin{bmatrix} \Delta F_1(k) \\ \Delta F_2(k) \\ \vdots \\ \Delta F_p(k) \end{bmatrix},$$

$$X_a(k) := \begin{bmatrix} X_{1a}(k) \\ X_{2a}(k) \\ \vdots \\ X_{pa}(k) \end{bmatrix}, \quad X_b(k) := \begin{bmatrix} X_{1b}(k) \\ X_{2b}(k) \\ \vdots \\ X_{pb}(k) \end{bmatrix},$$

$$X^H(k) := [X_1(k), X_2(k), \dots, X_p(k)]^T,$$

$$F^H := [F_1, F_2, \dots, F_p]^T,$$

$$\Delta F^H(k) := [\Delta F_1(k), \Delta F_2(k), \dots, \Delta F_p(k)]^T,$$

$$I_{np \times n} := [I_n, I_n, \dots, I_n]^T.$$

By using the star product properties, a more compact form of the gradient based iterative algorithm in (29)-(31) can be written as

$$\begin{aligned} X_a(k) &= X(k-1) + \mu S_A^T \star \left\{ F - [S_A \star X(k-1) \star S_B \right. \\ &\quad \left. + S_C \star X^H(k-1) \star S_D] I_{np \times n} \right\} \star S_B^T I_{np \times n} \\ &= X(k-1) + \mu S_A^T \star \Delta F(k) \star S_B^T I_{np \times n}, \end{aligned} \quad (33)$$

$$\begin{aligned} X_b(k) &= X(k-1) + \mu S_D^T \star \left\{ F^H - (S_B^T \star X^H(k-1) \star S_A^T \right. \\ &\quad \left. + S_D^T \star X(k-1) \star S_C^T) I_{np \times n} \right\} \star S_C^T I_{np \times n} \\ &= X(k-1) + \mu S_D^T \star \Delta F^H(k) \star S_C^T I_{np \times n}, \end{aligned} \quad (34)$$

$$X(k) = \frac{X_a(k) + X_b(k)}{2}. \quad (35)$$

IV. NUMERICAL EXAMPLES

In this section, we provide two examples to validate the effectiveness of the proposed algorithms.

Example 1 Consider the coupled matrix equation in (1) with

$$\begin{aligned} A_1 &= \begin{bmatrix} 1 & 0 \\ 2 & -3 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & 2 \\ -4 & 1 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 4 & 1 \\ -3 & 5 \end{bmatrix}, \\ D_1 &= \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 5 & 0 \\ -4 & 1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 3 & -1 \\ 2 & -2 \end{bmatrix}, \\ C_2 &= \begin{bmatrix} -1 & 3 \\ -1 & 2 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 3 & 1 \\ -5 & -1 \end{bmatrix}, \\ F_1 &= \begin{bmatrix} 0 & 40 \\ -11 & 3 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 20 & 13 \\ 8 & -9 \end{bmatrix}. \end{aligned}$$

Using Lemma 2, the unique solutions of this equation can be given by

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} 2 & 8 \\ -1 & 5 \end{bmatrix}.$$

Taking $X(0) = Y(0) = 10^{-6} \mathbf{1}_{2 \times 2}$ and applying the gradient based iterative algorithm in (15)-(22) to compute $X(k)$ and $Y(k)$, the iterative errors $\delta(k) := \sqrt{(\|X(k) - X\|^2 + \|Y(k) - Y\|^2) / (\|X\|^2 + \|Y\|^2)}$ versus k with different convergence factor μ is illustrated in Fig. 1. Specially, when $\mu = 1/37$ and $k = 100$, the iterative results are

$$\begin{aligned} X(100) &= \begin{bmatrix} 0.99991 & 1.99988 \\ 3.00406 & 4.00039 \end{bmatrix}, \\ Y(100) &= \begin{bmatrix} 2.00154 & 7.99689 \\ -0.99511 & 4.99945 \end{bmatrix}, \end{aligned}$$

and $\delta(100) = 6.53519 \times 10^{-4}$.

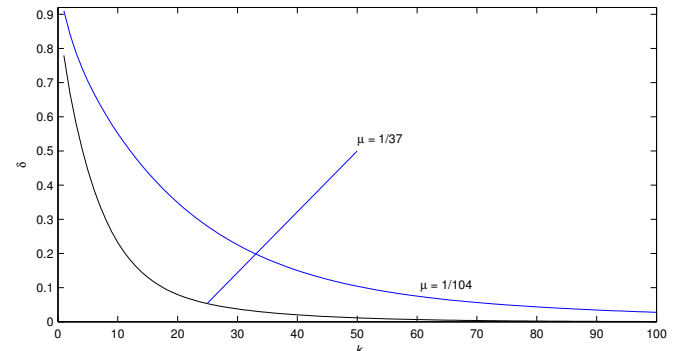


Fig. 1. The errors $\delta(k)$ versus k of Example 1

As depicted in Fig. 1, the error $\delta(k)$ decreases and converges to zero as k increases, which verifies the effectiveness of the proposed algorithm. In addition, the convergence performance associated with $\mu = 1/37$ is better than that associated with $\mu = 1/104$. This indicates that the sufficient condition given in (22) to ensure the convergence of the algorithm is very conservative, and how to choose an optimal convergence factor is the focus of our future work.

Example 2 Consider a general coupled matrix equation with the form of (23), where the coefficient matrices A_{ij} , B_{ij} , C_{ij} , D_{ij} , F_i and the unique solution X_j with $i = 1, 2$ and $j = 1, 2$ are given by

$$\begin{aligned} A_{11} &= \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}, B_{11} = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}, C_{11} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}, \\ D_{11} &= \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}, A_{12} = \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}, B_{12} = \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}, \\ C_{12} &= \begin{bmatrix} 2 & -3 \\ 0 & 2 \end{bmatrix}, D_{12} = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, A_{21} = \begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix}, \\ B_{21} &= \begin{bmatrix} 1 & 1 \\ 3 & -5 \end{bmatrix}, C_{21} = \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix}, D_{21} = \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix}, \\ A_{22} &= \begin{bmatrix} -1 & 3 \\ -1 & 2 \end{bmatrix}, B_{22} = \begin{bmatrix} -2 & 3 \\ -1 & 3 \end{bmatrix}, C_{22} = \begin{bmatrix} -1 & 1 \\ -7 & 0 \end{bmatrix}, \\ D_{22} &= \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, F_1 = \begin{bmatrix} -63 & 23 \\ 61 & 9 \end{bmatrix}, F_2 = \begin{bmatrix} 4 & -7 \\ 23 & -24 \end{bmatrix}, \\ X_1 &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, X_2 = \begin{bmatrix} 2 & 8 \\ -1 & 5 \end{bmatrix}. \end{aligned}$$

The gradient based iterative algorithm in (29)-(32) for initial values $X_1(0) = X_2(0) = 10^{-6} \mathbf{1}_{2 \times 2}$ is applied to solve this coupled matrix equation. For different convergence factor $\mu = 1/1470$ and $1/185$, the simulation results are shown in Fig. 2.

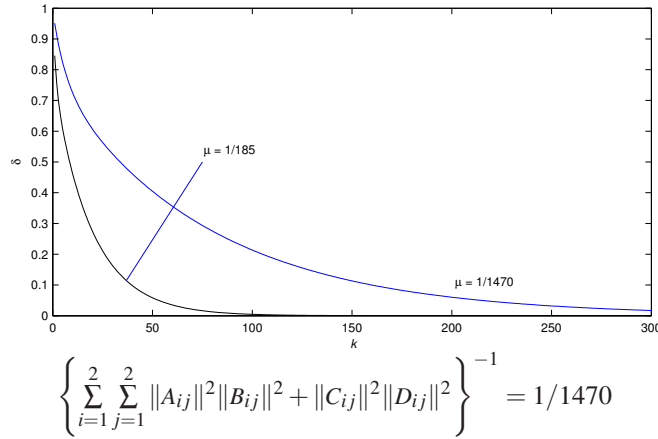


Fig. 2. The errors $\delta(k)$ versus k of Example 2

From Fig. 2, we can see that the convergence rate of the algorithm depends on the convergence factor μ , and a larger value leads to a faster rate. For $\mu = 1/185$, we have

$$\begin{aligned} X(300) &= \begin{bmatrix} 1.00000 & 1.99994 \\ 3.00024 & 3.99986 \end{bmatrix}, \\ Y(300) &= \begin{bmatrix} 2.00000 & 8.00001 \\ -1.00005 & 4.99996 \end{bmatrix}, \end{aligned}$$

and $\delta(300) = 2.62284 \times 10^{-5}$.

V. CONCLUSIONS

The gradient based iterative algorithms are developed to solve the general coupled matrix equations with real coefficients by applying the hierarchical identification principle and the gradient search method. The given simulation results

well demonstrate that the proposed algorithms have good convergence properties and high accuracy. Moreover, the problem studied in this paper is quite general, thus the proposed algorithms are also applicable to solving its special cases, such as ones in [18], [19].

The basic idea of the proposed algorithm can be applied to study identification problems of time-varying systems [38], nonlinear systems [39]–[43], dual-rate/multirate systems [44]–[67], as well as to design filters [68].

REFERENCES

- [1] L. Xie, Y.J. Liu, H.Z. Yang, Gradient based and least squares based iterative algorithms for matrix equations $AXB+CX^TD=F$, *Applied Mathematics and Computation* 217 (5) (2010) 2191-2199.
- [2] J. Ding, Y.J. Liu, F. Ding, Iterative solutions to matrix equations of form $AiXB_i=F_i$, *Computers & Mathematics with Applications* 59 (11) (2010) 3500-3507.
- [3] F. Ding, Transformations between some special matrices, *Computers & Mathematics with Applications* 59 (8) (2010) 2676-2695.
- [4] M. Dehghan, M. Hajarian, An efficient algorithm for solving general coupled matrix equations and its application, *Mathematical and Computer Modelling* 51 (9-10) (2010) 1118-1134.
- [5] M. Dehghan, M. Hajarian, An iterative method for solving the generalized coupled Sylvester matrix equations over generalized bisymmetric matrices, *Applied Mathematics Modelling* 34 (3) (2010) 639-654.
- [6] Z.Y. Li, Y. Wang, B. Zhou, G.R. Duan, Least squares solution with the minimum-norm to general matrix equations via iteration, *Applied Mathematics and Computation* 215 (10) (2010) 3547-3562.
- [7] A.G. Wu, B. Li, Y. Zhang, G.R. Duan, Finite iterative solutions to coupled Sylvester-conjugate matrix equations, *Applied Mathematics Modelling* 35 (3) (2011) 1065-1080.
- [8] F.L. Li, L.S. Gong, X.Y. Hu, L. Zhang, Successive projection iterative method for solving matrix equation $AX=B$, *Journal of Computational and Applied Mathematics* 234 (8) (2010) 2405-2410.
- [9] D.F. Li, C.J. Zhang, Split Newton iterative algorithm and its application, *Applied Mathematics and Computation* 217 (5) (2010) 2260-2265.
- [10] M. Dehghan, M. Hajarian, The general coupled matrix equations over generalized bisymmetric matrices, *Linear Algebra and its Applications* 432 (6) (2010) 1531-1552.
- [11] M. Dehghan, M. Hajarian, Finite iterative algorithms for the reflexive and anti-reflexive solutions of the matrix equation $A_1X_1B_1+A_2X_2B_2=C$, *Mathematical and Computer Modelling* 49 (9-10) (2009) 1937-1959.
- [12] B. Zhou, Z.Y. Li, G.R. Duan, Y. Wang, Weighted least squares solutions to general coupled Sylvester matrix equations, *Journal of Computational and Applied Mathematics* 224 (2) (2010) 759-776.
- [13] A.G. Wu, G. Feng, G.R. Duan, W.J. Wu, Finite iterative solutions to a class of complex matrix equations with conjugate and transpose of the unknowns, *Mathematical and Computer Modelling* 52 (9-10) (2010) 1463-1478.
- [14] F. Ding, P.X. Liu, J. Ding, Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle, *Applied Mathematics and Computation* 197 (1) (2008) 41-50.
- [15] A.G. Wu, X.L. Zeng, G.R. Duan, W.J. Wu, Iterative solutions to the extended Sylvester-conjugate matrix equations, *Applied Mathematics and Computation* 217 (1) (2010) 130-142.
- [16] F. Ding, T. Chen, Gradient based iterative algorithms for solving a class of matrix equations, *IEEE Transactions on Automatic Control* 50 (8) (2005) 1216-1221.
- [17] L. Xie, J. Ding, F. Ding, Gradient based iterative solutions for general linear matrix equations, *Computers & Mathematics with Applications* 58 (7) (2009) 1441-1448.
- [18] F. Ding, T. Chen, Iterative least squares solutions of coupled Sylvester matrix equations, *Systems & Control Letters* 54 (2) (2005) 95-107.
- [19] F. Ding, T. Chen, On iterative solutions of general coupled matrix equations, *SIAM Journal on Control and Optimization* 44 (6) (2006) 2269-2284.
- [20] F. Ding, T. Chen, Performance analysis of multi-innovation gradient type identification methods, *Automatica* 43 (1) (2007) 1-14.

- [21] F. Ding, Several multi-innovation identification methods, *Digital Signal Processing* 20 (4) (2010) 1027-1039.
- [22] D.Q. Wang, F. Ding, Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems, *Digital Signal Processing* 20 (3) (2010) 750-762.
- [23] D.Q. Wang, Y.Y. Chu, F. Ding, Auxiliary model-based RELS and MI-ELS algorithms for Hammerstein OEMA systems, *Computers & Mathematics with Applications* 59 (9) (2010) 3092-3098.
- [24] L.L. Han, F. Ding, Identification for multirate multi-input systems using the multi-innovation identification theory, *Computers & Mathematics with Applications* 57 (9) (2009) 1438-1449.
- [25] L.L. Han, F. Ding, Multi-innovation stochastic gradient algorithms for multi-input multi-output systems, *Digital Signal Processing* 19 (4) (2009) 545-554.
- [26] J.B. Zhang, F. Ding, Y. Shi, Self-tuning control based on multi-innovation stochastic gradient parameter estimation, *Systems & Control Letters* 58 (1) (2009) 69-75.
- [27] F. Ding, P.X. Liu, G. Liu, Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises, *Signal Processing* 89 (10) (2009) 1883-1890.
- [28] Y.J. Liu, Y.S. Xiao, X.L. Zhao, Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model, *Applied Mathematics and Computation* 215 (4) (2009) 1477-1483.
- [29] L. Xie, H.Z. Yang, F. Ding, Modeling and identification for non-uniformly periodically sampled-data systems, *IET Control Theory & Applications* 4 (5) (2010) 784-794.
- [30] F. Ding, P.X. Liu, G. Liu, Multi-innovation least squares identification for system modeling, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 40 (3) (2010) 767-778.
- [31] Y.J. Liu, L. Yu, F. Ding, Multi-innovation extended stochastic gradient algorithm and its performance analysis, *Circuits, Systems and Signal Processing* 29 (4) (2010) 649-667.
- [32] J. Chen, Y. Zhang, R.F. Ding, Auxiliary model based multi-innovation algorithms for multivariable nonlinear systems, *Mathematical and Computer Modelling* 52 (9-10) (2010) 1428-1434.
- [33] F. Ding, P.X. Liu, G. Liu, Gradient based and least-squares based iterative identification methods for OE and OEMA systems, *Digital Signal Processing* 20 (3) (2010) 664-677.
- [34] F. Ding, T. Chen, Identification of Hammerstein nonlinear ARMAX systems, *Automatica* 41 (9) (2005) 1479-1489.
- [35] Y.J. Liu, D.Q. Wang, F. Ding, Least-squares based iterative algorithms for identifying Box-Jenkins models with finite measurement data, *Digital Signal Processing* 20 (5) (2010) 1458-1467.
- [36] D.Q. Wang, G.W. Yang, R.F. Ding, Gradient-based iterative parameter estimation for Box-Jenkins systems, *Computers & Mathematics with Applications* 60 (5) (2010) 1200-1208.
- [37] X.G. Liu, J. Lu, Least squares based iterative identification for a class of multirate systems, *Automatica* 46 (3) (2010) 549-554.
- [38] F. Ding, T. Chen, Performance bounds of the forgetting factor least squares algorithm for time-varying systems with finite measurement data, *IEEE Transactions on Circuits and Systems-I: Regular Papers* 52 (3) (2005) 555-566.
- [39] D.Q. Wang, Y.Y. Chu, G.W. Yang, F. Ding, Auxiliary model-based recursive generalized least squares parameter estimation for Hammerstein OEAR systems, *Mathematical and Computer Modelling* 52 (1-2) (2010) 309-317.
- [40] F. Ding, Y. Shi, T. Chen, Auxiliary model based least-squares identification methods for Hammerstein output-error systems, *Systems & Control Letters* 56 (5) (2007) 373-380.
- [41] F. Ding, T. Chen, Z. Iwai, Adaptive digital control of Hammerstein nonlinear systems with limited output sampling, *SIAM Journal on Control and Optimization* 45 (6) (2006) 2257-2276.
- [42] F. Ding, P.X. Liu, G. Liu, Identification methods for Hammerstein nonlinear systems, *Digital Signal Processing* 21 (2) (2011) 215-238.
- [43] D.Q. Wang, F. Ding, Least squares based and gradient based iterative identification for Wiener nonlinear systems, *Signal Processing* 91 (5) (2011) 1182-1189.
- [44] F. Ding, T. Chen, Combined parameter and output estimation of dual-rate systems using an auxiliary model, *Automatica* 40 (10) (2004) 1739-1748.
- [45] F. Ding, T. Chen, Parameter estimation of dual-rate stochastic systems by using an output error method, *IEEE Transactions on Automatic Control* 50 (9) (2005) 1436-1441.
- [46] F. Ding, T. Chen, Hierarchical identification of lifted state-space models for general dual-rate systems, *IEEE Transactions on Circuits and Systems-I: Regular Papers* 52 (6) (2005) 1179-1187.
- [47] J. Ding, Y. Shi, H.G. Wang, F. Ding, A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems, *Digital Signal Processing* 20 (4) (2010) 1238-1249.
- [48] L.L. Han, J. Sheng, F. Ding, Y. Shi, Auxiliary models based recursive least squares identification for multirate multi-input systems, *Mathematical and Computer Modelling* 50 (7-8) (2009) 1100-1106.
- [49] Y.J. Liu, L. Xie, F. Ding, An auxiliary model based recursive least squares parameter estimation algorithm for non-uniformly sampled multirate systems, *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 223 (4) (2009) 445-454.
- [50] F. Ding, L. Qiu, T. Chen, Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems, *Automatica* 45 (2) (2009) 324-332.
- [51] F. Ding, G. Liu, X.P. Liu, Partially coupled stochastic gradient identification methods for non-uniformly sampled systems, *IEEE Transactions on Automatic Control*, 2010, 55(8): 1976-1981.
- [52] Y. Shi, F. Ding, T. Chen, Multirate crosstalk identification in xDSL systems, *IEEE Transactions on Communications* 54 (10) (2006) 1878-1886.
- [53] J. Ding, F. Ding, The residual based extended least squares identification method for dual-rate systems, *Computers & Mathematics with Applications* 56 (6) (2008) 1479-1487.
- [54] Y.S. Xiao, F. Ding, Y. Zhou, M. Li, J.Y. Dai, On consistency of recursive least squares identification algorithms for controlled autoregression models, *Applied Mathematical Modelling* 32 (11) (2008) 2207-2215.
- [55] J. Ding, L.L. Han, X.M. Chen, Time series AR modeling with missing observations based on the polynomial transformation, *Mathematical and Computer Modelling* 51 (5-6) (2010) 527-536.
- [56] F. Ding, H.Z. Yang, F. Liu, Performance analysis of stochastic gradient algorithms under weak conditions, *Science in China Series F-Information Sciences* 51 (9) (2008) 1269-1280.
- [57] F. Ding, P.X. Liu, H.Z. Yang, Parameter identification and intersample output estimation for dual-rate systems, *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* 38 (4) (2008) 966-975.
- [58] Y.J. Liu, J. Sheng, R.F. Ding, Convergence of stochastic gradient estimation algorithm for multivariable ARX-like systems, *Computers & Mathematics with Applications* 59 (8) (2010) 2615-2627.
- [59] F. Ding, Y. Shi, T. Chen, Performance analysis of estimation algorithms of non-stationary ARMA processes, *IEEE Transactions on Signal Processing* 54 (3) (2006) 1041-1053.
- [60] J. Ding, F. Ding, S. Zhang, Parameter identification of multi-input, single-output systems based on FIR models and least squares principle, *Applied Mathematics and Computation* 197 (1) (2008) 297-305.
- [61] Y.S. Xiao, H.B. Chen, F. Ding, Identification of multi-input systems based on the correlation techniques, *International Journal of Systems Science* 42 (1) (2011) 139-147.
- [62] F. Ding, J. Ding, Least squares parameter estimation with irregularly missing data, *International Journal of Adaptive Control and Signal Processing* 24 (7) (2010) 540-553.
- [63] L. Chen, J.H. Li, R.F. Ding, Identification of the second-order systems based on the step response, *Mathematical and Computer Modelling* 53 (5-6) (2011) 1074-1083.
- [64] B. Bao, Y.Q. Xu, J. Sheng, R.F. Ding, Least squares based iterative parameter estimation algorithm for multivariable controlled ARMA system modelling with finite measurement data, *Mathematical and Computer Modelling* 53 (9-10) (2011) 1664-1669.
- [65] H.Q. Han, G.L. Song, Y.S. Xiao, Y.W. Liao, R.F. Ding, Performance analysis of the AM-SG parameter estimation for multivariable systems, *Applied Mathematics and Computation* 217 (12) (2011) 5566-5572.
- [66] H.H. Yin, Z.F. Zhu, F. Ding, Model order determination using the Hankel matrix of impulse responses, *Applied Mathematics Letters* 24 (5) (2011) 797-802.
- [67] Z.N. Zhang, F. Ding, X.G. Liu, Hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems, *Computers & Mathematics with Applications* 61 (3) (2011) 672-682.
- [68] Y. Shi, F. Ding, T. Chen, 2-Norm based recursive design of transmultiplexers with designable filter length, *Circuits, Systems and Signal Processing* 25 (4) (2006) 447-462.