

SUBSPACE IDENTIFICATION METHOD FOR COMBINED DETERMINISTIC-STOCHASTIC BILINEAR SYSTEMS

Huixin Chen and Jan Maciejowski¹

*Department of Engineering
University of Cambridge
Cambridge CB2 1PZ, U.K.*

Abstract: In this paper, a ‘four-block’ subspace system identification method for combined deterministic-stochastic bilinear systems is developed. Estimation of state sequences, followed by estimation of system matrices, is the central component of subspace identification methods. The prominent difference of our new approach is a ‘four-block’ arrangement of data matrices which leads to a linearization of the system state equation, when written in block form. A major advantage of this approach, over a previous bilinear subspace algorithm, is that the measured input is not restricted to be white. We show that, providing a certain data-dependent eigenvalue condition is met, our algorithm provides asymptotically unbiased estimates, and we indicate the rate at which the bias decreases. Simulation results show that this algorithm requires a smaller sample size than earlier algorithms (for comparable performance) and that the computational complexity is significantly lower. *Copyright*©2000 *IFAC*

Keywords: System Identification, Bilinear Systems, Subspace Methods, State Sequence Estimation, State-space Models.

1. INTRODUCTION

There has been much study of the identification problem of bilinear systems. A three-block subspace identification method for deterministic bilinear system is presented in (Chen and Maciejowski, 1999). In this paper we present a very closely related method for identifying a combined deterministic-stochastic bilinear state-space model. A method for subspace identification of combined deterministic-stochastic bilinear systems subject to white noise inputs was derived in (Favoreel *et al.*, 1996), (Favoreel

et al., 1997), (Favoreel *et al.*, 1999), which was the first effort to use subspace methodology to identify the state-space model of a bilinear system in the stochastic case. A subspace identification algorithm which can be used with a non-white input was proposed in (Favoreel and De Moor, 1998), but it results in biased estimates of the system. A ‘three-block’ subspace identification algorithm is given in (Chui and Maciejowski, 1998), which can remove the effects of unmeasured noise sources and obtain accurate estimates. In the linear system case, the three-block algorithm is linked with the system’s Markov parameters. Unfortunately, the realization

¹ Corresponding author.

theory for bilinear systems is more complicated than for linear systems. In our ‘four-block’ subspace method for bilinear systems, the four-block data matrix arrangement and the observation matrix equation are used to linearise the system equation in the block matrix form using linear and bilinear algebra. This allows bilinear models to be obtained from row and column spaces of certain matrices, calculated from the input-output data by means of some bilinear-algebraic operations. All the proofs are omitted here, but can be found in (Chen and Maciejowski, 2000).

2. NOTATIONS

The use of much specialised notation seems to be unavoidable in the current context. Mostly we follow the notation used in (Favoreel *et al.*, 1999), (Chen and Maciejowski, 1999), but we introduce all the notation here for completeness.

We use \otimes to denote the Kronecker product and \odot the Khatri-Rao product of two matrices with $F \in \mathbf{R}^{t \times p}$ and $G \in \mathbf{R}^{u \times p}$ defined in (Khatri and Rao, 1968), (Suda *et al.*, 1973):

$$F \odot G \triangleq [f_1 \otimes g_1, f_2 \otimes g_2, \dots, f_p \otimes g_p]$$

$+$, \oplus and \cap denote the sum, the direct sum and the intersection of two vector spaces, \cdot^\perp denotes the orthogonal complement of a subspace with respect to the predefined ambient space, the Moore-Penrose inverse is written as \cdot^\dagger , and the Hermitian as \cdot^* .

In this paper we consider combined deterministic-stochastic time-invariant bilinear system of the form:

$$\begin{aligned} x_{t+1} &= Ax_t + Nu_t \otimes x_t + Bu_t + w_t \\ y_t &= Cx_t + Du_t + v_t \end{aligned} \quad (1)$$

where $x_t \in \mathbf{R}^n$, $y_t \in \mathbf{R}^l$, $u_t \in \mathbf{R}^m$, and $N = [N_1 \ N_2 \ \dots \ N_m] \in \mathbf{R}^{n \times nm}$, $N_i \in \mathbf{R}^{n \times n}$ ($i = 1, \dots, m$).

The input u_t is assumed to be independent of the measurement noise v_t and the process noise w_t . The covariance matrix of w_t and v_t is:

$$\mathbf{E} \left[\begin{pmatrix} w_p \\ v_p \end{pmatrix} \begin{pmatrix} w_q \\ v_q \end{pmatrix}^T \right] = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{pq} \geq 0$$

We assume that the sample size is \tilde{N} , namely that input-output data $\{u(t), y(t) : t = 0, 1, \dots, \tilde{N}\}$ are available. For arbitrary t we define

$$X_t \triangleq [x_t \ x_{t+1} \ \dots \ x_{t+j-1}] \in \mathbf{R}^{n \times j}$$

but for the special cases $t = 0$, $t = k$, $t = 2k$ and $t = 3k$ we define, with some abuse of notation,

$$\begin{aligned} X_p &\triangleq [x_0 \ x_1 \ \dots \ x_{j-1}] \in \mathbf{R}^{n \times j} \\ X_c &\triangleq [x_k \ x_{k+1} \ \dots \ x_{k+j-1}] \in \mathbf{R}^{n \times j} \\ X_f &\triangleq [x_{2k} \ x_{2k+1} \ \dots \ x_{2k+j-1}] \in \mathbf{R}^{n \times j} \\ X_r &\triangleq [x_{3k} \ x_{3k+1} \ \dots \ x_{3k+j-1}] \in \mathbf{R}^{n \times j} \end{aligned}$$

where k is the *row block size*. The suffices p , c , f and r are supposed to be mnemonic, representing ‘past’, ‘current’, ‘future’ and ‘remote future’ respectively. This division of the state history into four (overlapping) segments is the reason for referring to our algorithm as a ‘four-block’ method. In the deterministic case the ‘remote future’ segment is not necessary, and a simpler ‘three-block’ method results (Chen and Maciejowski, 1999).

We define U_t , U_p , U_f , U_r , Y_t , Y_p , Y_f , Y_r , W_t , U_p , W_f , W_r , V_t , V_p , V_f , V_r , similarly. These matrices will later be used to construct larger matrices with a ‘generalised block-Hankel’ structure. In order to use all the available data in these, the number of columns j is such that $\tilde{N} = 3k + j - 1$ and let $d_i = \sum_{p=1}^i (m+1)^{p-1} l$, $e_i = \sum_{p=1}^i (m+1)^{p-1} m$ and $f_k = e_k + \frac{m}{2}(m+1)^k + l[(m+1)^k - 1] + e_k^2$.

For arbitrary q and $i \geq q + 2$, we define

$$\begin{aligned} X_{q|q} &\triangleq \begin{pmatrix} X_q \\ U_q \odot X_q \end{pmatrix} \in \mathbf{R}^{(m+1)n \times j} \\ X_{i-1|q} &\triangleq \begin{pmatrix} X_{i-2|q} \\ U_{i-1} \odot X_{i-2|q} \end{pmatrix} \in \mathbf{R}^{(m+1)^{i-q}n \times j} \\ Y_{q|q} &\triangleq Y_q \\ Y_{i-1|q} &\triangleq \begin{pmatrix} Y_{i-1} \\ Y_{i-2|q} \\ U_{i-1} \odot Y_{i-2|q} \end{pmatrix} \in \mathbf{R}^{d_{i-q} \times j} \\ U_{q|q}^+ &\triangleq U_q \\ U_{i-1|q}^+ &\triangleq \begin{pmatrix} U_{i-2}^+ \\ U_{i-1} \\ U_{i-1} \odot U_{i-2|q}^+ \end{pmatrix} \in \mathbf{R}^{((m+1)^{i-q}-1) \times j} \\ U_{q|q}^{++} &\triangleq \begin{pmatrix} U_{q,1} \odot U_q \\ U_{q,2} \odot U_q(2:m,:) \\ U_{q,3} \odot U_q(3:m,:) \\ \vdots \\ U_{q,m} \odot U_{q,m} \end{pmatrix} \in \mathbf{R}^{\frac{m(m+1)}{2} \times j} \\ U_{i-1|q}^{++} &\triangleq \begin{pmatrix} U_{i-2|q}^{++} \\ U_{i-1} \odot U_{i-2|q}^{++} \end{pmatrix} \in \mathbf{R}^{\frac{m}{2}(m+1)^{i-q} \times j} \end{aligned}$$

$$U_{i-1|q}^y \triangleq U_{i-1|q}^+ \odot Y_q$$

$$U_{i+k-1|k+q}^{u,y} \triangleq \begin{pmatrix} U_{i+k-1|k+q}^{++} \\ U_{i+k-1|k+q}^y \\ U_{i+k-1|k+q}^+ \odot U_{i-1|q} \end{pmatrix}$$

$U_{i-1|q}$, $W_{i-1|q}$ and $V_{i-1|q}$ and can be defined similarly.

Remark 1. The meaning of $U_{i-1|q}^+$ is different from that in (Favoreel *et al.*, 1996). $U_{i-1|q}^{++}$ is newly introduced here.

$$X^c \triangleq X_{2k-1|k}, X^f \triangleq X_{3k-1|2k}, X^r \triangleq X_{4k-1|3k}$$

$$U^p \triangleq U_{k-1|0}, U^c \triangleq U_{2k-1|k}, U^f \triangleq U_{3k-1|2k}$$

$$U^{p,y} \triangleq U^p \odot Y_p, U^{c,y} \triangleq U^c \odot Y_c$$

$$U^{f,y} \triangleq U^f \odot Y_f, U^{r,y} \triangleq U^r \odot Y_r,$$

$$U^{c,u,y} \triangleq \begin{pmatrix} U^c \\ U^{++c} \\ U^{c,y} \\ U^c \odot U^p \end{pmatrix}, U^{f,u,y} \triangleq \begin{pmatrix} U^f \\ U^{++f} \\ U^{f,y} \\ U^f \odot U^c \end{pmatrix}$$

$U^r, Y^p, Y^c, Y^f, Y^r, W^c, W^f, W^r, V^c, V^f, V^r, U^c, U^f, U^r, U^{++c}, U^{++f}, U^{++r}$ and $U^{r,u,y}$ can be defined similarly. Finally, we denote by \mathcal{U}_p the space spanned by all the rows of the matrix U_p . That is,

$$\mathcal{U}_p := \text{span}\{\alpha^* U_p, \alpha \in \mathbf{R}^{km}\}$$

$\mathcal{U}_c, \mathcal{U}_f, \mathcal{U}_r, \mathcal{Y}_p, \mathcal{Y}_c, \mathcal{Y}_f, \mathcal{Y}_r, \mathcal{U}^p, \mathcal{Y}^p, \mathcal{U}^f, \mathcal{Y}^f, \mathcal{U}^{p,u,y}, \mathcal{U}^{f,u,y}$ etc are defined similarly.

3. ANALYSIS

Lemma 1. The system (1) can be rewritten in the following matrix equation form:

$$X_{t+1} = AX_t + NU_t \odot X_t + BU_t + W_t$$

$$Y_t = CX_t + DU_t + V_t \quad (2)$$

Lemma 2. For $j \geq 0$, and the block size k , we have

$$X_{k-1+j|j} = \begin{pmatrix} X_j \\ U_{k-1+j|j}^+ \odot X_j \end{pmatrix}$$

Lemma 3. For F, G, H, J of compatible dimensions, $F \in \mathbf{R}^{k \times l}$, $G \in \mathbf{R}^{l \times m}$, $H \in \mathbf{R}^{p \times l}$, $J \in \mathbf{R}^{l \times m}$:

$$(FG \otimes HJ) = (F \otimes H)(G \otimes J)$$

$$(FG \odot HJ) = (F \otimes H)(G \odot J)$$

Lemma 4. (Input-Output Equation). For the combined deterministic-stochastic system (1) and $j \geq 0$, we have the following Input-Output Equation

$$X_{k+j} = \Delta_k^X X_{k-1+j|j} + \Delta_k^U U_{k-1+j|j} + \Delta_k^W W_{k-1+j|j}$$

$$Y_{k-1+j|j} = \mathcal{L}_k^X X_{k-1+j|j} + \mathcal{L}_k^U U_{k-1+j|j} + \mathcal{L}_k^W W_{k-1+j|j} + \mathcal{L}_k^V V_{k-1+j|j}$$

where

$$\Delta_n^X \triangleq [A\Delta_{n-1}^X, N_1\Delta_{n-1}^X, \dots, N_m\Delta_{n-1}^X]$$

$$\Delta_1^X \triangleq [A, N_1, \dots, N_m]$$

$$\Delta_n^U \triangleq [B, A\Delta_{n-1}^U, N_1\Delta_{n-1}^U, \dots, N_m\Delta_{n-1}^U], \Delta_1^U \triangleq B$$

$$\Delta_n^W \triangleq [I_{n \times n}, A\Delta_{n-1}^W, N_1\Delta_{n-1}^W, \dots, N_m\Delta_{n-1}^W]$$

$$\Delta_1^W \triangleq I_{n \times n}$$

$$\mathcal{L}_k^X \triangleq \begin{bmatrix} C\Delta_{k-1}^X & 0 \\ \mathcal{L}_{k-1}^X & 0 \\ 0 & \mathcal{L}_{k-1}^X \end{bmatrix}, \mathcal{L}_k^U \triangleq \begin{bmatrix} D & C\Delta_{k-1}^U & 0 \\ 0 & \mathcal{L}_{k-1}^U & 0 \\ 0 & 0 & \mathcal{L}_{k-1}^U \end{bmatrix}$$

$$\mathcal{L}_k^W \triangleq \begin{bmatrix} 0 & C\Delta_{k-1}^W & 0 \\ 0 & \mathcal{L}_{k-1}^W & 0 \\ 0 & 0 & \mathcal{L}_{k-1}^W \end{bmatrix}, \mathcal{L}_k^V \triangleq \begin{bmatrix} I_{l \times l} & 0 & 0 \\ 0 & \mathcal{L}_{k-1}^V & 0 \\ 0 & 0 & \mathcal{L}_{k-1}^V \end{bmatrix}$$

with

$$\mathcal{L}_1^X \triangleq [C, 0_{l \times m}], \mathcal{L}_1^U \triangleq D, \mathcal{L}_1^W \triangleq 0_{l \times n}, \mathcal{L}_1^V \triangleq I_{l \times l}$$

Lemma 5. For system (1), if

$$\lambda = \max_{j=0, \dots, \bar{N}} |\text{eig}(A + \sum_{i=1}^n u_{j,i} N_i)| < 1, \quad (3)$$

then

$$X_c = E(Y_c - DU_c - V_c) + (I - EC)\Delta_n^U U^p + o(\lambda^k)$$

$$X_f = E(Y_f - DU_f - V_f) + (I - EC)\Delta_n^U U^c + o(\lambda^k)$$

$$X_r = E(Y_r - DU_r - V_r) + (I - EC)\Delta_n^U U^f + o(\lambda^k)$$

where $o(\lambda^k)$ is used to denote a matrix M , such that $\|M\|_1 = o(\lambda^k)$.

Remark 2 This holds for any matrix E of compatible dimensions. In particular, it holds for $E = C^\dagger$, where $CC^\dagger = I$, and if $l \geq n$, then $I - C^\dagger C = 0$ and the expression become exact. In the sequel, we will assume the Moore-Penrose pseudo-inverse is used.

Theorem 1. The system (1) can be written in the following form if the condition (3) holds:

$$Y^c = \mathcal{O}_k X_c + \mathcal{T}_k^u U^{c,u,y} + \mathcal{T}_k^v U^{+c} \odot V_c + \mathcal{L}_k^W W^c + \mathcal{L}_k^V V^c + o(\lambda^k) \quad (4)$$

$$Y^f = \mathcal{O}_k X_f + \mathcal{T}_k^u U^{f,u,y} + \mathcal{T}_k^v U^{+f} \odot V_f + \mathcal{L}_k^W W^f + \mathcal{L}_k^V V^f + o(\lambda^k) \quad (5)$$

$$Y^r = \mathcal{O}_k X_r + \mathcal{T}_k^u U^{r,u,y} + \mathcal{T}_k^v U^{+r} \odot V_r + \mathcal{L}_k^W W^r + \mathcal{L}_k^V V^r + o(\lambda^k) \quad (6)$$

$$X_f = \mathcal{F}_k X_c + \mathcal{G}_k^u U^{c,u,y} + \mathcal{G}_k^v U^{+c} \odot V_c + \Delta_k^W W^c + o(\lambda^k)$$

$$X_r = \mathcal{F}_k X_f + \mathcal{G}_k^u U^{f,u,y} + \mathcal{G}_k^v U^{+f} \odot V_f + \Delta_k^W W^f + o(\lambda^k)$$

where $\mathcal{O}_k, \mathcal{T}_k^u, \mathcal{T}_k^v, \mathcal{F}_k, \mathcal{G}_k^u$ and \mathcal{G}_k^v are system-dependent constant matrices.

We make the assumption that due to ergodicity and stationarity of the variables, all the covariances used in this paper can be estimated by replacing ensemble means by time means. Therefore, we define the expectation operator \mathbf{E} is equivalent to the operator \mathbf{E}_j which is defined as follows:

$$\mathbf{E}[\bullet] = \lim_{j \rightarrow \infty} \mathbf{E}_j[\bullet] = \lim_{j \rightarrow \infty} \frac{1}{j} \sum_{k=0}^j [\bullet]$$

The orthogonal projection operator $\Pi_B A$ is defined as in (Favoreel *et al.*, 1996): $\Pi_B A \triangleq AB^T(BB^T)^\dagger B$

Theorem 2. If the linear part of the system (1) is observable and

$$\begin{pmatrix} Y^c \\ U^{c,u,y} \\ U^{f,u,y} \\ U^{r,u,y} \end{pmatrix} \quad (7)$$

is a full row rank matrix, denoting $\mathcal{S} := \mathcal{Y}^c + \mathcal{U}^{c,u,y} + \mathcal{U}^{f,u,y} + \mathcal{U}^{r,u,y}$ and $\mathcal{R} = \Pi_S \mathcal{Y}_f + \mathcal{U}^{f,u,y}$, then

$$\Pi_{\mathcal{R}^\perp} \Pi_S \mathcal{Y}^r = \mathcal{T}_k^u \Pi_{\mathcal{R}^\perp} U^{r,u,y} + o(\lambda^k) \quad (8)$$

4. ALGORITHM

1. Decompose Y^r into $\mathcal{O}_k X_r$ and $\mathcal{T}_k^u U^{r,u,y}$ using orthogonal projection: from (8) of Theorem 2, estimated $\hat{\mathcal{T}}_k^u$

$$\hat{\mathcal{T}}_k^u = (\Pi_{\mathcal{R}^\perp} \Pi_S \mathcal{Y}^r) (\Pi_{\mathcal{R}^\perp} U^{r,u,y})^\dagger \quad (9)$$

2. Obtain the SVD decomposition and partition as

$$\begin{aligned} & [\Pi_S Y_{2k-1|k} \Pi_S Y_{2k|k+1}] - \hat{\mathcal{T}}_k^u \begin{bmatrix} U_{2k-1|k}^{u,y} & U_{2k|k+1}^{u,y} \end{bmatrix} \\ & =: \Gamma \Sigma \Omega^* = \begin{bmatrix} \Gamma_1 & \Gamma_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \Omega_1^* \\ \Omega_2^* \end{bmatrix} \end{aligned}$$

Since we expect

$$\Gamma \Sigma \Omega^* = \Gamma_1 \Sigma_1 \Omega_1^* = \mathcal{O}_k \begin{bmatrix} X_{k-1} & X_k \end{bmatrix}$$

from (4-6) ($\text{rank}(\Sigma_1) = n$ and $\text{rank}(\Sigma_2) = 0$), form the estimates $\hat{\mathcal{O}}_k = \Gamma_1 \Sigma_1^{1/2}$ and $[\hat{X}_{k-1} \ \hat{X}_k] = \Sigma_1^{1/2} \Omega_1^*$, retaining only \hat{n} significant singular values in Σ_1 . ($\hat{\mathcal{O}}_k$ is not needed later.)

3. Estimate the parameters A, B, C, D, N on the basis of equation (2), as

$$\begin{bmatrix} \hat{X}_k \\ Y_{k-1} \end{bmatrix} = \begin{bmatrix} A & N & B \\ C & 0 & D \end{bmatrix} \begin{bmatrix} \hat{X}_{k-1} \\ U_{k-1} \odot \hat{X}_{k-1} \\ U_{k-1} \end{bmatrix} \quad (10)$$

in a least-squares sense.

Remark 3 Other estimates could be obtained by using other right-inverses in steps 1 and 3, and another factorisation in step 2. In (Favoreel *et al.*, 1999) it is suggested that constrained least-squares could be used in step 3, because of the known structure of the solution. Our initial simulation experience is that this does not have much effect on the estimated eigenvalues of matrix A and N .

Estimate the covariance matrix by calculating

$$\begin{aligned} \begin{bmatrix} \epsilon_w \\ \epsilon_v \end{bmatrix} &= \begin{bmatrix} \hat{X}_{k+1} \\ Y_k \end{bmatrix} - \begin{bmatrix} \hat{A} & \hat{N} & \hat{B} \\ \hat{C} & 0 & \hat{D} \end{bmatrix} \begin{bmatrix} \hat{X}_k \\ U_k \odot \hat{X}_k \\ U_k \end{bmatrix} \\ \begin{bmatrix} \hat{Q} & \hat{S} \\ \hat{S}^T & \hat{R} \end{bmatrix} &= \mathbf{E}_j \left[\begin{pmatrix} \epsilon_w \\ \epsilon_v \end{pmatrix} \begin{pmatrix} \epsilon_w \\ \epsilon_v \end{pmatrix}^* \right] \end{aligned}$$

5. EXAMPLES

In this section, two simple second order bilinear systems introduced in (Favoreel *et al.*, 1996), (Favoreel *et al.*, 1999) are given to see how the new algorithm works and how it compares with the algorithm in (Favoreel *et al.*, 1996), (Favoreel *et al.*, 1999).

Example 1. The system matrices are

$$A = \begin{pmatrix} 0 & 0.5 \\ -0.5 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, C = (1 \ 1),$$

$$D = 2, \quad N_1 = \begin{pmatrix} 0.4 \\ 0 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 0 \\ 0.3 \end{pmatrix}$$

and the noise covariance matrices

$$Q = \begin{pmatrix} 0.16 & 0 \\ 0 & 0.04 \end{pmatrix}, R = 0.09, S = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

In (Favoreel *et al.*, 1996), the input is a white noise series and $k = 3, j = 8191$. In both Table 1 and Table 2, the system input of case I has a uniform distribution with mean value zero and variance 1 and $\lambda = 0.7809$. In case II, we adjust the system noise as follows:

$$Q = \begin{pmatrix} 0.0016 & 0 \\ 0 & 0.0004 \end{pmatrix}, R = 0.0009, S = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

to increase the noise ratio between the input signal and system noise. The estimation results are also given in Table 1. It is shown that the greater the ratio, the better the convergence. For cases III and IV, we introduce a colored noise input signal z , with mean 0, standard deviation 1.1664, $\lambda = 0.7906$ and $r_q = Ez_k z_{k+q} = 0.5^q, \quad q = 0, 1, 2, \dots$ and the system noise is taken the same as in cases I and II respectively. The simulation results are also shown in Table 1. The values $k = 2, j = 595$ were used in all our simulations. These compare favourably with the values $k = 3, j = 8191$ used by (Favoreel *et al.*, 1996) for the bilinear N4SID algorithm.

Table 1. Example 1: Results with different inputs, noise ratios and algorithms

	eig(A)	eig(N)
True	$\pm 0.5i$	0.4, 0.3
N4SID	$-0.0027 \pm 0.4975i$	0.4011, 0.3055
Case I	$-0.0078 \pm 0.4864i$	0.4128, 0.3035
Case II	$0.0000 \pm 0.5000i$	0.4005, 0.3030
Case III	$0.0089 \pm 0.4945i$	0.3906, 0.3149
Case IV	$0.0005 \pm 0.4980i$	0.4006, 0.2976

The simulation results with different sample sizes for cases I and II are given in Table 2, to show the relationship between the accuracy and the sample size.

Example 2. The system matrices and the noise covariance matrices are

Table 2. Example 1: Effect of sample size

	eig(A)	eig(N)
True	$\pm 0.5i$	0.4, 0.3
j=95 (I)	$-0.0027 \pm 0.5018i$	0.3729, 0.3473
j=95 (II)	$0.0059 \pm 0.4982i$	0.4026, 0.3095
j=295 (I)	$0.0074 \pm 0.4966i$	0.4298, 0.3022
j=295 (II)	$0.0031 \pm 0.5008i$	0.4012, 0.2988
j=595 (I)	$0.0078 \pm 0.4864i$	0.4128, 0.3035
j=595 (II)	$0.0000 \pm 0.5000i$	0.4005, 0.3030

$$A = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.3 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N_1 = \begin{pmatrix} 0.6 & 0 \\ 0 & 0.4 \end{pmatrix}, N_2 = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.5 \end{pmatrix}$$

$$Q = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}, R = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}, S = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The input we selected here has a two-dimensional uniform distribution. Table 3 shows the true and estimated eigenvalues of A and N , as found by (Favoreel *et al.*, 1999) using the bilinear N4SID algorithms with $k = 2, j = 4095$, and using our new algorithm with $k = 2, j = 597, \lambda = 0.7803$, and solving equation (10) using both ordinary least-squares (Model I) and constrained least-squares (Model II).

Table 3. Example 2: Comparisons with different algorithms.

	eig(A)	eig(N ₁)	eig(N ₂)
True	0.5, 0.3	0.6, 0.4	0.2, 0.5
N4SID	0.5001	0.5994	0.5016
(j=4095)	0.2979	0.4020	0.1914
Model I	0.4998	0.5998	0.5000
(j=597)	0.3002	0.4000	0.2001
Model II	0.5004	0.5998	0.4999
(j=597)	0.2990	0.3997	0.1997

It is seen that the eigenvalues of the system matrices of the model and the original system are very close to each other. From Table 1 and Table 3, it is seen that the algorithm presented here has a quicker convergence rate and smaller sample size requirement compared to the bilinear N4SID algorithm. As stated in (Favoreel *et al.*, 1996), (Favoreel *et al.*, 1999), the noise covariance estimates are not very accurate since the k -parameter is too small ($k = 2$).

Remark 4 Our new algorithm has considerably lower computational complexity than the one given in (Favoreel and De Moor, 1998). The major computational load is involved in finding the right-inverse in (9). The matrix involved here has dimensions $f_k \times j$, where $f_k = e_k + (m/2)(m+1)^k + l[(m+1)^k - 1] + e_k^2$. For Example 2, with $k = 2$, we

have $f_k = 136$. In (Favoreel and De Moor, 1998), equation (9), the matrix whose right-inverse has to be found has dimensions $(d_k + 2e_k + e_k d_k + e_k^2) \times j$. The row dimension for both algorithms increases exponentially with k . For Example 2 with $k = 2$ this row dimension is 152. Furthermore, since our algorithm seems to require much smaller values of j for comparable performance, the column dimension is also much smaller for our algorithm in practice.

6. CONCLUSION

A new subspace identification algorithm has been presented for the identification of a general combined deterministic-stochastic bilinear system. The main result in this paper is that the system input does not have to be white noise, which is the main restriction and assumption in the current bilinear subspace system identification algorithms. Simulation results and comparisons with subspace identification algorithm have been given.

7. ACKNOWLEDGEMENT

The work reported in this paper was supported by the UK Engineering and Physical Science Research Council under the GrantGR/MO8332, and by the European Research Network on System Identification (ERNSI) under TMR contract ERB FMRX CT98 0206.

8. REFERENCES

- Chen, H. and J.M. Maciejowski (1999). Subspace identification of deterministic bilinear systems. Accepted for ACC 2000, Chicago, Illinois, USA.
- Chen, H. and J.M. Maciejowski (2000). New subspace identification method for bilinear systems. Technical Report CUED/F-INFENG/TR.357. Department of Engineering, University of Cambridge.
- Chui, N.L.C. and J.M. Maciejowski (1998). Subspace identification – a Markov parameter approach. Technical Report CUED/F-INFENG/TR. 337. Department of Engineering, University of Cambridge.
- Favoreel, W. and B. De Moor (1998). Subspace identification of bilinear systems. In: *Proc. MTNS*. pp. 787–790.
- Favoreel, W., B. De Moor and P. Van Overschee (1996). Subspace identification of bilinear systems subject to white inputs. Technical report. Katholieke Universiteit Leuven.
- Favoreel, W., B. De Moor and P. Van Overschee (1997). Subspace identification of balanced deterministic bilinear systems subject to white inputs. In: *ECC*.
- Favoreel, W., B. De Moor and P. Van Overschee (1999). Subspace identification of bilinear systems subject to white inputs. *IEEE Trans on AC* **44**(6), 1157–1165.
- Khatri, C.G. and C.R. Rao (1968). Solutions to some functional equations and their applications to characterization of probability distributions. *The Indian J. Stat. series A*. **30**, 167–180.
- Suda, N., S. Kodama and M. Ikeda (1973). *Matrix Theory in Automatical Control*. Japanese Automatical Control Association.