IEEE1086

Subspace Identification of Deterministic Bilinear Systems

Huixin Chen and **Jan Maciejowski** *

Department of Engineering
University of Cambridge
Cambridge CB2 1PZ U.K.
Tel. +44 1223 332732
Fax. +44 1223 332662
Email: hc240@eng.cam.ac.uk, jmm@eng.cam.ac.uk

Keywords:

System identification, Bilinear systems, Subspace methods, Coloured inputs.

Abstract

In this paper, a subspace method for the identification of deterministic bilinear systems is developed. The input signal to the system does not have to be white, which is a major advantage over an existing subspace method for bilinear systems. Simulation results also show that the new algorithm converges much more rapidly (with sample size) than the existing method, and hence is more effective with small sample sizes. The faster convergence is presumably due to the insensitivity of the algorithm to the sample-spectrum of the input signal. These advantages are achieved by a different arrangement of the input-output equations into 'blocks', and projections onto different spaces than the ones used in the existing method. A further advantage of our algorithm is that the dimensions of the matrices involved are significantly smaller, so that the computational complexity is lower.

1 Introduction

Bilinear systems are attractive models for many dynamical processes, because they allow a significantly larger class of behaviours than linear systems, yet retain a rich theory which is closely related to the familiar theory of linear systems [11]. They exhibit phenomena encountered in many engineering systems, such as amplitude-dependent time constants. Many practical system models are bilinear, and more general nonlinear systems can often be well approximated by bilinear models [13].

Most studies of the identification problem of bilinear systems have assumed an input-output formulation. Standard methods such as recursive least squares, extended least squares, recursive auxiliary variable and recursive prediction error algorithms, have been applied to identifying bilinear systems. Simulation studies have been undertaken [10], and some statistical results (strong consistency and parameter estimate convergence rates) are also available [3].

In this paper, we consider the identification of (multivariable) bilinear systems in state-space form. There are many advantages of using state-space models, particularly in the multivariable case [4]. In recent years 'subspace' methods have been developed which have proved to be extremely effective for the identification of linear systems [5, 14, 16, 17]. In [6, 7, 9] extension of such methods were given for bilinear systems, but the algorithm presented there was effective only if the measured input signal to the system being identified is white. To our knowledge this was the first extension of the subspace approach to bilinear systems. In [8] another subspace algorithm for bilinear systems was presented by the same authors, which apparently does not require a white input signal. However this algorithm is known to give biased results, and it must therefore be questioned whether it can really be considered to be an effective algorithm for the case of non-white inputs.

In this paper an alternative subspace algorithm for identifying bilinear systems is proposed. It does not require the measured input to be white, and the matrices which need to be constructed and operated upon are much smaller than those which appear in [9]. Simulations show that it works well when the input signal is not white; they also show that if the input signal is white, then good results are obtained with much smaller sample sizes than are required for the algorithm of [9].

This paper deals only with the deterministic case,

*Corresponding author

in which all the inputs to the system are measured, and there are no stochastic disturbances or other unmeasured inputs. In the companion paper [2] we deal with the combined deterministic-stochastic case (which is the case treated in [9]). All proofs are omitted here, but can be found in [1].

2 Notation

The use of much specialised notation seems to be unavoidable in the current context. Mostly we follow the notation used in [9], 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 [12, 15]:

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

$+, \oplus$ 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 $\cdot^{*}$.

We consider deterministic time-invariant bilinear systems of the form:

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

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

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

For arbitrary t we define

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

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

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

where k is the *row block size*. The suffices p and f are supposed to be mnemonic, representing 'past' and 'future' respectively.

We define U_t, U_p, U_f, Y_t, Y_p, and Y_f similarly:

$$\begin{aligned} U_t &\triangleq [u_t\ u_{t+1}\ ...\ u_{t+j-1}] \in \mathbf{R}^{m\times j} \\ U_p &\triangleq [u_0\ u_1\ ...\ u_{j-1}] \in \mathbf{R}^{m\times j} \\ U_f &\triangleq [u_k\ u_{k+1}\ ...\ u_{k+j-1}] \in \mathbf{R}^{m\times j} \\ Y_t &\triangleq [y_t\ y_{t+1}\ ...\ y_{t+j-1}] \in \mathbf{R}^{l\times j} \\ Y_p &\triangleq [y_0\ y_1\ ...\ y_{j-1}] \in \mathbf{R}^{l\times j} \\ Y_f &\triangleq [y_k\ y_{k+1}\ ...\ y_{k+j-1}] \in \mathbf{R}^{l\times j} \end{aligned}$$

These matrices will later be used to construct larger matrices with a 'generalised block-Hankel' structure such as (2). In order to use all the available data in these, the number of columns j is such that $\tilde{N} = 2k + j - 1$.

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

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

Remark 1 The meaning of $U^{+}_{i-1|q}$ is different from that in [6]. $U^{++}_{i-1|q}$ is newly introduced here.

$$\begin{aligned}
U^{u}_{i-1|q} &\triangleq U^{++}_{i-1|q} \odot U_q \\
U^{y}_{i-1|q} &\triangleq U^{+}_{i-1|q} \odot Y_q \\
U^{u,y}_{i-1|q} &\triangleq \begin{pmatrix} U_{i-1|q} \\ U^{u}_{i-1|q} \\ U^{y}_{i-1|q} \end{pmatrix} \\
X^p &\triangleq X_{k-1|0} \in \mathbf{R}^{(m+1)^k n\times j} \\
X^f &\triangleq X_{2k-1|k} \in \mathbf{R}^{(m+1)^k n\times j} \\
U^p &\triangleq U_{k-1|0} \in \mathbf{R}^{e_k\times j} \\
U^f &\triangleq U_{2k-1|k} \in \mathbf{R}^{e_k\times j} \\
Y^p &\triangleq Y_{k-1|0} \in \mathbf{R}^{d_k\times j} \\
Y^f &\triangleq Y_{2k-1|k} \in \mathbf{R}^{d_k\times j} \\
U^{++p} &\triangleq U^{++}_{k-1|0} \in \mathbf{R}^{\frac{m}{2}(m+1)^k\times j} \\
U^{++f} &\triangleq U^{+}_{2k-1|k} \in \mathbf{R}^{\frac{m}{2}(m+1)^k\times j}
\end{aligned}$$

$$
\begin{aligned}
U^{p,y} &\triangleq U^{+p} \odot Y_p \in \mathbf{R}^{l[(m+1)^k-1]\times j} \\
U^{f,y} &\triangleq U^{+f} \odot Y_f \in \mathbf{R}^{l[(m+1)^k-1]\times j} \\
U^{p,u,y} &\triangleq \begin{pmatrix} U^p \\ U^{++p} \\ U^{p,y} \end{pmatrix} \in \mathbf{R}^{f_k \times j} \\
U^{f,u,y} &\triangleq \begin{pmatrix} U^f \\ U^{++f} \\ U^{f,y} \end{pmatrix} \in \mathbf{R}^{f_k \times j}
\end{aligned}
$$

where $d_i = \Sigma_{p=1}^{i}(m+1)^{p-1}l$, $e_i = \Sigma_{p=1}^{i}(m+1)^{p-1}m$ and $f_k = e_k + \frac{m}{2}(m+1)^k + l[(m+1)^k - 1]$. Finally, we denote by $\mathcal{U}_p$ the space spanned by all the rows of the matrix U_p. That is,

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

$\mathcal{U}_f$, $\mathcal{Y}_p$, $\mathcal{Y}_f$, $\mathcal{U}^p$, $\mathcal{Y}^p$, $\mathcal{U}^f$, $\mathcal{Y}^f$, $\mathcal{U}^{p,u,y}$, $\mathcal{U}^{f,u,y}$ etc can be defined similarly.

3 Analysis

Lemma 1

$$X^p = \begin{pmatrix} X_p \\ U^{+p} \odot X_p \end{pmatrix}$$

$$X^f = \begin{pmatrix} X_f \\ U^{+f} \odot X_f \end{pmatrix}$$

Lemma 2 *¿From (1) we have, modulo a state-coordinate transformation,*

$$X_p = C^\dagger(Y_p - DU_p)$$
$$X_f = C^\dagger(Y_f - DU_f)$$

Remark 2 This holds for any right inverse of C. Different choices of right inverse correspond to different choices of state coordinates. Note that the spaces $\mathcal{X}_p$ and $\mathcal{X}_f$ do not depend on this choice. We will assume that the Moore-Penrose pseudo-inverse of C is used.

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

$$
\begin{aligned}
(FG \otimes HJ) &= (F \otimes H)(G \otimes J) \\
(FG \odot HJ) &= (F \otimes H)(G \odot J)
\end{aligned}
$$

Lemma 4 (Input-Output Equation)

$$
\begin{aligned}
Y^p_{k-1|0} &= \overline{\Gamma}_k X^p_{k-1|0} + H_k U^p_{k-1|0} &\quad (2) \\
X_f &= \overline{A}_k X^p_{k-1|0} + \triangle^d_k U^p_{k-1|0} &\quad (3)
\end{aligned}
$$

where

$$
\begin{aligned}
\overline{\Gamma}_i &\triangleq \begin{pmatrix} C\overline{A}_{i-1} & 0 & \dots & 0 \\ \overline{\Gamma}_{i-1} & 0 & \dots & 0 \\ 0 & \overline{\Gamma}_{i-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \overline{\Gamma}_{i-1} \end{pmatrix} \in \mathbf{R}^{d_i \times (m+1)^i n} \\
\overline{\Gamma}_1 &\triangleq (C \quad 0_{l\times(m+1)n}) \\
H_i &\triangleq \begin{pmatrix} D & C\triangle^d_{i-1} & 0 & \dots & 0 \\ 0 & H_{i-1} & 0 & \dots & 0 \\ 0 & 0 & H_{i-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & H_{i-1} \end{pmatrix} \in \mathbf{R}^{d_i \times e_i} \\
H_1 &\triangleq D,
\end{aligned}
$$

where

$$
\begin{aligned}
\triangle^d_n &\triangleq \begin{bmatrix} B & A\triangle^d_{n-1} & N_1\triangle^d_{n-1} \dots, & N_m\triangle^d_{n-1} \end{bmatrix} \\
\triangle^d_1 &\triangleq B \\
\overline{A}_i &\triangleq (A\overline{A}_{i-1} \quad N_1\overline{A}_{i-1} \quad \dots \quad N_m\overline{A}_{i-1}), \\
\overline{A}_0 &\triangleq I_{n\times n}
\end{aligned}
$$

Theorem 1 (Input-Output Equation in Two Block Form) *The system (1) can be written in the following 'two block' form:*

$$Y^p = \mathcal{C}_k X_p + \mathcal{D}_k U^{p,u,y} \quad (4)$$
$$Y^f = \mathcal{C}_k X_f + \mathcal{D}_k U^{f,u,y} \quad (5)$$

$$X_f = \mathcal{A}_k X_p + \mathcal{B}_k U^{p,u,y} \quad (6)$$

Theorem 2 *If the following condition is satisfied*

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

is a full row rank matrix, then

$$
\begin{aligned}
\mathcal{X}_f &\subset \mathcal{Y}^p + \mathcal{U}^{p,u,y} \\
\mathcal{X}_f + \mathcal{U}^{f,u,y} &= \mathcal{X}_f \oplus \mathcal{U}^{f,u,y} \\
&\subset (\mathcal{Y}^p + \mathcal{U}^{p,u,y}) \oplus \mathcal{U}^{f,u,y} &\quad (8)
\end{aligned}
$$

4 Algorithm

1. Decompose Y^f into $\mathcal{C}_k X_f$ and $\mathcal{D}_k U^{f,y,u}$ using orthogonal projection: from (5) and (8) it follows that

$$\Pi_{\Omega^\perp} Y^f = \mathcal{D}_k \Pi_{\Omega^\perp} U^{f,u,y} \quad (9)$$

where $\Omega = \mathcal{Y}^p + \mathcal{U}^{p,u,y}$. Determine $\mathcal{D}_k \in \mathbf{R}^{d_k \times (e_k + \frac{m}{2}(m+1)^k + l[(m+1)^k - 1])}$ from

$$\mathcal{D}_k = \left(\Pi_{\Omega^\perp} Y^f\right)\left(\Pi_{\Omega^\perp} U^{f,u,y}\right)^\dagger \tag{10}$$

2. Obtain the SVD decompositon and partition accordingly by selecting an model order

$$\begin{aligned}&\left[Y_{2k-2|k-1} \quad Y_{2k-1|k}\right] - \mathcal{D}_k \left[U^{u,y}_{2k-2|k-1} \quad U^{u,y}_{2k-1|k}\right] \\ &=: \Gamma\Sigma\Omega^* = \left[\begin{array}{cc} \Gamma_1 & \Gamma_2 \end{array}\right] \left[\begin{array}{cc} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{array}\right] \left[\begin{array}{c} \Omega_1^* \\ \Omega_2^* \end{array}\right]\end{aligned} \tag{11}$$

Since we expect

$$\Gamma\Sigma\Omega^* = \Gamma_1\Sigma_1\Omega_1^* = \mathcal{C}_k\left[X_{k-1} \quad X_k\right] \tag{12}$$

from (4) and (5) ($rank(\Sigma_1) = n$ and $rank(\Sigma_2) = 0$), form the estimates $\hat{\mathcal{C}}_k = \Gamma_1\Sigma_1^{1/2}$ and $[\hat{X}_{k-1} \quad \hat{X}_k] = \Sigma_1^{1/2}\Omega_1^*$, retaining only significant singular values in Σ_1. ($\hat{\mathcal{C}}_k$ is not needed later.)

3. Estimate the parameters A, B, C, D, N by solving

$$\left[\begin{array}{c} \hat{X}_k \\ Y_{k-1} \end{array}\right] = \left[\begin{array}{ccc} A & N & B \\ C & 0 & D \end{array}\right] \left[\begin{array}{c} \hat{X}_{k-1} \\ U_{k-1} \odot \hat{X}_{k-1} \\ U_{k-1} \end{array}\right] \tag{13}$$

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 [9] 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 have not much effect the eigenvalues of matrices A and N in [2].

5 Examples

In this section, two second-order bilinear systems introduced in [6, 9] are used to see how the new algorithm works, and how it compares with existing algorithms.

Example 1 The system matrices are

$$A = \left(\begin{array}{cc} 0 & 0.5 \\ -0.5 & 0 \end{array}\right),\ B = \left(\begin{array}{c} 1 \\ 1 \end{array}\right),\ C = \left(\begin{array}{cc} 1 & 1 \end{array}\right),$$

$$D = 2, \quad N_1 = \left(\begin{array}{c} 0.4 \\ 0 \end{array}\right) \quad N_2 = \left(\begin{array}{c} 0 \\ 0.3 \end{array}\right)$$

Table 1 shows the eigenvalues of the estimated A and N in various cases. The row labelled 'N4SID' gives the results obtained in [6], with a white input and $k = 3$, $j = 8191$. 'Case I' is for a white input, with uniform distribution in the interval [0,0.01] and $k = 2$. 'Case II' is for a white input with normal distribution $N(0, 0.1^2)$ and $k = 6$. 'Case III' is for a coloured input z with mean 0, standard deviation 3.3e-05 and $r_q = Ez_kz_{k+q} = 0.5^q$, $q = 0, 1, 2, ...$ and $k = 2$. 'Case IV' is for a white input with exponential distribution with parameter $\lambda = 0.04$ and $k = 2$. In all the cases I–IV the number of columns is only $j = 597$, compared with $j = 8191$ for the N4SID case. It is seen that the eigenvalues of the true and estimated matrices are very close to each other.

	A	N
True	$\pm 0.5i$	0.4, 0.3
N4SID	$-0.0027 \pm 0.4975i$	0.4011, 0.3055
Case I	$0.0000 \pm 0.5001i$	0.3994, 0.2952
Case II	$0.0000 \pm 0.4993i$	0.4019 0.3064
Case III	$0.0000 \pm 0.5000i$	0.3953 0.3085
Case IV	$0.0006 \pm 0.5011i$	0.3934, 0.3007

Table 1: Example 1: Results with different inputs and algorithms

Table 2 shows how the eigenvalues of the estimated A and N depend on j, in each case with $k = 2$.

	A	N
True	$\pm 0.5i$	0.4, 0.3
j=97	$0.0000 \pm 0.5003i$	0.4092, 0.2789
j=297	$0.0000 \pm 0.5002i$	0.4043, 0.2914
j=597	$0.0000 \pm 0.5001i$	0.3994, 0.2952

Table 2: Example 1: Effect of sample size

Example 2 The system matrices are

$$A = \left(\begin{array}{cc} 0.5 & 0 \\ 0 & 0.3 \end{array}\right), B = \left(\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right), C = \left(\begin{array}{cc} 1 & 0 \\ 0 & 2 \end{array}\right)$$

$$D = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right), N_1 = \left(\begin{array}{cc} 0.6 & 0 \\ 0 & 0.4 \end{array}\right), N_2 = \left(\begin{array}{cc} 0.2 & 0 \\ 0 & 0.5 \end{array}\right)$$

Table 3 shows the results obtained in [9] with $j = 4095$ and $k = 2$, but *with stochastic inputs*, and the results obtained by our new algorithm with $j = 597$ and $k = 2$ in the deterministic case. A fairer comparison is available in [1, 2]. In both cases the input signal was white, with a uniform two-dimensional distribution.

	True	N4SID	New
A	0.5, 0.3	0.5001, 0.2979	0.5000, 0.3000
N_1	0.6, 0.4	0.5994, 0.4020	0.6000, 0.4000
N_1	0.2, 0.5	0.1914, 0.5016	0.2000, 0.5000

Table 3: Example 2: Results with different inputs and algorithms

Remark 4 Our new algorithm has considerably lower computational complexity than the one given in [8]. The major computational load is involved in finding the right-inverse in (10). The matrix involved here has dimensions $f_k \times j$, where $f_k = e_k + (m/2)(m+1)^k + l[(m+1)^k - 1]$. For Example 2, with $k = 2$, we have $f_k = 33$. In [8], 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 increases exponentially quickly 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 algorithm for the identification of deterministic bilinear systems has been developed. Its main advantage is that the system input does not have to be white. It also has lower computational complexity than the previously proposed algorithm, because the dimensions of the matrices involved in it are much smaller. Its wider applicability has been demonstrated by two examples, which also show that even when coloured inputs are available the new algorithm converges to correct estimates relatively quickly. The presumed reason for this is that, since the algorithm does not depend on whiteness of the input, it is insensitive to the large errors in the sample spectrum which are inevitable with small sample sizes.

7 Acknowledgement

The work reported in this paper was supported by the UK Engineering and Physical Science Research Council under Research Grant GR/M08332, and by the European Research Network on System Identification (ERNSI) under TMR contract ERB FMRX CT98 0206.

References

[1] Chen H. and Maciejowski J.M., New Subspace Identification Method for Bilinear Systems, *Technical Report CUED/F-INFENG/TR.357 Cambridge University Engineering Department*, 1999.

[2] Chen H. and Maciejowski J.M., Subspace identification of combined deterministic-stochastic bilinear systems, *Submitted to IFAC Symp. on System Identification, SYSID 2000, Santa Barbara*, June 2000.

[3] Chen H., Zinober A.S.I. and Ruan R., Strong Consistency and Convergence Rate of Parameter Identification for Bilinear Systems, *Int.J.Control*, 1996, vol. **63**, No.5 pp. 907-919

[4] Chou C.T. and Maciejowski J.M., System Identification Using Balanced Parameterizations, *IEEE Trans. Auto. Contr.*, Vol. **42**, No. 7. July, 1997 pp.956-974.

[5] Chui N.L.C. and Maciejowski J.M., Subspace Identification – a Markov parameter approach, Submitted to *IEEE Trans. Auto. Contr.*. Also *Technical ReportCUED/F-INFENG/TR.337, Cambridge University Engineering Department*, December, 1998.

[6] Favoreel W, De Moor B. and Van Overschee P., Subspace identification of bilinear systems subject to white inputs, *Technical Report ESAT-SISTA/TR 1996-53I, Katholieke Universiteit Leuven*, 1996.

[7] Favoreel W, De Moor B. and Van Overschee P., Subspace identification of Balanced deterministic bilinear systems subject to white inputs, *Proc. European Ccontrol Conf, Brussels*, 1997.

[8] Favoreel W and De Moor B, Subspace identification of bilinear systems, *Proc. MTNS, Padova*, 1998. pp. 787-790.

[9] Favoreel W, De Moor B. and Van Overschee P., Subspace identification of bilinear systems subject to white inputs, *IEEE Trans. Auto. Contr.*, Vol. **44**, No. 6, June 1999. pp. 1157–1165.

[10] Fnaiech F. and Ljung L., Recursive Identification of Bilinear System, *Int.J. Control*, vol. **45**, pp. 453-470., 1987.

[11] Isidori, A, *Nonlinear Control Systems (3rd ed)*, Berlin: Springer-Verlag, 1995.

[12] Khatri C.G. and Rao C.R., Solutions to some functional equations and their applications to characterization of probability distributions, *Sankhya : The Indian J. Stat.*, series A, **30**, pp. 167-180, 1968.

[13] Krener A.J., Bilinear and nonlinear realizations of input-output maps. *SIAM Journal on Control*, vol. **13**, pp.827-834., 1975.

[14] Larimore,W.E, System identification, reduced-order filtering and modeling via canonical variate analysis, *Proc. American Contr.Conf., San Francisco*, (1983).

[15] Suda N., Kodama S. and Ikeda M., *Matrix Theory in Automatical Control,* Japanese Automatical Control Association, 1973.

[16] Van Overschee P. and De Moor B., N4SID: subspace algorithms of combined deterministic and stochastic systems, *Automatica*, vol. **30**, No. 1, pp. 75-93, 1994

[17] Verhaegen,M, and Dewilde,P.M, Subspace model identification, Parts 1 and 2, *Int. J. Contr.*, **56**, 1187-1241, (1992).