Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

A Subspace Approach for Identifying Bilinear Systems with Deterministic Inputs

J. A. Ramos
Department of Electrical & Computer Engineering
Purdue University at Indianapolis
723 West Michigan Street, SL 164-E
Indianapolis, IN 46202-5132

and

P. Lopes dos Santos
Faculdade de Engenharia da Universidade do Porto
Departamento de Engenharia Electrotécnica e de Computadores
Rua Dr. Roberto Frias
4200 465 Porto, Portugal

Abstract— **In this paper we introduce an identification algorithm for MIMO bilinear systems subject to deterministic inputs. The new algorithm is based on an expanding dimensions concept, leading to a rectangular, dimension varying, linear system. In this framework the observability, controllability, and Markov parameters are similar to those of a time-varying system. The fact that the system is time invariant, leads to an equaivaleet linear deterministic subspace algorithm. Provided a rank condition is satisfied, the algorithm will produce unbiased parameter estimates. This rank condition can be guaranteed to hold if the ratio of the number of outputs to the number of inputs is larger than the system order. This is due to the typical exponential blow-out in the dimensions of the Hankel data matrices of bilinear systems, in particular for deterministic inputs since part of the input subspace cannot be projected out. Other algorithms in the literature, based on Walsh functions, require that the number of outputs is at least equal to the system order. For ease of notation and clarification, the algorithm is presented as an intersection based subspace algorithm. Numerical results show that the algorithm reproduces the system parameters very well, provided the rank condition is satisfied. When the rank condition is not satisfied, the algorithm will return biased parameter estimates, which is a typical bottleneck of bilinear system identification algorithms for deterministic inputs.**

I. Introduction

The identification of subspace models in the state space domain has received considerable attention during the past decade due to the plethora of algorithms emerging from the subspace identification community, most notably, MOESP, N4SID, and CVA, to name only a few [29], [35], [16]. Centering around these three main algorithms, there has been plenty of work documenting extensions of these, and the interested reader is referred to [18], [5], [6], [8], [24] for more details. These algorithms have been quite successful with linear models, thus being capable of handling a variety of input/output conditions such as deterministic inputs with no noise in the system [20], [30], deterministic inputs with noise in the system (referred to as the deterministic-stochastic case) [29], [35], [16], and the no input, purely stochastic case (referred to as the stochastic realization problem) [33], [7], [1]. For nonlinear systems, the class of subspace algorithms is far more restrictive and, as expected, there are fewer algorithms than those for linear systems. Nevertheless, extensions of MOESP, N4SID, and CVA for nonlinear systems have appeared scattered in the literature. Most algorithms in this category are for Hammerstein systems [31], Wiener systems [32], bilinear systems [26], [21], [25], [27], [11], [12], [3], [4], [22], and for linear parameter varying systems [28]. Concerning bilinear systems, one main concern with bilinear subspace algorithms is that the Hankel data matrices grow exponentially. That is, suppose $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^\ell$ denote, respectively, the input and output at time k. Then the past and future Hankel data matrices have dimensions $Y_\bullet \in \mathbb{R}^{\frac{\ell}{m}[(m+1)^i-1]\times j}$ and $U_\bullet \in \mathbb{R}^{[(m+1)^i-1]\times j}$, respectively, where i is the number of block rows in the Hankel data matrices and j is the number of columns. When the system contains general inputs (i.e., non white noise case), the rank condition needed for determining the observable subspace is violated, thus leading to biased parameter estimates. Some authors have attempted to overcome this limitation, the results of which can be found in [27], [4]. The case where the inputs are a white noise process has been well studied and presented in [11], [12], [26], [21], [25], [3], [4]. The reason why white noise inputs are almost a necessity is because the observable subspace can be obtained similarly to the linear case by performing certain projections. This has led to bilinear identification algorithms quite similar to the linear identification algorithms [29], [35], [16]. When the noise is not white, then the projections still contain information that would otherwise be nulled, thus increasing the rank of the observability matrix. A recent bilinear subspace algorithm was presented by the authors for the case of white noise inputs, where the identification is done by successive linear subspace identification [10]. As such, the dimensions are equivalent to those of a linear subspace algorithm, at the expense of performing an iterative, but convergent, procedure.

It is well known in system theory that a bilinear system can approximate a nonlinear system fairly well, via a finite sum of the Volterra series expansion between the inputs and outputs of the system [2], [14], [9], [13]. However, for more general nonlinear systems, a finite sum of Volterra kernels may no longer hold. Thus, bilinear subspace algorithms can only solve a limited class of nonlinear system identification problems. Despite this limitation, they still provide a higher degree of approximation to nonlinear models than traditional linear models. Furthermore, from a system theory point of view, bilinear models behave similarly to linear models, which is not the case with more general nonlinear models [13]. Finally, a full understanding of the bilinear model is important before attempting to model more general nonlinear systems such as linear parameter varying models [28], which is a direct extension of the bilinear case.

In this paper we present a bilinear subspace identification algorithm for the purely deterministic case. The motivation for the paper follows from the works of [9], [14], [2], [12], [21]. Here we take a slightly different approach by considering the Kronecker product $x_k \otimes u_k$ instead of the usual $u_k \otimes x_k$. This allows us to model the bilinear system as a rectangular linear system with varying dimensions. In this framework, the system evolves like a time-varying linear system of much higher dimensions. Nevertheless, system properties such as the observability and controllability matrices, as well as the Toeplitz matrix of Markov parameters, show up identically to the linear case. In essence we are considering a bilinear system to be a linear system in a much higher dimensional space. The rest of the paper is outlined as follows: In section 2 we introduce our linear framework for deterministic bilinear systems. In section 3 we present the deterministic bilinear algorithm as an extension of the work of [20] for bilinear systems. In section 4 we present some numerical results, and finally, in section 5, we draw some conclusions of our work.

II. Modeling the Bilinear System as a Rectangular Linear System

Consider the general bilinear system with deterministic inputs, given by

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Ex_k \otimes u_k & (1)\\ y_k &= Cx_k + Du_k + Fx_k \otimes u_k, & (2) \end{aligned}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, and $y_k \in \mathbb{R}^\ell$ are, respectively, the state, input, and output vectors at time k. The unknown parameter matrices have dimensions $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{\ell\times n}$, $D \in \mathbb{R}^{\ell\times m}$, $E = \left[\begin{array}{ccc} E_1 & \cdots & E_n \end{array}\right]$, and $F = \left[\begin{array}{ccc} F_1 & \cdots & F_n \end{array}\right]$, where $E_i \in \mathbb{R}^{n\times m}$ and $F_k \in \mathbb{R}^{\ell\times m}$, for $k = 1, \ldots, n$. Let the input, output, and state vectors be initially denoted by

$$\begin{aligned} u_k^0 &= u_k \in \mathbb{R}^m\\ y_k^0 &= y_k \in \mathbb{R}^\ell\\ x_k^0 &= x_k \in \mathbb{R}^n, \end{aligned}$$

and consider augmenting these according to the following dimension expansion rule

$$\begin{aligned} u_k^i &= \left[\begin{array}{c} u_k^{i-1} \\ u_k^{i-1} \otimes u_{k+i} \end{array}\right] \in \mathbb{R}^{m(m+1)^i}\\ y_k^i &= \left[\begin{array}{c} y_k^{i-1} \\ y_k^{i-1} \otimes u_{k+i} \end{array}\right] \in \mathbb{R}^{\ell(m+1)^i}\\ x_k^i &= \left[\begin{array}{c} x_k^{i-1} \\ x_k^{i-1} \otimes u_{k+i-1} \end{array}\right] \in \mathbb{R}^{n(m+1)^i}. \end{aligned}$$

Let us now re-write $(1)-(2)$ using the above notation, i.e.,

$$\begin{aligned} x_{k+1}^0 &= \left[\begin{array}{c|c} A & E \end{array}\right] \left[\begin{array}{c} x_k \\ x_k \otimes u_k \end{array}\right] + Bu_k^0\\ &= A_0 x_k^1 + B_0 u_k^0\\ y_k^0 &= \left[\begin{array}{c|c} C & F \end{array}\right] \left[\begin{array}{c} x_k \\ x_k \otimes u_k \end{array}\right] + Du_k^0\\ &= C_0 x_k^1 + D_0 u_k^0, \end{aligned}$$

where the initial parameter matrices are given by

$$\begin{aligned} A_0 &= \left[\begin{array}{c|c} A & E \end{array}\right] \in \mathbb{R}^{n\times n(m+1)}\\ B_0 &= B \in \mathbb{R}^{n\times m}\\ C_0 &= \left[\begin{array}{c|c} C & F \end{array}\right] \in \mathbb{R}^{\ell\times n(m+1)}\\ D_0 &= D \in \mathbb{R}^{\ell\times m}. \end{aligned}$$

We now apply the first update using the dimension expansion rule and some properties of Kronecker products [17].

$$x_k^1 = \left[\begin{array}{c} x_k^0 \\ x_k^0 \otimes u_k \end{array}\right],$$

thus leading to

$$\begin{aligned} x_{k+1}^1 &= A_1 x_k^2 + B_1 u_k^1\\ y_k^1 &= C_1 x_k^2 + D_1 u_k^1, \end{aligned}$$

where the parameter matrices now become

$$\begin{aligned} A_1 &= \left[\begin{array}{c|c} A_0 & 0_{n\times nm(m+1)} \\ \hline 0_{nm\times n(m+1)} & A_0 \otimes I_m \end{array}\right]\\ &\in \mathbb{R}^{n(m+1)\times n(m+1)^2}\\ B_1 &= \left[\begin{array}{c|c} B_0 & 0_{n\times m^2} \\ \hline 0_{nm\times m} & B_0 \otimes I_m \end{array}\right]\\ &\in \mathbb{R}^{n(m+1)\times m(m+1)}\\ C_1 &= \left[\begin{array}{c|c} C_0 & 0_{\ell\times nm(m+1)} \\ \hline 0_{\ell m\times n(m+1)} & C_0 \otimes I_m \end{array}\right]\\ &\in \mathbb{R}^{\ell(m+1)\times n(m+1)^2}\\ D_1 &= \left[\begin{array}{c|c} D_0 & 0_{\ell\times m^2} \\ \hline 0_{\ell m\times m} & D_0 \otimes I_m \end{array}\right]\\ &\in \mathbb{R}^{\ell(m+1)\times m(m+1)}. \end{aligned}$$

If we continue expanding the system, we obtain a linear, rectangular, state space system with varying dimensions, i.e.,

$$x_{k+1}^{i-1} = A_{i-1}x_k^i + B_{i-1}u_k^{i-1} \quad (3)$$
$$y_k^{i-1} = C_{i-1}x_k^i + D_{i-1}u_k^{i-1}, \quad (4)$$

where, in general, the parameter matrices are recursively computed from

$$A_i = \left[\begin{array}{c|c} A_{i-1} & 0_{n\cdot m_{i-1}\times n\cdot m_i} \\ \hline 0_{n\cdot m_{i-1}\times n\cdot m_i} & A_{i-1}\otimes I_m \end{array}\right]$$
$$B_i = \left[\begin{array}{c|c} B_{i-1} & 0_{n_{i-1}\times m\cdot m_{i-1}} \\ \hline 0_{n\cdot m_{i-1}\times m_{i-1}} & B_{i-1}\otimes I_m \end{array}\right]$$
$$C_i = \left[\begin{array}{c|c} C_{i-1} & 0_{\ell_{i-1}\times n\cdot m_{i-1}} \\ \hline 0_{\ell\cdot m_{i-1}\times n\cdot m_i} & C_{i-1}\otimes I_m \end{array}\right]$$
$$D_i = \left[\begin{array}{c|c} D_{i-1} & 0_{\ell_{i-1}\times m\cdot m_{i-1}} \\ \hline 0_{\ell\cdot m_{i-1}\times m_{i-1}} & D_{i-1}\otimes I_m \end{array}\right],$$

where we have made use of the following compact notation for the dimensions: $n_i = n(m+1)^i$, $m_i = m(m+1)^i$, and $\ell_i = \ell(m+1)^i$. The new parameter matrices now have dimensions

$$A_i \in \mathbb{R}^{n_i\times n_{i+1}}$$
$$B_i \in \mathbb{R}^{n_i\times m_i}$$
$$C_i \in \mathbb{R}^{\ell_i\times n_{i+1}}$$
$$D_i \in \mathbb{R}^{\ell_i\times m_i}.$$

One could say that $(3)-(4)$ is a linear, dimension-varying system that resembles a descriptor system, i.e.,

$$\left[\begin{array}{c|c} I_{n_{i-1}} & 0_{n_{i-1}\times m\cdot n_{i-1}} \\ \hline 0_{m\cdot n_{i-1}\times n_{i-1}} & 0_{m\cdot n_{i-1}\times m\cdot n_{i-1}} \end{array}\right] x_{k+1}^i$$
$$= \left[\begin{array}{c} A_{i-1} \\ \hline 0_{m\cdot n_{i-1}\times n_i} \end{array}\right] x_k^i + \left[\begin{array}{c} B_{i-1} \\ \hline 0_{m\cdot n_{i-1}\times m_{i-1}} \end{array}\right] u_k^{i-1}$$
$$y_k^{i-1} = C_{i-1}x_k^i + D_{i-1}u_k^{i-1}.$$

III. Deterministic Bilinear Subspace Identification Algorithm

Let us now define the Hankel like data matrices as it is customary in linear deterministic systems terminology [20], [30]

$$Y_1 = \begin{bmatrix} y_k^{i-1} & y_{k+1}^{i-1} & \cdots & y_{k+j-1}^{i-1} \\ y_{k+1}^{i-2} & y_{k+2}^{i-2} & \cdots & y_{k+j}^{i-2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{k+i-1}^0 & y_{k+i}^0 & \cdots & y_{k+i+j-2}^0 \end{bmatrix}$$

$$Y_2 = \begin{bmatrix} y_{k+i}^{i-1} & y_{k+i+1}^{i-1} & \cdots & y_{k+i+j-1}^{i-1} \\ y_{k+i+1}^{i-1} & y_{k+i+2}^{i-1} & \cdots & y_{k+i+j}^{i-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{k+2i-1}^0 & y_{k+2i}^0 & \cdots & y_{k+2i+j-2}^0 \end{bmatrix}$$

$$U_1 = \begin{bmatrix} u_k^{i-1} & u_{k+1}^{i-1} & \cdots & u_{k+j-1}^{i-1} \\ u_{k+1}^{i-2} & u_{k+2}^{i-2} & \cdots & u_{k+j}^{i-2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{k+i-1}^0 & u_{k+i}^0 & \cdots & u_{k+i+j-2}^0 \end{bmatrix}$$

$$U_2 = \begin{bmatrix} u_{k+i}^{i-1} & u_{k+i+1}^{i-1} & \cdots & u_{k+i+j-1}^{i-1} \\ u_{k+i+1}^{i-2} & u_{k+i+2}^{i-2} & \cdots & u_{k+i+j}^{i-2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{k+2i-1}^0 & u_{k+2i}^0 & \cdots & u_{k+2i+j-2}^0 \end{bmatrix},$$

where $\{U_1, Y_1\}$ and $\{U_2, Y_2\}$ are, respectively, the past and future input-output Hankel data matrices. It can be shown that these satisfy the following subspace equations, much like in [26], [21], [11], [12], [3], [4]

$$Y_1 = \Gamma X_1^i + H_T U_1 \quad (5)$$
$$Y_2 = \Gamma X_2^i + H_T U_2 \quad (6)$$
$$X_2^0 = \Phi X_1^i + \Delta U_1 \quad (7)$$

where Γ is the observability matrix given by

$$\Gamma = \begin{bmatrix} C_{i-1} \\ C_{i-2}\Phi_{i-1,i-1} \\ C_{i-3}\Phi_{i-2,i-1} \\ \vdots \\ C_0\Phi_{1,i-1} \end{bmatrix},$$

H_T is a lower triangular Toeplitz matrix composed of Markov parameters, i.e.,

$$H_T = \begin{bmatrix} D_{i-1} & & & \\ C_{i-2}B_{i-1} & D_{i-2} & & \\ C_{i-3}\Phi_{i-2,i-2}B_{i-1} & C_{i-3}B_{i-2} & & \\ \vdots & \vdots & \ddots & \\ C_0\Phi_{1,i-2}B_{i-1} & C_0\Phi_{1,i-3}B_{i-2} & \cdots & D_0 \end{bmatrix},$$

the controllability matrix Δ is given by

$$\Delta = \begin{bmatrix} \Phi_{0,i-2}B_{i-1} & \Phi_{0,i-3}B_{i-2} & \cdots & B_0 \end{bmatrix},$$

and finally, the transition matrix is given by

$$\Phi_{i_1,i_2} = A_{i_1}A_{i_1+1}\cdots A_{i_2}$$

Here we use Φ to denote $\Phi_{0,i-1}$.

The past and future state matrices have dimensions

$$X_1^i = \begin{bmatrix} x_k^i \mid x_{k+1}^i \mid \cdots \mid x_{k+j-1}^i \end{bmatrix} \in \mathbb{R}^{n_i\times j}$$
$$X_2^i = \begin{bmatrix} x_{k+i}^i \mid x_{k+i+1}^i \mid \cdots \mid x_{k+i+j-1}^i \end{bmatrix} \in \mathbb{R}^{n_i\times j},$$

whereas the Hankel data matrices have dimensions

$$Y_1 \in \mathbb{R}^{\frac{\ell}{m}\left((m+1)^i-1\right)\times j}$$
$$Y_2 \in \mathbb{R}^{\frac{\ell}{m}\left((m+1)^i-1\right)\times j}$$
$$U_1 \in \mathbb{R}^{\left((m+1)^i-1\right)\times j}$$
$$U_2 \in \mathbb{R}^{\left((m+1)^i-1\right)\times j}.$$

In the remainder of this section we follow the intersection approach in [20] to determine the future state sequence X_2^0.

Assume for the time being that the system is observable. Later on we will see under what conditions this is true. Then from $(5)-(6)$ we can solve for X_1^i and X_2^i as follows

$$\begin{aligned} X_1^i &= \underbrace{\left[\begin{array}{c|c} -\Gamma^\dagger H_T & \Gamma^\dagger \end{array}\right]}_{T} \underbrace{\left[\begin{array}{c} U_1 \\ Y_1 \end{array}\right]}_{H_1} \\ &= TH_1 \\ X_2^i &= \underbrace{\left[\begin{array}{c|c} -\Gamma^\dagger H_T & \Gamma^\dagger \end{array}\right]}_{T} \underbrace{\left[\begin{array}{c} U_2 \\ Y_2 \end{array}\right]}_{H_2} \\ &= TH_2. \end{aligned}$$

Now substituting for X_1^i in (7), we get

$$\begin{aligned} X_2^0 &= \Phi \left[\begin{array}{c|c} -\Gamma^\dagger H_T & \Gamma^\dagger \end{array}\right] \left[\begin{array}{c} U_1 \\ Y_1 \end{array}\right] + \Delta U_1 \\ &= \underbrace{\left[\begin{array}{c|c} -\Phi\Gamma^\dagger H_T + \Delta & \Phi\Gamma^\dagger \end{array}\right]}_{T_2} \left[\begin{array}{c} U_1 \\ Y_1 \end{array}\right] \\ &= T_2 H_1. \end{aligned}$$

Notice that $X_2^i \in \mathbb{R}^{n_i \times j}$ and $X_2^0 \in \mathbb{R}^{n \times j}$ depend on both H_1 and H_2. Furthermore, they both have different dimensions, which means that X_2^0 can be obtained from the intersection of the row spaces of H_1 and H_2, i.e.,

$$\begin{aligned} X_2^i &= \underbrace{\left[\begin{array}{c} T_1 \\ \hline \\ \end{array}\right]}_{T} \left[\begin{array}{c} U_2 \\ Y_2 \end{array}\right] \\ X_2^0 &= T_1 H_2 \\ X_2^0 &\in \text{span}_{\text{row}}\{H_1\} \cap \text{span}_{\text{row}}\{H_2\}, \end{aligned}$$

where

$$\begin{aligned} H_1 &= \left[\begin{array}{c} U_1 \\ Y_1 \end{array}\right] \\ H_2 &= \left[\begin{array}{c} U_2 \\ Y_2 \end{array}\right]. \end{aligned}$$

If we now concatenate the data matrices H_1 and H_2 as

$$H = \left[\begin{array}{c} H_1 \\ H_2 \end{array}\right],$$

we then obtain the following rank condition [20]:

$$\begin{aligned} \text{rank}\{H\} &= \text{rank}\left\{\left[\begin{array}{c} H_1 \\ H_2 \end{array}\right]\right\} \\ &= 2(m+1)^i(n+1) - (n+2). \end{aligned}$$

Thus, X_2^0 can be obtained from the intersection of the row spaces of H_1 and H_2, provided the following rank condition is satisfied

$$\text{rank}\{\Gamma\} = n(m+1)^i,$$

where $\Gamma \sim \frac{\ell}{m}\left[(m+1)^i - 1\right] \times n(m+1)^i$. This rank condition places a restriction on the number of inputs and outputs the system can handle. That is,

$$\begin{aligned} \frac{\ell}{m}\left[(m+1)^i - 1\right] &> n(m+1)^i \\ \frac{\ell}{m}\left[1 - \frac{1}{(m+1)^i}\right] &> n \\ \frac{\ell}{m} &> n, \text{ for } i \text{ large.} \end{aligned}$$

This condition may be very restrictive for non-white noise inputs. However, when the inputs are white noise one can apply a projection algorithm as in [11], [12] to relax this condition. If we were to alleviate this condition for the general input case, then we would obtain biased parameter estimates as was pointed out in [12]. For industrial processes this condition would mean placing extra sensors/actuators in order to estimate the observability subspace that is required for the identification. On the other hand, there are tradeoffs to consider. We could ask the question, should we account for more outputs in order to gain the freedom of using more general inputs? An alternate question could be, is a white noise input an accurate description of an industrial processes? Answers to these questions remain to be found. Finally, one way to increase the dimension of the outputs without adding sensors is to use a B-splines transformation of the output data as in [23].

IV. Deterministic Bilinear Subspace Algorithm

The algorithm can be implemented as a deterministic version of MOESP, CVA, or N4SID [29], [16], [35]. However, for convenience of notation, here we follow the work of [20]. The algorithm can be described as follows:

Step 0: Form the Hankel data matrices Y_1, Y_2, U_1, and U_2. Form the concatenated Hankel matrix

$$H = \left[\begin{array}{c} H_1 \\ H_2 \end{array}\right].$$

Step 1: Compute the singular value decomposition of H

$$\begin{aligned} H &= USV^T \\ &= \left[\begin{array}{cc} U_{11} & U_{12} \\ U_{21} & U_{22} \end{array}\right] \left[\begin{array}{cc} S_{11} & 0_{12} \\ 0_{21} & 0_{22} \end{array}\right] V^T, \end{aligned}$$

where $mi = (m+1)^i - 1$, $\ell i = \frac{\ell}{m}[(m+1)^i - 1]$,

$$\begin{aligned} U_{11} &\in \mathbb{R}^{(mi+\ell i)\times(2mi+n)} \\ U_{12} &\in \mathbb{R}^{(mi+\ell i)\times(2\ell i-n)} \\ U_{21} &\in \mathbb{R}^{(mi+\ell i)\times(2mi+n)} \\ U_{22} &\in \mathbb{R}^{(mi+\ell i)\times(2\ell i-n)} \\ S_{11} &\in \mathbb{R}^{(2mi+n))\times(2mi+n)} \\ V &\in \mathbb{R}^{j\times j} \\ 0_{12} &= 0_{(2mi+n)\times(j-2mi-n)} \\ 0_{21} &= 0_{(2\ell i-n)\times(2mi+n)} \\ 0_{22} &= 0_{(2\ell i-n)\times(j-2mi-n)}, \end{aligned}$$

and

$$\left[\begin{array}{c|c} U_{12}^T & U_{22}^T \end{array}\right] \left[\begin{array}{c} H_1 \\ H_2 \end{array}\right] = 0_{(2\ell i-n)\times j}.$$

Step 2: Compute the compression matrix $U_q \in \mathbb{R}^{(2\ell i-n)\times n}$ by computing the singular value decomposition of

$$U_{12}^T U_{11} S_{11} = \left[\begin{array}{c|c} U_q & U_q^{\perp} \end{array}\right] \left[\begin{array}{cc} S_q & 0_{12} \\ 0_{21} & 0_{22} \end{array}\right] V_q^T,$$

where

$$\begin{aligned} S_q &\in \mathbb{R}^{n\times n} \\ 0_{12} &= 0_{n\times 2mi} \\ 0_{21} &= 0_{2(\ell i-n)\times n} \\ 0_{22} &= 0_{2(\ell i-n)\times 2mi}. \end{aligned}$$

Step 3: Compute the matrix of states X_2^0, defined as

$$X_2^0 = \left[\begin{array}{c|c|c|c} x_{k+i} & x_{k+i+1} & \cdots & x_{k+i+j-1} \end{array}\right],$$

using the compression matrix

$$X_2^0 = U_q^T U_{12}^T H_1.$$

Step 4: Compute $\{A, B, C, D, E, F\}$ by solving the following overdetermined system of equations:

$$\left[\begin{array}{c} \mathcal{X}_2^0 \\ \mathcal{Y}_1^0 \end{array}\right] = \left[\begin{array}{ccc} A & B & E \\ C & D & F \end{array}\right] \left[\begin{array}{c} \mathcal{X}_1^0 \\ \mathcal{U}_1^0 \end{array}\right],$$

where

$$\begin{aligned} \left[\begin{array}{c} \mathcal{X}_2^0 \\ \mathcal{Y}_1^0 \end{array}\right] &= \left[\begin{array}{c|c|c|c} x_{k+1} & x_{k+2} & \cdots & x_{k+i+j-1} \\ \hline y_k & y_{k+1} & \cdots & y_{k+i+j-2} \end{array}\right] \\ \left[\begin{array}{c} \mathcal{X}_1^0 \\ \mathcal{U}_1^0 \end{array}\right] &= \left[\begin{array}{c|c|c|c} x_k & x_{k+1} & \cdots & x_{k+i+j-2} \\ \hline u_k & u_{k+1} & \cdots & u_{k+i+j-2} \end{array}\right]. \end{aligned}$$

V. Numerical Results

Consider the following bilinear system

$$\begin{aligned} x_{k+1} &= A x_k + B u_k + E x_k \otimes u_k \\ y_k &= C x_k + D u_k + F x_k \otimes u_k, \end{aligned}$$

with actual parameter matrices

$$\begin{aligned} A &= \left[\begin{array}{c|c} 0.1000 & 0.0000 \\ \hline 1.0000 & 0.2000 \end{array}\right] \\ B &= \left[\begin{array}{c} 1.0000 \\ \hline 0.0000 \end{array}\right] \\ C &= \left[\begin{array}{c|c} 0.0000 & 1.0000 \\ \hline 1.0000 & 1.0000 \\ \hline 0.0000 & 1.0000 \end{array}\right] \\ D &= \left[\begin{array}{c} 0.5000 \\ \hline 0.0000 \\ \hline 1.0000 \end{array}\right] \\ E &= \left[\begin{array}{c|c} 0.2000 & 0.1000 \\ \hline 0.0000 & -0.5000 \end{array}\right] \\ F &= \left[\begin{array}{c|c} 0.1000 & 0.1000 \\ \hline 1.0000 & 0.0000 \\ \hline 0.0000 & 1.0000 \end{array}\right], \end{aligned}$$

where $u_k \in \mathbb{R}$, $y_k \in \mathbb{R}^3$, and $x_k \in \mathbb{R}^2$. The input was considered a random binary signal generated using the following MATLAB command to create a random binary inpute signal, >> u=sign(randn(N,1));, where $N = 450$ is the number of observations, $i = 4$ and $j = 400$ are, respectively, the number of row blocks and columns of the Hankel data matrices. The dimensions of the Hankel data matrices are $Y_\bullet \in \mathbb{R}^{45\times 400}$ and $U_\bullet \in \mathbb{R}^{15\times 400}$. The rank of H is $2mi + n = 32$. After applying the deterministic bilinear subspace identification algorithms, we obtained the following parameter matrices:

$$\begin{aligned} A_b &= \left[\begin{array}{c|c} 0.4422 & 0.9973 \\ \hline -0.0831 & -0.1422 \end{array}\right] \\ B_b &= \left[\begin{array}{c} 0.3309 \\ \hline 0.2297 \end{array}\right] \\ C_b &= \left[\begin{array}{c|c} 1.2965 & -0.6155 \\ \hline 1.6616 & 0.8880 \\ \hline 1.2965 & -0.6155 \end{array}\right] \\ D_b &= \left[\begin{array}{c} 0.1662 \\ \hline 0.3651 \\ \hline 1.2965 \end{array}\right] \\ E_b &= \left[\begin{array}{c|c} 0.2806 & 0.2831 \\ \hline 0.0909 & 0.5963 \end{array}\right] \\ F_b &= \left[\begin{array}{c|c} 0.0888 & 0.5000 \\ \hline 1.5036 & 0.0000 \\ \hline -0.6155 & 1.0000 \end{array}\right]. \end{aligned}$$

Table 1 compares the eigenvalues of A and E for both simulated and identified models.

Table 1. Eigenvalues of observed and identified system eigenvalues.

	$\lambda_1(\bullet)$	$\lambda_2(\bullet)$	$\hat{\lambda}_1(\bullet)$	$\hat{\lambda}_2(\bullet)$
A	0.2000	0.10000	0.2000	0.1000
E	0.2000	−0.5000	−0.5000	0.2000

As can be seen, the identification algorithm performed as expected, reproducing the exact eigenvalues of both the A and E matrices. Finally, the first five outputs and their fitted values are given in Table 2.

Table 2. Observed and fitted output values.

k	y_{1_k}	y_{2_k}	y_{3_k}	$\hat{y}_{1_k}$	$\hat{y}_{2_k}$	$\hat{y}_{3_k}$
1	0.48	0.00	0.96	0.48	−0.00	0.96
2	−0.61	−0.03	−1.03	−0.61	−0.03	−1.03
3	0.47	1.00	−1.06	0.47	1.00	−1.06
4	−0.09	−2.42	−0.10	−0.09	−2.42	−0.10
5	−0.42	0.35	−0.77	−0.42	0.35	−0.77

VI. Conclusions

In this paper we have formulated a bilinear system as a rectangular linear system of variable dimensions. Within this framework, we were able to model the bilinear system as an equivalent linear system in a higher dimensional space. This led to an equivalent deterministic system identification algorithm as in [20]. Here, the Hankel data matrices are very similar to those in [12] but the subspace equations come out in such a way that the linear systems structure is evident. The algorithm requires a condition on the number of inputs and outputs. We simulated data from a single input, three output bilinear system and used this data to identify the model. The results show that the algorithm was able to satisfy the rank condition inherent of subspace algorithms, thus producing very accurate results. It still remains an open problem to find ways of overcoming the curse of dimensionality, as well as the conditions on the number of inputs and outputs.

REFERENCES

[1] Arun, K. S. and S. Y. Kung, "Balanced Approximation of stochastic systems,"*SIAM Journal on Matrix Analysis and Applications*, Vol. 11, No. 1, pp. 42-68, 1990.

[2] Bruni, C. Dipillo, G., and G. Koch, "Bilinear systems: An appealing class of nearly linear systems in theory and applications," *IEEE Trans. Automat. Contr.*, **AC-19**, No. 8, pp. 334-348, 1974.

[3] Chen, H. and J. Maciejowski, "Subspace identification of combined deterministic-stochastic bilinear systems," Proceedings of the IFAC SYSID 2000, Santa Barbara, USA, pp. , 2000.

[4] Chen, H. and J. Maciejowski, "An improved subspace identification method for bilinear systems," Proceedings of the 39^{th} IEEE Conference on Decision and Control, Sidney, Australia, 2000.

[5] Cho, Y. M., Xu, G., and T. Kailath, "Fast identification of state-space models via exploitation of displacement structure," *IEEE Trans. Automat. Contr.*, **AC-39**, No. 10, pp. –, 1994.

[6] Chun, T. C., "*Geometry of Linear Systems and Identification*," Ph.D. dissertation, Cambridge University, England, 1994.

[7] Desai, U. B., Pal, D., and R. Kirkpatrick, "A realization approach to stochastic model reduction," *Int. Journal of Control*, Vol. 42, No. 4, pp. 821-838, 1985.

[8] Deistler, M., Peternell, K., and W. Scherrer, "Consistency and relative efficiency of subspace methods," Proc. SYSID'94, Vol. 2, pp. 157-163, 4-6 July, Copenhagen, Denmark, 1994.

[9] Desai, U. B., "Realization of bilinear stochastic systems," *IEEE Trans. Automat. Contr.*, **AC-31**, No. 2, pp. 189-192, 1986.

[10] Dos Santos, P.L., Ramos, J.A., and J.L. Martins do Carvalho, "Identification of bilinear subspace systems using an iterative deterministic-stochastic subspace approach," Proceedings 44^{th} IEEE Conference for Decision and Control, March 2005.

[11] Favoreel, W., De Moor, B., and P. Vanoverschee, "Subspace identification of bilinear systems subject to white inputs," *IEEE Trans. Automat. Contr.*, **AC-44**, No. 6, pp. 1157-1169, 1999.

[12] Favoreel, W., "Subspace methods for identification and control linear and bilinear systems," Ph.D. Thesis, Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium, 1999.

[13] Hsu, C. S., Desai, U. B., and C. A. Crawley, "Realization and approximation of discrete bilinear systems," In: Applied Digital Control, S. G. Tzafestas (ed.), Elsevier Science Publishers, North-Holland, pp. 171-187, 1985.

[14] Isidori, A., "Direct construction of minimal bilinear realizations from nonlinear input-output maps," *IEEE Trans. Automat. Contr.*, **AC-18**, No. 6, pp. 626-631, 1973.

[15] 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.

[16] Larimore, W. E., "The optimality of canonical variate identification by example, filtering, and adaptive control," Proc. SYSID'94, Vol. 2, pp. 151-156, 4-6 July, Copenhagen, Denmark, 1994.

[17] Magnus, J. R. and H. Neudecker, "*Matrix differential calculus with applications in statistics and economics,*" Wiley Publisher, New York, 1999.

[18] McKelvey, T., "*On State-Space Models in System Identification*," Ph.D. dissertation, Linköping University, Sweden, 1994.

[19] Mohler, R. and W. J. Kolodziej, "An overview of bilinear system theory and applications," *IEEE Trans. Man and Cybernetics*, vol. 10, no. 10, pp. 683-688, 1980.

[20] Moonen, M., De Moor, B., Vandenberghe, L., and J. Vandewalle, "On- and off-line identification of linear state space models," *Int. Journal of Control*, Vol. 49, No. 1, pp. 219-232, 1989.

[21] Verdult, V. and M. Verhaegen, "Subspace based identification of MIMO bilinear systems, Proc. of the ECC, Karlruhe, Germany, 1999.

[22] Ramos, J. A., Moonen, M., and D. Mallant, "Modeling the rainfall-runoff process with a bilinear state-space model," Proc. 1991 European Conference, Ghent, Belgium, November 1991.

[23] Ramos, J. A. and J. F. Durand, "Identification of nonlinear systems using a B-splines parametric subspace approach," Proceedings of the 1999 IEEE American Control Conference, San Diego, CA, June 1999.

[24] Swindlehurst, A., Roy, R., Ottersten, B., and T. Kailath, "System identification via weighted subspace fitting," Proc. American Contr. Conf., pp. 2158-2163, 1992.

[25] Verdult, V., Verhaegen, M., Chou, C. T., and M. Lovera, "Efficient and systematic identification of MIMO bilinear state space models, Proc. of the 37^{th} IEEE Conference on Decision and Control, Tampa, Florida, 1998.

[26] Verdult, V. and M. Verhaegen, "Bilinear state space systems for nonlinear dynamics," *Theory of Biosciences*, vol. 119, pp. 1-9, 2000.

[27] Verdult, V. and M. Verhaegen, "A kernel method for subspace identification of multivariable bilinear systems," Proc. of the 42^{nd} IEEE Conf. on Decision and Control, Maui, Hawaii, USA, December 2003.

[28] Verdult, V. and M. Verhaegen, "Subspace identification of multivariable linear parameter-varying systems," *Automatica*, Vol. 38, pp. 805-814, 2002.

[29] Verhaegen, M., "Subspace model identificatio, Part III: Analysis of ordinary output-error state space model identification class of algorithms," *Int. Journal of Control*, Vol. 58, No. 3, pp. 555-586, 1993.

[30] Verhaegen, M., "Identification of the deterministic part of MIMO state space models given in innovations form from input-output data," *Automatica*, vol. 30, no. 1, pp. 61-74, 1994.

[31] Verhaegen, M. and D. Westwick, "Identifying MIMO Hammerstein systems in the context of subspace model identification methods," *Int. Journal of Control*, Vol. 68, No. 2, pp. 331-349, 1996.

[32] Westwick, D. and M. Verhaegen, "Identifying MIMO Wiener systems using subspace model identification methods," *Signal Processing*, Vol. 52, No. 2, pp. 235-258, 1996.

[33] Van Overschee, P. and B. De Moor, "Subspace algorithms for the stochastic identification problem," *Automatica*, Vol. 29, No. 3, pp. 649-660, 1993.

[34] Van Overschee, P. and B. De Moor, "*Subspace identification for linear systems*," Kluwer Academic Publishers, Boston, MA, 1996.

[35] Van Overschee, P. and B. De Moor, "N4SID: Subspace algorithms for the identification of combined deterministic and stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75-93, 1994.