2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Subspace System Identification of Separable-in-Denominator 2-D Stochastic Systems

José A. Ramos[1] and Paulo J. Lopes dos Santos[2]

[1] Nova Southeastern University
Farquhar College of Arts and Sciences
Division of Mathematics, Science, and Technology
3301 College Avenue, Fort Lauderdale, FL 33314, Email: jr1284@nova.edu

[2] Faculdade de Engenharia da Universidade do Porto
Departamento de Engenharia Electrotécnica e de Computadores
Rua Dr. Roberto Frias, Porto, Portugal 4200 465, Email: pjsantos@fe.up.pt

Abstract— **The fitting of a causal dynamic model to an image is a fundamental problem in image processing, pattern recognition, and computer vision. There are numerous other applications that require a causal dynamic model, such as in scene analysis, machined parts inspection, and biometric analysis, to name only a few. There are many types of causal dynamic models that have been proposed in the literature, among which the autoregressive moving average (ARMA) and state-space models are the most widely known. In this paper we introduce a 2-D stochastic state-space system identification algorithm for obtaining stochastic 2-D, causal, recursive, and separable-in-denominator (CRSD) models in the Roesser state-space form. The algorithm is tested with a real image and the reconstructed image is shown to be almost indistinguishable to the true image.**

I. INTRODUCTION

This paper deals with the parametric identification of a quarter-plane (QP) causal linear space-invariant (LSI) 2-D system excited by an unknown zero-mean white Gaussian noise process. There has been a significant amount of work in 2-D system identification and parametric modeling of 2-D stationary random processes during the last two decades. However, most of the work deals with parametric 2-D autoregressive moving average models, which have received great attention in a wide range of image and signal processing applications. These include image restoration, image compression, stochastic texture analysis, modeling of 2-D data, and spectrum estimation of 2-D data [1], [2], [3], [10], [12], [13], to name only a few. Lashgari, et al. [8] introduced an algorithm for image enhancement using notions of state-space and developed a minimum variance estimation algorithm. Fraanje, et al. [5] also developed a deterministic canonical 2-D subspace algorithm.

Here we solve the 2-D stochastic subspace identification problem for the CRSD model, along the lines of [7], [9], [11], [14], [16], and a new algorithm is introduced. The rest of the paper is organized as follows. In Section 2 the problem is briefly formulated. In Section 3 we derive the horizontal subspace equations, whereas in Section 4 we derive the vertical subspace equations. In section 5 we present the 2-D stochastic subspace system identification algorithm. In Section 6 we present a case study involving a real color image. In order to deal with color images, we have taken the red, green, and blue (RGB) images as the outputs of the system. Thus the data becomes a multivariate stochastic process of dimensions $\boldsymbol{Y} \in \mathbb{R}^{\ell(N+1)\times(M+1)}$, where $\ell = 3$. Conclusions are then drawn in Section 7.

II. PROBLEM FORMULATION

Consider the 2-D quarter plane causal, recursive, and separable-in-denominator (CRSD) stochastic system given in the Roesser state-space model form

$$\begin{aligned} x^h_{r+1,s} &= A_1 x^h_{r,s} + A_2 x^v_{r,s} + w^h_{r,s} & (1)\\ x^v_{r,s+1} &= A_4 x^v_{r,s} + w^v_{r,s} & (2)\\ y_{r,s} &= C_1 x^h_{r,s} + C_2 x^v_{r,s} + v_{r,s}, & (3)\end{aligned}$$

where $x^h_{r,s} \in \mathbb{R}^{n_h}$, $x^v_{r,s} \in \mathbb{R}^{n_v}$, and $y_{r,s} \in \mathbb{R}^{\ell}$ denote, respectively, the local horizontal state, local vertical state, and output vectors at the $(r,s)^{th}$ location of a finite spatial domain $\mathcal{D}$. The system matrices $\{A, C\}$ have partitioned dimensions $A_1 \in \mathbb{R}^{n_h\times n_h}$, $A_2 \in \mathbb{R}^{n_h\times n_v}$, $A_4 \in \mathbb{R}^{n_v\times n_v}$, $C_1 \in \mathbb{R}^{\ell\times n_h}$, and $C_2 \in \mathbb{R}^{\ell\times n_v}$. The noise vectors $w^h_{r,s} \in \mathbb{R}^{n_h}$, $w^v_{r,s} \in \mathbb{R}^{n_v}$, and $v_{r,s} \in \mathbb{R}^{\ell}$ are white noise processes with zero mean and joint covariance matrix given by

$$\text{cov}\left\{\begin{array}{c} w^h_{r,s}\\ w^v_{r,s}\\ v_{r,s}\end{array}\right\} = \left[\begin{array}{cc|c} Q_{hh} & Q_{hv} & S_h\\ Q_{vh} & Q_{vv} & S_v\\ \hline S_h^T & S_v^T & R\end{array}\right] = \left[\begin{array}{c|c} Q & S\\ \hline S^T & R\end{array}\right],$$

where M^T denotes the transpose of M, and $\{Q, R, S\}$ are covariance matrices of appropriate dimensions. The noise and state vectors are uncorrelated with each other, i.e., $\mathbb{E}\left\{x^h_{r,s}\left[\begin{array}{c|c}(w_{r',s'})^T & (v_{r',s'})^T\end{array}\right]\right\} = 0_{n_h\times(n+\ell)}$ and $\mathbb{E}\left\{x^v_{r,s}\left[\begin{array}{c|c}(w_{r',s'})^T & (v_{r',s'})^T\end{array}\right]\right\} = 0_{n_v\times(n+\ell)}$, $\forall\ r' \geq r$ and $s' \geq s$, where $w = \left[\begin{array}{c} w^h_{r,s}\\ w^v_{r,s}\end{array}\right]$, $n = n_h + n_v$, $\mathbb{E}$ is the expectation operator, and $0_{m\times n}$ denotes a zero matrix of dimensions $(m\times n)$. Furthermore, the states $x^h_{r,s}$ and $x^v_{r,s}$ have zero mean.

The positive definite state covariance is given by [4]

$$\Pi = \text{cov}\left\{ \begin{array}{c} x^h_{r,s} \\ x^v_{r,s} \end{array} \right\} = \left[\begin{array}{c|c} \Pi_h & 0_{n_h\times n_v} \\ \hline 0_{n_v\times n_h} & \Pi_v \end{array} \right].$$

If we now define the covariance of the state update as

$$\Pi' = \text{cov}\left\{ \begin{array}{c} x^h_{r+1,s} \\ x^v_{r,s+1} \end{array} \right\} = \left[\begin{array}{c|c} \Pi_h & \Pi_{hv} \\ \hline \Pi^T_{hv} & \Pi_v \end{array} \right],$$

where $\Pi_{hv} = A_2\Pi_v A_4^T + Q_{hv}$ and $\Pi_{vh} = \Pi^T_{hv}$. Then by taking the expectation on both sides of (1) – (2), we obtain the state covariance update equation as

$$\Pi' \;=\; A\Pi A^T + Q, \tag{4}$$

where $\Pi = \Pi^T$ and $\Pi' = (\Pi')^T$. Note that (4) is not a matrix Lyapunov state covariance equation since $\Pi' \neq \Pi$. However, by partitioning (4) one can decompose it into a pair of horizontal and vertical matrix Lyapunov type equations of the form (5) – (6) shown below. If we then add the symmetry constraints, we obtain the system

$$\Pi_h \;=\; A_1\Pi_h A_1^T + A_2\Pi_v A_2^T + Q_{hh} \tag{5}$$
$$\Pi_v \;=\; A_4\Pi_v A_4^T + Q_{vv},\ \Pi_h \;=\; \Pi_h^T,\ \ \Pi_v \;=\; \Pi_v^T. \tag{6}$$

Now vectorizing (5) – (6), one can then find vec$\{\Pi_h\}$ and vec$\{\Pi_v\}$ by solving the following system of equations [6]

$$\left[\begin{array}{c|c} I_{n_h^2} - A_1\otimes A_1 & A_2\otimes A_2 \\ \hline 0_{n_v^2\times n_h^2} & I_{n_v^2} - A_4\otimes A_4 \\ \hline I_{n_h^2} - \Theta_h & 0_{n_h^2\times n_v^2} \\ \hline 0_{n_v^2\times n_h^2} & I_{n_v^2} - \Theta_v \end{array} \right] \cdot \left[\begin{array}{c} \text{vec}\{\Pi_h\} \\ \text{vec}\{\Pi_v\} \end{array} \right] = \left[\begin{array}{c} \text{vec}\{Q_{hh}\} \\ \hline \text{vec}\{Q_{vv}\} \\ \hline 0_{n_h^2\times 1} \\ \hline 0_{n_v^2\times 1} \end{array} \right], \tag{7}$$

where I_k denotes a $k\times k$ identity matrix, $\Theta_h \in \mathbb{R}^{n_h^2\times n_h^2}$ and $\Theta_v \in \mathbb{R}^{n_v^2\times n_v^2}$ are permutation matrices such that vec$\{\Pi_h^T\} = \Theta_h$vec$\{\Pi_h\}$ and vec$\{\Pi_v^T\} = \Theta_v$vec$\{\Pi_v\}$, respectively, and $\otimes$ denotes the matrix Kronecker product.

Throughout the rest of the paper we will use the symbol > 0 (≥ 0) to indicate that a matrix is positive definite (positive semi-definite). Model (1) – (3) then satisfies the following constraints, also known as the positive real conditions: $\left[\begin{array}{c|c} Q & S \\ \hline S^T & R \end{array} \right] \geq 0$, $Q \geq 0$, $R > 0$, and $\Pi > 0$.

Now define G_1 and G_2 as the horizontal and vertical partitions of the matrix $G \in \mathbb{R}^{n\times\ell}$, as in [4], [14]

$$G \;=\; \mathbb{E}\left\{ \left[\begin{array}{c} x^h_{r+1,s} \\ x^v_{r,s+1} \end{array} \right] y^T_{r,s} \right\} = A\Pi C^T + S = \left[\begin{array}{c} G_1 \\ G_2 \end{array} \right]$$
$$G_1 \;=\; A_1\Pi_h C_1^T + A_2\Pi_v C_2^T + S_h$$
$$G_2 \;=\; A_4\Pi_v C_2^T + S_v,$$

then one can show that the 2-D output autocovariance sequence is given in terms of the Markov parameters of the system, i.e.,

$\Lambda_{k,l} = \mathbb{E}\left\{y_{r+k,s+l}y^T_{r,s}\right\}$	
$= C_1\Pi_h C_1^T + C_2\Pi_v C_2^T + R,$	if $k = 0,\ l = 0$
$= C_1 A_1^{k-1} G_1,$	if $k \geq 1,\ l = 0$
$= C_2 A_4^{l-1} G_2,$	if $k = 0,\ l \geq 1$
$= C_1 A_1^{k-1} A_2 A_4^{l-1} G_2,$	if $k \geq 1,\ l \geq 1.$

The problem can be stated as follows: Given a set of ℓ distinct $(N+1)\times(M+1)$ images whose (r,s) pixels are the output sequence $y_{r,s} \in \mathbb{R}^\ell$, for $r \in [0,N]$ and $s \in [0,M]$, find $\{n, A, C, G, \Pi, Q, R, S\}$. This will lead to a 2-D Kalman filter innovations model of (1) – (3).

III. Horizontal Data Processing

In order to save space, throughout the rest of the paper we will denote by hankel$\{(a_0, a_1, \ldots, a_k), n_1, n_2\}$ an $(n_1 \times n_2)$ Hankel matrix composed of the sequence $(a_0, a_1, \ldots, a_k)$ [7], [14], [16]. Let us now define the past and future state matrices, output, and innovations Hankel matrices for $k = 0, 1, \ldots, M$ and $N = 2i + j - 2$, as:

$$\begin{aligned}
\hat{X}^h_p(k) &= \left[\begin{array}{cccc} \hat{x}^h_{0,k} & \hat{x}^h_{1,k} & \cdots & \hat{x}^h_{j-1,k} \end{array} \right] \\
\hat{X}^h_f(k) &= \left[\begin{array}{cccc} \hat{x}^h_{i,k} & \hat{x}^h_{i+1,k} & \cdots & \hat{x}^h_{i+j-1,k} \end{array} \right] \\
\hat{X}^{hv}_p(k) &= \text{hankel}\{(\hat{x}^v_{0,k}, \hat{x}^v_{1,k}, \ldots, \hat{x}^v_{i+j-2,k}), n_v i, j\} \\
\hat{X}^{hv}_f(k) &= \text{hankel}\{(\hat{x}^v_{i,k}, \hat{x}^v_{i+1,k}, \ldots, \hat{x}^v_{2i+j-2,k}), n_v i, j\} \\
Y^h_p(k) &= \text{hankel}\{(y_{0,k}, y_{1,k}, , \ldots, y_{i+j,k}), \ell i, j\} \\
Y^h_f(k) &= \text{hankel}\{(y_{i,k}, y_{i+1,k}, \ldots, y_{2i+j-2,k}), \ell i, j\} \\
E^h_p(k) &= \text{hankel}\{(e_{0,k}, e_{1,k}, \ldots, e_{i+j-2,k}), \ell i, j\} \\
E^h_f(k) &= \text{hankel}\{(e_{i,k}, e_{i+1,k}, \ldots, e_{2i+j-2,k}), \ell i, j\}.
\end{aligned}$$

Note that the above definitions correspond to the 2-D CRSD innovations model of (1) – (3), and has the form

$$\hat{x}^h_{r+1,s} \;=\; A_1\hat{x}^h_{r,s} + A_2\hat{x}^v_{r,s} + K_h e_{r,s} \tag{8}$$
$$\hat{x}^v_{r,s+1} \;=\; A_4\hat{x}^v_{r,s} + K_v e_{r,s} \tag{9}$$
$$y_{r,s} \;=\; C_1\hat{x}^h_{r,s} + C_2\hat{x}^v_{r,s} + e_{r,s}, \tag{10}$$

where $\{\hat{x}^h_{r,s}, K_h\}$ and $\{\hat{x}^v_{r,s}, K_v\}$ are, respectively, the horizontal and vertical {state estimates, Kalman gain}, and $e_{r,s}$ is the innovations vector.

A. Horizontal Subspace Equations

Since the past and future vertical state matrices can be decoupled from the horizontal states, we will now derive the horizontal subspace equations. For this we need to define the following past and future extended state estimates, output, and innovations matrices

$$\hat{\boldsymbol{X}}^h_p \;=\; \left[\hat{X}^h_p(0) \,\middle|\, \hat{X}^h_p(1) \,\middle|\, \cdots \,\middle|\, \hat{X}^h_p(M) \right] \tag{11}$$
$$\hat{\boldsymbol{X}}^h_f \;=\; \left[\hat{X}^h_f(0) \,\middle|\, \hat{X}^h_f(1) \,\middle|\, \cdots \,\middle|\, \hat{X}^h_f(M) \right] \tag{12}$$
$$\hat{\boldsymbol{X}}^{hv}_p \;=\; \left[\hat{X}^{hv}_p(0) \,\middle|\, \hat{X}^{hv}_p(1) \,\middle|\, \cdots \,\middle|\, \hat{X}^{hv}_p(M) \right] \tag{13}$$

$$\hat{\boldsymbol{X}}_f^{hv} = \left[\ \hat{X}_f^{hv}(0) \mid \hat{X}_f^{hv}(1) \mid \cdots \mid \hat{X}_f^{hv}(M)\ \right] \quad (14)$$
$$\boldsymbol{Y}_p^h = \left[\ Y_p^h(0) \mid Y_p^h(1) \mid \cdots \mid Y_p^h(M)\ \right] \quad (15)$$
$$\boldsymbol{Y}_f^h = \left[\ Y_f^h(0) \mid Y_f^h(1) \mid \cdots \mid Y_f^h(M)\ \right] \quad (16)$$
$$\boldsymbol{E}_p^h = \left[\ E_p^h(0) \mid E_p^h(1) \mid \cdots \mid E_p^h(M)\ \right] \quad (17)$$
$$\boldsymbol{E}_f^h = \left[\ E_f^h(0) \mid E_f^h(1) \mid \cdots \mid E_f^h(M)\ \right]. \quad (18)$$

If we now define the extended block matrices [11], [14]

$$\Gamma_i^h = \left[\ C_1^T \mid (C_1A_1)^T \mid \cdots \mid (C_1A_1^{i-1})^T\ \right]$$
$$\Gamma_i^{hv} = \begin{bmatrix} C_2 & & & \\ C_1A_2 & C_2 & & \\ \vdots & \vdots & \ddots & \\ C_1A_1^{i-2}A_2 & C_1A_1^{i-3}A_2 & \cdots & C_2 \end{bmatrix}$$
$$H_i^h = \begin{bmatrix} I_\ell & & & \\ C_1K_h & I_\ell & & \\ \vdots & \vdots & \ddots & \\ C_1A_1^{i-2}K_h & C_1A_1^{i-3}K_h & \cdots & I_\ell \end{bmatrix}$$
$$\mathcal{C}_i^{hv} = \left[\ A_1^{i-1}A_2 \mid A_1^{i-2}A_2 \mid \cdots \mid A_2\ \right]$$
$$\mathcal{C}_i^h = \left[\ A_1^{i-1}K_h \mid A_1^{i-2}K_h \mid \cdots \mid K_h\ \right],$$

then one can show that the following horizontal subspace equations hold

$$\boldsymbol{Y}_p^h = \Gamma_i^h \hat{\boldsymbol{X}}_p^h + \Gamma_i^{hv} \hat{\boldsymbol{X}}_p^{hv} + H_i^h \boldsymbol{E}_p^h \quad (19)$$
$$\boldsymbol{Y}_f^h = \Gamma_i^h \hat{\boldsymbol{X}}_f^h + \Gamma_i^{hv} \hat{\boldsymbol{X}}_f^{hv} + H_i^h \boldsymbol{E}_f^h \quad (20)$$
$$\hat{\boldsymbol{X}}_f^h = A_1^i \hat{\boldsymbol{X}}_p^h + \mathcal{C}_i^{hv} \hat{\boldsymbol{X}}_p^{hv} + \mathcal{C}_i^h \boldsymbol{E}_p^h. \quad (21)$$

Equations (19) – (21) are the heart of the horizontal portion of the 2-D stochastic subspace identification algorithm, which will be presented in Section 6.

B. *Horizontal Projections $\boldsymbol{Y}_f^h|\boldsymbol{Y}_p^h$ and $\boldsymbol{Y}_p^h|\boldsymbol{Y}_f^h$ [14]*

Assuming rank$\{\boldsymbol{\Gamma}_i^h\} = n_h$, we will now derive the horizontal portion of the 2-D stochastic subspace identification algorithm. First we need to introduce the horizontal past and future output covariance matrices, $\boldsymbol{R}_{pp}^h$ and $\boldsymbol{R}_{ff}^h$, along with the horizontal cross covariance matrix between the future and the past, $\boldsymbol{H}_{fp}^h$. The above mentioned covariances are, respectively, defined as $\boldsymbol{R}_{pp}^h = \boldsymbol{Y}_p^h \boldsymbol{D} (\boldsymbol{Y}_p^h)^T$, $\boldsymbol{R}_{ff}^h = \boldsymbol{Y}_f^h \boldsymbol{D} (\boldsymbol{Y}_f^h)^T$, and $\boldsymbol{H}_{fp}^h = \boldsymbol{Y}_f^h \boldsymbol{D} (\boldsymbol{Y}_p^h)^T$, where $\boldsymbol{D} = \text{diag}\{d_1, d_2, \ldots, d_{(M+1)j}\}$ is a set of weights such that $\sum_{k=1}^{(M+1)j} d_k = 1$, and $d_k > 0$, for $k = 1, 2, \ldots, (M+1)j$. Here we use $d_k = \frac{1}{(M+1)j}$, for $k = 1, 2, \ldots, (M+1)j$. Now define the projection of the future onto the past as

$$\begin{aligned} \boldsymbol{Y}_f^h|\boldsymbol{Y}_p^h &= \left[\boldsymbol{Y}_f^h \boldsymbol{D} (\boldsymbol{Y}_p^h)^T\right] \cdot \left[\boldsymbol{Y}_f^h \boldsymbol{D} (\boldsymbol{Y}_p^h)^T\right]^{-1} \cdot \boldsymbol{Y}_p^h \\ &= \boldsymbol{H}_{fp}^h (\boldsymbol{R}_{pp}^h)^{-1} \boldsymbol{Y}_p^h = \Gamma_i^h \cdot \Delta_i^h (\boldsymbol{R}_{pp}^h)^{-1} \boldsymbol{Y}_p^h \\ &= \Gamma_i^h \cdot \hat{\boldsymbol{X}}_f^h, \end{aligned} \quad (22)$$

where $\Delta_i^h = \left[\ A_1^{i-1}G_1 \mid A_1^{i-2}G_1 \mid \cdots \mid G_1\ \right]$ and $\boldsymbol{H}_{fp}^h = \Gamma_i^h \cdot \Delta_i^h$.

IV. Vertical Data Processing

Here we use the vertical model

$$\begin{aligned} \hat{x}_{r,s+1}^v &= A_4 \hat{x}_{r,s}^v + K_v e_{r,s} \\ e_{r,s}^h &= y_{r,s} - C_1 \hat{x}_{r,s}^h = C_2 \hat{x}_{r,s}^v + e_{r,s}. \end{aligned}$$

Since C_1 and $\hat{\boldsymbol{X}}_f^h = \{\hat{x}_{i,s}^h, \hat{x}_{i+1,s}^h, \hat{x}_{i+2,s}^h, \ldots, \hat{x}_{i+j-1,s}^h\}$, for $s = 0, 1, 2, \ldots, M$, can be determined from the horizontal model, we are now restricted to use $e_{r,s}^h$, for $r = i, i+1, \ldots, i+j-1$ and $s = 0, 1, 2, \ldots, M$ as the vertical data, where $M = 2i + j - 2$. Let us now define the past and future state, output, and innovations Hankel matrices in the vertical direction, for $k = i, i+1, \ldots, i+j-1$, as

$$\begin{aligned} \hat{X}_p^v(k) &= \left[\ \hat{x}_{k,0}^v \quad \hat{x}_{k,1}^v \quad \cdots \quad \hat{x}_{k,j-1}^v\ \right] \\ \hat{X}_f^v(k) &= \left[\ \hat{x}_{k,i}^v \quad \hat{x}_{k,i+1}^v \quad \cdots \quad \hat{x}_{k,i+j-1}^v\ \right] \\ Y_p^v(k) &= \text{hankel}\{(e_{k,0}^h, e_{k,1}^h, \ldots, e_{k,i+j-2}^h), \ell i, j\} \\ Y_f^v(k) &= \text{hankel}\{(e_{k,i}^h, e_{k,i+1}^h, \ldots, e_{k,2i+j-2}^h), \ell i, j\} \\ E_p^v(k) &= \text{hankel}\{(e_{k,0}, e_{k,1}, \ldots, e_{k,i+j-2}), \ell i, j\} \\ E_f^v(k) &= \text{hankel}\{(e_{k,i}, e_{k,i+1}, \ldots, e_{k,2i+j-2}), \ell i, j\}. \end{aligned}$$

Let us now concatenate all state, noise, and data matrices for $k = i, i+1, \ldots, i+j-1$. That is,

$$\begin{aligned} \hat{\boldsymbol{X}}_p^v &= \left[\ \hat{X}_p^v(i) \mid \hat{X}_p^v(i+1) \mid \cdots \mid \hat{X}_p^v(i+j-1)\ \right] \\ \hat{\boldsymbol{X}}_f^v &= \left[\ \hat{X}_f^v(i) \mid \hat{X}_f^v(i+1) \mid \cdots \mid \hat{X}_f^v(i+j-1)\ \right] \\ \boldsymbol{E}_p^v &= \left[\ E_p^v(i) \mid E_p^v(i+1) \mid \cdots \mid E_p^v(i+j-1)\ \right] \\ \boldsymbol{E}_f^v &= \left[\ E_f^v(i) \mid E_f^v(i+1) \mid \cdots \mid E_f^v(i+j-1)\ \right] \\ \boldsymbol{Y}_p^v &= \left[\ Y_p^v(i) \mid Y_p^v(i+1) \mid \cdots \mid Y_p^v(i+j-1)\ \right] \\ \boldsymbol{Y}_f^v &= \left[\ Y_f^v(i) \mid Y_f^v(i+1) \mid \cdots \mid Y_f^v(i+j-1)\ \right]. \end{aligned}$$

Finally, we define the extended vertical parameter matrices as

$$\Gamma_i^v = \left[\ (C_2)^T \mid (C_2A_4)^T \mid \cdots \mid (C_2A_4^{i-1})^T\ \right]$$
$$H_i^v = \begin{bmatrix} I_\ell & & & \\ C_2K_v & I_\ell & & \\ \vdots & \vdots & \ddots & \\ C_2A_4^{i-2}K_v & C_2A_4^{i-3}K_v & \cdots & I_\ell \end{bmatrix}$$
$$\mathcal{C}_i^v = \left[\ A_4^{i-1}K_v \mid A_4^{i-2}K_v \mid \cdots \mid K_v\ \right].$$

We are now ready to define the subspace equations in the vertical direction. Here the matrices $Y_p^v(k)$ and $Y_f^v(k)$, for $k = i, i+1, \ldots, i+j-1$, will be the data used in the vertical portion of the 2-D stochastic subspace identification algorithm.

A. *Vertical Subspace Equations*

The subspace equations in the vertical direction can now be written as

$$\boldsymbol{Y}_p^v = \Gamma_i^v \hat{\boldsymbol{X}}_p^v + H_i^v \boldsymbol{E}_p^v \quad (23)$$
$$\boldsymbol{Y}_f^v = \Gamma_i^v \hat{\boldsymbol{X}}_f^v + H_i^v \boldsymbol{E}_f^v \quad (24)$$
$$\hat{\boldsymbol{X}}_f^v = A_4^i \hat{\boldsymbol{X}}_p^v + \mathcal{C}_i^v \boldsymbol{E}_p^v. \quad (25)$$

We will now develop the projection equations that will allow us to identify the parameters of the vertical model.

B. Vertical Projections $\boldsymbol{Y}^v_f|\boldsymbol{Y}^v_p$ and $\boldsymbol{Y}^v_p|\boldsymbol{Y}^v_f$ [14]

As pointed out earlier, the 2-D stochastic subspace identification algorithm will depend on a fundamental rank condition from which the system orders can be determined. Toward this end and assuming $\text{rank}\{\boldsymbol{\Gamma}^v_i\} = n_v$, we will now derive the vertical portion of the 2-D stochastic subspace identification algorithm. First we need to introduce the vertical past and future output covariance matrices, $\boldsymbol{R}^v_{pp}$ and $\boldsymbol{R}^v_{ff}$, along with the vertical cross covariance matrix between the future and the past, $\boldsymbol{H}^v_{fp}$. These are respectively defined as $\boldsymbol{R}^v_{pp} = \boldsymbol{Y}^v_p\boldsymbol{D}(\boldsymbol{Y}^v_p)^T$, $\boldsymbol{R}^v_{ff} = \boldsymbol{Y}^v_f\boldsymbol{D}(\boldsymbol{Y}^v_f)^T$, and $\boldsymbol{H}^v_{fp} = \boldsymbol{Y}^v_f\boldsymbol{D}(\boldsymbol{Y}^v_p)^T$. We now define the projection of the future onto the past, $\boldsymbol{Y}^v_f|\boldsymbol{Y}^v_p$, as

$$\begin{aligned}\boldsymbol{Y}^v_f|\boldsymbol{Y}^v_p &= \left[\boldsymbol{Y}^v_f\boldsymbol{D}(\boldsymbol{Y}^v_p)^T\right]\cdot\left[\boldsymbol{Y}^v_f\boldsymbol{D}(\boldsymbol{Y}^v_p)^T\right]^{-1}\cdot\boldsymbol{Y}^v_p \\ &= \Gamma^v_i\cdot\hat{\boldsymbol{X}}^v_f, \end{aligned}\tag{26}$$

where $\Delta^v_i = \left[\begin{array}{c|c|c|c} A_4^{i-1}G_2 & A_4^{i-2}G_2 & \cdots & G_2\end{array}\right]$ and $\boldsymbol{H}^v_{fp} = \Gamma^v_i\cdot\Delta^v_i$.

V. 2-D Stochastic Identification Algorithm

In this section we present an algorithm for the 2-D stochastic subspace system identification problem. For further details see [7], [?], [9], [11], [14], [16].

A. Stochastic 4SID Algorithm

1) Assemble the horizontal data matrices $\{\boldsymbol{Y}^h_p, \boldsymbol{Y}^h_f\}$ and compute the LQ decomposition of

$$\frac{1}{\sqrt{j(M+1)}}\begin{bmatrix}\boldsymbol{Y}^h_p\\ \boldsymbol{Y}^h_f\end{bmatrix} = \begin{bmatrix}L^h_{11} & 0_{\ell i\times\ell i}\\ L^h_{21} & L^h_{22}\end{bmatrix}\begin{bmatrix}Q^h_1\\ Q^h_2\end{bmatrix}.$$

2) Compute the orthogonal projection $\boldsymbol{Y}^h_f|\boldsymbol{Y}^h_p$ from

$$\begin{aligned}\boldsymbol{Y}^h_f|\boldsymbol{Y}^h_p &= [\boldsymbol{Y}^h_f\boldsymbol{D}(\boldsymbol{Y}^h_p)^T][\boldsymbol{Y}^h_p\boldsymbol{D}(\boldsymbol{Y}^h_p)^T]^{-1}\boldsymbol{Y}^h_p \\ &= L^h_{21}\left(L^h_{11}\right)^{-1}\boldsymbol{Y}^h_p.\end{aligned}$$

3) Perform the singular value decomposition (SVD)

$$\begin{aligned}\boldsymbol{Y}^h_f|\boldsymbol{Y}^h_p &= \left[\begin{array}{c|c} U_h & U_h^\perp\end{array}\right]\begin{bmatrix}\Sigma_h & \times\\ \times & \times\end{bmatrix}\begin{bmatrix}V_h^T\\ (V_h^\perp)^T\end{bmatrix} \\ &= U_h\Sigma_h^{\frac{1}{2}}\cdot\Sigma_h^{\frac{1}{2}}V_h^T = \Gamma^h_i\cdot\hat{\boldsymbol{X}}^h_f,\end{aligned}$$

where $U_h\in\mathbb{R}^{\ell i\times n_h}$, $\Sigma_h\in\mathbb{R}^{n_h\times n_h}$, $V_h\in\mathbb{R}^{n_h\times j(M+1)}$, $\times$ denotes a zero matrix of appropriate dimensions, $\Sigma_h = \text{diag}\{\sigma^h_1, \sigma^h_2, \ldots, \sigma^h_{n_h}\}$ denotes the n_h nonzero singular values in descending order.

4) Recover $\hat{\boldsymbol{X}}^h_f = \left(\Gamma^h_i\right)^\dagger L^h_{21}\left(L^h_{11}\right)^{-1}\boldsymbol{Y}^h_p$, where $\left(\Gamma^h_i\right)^\dagger$ denotes the pseudo inverse of Γ^h_i.

5) Compute the first ℓ rows of $\boldsymbol{E}^h_f = \boldsymbol{Y}^h_f - \boldsymbol{Y}^h_f|\boldsymbol{Y}^h_p$, i.e., $\boldsymbol{E}^h_f(1:\ell,:) = \left[\begin{array}{c|c|c|c}\boldsymbol{e}^h(0) & \boldsymbol{e}^h(1) & \cdots & \boldsymbol{e}^h(M)\end{array}\right]$, which provides the residuals to be used in the vertical model since $C_1 = \Gamma^h_i(1{:}\ell,:)$ and for $k\in[0,M]$, we have $\boldsymbol{e}^h(k) = \boldsymbol{y}^h(k) - C_1\hat{\boldsymbol{x}}^h(k)$, and

$$\begin{aligned}\boldsymbol{e}^h(k) &= \left[\begin{array}{c|c|c|c} e^h_{i,k} & e^h_{i+1,k} & \cdots & e^h_{i+j-1,k}\end{array}\right]\\ \boldsymbol{y}^h(k) &= \left[\begin{array}{c|c|c|c} y_{i,k} & y_{i+1,k} & \cdots & y_{i+j-1,k}\end{array}\right]\\ \hat{\boldsymbol{x}}^h(k) &= \left[\begin{array}{c|c|c|c} \hat{x}_{i,k} & \hat{x}_{i+1,k} & \cdots & \hat{x}_{i+j-1,k}\end{array}\right].\end{aligned}$$

6) Using the residuals $\{e^h_{i,k}, e^h_{i+1,k}, \ldots, e^h_{i+j-1,k}\}$ for $k = 0,1,\ldots,M$, assemble the vertical data matrices $\{\boldsymbol{Y}^v_p, \boldsymbol{Y}^v_f\}$ and compute the LQ decomposition of

$$\frac{1}{j^2}\begin{bmatrix}\boldsymbol{Y}^v_p\\ \boldsymbol{Y}^v_f\end{bmatrix} = \begin{bmatrix}L^v_{11} & 0_{\ell i\times\ell i}\\ L^v_{21} & L^v_{22}\end{bmatrix}\begin{bmatrix}Q^v_1\\ Q^v_2\end{bmatrix}.$$

7) Compute the orthogonal projection $\boldsymbol{Y}^v_f|\boldsymbol{Y}^v_p$ from

$$\begin{aligned}\boldsymbol{Y}^v_f|\boldsymbol{Y}^v_p &= [\boldsymbol{Y}^v_f\boldsymbol{D}(\boldsymbol{Y}^v_p)^T][\boldsymbol{Y}^v_p\boldsymbol{D}(\boldsymbol{Y}^v_p)^T]^{-1}\boldsymbol{Y}^v_p \\ &= L^v_{21}\left(L^v_{11}\right)^{-1}\boldsymbol{Y}^v_p.\end{aligned}$$

8) Perform the singular value decomposition (SVD)

$$\begin{aligned}\boldsymbol{Y}^v_f|\boldsymbol{Y}^v_p &= \left[\begin{array}{c|c} U_v & U_v^\perp\end{array}\right]\begin{bmatrix}\Sigma_v & \times\\ \times & \times\end{bmatrix}\begin{bmatrix}V_v^T\\ (V_v^\perp)^T\end{bmatrix} \\ &= U_v\Sigma_v^{\frac{1}{2}}\cdot\Sigma_v^{\frac{1}{2}}V_v^T = \Gamma^v_i\cdot\hat{\boldsymbol{X}}^v_f,\end{aligned}$$

where $U_v\in\mathbb{R}^{\ell i\times n_v}$, $\Sigma_v\in\mathbb{R}^{n_v\times n_v}$, $V_v\in\mathbb{R}^{n_v\times j^2}$, $\Sigma_v = \text{diag}\{\sigma^v_1, \sigma^v_2, \ldots, \sigma^v_{n_v}\}$ denotes the n_v nonzero singular values in descending order.

9) Recover $\hat{\boldsymbol{X}}^v_f = (\Gamma^v_i)^\dagger L^v_{21}\left(L^v_{11}\right)^{-1}\boldsymbol{Y}^v_p$, where $(\Gamma^v_i)^\dagger$ denotes the pseudo inverse of Γ^v_i.

10) Assemble the state matrices $\hat{\boldsymbol{X}}^h_f\in\mathbb{R}^{n_h\times j(M+1)}$ and $\hat{\boldsymbol{X}}^v_f\in\mathbb{R}^{n_v\times j^2}$ into $\hat{\boldsymbol{X}}^h = \text{reshape}\{\hat{\boldsymbol{X}}^h_f, n_hj, M+1\}$ and $\hat{\boldsymbol{X}}^v = \text{reshape}\{\hat{\boldsymbol{X}}^v_f, n_vj, j\}$, respectively, where $\text{reshape}\{M, n_1n_2, n_3\}$ takes the columns of $M\in\mathbb{R}^{n_1\times n_2n_3}$ and first converts them into $(n_1n_2\times 1)$ block rows, then stacks these columnwise into an $(n_1n_2\times n_3)$ matrix. Extract the same common information from the raw image matrix $\boldsymbol{Y}\in\mathbb{R}^{\ell(N+1)\times(M+1)}$. The i through $i+j-1$ block rows and i through $i+j-1$ columns are the ones in common for all matrices, thus we extract these as $\hat{\boldsymbol{X}}^h_c = \hat{\boldsymbol{X}}^h(1:n_hj, i+1:i+j)$, $\hat{\boldsymbol{X}}^v_c = \hat{\boldsymbol{X}}^v(1:n_vj, 1:j)$, and $\boldsymbol{Y}_c = \boldsymbol{Y}(\ell i+1:\ell j, i+1:i+j)$.

11) Solve the overdetermined system of equations

$$\begin{bmatrix}\hat{\boldsymbol{x}}^h_2\\ \hat{\boldsymbol{x}}^v_2\\ \boldsymbol{y}_1\end{bmatrix} = \left[\begin{array}{c|c} A_1 & A_2\\ \hline 0_{n_v\times n_h} & A_4\\ \hline C_1 & C_2\end{array}\right]\begin{bmatrix}\hat{\boldsymbol{x}}^h_1\\ \hat{\boldsymbol{x}}^v_1\end{bmatrix} + \begin{bmatrix}\boldsymbol{w}^h\\ \boldsymbol{w}^v\\ \boldsymbol{v}\end{bmatrix},$$

where $\hat{\boldsymbol{x}}^h_1 = \text{reshape}\{\hat{\boldsymbol{X}}^h_c, n_h, j^2\}$, $\hat{\boldsymbol{x}}^v_1 = \text{reshape}\{\hat{\boldsymbol{X}}^v_c, n_v, j^2\}$, $\boldsymbol{y}_1 = \text{reshape}\{\boldsymbol{Y}_c, \ell, j^2\}$, $\hat{\boldsymbol{x}}^h_2 = \text{reshape}\{\hat{\boldsymbol{X}}^h_c, n_h, j^2\}$, and $\hat{\boldsymbol{x}}^v_2 = \text{reshape}\{\hat{\boldsymbol{X}}^v_c, n_v, j^2\}$. The least squares solution is given by

$$\left[\begin{array}{c|c} \hat{A}_1 & \hat{A}_2\\ \hline 0_{n_v\times n_h} & \hat{A}_4\\ \hline \hat{C}_1 & \hat{C}_2\end{array}\right] = \begin{bmatrix}\hat{\boldsymbol{x}}^h_2\\ \hat{\boldsymbol{x}}^v_2\\ \boldsymbol{y}_1\end{bmatrix}\cdot\begin{bmatrix}\hat{\boldsymbol{x}}^h_1\\ \hat{\boldsymbol{x}}^v_1\end{bmatrix}^\dagger.$$

12) Compute residuals from step 11, . $\hat{\boldsymbol{w}}^h$, $\hat{\boldsymbol{w}}^v$, and $\hat{\boldsymbol{v}}$.

13) Compute the noise covariance matrices from

$$\left[\begin{array}{cc|c} Q_{hh} & Q_{hv} & S_h \\ Q_{vh} & Q_{vv} & S_v \\ \hline S_h^T & S_v^T & R \end{array}\right] = \operatorname{cov}\left\{\left[\begin{array}{c} \hat{\boldsymbol{w}}^h \\ \hat{\boldsymbol{w}}^v \\ \hat{\boldsymbol{v}} \end{array}\right]\right\}.$$

14) Solve the pair of Lyapunov equations from (7) to obtain Π_h and Π_v.

15) Compute G and $\Lambda_{0,0}$ from

$$G = A\Pi C^T + S \text{ and } \Lambda_{0,0} = C\Pi C^T + R.$$

16) Solve the Riccati equations for P_h and P_v from

$$\begin{aligned} P_h &= A_1 P_h A_1^T + A_2 P_v A_2^T \\ &\quad + (G_1 - A_1 P_h C_1^T - A_2 P_v C_2^T) \\ &\quad \times (\Lambda_{0,0} - C_1 P_h C_1^T - C_2 P_v C_2^T)^{-1} \\ &\quad \times (G_1 - A_1 P_h C_1^T - A_2 P_v C_2^T)^T \\ P_v &= A_4 P_v A_4^T + (G_2 - A_4 P_h C_2^T) \\ &\quad \times (\Lambda_{0,0} - C_1 P_h C_1^T - C_2 P_v C_2^T)^{-1} \\ &\quad \times (G_2 - A_4 P_v C_2^T)^T. \end{aligned}$$

17) Compute the Kalman gain matrices K_h and K_v from

$$\begin{aligned} K_h &= (G_1 - A_1 P_h C_1^T - A_2 P_v C_2^T) \\ &\quad \times (\Lambda_{0,0} - C_1 P_h C_1^T - C_2 P_v C_2^T)^{-1} \\ K_v &= (G_2 - A_4 P_h C_2^T) \\ &\quad \times (\Lambda_{0,0} - C_1 P_h C_1^T - C_2 P_v C_2^T)^{-1}. \end{aligned}$$

18) Using initial conditions $\hat{x}^h_{0,s} = 0_{n_h \times 1}$ for $s = 0, 1, \ldots, M$ and $\hat{x}^v_{r,0} = 0_{n_v \times 1}$ for $r = 0, 1, \ldots, N$, compute the enhanced image $\hat{\boldsymbol{Y}} \in \mathbb{R}^{\ell(N+1)\times(M+1)}$ from the 2-D Kalman filter:

$$\begin{aligned} \hat{y}_{r,s} &= C_1 \hat{x}^h_{r,s} + C_2 \hat{x}^v_{r,s} \\ e_{r,s} &= y_{r,s} - \hat{y}_{r,s} \\ \hat{x}^h_{r+1,s} &= A_1 \hat{x}^h_{r,s} + A_2 \hat{x}^v_{r,s} + K_h e_{r,s} \\ \hat{x}^v_{r,s+1} &= A_4 \hat{x}^v_{r,s} + K_v e_{r,s}. \end{aligned}$$

19) End 4SID.

Note that 12)–17) are the purely stochastic version of N4SID [14].

VI. Case Study

An image can be modeled as a 2-D stochastic process of the form

$$y_{r,s} = y^{true}_{r,s} + v_{r,s}, \qquad (27)$$

where $y_{r,s}$ is the measured image, $y^{true}_{r,s}$ is the unknown true image, and $v_{r,s}$ is a white noise process. We want to make a clear distinction here between the deblurring problem and the image modeling problem. The former is a deconvolution problem, whereas the latter is a decomposition of a stochastic process into a true process and additive noise, and acts like a measurement device. The algorithm was tested with the classical Lena image, obtained from the University of Southern California image repository [15]. The original Lena image is shown in Figure 1 and the reconstructed image resulting from the application of the algorithm is shown in Figure 2. The innovations were plotted as an inverted image and is shown in Figure 3. As can be seen from Figure 1 – 3, the 2-D Kalman filter model, whose horizontal and vertical orders were $n_h = 8$ and $n_v = 7$, respectively, performed fairly well in recovering the original image. The variance accounted for (VAF) was 99.57 and the equivalent signal-to-noise-ratio (SNR) between the original and residual image was 24.88. These two statistics indicate a very good performance of the algorithm. Other performance measures used to assess the algorithm are graphs of the original and fitted images, along with the innovations, all plotted as time series. These are shown in Table 1 and indicate an excellent performance of the algorithm. A whiteness test done on the autocorrelations of the innovations verified the model assumptions. Finally, the horizontal and vertical singular value plots are shown in Table 1.

Fig. 1. Original Lena image.

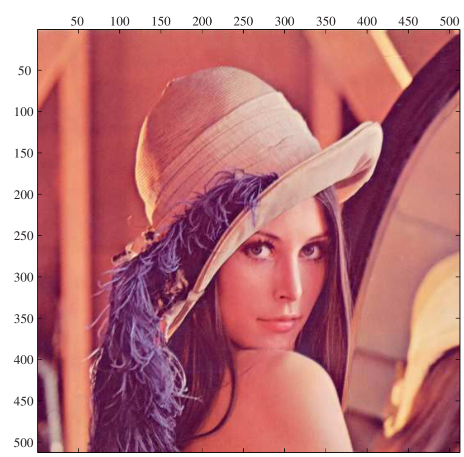


Fig. 2. Reconstructed Lena image

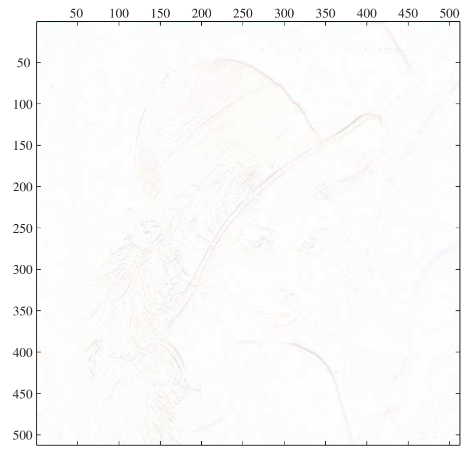
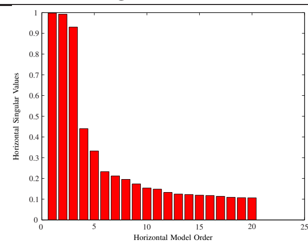
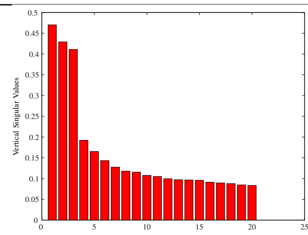
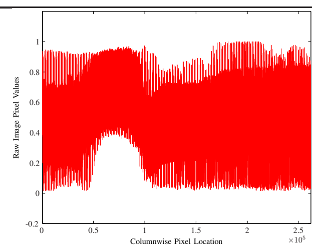
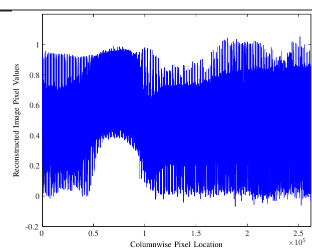
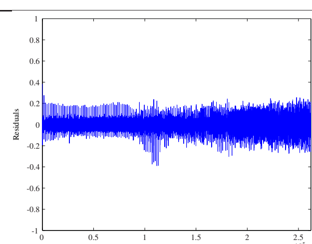


Fig. 3. Residual Lena image.

Property	Lena Image
Horizontal Singular Values	
Vertical Singular Values	
Raw Image Trace	
Reconstructed Image Trace	
Innovations	

VII. CONCLUSIONS

We have introduced a new 2-D stochastic state space subspace system identification (4SID) algorithm for modeling 2-D stochastic processes. The algorithm is based on a causal, recursive, separable-in-denominator (CRSD) model and takes advantage of this structure to decompose the problem into a cascade of horizontal and vertical system identification sub-algorithms. At this stage this is a reasonable assumption due to the fact that the states of a general 2-D stochastic Roesser model are coupled. Instead we approached the problem by a separable-in-denominator structure, thus leading to a CRSD 2-D Kalman filter model for processing 2-D stochastic processes. The algorithm was tested with a real image and it performed very well.

REFERENCES

[1] S. Attasi, "Modelling and recursive estimation for double indexed sequences," In: R.K. Mehra and D.G. Lainiotis (eds.), *System Identification: Advances and Case Studies*, pp. 289-348, Academic Press, New York, 1976.

[2] J. Biemond and F. Van der Putten, "Image restoration using a parallel identification and filtering procedure," IEEE International Conference on ICASSP-85, vol. 10, pp. 660663, March 1985.

[3] J. A. Cadzow and K. Ogino, "Two-dimensional spectral estimation," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 29, no. 3, 396401, 1981.

[4] P. E. Caines, *Linear Stochastic Systems*, Wiley, New York, 1987.

[5] R. Fraanje and M. Verhaegen, "A spatial canonical approach to multidimensional state-space identification for distributed parameter systems," *4th International Workshop on Multidimensional Systems* (NDS 2005), Wuppertal, Germany, pp. 217-222, 2005.

[6] T. Kailath, *Linear Systems*, Prentice Hall, Englewood Cliffs, N. J., 1980.

[7] T. Katayama, *Subspace Methods for System Identification*, Springer Verlag, London, 2005.

[8] B. Lashgari, L. Silverman, and J. Abramatic, "Approximation of 2-D separable in denominator filters," *IEEE Trans. Circuits Systems*, vol. 30, no. 2, pp. 107-121, 1983.

[9] M. Moonen, B. De Moor, L. Vandenberghe, and J. Vandewalle, "On- and off-line identification of linear state space models," *International Journal of Control*, vol. 49, 219-232, 1989.

[10] Y. W. Nijim, S. D. Stearns, and W. B. Mikhael, "Linear ARMA predictors for the lossless compression of two-dimensional signals," *Digital Signal Process.*, vol. 7, no. 2, 120126 1997.

[11] Ramos, J., Alenany, A., Shang, H. and Lopes dod Ssantos, P., "Subspace Algorithms for Identifying Separable in Denominator 2-D Systems with Deterministic Inputs," to appear in *IET Journal of Control Applications*.

[12] R. P. Roesser, "A discrete-state-space model for linear image processing," *IEEE Trans. Automatic Control*, vol. 20, pp. 1-10, 1975.

[13] A. Rosenfeld, *Image Modelling*, Academic Press, New York, 1982.

[14] P. Van Overschee and B. De Moor, *Subspace identification for linear Systems*. Kluwer Academic Publisher, 1996.

[15] The USC-SIPI Image Database http://sipi.usc.edu/database.

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