Pergamon

0005-1098(94)00091-3

Automatica, Vol. 31, No. 2, pp. 201–216, 1995
Elsevier Science Ltd
Printed in Great Britain
0005-1098/95 $9.50 + 0.00

A Class of Subspace Model Identification Algorithms to Identify Periodically and Arbitrarily Time-varying Systems*

MICHEL VERHAEGEN† and XIAODE YU†

The problem of identifying linear time-varying systems is solved in a subspace model identification framework making use of an ensemble of input–output data. This strategy is applied to the identification of periodically time-varying systems using a single input–output data batch.

Key Words—Time-varying systems; periodically time-varying systems; state space model; ensemble identification; linear algebra; subspace model identification.

Abstract—Subspace model identification algorithms that allow the identification of a linear, time-varying (LTV) state space model from an ensemble set of input–output measurements are presented in this paper. Each pair of input and output sequences in this ensemble is recorded when the underlying system to be identified undergoes the same time-varying behavior. The algorithms operate directly on the available ensemble of input–output data and are a generalization of the recently proposed Multivariable Output Error State sPace (MOESP) class of algorithms to this ensemble type of identification problems. A special case is considered in this paper, where the repetition of this time-varying behavior is intrinsic, namely in periodically time-varying systems. An example of identifying a multirate sampled data system from a recorded input and output sequence demonstrates some of the capabilities of the presented subspace model identification algorithms.

1. INTRODUCTION

In the area of system identification, much effort has been spent in the development of identification schemes to identify linear time-invariant dynamical systems (Ljung, 1987; Söderström and Stoica, 1989). However, in reality, most systems demonstrate time-varying and/or nonlinear behavior. Since it is common practice in system engineering to treat non-linear systems as linear time-varying systems, the identification of linear time-varying systems should be an important topic in system identification.

As outlined in Kearney *et al.* (1991) and MacNeil *et al.* (1992) there are two classes of approaches to identify linear time-varying systems. The main characteristic of the first class is that it relies on a single time sequence of input and output quantities, of which important representatives are recursive identification schemes equipped with a mechanism to 'forget' past data and 'functional series' modeling techniques (Niedzwiecki, 1990). The second class is characterized by the use of an ensemble of input and output sequences, each exhibiting the same underlying time-varying behavior.

The methods of the first class have difficulties in identifying accurate models when no precise a priori information is available about the nature of the time-variation and/or the change in the dynamics of the underlying system occurs at a rate comparable to the sampling rate. For recursive identification schemes, with for example exponential forgetting, it is well-known that the use of a small forgetting factor can easily result in inaccurate model estimates. Furthermore, it has been shown in Niedzwiecki (1990) that when there is a slight error in the a priori information necessary to determine the structure of the series in functional series expansion methods, the latter approach may completely fail. Therefore, when there is only little information about the nature of the time-variation and/or the behavior changes rapidly, the first class of methods is inappropriate to use to estimate accurate models. When accurate models are required in these circumstances, a possible solution can be found via the second class of time-varying identification methods. This recommendation is also made in Widrow and Walach (1984), Niedzwiecki (1990), Kearney *et al.* (1991) and Yu and Verhaegen (1993).

* Received 21 October 1992; revised 6 August 1993; revised 31 January 1994; received in final form 10 June 1994. This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Paul van den Hof under the direction of Editor T. Söderström. Corresponding author Dr Michel Verhaegen. Tel. +31 15 78 14 42; Fax +53 15 62 32 71; E-mail verhaege@dutentb.et.tudelft.nl.

† Department of Electrical Engineering, Delft University of Technology, P.O. Box 5031, Delft, The Netherlands.

Despite this recommendation, ensemble identification techniques are not frequently used. Possible reasons are the lack of algorithms and the practical difficulties associated with acquiring, storing and analyzing the necessary, large data sets.

Nevertheless, in for example biomedical applications, such as reported in Kearney *et al.* (1991), MacNeil *et al.* (1992) and Yu and Verhaegen (1993), one is prepared to face these practical difficulties in order to obtain models which accurately represent the input/output (i/o) behavior. Other possible applications include the identification of *non-linear* systems operating along a particular trajectory, such as a robot arm executing repetitive maneuvers, and exhibiting fast dynamical changes as well as *periodically* time-varying systems. In the latter class of systems the collection of an ensemble of i/o sequences each exhibiting the same time-varying behavior is intrinsic.

The extension of existing time-invariant identification schemes to the ensemble identification problem first of all requires a proper redefinition of the structured matrices that are processed by the identification schemes in a time-varying context. For the Hankel matrix such a generalization is given for example in Shokoohi and Silverman (1987) and Dewilde and Van der Veen (1993). In Shokoohi and Silverman (1987) a solution was presented to identify a state space model from a collection of impulse responses, each being the response to an input impulse given at consecutive time instances. Combining this solution with the estimation of the impulse responses from an ensemble of i/o data supplies a possible solution to the identification of a state space model from an ensemble of i/o data. The latter problem is addressed in this paper and is referred to as the *ensemble state space model identification problem*.

In Kearney *et al.* (1991), a numerical scheme was proposed to identify a time-varying Finite Impulse Response (FIR) model from an ensemble of i/o data. However, the estimation of a FIR model followed by the calculation of a state space realization will produce inaccurate results when the underlying system is marginally stable. This was, for example indicated in Verhaegen and Dewilde (1992b) and Van Overschee and De Moor (1994), for the time-invariant case. A better alternative is to use Subspace Model Identification (SMI) techniques (Verhaegen and Dewilde, 1992a). Therefore, we extend in this paper a particular variant of recently developed SMI schemes, namely the Multivariable Output-Error State sPace (MOESP) class of schemes (Verhaegen and Dewilde, 1992a; Verhaegen, 1993a–c), to the ensemble state space model identification problem.

The reasons for extending this particular approach are twofold. First, the MOESP approach allows the same classes of time-invariant identification problems to be addressed as related SMI approaches, such as Moonen *et al.* (1989), Larimore (1990) and Van Overschee and De Moor (1994). Therefore, extending the MOESP approach in no way restricts the range of problems that can be tackled. Second, the close relationship between the MOESP approach and the related approaches in Larimore (1990) and Van Overschee and De Moor (1994) would enable the extension of these other approaches without much difficulty when following the strategy for the MOESP approach outlined in this paper.

The present paper elaborates on the work presented in Verhaegen (1991). There it was shown that ensemble identification problems are a natural way of formulating identification problems in the non-stationary (operator-theoretic) system theory that is currently reaching a higher degree of maturity, see, for example Dewilde *et al.* (1993). In Verhaegen (1991) we continued using the notation established in the pioneering papers (Dewilde and Dym, 1992; Dewilde and Van der Veen, 1993). In this notation the state, input and output vectors are row vectors instead of column vectors. Also, their dimension is allowed to vary. Here we continue to use this general set-up for the sake of consistency.

The variation of the state dimension has been observed in an application of the algorithms applied to the identification of a biomedical system (see Yu and Verhaegen, 1993). The variation of the dimensions of the input and output quantities occurs easily in the treatment of periodically time-varying systems, such as multirate sampled data systems. Special attention is given to the latter class of systems.

The organization of this paper is as follows: Section 2 summarizes some general notation used throughout the paper and quickly reviews some properties of random variables in a time-varying set-up. Section 3 presents the ensemble state space model identification problem when there are no errors in the i/o data and contains definitions such as the persistency of excitation of the input, relevant to solving the ensemble identification problem. The solution to the ensemble identification problem when the output is disturbed by zero mean errors of white coloring or arbitrary coloring is presented in

Section 4. In Section 5 we specialize this solution to the identification of periodically time-varying systems. We also show some of the capabilities of the derived algorithms by reporting their application in simulation to a multirate sampled data system. Finally, Section 6 contains some concluding remarks.

2. NOTATIONS AND STATISTICAL FRAMEWORK OF ANALYSIS

In this section we list some frequently used notation and briefly review the statistical concepts used throughout the paper.

(1) $\rho(A)$, rank of matrix A.
(2) I_n denotes the identity matrix of order n.
(3) The matrix inequality $A(\geq) > B$ means that $A - B$ is (semi-)positive definite.
(4) $\mathbb{R}^n$ denotes the n-dimensional vector space over the field of real numbers and $\mathbb{R}^{n\times m}$ denotes the space of $n \times n$ matrices with entries in $\mathbb{R}$. $\mathbb{Z}$ denotes the set of integers and $\mathbb{Z}^+$ denotes the set of positive integers.
(5) Let $A \in \mathbb{R}^{n\times m}$, then A^{T} denotes the transpose of the matrix A.
(6) The QR factorization: the QR factorization of a matrix $A \in \mathbb{R}^{n\times m}$ is a factorization of this matrix into an orthonormal column matrix $Q \in \mathbb{R}^{n\times n}$, satisfying $Q^{\mathrm{T}}Q = I_n$, and an upper trapezoidal matrix $R \in \mathbb{R}^{n\times m}$ such that:

$$A = QR.$$

Sometimes, when $n \geq m$, we consider a partial QR factorization in which only the first m columns of the orthonormal matrix Q and the first m rows of the matrix R are retained.
(7) The vectors considered in this paper are all row vectors.
(8) Matrix partitioning; standard Matlab (Moler *et al.*, 1987) notation is used to denote the partitioning of a matrix or vector.
(9) The statistical framework of analysis.

For a discrete stochastic process, for example, v_t, $v_{j,t}$ presents the observation in the jth experiment at time instant t. The ensemble of v_t is a family of time histories $v_{j,t}$ for $j \in [1, n]$, $t \in [t_0, t_0 + T - 1]$. For a specific time t, v_t is a random variable (RV) (Anderson, 1966; Papoulis, 1991), and it is assumed that the ensemble sample average and the ensemble sample covariance are asymptotically unbiased estimates of the true mean and covariance of these RVs. For example, for the mean this gives the following equality:

$$E[v_t] = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} v_{j,t} \quad \text{or}$$

$$E[v_t] = \frac{1}{n} \sum_{j=1}^{n} v_{j,t} + O_n(\varepsilon),$$

where $O_n(\varepsilon)$ is a matrix quantity of norm ε which vanishes as $n \to \infty$.

3. PRELIMINARIES, PROBLEM STATEMENT AND DEFINITIONS

3.1. *Model description and the ensemble identification problem*

Here we consider the deterministic part of the dynamic relationships between input and output quantities of the system to be identified to be given by the following multi-input, multi-output (MIMO) discrete linear time-varying state space model:

$$x_{t+1} = x_t A_t + u_t B_t, \tag{1}$$

$$y_t = x_t C_t + u_t D_t, \tag{2}$$

where $u_t \in \mathbb{R}^{m_t}$, $y_t \in \mathbb{R}^{l_t}$, $x_t \in \mathbb{R}^{N_t}$ and the system matrices are of the following sizes:

$$\begin{matrix} \begin{bmatrix} A_t & C_t \\ B_t & D_t \end{bmatrix} & \begin{matrix} N_t \\ m_t \end{matrix} \\ \begin{matrix} N_{t+1} & l_t \end{matrix} & \end{matrix}.$$

The solution to equation (1) is given by:

$$x_t = x_0 \Phi(t, t_0) + \sum_{i=t_0}^{t-1} u_i B_i \Phi(t, i-1), \tag{3}$$

where $\Phi(t, t_0)$ is the transition matrix, satisfying:

$$\Phi(t, t_0) \triangleq \begin{cases} A_{t_0} A_{t_0+1} \cdots A_{t-1} & t > t_0 \\ I & t = t_0 \\ \text{undefined} & t < t_0 \end{cases}. \tag{4}$$

Definition 1. The system state space quadruple $(\alpha_t, \beta_t, \gamma_t, \delta_t)$ is said to be *similarly equivalent* to (A_t, B_t, C_t, D_t) if a transformation $T_t \in \mathbb{R}^{N_t \times N_t}$ exists such that T_t and T_t^{-1} are bounded for all $t \in \mathbb{Z}$ and $(\alpha_t, \beta_t, \gamma_t, \delta_t)$ such that:

$$\begin{bmatrix} \alpha_t & \gamma_t \\ \beta_t & \delta_t \end{bmatrix} = \begin{bmatrix} T_t^{-1} & 0 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} A_t & C_t \\ B_t & D_t \end{bmatrix} \begin{bmatrix} T_{t+1} & 0 \\ 0 & I_l \end{bmatrix}. \tag{5}$$

□

If the representation (A_t, B_t, C_t, D_t) is bounded ($A_t(.), B_t(.), C_t(.)$ is bounded) and stable ($\Phi(t, t_0) \to 0$ as $t \to \infty$), then the transformation T_t preserves the boundedness of the transformed system by equation (5), and the stability by the following equation (Shokoohi and Silverman, 1987):

$$\Phi'(t, t_0) = T_{t_0}^{-1} \Phi(t, t_0) T_t, \tag{6}$$

where $\Phi'(t, t_0) = \alpha_{t_0} \alpha_{t_0+1} \cdots \alpha_{t-1}$.

The deterministic ensemble (state space model) identification problem: let the index set j, t in $u_{j,t}$ indicate the input sample at time t of the jth experiment conducted with the system (2) and (3). If thus, $j \in [j_0, j_0 + n - 1]$ and $t \in [t_0, t_0 + T - 1]$ (where j_0 indicates the first experiment, t_0 indicates the first time instant, n indicates the total number of experiments and T indicates the total measuring time a single experiment lasts), the problem is to determine a state space description:

$$\xi_{j,t+1} = \xi_{j,t}\alpha_t + u_{j,t}\beta_t, \tag{7}$$

$$y_{j,t} = \xi_{j,t}\gamma_t + u_{j,t}\delta_t, \tag{8}$$

such that $(\alpha_t, \beta_t, \gamma_t, \delta_t)$ is similarly equivalent to the state space representation (A_t, B_t, C_t, D_t) based on the following ensemble of output data sequences:

$$\begin{pmatrix} y_{j_0,t_0} & y_{j_0,t_0+1} & \cdots & y_{j_0,t_0+T-1} \\ y_{j_0+1,t_0} & y_{j_0+1,t_0+1} & \cdots & y_{j_0+1,t_0+T-1} \\ \vdots & \vdots & \cdots & \vdots \\ y_{j_0+n-1,t_0} & y_{j_0+n-1,t_0+1} & \cdots & y_{j_0+n-1,t_0+T-1} \end{pmatrix} \tag{9}$$

and input sequences $u_{j,t}$ for the same series of experiments and over the same time interval.

From the state space representation (2) and (3), it is easy to find the following relationship:

$$\begin{pmatrix} y_{j_0,t} & y_{j_0,t+1} & \cdots & y_{j_0,t+s-1} \\ y_{j_0+1,t} & y_{j_0+1,t+1} & \cdots & y_{j_0+1,t+s-1} \\ \vdots & \vdots & \cdots & \vdots \\ y_{j_0+n-1,t} & y_{j_0+n-1,t+1} & \cdots & y_{j_0+n-1,t+s-1} \end{pmatrix}$$

$$= \begin{pmatrix} x_{j_0,t} \\ x_{j_0+1,t} \\ \vdots \\ x_{j_0+n-1,t} \end{pmatrix} \times \begin{bmatrix} C_t & A_tC_{t+1} & \cdots & A_tA_{t+1}\cdots A_{t+s-2}C_{t+s-1} \end{bmatrix}$$

$$+ \begin{pmatrix} u_{j_0,t} & u_{j_0,t+1} & \cdots & u_{j_0,t+s-1} \\ u_{j_0+1,t} & u_{j_0+1,t+1} & \cdots & u_{j_0+1,t+s-1} \\ \vdots & \vdots & \cdots & \vdots \\ u_{j_0+n-1,t} & u_{j_0+n-1,t+1} & \cdots & u_{j_0+n-1,t+s-1} \end{pmatrix}$$

$$\times \begin{pmatrix} D_t & B_tC_{t+1} & \cdots & B_tA_{t+1}\cdots A_{t+s-2}C_{t+s-1} \\ & D_{t+1} & \cdots & \vdots \\ & & \ddots & \\ & & & D_{t+s-1} \end{pmatrix}. \tag{10}$$

This equation is denoted more compactly as:

$$Y_{t,s} = X_t\Theta_{t,s} + U_{t,s}\Delta_{t,s}. \tag{11}$$

Following the work of Shokoohi and Silverman (1987) and Dewilde and Van der Veen (1993), the matrices $Y_{t,s}$ and $U_{t,s}$ are referred to as generalized Hankel matrices.

The index set of these generalized Hankel matrices should be j_0, t, s, n. This set would allow a precise definition of which part of the recorded output sequences in equation (9) and the corresponding input sequences are stored in both matrices. However, since the experiment index will not be relevant in the solution presented to the ensemble identification problem and since we assume the number of experiments n to be fixed by the experimental circumstances, we restrict the index set of both matrices to t and s. Here t indicates the time index of their top left entry and s determines their column width in the following way. Let

$$L_t^s = \sum_{\tau=t}^{t+s-1} l_\tau \quad \text{and} \quad M_t^s = \sum_{\tau=t}^{t+s-1} m_\tau, \tag{12}$$

then $Y_{t,s} \in \mathbb{R}^{n \times L_t^s}$ and $U_{t,s} \in \mathbb{R}^{n \times M_t^s}$ and as a consequence $X_t \in \mathbb{R}^{n \times N_t}$, $\Theta_{t,s} \in \mathbb{R}^{N_t \times L_t^s}$ and $\Delta_{t,s} \in \mathbb{R}^{M_t^s \times L_t^s}$.

Based on the data representation, the solution to the deterministic ensemble identification problem will be given by subsequently treating the following subproblems:

(1) given the data matrices $Y_{t,s}$ and $U_{t,s}$, determine the condition on the input sequences to retrieve the row space of the observability matrix $\Theta_{t,s}$ and determine this row space;
(2) determine the matrices $[\alpha_t, \gamma_t]$ from the row space of $\Theta_{t,s}$;
(3) determine the matrices $[\beta_t, \delta_t]$; and
(4) determine the minimum μ and χ such that $[\alpha_t, \beta_t, \gamma_t, \delta_t]$ can be calculated in the time interval $t \in [t_0 + \mu, t_0 + T - 1 - \chi]$.

The division of the overall identification problem into the above subproblems is inspired by that used in tackling subspace identification problems for LTI (linear time-invariant) systems within the MOESP class of algorithms (see, e.g. Verhaegen and Dewilde, 1992a; Verhaegen, 1993b). Here we remark that Subproblem 4 is an additional problem of concern only in the LTV case.

3.2. *Basic lemmas and important definitions*

In this paper, we will use the following lemma:

Lemma 1. Sylvester's inequality (Kailath, 1980, p. 655). Let $M_1 \in \mathbb{R}^{m \times n}$ and $M_2 \in \mathbb{R}^{n \times p}$, then:

$$\rho(M_1) + \rho(M_2) - n \leq \rho(M_1M_2) \leq \min\{\rho(M_1), \rho(M_2)\}. \tag{13}$$

Furthermore, the following series of definitions are demonstrated to be relevant in solving the ensemble identification problem. The first two

make use of the observability and controllability Gramian, denoted, respectively, by $G_o(t, t+\delta)$ and $G_c(t-\delta, t)$. Let the extended controllability matrix $\Psi_{t,\delta}^T$ be equal to

$$[B_{t-1}^T \quad A_{t-1}^T B_{t-1}^T \quad \cdots \quad A_{t-1}^T A_{t-2}^T \cdots A_{t-\delta+1}^T B_{t-\delta}^T],$$

the Gramians are given as:

$$G_o(t, t+\delta) = \Theta_{t,\delta}\Theta_{t,\delta}^T \quad G_c(t-\delta, t) = \Psi_{t,\delta}^T \Psi_{t,\delta}$$

Definition 2. (Jazwinski, 1970). The pair $[A_t, C_t]$ is uniformly observable if $\exists \delta \in \mathbb{Z}^+$, and positive constants b_1, b_2 such that:

$$0 < b_1 I \leq G_o(t, t+\delta) \leq b_2 I, \quad \forall t. \tag{14}$$

The least integer value δ, for which condition (14) holds is denoted by δ_o. For linear time-invariant systems, δ_o is referred to as the observability index (Kailath, 1980, p. 356). With the δ defined in Definition 2, it follows immediately via Lemma 1, that $\rho(\Theta_{t,\delta}) = N_t$ (and this implies that $L_t^\delta \geq N_t$). The dual of definition 2 is:

Definition 3. (Jazwinski, 1970). The pair $[A_t, B_t]$ is uniformly controllable if $\exists \delta \in \mathbb{Z}^+$, and positive constant a_1, a_2 such that:

$$0 < a_1 I \leq G_c(t-\delta, t) \leq a_2 I, \quad \forall t. \tag{15}$$

The least integer value of δ for which the condition (15) holds is denoted by δ_c and is called the controllability index when the system is time-invariant (Kailath, 1980). When the conditions in Definition 3 hold, Lemma 1 shows that $\rho(\Psi_{t,\delta}) = N_t$ and $M_t^\delta \geq N_t$.

Definition 4. A bounded realization (A_t, B_t, C_t, D_t) is said to be uniform if it is uniformly controllable and uniformly observable (Shokoohi and Silverman, 1987). Based on this definition we have the following lemma:

Lemma 2. Let (A_t, B_t, C_t, D_t) be a uniform realization, then its similarly equivalent realization $(\alpha_t, \beta_t, \gamma_t, \delta_t)$ is also a uniform realization.

Proof. The proof follows from the fact that the controllability and observability Gramians of both realizations are congruent (Strang, 1980). □

Definition 5. A system representation (A_t, B_t, C_t, D_t) is said to be uniformly balanced if the following two conditions hold: (i) (A_t, B_t, C_t) is uniform; and (ii) $\forall t: G_o(t, t+s) = G_c(t-s, t)$ and equal to a diagonal matrix with positive entries.

In Shokoohi and Silverman (1987), an algorithm is presented to compute a uniformly balanced realization for any given bounded uniform realization. The algorithm essentially proceeds in two steps. In the first step, a similarity transformation is performed such that the observability Gramian of the similarly equivalent realization is the identity matrix. In the second step, a similarity transformation is performed such that the controllability Gramian becomes diagonal and equal to the observability Gramian.

Our final definition, which is illustrated by an example in Appendix 1, is stated next.

Definition 6. The input sequences in the data matrix $U_{t,s}$ are locally persistent excitations at the time instant t if:

$$\rho([U_{t,s} \mid X_t]) = M_t^s + N_t.$$

The above definition is a generalization of a standard assumption made for LTI systems in the context of ensemble identification. Generally, as in the LTI case, and as shown in Yu and Verhaegen (1993) for the ensemble identification case, this assumption is fairly easy to satisfy when making use of different realizations of a stochastic process obtained by filtering a stationary white noise sequence through an LTI system.

4. A SUBSPACE MODEL IDENTIFICATION SOLUTION TO THE ENSEMBLE IDENTIFICATION PROBLEM

4.1. *The deterministic ensemble identification problem*

When there is no noise on the i/o data, the matrices $Y_{t,s}$ and $U_{t,s}$ are related as in equation (11). Based on this relationship, we address the subproblems described at the end of Section 3.1 in the following subsections.

4.1.1. *The determination of the row space of* $\Theta_{t,s}$. Let us denote the QR factorization of the compound matrix $[U_{t,s} \mid Y_{t,s}]$ as:

$$[U_{t,s} \mid Y_{t,s}] = [Q_{1,t} \mid Q_{2,t}] \begin{pmatrix} R_{11,t} & R_{12,t} \\ 0 & R_{22,t} \end{pmatrix} \begin{matrix} M_t^s \\ n - M_t^s \end{matrix}, \tag{16}$$
$$\qquad\qquad M_t^s \qquad L_t^s$$

where $Q_{1,t}^T Q_{1,t} = I_{M_t^s}$ and $Q_{1,t}^T Q_{2,t} = 0$, then we have the following theorem:

Theorem 1. Let the system be uniformly controllable and uniformly observable. In addition let the following conditions be satisfied:

(1) $s \geq \delta_0$;
(2) the input sequences in the data matrix $U_{t,s}$

are locally persistent excitations at time instant t;

(3) $n \geq (M_t^s + N_t)$;

(4) the QR factorization of matrix $[U_{t,s} \mid Y_{t,s}]$ is given and partitioned as in equation (16);

then $\rho(R_{22,t}) = \rho(\Theta_{t,s}) = N_t$, and the row space of $R_{22,t}$ equals the row space of the observability matrix $\Theta_{t,s}$.

Proof. Denote the QR factorization of the matrix pair $[U_{t,s} \mid X_t]$ as:

$$[U_{t,s} \mid X_t] = [Q_{1,t} \mid Q_{x,t}] \begin{pmatrix} R_{11,t} & R_{x1,t} \\ 0 & R_{x2,t} \end{pmatrix} \begin{matrix} M_t^s \\ N_t \end{matrix}, \quad \begin{matrix} M_t^s & N_t \end{matrix} \tag{17}$$

where $Q_{1,t}^{\mathrm{T}} Q_{x,t} = 0$. Since $n \geq (M_t^s + N_t)$, the columns of the matrix $[Q_{1,t} \mid Q_{x,t}]$ are orthogonal, condition (2) shows with the help of Lemma 1 that

$$\rho\left(\begin{bmatrix} R_{11,t} & R_{x1,t} \\ 0 & R_{x2,t} \end{bmatrix}\right) = \rho([U_{t,s} \mid X_t]) = M_t^s + N_t \quad \text{and} \quad \rho(R_{x2,t}) = N_t. \tag{18}$$

With the QR factorization in equation (16) we can express $Y_{t,s}$ as:

$$Y_{t,s} = Q_{1,t} R_{12,t} + Q_{2,t} R_{22,t}$$

and equation (11) as:

$$Y_{t,s} = X_t \Theta_{t,s} + Q_{1,t} R_{11,t} \Delta_{t,s}.$$

Using the expression for X_t as given in equation (17), the latter equation can also be denoted as:

$$Y_{t,s} = Q_{1,t} R_{x1,t} \Theta_{t,s} + Q_{x,t} R_{x2,t} \Theta_{t,s} + Q_{1,t} R_{11,t} \Delta_{t,s}.$$

Hence we have:

$$Q_{1,t} R_{12,t} + Q_{2,t} R_{22,t} = Q_{1,t} R_{x1,t} \Theta_{t,s} + Q_{x,t} R_{x2,t} \Theta_{t,s} + Q_{1,t} R_{11,t} \Delta_{t,s}. \tag{19}$$

Multiplying the left-hand side of equation (19) with $Q_{1,t}^{\mathrm{T}}$ and using the properties that $Q_{1,t}^{\mathrm{T}} Q_{1,t} = I_{M_t^s}$, $Q_{1,t}^{\mathrm{T}} Q_{2,t} = Q_{1,t}^{\mathrm{T}} Q_{x,t} = 0$ we obtain:

$$R_{12,t} = R_{x1,t} \Theta_{t,s} + R_{11,t} \Delta_{t,s}. \tag{20}$$

Substituting this relationship back into the right-hand side of equation (19) yields:

$$Q_{2,t} R_{22,t} = Q_{x,t} R_{x2,t} \Theta_{t,s}. \tag{21}$$

The right-hand side of equation (18), and the fact that $\rho(Q_{x,t}) = N_t$, $\rho(\Theta_{t,s}) = N_t$ (since $s \geq \delta_0$) show by Lemma 1 that

$$\rho(R_{22,t}) = \rho(\Theta_{t,s}) = N_t.$$

Furthermore, since $Q_{2,t}^{\mathrm{T}} Q_{x,t} R_{x2,t} \in R^{n - M_t^s \times N_t}$ and $\Theta_{t,s} \in \mathbb{R}^{N_t \times L_t^s}$ and since $n - M_t^s \geq N_t$ and $L_t^s \geq N_t$ by condition (2), the row space of $R_{22,t}$ equals that of $\Theta_{t,s}$. □

It should be remarked that when condition (3) has to hold for $\forall t$, the minimal number of experiments, denoted by $n_{\min}$, required is:

$$n_{\min} = \max_{t \in [t_0, t_0 + T - 1]} (M_t^s + N_t). \tag{22}$$

Therefore, we see that $n_{\min}$ depends on an upperbound of the order of the system, provided that s has been chosen such that conditions (1) and (2) of Theorem 1 are satisfied. When s is chosen of the same order of magnitude as the order of the system, as can be done in the time-invariant case, then $n_{\min}$ turns out to be of the same order in magnitude as s. However, as we see in Section 4.2, the presence of errors on the output measurements would require the number n to be very large (∞) when (consistent) estimates of high accuracy are required.

As a result of Theorem 1, we can find the row space of $\Theta_{t,s}$ through a SVD of $R_{22,t}$. Denote this decomposition as:

$$R_{22,t} = U_t \begin{pmatrix} S_{N_t} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_{N_t}^{\mathrm{T}} \\ (V_{N_t}^{\perp})^{\mathrm{T}} \end{pmatrix} \tag{23}$$

with $U_t \in \mathbb{R}^{(n - M_t^s) \times (n - M_t^s)}$, $S_{N_t} \in \mathbb{R}^{N_t \times N_t}$, $V_{N_t} \in \mathbb{R}^{L_t^s \times N_t}$ and $V_{N_t}^{\perp} \in \mathbb{R}^{L_t^s \times (L_t^s - N_t)}$. Let $T_t \in \mathbb{R}^{N_t \times N_t}$ be a square invertible matrix, then we denote the result of Theorem 1 as:

$$V_{N_t}^{\mathrm{T}} = T_t^{-1} \Theta_{t,s}. \tag{24}$$

4.1.2. *The determination of the matrices* $[\alpha_t, \gamma_t]$. From equation (24) and (5) we get:

$$V_{N_t}^{\mathrm{T}} = [\gamma_t \quad \alpha_t \gamma_{t+1} \quad \cdots \quad \alpha_t \cdots \alpha_{t+s-2} \gamma_{t+s-1}]. \tag{25}$$

Hence,

$$\gamma_t = V_{N_t}^{\mathrm{T}}(:, 1 : l_t) \tag{26}$$

and α_t can be computed as shown in the following lemma.

Lemma 3. Let the conditions in Theorem 1 hold, with condition (2) strengthened to $s > \delta_0$ and condition (3) replaced by $n \geq \max_{t \in [t_0, t_0 + T - 1]} (M_t^s + N_t)$ then α_t can be solved from the overdetermined set of equations as follows:

$$\begin{aligned} V_{N_t}^{\mathrm{T}}(:, l_t + 1 : L_t^s) &= \alpha_t [\gamma_{t+1} \quad \alpha_{t+1} \gamma_{t+2} \quad \cdots \\ &\qquad \alpha_{t+1} \cdots \alpha_{t+s-2} \gamma_{t+s-1}] \\ &= \alpha_t V_{N_{t+1}}^{\mathrm{T}}(:, 1 : L_{t+1}^{s+1}). \end{aligned} \tag{27}$$

Proof. By Theorem 1, we can determine the row space of $\Theta_{t+1,s}$ via a SVD of the matrix $R_{22,t+1}$. Denote this row space by V_{N+1}^{T}, then there exists a non-singular matrix T_{t+1} such that:

$$V_{N_{t+1}}^{\mathrm{T}} = T_{t+1}^{-1} \Theta_{t+1,s}.$$

Since $s > \delta_o$, $\rho(\Theta_{t+1,s-1}) = N_{t+1}$ the matrix $V^{\mathrm{T}}_{N_{t+1}}(:, 1:L^{s-1}_{t+1})$ has full row rank and we can solve equation (27) for α_t. □

From the above relationships we can obtain γ_t in the time interval $[t_0, t_0 + T - s]$ and α_t for $t \in [t_0, t_0 + T - 1 - s]$.

4.1.3. *The determination of the matrices* $[\beta_t, \delta_t]$.

Theorem 2. Let the system be uniformly controllable and observable. In addition, let the following conditions be satisfied:

(1) $s > \delta_o$;
(2) the input sequences in the series of data matrices $U_{t+1-i,s}$ for $i = 0:1:s$ are locally persistent excitations;
(3) $n \geq \max_{t \in [t_0, t_0+T-1]} (M^s_t + N_t)$;
(4) the QR factorization of matrix sequence $[U_{t+1-i,s} \mid Y_{t+1-i,s}]$ for $i = 0:s$ be given and partitioned as in equation (16);
(5) $\rho(V^{\perp}_{N_{t-i}}(L^{s-1}_{t-i} + 1 : L^{s}_{t-i}, :)) = l_{t-i+s-1}$ for $i = 0:1:s-1$;

then, β_t and δ_t can be solved from the equation:

$$(\delta_t \mid \beta_t)\begin{pmatrix} I_{l_t} & 0 \\ 0 & V^{\mathrm{T}}_{N_{t+1}}(:, 1:L^{s-1}_{t+1}) \end{pmatrix} \times \begin{bmatrix} V^{\perp}_{N_t} & \dfrac{V^{\perp}_{N_{t-1}}(L^1_{t-1} + 1 : L^s_{t-1}, :)}{0} \\ \cdots & \begin{array}{c} V^{\perp}_{N_{t-s+1}}(L^{s-1}_{t-s+1} + 1 : L^s_{t-s+1}) \\ \hline 0 \\ \hline \vdots \\ \hline 0 \end{array} \end{bmatrix}$$
$$= [\phi_t(1:m_t, :) \mid \phi_{t-1}(M^1_{t-1} + 1 : M^1_{t-1} + m_t, :) \mid \cdots \mid \phi_{t-s+1}(M^{s-1}_{t-s+1} + 1 : M^{s-1}_{t-s+1} + m_t, :)], \quad (28)$$

where $\phi_t = R^{-1}_{11,t} R_{12,t} V^{\perp}_{N_t}$.

Proof. Using Theorem 1, $\Theta_{t-i,s} V^{\perp}_{N_{t-i}} = 0$ for $i = 0:s-1$. As a consequence, multiplying equation (20) by $V^{\perp}_{N_t}$ reduces this equation to:

$$R_{12,t} V^{\perp}_{N_t} = R_{11,t} \Delta_{t,s} V^{\perp}_{N_t} \quad (29)$$

Denote the matrix product $R^{-1}_{11,t} R_{12,t} V^{\perp}_{N_t}$ as ϕ_t,

$$\phi_t = R^{-1}_{11,t} R_{12,t} V^{\perp}_{N_t} = \Delta_{t,s} V^{\perp}_{N_t}$$

then equation (29) is expressed explicitly as:

$$\begin{bmatrix} \phi_t(1:m_t, :) \\ \hline \phi_t(m_t + 1 : M^2_t, :) \\ \hline \vdots \\ \hline \phi_t(M^{s-1}_t + 1 : M^s_t, :) \end{bmatrix} = \begin{bmatrix} (\delta_t \mid \beta_t)\begin{pmatrix} I_{l_t} & 0 \\ 0 & V^{\mathrm{T}}_{N_{t+1}}(:, 1:L^{s-1}_{t+1}) \end{pmatrix} \\ \hline (\delta_{t+1} \mid \beta_{t+1})\begin{pmatrix} 0 & I_{l_{t+1}} & 0 \\ 0 & 0 & V^{\mathrm{T}}_{N_{t+2}}(:, 1:L^{s-2}_{t+2}) \end{pmatrix} \\ \hline \vdots \\ \hline (\delta_{t+s-1} \mid \beta_{t+s-1})\begin{pmatrix} 0 & \cdots & I_{l_{t+s-1}} \\ 0 & \cdots & 0 \end{pmatrix} \end{bmatrix} \times V^{\perp}_{N_t} \quad (30)$$

The matrices δ_t and β_t thus satisfy equation (28).

Since by Lemma 3, the matrix:

$$\begin{pmatrix} I_{l_t} & 0 \\ 0 & V^{\mathrm{T}}_{N_{t+1}}(:, 1:L^{s-1}_{t+1}) \end{pmatrix}$$

has full row rank and since condition (5) guarantees that the first matrix to the left of the equality in equation (28) has full row rank, Lemma 1 again demonstrates that we can solve δ_t and β_t from equation (28). □

The above analysis shows that when the conditions stipulated in Lemma 3 and Theorem 2 are satisfied over the time interval $[t_0, t_0 + T - 1]$, we can get δ_t and β_t for $t \in [t_0 + s - 1, t_0 + T - s]$. Hence, under these conditions, the calculation of the quadruple of system matrices $\begin{bmatrix} \alpha_t & \gamma_t \\ \beta_t & \delta_t \end{bmatrix}$ is possible in the time interval $t \in [t_0 + \mu, t_0 + T - 1 - \chi]$, with minimal values for μ and χ, respectively, $s - 1$ and s.

To conclude this part, we summarize the above results into a generalization of the ordinary MOESP algorithm (Verhaegen and Dewilde, 1992a) applicable to ensemble identification problems.

The ordinary MOESP algorithm for ensemble identification problems.

Given:

(1) a uniformly controllable and uniformly observable time-varying system;
(2) $s > \delta_0$, with δ_0 defined in the paragraph following Definition 2;
(3) an ensemble of input sequences such that input sequences in the series of data matrices

$U_{t,s}$ are locally persistent excitations for $t \in [t_0, t_0 + T - s - 1]$;

(4) $n \geq \max_{t \in [t_0, t_0+T-1]} (M_t^s + N_t)$, with n the number of experiments, M_t^s defined in equation (12) and N_t the order of the system at time instant t.

For $t = t_0 + s - 1 : t_0 + T - s - 1$ carryout the following:

Step 1: construct the sequence of generalized Hankel matrices $U_{t+1-i,s}$, $Y_{t+1-i,s}$ for $i = 0:1:s$;

Step 2: perform a QR factorization of the sequence of compound matrices $[U_{t+1-i,s} \mid Y_{t+1-i,s}]$ for $i = 0:1:s$, without storing Q, and partition the R factor as in equation (16);

Step 3: perform a SVD of the sequence of matrices $R_{22,t+1-i}$ for $i = 0:1:s$ as given in equation (23) and store the matrix sequences $V_{N_{t+1-i}}$, $V_{N_{t+1-i}}^{\perp}$ for $i = 0:1:s$; and

Step 4: solve equations (26) and (27) to obtain γ_t, α_t and when condition (5) of Theorem 2 is satisfied, then solve equation (28) to determine β_t and δ_t.

For the special case where the input and output dimensions are constants and denoted by m and l, respectively, equations (26)–(28) specialize to:

$$\gamma_t = V_{N_t}^{\mathrm{T}}(:, 1:l), \tag{31}$$

$$V_{N_t}^{\mathrm{T}}(:, l+1:sl) = \alpha_t V_{N_{t+1}}^{\mathrm{T}}(:, 1:(s-1)l), \tag{32}$$

$$(\delta_t \mid \beta_t)\left(\begin{array}{c|c} I_l & 0 \\ \hline 0 & V_{N_{t+1}}^{\mathrm{T}}(:, 1:(s-1)l) \end{array}\right) \times \left[\begin{array}{c|c|c|c} V_{N_t}^{\perp} & \dfrac{V_{N_{t+1}}^{\perp}(l+1:sl, :)}{0} & \cdots & \dfrac{V_{N_{t-s+1}}^{\perp}((s-1)l+1:sl, :)}{\begin{array}{c}0\\ \hline \vdots \\ \hline 0\end{array}} \end{array}\right] = [\phi_t(1:m, :) \mid \phi_{t-1}(m+1:2m, :) \mid \cdots \mid \phi_{t-s+1}((s-1)m+1:sm, :)], \tag{33}$$

respectively.

4.1.4. *A property of the realization calculated by the ordinary MOESP algorithm for ensemble identification problems.* If the system realization $[A_t, B_t, C_t, D_t]$ is a uniform realization, then $[\alpha_t, \gamma_t, \beta_t, \delta_t]$ is also a uniform realization (Lemma 2). Moreover, since the resulting observability Gramian equals the identity matrix, the first step of the algorithm presented in Shokoohi and Silverman (1987) to calculate a uniformly balanced realization of a time-variant system can be skipped when such a realization is needed.

4.2. *Additive errors in the output of the system to be identified*

In reality, the input–output data sequences are collected by measurement. As for the time-invariant case (Verhaegen and Dewilde, 1992a) we only consider additive errors on the output measurement. In this case, the system model (2) and (3) changes into:

$$x_{j,t+1} = x_{j,t}A_t + u_{j,t}B_t, \tag{34}$$

$$z_{j,t} = x_{j,t}C_t + u_{j,t}D_t + v_{j,t}, \tag{35}$$

where $z_{j,t}$ is the measured output and $v_{j,t}$ represents the noise. Let $V_{t,s}$ denote the generalized Hankel matrix constructed from the samples $v_{j,t}$ as $Y_{t,s}$ has been constructed from $y_{j,t}$, then the relevant relationship between input and output sequences, as given by (11) for the noise free case, changes into:

$$Z_{t,s} = X_t\Theta_{t,s} + U_{t,s}\Delta_{t,s} + V_{t,s}. \tag{36}$$

For these types of errors, we can state a generalizatoin of Theorem 1 of Verhaegen (1993c) within the present ensemble identification context.

Theorem 3. Let the system be uniformly controllable and observable. In addition let the following conditions be satisfied:

(1) $s > \delta_0$;

(2) the input sequences in the series of data matrices $U_{t,s}$ are locally persistent excitations, such that

$$\rho\left[\lim_{n\to\infty} \frac{1}{n}\begin{bmatrix} U_{t,s}^{\mathrm{T}} \\ X_t^{\mathrm{T}} \end{bmatrix}[U_{t,s} \quad X_t]\right] = M_t^s + N_t;$$

(3) the QR factorizations of $[U_{t,s} \mid Z_{t,s}]$ and $[U_{t,s} \mid X_t]$ be given as in equation (16) with $Y_{t,s}$ replaced by $Z_{t,s}$ and equation (17), respectively;

(4) the noise $v_{j,t}$ is independent of the input and has mean zero, hence the following limits hold:

$$\lim_{n\to\infty} \frac{1}{n} V_{t_1,s}^{\mathrm{T}} U_{t_2,r} = 0, \quad \forall t_1, t_2, \quad \text{and} \quad \forall r > 0, \quad \forall s > 0. \tag{37}$$

$$\lim_{n\to\infty} \frac{1}{n} V_{t,s}^{\mathrm{T}} V_{t,s} = R_v, \tag{38}$$

$$\lim_{n\to\infty} \frac{1}{n} V_{t,s}^{\mathrm{T}} X_t = 0, \tag{39}$$

$$\lim_{n\to\infty} \frac{1}{n} R_{x2}^{\mathrm{T}} R_{x2} = P_{x2}, \tag{40}$$

then

$$\lim_{n\to\infty}\frac{1}{n}R_{22,t}^{\mathrm{T}}R_{22,t}=\Theta_{t,s}^{\mathrm{T}}P_{x2}\Theta_{t,s}+R_v. \quad (41)$$

Proof. Taking into account the treatment of the statistical quantities as outlined in Section 2, the proof of this theorem can be given along the lines set up when proving Theorem 1 of Verhaegen (1993c). From the data equation (36) and the QR factorization defined in Item 3 of the theorem, we obtain the following relationships:

$$U_{t,s}=Q_{1,t}R_{11,t}, \quad (42)$$

$$Z_{t,s}=Q_{1,t}R_{12,t}+Q_{2,t}R_{22,t}=X_t\Theta_{t,s}+U_{t,s}\Delta_{t,s}+V_{t,s}, \quad (43)$$

$$X_t=Q_{1,t}R_{x1,t}+Q_{x,t}R_{x2,t}. \quad (44)$$

Substituting equations (42) and (44) into equation (43) yields:

$$\frac{1}{\sqrt{n}}(Q_{1,t}R_{12,t}+Q_{2,t}R_{22,t})$$
$$=\frac{1}{\sqrt{n}}(Q_{1,t}R_{x1,t}\Theta_{t,s}+Q_{x,t}R_{x2,t}\Theta_{t,s}$$
$$+Q_{1,t}R_{11,t}\Delta_{t,s}+V_{t,s}). \quad (45)$$

Equation (37) can be denoted as:

$$\frac{1}{n}V_{t,s}^{\mathrm{T}}U_{t,s}=\frac{1}{n}V_{t,s}^{\mathrm{T}}Q_{1,t}R_{11,t}=O_n(\varepsilon).$$

From Condition 2 of the theorem, it follows that there exists an $\bar{n}$ such that for $n\geq\bar{n}$ the matrix $\frac{1}{\sqrt{n}}R_{11,t}$ is invertible. Therefore, it follows from the prior equation that:

$$\frac{1}{\sqrt{n}}V_{t,s}Q_{1,t}=O_n(\varepsilon). \quad (46)$$

Inserting this result into equation (39), we obtain, since by condition 2 the matrix $\frac{1}{\sqrt{n}}R_{x2,t}$ is also invertible for $n\geq\bar{n}$, that:

$$\frac{1}{\sqrt{n}}V_{t,s}Q_{x,t}=O_n(\varepsilon). \quad (47)$$

Now multiplying equation (45) on the left by $Q_{1,t}^{\mathrm{T}}$ and using the orthogonality between the matrices $Q_{1,t}$, $Q_{2,t}$ and $Q_{x,t}$, and equation (46) yields:

$$\frac{1}{\sqrt{n}}R_{12,t}=\frac{1}{\sqrt{n}}(R_{x1,t}\Theta_{t,s}+R_{11,t}\Delta_{t,s})+O_n(\varepsilon)$$

and hence equation (45) reduces to:

$$\frac{1}{\sqrt{n}}Q_{2,t}R_{22,t}=\frac{1}{\sqrt{n}}(Q_{x,t}R_{x2,t}\Theta_{t,s}+V_{t,s})+O_n(\varepsilon).$$

Multiplying both sides of this relationship on the left by their transpose and using equation (47) yields:

$$\frac{1}{n}R_{22,t}^{\mathrm{T}}R_{22,t}=\frac{1}{n}\Theta_{t,s}^{\mathrm{T}}R_{x2,t}^{\mathrm{T}}\Theta_{t,s}+\frac{1}{n}V_{t,s}^{\mathrm{T}}V_{t,s}+O_n(\varepsilon).$$

Taking the limit $n\leftarrow\infty$ yields equation (41). □

In the following two subsections, we consider two specific types of perturbations $v_{j,t}$.

4.2.1. *Case* 1: *the additive errors are discrete zero-mean white noise errors.* For this case, we can state the following corollary to Theorem 3:

Corollary 1. Let the conditions of Theorem 3 hold, and let the noise v_t be zero-mean white noise, then the ordinary MOESP algorithm to solve ensemble identification problems determines the row space of the matrix $\Theta_{t,s}$ asymptotically unbiasedly.

Proof. Making use of the white noise property of v_t, equation (41) becomes:

$$\lim_{n\to\infty}\frac{1}{n}R_{22,t}^{\mathrm{T}}R_{22,t}=\Theta_{t,s}^{\mathrm{T}}P_{x2}\Theta_{t,s}+\sigma_{v_t}^2I,$$

where $\sigma_{v_t}^2$ is the variance of the *RV* v_t. With the SVD in equation (23), we can denote this relationship as,

$$\frac{1}{n}R_{22,t}^{\mathrm{T}}R_{22,t}$$
$$=(V_{N_t}\,|\,V_{N_t}^{\perp})\left(\begin{array}{c|c}\frac{1}{n}S_{N_t}^2+\sigma_{v_t}^2I & 0\\ \hline 0 & \sigma_{v_t}^2I\end{array}\right)\left(\begin{array}{c}V_{N_t}^{\mathrm{T}}\\ \hline (V_{N_t}^{\perp})^{\mathrm{T}}\end{array}\right)$$
$$+\mathrm{O_n}(\varepsilon),$$

where the column space of V_{N_t} and $\Theta_{t,s}$ coincide. Since the first term in the right-hand side is again a SVD of the matrix $\frac{1}{N}R_{22,t}^{\mathrm{T}}R_{22,t}$, and since $O_n(\varepsilon)$ vanishes for $n\to\infty$, the relationship shows that the assertion made in the corollary holds. □

Based on this corollary, we can show as in the time-invariant case (see Verhaegen, 1993c), that the quadruple of system matrices $\begin{bmatrix}\alpha_t & \gamma_t\\ \beta_t & \delta_t\end{bmatrix}$ can be computed consistently by the ordinary MOESP algorithm for ensemble identification problems. In Section 5, we illustrate this property by means of numerical simulation.

The proof of Corollary 1 shows that asymptotically the order of the system is also correctly displayed by the gap between the 'signal' singular values, those of the square root

of the matrix $\left(\frac{1}{n}S_{N_t}^2 + \sigma_{v_t}^2 I\right)$, and the 'noise' singular values, namely σ_{v_t}. For the finite data length case the detection of such a gap might not be trivial at all. One approach here is to apply the strategy suggested in Verhaegen (1993c) for LTI systems in the present LTV context. The strategy for LTI systems is to use the dependency of the computed singular values, ordered in decreasing order of magnitude, and the residuals, that is the norm of the difference between the measured output and the reconstructed one, on the model order. This dependency is used to detect an N_tth-order system when a 'big gap' between the N_t and (N_t+1)th singular value corresponds to a 'marginal' decrease of the residual for models of order N_t and larger. In the present LTV context, the singular values and the residuals are a function of both the model order and time. Therefore, in this case the LTI strategy should be applied at each time instant. The usefulness of this strategy in a real-life ensemble identification experiment is shown in Yu and Verhaegen (1993). Here we illustrate the strategy in the example treated in Section 5.

4.2.2. *Case* 2: *the additive errors are discrete zero-mean noise of arbitrary coloring.* When the noise is not zero-mean and white, the calculation of the row space of $\Theta_{t,s}$ in the ordinary MOESP algorithm for ensemble identification problems is according to the description above. A possible and well-known rescue in time-invariant system identification is to introduce instrumental variables. Also for ensemble identification problems, the instrumental variable idea can be used. This is demonstrated by extending the PI scheme of Verhaegen (1993c), in which Past Input (PI) data was used as instrumental variables, to the ensemble identification problem.

The key step in this generalization is presented in our final theorem, which is caused by the similarity with its time-invariant counterpart in Verhaegen (1993c) stated without proof.

Theorem 4. Let the conditions (1)–(4) in Theorem 3 hold, and let the following QR factorization be defined:

$$[U_{t,s} \mid Z_{t,s} \mid U_{t-s,s}] = [Q_{1,t} \mid Q_{2,t} \mid Q_{3,t}] \begin{bmatrix} R_{11,t} & R_{12,t} & R_{13,t} \\ 0 & R_{22,t} & R_{23,t} \\ 0 & 0 & R_{33,t} \end{bmatrix} \begin{matrix} M_t^s \\ L_t^s \\ M_t^s \end{matrix} \quad (48)$$

then we have:

$$\lim_{n\to\infty} \frac{1}{n} R_{23,t}^{\mathrm{T}} R_{22,t} = \lim_{n\to\infty} \frac{1}{n} U_{t-s,s}^{\mathrm{T}} Q_{x,t} R_{x2,t} \Theta_{t,s}. \quad (49)$$

When the input sequences $u_{j,t}$ are chosen such that for $n \to \infty$, the row spaces of the matrices $R_{23,t}^{\mathrm{T}} R_{22,t}$ and $\Theta_{t,s}$ coincide, we can again determine this row space via a SVD of the matrix $R_{23,t}^{\mathrm{T}} R_{22,t}$. Based on this result, we are able to generalize the PI scheme of Verhaegen (1993c) to ensemble identification problems. This extension, which will be summarized in the next paragraph, is then able to determine the quadruple of system matrices over the time-interval $t \in [t_0 + \mu, t_0 + T - 1 - \chi]$ for μ equal to $2s-1$ and χ equal to s, consistently.

The PI scheme for ensemble identification problems.

Given: the same information and conditions needed for the ordinary MOESP scheme for the ensemble identification problem.

For $t = t_0 + 2s - 1 : t_0 + T - s - 1$ carry out the following:

Step 1: construct the sequence of compound generalized Hankel matrices

$$[U_{t+1-i,s} \mid Z_{t+1-i,s} \mid U_{t+1-i-s,s}]$$

for $i = 0:1:s$;

Step 2: perform a QR factorization of the sequence of compound matrices

$$[U_{t+1-i,s} \mid Z_{t+1-i,s} \mid U_{t+1-i-s,s}]$$

for $i = 0:1:s$, without storing Q, and partition the R factor as in equation (48);

Step 3: perform a SVD of the sequence of matrices $R_{23,t+1-i}^{\mathrm{T}} R_{22,t+1-i}$ for $i = 0:1:s$ given as,

$$R_{23,t+1-i}^{\mathrm{T}} R_{22,t+1-i} = \left(\begin{array}{c|c} S_{N_{t+i-i}} & 0 \\ \hline 0 & 0 \end{array}\right) \begin{pmatrix} V_{N_{t+i-i}}^{\mathrm{T}} \\ (V_{N_{t+i-1}}^{\perp})^T \end{pmatrix}$$

and store the matrix sequences $V_{N_{t+1-i}}$, $V_{N_{t+1-i}}^{\perp}$ for $i = 0:1:s$; and

Step 4: similar to Step 4 of the ordinary MOESP scheme for the ensemble identification problem.

5. IDENTIFICATION OF DISCRETE PERIODICALLY TIME-VARYING SYSTEMS

Linear periodic systems constitute an important class of linear systems and many mechanical and chemical plants exhibit periodic behavior (Al-Rahmani and Franklin, 1989). Except for the intrinsic periodic systems, periodic discrete-time systems also naturally arise when performing multirate (MR) (see Section 5.2) or multi order (MO) (Kalman and Bertram, 1959) sampling on the output and/or input of a linear time-invariant continuous system.

It is common practice in the analysis of periodically time-varying systems to embed (or lift) them into a time-invariant system of a 'large' state dimension. This allows the use of techniques applicable to time-invariant systems. However, from a computational point of view such a lifting procedure results in very inefficient algorithms and should therefore be avoided. This assertion is supported by the authors of Lin and Kin (1993). In the context of the present paper, we give further evidence of this assertion in Remark 1.

As stated in Kalman and Bertram (1959), Meyer and Burns (1975), Araki and Yamamoto (1986) and Lin and Kin (1993), state space models are extremely well suited to analyze multirate sampled data systems. This, in combination with the fact that many system analysis and control design procedures are available for this class of systems, make it highly desirable to have access to proper identification schemes which allow the identification of time-varying state space models. With the techniques reported in the previous section as our basis we present a numerical scheme that allows us to address this task. First we specialize the notation of the state space model (2) and (3) into the present context.

Let $x_t \in R^{N_t}$, $u_t \in R^{m_t}$ and $y_t \in R^{l_t}$ be the state, input and output vectors in the state space realization (2) and (3) and let this system be periodically time-varying with period P, then the system matrices A_t, B_t, C_t, D_t in (2) and (3), satisfy the following additional constraint:

$$A_{iP+k} = A_k, \quad B_{iP+k} = B_k, \quad C_{iP+k} = C_k, \quad D_{iP+k} = D_k, \tag{50}$$

where $1 \leq k < P$, $i \in \mathbb{Z}$.

5.1. *The ordinary MOESP scheme for periodically time-varying systems*

Let us consider $t_0 = 1$ and let the following single sequences of input and output data of the periodic system to be identified be given by:

$$[u_1, u_2, \ldots u_t \cdots u_P, u_{P+1} \cdots u_{M_{\text{tot}}}]$$
$$[y_1, y_2, \ldots y_t \cdots y_P, y_{P+1} \cdots y_{M_{\text{tot}}}].$$

Then we can define an ensemble of input (and output) sequences as in the ensemble identification problem, i.e.

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_{M_{\text{tot}}-(n-1)P} \\ u_{P+1} & u_{P+2} & \cdots & \cdot \\ \vdots & & & \vdots \\ u_{(n-2)P+1} & \cdot & \cdots & \cdot \\ u_{(n-1)P+1} & & \cdots & u_{M_{\text{tot}}} \end{bmatrix} \tag{51}$$

and it becomes possible to use the developed algorithmic schemes to identify the periodically time-varying system. However, because of the periodicity, two additional remarks need to be made.

First, we want to have more insight into the minimal value of the length of the sequences, namely what is $(M_{\text{tot}})_{\min}$? The periodicity reduces the minimal number of sequences as given in equation (22) to:

$$n_{\min} = \max_{t \in [1,P]} (M_t^s + N_t).$$

To obtain the last required row space, namely $\Theta_{P,s}$, the input (and output) generalized Hankel matrices $U_{P,s}$ (and $Y_{P,s}$) are needed. Let us recall that, e.g. the matrix $U_{P,s}$, equals

$$\begin{bmatrix} u_P & u_{P+1} & \cdots & u_{P+s-1} \\ u_{2P} & u_{2P+1} & \cdots & \cdot \\ \vdots & \cdot & & \vdots \\ u_{nP} & u_{nP+1} & \cdots & u_{nP+s-1} \end{bmatrix}.$$

As a consequence, this requires that:

$$(M_{\text{tot}})_{\min} = \left[\max_{t \in [1,P]} (M_t^s + N_t) \right] P + s - 1.$$

Second, we should guarantee that the calculated state space realization is also periodic. This can simply be done by using the appropriate row space of the extended observability matrix $\Theta_{t,s}$ at various time instances. We only highlight this for the calculation of the system matrix α_t, since the same strategy applies in guaranteeing the periodicity of β_t, γ_t and δ_t.

The sequence of system matrices $\{\alpha_1, \alpha_2, \ldots, \alpha_P, \alpha_{P+1}\}$ are similarly equivalent with $\{A_1, A_2, \ldots, A_P, A_{P+1}\}$, denoted explicitly as:

$$\begin{aligned} \alpha_1 &= T_1^{-1} A_1 T_2, \ldots, \\ \alpha_{P+1} &= T_{P+1}^{-1} A_{P+1} T_{P+2} \\ &= T_{P+1}^{-1} A_1 T_{P+2}, \ldots. \end{aligned} \tag{52}$$

When, now, α_1 has to be equal to α_{P+1}, we conclude that T_{iP+k}, for $i \geq 0$ and $1 \leq k < P$, must be equal to T_k. In the identification scheme, this can be done by setting the row space of the extended observability matrices $\Theta_{iP+k,s}$ for $i \geq 0$ and $1 \leq k < P$ equal to that of $\Theta_{k,s}$ and the same for the orthogonal complements.

This analysis yields to the following algorithm to identify periodically time-varying systems: The ordinary MOESP algorithm to identify periodically time-varying systems.
Given:

(1) a uniformly controllable and observable periodic system;
(2) $s > \delta_0$;

(3) input and output sequences, which can be stacked as in equation (51) such that the series of matrices $U_{1,s} \cdots U_{P,s}$, are locally persistent excitations. This requirement rules out the use of periodic inputs with period equal to P;

(4) $M_{\text{tot}} \geq \left[\max_{t \in [1,P]} (M_t^s + N_t) \right] P + s - 1.$

Carry out the following:

Step 1: construct the input and output matrix $U_{t,s}$ and $Y_{t,s}$ for $t = 1:P$;

Step 2: implement steps 2 and 3 of the ordinary MOESP algorithm for ensemble identification problems to obtain for $t = 1:P$ the row spaces of $\Theta_{t,s}$ denoted by $V_{N_t}^{\mathrm{T}}$ and their orthogonal complements $V_{N_t}^{\perp}$;

Step 3: compute α_t by solving equation (26), and γ_t by equation (27), noting that $V_{N_{P+1}} = V_{N_1}$; and

Step 4: compute β_t and δ_t by solving equation (28), noting that $V_{N_0}^{\perp} = V_{N_P}^{\perp}$, $V_{N_{-1}}^{\perp} = V_{N_{P-1}}^{\perp}, \ldots$.

Finally, we mention that when the output is disturbed by additive zero-mean white noise, consistent estimates are obtained with $M_{\text{tot}} \to \infty$. For zero-mean additive errors of arbitrary color, in order to guarantee consistency, instrumental variables might be used in a very much the same way as outlined in the PI scheme for ensemble identification problems.

Remark 1. Making the assumption that the input and output dimensions are constant and equal to m and l, respectively, the computational complexity of the series of QR factorizations required in Step 2 of the above algorithm is $P \cdot O(n(m+l)^2 s^2)$. However, when lifting the periodically time-variant system into a time-invariant one, one requires the single QR factorization of the compound block Hankel matrix:

$$\begin{bmatrix} [u_1 \cdots u_P] & \cdots & [u_{(s-1)P+1} \cdots u_{sP}] \\ [u_{P+1} \cdots u_{2P}] & \cdots & \\ \vdots & & \\ [u_{(n-1)P+1} \cdots u_{nP}] & \cdots & \end{bmatrix}$$

$$\begin{bmatrix} [y_1 \cdots y_P] & \cdots & [y_{(s-1)P+1} \cdots y_{sP}] \\ [y_{P+1} \cdots y_{2P}] & \cdots & \\ \vdots & & \\ [y_{(n-1)P+1} \cdots y_{nP}] & & \end{bmatrix}.$$

The computational complexity of this step is $O(n(m+l)^2 s^2 P^2)$ (which clearly is an order of magnitude P, which is the length of the period) larger than the algorithm above. Similarly, we may conclude that similar savings hold in the other steps of the above algorithm, even for the case when the input and output dimensions vary in time. □

5.2. *A multirate sampling system simulation study*

To demonstrate the operation of the extended variants of the ordinary MOESP algorithm, we perform a number of experiments with a multirate sampled data system where the ratios of the sampling rates are rational numbers, that is when all the sampling periods are integer multiples of a Smallest Time Period (STP). When following such a sampling policy for a MIMO linear time-invariant system, we obtain a periodically time-varying system.

The analysis of such an example will demonstrate the capabilities of the ordinary MOESP scheme for periodically time-varying systems. However, because of the close relationship with the variant for the ensemble identification problem it also demonstrates the capabilities of the latter scheme.

The mathematical model used in the simulation. Let T now denote the STP and let the Basic Time Period (BTP) be the smallest common multiple of all sampling periods. The mathematical model of the system to be identified is given by the following second-order state space model:

$$\frac{\mathrm{d}}{\mathrm{d}t}[x_1 \quad x_2] = [x_1 \quad x_2]\begin{bmatrix} 0 & -K(t) \\ 1 & -C(t) \end{bmatrix} + f(t)[0 \quad 1],$$

where $f(t)$ is the input, x_1 and x_2 are the outputs, and $C(t)$ and $K(t)$ are time-varying scalars specified next. In the experiment we used the following numerical data:

$T = 0.01$ s,
Sampling period of input $f(t)$: T,
Sampling period of output x_2: $2T$,
Sampling period of output x_1: $3T$,
BTP $= 6T$.

$C(t)$ and $K(t)$ are periodic time functions with period BTP, and specified over one period as:

$$C(t) = 1 + 0.3 \sin\left(\frac{\pi}{6} t\right) \quad 0 \leq t \leq 6T,$$

$$K(t) = 0.5 + 0.02t.$$

We remark that identifying the above system is a multiple output identification problem. The above sampling schedule is illustrated in Fig. 1. This MR sampling system can be easily described and simulated as a SIMO system in a state framework in which the input and output quantities are permitted to have varying dimensions. The simulation experiment is carried out on the Matlab (Moler *et al.*, 1987). By using a zero-order hold for the input, the discrete-time equivalent has the following form:

$$x_{k+1} = x_k A_k + f_k B_k,$$

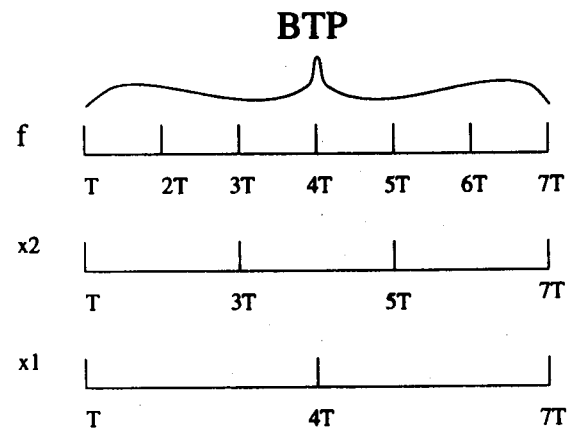

Fig. 1. Sampling periods for input and output signals in the simulation example.

where the matrix A_k is the solution of the matrix differential equation:

$$\dot{\Phi}(t, kT) = \Phi(t, kT)\begin{bmatrix} 0 & -K(t) \\ 1 & -C(t) \end{bmatrix},$$

over the time interval $t \in [kT, (k+1)T]$ with initial conditions $\Phi(kT, kT) = I$. The matrix B_k is equal to:

$$\int_{kT}^{(k+1)T} B(\sigma)\Phi((k+1)T, \sigma)\, d\sigma.$$

The output matrices of the multirate sampled data system used in the simulation are:

$$C_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C_2 = [\ ], \quad C_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$C_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C_5 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_6 = [\ ].$$

The experiments. We set up two Monte Carlo simulation tests with both SNR equal to 30 dB. The input f is a white noise sequence with variance equal to 1. Let s be equal to 10. In this example, $P = 6$, $\max_{t \in (1,P)} (M_t^{10} + N_t) = 12$, then the $\min(M_{tot}) = 81$. We need two different M_{tot} to test the properties of unbiasedness and consistency of the algorithm: $M_{tot_1} = 150$ and $M_{tot_2} = 600$. One hundred runs were performed in each test.

The quantities x_1 and x_2 denote the noiseless output, v_1 and v_2 denote the added output noise and they are zero-mean white. The output measurements are:

$$x_{1,m} = x_1 + v_1,$$
$$x_{2,m} = x_2 + v_2.$$

The identified second-order models were compared with two different aspects as discussed below.

First, we consider the transition matrix of the given discretized system over one period, namely $A_1A_2 \cdots A_6$ and the same form for the identified model, namely $\alpha_1\alpha_2 \cdots \alpha_6$. Using the relationship between α_i and A_i as displayed in equation (52) and the fact that the identification algorithm assures that $T_1 = T_{1+P}$, yields that:

$$A_1A_2 \cdots A_6 = T_1^{-1}\alpha_1\alpha_2 \cdots \alpha_6 T_1.$$

Therefore, in the noise-free case, the eigenvalues of these two matrix products coincide. In the noisy data case, we can only expect the eigenvalues of the matrix $\alpha_1\alpha_2 \cdots \alpha_6$ to be asymptotically unbiased estimates of those of the matrix $A_1A_2 \cdots A_6$. Therefore, as a first measure of the performance of the developed procedure, we depict in Fig. 2, the eigenvalues of both matrices. In this figure, the center of the cross denotes the true eigenvalue location. From the figure we see that when keeping s constant, the estimated eigenvalues are unbiased and have a reduced variance when M_{tot} increases. Thus this experiment confirms Corollary 1.

Second, we compare the output error that is the error between the measured outputs $x_{2,m}$,

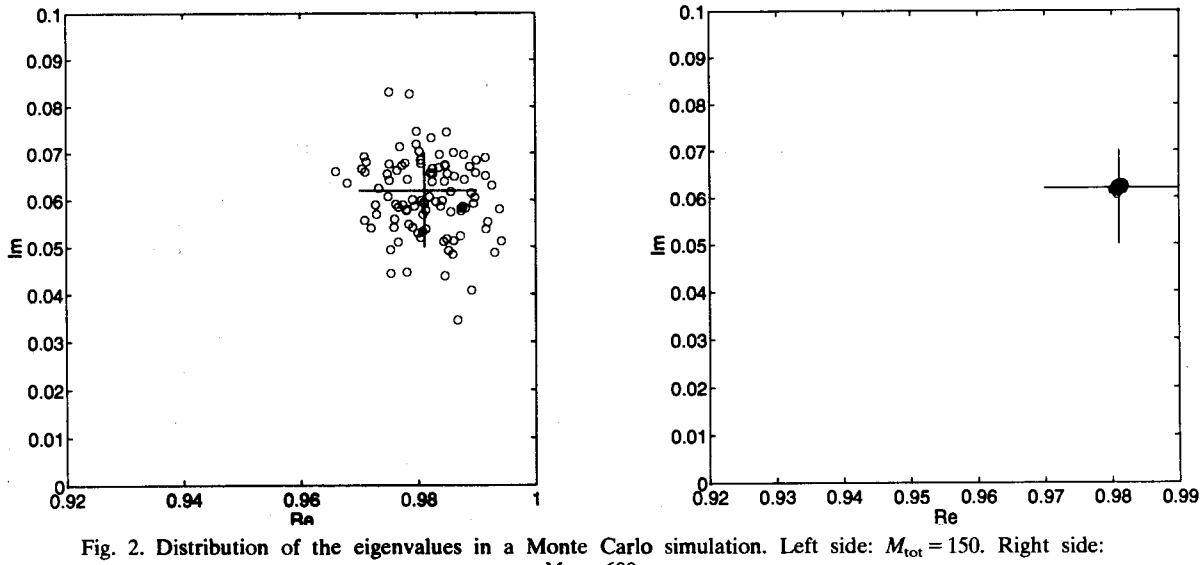

Fig. 2. Distribution of the eigenvalues in a Monte Carlo simulation. Left side: $M_{tot} = 150$. Right side: $M_{tot} = 600$.

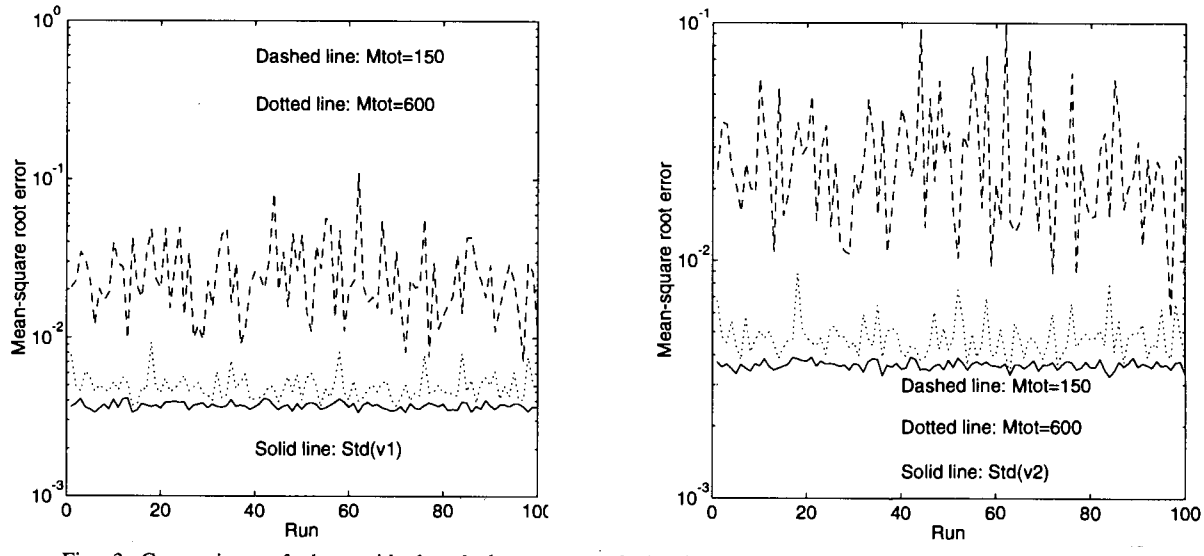

Fig. 3. Comparison of the residuals of the output of the identified second-order system with the measurement errors. Left side: std (v_1), std $(\hat{x}_1 - x_{1,m})$. Right side: std (v_2), std $(\hat{x}_2 - \hat{x}_{2,m})$.

$x_{1,m}$ and the reconstructed outputs $\hat{x}_2$, $\hat{x}_1$ using the estimated periodic data model over a time interval of 600 input samples. The results of this comparison are presented in Fig. 3. Here, the solid lines represent the noise level on the output as indicated by std (v_1) and std (v_2), where std denotes the standard deviation. The other types of lines show the quantities of std $(\hat{x}_2 - x_{2,m})$ and std $(\hat{x}_1 - x_{1,m})$. When the reconstruction could be done with the original system the latter quantities are then equal to the noise level. Also these results now confirm the consistency of the full identified model.

To end this section, we illustrate how the order can be detected following the remark made at the end of Section 4.2.1. We, therefore, use the singular values of the matrix $R_{22,t}$ for $t = 1:1:6$ and the residuals v_1, v_3, v_4 and v_5, defined as:

$$v_1^2 = \sum_{j=0}^{\mathrm{fix}(M_{\mathrm{tot}}/P-1)} \left\| \begin{bmatrix} x_{1,m}(jP) - \hat{x}_1(jP) \\ x_{2,m}(jP) - \hat{x}_2(jP) \end{bmatrix} \right\|^2,$$

$$v_3^2 = \sum_{j=0}^{\mathrm{fix}(M_{\mathrm{tot}}/P-1)} (x_{2,m}(jP+2) - \hat{x}_2(jP+2))^2,$$

$$v_4^2 = \sum_{j=0}^{\mathrm{fix}(M_{\mathrm{tot}}/P-1)} (x_{1,m}(jP+3) - \hat{x}_1(jP+3))^2,$$

$$v_5^2 = \sum_{j=0}^{\mathrm{fix}(M_{\mathrm{tot}}/P-1)} (x_{2,m}(jP+4) - \hat{x}_2(jP+4))^2,$$

where fix $(M_{\mathrm{tot}}/P - 1)$ is the integer number closest to $M_{\mathrm{tot}}/P - 1$ towards 0. The singular values are displayed in Fig. 4 for M_{tot} equal to 150 on the left-hand side and 600 on the right-hand side. The residuals are displayed in our final figure. For both cases of M_{tot} we observe a clear gap between the second and third largest singular value in combination with no significant decrease in the residuals for models of order 2 and larger. For the case $M_{\mathrm{tot}} = 150$, the residuals even increase for model orders 3 and 4 and, therefore, we restrict displaying the residuals only up to models of order 3 in the left-hand side of Fig. 5. Hence the outlined strategy allows the detection of the correct system order, namely 2.

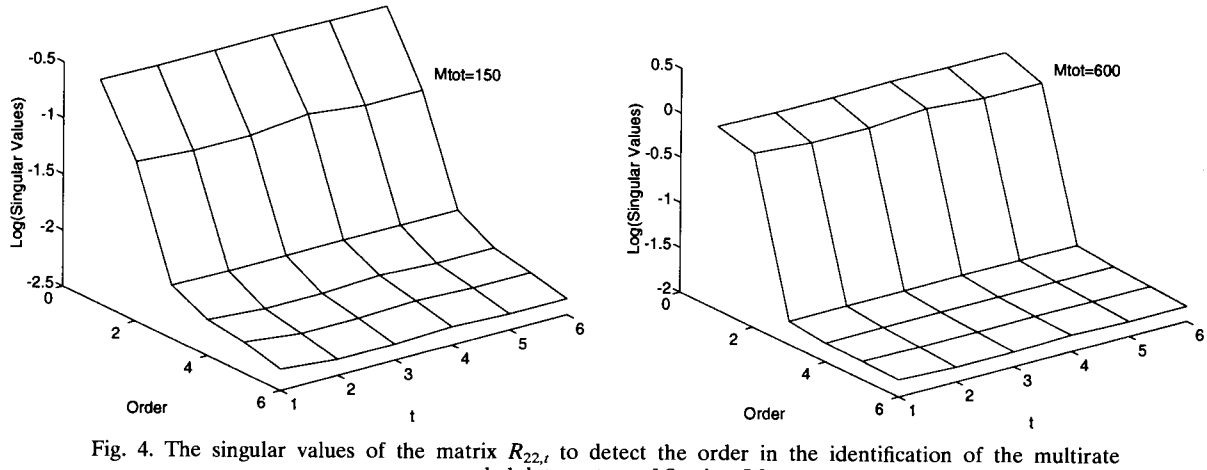

Fig. 4. The singular values of the matrix $R_{22,t}$ to detect the order in the identification of the multirate sampled data system of Section 5.2.

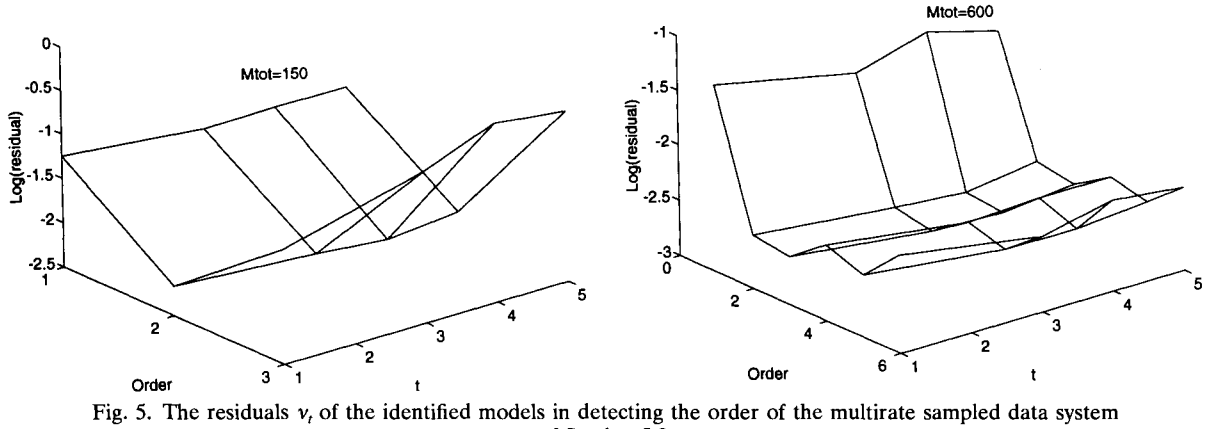

Fig. 5. The residuals v_t of the identified models in detecting the order of the multirate sampled data system of Section 5.2.

6. CONCLUSIONS

In this paper, algorithms for identifying a discrete, linear time-varying state space model have been introduced. The algorithms require a multiple series of experiments, each input and output sequences is recorded when the underlying system undergoes the same time-varying behavior. The algorithms were extended for periodic discrete-time systems. Here the repetition of the experiment is intrinsic. An application of the developed schemes to the identification of a multivariable multirate sampled data system demonstrates the usefulness of these schemes. The results of the simulation experiment showed that the algorithm allows us to consistently estimate a state space model for such systems when the noise on the output is zero-mean white.

The additional usefulness of the schemes developed in this paper was demonstrated in Yu and Verhaegen (1993). In the latter paper, a realistic identification problem, namely the identification of the human joint dynamics, was considered. For this practical application, especially the PI scheme for the ensemble identification problem, allowed the identification of low-order state space models, with the order changing from 2 to 3, which lead to an accurate reconstruction of the output. In the same paper, the results were compared with recursive output error identification schemes using exponential forgetting factors between 0.9 and 0.95 and this comparison study clearly showed that the latter approach was completely inadequate in identifying the underlying foot dynamics.

Further research on the use of the numerical schemes developed is proposed. A first topic would be the application of the schemes to the identification of non-linear systems and linear time-varying systems with poor a prior information on the nature of the changing time-varying behavior. Another topic would be to increase the computational efficiency of the algorithms proposed. Research on this topic might be performed along the lines of the approach followed in Cho *et al.* (1994) however, now making use of the displacement theory in a time-varying context such as developed in Constantinescu *et al.* (1993).

Acknowledgements—The research of Dr Michel Verhaegen has been made possible by the funding of a senior research fellowship from the Royal Dutch Academy of Arts and Sciences. This research was supported in part by the Commission of the EC under the ESPRIT BRA program 6632 (NANA2). The authors wish to thank the reviewers of this paper for many constructive suggestions.

REFERENCES

Al-Rahmani, H. M. and G. F. Franklin (1989). Linear periodic systems: eigenvalue assignment using discrete feedback. *IEEE Trans. Autom. Contr.*, **34**, 99–103.

Anderson, T. W. (1966). *An Introduction to Multivariate Statistical Analysis*. J. Wiley and Sons, Chichester.

Araki, M. and K. Yamamoto (1986). Multivariable multirate sampled-data systems: state-space description, transfer characteristics, and Nyquist criterion. *IEEE Trans. Autom. Contr.*, **31**, 145–154.

Cho, Y. M., G. Xu and T. Kailath (1994). Fast recursive identification of state space models via exploitation of displacement structure. *Automatica*, **30**.

Constantinescu, T., A. H. Sayed and T. Kailath (1993). Structured Matrices and Moment Problems. In P. Dewilde, M. A. Kaashoek and M. Verhaegen (Eds), *Challenges of a Generalized System Theory*. North Holland, Amsterdam.

Dewilde, P. and H. Dym (1992). Interpolation for upper triangular operators. In I. Gohberg, (Ed.), *Operator Theory: Advances and Applications*, Vol. OT 56, pp. 153–260. Birkhaüser, Basel.

Dewilde, P. M. and A. J. Van der Veen (1993). On the Hankel–Norm approximation of upper-triangular operators and matrices. *Integral Equations Operator Theory*, **17**, 1–45.

Dewilde, P., M. A. Kaashoek and M. Verhaegen (Eds) (1993). *Challenges of a Generalized System Theory*. Essays of the Royal Dutch Academy of Sciences. North-Holland, Amsterdam.

Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press, New York.

Kailath, T. (1980). *Linear Systems*. Prentice-Hall, Englewood Cliffs, N.J.

Kalman, R. E. and J. E. Bertram (1959). A unified approach to the theory of sampling systems. *J. Franklin Inst.*, 405–436.

Kearney, R. E., R. E. Kirsch, J. B. MacNeil and I. W. Hunter (1991). An ensemble time-varying identification

technique: theory and biomedical applications. In *Preprints of the 9th IFAC/FORS Symposium on Identification and System Parameter Estimation*, pp. 191–196.
Larimore, W. (1990). Canonical variate analysis in identification, filtering and adaptive control. In *Proc. 29th OEEE Conf. Decision and Contr.*, pp. 596–604.
Lin, C. A. and C. W. Kin (1993). Minimal periodic realizations of transfer matrices. *IEEE Trans. Autom. Contr.*, **38**, 462–466.
Ljung, L. (1987). *System Identification: Theory for the User.* Prentice-Hall, Englewood Cliffs, N.J.
MacNeil, J. B., R. E. Kearney and I. W. Hunter (1992). Identification of time-varying biological systems from ensemble data. *IEEE Trans. Biomed. Engng*, **39**, 1213–1225.
Meyer, R. A. and C. S. Bunus (1975). A unified analysis of multirate and periodically time-varying digital filters. *IEEE Trans. Circuits and Syst.*, **CAS-22**, 163–168.
Moonen, M. B., De Moor, L. Vandenberghe and J. Vandewalle (1989). On- and off-line identification of linear state space model. *Int. J. Contr.*, **49**, 219–232.
Moler, C., J. Little and S. Bangert (1987). PRO-MATLAB User's Guide. The MathWorks.
Niedzwiecki, M. (1990). Recursive functional series modeling estimators for identification of time-varying plants—more bad news than good? *IEEE Trans. Autom. Contr.*, **35**, 610–616.
Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, New York.
Shokoohi, S. and L. M. Silverman (1987). Identification and model reduction of time-varying discrete-time systems. *Automatica*, **23**, 509–522.
Söderström, T. and P. Stoica (1989). *System Identification.* Prentice-Hall, Englewood Cliffs, N.J.
Strang, G. (1980). *Linear Algebra and its Applications.* Academic Press, New York.
Van Overschee, P. and B. De Moor (1994). N4SID: subspace algorithm for the identification of combined deterministic-stochastic systems. *Automatica*, **30**, 75–93.
Verhaegen, M. (1991). Identification of time-varying state space models from input–output Data. In *Proc. of the Workshop on Advanced Algorithms and Their Realization*, Bonas, August 1991, Chapter XI.
Verhaegen, M. (1993a). Application of a subspace model identification technique to identify LTI systems operating in closed loop. *Automatica*, **29**, 1027–1040.
Verhaegen, M. (1993b). Identification of the deterministic part of MIMO state space models given in innovation form from input–output data. *Automatica*, **30**, 61–74.
Verhaegen, M. (1993c). Subspace model identification. Part III: analysis of the ordinary output-error state space model identification algorithm. *Int. J. Contr.*, **58**, 555–586.
Verhaegen, M. and P. Dewilde (1992a). Subspace model identification. Part I: the output-error state space model identification class of algorithms, *Int. J. Contr.*, **56**, 1187–1210.
Verhaegen, M. and P. Dewilde (1992b). Subspace model identification. Part II: analysis of the elementary output-error state space model identification algorithm. *Int. J. Contr.*, **56**, 1211–1241.
Widrow, B. and E. Walach (1984). On the statistical efficiency of the LMS algorithm with nonstationary inputs. *IEEE Trans. Inf. Theory*, **IT-30**.
Yu, X. and M. Verhaegen (1993). Application of a time-varying subspace model identification scheme to the identification of the human joint dynamics. In *Proc. of European Contr. Conf.*, Groningen, The Netherlands, Vol. 2, pp. 603–608.

APPENDIX

An illustration of Definition 6

Assume there is a uniformly controllable, linear and time-varying system, the input dimension is 1 and the input sequences are:

$$u_{j,t} = \begin{cases} 1 & j = t \\ 0 & \text{otherwise} \end{cases}$$

With $j \in [1, n]$ and $t \in [1, T]$, then the input sequence matrix $U_{1,T}$ for $T > n$ is of the type:

$$U_{1,T} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}. \qquad \text{(A.1)}$$

For simplicity, suppose that for all the experiments performed the initial state is zero, the order of the system is 2 in the interval $[1, T]$, $\delta_c = 2$ and the window length is chosen such that $n - s \geq \delta_c$, in this example, we look at $s = 4$. Then the matrix $[U_{1,4} \mid X_{(1,1),n}]$ reads:

$$[U_{1,4} \mid X_1] = \left[\begin{array}{cccccc} 1 & & & & 0 & 0 \\ & 1 & & & 0 & 0 \\ & & 1 & & 0 & 0 \\ & & & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & & & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

For the zero initial condition, the rank of the matrix is 4 and the sequences stacked in it are not locally persistent excitations at time instant $t = 1$.

Next, for the same experiment, consider the matrix $[U_{3,4} \mid X_3]$:

$$[U_{3,4} \mid X_3] = \left[\begin{array}{cccc:c} 0 & 0 & 0 & 0 & B_1A_2 \\ 0 & 0 & 0 & 0 & B_2 \\ \hline 1 & & & & 0 \\ & 1 & & & 0 \\ & & 1 & & 0 \\ & & & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Since $\delta_c = 2$, Lemma 3 shows that the submatrix $\begin{bmatrix} B_1A_2 \\ B_2 \end{bmatrix}$ has full rank, and therefore the matrix $[U_{3,4} \mid X_3]$ has full column rank.

Since $n - s \geq \delta_c$, Lemma 3 shows that the last matrix in the sequence $\{[U_{t,4} \mid X_4]\}$ for $t > 1$ which has full column rank is the matrix:

$$[U_{n-s+1,4} \mid X_{n-s+1}] = \left[\begin{array}{cccc:c} & & & & B_1A_2 \cdots A_{n-s} \\ & & & & \vdots \\ & & & & B_{n-s}A_{n-s} \\ & & & & B_{n-s} \\ 1 & & & & 0 \\ & 1 & & & 0 \\ & & 1 & & 0 \\ & & & 1 & 0 \end{array}\right].$$

Summarizing, when using Dirac impulse as input and zero initial state condition, the matrices $[U_{t,s} \mid X_t]$ for $t \in [\delta_c + 1, n - s + 1]$ are full rank matrices for $n - s \geq \delta_c$.