



**HAL**  
open science

# Convergence analysis of instrumental variable recursive subspace identification algorithms

Guillaume Mercère, Marco Lovera

► **To cite this version:**

Guillaume Mercère, Marco Lovera. Convergence analysis of instrumental variable recursive subspace identification algorithms. 14th IFAC Symposium on System Identification, SYSID'2006, Mar 2006, Newcastle, Australia. pp.CDROM. hal-00098152

**HAL Id: hal-00098152**

**<https://hal.science/hal-00098152>**

Submitted on 25 Sep 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# CONVERGENCE ANALYSIS OF INSTRUMENTAL VARIABLE RECURSIVE SUBSPACE IDENTIFICATION ALGORITHMS

G. Mercère\* M. Lovera\*\*

\* *Laboratoire d'Automatique et d'Informatique Industrielle*  
40 avenue du recteur Pineau, 86022 Poitiers, France  
Phone: + 33 (0)5 49453509; Fax: +33 (0)5 49454034  
Email: [guillaume.mercere@univ-poitiers.fr](mailto:guillaume.mercere@univ-poitiers.fr)

\*\* *Dipartimento di Elettronica e Informazione, Politecnico di Milano*  
Piazza Leonardo da Vinci 32, 20133 Milano, Italy  
Phone: + 39 (0)2 23993592; Fax: +39 (0)2 23993412  
Email: [lovera@elet.polimi.it](mailto:lovera@elet.polimi.it)

Abstract: The convergence properties of a recently developed recursive subspace identification algorithm are investigated in this paper. The algorithm operates on the basis of an extended instrumental variable (EIV) version of the propagator method for signal subspace estimation. It is proved that, under weak conditions on the input signal and the identified system, the considered MOESP class of recursive subspace identification algorithm converges to a consistent estimate of the propagator and, by extension, of the state space system matrices.

Keywords: Subspace methods, Recursive algorithms, Identification algorithms, State space models, Tracking systems.

## 1. INTRODUCTION

Algorithms for recursive subspace model identification (RSMI) have been extensively studied in recent years (see, e.g., (Verhaegen and Deprettere, 1991; Cho *et al.*, 1994; Gustafsson, 1997; Lovera *et al.*, 2000; Oku and Kimura, 2002; Mercère *et al.*, 2004b)).

So far, two main approaches to the RSMI problem have been developed in the literature. First, some works have proposed adaptations of subspace model identification (SMI) algorithms in order to update the singular value decomposition (SVD) (Verhaegen and Deprettere, 1991; Cho *et al.*, 1994). Unfortunately, these methods have the drawback of requiring the disturbances acting on the system output to be spatially and temporally white, which is obviously restrictive in practice. The second approach (Lovera *et al.*, 2000; Oku and Kimura, 2002; Mercère *et al.*, 2004b) relies on the strong analogies between RSMI and signal processing techniques dedicated to direction of arrival

(DOA) estimation. Two points of view have been more precisely suggested to find SVD alternatives in a recursive framework:

- The first one consists in adapting the so-called Yang's criterion (Yang, 1995) to the recursive update of the observability matrix (Lovera *et al.*, 2000; Oku and Kimura, 2002; Lovera, 2003). In particular, DOA estimation algorithms have been adjusted in order to deal with more general types of perturbations than the ones arising in the DOA framework thanks to the use of instrumental variables.
- The second one rests on the adaptation of another array signal processing technique: the propagator method (Munier and Delisle, 1991). The advantage of this approach over the previous conception lies in the use of a linear operator and unconstrained and unapproximated quadratic criteria which lead to easy recursive least squares

algorithms (Mercère *et al.*, 2003; Mercère *et al.*, 2004b; Mercère *et al.*, 2005).

While a significant level of maturity has now been reached on the algorithmic side, very limited attention has been dedicated to the analysis of the convergence properties of the proposed methods for the recursive update of the subspace estimates. In (Oku and Kimura, 2002), the developed gradient-based RSMTI technique was studied and conditions on the gain of the gradient iteration were derived. The convergence study, however, is based on assumptions on the signal-to-noise ratio which limit the validity of the results.

In the light of the above discussion, the aim of this paper is to investigate the convergence properties of recursive implementations of the MOESP class (Verhaegen, 1994) of subspace identification methods. In particular, the RSMTI algorithms which operate on the basis of the IV version of the propagator method for signal subspace estimation will be considered and a convergence result for such techniques will be derived and discussed.

## 2. PROBLEM FORMULATION AND NOTATION

Assume that the true system can be described by the discrete-time linear time-invariant state space model in innovation form

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{K}\mathbf{e}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{e}(t) \end{aligned} \quad (1)$$

with  $n_y$  outputs  $\mathbf{y}$ ,  $n_u$  inputs  $\mathbf{u}$ ,  $n_x$  states  $\mathbf{x}$ . Assume also that:

- A1:** the system (1) is asymptotically stable;
- A2:** the pairs  $\{\mathbf{A}, \mathbf{C}\}$  and  $\{\mathbf{A}, [\mathbf{B} \ \mathbf{K}]\}$  are respectively observable and reachable;
- A3:** the input  $\mathbf{u}$  is a quasi stationary deterministic sequence with correlation function (Ljung, 1999)

$$\mathbf{R}_{\mathbf{u}}(\tau) = \bar{\mathbb{E}} [\mathbf{u}(t+\tau)\mathbf{u}^T(t)] \quad (2)$$

where

$$\bar{\mathbb{E}}[\cdot] = \lim_{N \rightarrow \infty} \sum_{t=1}^N \mathbb{E}[\cdot] \quad (3)$$

and  $\mathbb{E}$  is the expectation operator.

The algorithm considered in this paper recursively estimate the state space matrices  $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$  at each new data acquisition. The proposed method is based on the estimation of a basis for the observability subspace from the input output (I/O) relation (Mercère *et al.*, 2004b)

$$\mathbf{y}_f(t) = \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{H}_f \mathbf{u}_f(t) + \mathbf{b}_f(t) \quad (4)$$

where the stacked input and output vectors ( $\mathbf{u}_f$  and  $\mathbf{y}_f$ , respectively) are defined as

$$\begin{aligned} \mathbf{u}_f(t) &= [\mathbf{u}^T(t) \ \cdots \ \mathbf{u}^T(t+f-1)]^T \in \mathbb{R}^{n_u f \times 1} \\ \mathbf{y}_f(t) &= [\mathbf{y}^T(t) \ \cdots \ \mathbf{y}^T(t+f-1)]^T \in \mathbb{R}^{n_y f \times 1} \end{aligned} \quad (5)$$

with  $f > n_x$ ,  $\mathbf{\Gamma}_f$  is the observability matrix

$$\mathbf{\Gamma}_f = [\mathbf{C}^T \ (\mathbf{C}\mathbf{A})^T \ \cdots \ (\mathbf{C}\mathbf{A}^{f-1})^T]^T, \quad (6)$$

$\mathbf{H}_f$  is the block Toeplitz matrix of the impulse responses from  $\mathbf{u}$  to  $\mathbf{y}$  and  $\mathbf{b}_f = \mathbf{G}_f \mathbf{e}_f$  with  $\mathbf{G}_f$  the block Toeplitz matrix of the impulse responses from  $\mathbf{e}$  to  $\mathbf{y}$ . The class of algorithms considered herein is based on the application of the so-called propagator method (Munier and Delisle, 1991) (first proposed in the array signal processing literature) to the recursive estimation of  $\mathbf{\Gamma}_f$ . To this purpose, note that letting

$$\mathbf{z}_f(t) = \mathbf{y}_f(t) - \mathbf{H}_f \mathbf{u}_f(t), \quad (7)$$

equation (4) can be written as

$$\mathbf{z}_f(t) = \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{b}_f(t). \quad (8)$$

From this relation, a two-step procedure for the recursive estimation of the system matrices can be devised:

- (1) update of the ‘‘observation vector’’  $\mathbf{z}_f$  from the I/O measurements by using (7) (see Section 3);
- (2) estimation of a basis of  $\mathbf{\Gamma}_f$  from this observation vector by using (8) (see Section 4).

## 3. RECURSIVE ESTIMATION OF THE OBSERVATION VECTOR

The problem of estimating the observation vector can be solved by adjusting ideas from offline subspace identification to the recursive framework. For, let

$$\mathbf{Z}_f = \mathbf{Y}_f - \mathbf{H}_f \mathbf{U}_f \quad (9)$$

where  $\mathbf{Y}_f \in \mathbb{R}^{n_y f \times N}$  and  $\mathbf{U}_f \in \mathbb{R}^{n_u f \times N}$  are the Hankel I/O data matrices defined as:

$$\mathbf{Y}_f(\bar{t}) = [\mathbf{y}_f(t) \ \cdots \ \mathbf{y}_f(t+N-1)] \quad (10)$$

with  $N \gg f > n$  and  $\bar{t} = t + N - 1$ . Then, it is easy to show that:

$$\mathbf{Z}_f(\bar{t}) = \mathbf{Y}_f(\bar{t}) - \mathbf{H}_f \mathbf{U}_f(\bar{t}) = [\mathbf{Z}_f(\bar{t}-1) \ \mathbf{z}_f(\bar{t})] \quad (11)$$

which proves that the update of  $\mathbf{Z}_f$  leads to the observation vector at time  $\bar{t}$ . Since, in offline subspace identification,  $\mathbf{Z}_f$  is calculated from the orthogonal projection on the kernel of  $\mathbf{U}_f$

$$\mathbf{Z}_f = \mathbf{Y}_f \mathbf{\Pi}_{\mathbf{U}_f^\perp}, \quad (12)$$

the proposed method consists in recursively updating this projection at each iteration. Several algorithms for the computation of this update have been developed in the literature (see, e.g., (Lovera *et al.*, 2000; Oku and Kimura, 2002; Mercère *et al.*, 2004a)). In this paper, an approach based on the matrix inversion lemma will be used, which has the advantage of providing an explicit expression of the observation vector in terms of the I/O data (Mercère *et al.*, 2004a). The idea is to recursively update the quantity

$$\begin{aligned} &\mathbf{Y}_f(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f(\bar{t})^\perp} = \\ &\mathbf{Y}_f(\bar{t}) \left\{ \mathbf{I} - \mathbf{U}_f^T(\bar{t}) (\mathbf{U}_f(\bar{t}) \mathbf{U}_f^T(\bar{t}))^{-1} \mathbf{U}_f^T(\bar{t}) \right\} \end{aligned} \quad (13)$$

at each new data acquisition, knowing that

$$\mathbf{U}_f(\bar{t}) = [\mathbf{U}_f(\bar{t}-1) \mathbf{u}_f(\bar{t})] \quad (14)$$

$$\mathbf{Y}_f(\bar{t}) = [\mathbf{Y}_f(\bar{t}-1) \mathbf{y}_f(\bar{t})] \quad (15)$$

by applying the matrix inversion lemma to  $(\mathbf{U}_f \mathbf{U}_f^T)^{-1}$ . It can be shown that the observation vector can be recursively estimated with the following algorithm (Mercère *et al.*, 2004a):

*Algorithm 3.1.* Assume that  $\mathbf{W}_f = (\mathbf{U}_f \mathbf{U}_f^T)^{-1}$  and  $\mathbf{V}_f = \mathbf{Y}_f \mathbf{U}_f^T$  have been estimated at time  $\bar{t}-1$ . Then, when a new I/O data sequence  $\{\mathbf{u}_f(\bar{t}), \mathbf{y}_f(\bar{t})\}$  is acquired, the observation vector is updated by the following recursion:

$$\boldsymbol{\beta}_f(\bar{t}) = \mathbf{W}_f(\bar{t}-1) \mathbf{u}_f(\bar{t}) \quad (16a)$$

$$\delta_f(\bar{t}) = \mathbf{u}_f^T(\bar{t}) \boldsymbol{\beta}_f(\bar{t}) \quad (16b)$$

$$\alpha_f(\bar{t}) = \frac{1}{1 + \delta_f(\bar{t})} \quad (16c)$$

$$\mathbf{z}_f(\bar{t}) = \alpha_f(\bar{t}) (\mathbf{y}_f(\bar{t}) - \mathbf{V}_f(\bar{t}-1) \boldsymbol{\beta}_f(\bar{t})) \quad (16d)$$

$$\mathbf{V}_f(\bar{t}) = \mathbf{V}_f(\bar{t}-1) + \mathbf{y}_f(\bar{t}) \mathbf{u}_f^T(\bar{t}) \quad (16e)$$

$$\mathbf{W}_f(\bar{t}) = (\mathbf{W}_f(\bar{t}-1) - \alpha_f(\bar{t}) \boldsymbol{\beta}_f(\bar{t}) \boldsymbol{\beta}_f^T(\bar{t}))^{-1}. \quad (16f)$$

*Remark 3.1.* Letting  $\mathbf{X}(\bar{t}) = [\mathbf{x}(t) \cdots \mathbf{x}(\bar{t})]$  and replacing in (16d)  $\mathbf{y}_f$  with (4) and  $\mathbf{Y}_f$  with (Verhaegen, 1994)

$$\mathbf{Y}_f(\bar{t}-1) = \mathbf{G}_f \mathbf{X}(\bar{t}-1) + \mathbf{H}_f \mathbf{U}_f(\bar{t}-1) + \mathbf{B}_f(\bar{t}-1), \quad (17)$$

it is easy to show that

$$\begin{aligned} \mathbf{z}_f(\bar{t}) &= \alpha_f(\bar{t}) \mathbf{G}_f (\mathbf{x}(\bar{t}) - \mathbf{X}(\bar{t}-1) \mathbf{U}_f^T(\bar{t}-1) \boldsymbol{\beta}_f(\bar{t})) \\ &+ \alpha_f(\bar{t}) (\mathbf{b}_f(\bar{t}) - \mathbf{B}_f(\bar{t}-1) \mathbf{U}_f^T(\bar{t}-1) \boldsymbol{\beta}_f(\bar{t})) \\ &= \alpha_f(\bar{t}) (\mathbf{G}_f \tilde{\mathbf{x}}(\bar{t}) + \tilde{\mathbf{b}}_f(\bar{t})). \end{aligned} \quad (18)$$

It is apparent from the above equation that  $\mathbf{z}_f$  belongs to  $\text{span}_{\text{col}} \{\mathbf{G}_f\}$ .

*Remark 3.2.* In the following, the factor  $\alpha_f$ , which appears as a scaling factor in the expression of  $\mathbf{z}_f$ , will be neglected for simplicity *i.e.*

$$\mathbf{z}_f(\bar{t}) = \mathbf{G}_f \tilde{\mathbf{x}}(\bar{t}) + \tilde{\mathbf{b}}_f(\bar{t}). \quad (19)$$

Note that this simplification does not affect the properties of the algorithm since (19) provides all the information needed to estimate  $\text{span}_{\text{col}} \{\mathbf{G}_f\}$ .

#### 4. RECURSIVE UPDATE OF THE OBSERVABILITY MATRIX

Once the observation vector is estimated, the second step of the recursive subspace identification procedure consists in online updating the observability matrix. Unlike previous approaches (see, e.g., (Lovera *et al.*,

2000; Oku and Kimura, 2002)), in this paper, the focus will be on updating algorithms based on the propagator concept (see also (Mercère *et al.*, 2004b)).

Under assumption A2, since  $\mathbf{G}_f \in \mathbf{R}^{n_y f \times n_x}$  with  $n_y f > n_x$ ,  $\mathbf{G}_f$  has at least  $n_x$  linearly independent rows, which can be gathered in a submatrix  $\mathbf{G}_{f_1}$ . Then, the complement  $\mathbf{G}_{f_2}$  of  $\mathbf{G}_{f_1}$  can be expressed as a linear combination of these  $n_x$  rows. So, there is a unique linear operator  $\mathbf{P}_f \in \mathbf{R}^{n_x \times (n_y f - n_x)}$ , named propagator (Munier and Delisle, 1991), such that

$$\mathbf{G}_{f_2} = \mathbf{P}_f^T \mathbf{G}_{f_1}. \quad (20)$$

Furthermore, it is easy to verify that

$$\mathbf{G}_f = \begin{bmatrix} \mathbf{G}_{f_1} \\ \mathbf{G}_{f_2} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{f_1} \\ \mathbf{P}_f^T \mathbf{G}_{f_1} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} \mathbf{G}_{f_1} = \mathbf{E}_o \mathbf{G}_{f_1}. \quad (21)$$

Thus, since  $\text{rank} \{\mathbf{G}_{f_1}\} = n_x$ ,

$$\text{span}_{\text{col}} \{\mathbf{G}_f\} = \text{span}_{\text{col}} \{\mathbf{E}_o\}. \quad (22)$$

Equation (22) implies that it is possible to estimate the observability matrix (in a particular basis) by estimating the propagator. This operator can be determined from (8). Indeed, applying a data reorganization so that the first  $n_x$  rows of  $\mathbf{G}_f$  are linearly independent, (8) can be partitioned as

$$\mathbf{z}_f(t) = \begin{bmatrix} \mathbf{z}_{f_1}(t) \\ \mathbf{z}_{f_2}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} \mathbf{G}_{f_1} \mathbf{x}(t) + \begin{bmatrix} \mathbf{b}_{f_1}(t) \\ \mathbf{b}_{f_2}(t) \end{bmatrix} \quad (23)$$

where  $\mathbf{z}_{f_1} \in \mathbf{R}^{n_x \times 1}$  and  $\mathbf{z}_{f_2} \in \mathbf{R}^{(n_y f - n_x) \times 1}$  are the components of  $\mathbf{z}_f$  respectively corresponding to  $\mathbf{G}_{f_1}$  and  $\mathbf{G}_{f_2}$ . In the ideal noise free case, it is easy to show that

$$\mathbf{z}_{f_2} = \mathbf{P}_f^T \mathbf{z}_{f_1}. \quad (24)$$

In the presence of noise, this relation no longer holds. However, an estimate of  $\mathbf{P}_f$  can be obtained by minimising the cost function

$$J(\mathbf{P}_f) = \mathbb{E} \|\mathbf{z}_{f_2} - \mathbf{P}_f^T \mathbf{z}_{f_1}\|^2. \quad (25)$$

It is easy to see from (23) that the estimate of  $\mathbf{P}_f$  obtained by minimising (25) is biased (Mercère *et al.*, 2004b). This issue is normally circumvented in the array signal processing literature by assuming that the noise vector  $\mathbf{b}_f$  is spatially and temporally white and simultaneously estimating the propagator and the noise variance (see, e.g., (Marcos *et al.*, 1995)). Unfortunately, it is apparent from (18) that the noise  $\tilde{\mathbf{b}}_f$  is not white. So, it is necessary to modify the criterion (25) to ensure that the propagator approach provides unbiased estimates in this more general case. This is obtained by introducing an instrumental variable  $\boldsymbol{\xi} \in \mathbf{R}^{n_\xi \times 1}$  in (25), assumed to be uncorrelated with the noise but sufficiently correlated with the state vector  $\mathbf{x}$ , and by defining the new cost function

$$J_{IV}(\mathbf{P}_f) = \left\| \mathbf{R}_{\mathbf{z}_{f_2}} \boldsymbol{\xi} - \mathbf{P}_f^T \mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi} \right\|_F^2. \quad (26)$$

Four algorithms (IVPM, EIVPM, EIVsqrtPM and COIVPM (Mercère *et al.*, 2003; Mercère *et al.*, 2004a; Mercère *et al.*, 2005)) have been developed to minimise this criterion according to the number of instruments in  $\boldsymbol{\xi}$ . In this paper, in order to study the

convergence properties of the propagator associated with an instrumental variable, the EIVPM algorithm is considered. This technique requires to construct an instrumental variable such that  $n_\xi \geq n_x$ . By assuming that the input is sufficiently ‘‘rich’’ (see Section 5) so that  $\mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}$  is full rank, the asymptotic least squares estimate of the propagator is given by

$$\hat{\mathbf{P}}_f^T = \mathbf{R}_{\mathbf{z}_{f_2}} \boldsymbol{\xi} \mathbf{R}_{\mathbf{z}_{f_1}}^\dagger \boldsymbol{\xi}. \quad (27)$$

Then, a recursive version of (27) can be obtained by adapting the overdetermined instrumental variable technique first proposed in (Friedlander, 1984). The resulting algorithm is given by

$$\mathbf{g}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_2}} \boldsymbol{\xi}(t) \boldsymbol{\xi}(t) & \mathbf{z}_{f_2}(t) \end{bmatrix} \quad (28a)$$

$$\boldsymbol{\Lambda}(t) = \begin{bmatrix} -\boldsymbol{\xi}^T(t) \boldsymbol{\xi}(t) & \lambda \\ \lambda & 0 \end{bmatrix} \quad (28b)$$

$$\boldsymbol{\Psi}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t-1) \boldsymbol{\xi}(t) & \mathbf{z}_{f_1}(t) \end{bmatrix} \quad (28c)$$

$$\mathbf{K}_f(t) = (\boldsymbol{\Lambda}(t) + \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t))^{-1} \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \quad (28d)$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + (\mathbf{g}_f(t) - \mathbf{P}_f^T(t-1) \boldsymbol{\Psi}_f(t)) \mathbf{K}_f(t) \quad (28e)$$

$$\mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t) = \lambda \mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t-1) + \mathbf{z}_{f_1}(t) \boldsymbol{\xi}^T(t) \quad (28f)$$

$$\mathbf{R}_{\mathbf{z}_{f_2}} \boldsymbol{\xi}(t) = \lambda \mathbf{R}_{\mathbf{z}_{f_2}} \boldsymbol{\xi}(t-1) + \mathbf{z}_{f_2}(t) \boldsymbol{\xi}^T(t) \quad (28g)$$

$$\mathbf{L}_f(t) = \frac{1}{\lambda^2} (\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1) \boldsymbol{\Psi}(t) \mathbf{K}_f(t)) \quad (28h)$$

$$\text{with } \mathbf{L}_f(t) = (\mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t))^{-1}. \quad (28i)$$

where  $0 < \lambda \leq 1$  is a forgetting factor.

## 5. CONVERGENCE ANALYSIS OF EIVPM

In this Section, the convergence conditions and the asymptotic distribution of the estimation error of the EIVPM algorithm are derived. For, note that assuming  $\lambda = 1$ , the quadratic criterion (26) is minimized by  $\hat{\mathbf{P}}_f$  such that

$$\hat{\mathbf{P}}_f^T(t) \mathbf{R}(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_2}} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t) \end{bmatrix} \quad (29)$$

with

$$\mathbf{R}(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t) \end{bmatrix}. \quad (30)$$

The right hand side of (29) can be equivalently written as

$$\mathbf{R}_{\mathbf{z}_{f_2}} \boldsymbol{\xi} \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi} = \left[ \frac{1}{t} \sum_{\tau=1}^t \mathbf{z}_{f_2}(\tau) \boldsymbol{\xi}^T(\tau) \right] \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}. \quad (31)$$

Now, writing (19) in the propagator basis, we get

$$\begin{bmatrix} \mathbf{z}_{f_1}(t) \\ \mathbf{z}_{f_2}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} \tilde{\mathbf{x}}(t) + \begin{bmatrix} \tilde{\mathbf{b}}_{f_1}(t) \\ \tilde{\mathbf{b}}_{f_2}(t) \end{bmatrix}. \quad (32)$$

Therefore, it is possible to write  $\mathbf{z}_{f_2}$  in terms of the true propagator  $\mathbf{P}_f^T$

$$\mathbf{z}_{f_2}(t) = \mathbf{P}_f^T \mathbf{z}_{f_1}(t) + \underbrace{(\tilde{\mathbf{b}}_{f_2}(t) - \mathbf{P}_f^T \tilde{\mathbf{b}}_{f_1}(t))}_{\boldsymbol{\eta}(t)}. \quad (33)$$

By introducing this equation in (31), we get

$$\begin{aligned} \mathbf{R}_{\mathbf{z}_{f_2}} \boldsymbol{\xi} \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi} &= \frac{1}{t} \sum_{\tau=1}^t (\mathbf{P}_f^T \mathbf{z}_{f_1}(\tau) + \boldsymbol{\eta}(\tau)) \boldsymbol{\xi}^T(\tau) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi} \\ &= \mathbf{P}_f^T \mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t) + \mathbf{R}_{\boldsymbol{\eta}} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t) \end{aligned} \quad (34)$$

and finally, from (29) and (34) we have

$$(\hat{\mathbf{P}}_f^T(t) - \mathbf{P}_f^T) \mathbf{R}(t) = \mathbf{R}_{\boldsymbol{\eta}} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t). \quad (35)$$

Along the lines of (Ljung and Söderström, 1983, Chapter 4), the convergence analysis of EIVPM is based on the analysis of (35), *i.e.*

- (1) proving that  $(\hat{\mathbf{P}}_f^T(t) - \mathbf{P}_f^T) \mathbf{R}(t) \rightarrow 0$  w.p.1 as  $t \rightarrow \infty$ ;
- (2) deriving conditions under which  $\mathbf{R}(t) \rightarrow \mathbf{R}_0$  as  $t \rightarrow \infty$ , with  $\mathbf{R}_0$  full rank.

These two steps are considered in the following subsections.

### 5.1 Convergence of $(\hat{\mathbf{P}}_f^T(t) - \mathbf{P}_f^T) \mathbf{R}(t)$

From (35), it is easy to establish the following proposition:

*Proposition 5.1.* Consider algorithm (28) and assume that

- the input  $\mathbf{u}$  is uncorrelated with the innovation  $\mathbf{e}$  (system in open loop);
- $\boldsymbol{\xi} \in \mathbb{R}^{n_\xi \times 1}$  ( $n_\xi \geq n_x$ ) is uncorrelated with the noise but sufficiently correlated with the state vector  $\mathbf{x}$ .

Then  $(\hat{\mathbf{P}}_f^T(t) - \mathbf{P}_f^T) \mathbf{R}(t) \rightarrow 0$  w.p.1 as  $t \rightarrow \infty$ .

*Proof 5.1.* Note that from (18)

$$\frac{1}{t} \sum_{\tau=1}^t \tilde{\mathbf{b}}_f(\tau) \boldsymbol{\xi}^T(\tau) = \frac{1}{t} \sum_{\tau=1}^t [\mathbf{G}_f(\mathbf{e}_f(\tau) - \mathbf{E}_f(\tau-1) \mathbf{U}_f^T(\tau-1) \boldsymbol{\beta}_f(\tau)) \boldsymbol{\xi}^T(\tau)] \quad (36)$$

tends to

$$\mathbf{G}_f (\mathbf{R}_{\mathbf{e}_f} \boldsymbol{\xi} - \mathbf{R}_{\mathbf{e}_f \mathbf{u}_f} \mathbf{R}_{\mathbf{u}_f}^{-1} \mathbf{R}_{\mathbf{u}_f} \boldsymbol{\xi}) \quad (37)$$

w.p.1 when  $t \rightarrow \infty$  where

$$\mathbf{R}_{\mathbf{e}_f} \boldsymbol{\xi} = \mathbb{E} [\mathbf{e}_f(\tau) \boldsymbol{\xi}^T(\tau)] = \mathbf{0} \quad (38)$$

$$\mathbf{R}_{\mathbf{e}_f \mathbf{u}_f} = \mathbb{E} [\mathbf{e}_f(\tau-1) \mathbf{u}_f^T(\tau-1)] = \mathbf{0} \quad (39)$$

$$\mathbf{R}_{\mathbf{u}_f} = \mathbb{E} [\mathbf{u}_f(\tau-1) \mathbf{u}_f^T(\tau-1)] \quad (40)$$

$$\mathbf{R}_{\mathbf{u}_f} \boldsymbol{\xi} = \mathbb{E} [\mathbf{u}_f(\tau) \boldsymbol{\xi}^T(\tau)] \quad (41)$$

according to the assumptions on  $\boldsymbol{\xi}$  and  $\mathbf{u}$ . Thus, we have

$$\frac{1}{t} \sum_{\tau=1}^t \tilde{\mathbf{b}}_f(\tau) \boldsymbol{\xi}^T(\tau) \rightarrow 0 \text{ w.p.1 as } t \rightarrow \infty \quad (42)$$

and, by extension

$$\frac{1}{t} \sum_{\tau=1}^t \boldsymbol{\eta}(\tau) \boldsymbol{\xi}^T(\tau) \rightarrow 0 \text{ w.p.1 as } t \rightarrow \infty. \quad (43)$$

## 5.2 Convergence of $\mathbf{R}(t)$

In order to complete the proof, we need to study under which conditions  $\mathbf{R}(t)$  converges to a full rank matrix in order to conclude that  $\hat{\mathbf{P}}_f^T(t) \rightarrow \mathbf{P}_f^T$  w.p.1 as  $t \rightarrow \infty$ . To this purpose, we just need to analyse the rank of  $\mathbf{R}_{\mathbf{z}_{f_1}\xi}$  as  $t \rightarrow \infty$ . Note that writing (19) in the propagator basis gives (see (32))

$$\mathbf{z}_{f_1}(t) = \bar{\mathbf{x}}(t) + \tilde{\mathbf{b}}_{f_1}(t). \quad (44)$$

Then, recalling the definition of  $\bar{\mathbf{x}}$  given in (18),  $\mathbf{R}_{\mathbf{z}_{f_1}\xi}(t)$  can be written as

$$\begin{aligned} \mathbf{R}_{\mathbf{z}_{f_1}\xi}(t) &= \frac{1}{t} \sum_{\tau=1}^t \tilde{\mathbf{b}}_{f_1}(\tau) \boldsymbol{\xi}^T(\tau) + \\ &\frac{1}{t} \sum_{\tau=1}^t \left[ \left( \mathbf{x}(\tau) - \mathbf{X}(\tau-1) \mathbf{U}_f^T(\tau-1) \boldsymbol{\beta}_f(\tau) \right) \boldsymbol{\xi}^T(\tau) \right]. \end{aligned} \quad (45)$$

Since (42) holds, the first term of the right hand side of (45) can be neglected. Therefore, we have that

$$\mathbf{R}_{\mathbf{z}_{f_1}\xi}(t) \rightarrow \mathbf{R}_0^{1/2} = \mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{x}\mathbf{u}_f} \mathbf{R}_{\mathbf{u}_f}^{-1} \mathbf{R}_{\mathbf{u}_f\xi} \quad (46)$$

w.p.1 when  $t \rightarrow \infty$  where

$$\mathbf{R}_{\mathbf{x}\xi} = \mathbb{E} \left[ \mathbf{x}(\tau) \boldsymbol{\xi}^T(\tau) \right] \quad (47)$$

$$\mathbf{R}_{\mathbf{x}\mathbf{u}_f} = \mathbb{E} \left[ \mathbf{x}(\tau-1) \mathbf{u}_f^T(\tau-1) \right] \quad (48)$$

$$\mathbf{R}_{\mathbf{u}_f\xi} = \mathbb{E} \left[ \mathbf{u}_f(\tau) \boldsymbol{\xi}^T(\tau) \right]. \quad (49)$$

According to the Schur Lemma (Golub and Van Loan, 1996), a sufficient condition that guarantees that  $\mathbf{R}_0^{1/2}$  is full rank is given by

$$\begin{aligned} \text{rank} \left\{ \begin{bmatrix} \mathbf{R}_{\mathbf{x}\xi} & \mathbf{R}_{\mathbf{x}\mathbf{u}_f} \\ \mathbf{R}_{\mathbf{u}_f\xi} & \mathbf{R}_{\mathbf{u}_f} \end{bmatrix} \right\} = \\ \text{rank} \left\{ \mathbb{E} \left[ \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}_f(t) \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}(t) \\ \mathbf{u}_f(t) \end{bmatrix}^T \right] \right\} = n_x + f n_u. \end{aligned} \quad (50)$$

If the instrumental variable vector is constructed from past input and output data, such as

$$\begin{aligned} \boldsymbol{\xi}(t) &= [\mathbf{y}_p^T(t) \ \mathbf{u}_p^T(t)]^T \\ &= [\mathbf{y}^T(t-p) \ \dots \ \mathbf{y}^T(t-1) \ \mathbf{u}^T(t-p) \ \dots \ \mathbf{u}^T(t-1)]^T, \end{aligned} \quad (51)$$

$\boldsymbol{\xi}(t) \in \mathbb{R}^{(n_y+n_u)p \times 1}$ , then equation (50) corresponds to the so-called *critical relation* for the consistency of IV subspace identification algorithms, first derived in the classical paper (Jansson and Wahlberg, 1998). In particular, conditions under which (50) holds have been derived in the cited paper and lead to the following general theorem for convergence of the EIVPM algorithm.

*Proposition 5.2.* Consider the EIVPM algorithm (28) and assume that:

- $(\mathbf{A}, \mathbf{B})$  is reachable;
- $\mathbf{K} = 0$  (i.e., no process noise);
- $p \geq n_x$ ;

- the input  $\mathbf{u}$  is persistently exciting of order  $f + p + n_x$ ;
- the input  $\mathbf{u}$  is uncorrelated with the innovation  $\mathbf{e}$  (system in open loop).

Then:

$$\hat{\mathbf{P}}_f^T(t) \rightarrow \mathbf{P}_f^T \text{ w.p.1 as } t \rightarrow \infty \quad (52)$$

In addition, convergence is also guaranteed if  $\mathbf{K} \neq 0$  in a number of special cases (single input systems, white noise input, ARMA input signal), see again (Jansson and Wahlberg, 1998) for details.

## 6. SIMULATION EXAMPLE

Consider the fourth-order system (Van Overschee and De Moor, 1995):

$$\begin{aligned} \mathbf{x}(t+1) &= \begin{bmatrix} 0.67 & 0.67 & 0 & 0 \\ -0.67 & 0.67 & 0 & 0 \\ 0 & 0 & -0.67 & -0.67 \\ 0 & 0 & 0.67 & -0.67 \end{bmatrix} \mathbf{x}(t) \\ &+ \begin{bmatrix} 0.6598 \\ 1.9698 \\ 4.3171 \\ -2.6436 \end{bmatrix} u(t) + \begin{bmatrix} -0.1027 \\ 0.5501 \\ 0.3545 \\ -0.5133 \end{bmatrix} e(t) \end{aligned} \quad (53)$$

$$\begin{aligned} y(t) &= [-0.5749 \ 1.0751 \ -0.5225 \ 0.183] \mathbf{x}(t) \\ &\quad -0.7139 u(t) + 0.9706 e(t). \end{aligned} \quad (54)$$

The input  $u$  and the innovation  $e$  are white Gaussian noises with zero mean and variance 1 and 9 respectively. This leads to a signal to noise ratio (of variances) at the output of 2.2. The initial estimates of the system matrices are randomly generated under the constraint that the absolute value of the maximum eigenvalue of  $\hat{\mathbf{A}}(0)$  is less than 1 (stability requirement). The forgetting factor is fixed at 1 in order to meet the assumptions of the convergence study, while the  $f$  and  $p$  parameters are equal to 6.

A simple way of checking the accuracy of the estimated model consists in comparing the real parts estimated eigenvalues of  $\hat{\mathbf{A}}$  with the true ones, i.e.,  $\pm 0.67$ . Figure 1 illustrates this comparison by showing the trajectories of the estimated eigenvalues obtained with EIVPM. Despite the lack of forgetting, the algorithm provides consistent estimates of the eigenvalues as expected. Figure 2 shows the evolution of the singular values of the matrix  $\mathbf{R}(t)$  on a sizeable lapse of time relatively to the convergence speed of EIVPM. As can be seen from this figure, the smallest singular value of  $\mathbf{R}(t)$  is always nonzero, which reinforces the fact that  $\mathbf{R}(t)$  converges to a nonsingular matrix.

## 7. CONCLUDING REMARKS

The class of recursive subspace identification algorithms of the MOESP class based on the EIV version of the propagator method for signal subspace estimation has been analysed and a convergence proof for the EIVPM algorithm has been derived. A simulation example is used to illustrate the validity of the underlying assumptions.

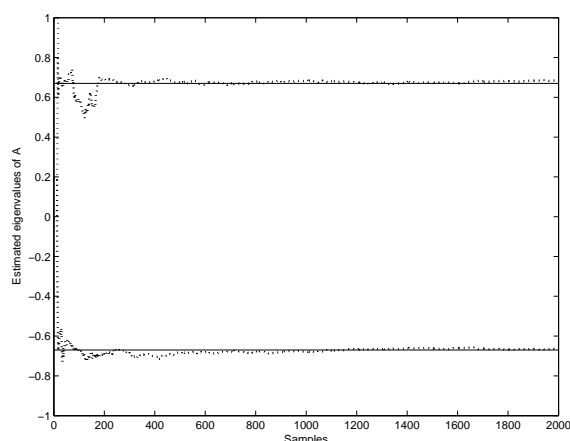


Fig. 1. Real parts of estimated eigenvalues of  $A$  with the EIVPM algorithm.

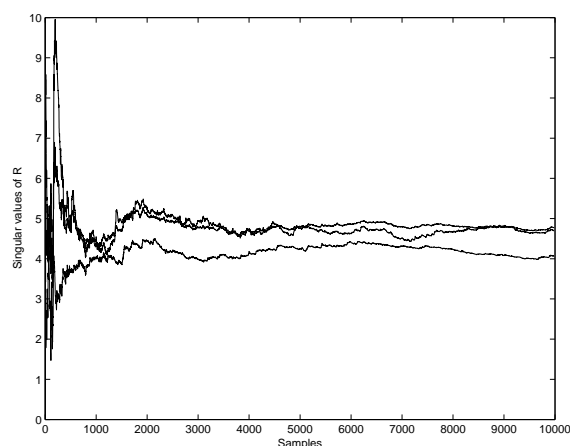


Fig. 2. Singular values of  $R(t)$ .

## 8. REFERENCES

- Cho, Y. M., G. Xu and T. Kailath (1994). Fast recursive identification of state space models via exploitation displacement structure. *Automatica* **30**, 45–60.
- Friedlander, B. (1984). The overdetermined recursive instrumental variable method. *IEEE Transactions on Automatic Control* **4**, 353–356.
- Golub, G. H. and C. F. Van Loan (1996). *Matrix computations*. 3<sup>rd</sup> ed.. John Hopkins University Press. Baltimore MD.
- Gustafsson, T. (1997). Recursive system identification using instrumental variable subspace tracking. In: *11th IFAC Symposium on System Identification*. Fukuoka, Japan.
- Jansson, M. and B. Wahlberg (1998). On consistency of subspace methods for system identification. *Automatica* **34**, 1507–1519.
- Ljung, L. (1999). *System identification. Theory for the user*. 2<sup>nd</sup> ed.. PTR Prentice Hall Information and System Sciences Series. T. Kailath, Series Editor. Upper Saddle River.
- Ljung, L. and T. Söderström (1983). *Theory and practice of recursive identification*. MIT Press. Cambridge.
- Lovera, M. (2003). Recursive subspace identification based on projector tracking. In: *13th IFAC Symposium on System Identification*. Rotterdam, The Netherlands.
- Lovera, M., T. Gustafsson and M. Verhaegen (2000). Recursive subspace identification of linear and non linear Wiener statespace models. *Automatica* **36**, 1639–1650.
- Marcos, S., A. Marsal and M. Benidir (1995). The propagator method for source bearing estimation. *Signal Processing* **42**, 121–138.
- Mercère, G., S. Lecœuche and C. Vasseur (2003). A new recursive method for subspace identification of noisy systems: EIVPM. In: *13th IFAC Symposium on System Identification*. Rotterdam, The Netherlands.
- Mercère, G., S. Lecœuche and C. Vasseur (2004a). Robust adaptation of the propagator method to recursive subspace identification (in french). In: *Conférence Internationale Francophone d'Automatique*. Douz, Tunisia.
- Mercère, G., S. Lecœuche and C. Vasseur (2005). Sequential correlation based propagator algorithm for recursive subspace identification. In: *16th IFAC World Congress*. Praha, Czech Republic.
- Mercère, G., S. Lecœuche and M. Lovera (2004b). Recursive subspace identification based on instrumental variable unconstrained quadratic optimization. *International Journal of Adaptive Control and Signal Processing* **18**, 771–797.
- Munier, J. and G. Y. Delisle (1991). Spatial analysis using new properties of the cross spectral matrix. *IEEE Transactions on Signal Processing* **39**, 746–749.
- Oku, H. and H. Kimura (2002). Recursive 4SID algorithms using gradient type subspace tracking. *Automatica* **38**, 1035–1043.
- Van Overschee, P. and B. De Moor (1995). A unifying theorem for three subspace system identification algorithms. *Automatica* **31**, 1853–1864.
- Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input output data. *Automatica* **30**, 61–74.
- Verhaegen, M. and E. Deprettere (1991). A fast recursive MIMO state space model identification algorithm. In: *30th IEEE Conference on Decision and Control*. Brighton, United Kingdom.
- Yang, B. (1995). Projection approximation subspace tracking. *IEEE Transactions on Signal Processing* **43**, 95–107.