



## Identification of switching systems using change detection technique in the subspace framework

Komi Midzodzi Pekpe, Gilles Mourot, Komi Gasso, José Ragot

### ► To cite this version:

Komi Midzodzi Pekpe, Gilles Mourot, Komi Gasso, José Ragot. Identification of switching systems using change detection technique in the subspace framework. 43rd IEEE Conference on Decision and Control, CDC 2004, Dec 2004, Atlantis, United States. pp.CDROM. hal-00976444

**HAL Id: hal-00976444**

**<https://hal.science/hal-00976444>**

Submitted on 9 Apr 2014

**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.

# Identification of switching systems using change detection technique in the subspace framework

PEKPE Komi Midzodzi<sup>†</sup>, MOUROT Gilles<sup>†</sup>, GASSO Komi<sup>‡</sup> and RAGOT José<sup>†</sup>

<sup>†</sup> Centre de Recherche en Automatique de Nancy - CNRS UMR 7039

CRAN-INPL, 2 Avenue de La Forêt de Haye. 54516 Vandoeuvre-les-Nancy Cedex

<sup>‡</sup> PSI - FRE 2645

INSA Rouen, Avenue de l'Université - BP 08. 76801 Saint-Etienne du Rouvray Cedex

(kpekpe, gmourot, jragot)@ensem.inpl-nancy.fr, kgasso@insa-rouen.fr

**Abstract**—The paper describes an identification technique of switching system. The considered system is represented as a weighted sum of local models. To estimate the switching times, a change detection technique is applied. It provides the weights associated to the local models. The Markov parameters of these models are identified by a subspace method. This calculation can yield similar local models which are merged. The procedure of parameter identification and models merging is repeated until convergence. The performance of the approach is investigated on a simulation example.

**Index Terms**—switching system, identification, detection of abrupt model changes, Markov parameters, subspace method

## I. INTRODUCTION

The paper focuses on switching systems modelled as a weighted combination of linear local state space models. The transition of a model to another is assumed to be a jump. Consequently, the weights associated to the local models are binary : only one model is active all the time. The switching systems constitute an important class of non-linear systems (especially in system analysis and control). A detailed description of existing identification methods is proposed in [5]. Identification of such a system involves typically these tasks : 1) find the number of local models i.e. the number of different dynamics of the system, 2) partition the data into groups corresponding to the different operating regimes (dynamics) of the system, 3) identify the parameters of each local model. In the paper, we assume that the local models are stable and it exists a minimal separation time  $\tau$  between two consecutive switches. This separation time is referred in the literature as *dwell time* [1]. The estimation of the switching instants is realised by coupling a subspace method based on matrix projection with a statistical test. The method gives a coarse classification of the data according to the dynamics of the system. But, as the test leads to a delay for detection, the data immediately before the commutation times cannot be affected with certainty to a particular local model : these data are dubious and are not used to estimate the parameters of the local models. The classification of these data is performed after the parameters estimation. This estimation consists in determining the Markov parameters of the previously detected dynamics and can reveal similar local models which are merged. A new parameter identification is performed. This imbrication of

parameters estimation and models merging is repeated until convergence. A similar approach was proposed in [4] but it is limited to systems without process noises and cannot deal with the delay for detection problem, contrary to the method presented herein.

The section II formulates the switching system identification problem whereas the matrix notations and definitions are exposed in section III. The switching times detection using a subspace method is detailed in the next section. In section V the identification of the Markov parameters of the local models is tackled. The section VI deals with the determination and the merging of similar local models. Finally, the identification of a minimal and balanced realisation of the obtained local models is presented.

## II. PROBLEM FORMULATION

The output of the switching system is represented as a weighted sum of the outputs of local models :

$$y_k = \sum_{s=1}^h p_{s,k} y_{s,k} \quad (1)$$

where  $k$  represents the time index,  $y_k \in \mathbb{R}^\ell$  the output of the system,  $y_{s,k} \in \mathbb{R}^\ell$  the output of local model  $s$  and  $p_{s,k}$  the weight associated to  $y_{s,k}$ . For  $q$  input-output samples, the weights verify the following condition :

$$p_{s,k} \in \{0, 1\} \text{ and } \sum_{s=1}^h p_{s,k} = 1, \quad \forall k \in [1, q] \quad (2)$$

Each local model is supposed to be linear of order  $n_s$  and described by the state space model :

$$\begin{aligned} x_{s,k+1} &= A_s x_{s,k} + B_s u_k + v_{s,k} \\ y_{s,k} &= C_s x_{s,k} + D_s u_k + w_{s,k} \end{aligned} \quad (3)$$

where the process noises  $v_{s,k} \in \mathbb{R}^{n_s}$  and the measurement noises  $w_{s,k} \in \mathbb{R}^\ell$  of the  $s^{th}$  local model are independent gaussian white noises ( $\mathbf{E}(w_{s,k} w_{s,k}^T) = \sigma_{y_s} I_\ell$ ) and uncorrelated with the inputs  $u_k \in \mathbb{R}^m$ .  $x_{s,k} \in \mathbb{R}^{n_s}$  represents the state vector of model  $s$ . The  $h$  local models are assumed to be stable and active during a minimal time  $\tau$  [1].

From  $q$  samples of inputs  $u_k$  and outputs  $y_k$ , the aim is the determination of : - the number  $h$  of local models, - the weights  $p_{s,k}$  corresponding to each local model, - the order  $n_s$  and a realisation  $(A_s, B_s, C_s, D_s)$  of each local model.

### III. SYSTEM MATRICES

The matrix of outputs  $\bar{y}_{k,f}$  of the system and the input block-Hankel matrix  $U_{k,f} \in \mathbb{R}^{mi \times f}$  are defined as :

$$\bar{y}_{k,f} = \begin{pmatrix} y_k & y_{k+1} & \dots & y_{k+f-1} \end{pmatrix} \in \mathbb{R}^{\ell \times f} \quad (4)$$

$$U_{k,f} = \begin{pmatrix} u_k^{(i)} & u_{k+1}^{(i)} & \dots & u_{k+f-1}^{(i)} \end{pmatrix} \quad (5)$$

$$u_k^{(i)} = \begin{pmatrix} u_{k-i+1} \\ u_{k-i+2} \\ \vdots \\ u_k \end{pmatrix} \quad (6)$$

The matrix of local outputs  $\bar{y}_{s,k,f} \in \mathbb{R}^{\ell \times f}$ , the matrix of measurement noises  $\bar{w}_{s,k,f} \in \mathbb{R}^{\ell \times f}$  as well as the block-Hankel matrix of system noises  $V_{s,k,f} \in \mathbb{R}^{n_s i \times f}$  for the  $s^{th}$  local model are defined similarly. The integers  $i$  and  $f$  satisfy the constraints :

$$\begin{cases} mi < f < q \\ 2(n_s + 1) < \min(i/m, i/\ell), \quad 1 \leq s \leq h \end{cases} \quad (7)$$

The latter condition means that we have a sufficient number of Markov parameters in order to identify a realisation of the local models (see section VII).

#### Remark 1

The integer  $f$  is a generic notation. In the model change detection method using subspace method, it is replaced by  $L$  that represents the width of the sliding window. It is equal to  $j$  in the parameters estimation stage where it indicates the number of samples for identifying the local models.

The matrix of the state sequences  $X_{s,k,f} \in \mathbb{R}^{n_s \times f}$  and the matrix of weights  $P_s \in \mathbb{R}^{f \times f}$  of the local model  $s$  are :

$$X_{s,k,f} = \begin{pmatrix} x_{s,k} & x_{s,k+1} & \dots & x_{s,k+f-1} \end{pmatrix} \quad (8)$$

$$P_{s,k,f} = \begin{pmatrix} p_{s,k} & 0 & 0 & 0 \\ 0 & p_{s,k+1} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & p_{s,k+f-1} \end{pmatrix} \quad (9)$$

For simplicity sake, the notations  $V_{s,f}$ ,  $X_{s,f}$ ,  $P_s$  and  $U_f$  will be used instead of  $V_{s,k,f}$ ,  $X_{s,k,f}$ ,  $P_{s,k,f}$  and  $U_{k,f}$ . The observability matrix  $\Gamma_{i,s}$  of local model  $s$  and its controllability matrix  $C_{i,s}$  are respectively defined as :

$$\Gamma_{i,s} = \begin{pmatrix} (C_s)^T & (C_s A_s)^T & \dots & (C_s A_s^{i-1})^T \end{pmatrix}^T \in \mathbb{R}^{\ell i \times n_s} \quad (10)$$

$$C_{i,s} = \begin{pmatrix} B_s & A_s B_s & \dots & A_s^{i-1} B_s \end{pmatrix} \in \mathbb{R}^{n_s \times mi} \quad (11)$$

Finally, the Markov parameter matrices of the determinist part  $H_{s,i}^d \in \mathbb{R}^{\ell \times mi}$  and the stochastic part  $H_{s,i}^v \in \mathbb{R}^{n_s \times mi}$  of the  $s^{th}$  local model are given by :

$$\begin{aligned} H_{s,i}^d &= \begin{pmatrix} C_s A_s^{i-2} B_s & C_s A_s^{i-3} B_s & \dots & C_s B_s & D_s \end{pmatrix} \\ H_{s,i}^v &= \begin{pmatrix} C_s A_s^{i-2} & C_s A_s^{i-3} & \dots & C_s & 0 \end{pmatrix} \end{aligned} \quad (12)$$

### IV. ESTIMATION OF THE WEIGHTS

The determination of the weights of the local models is achieved by a subspace technique of change detection of system dynamics. The technique allows the estimation of the switching times and the determination of the weights.

#### A. Change detection of the system dynamics

From the input-output data, a residual vector is generated using a matrix projection. Indeed, the outputs of the  $s^{th}$  local model can be expressed as a function of the Markov parameters and the input block-Hankel matrix. The influence of the inputs is annihilated by projecting the outputs on a space orthogonal to the inputs. The projection gives a residual matrix whose last column is chosen as residual vector. The mean of the residual vector is null when no switching occurs and different from zero otherwise.

1) *Matrix input-output relation*: let  $s$  the active local model ; its output can be expressed as :

$$\begin{aligned} y_{s,k} &= C_s A_s^{i-1} x_{s,k-i} + C_s A_s^{i-2} B_s u_{k-i} + \dots + C_s B_s u_{k-1} \\ &\quad + D_s u_k + C_s A_s^{i-2} v_{s,k-i} + \dots + C_s v_{s,k-1} + w_{s,k} \end{aligned} \quad (13)$$

As the local model is assumed to be stable, the term  $C_s A_s^{i-1} x_{s,k-i}$  decays for high values of “ $i$ ”. This term is considered negligible if :

$$\|C_s A_s^{i-1} x_{s,k-i}\| < \sqrt{\text{variance}(w_{s,k})} \quad (14)$$

Thus, (13) becomes :

$$\begin{aligned} y_{s,k} &\simeq C_s A_s^{i-2} B_s u_{k-i} + \dots + C_s B_s u_{k-1} + D_s u_k \\ &\quad + C_s A_s^{i-2} v_{s,k-i} + \dots + C_s v_{s,k-1} + w_k \end{aligned} \quad (15)$$

All the local models being active over a duration greater than the *dwell time*  $\tau$ , if  $\tau \geq i + L$  ( $L$  is the width of the sliding window), we can stack  $i + L - 1$  measurements according to the relation (15). This yields :

$$\bar{y}_{s,k,L} \simeq H_{s,i}^d U_L + H_{s,i}^v V_{s,L} + \bar{w}_{s,k,L} \quad (16)$$

Notice that as only the local model  $s$  is active over the interval of width  $i + L - 1$ , the output matrix  $\bar{y}_{k,L}$  of the switching system is equal to the output matrix  $\bar{y}_{s,k,L}$  of the active local model in this interval.

2) *Residual generation*: the influence of the input terms is removed from (16) by projecting the row space of the matrix  $\bar{y}_{s,k,L}$  onto the space orthogonal to the row space of the input block-Hankel matrix  $U_L$  :

$$\bar{y}_{s,k,L} \Pi_{(U_L)^\perp} \simeq H_{s,i}^v V_{s,L} \Pi_{(U_L)^\perp} + \bar{w}_{s,k,L} \Pi_{(U_L)^\perp} \quad (17)$$

The matrix  $\Pi_{(U_L)^\perp}$  is an operator which realises the projection of the row space of a matrix orthogonally to the row space of the matrix  $U_L$ . It is defined as :

$$\Pi_{(U_L)^\perp} = I - U_L^{(-)} U_L \quad (18)$$

where  $U_L^{(-)}$  is the Moore-Penrose pseudo-inverse of matrix  $U_L$ .

The relation (17) provides a residual matrix containing only the noise terms. Any column of this matrix can be chosen as residual vector. In the sequel, we select arbitrarily the last column as residual vector (knowing that other choices can be made). Therefore, the residual vector is given by the following relation :

$$\begin{aligned}\varepsilon_{s,k} &= \bar{y}_{s,k,L} \Pi_{(U_L)^\perp} Z \\ \varepsilon_{s,k} &= H_{s,i}^v V_{s,L} \Pi_{(U_L)^\perp} Z + \bar{w}_{s,k,L} \Pi_{(U_L)^\perp} Z\end{aligned}\quad (19)$$

where  $Z = \begin{pmatrix} 0 & \dots & 0 & 1 \end{pmatrix}^T \in \mathbb{R}^{L \times 1}$  is a selection vector.  $\varepsilon_{s,k}$  is a zero mean white gaussian vector because it is extracted from a linear combination of matrices of zero mean gaussian white noises  $V_{s,L}$  and  $\bar{w}_{s,k,L}$ .

3) *Sensibility of the residual vector to a model switching*: we analyse the expression of the residual vector  $\varepsilon_{s,k}$  when a switch occurs in order to show  $\varepsilon_{s,k}$  has not zero mean. Under the assumption that the change occurs at the time  $r-i+1$ , the relation (16) can also be written (by considering one by one the columns of the matrix  $U_L$ ) :

$$\bar{y}_{s,k,L} \simeq \sum_{t=k}^{r-i} H_{s,i}^d U_L(t) + \sum_{t=r-i+1}^{L+k} \tilde{H}_{s,i}^d(t) U_L(t) + \zeta_s \quad (20)$$

with the following definitions :

- $\zeta_s = H_{s,i}^v V_{s,L} + \bar{w}_{s,k,L}$  : zero mean matrix of noises,
- $\tilde{H}_{s,i}^d(t)$ ,  $t \in [r-i+1, L+k]$  : matrix of Markov parameters after the change of dynamic,
- $U_L(t)$  : matrix with  $t^{th}$  column is equal to  $U_L(:, t)$  ( $t^{th}$  column of the matrix  $U_L$ ) and zeros elsewhere

$$U_L(t) = \begin{pmatrix} 0 & \dots & 0 & U_L(:, t) & 0 & \dots & 0 \end{pmatrix} \quad (21)$$

Then, the relation (20) becomes :

$$\bar{y}_{s,k,L} \simeq H_{s,i}^d U_L + \sum_{t=r-i+1}^{L+k} (\tilde{H}_{s,i}^d(t) - H_{s,i}^d) U_L(t) + \zeta_s \quad (22)$$

and the residual vector is given by :

$$\begin{aligned}\bar{y}_{s,k,L} \Pi_{(U_L)^\perp} Z &\simeq \\ \sum_{t=k-i+s}^{L+k} (\tilde{H}_{s,i}^d(t) - H_{s,i}^d) U_L(t) \Pi_{(U_L)^\perp} Z &+ \zeta_s \Pi_{(U_L)^\perp} Z\end{aligned}\quad (23)$$

It is easy to see from (23) that the residual vector is not zero mean vector if a model change occurs.

### B. Estimation of the switching times and the weights

Previously, it was shown that the residual vector  $\varepsilon_{s,k}$  is a zero mean gaussian vector with variance matrix  $R_\varepsilon$  in absence of model switching. If a change occurs, the mean value of  $\varepsilon_{s,k}$  is not null. Therefore, a  $\chi^2$  test is applied to detect these changes in the mean of the residual vector.

Let  $\mu$  the mean of the residual vector  $\varepsilon_{s,k}$ . The residual vector being gaussian, the term  $\varpi_{s,k} = \varepsilon_{s,k}^T R_\varepsilon^{-1} \varepsilon_{s,k}$  follows a  $\chi^2$  distribution with  $\ell$  degrees of freedom. Thus, it verifies the following test with the confidence threshold  $\alpha$  :

$$\varpi_{s,k} \begin{cases} < \chi_{\ell, \alpha}^2, \text{ no change has occurred } (\mu = 0) \\ \geq \chi_{\ell, \alpha}^2, \text{ a change has occurred } (\mu \neq 0) \end{cases} \quad (24)$$

After the estimation of the switching times, the weights  $p_{s,k}$  of the local models are deduced. Only one local model (model  $s$ ) is active between two consecutive switching times  $k_1$  and  $k_2$ . The classification (according to the operating regimes of the system) of the data collected during the interval  $[k_1, k_2]$  is delicate because of the delay for detection. If  $t_{max}$  is the maximal delay for detection, the data belonging to  $]k_2 - t_{max}, k_2[$  cannot be affected to a particular operating regime. The class of these data corresponds either to the active local model before the commutation (model  $s-1$ ) or to the active model after the switch (model  $s$ ). The classification of these dubious data is performed after the first estimation of the Markov parameter of the local models.

The calculation of the weights on the interval  $[k_1, k_2]$  is carried out as follows :

- the weights of the supposed active local model are set to 1 on  $[k_1, k_2 - t_{max}[$  ( $p_{s,k} \leftarrow 1$ ,  $k_1 \leq k < k_2 - t_{max}$ ),
- the weights of the other local models are set to 0 on the same interval ( $p_{r,k} \leftarrow 0$ ,  $k_1 \leq k < k_2 - t_{max}$ ,  $\forall r \neq s$ ),
- the weights of all the local models are set to 0 for the dubious data ( $p_{r,k} \leftarrow 0$ ,  $k_2 - t_{max} + 1 \leq k < k_2$ ,  $\forall r$ ) and the index  $s$  of the active model is associated to these points.

## V. IDENTIFICATION OF THE MARKOV PARAMETERS OF THE LOCAL MODELS

The determined weights are used to estimate the Markov parameters of the local models. This estimation relies on a matrix input-output relation exposed in the next subsection.

### A. Local input-output matrix relation

The relation allows to express the outputs of a local model according to the inputs and its state vector.

#### Matrix input-output relation

$$C_s A_s^{i-1} X_{s,k-i,j} P_s + H_{s,i}^d U_j P_s + H_{s,i}^v V_{s,j} P_s + \bar{w}_{s,k,j} P_s \quad (25)$$

The proof is given in appendix 1.

We recall that according to (4),  $f$  is equal to  $j$  and verifies the condition (7). In fact, the number of data is  $q = j + i - 1$ .

The estimation of the Markov parameters  $H_{s,i}^d$  is deduced from (25). As the relation is based on the matrix of local outputs  $\bar{y}_{s,k,j}$ , it is necessary to estimate this matrix.

### B. Determination of the local model outputs

The weights verifying (2), the following property holds :

$$(p_{g,k})^2 = p_{g,k} \text{ and } p_{g,k} \times p_{s,k} = 0, \text{ (if } g \neq s) \quad (26)$$

and the local outputs are obtained from the system outputs :

$$p_{g,k} y_k = p_{g,k} \sum_{s=1}^h p_{s,k} y_{s,k} = p_{g,k} y_{g,k} \quad \forall k \in [1, q] \quad (27)$$

By stacking the previous relation, we obtain :

$$\bar{y}_{s,k,j} P_s = \bar{y}_{k,j} P_s \quad (28)$$

### C. Identification of the Markov parameters

From the knowledge of the outputs of a local model, the corresponding Markov parameters are estimated by least squares. The estimation technique is summarised by the following theorem.

**Theorem 1** - Under the assumptions :

- 1) the local model  $s$  is stable,
- 2) the matrix  $U_j P_s$  has full rank,
- 3) the variable  $i$  is sufficient enough such as the term  $C_s A_s^{i-1} X_{s,k-i,j} P_s$  is negligible,

we have :

$$\bar{y}_{s,k,j} P_s U_j^T (U_j P_s U_j^T)^{-1} \xrightarrow{j \rightarrow \infty} H_{s,i}^d \quad (29)$$

**Proof** : the local model  $s$  being stable, the term  $C_s A_s^{i-1} X_{s,k-i,j} P_s$  in (25) is negligible for appropriate values of  $i$  (see section III). This gives :

$$\bar{y}_{s,k,j} P_s \simeq H_{s,i}^d U_j P_s + H_{s,i}^v V_{s,j} P_s + \bar{w}_{s,k,j} P_s$$

As the matrices of noises  $V_{s,j} P_s$  and  $\bar{w}_{s,k,j} P_s$  are uncorrelated with the inputs, an unbiased estimation of the matrix of Markov parameter  $H_{s,i}^d$  is obtained by least squares :

$$\hat{H}_{s,i}^d = \bar{y}_{s,k,j} P_s U_j^T (U_j P_s U_j^T)^{-1} \text{ with } \bar{y}_{s,k,j} P_s = \bar{y}_{k,j} P_s$$

■

### D. Classification of dubious data

Between two consecutive switching times  $k_1$  and  $k_2$  ( $k_1 < k_2$ ), the data belonging to the interval  $]k_2 - t_{\max}, k_2[$  cannot be classified because of the delay for detection. The class of the dubious data is determined now, from the knowledge of the Markov parameters of the local models :

- if  $e_{s,k} \leq e_{s-1,k} \forall k \in ]k_2 - t_{\max}, k_2[$ , then the local model  $s$  is active ( $p_{s,k} \leftarrow 1, p_{r,k} \leftarrow 0 \forall r$ ),
- if  $e_{s,k} > e_{s-1,k} \forall k \in ]k_2 - t_{\max}, k_2[$ , then the local model  $s-1$  is active ( $p_{s-1,k} \leftarrow 1, p_{r,k} \leftarrow 0 \forall r$ ),

where  $e_{r,k} = |y_k - \hat{H}_{r,i}^d u_k^{(i)}|$  is the equation error and  $s$  is the index associated to the unclassified points of the interval  $]k_2 - t_{\max}, k_2[$

## VI. MERGING OF THE SIMILAR LOCAL MODELS

After the first estimation of the local model parameters, we search “similar” local models. The proposed approach consists in comparing the Markov parameter matrices  $H_{s,i}^d$ . Let us estimate the variance matrix (used further in the merging algorithm) of the Markov parameters.

### A. Variance of the Markov parameters $H_{s,i}^d$

An unbiased estimation of the variance  $\sigma_{y,s}^2$  of the output error of the local model  $s$  is given by :

$$\hat{\sigma}_{y,s}^2 = \frac{\text{trace}(\varepsilon_s \varepsilon_s^T)}{i(j - mi)}, \quad \varepsilon_s = \bar{y}_{s,k,j} P_s - \hat{y}_{s,k,j} P_s \quad (30)$$

with  $\hat{y}_{s,k,j}$  the estimation of  $\bar{y}_{s,k,j}$ .

The proof is similar to the one of the lemma II.1 of [3]. The variance  $\Sigma_{H_{s,i}^d}$  of  $H_{s,i}^d$  is obtained from lemma 1.

### Lemma 1

An unbiased estimation of the variance matrix  $\Sigma_{H_{s,i}^d}$  is :

$$\hat{\Sigma}_{H_{s,i}^d} = i \hat{\sigma}_{y,s}^2 [U_j P_s U_j^T]^{-1} \quad (31)$$

### B. Determination of the “similar” local models

The similarity between two local models is measured by the distance between their Markov parameters : the distance between two matrices  $\hat{H}_{s,i}^d$  and  $\hat{H}_{r,i}^d$  is defined as :

$$d(\hat{H}_{s,i}^d, \hat{H}_{r,i}^d) = \text{trace} \left( (\hat{H}_{s,i}^d - \hat{H}_{r,i}^d) \hat{\Sigma}_{H_{s,r}}^{-1} (\hat{H}_{s,i}^d - \hat{H}_{r,i}^d)^T \right) \quad (32)$$

$$\text{with } \hat{\Sigma}_{H_{s,r}} = \frac{N_s}{N_s + N_r} \hat{\Sigma}_{H_{s,i}^d} + \frac{N_r}{N_s + N_r} \hat{\Sigma}_{H_{r,i}^d},$$

where  $N_s = \sum_{k=1}^j p_{s,k}$  and  $N_r = \sum_{k=1}^j p_{r,k}$ .

Considering  $\beta$  as the threshold of a  $\chi^2$  test with  $\ell$  degrees of freedom, the similarity test is described below.

### Similarity test

- $d(\hat{H}_{s,i}^d, \hat{H}_{r,i}^d) \geq \beta$  : the models  $s$  and  $r$  are not similar,
- $d(\hat{H}_{s,i}^d, \hat{H}_{r,i}^d) < \beta$  : the models  $s$  and  $r$  are similar.

### C. Merging of the local models

After the similarity test, the local models which are revealed identical are merged. The Markov parameters of the new (merged) local models are re-estimated, followed by a new search of similar local models. Notice that this parameter estimation gives more accurate results because the number of data used to estimate the parameters is increased due to the merging. Notice also that the application of the similarity test after the merging is useful because it is possible that some similar models are not recognized at the latter stage because of the inaccuracy of their parameters due the insufficient number of data used to identify them. The procedure alternates parameter estimation and similar models merging and is repeated until the number of local models converges.

The set of local models similar to the model  $s_0$  is noted  $\mathbb{E}_{s_0}$ . In the merging algorithm, all the local models belonging to  $\mathbb{E}_{s_0}$  are replaced by a new local model  $s_0^*$ . The weighting function of the model  $s_0^*$  is derived from :

$$p_{s_0^*,k} = \sum_{s \in \mathbb{E}_{s_0}} p_{s,k}, \quad k \in [1, q] \quad (33)$$

The merging technique is summarised by algorithm 1.

### Algorithm 1 (merging of local models)

- **Step 1 (initialisation)** :  $st \leftarrow 0$ .
- **Step 2 (estimation of the Markov parameters)**  
Compute the matrices  $H_{s,i}^d$  using theorem 1.
- **Step 3 (similarity test)**

Apply the similarity test and build the sets  $\mathbb{E}_s$  of similar models .

**- Step 4 (calculation of the weights after merging)**

For each set  $\mathbb{E}_s$ , if  $\text{card}(\mathbb{E}_s) > 1$  then :

- 1)  $st \leftarrow 1$
- 2)  $p_{s,k} \leftarrow p_{s,k}^*$ ; remove the local models  $r$ ,  $r \in \mathbb{E}_s$ ,  $r \neq s$ .

**- step 5 (re-organisation of the index of the remaining local models (merged and original models))**

**- Step 6 (stopping test)**

If  $st = 1$  then go to step 1 otherwise stop the algorithm.

The algorithm permits the automatic determination of the number of local models (or equivalently the number of different dynamics) of the switching system. Let compute now a state space realisation of each local model.

**VII. IDENTIFICATION OF A REALISATION OF THE LOCAL MODELS**

The aim here is the estimation of the order  $n_s$  and a realisation  $(A_s, B_s, C_s, D_s)$  of the local models from the knowledge of their Markov parameters.

The matrix  $D_s$  is extracted directly from the estimation of  $H_{s,i}^d$  (see 12). The Eigenvalue Realization Algorithm (ERA) [2] is used to determine the order and a minimal and balanced realisation.

**Algorithm 2 : ERA**

- Build the Hankel matrices  $\mathcal{H}_{\nu,s}^0$  and  $\mathcal{H}_{\nu,s}^1$  which contain the Markov parameters and are defined by the general relation : (with  $\nu = \text{int}(i/2)$ , the integer part of  $i/2$ ) :

$$\mathcal{H}_{\nu,s}^\gamma = \begin{pmatrix} C_s A_s^\gamma B_s & \cdots & C_s A_s^{\gamma+\nu-1} B_s \\ C_s A_s^{\gamma+1} B_s & \cdots & C_s A_s^{\gamma+\nu} B_s \\ \vdots & \vdots & \vdots \\ C_s A_s^{\gamma+\nu-1} B_s & \cdots & C_s A_s^{\gamma+2\nu-2} B_s \end{pmatrix} \quad (34)$$

Recall that the condition (7) implies that the matrices  $\mathcal{H}_{\nu,s}^0$  and  $\mathcal{H}_{\nu,s}^1$  have a dimension greater than  $n_s$ .

- Perform the singular values decomposition :

$$\mathcal{H}_{\nu,s}^0 = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} \simeq U_1 S_1 V_1^T \quad (35)$$

where  $S_2$  contains the neglected singular values.

- The order of the local model is equal to the number of singular values in  $S_1$ .
- Compute the observability matrix  $\Gamma_{\nu,s} = U_1 S_1^{1/2}$  and the controllability matrix  $\mathcal{C}_{\nu,s} = S_1^{1/2} V_1^T$ .
- The matrix  $C_s$  can be determined from the first  $\ell$  rows of  $\Gamma_{\nu,s}$  (see 10).
- The matrix  $B_s$  can be determined from the first  $m$  columns of  $\mathcal{C}_{\nu,s}$  (see 11).
- The matrix  $A_s$  is given by :  
 $A_s = S_1^{-1/2} U_1^T \mathcal{H}_{\nu,s}^1 V_1 S_1^{-1/2}$ .

**VIII. ILLUSTRATION EXAMPLE**

Consider a switching system which is the sum of two third order models. The changes of the system dynamics are summarised as follows :  $s = 1$  over the intervals  $[1, 999]$  and  $[1800, 2499]$  and  $s = 2$  over the intervals  $[1000, 1799]$  and  $[2500, 3500]$ . The matrices of the two submodels are defined below :

$$A_1 = \begin{pmatrix} 0.32 & 0.31 & 0 \\ -0.32 & 0.31 & 0 \\ 0 & 0 & -0.18 \end{pmatrix}, B_1 = \begin{pmatrix} 0.9 & -0.7 \\ 0.71 & -0.5 \\ 0.8 & 0.47 \end{pmatrix}$$

$$C_1 = \begin{pmatrix} -0.55 & 0.2 & 0.8 \\ 0.45 & 0.3 & 0.58 \end{pmatrix}, D_1 = \begin{pmatrix} 0.97 & 0.63 \\ -0.32 & 0.95 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} -0.1 & -0.4 & 0 \\ 0.5 & -0.4 & 0 \\ 0 & 0 & 0.26 \end{pmatrix}, B_2 = \begin{pmatrix} 0.1 & -0.6 \\ 0.32 & -0.66 \\ 0.3 & 0.82 \end{pmatrix}$$

$$C_2 = \begin{pmatrix} -0.8 & -0.1 & 0.7 \\ 0.3 & 0.48 & 0.9 \end{pmatrix}, D_2 = \begin{pmatrix} 0.5 & 0.3 \\ -0.2 & -0.5 \end{pmatrix}.$$

The inputs  $u_k$  are Pseudo-Random Binary Sequences (PRBS) with variable amplitudes. The signal to noise ratio of the outputs with relation to the measurement noises is 30,3db. The process noises  $v_k$  have a covariance  $\text{var}(v_k) \simeq 7 \times 10^{-4} I_3$ ; the covariance matrix of the inputs is  $\text{var}(u_k) \simeq 7 \times 10^{-2} I_2$ .  $q = 3500$  input-output data points have been processed on the system.

**A. First computation of the weights**

The changes of the system dynamics are detected using the procedure proposed in section IV with  $L = 38$ ,  $i = 13$  and  $t_{\max} = 19$  (half-width of the sliding window). The estimations of the switching times are : 1000, 1801 and 2501. The data belonging to the intervals  $]981, 999[$ ,  $]1782, 1801[$  and  $]2482, 2501[$  are not classified. Therefore, we suppose that the first model is active in the interval  $[1, 981[$ , the second one in  $[1000, 1782]$ , the third model in  $[1801, 2482]$  and the last one in the interval  $[2501, 3500]$ .

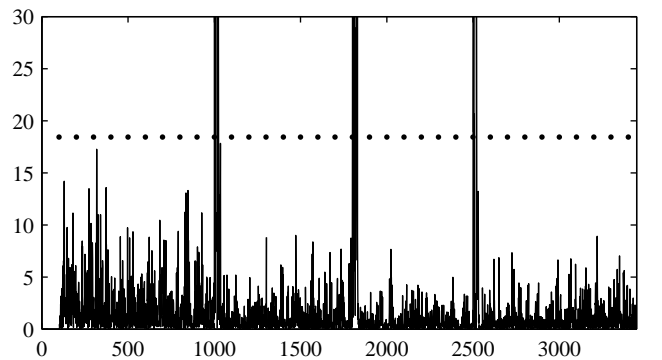


Fig. 1. Estimation of the switching times

Using theorem 1, the Markov parameters of these four local models are computed with  $i = 13$ . Then, the classes of the unclassified points are specified. The poles of each local

model are computed during 100 Monte carlo realisations. The estimated poles are portrayed in figure 2. This figure shows that the local models 1 and 3 and the local models 2 and 4 are similar.

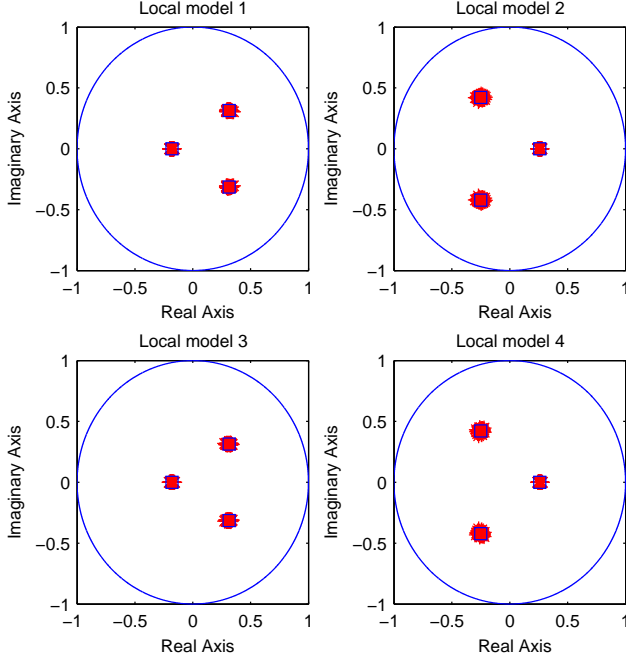


Fig. 2. Poles of the 4 local models for 100 Monte Carlo realisations

### B. Merging of local models (algorithm 1)

The similarity remark is confirmed by the similarity test. For 100 Monte carlo realisations, the following results are obtained :  $d(\hat{H}_{1,i}^d, \hat{H}_{2,i}^d) \in [28000, 38500]$ ,  $d(\hat{H}_{1,i}^d, \hat{H}_{3,i}^d) \in [4, 12]$ ,  $d(\hat{H}_{1,i}^d, \hat{H}_{4,i}^d) \in [27500, 38500]$ ,  $d(\hat{H}_{2,i}^d, \hat{H}_{4,i}^d) \in [3, 13]$  with the threshold  $\beta \simeq 18, 4$ . Therefore, the local models 1 and 3 are merged as well as the models 2 and 4. The Markov parameters of the two remaining models are computed again and their poles are estimated. We can notice the improvement of the accuracy of the poles for 100 Monte carlo realisations (see figure 3).

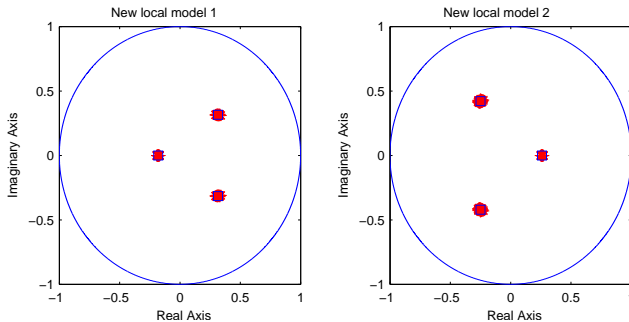


Fig. 3. Poles of the merged local models for 100 Monte Carlo realisations

## IX. CONCLUSION

A subspace identification technique of switching system was presented in the paper. The technique estimates first the switching times and permits an easy determination of the weights associated to the local models. Then, an estimation of the Markov parameters of the detected models is computed. A merging algorithm is proposed to combine the similar local models and is based on the comparison of the Markov parameters. The parameters of the merged models are estimated once. The procedure of models merging and parameters estimation is repeated until the number of models converges. Therefore, a minimal and balanced state space realisation of the remaining models can be determined. The overall approach is illustrated on a simulation example with shows the feasibility of the method. The application of the method on real process data will be studied as a perspective of this work.

## REFERENCES

- [1] J. P. Hespanha and A. S. Morse, Stability of switched systems with average dwell-time, in *38th CDC*, Phoenix, USA, (3), 1999, pp.2655-2660.
- [2] C. Juang. *Applied system identification*, Prentice-Hall, Englewood Cliffs; NJ, 1994.
- [3] L. Ljung. *System Identification*, 2nd Ed. Prentice Hall, New Jersey; 1999.
- [4] K. M. Pekpe, K. Gasso, G. Mourot and J. Ragot, Subspace identification of switching model, in *13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, 2003.
- [5] J. Roll. *Local and Piecewise Affine Approaches to System Identification*, PhD. Thesis, Linköping University, Division of Automatic Control, Sweden, 2003.

## X. APPENDIX 1

The proof of the matrix input-output relation is established herein for deterministic subsystem. The extension to a stochastic system is straightforward. The expression of  $\bar{y}_{s,k,j} \times P_s$  is given by :

$$\begin{aligned} \bar{y}_{s,k,j} \times P_s &= \begin{pmatrix} p_{s,k} y_{s,k} & p_{s,k+1} y_{s,k+1} & \dots & p_{s,k+j-1} y_{s,k+j-1} \end{pmatrix} \\ \bar{y}_{s,k,j} \times P_s &= \begin{pmatrix} p_{s,k} (C_s A_s^{i-1} x_{s,k-i} + C_s A_s^{i-2} B_s u_{k-i} + \dots + D_s u_k) & \dots \\ p_{s,k+j-1} (C_s A_s^{i-1} x_{s,k+j-1-i} + \dots + D_s u_{k+j-1}) \end{pmatrix} \\ \bar{y}_{s,k,j} \times P_s &= \begin{pmatrix} p_{s,k} C_s A_s^{i-1} x_{s,k-i} & \dots & p_{s,k+j-1} C_s A_s^{i-1} x_{s,k+j-1-i} \\ + (p_{s,k} (C_s A_s^{i-2} B_s u_{k-i} + \dots + D_s u_k) & \dots \\ p_{s,k+j-1} (C_s A_s^{i-2} B_s u_{k+j-1-i} + \dots + D_s u_{k+j-1}) \end{pmatrix} \\ \bar{y}_{s,k,j} \times P_s &= C_s A_s^{i-1} \begin{pmatrix} p_{s,k} x_{s,k-i} & \dots & p_{s,k+j-1} x_{s,k+j-1-i} \end{pmatrix} \\ &+ \begin{pmatrix} p_{s,k} (C_s A_s^{i-2} B_s u_{k-i} + \dots + D_s u_k) & \dots \\ p_{s,k+j-1} (C_s A_s^{i-2} B_s u_{k+j-1-i} + \dots + D_s u_{k+j-1}) \end{pmatrix} \end{aligned}$$

Finally, one obtains :

$$\bar{y}_{s,k,j} \times P_s = C_s A_s^{i-1} X_{s,k-i,j} P_s + H_{s,i}^d U_j P_s \quad (36)$$