



**AALBORG UNIVERSITY**  
DENMARK

**Aalborg Universitet**

## **A clustering- based bounded-error approach for identification of PWA hybrid systems**

Tabatabaei-Pour, Mojtaba

*DOI (link to publication from Publisher):*  
[10.1109/ICARCV.2006.345307](https://doi.org/10.1109/ICARCV.2006.345307)

*Publication date:*  
2006

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*

Tabatabaei-Pour, M. (2006). *A clustering- based bounded-error approach for identification of PWA hybrid systems*. Paper presented at Control, Automation, Robotics and Vision, 2006. ICARCV '06. 9th International Conference on, Singapore. <https://doi.org/10.1109/ICARCV.2006.345307>

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### **Take down policy**

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# A Clustering-Based Bounded-Error Approach for Identification of PWA Hybrid Systems

M. Tabatabaei-Pour, M. Gholami, K. Salahshoor  
Department of Automation and Instrumentation  
Petroleum University of Technology  
Tehran, Iran  
Email: mtabatabaei, mgholami, ksalashoor@put.ac.ir

H. R. Shaker  
Department of Electrical Engineering  
Shahid Bahonar University of Kerman  
Kerman, Iran  
Email: h.r.shaker@uk.ac.ir

**Abstract** – A new bounded-error approach for the identification of discrete time hybrid systems in the Piece-Wise Affine (PWA) form is introduced. The PWA identification problem involves the estimation of the number of affine submodels, the parameters of affine submodels and the partition of the PWA map from data. By imposing a bound on the identification error, we formulate the PWA identification problem as a MIN PFS problem (Partition into a MINimum Number of Feasible Subsystems) and propose a greedy clustering-based method for tackling it. The proposed approach yields to better results than the greedy randomized relaxation algorithm used in previous methods. Also, it is not sensitive to the overestimation of model orders and changes in the tuning parameters and therefore finding a right combination of the tuning parameters of the algorithm to get a model with prescribed bounded prediction error is simple.

**Index Terms** – Nonlinear Identification, Hybrid systems, PWA systems, MIN PFS problem, Clustering.

## I. INTRODUCTION

In recent years, the interest in hybrid systems has grown widely. Hybrid models describe processes that evolve according to dynamic equations and logic rules [1]. There are many frameworks for modeling and control of hybrid systems. PieceWise Affine (PWA) systems represent an attractive model structure especially for identification of hybrid systems due to their universal approximation properties and their equivalencies to several classes of hybrid systems [2], [3]. The problem of identification of hybrid systems in the PWA form using experimental data has investigated a few in recent years. The main intricacy of this problem is that the problem of estimating the regions cannot be decoupled from the identification of each submodel. In [4], an overview of different approaches for the identification of PWA systems is presented. Reference [5], proposes an approach that exploits combined use of clustering, linear identification, classification techniques for identification in the PieceWise affine AutoRegressive eXogeneous (PWARX) form. Also the algebraic procedure [6], the adapted weights procedure [7], the Bayesian procedure [8] the k-plane clustering procedure [9],[10], and the bounded-error procedure [11],[12], have been proposed for identification in this form.

In this paper, by imposing a bound on the identification

error, the identification problem is stated as a MIN PFS problem. For tackling the solution of the MIN PFS problem the idea of greedy algorithm [13], which solves the problem by solving consecutive MAX FS subproblems (MAXimum Feasible Subsystems) is used, but instead of using TRR (Randomized and Thermal Relaxation) method proposed in [13] or the modified version [12], we exploit a new clustering-based method. The main drawbacks of both the original [13],[11] and the modified [12] TRR algorithm is that it was observed in extensive trials that both the variance of the results can be quite large (*i.e.*, the number of extracted subsystems may differ considerably from trial to trial), and the average number of extracted subsystems can be rather far from the minimum. Although in the modified version of the algorithm, these properties are reduced, but they are still present. In this paper, we propose a new clustering-based algorithm for tackling the MAX FS problem that does not have the randomness of the TRR method (*i.e.* the variance of the results are zero) and the number of the extracted subsystems are very similar to the minimum. Also, the proposed algorithm is very less sensitive to changes in the tuning parameters and overestimation of model orders and therefore finding a right combination of the tuning parameters of the algorithm to get a model with prescribed bounded prediction error is simple.

After estimating the number of affine submodels and parameters of each affine submodels by solving the corresponding MIN PFS problem, the region estimation are performed using two class [13], [14] or multi-class [15], [16] separation techniques.

The paper is organized as follows. The Identification problem is formulated in section II. In section III, the proposed algorithm is presented and the results are compared with the greedy randomized and thermal relaxation algorithm. In section IV, the effect of changes in its tuning parameters is investigated. Also, the effect of overestimation of models orders are considered and compared with the method proposed in [5] and the modified greedy TRR algorithm [12]. In section V, we discuss the proposed procedure highlighting future research and possible modifications.

## II. PROBLEM FORMULATION

A PWA map is defined by the equation

$$f(\mathbf{x}) = \begin{cases} \varphi' \theta_1 & \text{if } \mathbf{x} \in \mathcal{X}_1 \\ \vdots \\ \varphi' \theta_s & \text{if } \mathbf{x} \in \mathcal{X}_s \end{cases}, \quad (1)$$

where  $\mathcal{X} \subset \mathbb{R}^n$  is a bounded polyhedron,  $\{\mathcal{X}_i\}_{i=1}^s$  is a polyhedral partition of  $\mathcal{X}$  (i.e.,  $\bigcup_{i=1}^s \mathcal{X}_i = \mathcal{X}$ ,  $\mathcal{X}_i \cap \mathcal{X}_j = \emptyset$  if  $i \neq j$  and each region  $\mathcal{X}_i$  and is a convex polyhedron, represented in the form  $\mathcal{X}_i = \{\mathbf{x} \in \mathbb{R}^n : H_i \varphi \leq 0\}$ , where  $H_i \in \mathbb{R}^{q_i \times (n+1)}$ ,  $\varphi$  is the extended vector  $\varphi = [\mathbf{x}' \ 1]'$  and  $\theta_i \in \mathbb{R}^{n+1}$  are parameter vectors. A PWA regression problem amounts to reconstructing a PWA map  $f$  from a finite set of datapoints  $(y_k, \mathbf{x}_k)$ ,  $k=1, \dots, N$  generated by the model

$$y_k = f(\mathbf{x}_k) + \varepsilon_k, \quad (2)$$

where  $\varepsilon_k$  are noise samples, which are Gaussian independent identically distributed random variables with zero mean and variance  $\sigma^2$ .

When considering the PWARX description of (2),  $\mathbf{x}_k$ , the vector of regressors, is denoted by

$$\mathbf{x}_k \triangleq [y_{k-1} \ y_{k-2} \ \dots \ y_{k-n_a} \ \mathbf{u}'_{k-1} \ \mathbf{u}'_{k-2} \ \dots \ \mathbf{u}'_{k-n_b}]', \quad (3)$$

where  $y \in \mathbb{R}$  and  $\mathbf{u} \in \mathbb{R}^m$  represent the input and the output of the system, respectively, and  $n_a$ ,  $n_b$  are the model orders.

The difficulty of the identification problem depends on which quantities in (1) assumed to be known. For example, if we know the set  $\bar{\mathcal{X}}_i$ ,  $i=1, 2, \dots, \bar{s}$  the problem complexity reduces to that of  $\bar{s}$  linear identification problems [5]. Here we consider the problem when the number of submodels,  $s$ , the parameters of each affine submodels,  $\{\theta_i\}_{i=1}^s$ , and the regions of the regressor space,  $\{\mathcal{X}_i\}_{i=1}^s$ , are unknown. During sections II and III, we assume that the model orders  $n_a, n_b$  are known, but in section IV we show that the proposed method is robust against the overestimation of model orders.

The bounded-error identification idea is to impose a bound  $\delta$  on the identification error (the error term between the identified PWARX model and the output of the true system) for all samples  $(y_k, \mathbf{x}_k)$  in the data set. So, the following condition must be satisfied:

$$|y_k - f(\mathbf{x}_k)| \leq \delta, \quad \forall k = 1, \dots, N, \quad (4)$$

for a given bound  $\delta > 0$ . The proposed algorithm does not require to fix the number of submodels *a priori* and estimates it during the identification procedure. If we measure the simplicity of the identified model as the number of submodels,

in order to find a model that is as simple as possible, we must find the minimum  $s$  such that (4) is satisfied. Then the identification problem can be stated as follows:

*Problem 1:* Given  $N$  data points  $(y_k, \mathbf{x}_k)$ ,  $k=1, \dots, N$ , and  $\delta > 0$ , estimate the minimum positive integer  $s$ , a set of parameter vectors  $\{\theta_i\}_{i=1}^s$ , and a polyhedral partition  $\{\mathcal{X}_i\}_{i=1}^s$  of the regressor set  $\mathcal{X}$ , such that the corresponding PWARX model (1)-(3) satisfies condition (4).

Note that choosing the bound  $\delta$  is a trade off between the complexity of the model and the accuracy of fit. The smaller  $\delta$ , the better the accuracy of fit at the cost of the complexity of model (i.e. the larger the number of submodels) and vice versa.

The identification procedure we used in this paper consists of the following steps

- 1) *Initialization*
- 2) *Dealing with undecidable datapoints*
- 3) *Region estimation*

Throughout the paper, the following example will be used in order to illustrate the identification procedure.

*Example 1.* Let the data be generated by the following PWARX system

$$y_k = \begin{cases} [-0.4 \ 1 \ 1.5] \varphi_k + e_k & \text{if } [4 \ -1 \ 10] \varphi_k < 0 \\ [0.5 \ -1 \ -0.5] \varphi_k + e_k & \text{if } \begin{bmatrix} -4 & 1 & -10 \\ 5 & 1 & -6 \end{bmatrix} \varphi_k \leq 0, \\ [-0.3 \ 0.5 \ -1.7] \varphi_k + e_k & \text{if } [-5 \ -1 \ 6] \varphi_k < 0 \end{cases}$$

for which  $\varphi_k = [y_{k-1} \ u_{k-1} \ 1]'$  and  $\bar{s} = 3$ ,  $N = 200$ ,  $\sigma^2 = 0.1$  and the input signal is distributed uniformly in  $[-4, 4]$ . The system and the data points are depicted in Fig. 1.

## III. THE IDENTIFICATION ALGORITHM

### A. Initialization via MIN PFS

The aim of the initialization step is to classify the data points into minimum number of affine submodels and estimate parameter vectors for each affine submodel. The above tasks are carried out by solving the following problem.

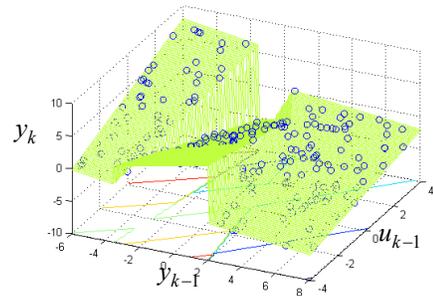


Fig. 1 The PWA systems and the data set for Example 1

TABLE I

THE PROPOSED ALGORITHM FOR THE MINPFS PROBLEM

**Algorithm 1**Given  $\delta > 0, \gamma > 0, c > 0, s = 0$  $\mathcal{D} = \{(y_k, \mathbf{x}_k) \mid k = 1, \dots, N\}$ WHILE  $\#\mathcal{D} > \gamma$ 

For each data point  $(y_i, \mathbf{x}_i)$  build a cluster  $\mathcal{C}_i$  that collects  $(y_i, \mathbf{x}_i)$  and  $c-1$  distinct data point  $(\tilde{y}, \tilde{\mathbf{x}})$  that satisfy

$$\forall (\tilde{y}, \tilde{\mathbf{x}}) \in \mathcal{C}_j, \left\| [y_j \ \mathbf{x}_j] - [\tilde{y} \ \tilde{\mathbf{x}}] \right\|^2 \leq \left\| [y_j \ \mathbf{x}_j] - [\hat{y} \ \hat{\mathbf{x}}] \right\|^2, \\ \forall (\hat{y}, \hat{\mathbf{x}}) \in \mathcal{D} \setminus \mathcal{C}_j$$

For each  $\mathcal{C}_i, i = 1, \dots, N$  find a parameter vector  $\theta_i$  using the robust least squares regression algorithm

For each  $\theta_i, i = 1, \dots, N$ 

$$\mathcal{I}_i = \{(y_k, \mathbf{x}_k) \in \mathcal{D} \mid |y_k - \phi'_k \theta_i| \leq \delta\}$$

$$\theta_{\text{best}} = \arg \max_{\theta_i, i=1, \dots, N} \#\mathcal{I}_i$$

 $t = 0$ WHILE  $\left\| \theta_{\text{best}}^t - \theta_{\text{best}}^{t-1} \right\| = 0$ 

$$\mathcal{I}_{\text{best}}^t = \{(y_k, \mathbf{x}_k) \in \mathcal{D} \mid |y_k - \phi'_k \theta_{\text{best}}^{t-1}| \leq \delta\}$$

For  $\mathcal{I}_{\text{best}}^t$  find  $\theta_{\text{best}}^t$  using the robust least squares regression algorithm

 $t = t + 1$ 

END

 $s = s + 1$ 

$$\theta_s = \theta_{\text{best}}^s$$

$$\mathcal{D}_s = \mathcal{I}_{\text{best}}^s$$

$$\mathcal{D} = \mathcal{D} \setminus \mathcal{D}_s$$

END

*Problem 2.* Given  $\delta > 0$ , find the smallest number of  $s$  of vector  $\theta_i, i = 1, \dots, s$ , and a mapping  $k \rightarrow i(k)$  such that  $|y_k - \phi'_k \theta_{i(k)}| \leq \delta$  for all  $k = 1, \dots, N$ .

In fact, solving the above problem is equivalent to partitioning a (possibly infeasible) system of  $N$  linear complementary inequalities  $|y_k - \phi'_k \theta| \leq \delta, k = 1, \dots, N$ , into a minimum number  $s$  of feasible subsystems considering the condition that two paired complementary inequalities must be simultaneously satisfied by the same parameter vector  $\theta$ . The above problem is known in the literature as MINPFS (Partition in to MINimum number of Feasible Subsystems).

The MINPFS problem is NP-hard. In [13], a greedy algorithm for solving it is proposed. The greedy algorithm subdivides the above problem into a sequence of subproblems. At each subproblem, the problem is to find a parameter vector  $\theta$  that satisfies a largest number of inequalities (MAXimum Feasible Subsystem, MAX FS). These feasible complementary inequalities (the solution of the MAX FS problem) are classified as the first cluster and extracted. Then the MAX FS problem is iteratively repeat on the remaining inequalities until the remaining subsystem is feasible.

The MAX FS problem is still NP-hard. In [10], [11] a modified version of the greedy algorithm for solving MIN PFS problem using a randomized and thermal relaxation method for the MAX FS problem is proposed. The main drawbacks of the modified greedy TRR algorithm is that it was observed in extensive trials that both the variance of the results can be large (*i.e.*, the number of extracted subsystems may differ from trial to trial), and the average number of extracted subsystems can be rather far from the minimum.

Here we propose a new algorithm (shown in Table I) that does not have the above drawbacks. In order to solve the MAX FS problem for PWA maps, we use the idea that a PWA map is locally linear. Therefore small subsets of neighboring points  $(y_k, \mathbf{x}_k)$  are likely to belong to the same submodels. For each datapoint  $(y_k, \mathbf{x}_k)$ , we build a cluster  $\mathcal{C}_i$  that collects the point  $(y_k, \mathbf{x}_k)$  and its  $c-1$  neighboring samples. Then, for each cluster  $\mathcal{C}_i$ , we find a parameter vector  $\theta_i$  using the robust least squares method. Using the parameter vector  $\theta_i, i = 1, \dots, N$ , we create a cluster  $\mathcal{D}_i$  such that:

$$\mathcal{I}_i = \{(y_k, \mathbf{x}_k), k = 1, \dots, N \mid |y_k - \phi'_k \theta_i| \leq \delta\}$$

In other words, the cluster  $\mathcal{I}_i$ , collects those linear complementary inequalities that satisfy the condition  $|y_k - \phi'_k \theta| \leq \delta$  where  $\theta = \theta_i$ . Then we consider the cluster  $\mathcal{I}_i$  that have the maximum cardinality as an initial solution for the MAX FS problem. Then we renew the  $\mathcal{I}_i$  and  $\theta_i$  in the internal WHILE loop (see Table I) repeatedly until there is no change in the parameter vector  $\theta_i$ . Finally, we consider  $\mathcal{I}_i^t$  as the final solution for the MAXFS problem, extract its datapoints and the corresponding inequalities and repeat the

above procedure iteratively until the number of the remaining data points be smaller than a threshold  $\gamma$ , which is tuned by the user as the minimum number of datapoints in a submodel.

Note that to find a parameter vector  $\theta_i$  for a cluster  $\mathcal{C}_i$  and also in updating  $\mathcal{I}_{\text{best}}^t$  and  $\theta_{\text{best}}^t$ , we use the robust least squares algorithm. Consider a mixed cluster  $\mathcal{I}_i$  (a cluster  $\mathcal{C}_i$  or  $\mathcal{I}_i$  is called mixed, if it contains datapoints that are belonged to more than one submodel), which most of its datapoints are belonged to a submodel and some datapoints are belonged to another submodel (outliers). In estimating the parameter vector  $\theta_i$ , the robust least square regression algorithm is less sensitive to outliers in comparison with the ordinary least square regression. Suppose that in the updating step, we use the ordinary least square regression. At each iteration, the result of the regression is affected by the outliers and hence, in the next iteration in constructing the  $\mathcal{I}_i^t$ , more

number of outlier will be added to  $\mathcal{I}_i^t$ , which will destroy the  $\theta_i$  in the next iteration and so on.

It is worthwhile here to point out that the proposed algorithm is deterministic and doesn't involve the randomness such as the TRR method and therefore the variance of the result is zero. We applied both our algorithm and the modified greedy TRR algorithm [12] to the example 1 with  $\delta = 0.6$  over 1000 runs and the results are depicted in Fig. 2. The greedy TRR algorithm extracts a number of feasible subsystems varying between 4 and 8, where the average number of the extracted subsystems is 6 (5.51), whereas our greedy clustering-based approach extracts 3 feasible subsystems for all runs and the number of extracted subsystems are exactly the same as the true system. We fixed parameters for the greedy algorithm as  $C = 20, T_0 = 100, \rho = 0.8$  and for our clustering-based approach as  $\gamma = 10, c = 10$ . However, here we've run the algorithm with previous tuning parameters, but in the next section we will investigate the effect of changes in the tuning parameters and we will see that the proposed algorithm represents a good robustness due to the changes in these parameters.

### B. Dealing with undecidable data

Undecidable datapoints are the data that are consistent with more than one submodels, which means for these points the condition  $|y_k - \phi_k' \theta_i| \leq \delta$  is satisfied for more than one index  $i$  (see Fig. 3.). Algorithm 1, will attribute these datapoints wrongly to a submodel that is extracted earlier than the other ones (the larger cluster). In order to classify these datapoints correctly, we use the spatial localization method proposed in [12].

### C. Region Estimation

During previous steps, we have obtained an estimate of each affine submodel in the PWARX representation. The last step of the identification procedure is to obtain an approximation for the unknown regions  $\{\mathcal{X}_i\}_{i=1}^s$  such that  $\mathbf{x}_k \in \mathcal{X}_i$  if  $(y_k, \mathbf{x}_k) \in \mathcal{D}_i$ . This problem is equivalent to that of separating  $s$  sets of points by means of linear classifiers (hyperplanes), which is widely investigated in the literature. Support Vector Machines (SVM) [14] and Robust Linear Programming (RLP) [15] methods can be used. Also, when the presence of 'holes' in the model can not be accepted, the extension of the above methods to multi-class cases (M-SVM and M-RLP) [16], [17] can be employed.

The final results for the estimated parameter vectors are shown in the Table II and the estimated separating hyperplanes are shown in Table III and depicted in Fig. 4.

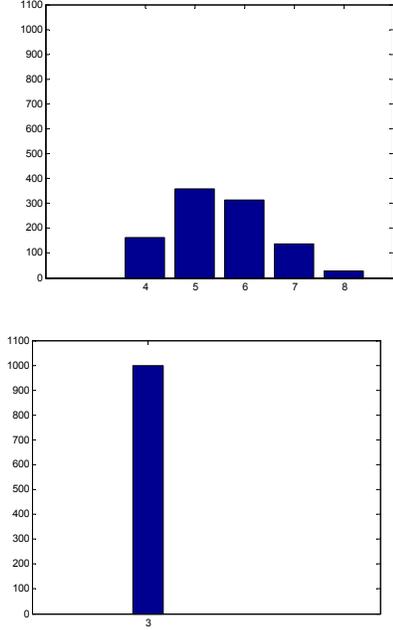


Fig. 2. Number of extracted feasible subsystems by application of the greedy TRR algorithm (Top) and the greedy clustering-based algorithm (Bottom) over 1000 runs for the MINPFS problem in example 1.

## IV. CHOOSING THE TUNING PARAMETERS AND THE EFFECT OF OVERESTIMATION OF MODEL ORDERS

In this section, we use the following overall quality measure introduced in [18] to investigate the effect of both the changes in the tuning parameters and overestimating of the model orders on the performance of the algorithm 1.

$$\sigma_\varepsilon^2 = \frac{1}{s} \sum_{i=1}^s \frac{\text{SSR}_{\mathcal{F}_i}}{|\mathcal{F}_i|}, \quad (5)$$

where set  $\mathcal{F}_i$  contains the datapoints classified to submodel  $i$  and the sum of squared residuals (SSR) of submodel  $i$  is defined as:

$$\text{SSR}_{\mathcal{F}_i} = \sum_{\mathbf{x}_k \in \mathcal{F}_i} (y_k - \phi_k' \theta_i)^2. \quad (6)$$

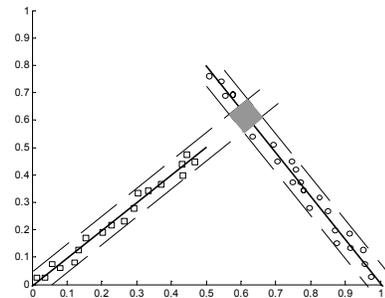


Fig. 3. A PWA map with two discrete modes. All datapoints in the grey area are undecidable for a fixed  $\delta$ .

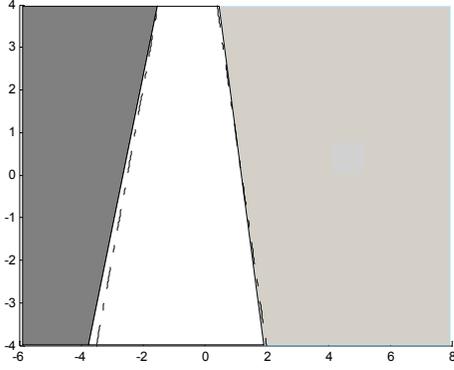


Fig. 4. The true (dashed lines) and the estimated (solid lines) partition of the regressor set in Example 1.

TABLE II

TRUE ( $\tilde{\theta}_i$ ) AND ESTIMATED ( $\theta_i$ ) PARAMETER VECTORS IN EXAMPLE 1

$\tilde{\theta}_1$	$\theta_1$	$\tilde{\theta}_2$	$\theta_2$	$\tilde{\theta}_3$	$\theta_3$
-0.4	-0.4232	0.5	0.4828	-0.3	-0.3060
1.0	0.9928	-1	-1.004	0.5	-0.4941
1.5	1.4543	-0.5	-0.4907	-1.7	-1.5936

TABLE III

TRUE ( $\tilde{h}_i$ ) AND ESTIMATED ( $h_i$ ) SEPARATING HYPERPLANES

$\tilde{h}_1$	$h_1$	$\tilde{h}_2$	$h_2$
-0.4	-0.4232	0.5	0.4828
1.0	0.9928	-1	-1.004
1.5	1.4543	-0.5	-0.4907

Note that the value of the  $\sigma_\varepsilon^2$  (averaged sum of squared residuals) should be small and about the expected noise of the identified system.

#### A. Choosing the tuning parameters

In this section, we will explain how to choose the tuning parameters and the effect of changes in the tuning parameters on the performance of the algorithm 1.

1) *On the choice of  $c$* : We suppose that  $c > n + 1$ . Because the robust least squares algorithm needs at least  $n + 2$  datapoin to perform the regression. The algorithm is not very sensitive to  $c$  because it will modify the selected cluster  $\mathcal{I}_i$  in the internal WHILE loop. On the other hand,  $c$  should not be too large such that all the clusters  $\mathcal{C}_i$  be "mixed" clusters.

For example 1, which the cardinality of the smallest subsystem is 43, we changed  $c$  from  $c = 4$  ( $n + 2$ ) to  $c = 40$ . The number of extracted feasible subsystems is 3 and  $\log \sigma_\varepsilon^2 \approx -1$  (which is consistent with the variance of noise  $\sigma^2 = 0.1$ ) for all the above values of  $c$  ( $\delta = 0.6$ ).

2) *On the choice of  $\gamma$* : The parameter  $\gamma$  is a guess of the minimum number of datapoints in the smallest submodel. For too small value of  $\gamma$ , algorithm 1 will extract a number of feasible subsystems more than the minimum number of

feasible subsystems such that some extracted subsystems contains only a few datapoints. On the other hand, we should not choose  $\gamma$  too large such that it becomes larger than the cardinality of the smallest submodel. In this case, we will loose a submodel.

For example 1, we fixed  $c = 10$ ,  $\delta = 0.6$  and changed  $\gamma$  from 10 to 40. The number of extracted feasible subsystems is 3 and  $\log \sigma_\varepsilon^2 = -1$  for all the above values of  $\gamma$ . But by choosing  $\gamma = 45$ , the number of extracted feasible subsystem is 2 and it is expectable because cardinality of the smallest subsystem is 43.

3) *On the choice of  $\delta$* : As we have explained previously, tuning the error bound  $\delta$  is a trade off between model complexity and quality of fit. Fore more explanation on the choice of  $\delta$  we refer the interested reader to [12].

#### B. Overestimation of model orders

During previous sections we assumed that the system orders  $n_a$  and  $n_b$  are known exactly, but usually in the practice this is not the case. In order to investigate the effect of overestimating the model orders, we will consider the following example with  $\sigma_\varepsilon^2 = 0.01$ ,  $y(0) = -10$ ,  $u(k) \sim \mathcal{U}[-10, 10]$ .

$$y(k+1) = \begin{cases} 2y(k) + 10 + e(k), & \text{if } y(k) \in [-10, 0) \\ -1.5y(k) + 10 + e(k), & \text{if } y(k) \in [0, 10] \end{cases}$$

The true model orders are  $n_a = 1$ ,  $n_b = 0$ . Identification procedure were applied for all combinations of  $n_a = 1, \dots, 4$  and  $n_b = 1, \dots, 5$ . We have used this example from [18] to be able to compare our results with the results reported in it. By applying algorithm 1 to the above example, it returns  $s = 2$  and  $\sigma_\varepsilon^2 \approx 0.01$  for all of the aforementioned combinations. Fig. 4 shows the values of the criterion  $\sigma_\varepsilon^2$  on the logarithmic scale for our procedure. The results show that overestimating the model orders doesn't spoil the performance of the algorithm. The results in [18] shows that the performance of the k-means clustering-based procedure [5] or the Bayesian procedure [8] rapidly declines by overestimating the model orders whereas the greedy TRR algorithm have no problem with the overestimated model orders.

#### V. CONCLUSION AND FURTHER RESEARCHES

In this paper a new procedure for solving the MINPFS problem for bonded-error identification of PWARX systems has been presented. The proposed algorithm doesn't have the randomness of the randomized and thermal relaxation method in extracting the feasible subsystems and therefore the variance of the results is zero and the number of the extracted subsystems is very similar to the minimum. Also, tuning the parameter of the algorithm is easy and the proposed procedure is not sensitive to overestimating of the model orders.

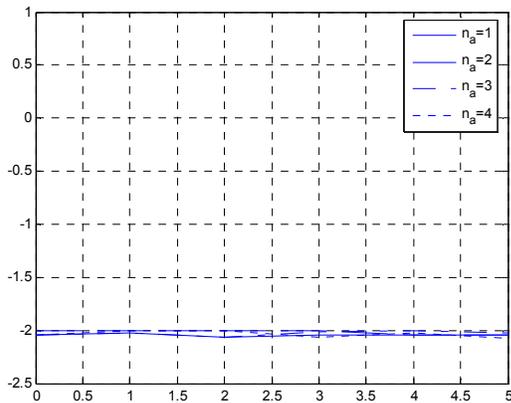


Fig. 5.  $\sigma_\epsilon^2$  for all combinations of  $n_a = 1, \dots, 4$  and  $n_b = 1, \dots, 5$

Further research will concern on the modifications on the algorithm to be able to identify separated regions with same parameter vectors as well as to identify submodels of different orders for each discrete mode.

The problem of choosing an appropriate input signal for identification of hybrid system is an open issue for further research. The main difficulty of this problem is that the input signal should be designed such that all reachable modes are excited sufficiently.

#### REFERENCES

- [1] P.J. Antsaklis, "A brief introduction to the theory and applications of hybrid systems," *Proc. IEEE, Special Issue on Hybrid Systems: Theory and Applications*, vol. 88, no. 7, pp 879–886, July 2000.
- [2] A. Bemporad, G. Ferrari-Trecate, and M. Morari, "Observability and controllability of piecewise affine and hybrid systems," *IEEE Trans. Autom. Contr.*, vol. 45, no. 10, pp. 1864–1876, 2000.
- [3] W. P. M. H. Heemels, B. De. Schutter, and A. Bemporad, "Equivalence of hybrid dynamical models," *Automatica*, vol. 37, no. 7, pp. 1085–1091, 2001.
- [4] J. Roll, "Local and piecewise affine approaches to system identification," Ph.D. dissertation, Department of Electrical Engineering, Linköping University, Sweden, 2003.
- [5] G. Ferrari-Trecate, M. Muselli, D. Liberati, and M. Morari, "A clustering technique for the identification of piecewise affine systems," *Automatica*, vol. 39, no. 2, pp. 205–217, 2003.
- [6] R. Vidal, S. Soatto, Y. Ma, and S. Sastry, "An algebraic geometric approach to the identification of a class of linear hybrid systems," in *Proc. 42nd IEEE Conf. Decision Contr.*, pp. 167–172, 2003.
- [7] J. Ragot, G. Mourt, and D. Maquin, "Parameter estimation of switching piecewise linear systems," in *Proc. 42nd IEEE Conf. Decision Contr.*, pp. 5783–5788, 2003.
- [8] A. Lj. Juloski, S. Wieland, and W. P. M. H. Heemels, "A Bayesian approach to identification of hybrid systems," in *Proc. 43rd IEEE Conf. on Decision and Cont.*, 2004.
- [9] M. Tabatabaei-Pour, K. Salahshoor, and B. Moshiri, "A modified k-plane clustering algorithm for identification of hybrid systems", to be appeared in *Proc. 6th World Congress on Intelligent Control and Automation*, 2006
- [10] M. Tabatabaei-Pour, K. Salahshoor, and B. Moshiri, "Identification of Piecewise Affine Hybrid Systems via k-plane Clustering algorithm", Accepted for publication in *WSEAS Transaction on Systems*, 2006.

- [11] A. Bemporad, A. Garulli, S. Paoletti, A. Vicino, "A greedy approach to identification of piecewise affine models" in *Hybrid Systems: Computation and Control. LNCS. Springer Verlag*, pp. 97-112, 2003.
- [12] A. Bemporad, A. Garulli, S. Paoletti, A. Vicino, "A bounded-error approach to piecewise affine system identification," *IEEE Trans. Automatic Control*, vol. 50, no. 10, pp. 1567-1580, 2005
- [13] E. Amaldi and M. Mattavelli, "The MIN PFS problem and piecewise linear model estimation," *Discrete Appl. Math.*, Vol.118, pp. 115-143, 2002
- [14] J. Roll, A. Bemporad, and L. Ljung, "Identification of piecewise affine systems via mixed-integer programming," *Automatica*, vol. 40, pp. 37–50, 2004.
- [15] K. P. Bennett and O. L. Mangasarian, "Robust linear programming discrimination of two linearly inseparable sets," *Optim. Methods Soft.*, vol. 1, pp. 23–34, 1992.
- [16] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.
- [17] K. P. Bennett and O. L. Mangasarian, "Multicategory discrimination via linear programming," *Optim. Methods Soft.*, vol. 3, pp. 27–39, 1994.
- [18] E. J. Bredensteiner and K. P. Bennett, "Multicategory classification by support vector machines," *Comput. Optim. Appl.*, vol. 12, pp. 53–79, 1999.
- [19] A.Lj. Juloski, W.P.M.H. Heemels, G. Ferrari-trecate, R. Vidal, S. paoletti, and J.H.G. Nissen, "Comparision of four procedures for the identification of hybrid systems," in *Proc. of 8th international workshop on hybrid systems: Computation and Control.* vol. 3414 of LNCS, pp. 354-369, Springer-Verlag, 2005.