

Clustering-based Identification of MIMO Piecewise Affine Systems

Nikola Hure, Mario Vašak

Laboratory for Renewable Energy Systems

University of Zagreb, Faculty of Electrical Engineering and Computing

Unska 3, HR-10000 Zagreb, Croatia

E-mail: nikola.hure@fer.hr, mario.vasak@fer.hr

Abstract—Piecewise Affine (PWA) models are used to approximate general nonlinear dynamics with an arbitrary precision. PWA model can be employed for a constrained optimal controller synthesis, whereas the complexity of the controller is in a large part determined with a complexity of the model. Among the prominent methods for a PWA system identification is the clustering-based identification, which is originally designed for identification of systems with a Multiple-Input Single-Output (MISO) structure. When applied for the Multiple-Input Multiple-Output (MIMO) system identification, previously used clustering-based approach implied independent estimation of PWA maps for each of the outputs, whereas the MIMO PWA model was constructed by merging the polyhedral partitions and parameters of each MISO model. PWA model obtained with the respective approach often contained a significant number of submodels, thus aggravating the controller design process. In this paper we propose a multivariate linear regression approach for the identification of a MIMO PWA model based on the clustering technique. The presented approach is a systematic extension and fully exploits all benefits of the clustering-based identification. The proposed approach is validated on a coupled MIMO system identification problem.

I. INTRODUCTION

Piecewise Affine (PWA) systems are defined with a set of affine submodels and convex regions that associate current state of the system with the corresponding affine submodel. PWA model identification is a challenging task since it involves simultaneous identification of regions in the regressor space and parameters of associated affine submodels. Algorithms that make the state of the art in the piecewise affine system identification are algebraic, Bayesian, bounded-error and clustering-based method [1].

Among the established methods for the piecewise affine identification, the clustering-based technique has shown to be robust to the effect of the measurement noise, but requires *a priori* knowledge about the model orders for the considered system. The method as presented in [2] requires the user to specify the number of submodels and number of points per local dataset which is used to estimate the local parameters of the model. Each local dataset is parameterised as a feature vector, which is afterwards used in the data classification. In [3] authors present a statistical clustering technique that employs consistent Akaike's Information Criterion (AIC) and the Minimum Description Length (MDL) criterion in order to estimate the number of submodels. The survey about the

piecewise affine system identification and its applications can be found in [4].

Clustering-based identification is originally designed for the systems with MISO structure. However, systems that are used for the optimal controller synthesis are often of MIMO structure and require MIMO PWA map identification. With that aim, natural extension of the clustering based technique is presented in this paper. The proposed approach is suitable for the identification of a MIMO PWA model.

The paper is organized as follows. In Section II, PWA identification problem in a general form is presented. In the following section, steps of the clustering-based identification procedure are described and application of the K-means++ algorithm [5] is proposed in order to improve the results of the feature vector clustering phase. In the same Section, feature vector transformation is introduced to reduce or completely omit partitioning in arbitrary dimensions of regressor vectors, e.g. when certain system variables are known to map linearly with respect to the system output. In Section IV, extension of the clustering-based procedure suitable for the MIMO PWA system identification is presented. Section V presents the results of the proposed approach on the coupled MIMO system identification problem.

II. PROBLEM DESCRIPTION

Assume that a dataset of N input-output data $\mathcal{S} = \{(x_i, y_i)\}_{i=1}^N$ is collected, where input vector x is drawn from the bounded multi-dimensional domain $\mathcal{X} \subset \mathbb{R}^n$ and output vector $y \in \mathbb{R}^m$ is generated according to a PWA model,

$$y_i = F(x_i) + \varepsilon_i, \quad (1)$$

where $\varepsilon_i \in \mathbb{R}^m$ is the noise vector sampled from the multivariate Gaussian distribution, whereas $E[\varepsilon_i \varepsilon_j'] = \mathbf{0}$, $\forall i \neq j$. The Piecewise Affine map F is in the general case defined with the equation

$$F(x) = \begin{cases} \theta_1 \begin{bmatrix} x \\ 1 \end{bmatrix} & \text{if } x \in \mathcal{X}_1, \\ \vdots \\ \theta_s \begin{bmatrix} x \\ 1 \end{bmatrix} & \text{if } x \in \mathcal{X}_s, \end{cases} \quad (2)$$

where θ is the parameter matrix of appropriate dimensions and \mathcal{X}_j is the j -th region of the input regressor domain \mathcal{X} which consists of disjoint convex sets, i.e. $\mathcal{X} = \cup_{j=1}^s \mathcal{X}_j$, $\mathcal{X}_i \cap \mathcal{X}_j = \emptyset$, $\forall i \neq j$.

In case of MIMO PWARX systems, inputs form regression vectors that are sampled in equidistant time instants

$$x_k = \begin{bmatrix} y_{(1),k-1} \cdots y_{(1),k-n_a,1} \cdots y_{(m),k-1} \cdots y_{(m),k-n_a,m} \\ u_{k-1}^\top \cdots u_{k-n_b}^\top \end{bmatrix}^\top, \quad (3)$$

where k is the discrete-time instant, $y_{(h)}$ is the h -th output, $u \in \mathbb{R}^r$ is the input vector of the model, n_a and n_b are model orders, whereas dimension of the regressor vector is $n = \sum_i n_{a,i} + r \cdot n_b$, $i = 1, \dots, m$.

PWA regression problem refers to reconstructing a PWA map F , which includes both partitioning of the input domain in s disjoint submodels $\mathcal{X}_1, \dots, \mathcal{X}_s$ and identifying parameters over respective submodels $\theta_1, \dots, \theta_s$.

III. CLUSTERING-BASED IDENTIFICATION OF PIECEWISE AFFINE SYSTEMS

Clustering-based identification [2] is presented for the following scenario:

- i) dimension of the system output vector m is 1 – suitable for MISO identification;
- ii) the number of submodels s is provided;
- iii) in the identification of PWARX models, model orders n_a and n_b are fixed and known *a priori*.

The main steps of the clustering-based identification are briefly discussed in the sequel. The notation which is used for the description of the clustering-based procedure is defined as follows. For MISO identification we consider the one-dimensional output $z \equiv y_{(h)}$, which is the h -th output of the MIMO model. Likewise, local datasets formed with one-dimensional output z are denoted with $\mathcal{C}_{(h),i}$ etc. For more details regarding the clustering-based identification algorithm refer to [2], [6]. The main steps of the algorithm are summarised in the following subsections.

A. Local Regression

Local Datasets (LDs) $\mathcal{C}_{(h),i}$, $i = 1, \dots, N$ are built. Every LD $\mathcal{C}_{(h),i}$ contains a datapoint $(x(i), z(i))$ and $c - 1$ distinct datapoints (\tilde{x}, \tilde{z}) nearest according to the euclidean metrics,

$$\|x(i) - \tilde{x}\|^2 \leq \|x(i) - \bar{x}\|^2, \quad \forall (\tilde{x}, \tilde{z}) \in \mathcal{S}_{(h)} \setminus \mathcal{C}_{(h),i}. \quad (4)$$

For each LD $\mathcal{C}_{(h),i}$ the affine model is identified by employing the least squares method,

$$\theta_{(h),i}^{LS} = (\Phi_i^\top \Phi_i)^{-1} \Phi_i^\top z_{\mathcal{C}_{(h),i}}, \quad (5)$$

$$\Phi_i = [\varphi_1 \varphi_2 \cdots \varphi_c]^\top,$$

where $\varphi_i = [x_i^\top \ 1]^\top$, $\{x_i\}_{i=1}^c$ are regressor vectors and $z_{\mathcal{C}_{(h),i}}$ is the vector of outputs contained in LD $\mathcal{C}_{(h),i}$. Parameter c should be properly tuned in order to avoid outliers in the parameter vectors $\theta_{(h),i}^{LS} \in \mathbb{R}^n$. Outliers are primarily a consequence of mixed LDs. In the clustering-based identification,

term mixed LDs refers to the local datasets containing output values generated by more than one submodel. Accordingly, local datasets containing only data points generated by one submodel are referred to as pure. In order to achieve good validation results, the ratio between mixed and pure LDs must be kept low.

B. Feature vectors construction

Each LD is parameterized as a feature vector $\xi_{(h),i} = [(\theta_{(h),i}^{LS})^\top m_i^\top]^\top$, $\forall i = 1, \dots, N$, where m_i is the mean value of regression vectors contained in a LD $\mathcal{C}_{(h),i}$,

$$m_i = \frac{1}{c} \sum_{(x,z) \in \mathcal{C}_{(h),i}} x, \quad i = 1, \dots, N. \quad (6)$$

In order to include confidence levels in the identification process, feature vectors are approximated with the realisation of Gaussian random vectors with mean ξ_i and variance

$$R_{(h),i} = \begin{bmatrix} V_{(h),i} & \mathbf{0} \\ \mathbf{0} & Q_i \end{bmatrix}, \quad (7)$$

where $V_{(h),i}$ is the empirical covariance matrix of the parameter vector $\theta_{(h),i}^{LS}$ [7],

$$V_{(h),i} = \frac{z_{\mathcal{C}_{(h),i}}^\top (I - \Phi_i (\Phi_i^\top \Phi_i)^{-1} \Phi_i^\top) z_{\mathcal{C}_{(h),i}} (\Phi_i^\top \Phi_i)^{-1}}{c - (n + 1)}, \quad (8)$$

where I is the identity matrix of suitable dimension and Q_i is the scatter matrix of the regression vectors in the LD $\mathcal{C}_{(h),i}$,

$$Q_i = \sum_{(x,z) \in \mathcal{C}_{(h),i}} (x - m_i)(x - m_i)^\top. \quad (9)$$

Moreover, scalar measure of the feature vector ξ_i confidence level is determined by employing

$$w_{(h),i} = \frac{1}{\sqrt{(2\pi)^{2n+1} \det(R_{(h),i})}}. \quad (10)$$

C. Clustering the feature vectors

If the identification problem is well posed, a PWA model with s submodels will result with s clusters of points in the feature vector space. Consequently, feature vectors are partitioned into s disjoint clusters $\{\mathcal{F}\}_{j=1}^s$ that minimise the cost functional quantifying displacement of the feature vectors from the identified centers of the clusters, whereas the centers of the clusters are optimised by employing a suitable algorithm. In [2] authors propose a variation of the batch K-means algorithm that exploits measures of confidence levels to reduce negative effects of outliers in the clustering process. Respective algorithm is computationally efficient but can be trapped in local minima in the case of a bad clustering centers initialisation whereas various algorithms have been proposed which guarantee optimality of the clustering procedure [6], [8].

Adapted K-means++ algorithm for the clustering centers initialisation. Algorithms that guarantee optimality of the clustering procedure exploit measure of a distance between

the sets of data within the dataset and require parameters which do not include the desired number of clusters. On the contrary, K-means allows specifying the desired number of clusters, whereas its optimality depends upon the clustering centers initialisation. In order to improve initialisation step of the K-means, here we adapt the K-means++ algorithm [5], which is defined as follows. Let $D(\xi_i)$ denote the shortest distance from a data point ξ_i to the closest center we have already chosen. The iterations of the dataset center selection are performed with the following steps of the initialisation algorithm:

- 1) take one center c_1 , chosen uniformly at random from the clustering dataset,
- 2) take new center c_j , whereas each data point ξ_i in the clustering dataset has the following probability to be drawn

$$p_i = \frac{(w_i D(\xi_i))^2}{\sum_{i=1, \dots, N} (w_i D(\xi_i))^2}, \quad (11)$$

- 3) if $j < s$, $j = j + 1$ and go to 2.

Transformation of the feature vectors for the discriminatory clustering. K-means clustering results with data clusters that minimise cumulative distance between the appurtenant feature vectors and center of the cluster. Note that the feature vector is comprised of the parameter vector and the mean of the regressors contained in a LD. In the specific identification scenarios, it may be useful to discriminate partitioning with respect to the spread of points in arbitrary dimensions, e.g. when mode switching does not occur with respect to certain variables of the system. Accordingly, one can employ the transformation of the regressor means m_i comprised in the feature vectors prior the identification process,

$$m'_i = T(m_i)m_i, \quad (12)$$

where $T(m_i)$ is an arbitrary transformation matrix.

D. Estimation of the submodel parameters

Since there is a bijective map between the i -th pair (x_i, z_i) and its feature vector ξ_i , data points $\{(x_i, z_i)\}_{i=1}^N$ can be classified in s data clusters $\{\mathcal{D}_{(h),j}\}_{j=1}^s$ according to

$$(x_i, z_i) \in \mathcal{D}_{(h),j}, \text{ if } \xi_i \in \mathcal{F}_{(h),j}. \quad (13)$$

The j -th parameter vector $\theta_{(h),j}$ is then obtained by employing the weighted least squares algorithm over the data in the set $\mathcal{D}_{(h),j}$, whereas the weights are chosen according to (10).

E. Estimation of the submodel regions

Region estimation problem consists in finding complete polyhedral partition of the regressor set $\mathcal{X}_{(h)}$, i.e. without "holes". The convex polyhedral regions $\mathcal{X}_{(h),j}$, $j = 1, \dots, s$, are found by solving a multicategory classification problem, a Multi-category Robust Linear Program (MRLP) [6] in particular. Clustered regressors are bounded with separating hyperplanes between clusters which are selected so as to minimize the cost function associated with misclassified regressors, using a single Linear Program (LP). Since there are large

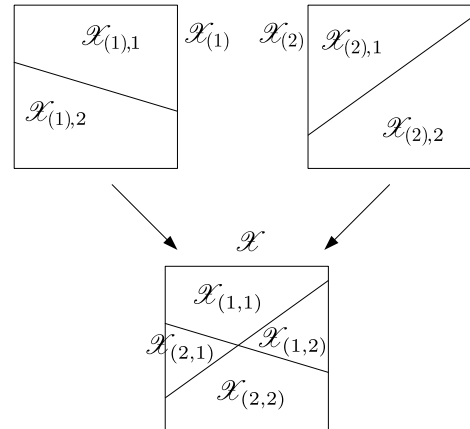


Fig. 1. Merging MISO partitions $\mathcal{X}_{(1)}$ and $\mathcal{X}_{(2)}$ into single MIMO partition \mathcal{X} [9]

memory requirements for solving respective LP which limits its applicability for larger dataset of regressors, a simplified problem is proposed in [9] to estimate the borders of the submodels.

IV. PWA IDENTIFICATION OF MIMO MODELS

Approach for the identification of a MIMO PWARX model [10] has been proposed in [9]. It is based on identifying m PWARX MISO models, one for each of the outputs which are finally merged in a single PWARX MIMO model as depicted in Fig. 1. While the discussed method is a natural way to extend the PWA MISO identification as described in [2] for the case of the PWA MIMO identification, it does not fully exploit the structure of the problem defined in Section II. The application of the respective approach for the identification of the wind turbine MIMO model is reported in [11].

In case of the MIMO PWA system identification, notice that the recognition criteria for each of the submodel regions is the parameter matrix θ_j which is concatenated of m parameter vectors, one for each of the outputs. It is reasonable to perform data classification in the extended space of feature vectors, containing parameter vectors for all of m outputs.

The benefits of the proposed approach are as follows. The approach enables user to select the number of submodels that form the identified PWA MIMO model. In the previously used approach [9] user would define the number of submodels for each of the PWA MISO models, whereas the resulting number of regions was ultimately determined by merging the identified MISO partitions into the single MIMO partition [9]. Moreover, the classification phase is expected to perform better when the considered number of attributes is larger. In the proposed approach the feature vector has $(s-1)(n+1)$ more attributes compared to the previously used approach. Furthermore, the proposed approach may alleviate the region over-generation effect that could be observed with the previously used approach.

Given that the polyhedral partition \mathcal{X} of a MIMO PWA model is formed by intersecting the independently identified

polyhedral partitions of MISO PWA models $\mathcal{X}_{(1)}, \dots, \mathcal{X}_{(m)}$ and some bordering constraints of the respective partitions coincide, the resulting partition \mathcal{X} is expected to contain the redundantly generated regions. It is a direct consequence of a non-ideal polyhedral partition estimation. If any of the common bordering constraints alters in the identification process, merging of the appurtenant partitions will result with region over-generation.

According to the discussion from above, we propose the clustering procedure that employs the feature vectors comprising the parameters of all submodels,

$$\theta_i = \left[\theta_{(1),i}^\top \quad \theta_{(2),i}^\top \quad \cdots \quad \theta_{(m),i}^\top \right]^\top. \quad (14)$$

Thus, a MIMO PWA map can be optimised in order to obtain the best approximation of nonlinearities for the specified number of regions s . Systematic framework for the identification of multiple-output linear systems is the multivariate linear regression [12], [13].

Likewise to the clustering-based identification scheme [2], authors in [3] perform the clustering in the space of regressors and multi-dimensional outputs. The advantage of their identification method, besides the ability to estimate MIMO PWARX models, is that the number of submodels can be estimated based on the information criteria such as CAIC and MDL. The approach that we propose has all stages of the clustering-based procedure [2], whereas some of them are reformulated to allow the identification of MIMO PWA systems.

A. Clustering-based identification procedure for MIMO models

To account for the PWA identification that is suitable for the identification of MIMO systems, the clustering-based procedure which is described in III is altered in the following.

Local regression. LDs are formed to contain input-output pairs $(x(k), y(k))$ likewise the discussion in Section III. For each LD \mathcal{C}_i all submodel parameters are identified employing the least squares method,

$$\begin{aligned} \theta_{(1),i}^{LS} &= (\Phi_i^\top \Phi_i)^{-1} \Phi_i^\top y_{(1),\mathcal{C}_i}, \\ &\vdots \\ \theta_{(m),i}^{LS} &= (\Phi_i^\top \Phi_i)^{-1} \Phi_i^\top y_{(m),\mathcal{C}_i}, \end{aligned} \quad (15)$$

where $y_{(h),\mathcal{C}_i}$ is a vector containing all h -th outputs of the local dataset \mathcal{C}_i and Φ_i is defined as in (5).

Feature vectors construction. Feature vectors are constructed of all submodel parameters and of the regressor vector mean m_i ,

$$\xi_i = \left[(\theta_{(1),i}^{LS})^\top \quad (\theta_{(2),i}^{LS})^\top \quad \cdots \quad (\theta_{(m),i}^{LS})^\top \quad m_i^\top \right]^\top. \quad (16)$$

Extended confidence matrix of the FV ξ_i is in the multivariate case given with

$$R_i = \begin{bmatrix} (\Phi_i^\top \Phi_i)^{-1} \otimes \Sigma & \mathbf{0} \\ \mathbf{0} & Q_i \end{bmatrix} \quad (17)$$

where the operation \otimes is the Kronecker product [12] and the estimate of Σ is given with

$$\hat{\Sigma} = \frac{y_{\mathcal{C}_i}^\top (I - \Phi_i (\Phi_i^\top \Phi_i)^{-1} \Phi_i^\top) y_{\mathcal{C}_i}}{c - (n + 1)}, \quad (18)$$

where $y_{\mathcal{C}_i} \in \mathbb{R}^{c \times m}$ contains the outputs of the model belonging to the local dataset \mathcal{C}_i .

Scalar measure of the feature vector ξ_i confidence level is determined with

$$w_k = \frac{1}{\sqrt{(2\pi)^{(m+1)n+m} \det(R_i)}}, \quad (19)$$

where w_k is the peak of the multivariate Gaussian distribution with covariance R_i .

Clustering of the feature vectors and estimation of the submodel parameters is conducted likewise the discussion in Section III.

V. EXAMPLE: IDENTIFICATION OF A PWARX MIMO MODEL

In order to validate the proposed extension of a clustering-based algorithm [2] suitable for the MIMO PWARX model identification, the test example of a coupled PWA system is constructed as depicted in Fig. 2. $G_{1,\lambda}$ and $G_{2,\lambda}$ are first order discrete transfer functions, $K_{12,\lambda}$ and $K_{21,\lambda}$ are gains, all of them varying with the decision parameter of mode switching λ ,

$$\lambda = \begin{cases} 1, & \text{if } y_{(1),k-1} > 0.5, y_{(2),k-1} > 0.5, \\ 2, & \text{if } y_{(1),k-1} \leq 0.5, y_{(2),k-1} > 0.5, \\ 3, & \text{if } y_{(1),k-1} \leq 0.5, y_{(2),k-1} \leq 0.5, \\ 4, & \text{if } y_{(1),k-1} > 0.5, y_{(2),k-1} \leq 0.5. \end{cases} \quad (20)$$

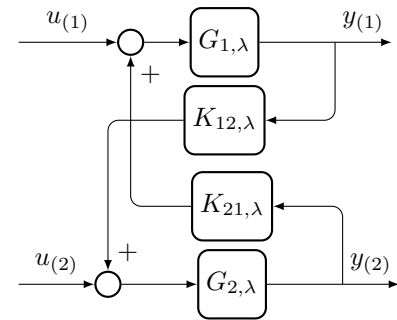


Fig. 2. Block diagram of the test example.

Equivalent MIMO PWARX model of the system described with the block diagram in Fig. 2 is defined with the following equations

$$y_k = \begin{cases} \begin{bmatrix} 0.9 & 0 & 0.1 & 0 \\ 0 & 0.6 & 0 & 0.4 \end{bmatrix} x_k, & \text{if } \lambda = 1, \\ \begin{bmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.15 & 0.6 & 0 & 0.4 \end{bmatrix} x_k, & \text{if } \lambda = 2, \\ \begin{bmatrix} 0.7 & -0.2 & 0.3 & 0 \\ 0.15 & 0.8 & 0 & 0.2 \end{bmatrix} x_k, & \text{if } \lambda = 3, \\ \begin{bmatrix} 0.9 & -0.2 & 0.1 & 0 \\ 0 & 0.8 & 0 & 0.2 \end{bmatrix} x_k, & \text{if } \lambda = 4, \end{cases} \quad (21)$$

where the regressor vector x_k is defined with

$$x_k = [y_{k-1}^\top, u_{k-1}^\top]^\top. \quad (22)$$

Regressor vectors are randomly distributed within the regressor set domain \mathcal{X} . \mathcal{X} has a shape of a hyperbox with each variable contained within the interval $(0, 1)$. By exploiting the *a priori* knowledge that the model switching is induced solely due to the output variables, means m_k are transformed prior the clustering procedure by employing the transformation matrix:

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (23)$$

The total number of regressor vectors that are used for the identification is $N = 8000$. In order to check the efficiency of the proposed algorithm for the identification of MIMO PWA systems, we compared its performance with the performance obtained by employing the algorithm in [9]. Mean square prediction errors of the identified PWA models for different values of identification parameters and signal-to-noise ratio (SNR) equal to 30 dB are summarised in Table I and Table II. SNR is determined by computing the ratio of the output measurement sum of squares to that of noise in decibels.

Table I shows results obtained by employing the algorithm published in [9], whereas the results for the herein proposed algorithm are shown in Table II. Since the algorithm in [9] does not allow specifying the number of regions for the PWA MIMO model, the results were obtained by considering three different scenarios for the number of regions comprised by each of the identified MISO PWA model partitions, denoted with a , b and c in the superscript of the number of submodels in Table II. In the scenario a we used 2 submodels to identify each of the MISO PWA model partitions. In the scenarios b and c we used 4 and 6 submodels for the identification of each MISO PWA partition, respectively. Final number of submodels is obtained by merging the identified MISO PWA partitions, as discussed in Section IV. According to the obtained results, the proposed algorithm achieves performance comparable to the previously used algorithm [9] by employing a significantly lower number of submodels.

TABLE I
RESULTS OF MIMO PWARX IDENTIFICATION [9] WITH RESPECT TO THE SIZE OF LDS c AND NUMBER OF SUBMODELS s , SNR= 30 dB

		c			
		12	14	16	18
s	4^a	1.09e-3	8.15e-4	7.99e-4	8.1e-4
	12^b	-	4.08e-4	-	-
	14^b	3.52e-4	-	-	-
	15^b	-	-	4.35e-4	-
	16^b	-	-	-	4.48e-4
	19^c	-	2.45e-4	-	-
	20^c	2.83e-4	-	-	-
	21^c	-	-	2.52e-4	-
	23^c	-	-	-	2.59e-4

Generally, the validation error of the model with respect to the value of the parameter c is closely related to the number of mixed LDS that enter the classification process. Given that

TABLE II
RESULTS OF THE PROPOSED MIMO PWARX IDENTIFICATION WITH RESPECT TO THE SIZE OF LDS c AND NUMBER OF SUBMODELS s , SNR= 30 dB

		c			
		12	14	16	18
s	2	1.43e-3	1.09e-3	0.99e-3	1.11e-3
	4	3.39e-4	3.71e-4	3.73e-4	4.04e-4
	6	3.94e-4	3.89e-4	4.15e-4	4.92e-4

the parameter c is fixed, the number of mixed LDS, hence the validation error, should be the smallest when the number of submodels used in the PWA model identification matches the real number of system modes. Regarding the results in Table II, in the scenario with the number of submodels set to $s = 2$ is the clustering algorithm unable to classify the data according to given submodels irrespective of the used parameter c . The smallest approximation error is achieved with $c = 12$ and $s = 4$. Respective parameters result with the smallest number of mixed LDS in the clustering-based procedure.

VI. CONCLUSION

Systematic approach to the MIMO PWA model identification is of crucial importance in a model-based controller synthesis for a large number of processes. The approach proposed here for the identification of PWA MIMO models exploits all benefits of the clustering-based identification [2]. Besides, K-means++ algorithm enhanced with the confidence level of the feature vectors is incorporated for the purpose of the clustering centers initialisation, resulting with the improved classification of the identification data. Moreover, the feature vector transformation is introduced in order to control the partitioning with respect to the arbitrary dimensions of the system, which is of particular importance when mode switching does not occur with respect to certain variables. The proposed extension is validated on an illustrative example. The results show the effectiveness of the approach.

ACKNOWLEDGMENT

This work has been supported by the Croatian Science Foundation under the project No. 6731 Control-based Hierarchical Consolidation of Large Consumers for Integration in Smart Grids (3CON, <http://www.fer.unizg.hr/3con>).

REFERENCES

- [1] A. L. Juloski, W. Heemels, G. Ferrari-Trecate, R. Vidal, S. Paoletti, and J. Niessen, "Comparison of four procedures for the identification of hybrid systems," in *International Workshop on Hybrid Systems: Computation and Control*. Springer, 2005, pp. 354–369.
- [2] 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.
- [3] H. Nakada, K. Takaba, and T. Katayama, "Identification of piecewise affine systems based on statistical clustering technique," *Automatica*, vol. 41, no. 5, pp. 905 – 913, 2005.
- [4] A. Garulli, S. Paoletti, and A. Vicino, "A survey on switched and piecewise affine system identification," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 344 – 355, 2012, 16th IFAC Symposium on System Identification.

- [5] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '07. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.
- [6] G. Ferrari-Trecate and M. Muselli, "Single-linkage clustering for optimal classification in piecewise affine regression," *IFAC Conference on the Analysis and Design of Hybrid Systems (ADHS 03)*, pp. 33–38, 2003.
- [7] L. Ljung, "System identification-theory for the user 2nd ed," *PTR Prentice Hall, Upper Saddle River, NJ*, 1999.
- [8] Z. Lassoued and K. Abderrahim, "New results on pwarx model identification based on clustering approach," *International Journal of Automation and Computing*, vol. 11, no. 2, pp. 180–188, 2014.
- [9] M. Vašak, D. Klanjčić, and N. Perić, "Piecewise affine identification of MIMO processes," in *Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, 2006 IEEE*. IEEE, 2006, pp. 1493–1498.
- [10] S. Paoletti, A. L. Juloski, G. Ferrari-Trecate, and R. Vidal, "Identification of hybrid systems a tutorial," *European Journal of Control*, vol. 13, no. 2, pp. 242 – 260, 2007.
- [11] H. Zhao, Q. Wu, Q. Guo, H. Sun, and Y. Xue, "Distributed model predictive control of a wind farm for optimal active power control. Part II: Implementation with clustering-based piece-wise affine wind turbine model," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 840–849, 2015.
- [12] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis, 3rd Edition*. Wiley, 2003.
- [13] S. Sawyer, "Multivariate linear models," University lecture, 2010.