A Bounded Switching Approach for Identification of Switched MIMO Systems

Mohammad Gorji Sefidmazgi
School of Information
The University of Arizona
Tucson, Arizona 85721

Mina Moradi Kordmahalleh,
Abdollah Homaifar, Ali Karimoddini
Autonomous Control and Information Technology Institute
Department of Electrical and Computer Engineering,
North Carolina A&T State University
Greensboro, North Carolina 27411

Edward Tunstel
Applied Physics Laboratory
Johns Hopkins University
Laurel, Maryland 20723

***Abstract*—This study considers offline identification of switched linear MIMO systems using measurements from their inputs and outputs. This is a class of non-convex optimization and ill-posed problems. To convert this optimization into a binary integer programming problem, the proposed approach assumes that the number of switches among the subsystems is upper-bounded. The state-space realization of each sub-system is found by the subspace identification. The proposed approach does not need the tuning of the moving window size or any penalization factor. The algorithm efficiency is evaluated through numerical simulations. The results indicate that the error of identification is small and the eigenvalues of sub-systems are estimated successfully.**

I. Introduction

Switched linear systems are one of the types of hybrid systems. They consist of several sub-systems which are activated by a switching mechanism [1]. The aim of switched system identification is to find the mathematical models of sub-systems and times of switchings from the measured noisy inputs/outputs. This is a challenging problem, since there is a strong coupling between estimation of the sub-system parameters and assigning data into the sub-systems. Another motivation for studying switched MIMO systems is identification and control of biological and power systems consisting multiple input/outputs [2], [3].

The model of each sub-system might be in autoregressive or state-space form. The switched system with autoregressive model could directly explain the relationship between inputs and outputs. A review of the methods for identification based on autoregressive models can be found in [4], [5] which are mostly for SISO systems. However, it is difficult to use the input-output models for identifying in switched MIMO systems. The dimension of corresponding regressor vector is high and the matrix in least square estimation is not full rank. As a result, proper parameter estimation is not possible [6], [7]. Modeling of MIMO systems in the form of state-space based on *subspace identification* is a common approach in control literature. There exists a well-matured theory for process control, observer design, system realization and stability analysis that relies on state-space models [8]. The identification of switched MIMO systems can therefore be made more effective when the sub-systems are described by minimal state-space models. However, finding the parameters of switched state-space models is difficult, since the states are generally unknown and thus the regression vector for least square optimization is not available.

Corresponding Author: A. Homaifar, Telephone: (336) 285–3271

The identification of state-space models for switched MIMO systems can be solved with the assumptions that the switching times are known [9], or the value of states are measurable [6]. However in many applications, the states are not available and these approaches are not suitable. In [10], an identification approach is proposed that can work online, but some datasets might be lost during the switching transition. In [11], the input-output data are projected into a higher dimensional space and then are clustered. In [12], a relaxation is added to convert the identification problem into a convex optimization with a need to tune a regularization parameter in noisy conditions. Several other studies have proposed methods based on a sliding window, e.g. machine learning techniques [13], chi-square test [14], projected subspace [15] and least square [16], where the data under the window are assigned to one of the sub-systems. However, choosing the length of the moving windows highly affects the accuracy of clustering.

This paper proposes an identification technique for switched MIMO systems. The proposed approach does not need tuning the length of sliding window. Moreover, it is not necessary to tune any real-valued hyper-parameters (such as the penalization factor). Instead, a little subset of integer-valued hyper-parameters from a finite set of possible structures are tuned. Thus, it does not require time-consuming model selection approaches such as *exhaustive grid search*. As a result, the proposed method is very suitable when minimal prior knowledge about the number of sub-systems and switches of the switched system is available. In the proposed method, we assume that the total number of switchings in the system is upper bounded. This assumption converts the identification problem into a binary integer programming problem. The proposed identification employs the *Bounded-Switching* (BS) clustering algorithm [17]. In our approach we analyze the system data in batch, and identifies the switching times and

978-1-5090-1897-0/16/$31.00 ©2016 IEEE

the sub-systems at the same time.

Therefore, the contribution of this paper is a new method proposed to identify the switched MIMO systems utilizing the BS-clustering and subspace identification, in order to find the parameters of sub-systems and also the switch times. The algorithm is tested through the numerical simulations. This paper is organized as follows. In the section II, we mathematically formulated the identification problem of switched linear MIMO systems. In Section III, the subspace identification method is explained which is used to find the sub-model parameters. Section IV shows the Bounded-Switching clustering that is applied to find the switch times. The algorithm performance is shown through MATLAB simulations in section V. Section VI provides the conclusions and future works.

II. Problem definition

A switched linear system could be represented as S MIMO sub-systems, and each sub-system is modeled as a state-space representation, which together generate a switched MIMO system. Each *sub-model* could be represented in discrete-time linear state-space time as below [16], [14]:

$$\begin{aligned} x(t+1) &= A_{s(t)}x_{s(t)}(t) + B_{s(t)}u(t) + w(t) \\ y(t) &= C_{s(t)}x_{s(t)}(t) + D_{s(t)}u(t) + v(t) \\ x(0) &= x_0; t = \{0,\dots,T-1\}; s(t) = \{1,\dots,S\} \end{aligned} \quad (1)$$

Here t is the time index, and T is the total length of the dataset. The $y \in \mathbb{R}^l$ and $u \in \mathbb{R}^m$ are the outputs and inputs of the system respectively. The $s(t) \in \{1,2,\dots,S\}$ and $x_{s(t)} \in \mathbb{R}^n$ are the indices and the state of active sub-system at time t. Furthermore, $w(t)$ and $v(t)$ are the additive process and measurement noises. Here, $\Psi_s = \{A_s, B_s, C_s, D_s\}$ are the parameters of the s-th sub-model. The activation path of sub-models and the matrices of parameters Ψ_s are unknown. Similar to [10], [16], [14], [13], we have assumed that sub-systems of the switched system may have different orders. In this case, the states of the switched system are not necessarily continuous at switch times.

Given the discrete pairs of generated data $\{u_t, y_t\}_{t=0}^{T-1}$ from the sub-systems of the switched system (1), our aim is to find the parameters Ψ_s of the switched system, the time instants at which switching between S sub-models occurs, and the activated sub-model at any given time t. In this problem, a clustering approach should be applied to the data $\{u_t, y_t\}_{t=0}^{T-1}$ to detect the sub-system that generated this output. In this paper, we assume the order of sub-models, n, is known and finite.

Assume that all of the output data, $y(t)$, $t = \{0,\dots,T-1\}$, is generated by only the s-th sub-model with parameters in the form of $\Psi_s = \{A_s, B_s, C_s, D_s\}$. In this case, the true states $\tilde{x}_s(t)$ and true outputs $\tilde{y}_s(t)$ (without any noise) of this hypothetical system are represented by:

$$\begin{aligned} \tilde{x}_s(t+1) &= A_s\tilde{x}_s(t) + B_s u(t) \\ \tilde{y}_s(t) &= C_s\tilde{x}_s(t) + D_s u(t) \\ \tilde{x}_s(0) &= \tilde{x}_{s0}; t = \{0,\dots,T-1\}; s = \{1,\dots,S\} \end{aligned} \quad (2)$$

However, only one of the sub-models from the set of $s = \{1,\dots,S\}$ is active at each time step t. Thus the problem of switched system identification can be defined as minimizing the distance between the measured output of the switched system $y(t)$ and outputs of S hypothetical systems in (2). This minimum of identification error can be achieved by finding the parameters of the switched system and also the switching times as formulated in (3) and (4):

$$\min_{\mu_s(t),\Psi_s} J(\mu_s(t), \Psi_s) \quad (3)$$

$$J(\mu_s(t), \Psi_s) = \sum_{t=0}^{T-1}\sum_{s=1}^{S} \mu_s(t) \left\| y(t) - \tilde{y}_s(t) \right\|^2 \quad (4)$$

In (4), $\mu_s(t) \in \{0,1\}$ is called as the sub-model activation function. The $\mu_s(t)$ shows that which one of the sub-models is active at time t, and Thus:

$$\sum_{s=1}^{S} \mu_s(t) = 1 \qquad t = \{0,\dots,T-1\} \quad (5)$$

Next we define the the *Distance Function* $h_s(t)$ as the Euclidean distance between the outputs of the switched system $y(t)$ and output of the s-th sub-system. The minimization problem of (4) is rewritten more compactly:

$$\begin{aligned} h_s(t) &= \left\| y(t) - \tilde{y}_s(t) \right\|^2 \\ & t = \{0,\dots,T-1\}; s = \{1,\dots,S\} \end{aligned} \quad (6)$$

Furthermore, we augment the values of sub-model activation functions and distance functions in two vectors M_s and H_s for $s = \{1,\dots,S\}$

$$\begin{aligned} H_s &= \begin{bmatrix} h_s(0) & h_s(1) & \dots & h_s(T-1) \end{bmatrix} \\ M_s &= \begin{bmatrix} \mu_s(0) & \mu_s(1) & \dots & \mu_s(T-1) \end{bmatrix} \end{aligned} \quad (7)$$

The optimization problem of (4) is rewritten as [17]:

$$\min_{H_s,M_s} \sum_{s=1}^{S} H_s.M_s^T \quad (8)$$

In order to find the optimal solution for this problem, it is necessary to search among all possible paths of $s(t)$, which is not a computationally efficient approach. In general, the cost function in (8) is a non-convex and ill-posed optimization problem. To solve this optimization problem, we use the *coordinate descent* algorithm [18] that iteratively solves this problem in two steps with respect to sub-model activation functions and parameters. The sub-model parameters Ψ_s are found by the subspace identification method, and the sub-model activation function $\mu_s(t)$ are found using the Bounded-Switching clustering.

III. Finding sub-models parameters

Suppose that $\mu_s(t)$ is known, thus the switching times among the S sub-systems can be determined. When the switches are separated by a certain minimum time, subspace identification coupled with a clustering technique can be used for identification [10]. Thus in order to find sub-model

parameters, similar to [9], [13], [16], we make the assumption that whenever the system visits a discrete state $s(t)$, it stays in it during a certain minimum time called as *dwell time*. If the sequence of discrete mode and switching times are observed and switching signal has suitable large dwell time, then the linear switched systems can be identified by identifying each linear subsystem separately [19].

Knowing the switch times, the time steps $t = \{0, \dots, T-1\}$ are partitioned into S segments, where the length of s-th segment is T_s. Each sub-model might be returning, and become activated several times in the sequence of the sub-models. As an example, a sub-system may be activated twice in $t = \{0, \dots, 100\}$ and $t = \{200, \dots, 300\}$. Thus these S time partitions might be non-continuous. Then, we separate all the pairs of $\{u(t), y(t)\}$ that belongs to the s-th time partitions, where $s \in \{1, \dots, S\}$. Assume that data for the s-th sub-model is $\{u_s(\tau), y_s(\tau)\}$ and $\tau = \{0, \dots, T_s - 1\}$. The goal is to find a state-space realization for this dataset using subspace identification technique. In this work, we use the *Past Outputs Multivariable Output-Error State-space* (PO-MOESP) method [20] to find the matrices A_s, B_s, C_s and D_s for the data assigned to the s-th sub-model. The input/output relation of the s-th sub-model is given by the following equations:

$$\begin{aligned} x_s(\tau+1) &= A_s x_s(\tau) + B_s u(\tau) + w(\tau) \\ y_s(\tau) &= C_s x_s(\tau) + D_s u(\tau) + v(\tau) \\ x_s(0) &= x_{s0}; \tau = \{0, \dots, T_s - 1\} \end{aligned} \quad (9)$$

For simplicity of notations, we removed the subscript s from these four matrices and $x_s(0)$. The input/output relation for the sub-model can be written for $\tau = \{0, \dots, T_s - 1\}$ [21]:

$$\begin{aligned} y(\tau) = CA^{\tau}x(0) &+ C\sum_{i=0}^{\tau-1} A^{\tau-i-1}Bu(i) \\ &+ C\sum_{i=0}^{\tau-1} A^{\tau-i-1}w(i) + Du(\tau) + v(\tau) \end{aligned} \quad (10)$$

The (10) can be rewritten in a matrix form as:

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ y(\rho-1) \end{bmatrix} = \underbrace{\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{\rho-1} \end{bmatrix}}_{\mathcal{O}_\rho} x(0)$$

$$+ \underbrace{\begin{bmatrix} D & 0 & 0 & \dots & 0 \\ CB & D & 0 & \dots & 0 \\ CAB & CB & D & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ CA^{\rho-2}B & CA^{\rho-3}B & \dots & CB & D \end{bmatrix}}_{\mathcal{T}_\rho} \begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ \vdots \\ u(\rho-1) \end{bmatrix}$$

$$+ \underbrace{\begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ C & 0 & 0 & \dots & 0 \\ CA & C & D & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ CA^{\rho-2}B & CA^{\rho-3}B & \dots & CB & D \end{bmatrix}}_{\mathcal{G}_\rho} \begin{bmatrix} w(0) \\ w(1) \\ w(2) \\ \vdots \\ w(\rho-1) \end{bmatrix}$$

$$+ \begin{bmatrix} v(0) \\ v(1) \\ v(2) \\ \vdots \\ v(\rho-1) \end{bmatrix} \quad (11)$$

where ρ is the integer *block size* such that $n < \rho \leq T_s$. The O_ρ is the extended observability matrix for the s-th sub-model and its size is $(l.\rho) \times n$. The T_ρ is the *block-Toeplitz* matrix with size $(l.\rho) \times (m.\rho)$.

Definition 1: The Hankel matrix of block size ρ constructed for a vector $f(t)$ is defined as:

$$F_{t,\rho,N} = \begin{bmatrix} f(t) & f(t+1) & \dots & f(t+N-1) \\ f(t+1) & f(t+2) & \dots & f(t+N) \\ \vdots & \vdots & \ddots & \vdots \\ f(t+\rho-1) & f(t+\rho) & \dots & f(t+N-\rho-2) \end{bmatrix} \quad (12)$$

Definition 2: The row vector $X_{t,N}$ is constructed from $x(t)$:

$$X_{t,N} = \begin{bmatrix} x(t) & x(t+1) & \dots & x(t+N-1) \end{bmatrix} \quad (13)$$

The first subscript of $X_{t,N}$ denotes the starting index and the second subscript is the length of the vector.

Since the underlying sub-model is time-invariant, the (11) can be rewritten for time-shifted inputs/outputs and states using Hankel matrices in $\tau = 0, \dots, T_s - 1$.

$$Y_{t,\rho,N} = \mathcal{O}_\rho . X_{t,N} + \mathcal{T}_\rho . U_{t,\rho,N} + \mathcal{G}_\rho . W_{t,\rho,N} + V_{t,\rho,N} \quad (14)$$

where $Y_{t,\rho,N}$, $U_{t,\rho,N}$, $W_{t,\rho,N}$ and $V_{t,\rho,N}$ are the Hankel matrices for vector of outputs and inputs, process noises and measurement noises. In the PO-MOESP method, proper projections are applied to the (14) such that all of the terms in the right hand side are removed except the term with the extended observability matrix $\mathcal{O}_\rho . X_{t,N}$. Deriving such projections are discussed in detail in [22]. After applying these projections, the extended observability matrix $\mathcal{O}_\rho$ can be found by the following procedure:

We define the past Hankel matrices for the inputs and output as $U_{0,\rho,N}$ and $Y_{0,\rho,N}$, and also the future input and output Hankel matrices $U_{\rho,\rho,N}$ and $Y_{\rho,\rho,N}$. Now define the instrumental variable $\mathcal{Z}$ as below:

$$\mathcal{Z} = \begin{bmatrix} U_{0,\rho,N} \\ Y_{0,\rho,N} \end{bmatrix} \quad (15)$$

The LQ factorization is applied as below:

$$\begin{bmatrix} U_{\rho,\rho,N} \\ \mathcal{Z} \\ Y_{\rho,\rho,N} \end{bmatrix} = \begin{bmatrix} U_{\rho,\rho,N} \\ U_{0,\rho,N} \\ Y_{0,\rho,N} \\ Y_{\rho,\rho,N} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 \\ L_{31} & L_{32} & L_{33} & 0 \\ L_{41} & L_{42} & L_{43} & L_{44} \end{bmatrix} \cdot \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \end{bmatrix} \quad (16)$$

The *singular value decomposition* (SVD) operator is applied to the vector $\begin{bmatrix} L_{42} & L_{43} \end{bmatrix}$:

$$\begin{bmatrix} L_{42} & L_{43} \end{bmatrix} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{17}$$

We take the first n columns of the $\mathbf{U}$ which is denoted by U_n where n is the order of the sub-system. It is proved that U_n is an estimate of $\mathcal{O}_\rho$. Thus, the estimation of matrices A and C can be found:

$$\hat{C} = \text{the first } l \text{ rows of } U_n \tag{18}$$

Take U_1 as the upper $(\rho-1)l$ rows of U_n and U_2 as the lower $(\rho-1)l$ rows of U_n. Then using pseudo-inverse †[22]:

$$\hat{A} = U_1^{\dagger}.U_2 \tag{19}$$

The matrices $\hat{B}$ and $\hat{D}$, together with the initial state $\hat{x}(0)$ can be computed by solving a least-squares problem. Let's define θ, the vector of unknown parameters:

$$\theta = \begin{bmatrix} x(0) & vec(B) & vec(D) \end{bmatrix}^T \tag{20}$$

Where $vec(.)$ is an operator that stacks the columns of a matrix on top of each other. The regressor vector is defined:

$$\phi(\tau)^T = \begin{bmatrix} \hat{C}\hat{A}^\tau & \sum_{i=0}^{\tau-1} u(i)^T \otimes \hat{C}\hat{A}^{\tau-i-1}) & u(\tau)^T \otimes I_l \end{bmatrix} \tag{21}$$

where $\otimes$ is the *Kronecker product* of matrices. Now $\hat{\theta}$ is found using least-square [22]:

$$\hat{\theta} = \arg\min_{\theta} \sum_{\tau=0}^{T_s-1} \left\| y(\tau) - \phi(\tau)^T.\theta \right\|^2 \tag{22}$$

In summary, the above procedure of subspace identification is applied to the data of S sub-models in order to find sub-model parameters $\hat{\Psi}_s = \{\hat{A}_s, \hat{B}_s, \hat{C}_s, \hat{D}_s\}$ and initial states $\hat{x}_s(0)$. After finding $\hat{\Psi}_s$, the error $h_s(t)$ is calculated using (2) and (6) which shows the distance between $y(t)$ to the models of s-th submodel. This distance will be used in the next section to assign the dataset to each of the S sub-models.

IV. Assigning Data to the Sub-Models

The time series analysis method with jumped models using the bounded-switching (BS) clustering was introduced in [17]. The basic idea of BS-clustering is to bound the total number of switchings in the time series and estimate the switch times. A similar idea was used before in the identification of switched SISO systems [23], [24] by penalizing the total variations of the switched system parameters over time. However in these methods, it is necessary to tune a real-valued penalization factor which is a time-consuming procedure. Instead in BS-clustering, it is assumed that the total number of switches is upper bound. This assumption make it possible to convert the ill-posed identification problem into a solvable optimization without need of a penalization factor.

Suppose that $\hat{\Psi}_s$ is available for all of the S sub-systems. The goal is to find $\mu_s(t)$ using a clustering algorithm. A function $q_s(t) \in \{0, 1\}$ is defined to count the total number of switches to/from the s-th sub-model

$$|\mu_s(t+1) - \mu_s(t)| \leq q_s(t) \qquad t = \{0, \ldots, T-2\}; s = \{1, \ldots, S\} \tag{23}$$

This equation can be extended as:

$$\begin{cases} \mu_s(t+1) - \mu_s(t) - q_s(t) \leq 0 \\ -\mu_s(t+1) + \mu_s(t) - q_s(t) \leq 0 \end{cases} \qquad t = \{0, \ldots, T-2\}; s = \{1, \ldots, S\} \tag{24}$$

During a switch times between two sub-systems e.g. s_2 and s_3, the value of $\mu_{s_2}(t)$ jumps from 1 to 0, while the value of $\mu_{s_3}(t)$ jumps from 0 to 1. To limit the total number of switching between the sub-models an integer value K, a constraint is added on the $q_s(t)$:

$$\sum_{s=1}^{S} \sum_{t=0}^{T-2} q_s(t) = 2K \tag{25}$$

where K is a known integer number. Next, we add the $q_s(t)$ to the set of unknown parameters, which should be found in the optimization problem. A vector Q_s is defined that contains all the values of $q_s(t)$ over time:

$$Q_s = \begin{bmatrix} q_s(0) & q_s(1) & \ldots & q_s(T-2) \end{bmatrix} \qquad s = \{1, \ldots, S\} \tag{26}$$

We define vector $\bar{M}$ containing all of the $\mu_s(t)$ and $q_s(t)$ of the S clusters (the M_s was defined in (7))

$$\bar{M} = \Big[\underbrace{M_1 \quad M_2 \quad \ldots \quad M_S}_{T \times S} \underbrace{Q_1 \quad Q_1 \quad \ldots \quad Q_S}_{(T-1) \times S} \Big] \tag{27}$$

and the vector $\bar{H}$ containing all the values of $h_s(t)$

$$\bar{H} = \Big[\underbrace{H_1 \quad H_2 \quad \ldots \quad H_S}_{T \times S} \underbrace{0 \quad 0 \quad \ldots \quad 0}_{(T-1) \times S} \Big] \tag{28}$$

Now, the optimization in (8) can be rewritten by $\bar{M}$ and $\bar{H}$ as [17]:

$$\min_{\bar{M}} \bar{H}.\bar{M}^T \tag{29}$$

The vector $\bar{M}$ has all of $\mu_s(t)$ and $q_s(t)$, could be determined by solving this optimization problem. All of the values of $\mu_s(t)$ and also $q_s(t)$ are binary, and thus the vector $\bar{M}$ has only binary values. As a result, the (29) is a binary integer programming problem, which is solvable by methods of convex optimization. Here, The optimization order is $(2T-1)S$. The (5), (24) and (25) form the constraints of the optimization problem, including $2S(T-1)$ inequality and $T+1$ equality constraints. In this work, we have used the toolbox of [25] that uses the branch-and-bound algorithm for solving the binary integer programming with computational complexity of $O(2^n)$, where n is the order of binary integer programming.

The procedure of proposed algorithm is as follows. In the first step, initial values are chosen for $\mu_s(t)$ and then the

iterative steps of sections III and IV should be followed to find Ψ_s and $\mu_s(t)$. The convergence of algorithm can be tested using the clustering error. The optimizations with respect to Ψ_s using subspace identification always has a solution if the minimum dwell time assumption is satisfied. Due to (5), the feasible region of the binary integer programming is non-empty. Thus, the optimization to find $\mu_s(t)$ always has a solution. In the result, the proposed identification method based on BS-clustering is monotone (with non-increasing error) and converges to at least a local solution. For obtaining the global solution of the originally non-convex optimization problem of (4), the procedure must be restarted using random initializations of $\mu_s(t)$.

The proposed algorithm needs tuning of the number of sub-models S, number of switches K and the block size ρ. The number of sub-models and switches can be selected from a finite set of possible structures based on prior knowledge about the the switched system. The block size is a number smaller than the dwell time [16].

V. Simulations

In this section, the proposed identification algorithm is tested using a simulation example borrowed from [16], where the switch time is a returning sequence for all the three sub-systems and the order of the sub-systems are different and thus it makes the estimation more difficult. The data with length of 900 is generated by three sub-models as below, and $D_s = 0$:

$$A_1 = \begin{bmatrix} 0.4 & 0.1 & 0 \\ 0.8 & 0.4 & 0 \\ 0 & 0 & 0.8 \end{bmatrix} \quad B_1 = \begin{bmatrix} 1.5 & 0.9 \\ 1 & -1 \\ -1.5 & 2.3 \end{bmatrix}$$

$$C_1 = \begin{bmatrix} 0.8 & 1.1 & 2 \\ -1.3 & 0.7 & 1.7 \\ 1.5 & 0.7 & -0.9 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0.4 & 0.6 \\ 0.5 & 0.1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 1.5 & 0.9 \\ 1 & -1 \end{bmatrix} \quad C_2 = \begin{bmatrix} 0.8 & 1.1 \\ -1.3 & 0.7 \\ 1.5 & 0.7 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0.3 & 0.2 & 0 \\ 0.8 & 0.2 & 0 \\ 0 & 0 & -0.75 \end{bmatrix} \qquad B_3 = B_1 \qquad C_3 = C_1 \tag{30}$$

Each sub-model is active in some periods of time:

$$s(t) = \begin{cases} 1 & \text{if } t \in [1, 199] \cup [400, 499] \cup [700, 799] \\ 2 & \text{if } t \in [200, 299] \cup [500, 599] \cup [800, 900] \\ 3 & \text{if } t \in [300, 399] \cup [600, 699] \end{cases} \tag{31}$$

Here, the sub-systems are returning, and they are activated more than once in the sequence. The inputs $u(t)$ are Pseudo-Random Binary Sequences (PRBS). The signal-to noise ratio (SNR) of the outputs with respect to the measurement noises is 20 db. The process noise $w(t)$ is a Gaussian white noise and has a covariance as $var(w(t)) = 7 \times 10^{-4} I_3$. We assign the block size as $\rho = 10$. In this example, the orders of the sub-models are not the same. However, for simplicity

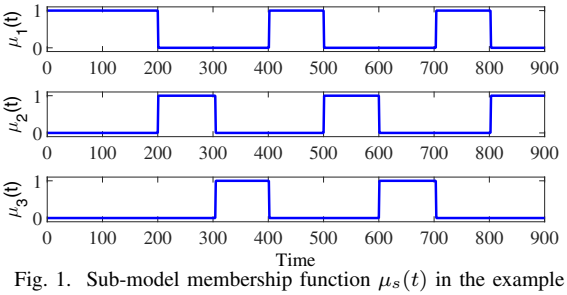

Fig. 1. Sub-model membership function $\mu_s(t)$ in the example

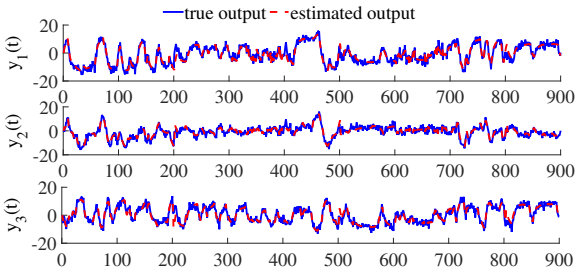

Fig. 2. Real and estimated outputs of the system in the example

of calculation, we estimate three sub-model parameters with the order $n = 3$. The proposed identification method is applied assuming $S = 3$ and $K = 7$ which coincides with the true configuration of the switched system. The iterative identification algorithm is converged after four iterations. The values of sub-model membership function $\mu_s(t)$ are shown in Figure 1. It can be seen that the switch times between three sub-models are found correctly. Figure 2 shows that the proposed approach has successfully approximated the real outputs of the switched system.

For testing the accuracy of detecting the true eigenvalues of three sub-systems, we run the algorithm 50 times, with different inputs and noises on the system. Figure 3 shows the estimated eigenvalues with blue dots and actual eigenvalues with red circles for three sub-models, where the eigenvalues are found with a good accuracy.

Next, we test the effect of measurement noise on the accuracy of identification by running the algorithm with different SNRs and calculating the the variance-accounted-for (VAF) values. Table 1 the identification approach is capable of capturing the dynamics of sub-models with different levels of noise.

$$VAF = max(1 - \frac{var(y(t) - \hat{y}(t))}{var(y(t))}, 0) \times 100 \tag{32}$$

In contrast to the solution of the same problem in [16], our algorithm does not need tuning of the sliding window length to reach an acceptable result. Instead, by knowing the number of sub-models and switches, the identification problem can be solved. If these hyper-parameters are unknown, only a finite set of models should be tested to reach the true structure of the switched system, without need of tuning any real-valued hyper-parameter. Furthermore, statistical tech-

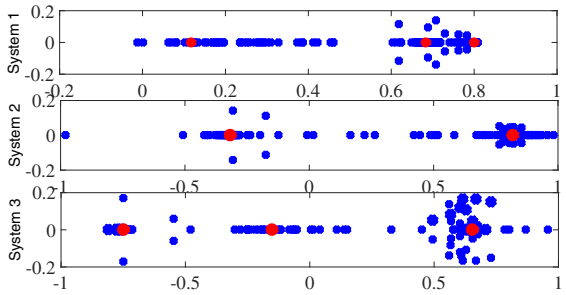

Fig. 3. Eigenvalues of true and estimated sub-models

TABLE I
EFFECT OF NOISE LEVEL ON ACCURACY OF IDENTIFICATION

SNR (dB)	15	20	25	30
VAF	92.5	95.8	96.2	98.9

niques such as *Akaike Information Criterion* can be used for selecting the model structure, which enables us to avoid overfitting/underfitting due to the incorrect settings of hyper-parameters.

VI. Conclusion

In this paper, we introduced an algorithm based on bounded-switching clustering and subspace identification technique for identification of switched linear MIMO systems. Compared to the conventional methods, our method does not need to tune hyper-parameters such as moving window length or penalization factor. The proposed identification method clusters the input/output pairs of the switched MIMO systems and utilizes the PO-MOESP method to find the state-space realization of each sub-system. The set of hyper-parameters which should be tuned are the number of sub-models and switches and also the block size. These parameters might be selected from a set of finite admissible structures for the switched system. Thus, the proposed method is very suitable when minimal *a priori* knowledge is available about the switched system. The identification method was tested through MATLAB simulations showing acceptable performance with small identification error. Future work will focus on the estimation of the order and number of the sub-models in switched MIMO systems.

Acknowledgment

This work is partially supported by the Expeditions in Computing by the National Science Foundation under Award Number CCF-1029731 and DOT under UTC-60040605 and Air Force Research Laboratory and OSD under agreement number FA8750-15-2-0116. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of NSF, DoT, Air Force Research Laboratory, OSD or the U.S. Government.

References

[1] D. Liberzon, *Switching in Systems and Control*. Birkhuser, 2003.
[2] J.-I. Imura, K. Kashima, M. Kusano, T. Ikeda, and T. Morohoshi, "Piecewise affine systems approach to control of biological networks." *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, vol. 368, no. 1930, pp. 4977–93, 2010.
[3] M. G. Kashani, S. Babaei, and S. Bhattacharya, "SVC and STATCOM application in Electric Arc Furnace efficiency improvement," in *4th IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, 2013, pp. 1–7.
[4] A. Garulli, S. Paoletti, and A. Vicino, "A survey on switched and piecewise affine system identification," in *16th IFAC Symposium on System Identification*, 2012, pp. 344 – 355.
[5] M. Gorji Sefidmazgi, M. Moradi Kordmahalleh, A. Homaifar, and A. Karimoddini, "Switched linear system identification based on bounded-switching clustering," Chicago, pp. 1806–1811, 2015.
[6] L. Bako, V. L. Le, F. Lauer, and G. Bloch, "Identification of MIMO switched state-space models," in *American Control Conference (ACC), 2013*, June 2013, pp. 71–76.
[7] L. Bako and R. Vidal, "Algebraic identification of MIMO SARX models," in *Proceedings of the 11th International Workshop on Hybrid Systems: Computation and Control*, ser. HSCC '08. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 43–57.
[8] W. S. Levine, Ed., *Control System Advanced Methods*. CRC Press, 2011.
[9] V. Verdult and M. Verhaegen, "Subspace identification of piecewise linear systems," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 4, Dec 2004, pp. 3838–3843 Vol.4.
[10] L. Bako, G. Mercre, and S. Lecoeuche, "On-line structured subspace identification with application to switched linear systems," *International Journal of Control*, vol. 82, no. 8, pp. 1496–1515, 2009.
[11] K. Huang, A. Wagner, and Y. Ma, "Identification of hybrid linear time-invariant systems via subspace embedding and segmentation (SES)," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 3, Dec 2004, pp. 3227–3234 Vol.3.
[12] L. Bako, G. Mercre, R. Vidal, and S. Lecoeuche, "Identification of switched linear state space models without dwell time," in *15th IFAC Symposium on System Identification*, 2009, pp. 569–574.
[13] R. Vilela Lopes, G. Araujo Borges, and J. Yoshiyuki Ishihara, "New algorithm for identification of discrete-time switched linear systems," in *American Control Conference (ACC), 2013*, June 2013, pp. 6219–6224.
[14] K. Pekpe, G. Mourot, and J. Ragot, "Identification of switching systems using change detection technique in the subspace framework," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 4, Dec 2004, pp. 3720–3725 Vol.4.
[15] J. Borges, V. Verdult, M. Verhaegen, and M. Botto, "A switching detection method based on projected subspace classification," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, Dec 2005, pp. 344–349.
[16] K. Pekpe and S. Lecuche, "Online clustering of switching models based on a subspace framework," *Nonlinear Analysis: Hybrid Systems*, vol. 2, no. 3, pp. 735 – 749, 2008.
[17] P. Metzner, L. Putzig, and I. Horenko, "Analysis of persistent non-stationary time series and applications," *Communications in Applied Mathematics and Computational Science*, vol. 7, no. 2, pp. 175–229, 2012.
[18] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
[19] L. H. Rajaoarisoa, J. F. Balmat, and N. K. M'Sirdi, "State estimation for MIMO hybrid dynamical model," in *IEEE Conference on Decision and Control and European Control Conference*. IEEE, Dec. 2011, pp. 5455–5460.
[20] M. Verhaegen, "Identification of the deterministic part of MIMO state space models given in innovations form from input-output data," *Automatica*, vol. 30, no. 1, pp. 61 – 74, 1994.
[21] B. Haverkamp, "State space identification: theory and practice," Ph.D. dissertation, Delft University of Technology, 2000.
[22] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach*. Cambridge University Press, 2012.
[23] M. Gorji Sefidmazgi, M. Moradi Kordmahalleh, A. Homaifar, and A. Karimoddini, "A finite element based method for identification of switched linear systems," in *American Control Conference (ACC)*, 2014, pp. 2644–2649.
[24] H. Ohlsson and L. Ljung, "Identification of switched linear regression models using sum-of-norms regularization," *Automatica*, vol. 49, no. 4, pp. 1045 – 1050, 2013.
[25] I. Gurobi Optimization, "Gurobi optimizer reference manual," 2015. [Online]. Available: http://www.gurobi.com