2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Realization of MIMO–SLSs from Markov parameters via forward/backward corrections

Fethi Bencherki[1], Semiha Türkay[2] and Hüseyin Akçay[2]

Abstract— **This paper serves as a first identification step in a two-step model-based control synthesis problem of switched linear systems (SLSs). More precisely, we present an algorithm that addresses the realization of the multi-input/multi-output MIMO-SLSs from Markov parameters under mild assumptions on the dwell-times and the submodels. A key point of the proposed approach is the introduction of the forward and backward correction operators, which relieves the dependence on the choice of basis vectors in computing state-space matrices of the realizations. A numerical example illustrates the derived results.**

I. INTRODUCTION

Hybrid systems characterize the interplay between discrete and continuous phenomena. They have seen broad application span, in robotics [1], communication networks [2], networked control systems [3], computer vision [4], etc. The SLSs form an important class of hybrid systems governed by an external switching signal. A large body of the literature on hybrid system identification has focused on single-input/single-output (SISO) models, for example, the piecewise auto-regressive exogenous (PWARX) and the switched auto-regressive exogenous (SARX) models, see the survey in [5]. For MIMO systems, state-space models are preferable since for such system classes there exist a plethora of elegant theories on controller design [6], observer design [7], balanced truncation [8] and so on.

The estimation problem of the SLSs in the state-space form faces extra difficulty compared to that in the input/output form, since first, in state-space the continuous-state is unknown, and second, the discrete-states estimated locally, for example by the subspace algorithms [9], usually lie in different state-bases. As a result, the discrete-state estimates cannot be used to predict outputs to prescribed inputs without performing a state basis correction [10]. The difficulties in estimating the SLSs require hybrid identification algorithms to operate under various assumptions on the dwell times [11], the observability of the discrete states [12], and in some cases, knowledge of the switching sequence [13].

[1]F. Bencherki is with the Department of Automatic Control, Lund University, Box 117, SE-221 00, Sweden. E-mail: fethi.bencherki@control.lth.se

[2]S. Türkay and H. Akçay are with the Department of Electrical and Electronics Engineering, Eskişehir Technical University, Eskişehir, 26555 Turkey. E-mails: {semihaturkay, huakcay}@eskisehir.edu.tr

The first author is a member of the ELLIIT Strategic Research Area at Lund University.

A. Previous work

Realization of the discrete-time SLSs from input-output maps was studied in [14]. It is based on the theory of rational formal power series in noncommutative variables. The SLSs can also be considered a type of linear parameter-varying (LPV) systems with non-smooth and abrupt changes in the dynamics. The relationship between the input-output maps and the state-space representations for the class of LPV system was recently explored in several works [15], [16], [17]. In [18], a four-stage algorithm for the realization of the MIMO-SLSs from Markov parameters was proposed. Balanced truncation of the SLSs was studied in [19]. The authors therein derived a bound on the approximation error. The Loewner framework for model reduction was studied in [20] for the SLSs. Through this procedure, one can derive state-space models directly from input-output data.

B. Motivation for the paper

Control and stabilization of switched systems have received considerable attention in the literature. This divides into two main streams: the more recent data-driven approaches [21], [22], [23] and the classical model-based approaches [24], [25], [26], [6]. While the first category of control synthesis approaches either completely bypasses or maybe be characterized by an implicit system identification step, the second requires a model that motivates the need to develop identification/realization algorithms for the SLSs.

Despite the recent advances in the realization theory for LPV systems, it is still incomplete. The work reported in [18] relies on the earlier linear time-varying (LTV) realization results [28]. A preliminary analysis in [18] showed that if the state-space matrices of the LTV realization are subjected to forward/backward corrections, more accurate segment estimates may be derived. These corrections eliminate the dependence on the basis vectors selected to compute the LTV realization as well. We show here that under mild assumptions on the discrete states and the switching sequence, a simple realization algorithm could be built around these forward/backward correction operators, therefore simplifying the procedure significantly as compared to that in [18]. Ultimately, we aspire to develop an algorithm for the identification of the MIMO-SLSs from the input-out data based on the results derived in this paper and [18], [27], [10].

C. Organization of the paper

This paper is structured as follows. In Section II, earlier work on the realization problem from Markov parameters is reviewed and the problem studied in this paper is introduced. In Section III, the forward and backward correction operators are defined. In Section IV, an algorithm that estimates the discrete states and the switches from the SLS Markov parameters is proposed and the main result of the paper is stated. In Section V, a numerical example showcases the effectiveness of the proposed approach. Lastly, Section VI concludes the paper.

D. Notation

I_n refers to the identity matrix of size $n \times n$. $\|X\|_F$ denotes the Frobenius norm of a given matrix X. $0_{m\times n}$ denotes the m by n matrix of zeros.

II. PROBLEM FORMULATION

Let us consider the SLS represented by the state-space model

$$\begin{aligned} x(k+1) &= A_{\varphi(k)}x(k)+B_{\varphi(k)}u(k), \\ y(k) &= C_{\varphi(k)}x(k)+D_{\varphi(k)}u(k) \end{aligned} \tag{1}$$

where $u(k)\in\mathbb{R}^m$, $y(k)\in\mathbb{R}^p$, $x(k)\in\mathbb{R}^n$ for $1\le k\le N$ are the input, the output, and the continuous-state vectors. φ, on the other hand, is an external switching sequence. Let the quadruplets $\mathscr{P}_{\varphi(k)} = (A_{\varphi(k)},B_{\varphi(k)},C_{\varphi(k)},D_{\varphi(k)})$, $1\le k\le N$ denote the discrete-states. We assume that n and the number of the discrete-states σ are unknown, but bounded above by some integers n^* and σ^*, that is, $n\le n^*$ and $\sigma\le\sigma^*$. Assuming that $\varphi([1\;\; N]) = \mathbb{S} = \{1,\cdots,\sigma\}$, let $\mathscr{P} = \left\{\mathscr{P}_j : j\in\mathbb{S}\right\}$ be the set of discrete-states. Thus, φ partitions $[1\;\; N]$ into the segments such that $\varphi(k)=\varphi(k_i)$, $\forall k\in[k_i\;\; k_{i+1}) = I(i)$ and $\forall i\in[0\;\; i^*]$ where for convenience, we set $k_0=1$ and $k_{i^*+1}=N+1$. Let $\chi=\{k_i : 1\le i\le i^*\}$ be the set of switches . The dwell-times are defined by $\delta_i = k_{i+1}-k_i$, $0\le i<i^*$. The minimum dwell-time is then $\delta_* = \min_{0\le i\le i^*}\delta_i$. The Markov parameters of the SLS (1) are defined as

$$h(k,t) = \begin{cases} D(k), & k=t \\ C(k)\Phi(k,t+1)B(t), & k>t \\ 0, & k<t \end{cases} \tag{2}$$

where $\Phi(k,t)$ is the state-transition matrix defined by $\Phi(k,k)=I_n$ and $\Phi(k,t)=A(k-1)\cdots A(t)$ if $k>t$. The following realization problem was formulated in [18].

Problem 1: Given n^*, σ^*, and the Markov parameters $h(k,t)$, $\forall 1\le t\le k\le N$, determine n, σ, φ, and $\mathscr{P}_j\in\mathscr{P}$, $\forall 1\le j\le\sigma$ uniquely up to one similarity transformation only. A realization algorithm that solves Problem 1 was proposed in [18] under the following two assumptions.

Assumption 1 (submodels set): Every $\mathscr{P}_j\in\mathscr{P}$, $j=1,\cdots,\sigma$ is bounded-input/bounded-output (BIBO) stable and has a MacMillan degree n.

Assumption 2 (Switching sequence): φ satisfies $\delta_*\ge n^*$.

An LTI system of state dimension n that is both observable and controllable, and thus minimal, has Macmillan degree n. This is also true for (1) when Assumptions 1–2 hold and $k\in(2n^*\;\; N-2n^*+1]$, see Lemma 3.1 in [18]. This result allows one to extract an LTV realization that is *topologically equivalent* to (1) by robust numerical linear algebra techniques. Recall that two realizations $(A(k),B(k),C(k),D(k))$ and $(\check{A}(k),\check{B}(k),\check{C}(k),\check{D}(k))$ have the same Markov parameters if they are topologically equivalent, i.e., there exists a bounded matrix $T(k)\in\mathbb{R}^{n\times n}$ with bounded inverse, called a *Lyapunov transformation*, such that for all k

$$\check{A}(k) = T(k+1)A(k)T^{-1}(k),\quad \check{B}(k)=T(k+1)B(k),$$
$$\check{C}(k) = C(k)T^{-1}(k),\quad \check{D}(k)=D(k).$$

Input-output maps of topologically equivalent realizations are invariant: They produce the same outputs as identical inputs.

A. An LTV Realization from Markov Parameters

For all $k>2n^*$, we define the *Hankel matrices* by

$$\mathscr{H}(k) = \begin{bmatrix} h(k,k-1) & \cdots & h(k,k-2n^*) \\ \vdots & \ddots & \vdots \\ h(k+2n^*,k-1) & \cdots & h(k+2n^*,k-2n^*) \end{bmatrix}$$

which factorize as $\mathscr{H}(k)=\mathscr{O}(k)\mathscr{R}(k-1)$ [28] where the *extended observability/controllability matrices* are defined by

$$\mathscr{O}(k) = \begin{bmatrix} C(k+1) \\ C(k+1)\Phi(k+1,k) \\ \vdots \\ C(k+2n^*)\Phi(k+2n^*,k) \end{bmatrix}, \tag{3}$$

$$\mathscr{R}(k-1) = \begin{bmatrix} B(k-1) & \cdots & \Phi(k,k-2n^*+1)B(k-2n^*) \end{bmatrix}. \tag{4}$$

Since $\mathscr{O}(k)$ and $\mathscr{R}(k-1)$ are full-rank matrices and $\mathrm{rank}(\mathscr{H}(k))=n$, an LTV realization which is topologically equivalent to the SLS (1) may be extracted from the pairs $\{\mathscr{H}(k),\mathscr{H}(k+1)\}$, $2n^*<k\le N-2n^*$ by performing singular value decompositions (SVD) [18].

Observe that $\mathscr{O}(k)$ and $\mathscr{R}(k-1)$ may be determined up to two similarity transformations. In fact, let $J_\uparrow$ and $J_\downarrow$ denote the shift matrices of block-row up/down and $J_\leftarrow$ and $J_\rightarrow$ denote the block-column left/right shift matrices defined by

$$\begin{aligned} J_\uparrow &= \left[0_{2n^*p\times p}\;\; I_{2n^*p}\right],\; J_\downarrow = \left[I_{2n^*p}\;\; 0_{2n^*p\times p}\right], \\ J_\leftarrow &= \begin{bmatrix} 0_{m\times(2n^*-1)m} \\ I_{(2n^*-1)m} \end{bmatrix},\; J_\rightarrow = \begin{bmatrix} I_{(2n^*-1)m} \\ 0_{m\times(2n^*-1)m} \end{bmatrix}. \end{aligned}$$

Then, from the factorization formula with k and $k+1$ plugged in, we derive two formulas $J_\uparrow\mathscr{O}(k) = \left(J_\downarrow\mathscr{O}(k+1)\right)A(k)$ and $\mathscr{R}(k)J_\leftarrow = A(k)\left(\mathscr{R}(k-1)J_\rightarrow\right)$. Hence, we retrieve $A(k)$ from $A(k) = \left(J_\downarrow\mathscr{O}(k+1)\right)^\dagger J_\uparrow\mathscr{O}_q(k)$ where $X^\dagger = (X^TX)^{-1}X^T$.

It remains to estimate $\mathscr{O}(k)$ and $\mathscr{O}(k+1)$. Apply SVD to $\mathscr{H}(k)$ and $\mathscr{H}(k+1)$ yielding $\mathscr{H}(k)=U(k)\Sigma(k)V^T(k)$ and $\mathscr{H}(k+1)=U(k+1)\Sigma(k+1)V^T(k+1)$. Let

$$\hat{\mathscr{O}}(k) = U(k)\Sigma^{1/2}(k),\quad \hat{\mathscr{R}}(k-1)=\Sigma^{1/2}(k)V^T(k),$$
$$\hat{\mathscr{O}}(k+1) = U(k+1)\Sigma^{1/2}(k+1),$$
$$\hat{\mathscr{R}}(k) = \Sigma^{1/2}(k+1)V^T(k+1).$$

Up to two non-singular matrices $T(k)$ and $T(k+1)$, $\hat{\mathscr{O}}(k)$ and $\hat{\mathscr{O}}(k+1)$ provide estimates of $\mathscr{O}(k)$ and $\mathscr{O}(k+1)$, that is, $\mathscr{O}(k)=\hat{\mathscr{O}}(k)T(k)$ and $\mathscr{O}(k+1)=\hat{\mathscr{O}}(k+1)T(k+1)$. Let

$$\hat{A}(k) = \left(J_\downarrow\hat{\mathscr{O}}(k+1)\right)^\dagger J_\uparrow\hat{\mathscr{O}}(k) = T(k+1)A(k)T^{-1}(k).$$

Note that the SVDs reveal n. To estimate $B(k)$ and $C(k)$, let

$$J_{\mathrm{C}} = [I_p \;\; O_{p\times 2np}], \quad J_{\mathrm{B}} = \begin{bmatrix} I_m \\ 0_{(2n-1)m\times m} \end{bmatrix}.$$

Then, $C(k) = J_{\mathrm{C}}\mathscr{O}(k)$, $B(k) = \mathscr{R}(k)J_{\mathrm{B}}$ as the estimates of $C(k)$ and $B(k)$, set $\hat{C}(k) = J_{\mathrm{C}}\hat{\mathscr{O}}(k)$, $\hat{B}(k) = \hat{\mathscr{R}}(k)J_{\mathrm{B}}$. Then, $\hat{C}(k) = C(k)T^{-1}(k)$, $\hat{B}(k) = T(k+1)B(k)$, $\hat{D}(k) = h(k,k)$ and $\hat{\mathscr{P}}(k) = (\hat{A}(k), \hat{B}(k), \hat{C}(k), \hat{D}(k))$ is topologically equivalent to $\mathscr{P}(k)$ on $(2n \;\; N-2n]$. This derivation forms Algorithm 1.

Algorithm 1: LTV realization from Markov parameters

Input: Markov parameters $h(k+i, k-j)$, $0 \le i \le 2n^*$, $1 \le j \le 2n^*$, and $k \in (2n^* \;\; N-2n^*+1]$

1: **while** $k \in (2n^* \;\; N-2n^*]$ **do**
2: Compute $\mathscr{H}(k)$ and $\mathscr{H}(k+1)$
3: Compute the SVDs and determine n
4: Compute the extended observability/controllability matrix estimates
5: Estimate $\hat{A}(k)$, $\hat{C}(k)$, $\hat{B}(k)$, and set $\hat{D}(k) = h(k,k)$
6: $\hat{\mathscr{P}}(k) = (\hat{A}(k), \hat{B}(k), \hat{C}(k), \hat{D}(k))$
7: $k \leftarrow k+1$
8: **end**
9: Restore the state dimension : $n^* = n$

Output: $\hat{\mathscr{P}}(k)$, $\forall k \in (2n^* \;\; N-2n^*]$

As step 9 in algorithm 1 indicates, once n is learned, n^* is brought down to n, which shall be assumed in the analysis that follows in the later sections.

In Algorithm 1, it is not necessary to calculate the discrete-state estimates $\hat{\mathscr{P}}(k)$ over a continuum of points in time, see Step 11 in Algorithm 2. The estimates $\hat{\mathscr{P}}(k)$ are not similar to $\mathscr{P}(k)$, but they match the Markov parameters of (1). If a fixed basis is used in the SVDs and $2n+1$ points from the left and/or the right end of each segment of sufficient length are removed, the similarity $\hat{\mathscr{P}}(k) \sim \mathscr{P}(k)$ will be enforced on the rest of the points. The fixed basis means that identical matrices result in identical SVDs. The segments are estimated by computing the differences $\mathscr{H}(k+1) - \mathscr{H}(k)$ as k is varied between its limits, see [18] for further details.

III. Forward/backward corrections for SLSs

Let $\hat{\mathscr{O}}(k)$, $\hat{\mathscr{R}}(k)$ be defined as in (3-4) by employing the triplet $(\hat{A}(k), \hat{B}(k), \hat{C}(k))$ instead. For all $k \in (2n \;\; N-2n]$, let

$$\begin{aligned} \mathscr{V}(k) &= \mathscr{O}^{\dagger}(k)\mathscr{O}(k+1), \\ \hat{\mathscr{V}}(k) &= \hat{\mathscr{O}}^{\dagger}(k)\hat{\mathscr{O}}(k+1). \end{aligned}$$

The latter is called the *forward correction operator* and can be computable from the input-output data since $\hat{\mathscr{O}}(k)$ and $\hat{\mathscr{O}}(k+1)$ can be as well [27]. Note that the relationship holds

$$\hat{\mathscr{V}}(k) = T(k)\mathscr{V}(k)T^{-1}(k+1).$$

Now, pre-multiply $\hat{A}(k)$ and $\hat{B}(k)$ with $\hat{\mathscr{V}}(k)$ and let

$$\begin{aligned} \hat{A}_{\mathrm{f}}(k) &= \hat{\mathscr{V}}(k)\hat{A}(k) = T(k)\mathscr{V}(k)A(k)T^{-1}(k), \\ \hat{B}_{\mathrm{f}}(k) &= \hat{\mathscr{V}}(k)\hat{B}(k) = T(k)\mathscr{V}(k)B(k) \end{aligned}$$

and leave $\hat{C}(k)$, $\hat{D}(k)$ the same. Let

$$\begin{aligned} \mathscr{P}_{\mathrm{f}}(k) &= (\mathscr{V}(k)A(k), \mathscr{V}(k)B(k), C(k), D(k)), \\ \hat{\mathscr{P}}_{\mathrm{f}}(k) &= (\hat{A}_{\mathrm{f}}(k), \hat{B}_{\mathrm{f}}(k), \hat{C}(k), \hat{D}(k)). \end{aligned}$$

Then, $T(k) : \mathscr{P}_{\mathrm{f}}(k) \mapsto \hat{\mathscr{P}}_{\mathrm{f}}(k)$ is a *forward time-varying similarity transformation* and if $\mathscr{V}(k) = I_n$, $\hat{\mathscr{P}}_{\mathrm{f}}(k) \sim \mathscr{P}_{\varphi(k)}$, i.e, the forwardly corrected realization is in similarity with the true discrete-state at time k, which is a key observation to retrieving the discrete-states as we will see later. Similarly for all all $k \in (2n \;\; N-2n]$, we introduce the *backward correction operator* given by

$$\begin{aligned} \hat{\mathscr{W}}(k) &= \hat{\mathscr{R}}(k-1)\hat{\mathscr{R}}^{\dagger}(k), \\ \mathscr{W}(k) &= \mathscr{R}(k-1)\mathscr{R}^{\dagger}(k). \end{aligned}$$

Note the relationship $\hat{\mathscr{W}}(k) = T(k)\mathscr{W}(k)T^{-1}(k+1)$. Post-multiply $\hat{A}(k)$ and $\hat{C}(k)$ with $\hat{\mathscr{W}}(k)$

$$\begin{aligned} \hat{A}_{\mathrm{b}}(k) &= \hat{A}(k)\hat{\mathscr{W}}(k) = T(k+1)A(k)\mathscr{W}(k)T^{-1}(k+1), \\ \hat{C}_{\mathrm{b}}(k) &= \hat{C}(k)\hat{\mathscr{W}}(k) = C(k)\mathscr{W}(k)T^{-1}(k+1) \end{aligned}$$

and leave $\hat{B}(k)$ and $\hat{D}(k)$ the same. Let

$$\begin{aligned} \mathscr{P}_{\mathrm{b}}(k) &= (A(k)\mathscr{W}(k), B(k), C(k)\mathscr{W}(k), D(k)), \\ \hat{\mathscr{P}}_{\mathrm{b}}(k) &= (\hat{A}_{\mathrm{b}}(k), \hat{B}(k), \hat{C}_{\mathrm{b}}(k), \hat{D}(k)). \end{aligned}$$

Thus, $T(k+1) : \mathscr{P}_{\mathrm{b}}(k) \mapsto \hat{\mathscr{P}}_{\mathrm{b}}(k)$ is a *backward time-varying similarity transformation* and if $\mathscr{W}(k) = I_n$, $\hat{\mathscr{P}}_{\mathrm{b}}(k) \sim \mathscr{P}_{\varphi(k)}$.

The advantage of working with the forward/backward correction operators is that one could use arbitrary basis vectors in computing the SVDs while maintaining the similarity relations. The following lemma summarizes the properties of the forward and backward correction operators on constant φ intervals. It appeared as Proposition 3.2 in [18].

Lemma 1: Suppose that Assumptions 1–2 hold and φ is constant on $[a \;\; b] \subset (2n \;\; N-2n]$. Then, $\mathscr{V}(k) = I_n$, $\mathscr{P}_{\mathrm{f}}(k) \sim \mathscr{P}_{\varphi(k)}$, $\forall k \in [a \;\; b-2n-1]$ if $b-a \ge 2n+1$ and $\mathscr{W}(k) = I_n$, $\mathscr{P}_{\mathrm{b}}(k) \sim \mathscr{P}_{\varphi(k)}$, $\forall k \in [a+2n \;\; b]$ if $b-a \ge 2n$.

Recall the relationships $\hat{\mathscr{V}}(k) = T(k)\mathscr{V}(k)T^{-1}(k+1)$ and $\hat{\mathscr{W}}(k) = T(k)\mathscr{W}(k)T^{-1}(k+1)$. Since $T(k) = T(k+1)$ does not necessarily hold in $[a \;\; b-2n-1]$ or $[a+2n \;\; b]$ because arbitrary vector bases might have been used in the SVDs, we can't expect $\hat{\mathscr{V}}(k) \sim \mathscr{V}(k) = I_n$ or $\hat{\mathscr{W}}(k) \sim \mathscr{W}(k) = I_n$ to hold on these intervals. Yet, $\mathscr{P}_{\mathrm{f}}(k) \sim \mathscr{P}_{\varphi(k)}$, $\forall k \in [a \;\; b-2n-1]$ and $\mathscr{P}_{\mathrm{b}}(k) \sim \mathscr{P}_{\varphi(k)}$, $\forall k \in [a+2n \;\; b]$. Furthermore, these are the relationships all we need to extract the discrete-states and detect the switches. In fact, we apply Lemma 1 to the following intervals when considering the forward time-varying similarity transformation

$$\begin{aligned} I_{\mathrm{f}}(0) &= [2n+1 \;\; k_1-2n-2] \text{ if } \delta_0(\chi) \ge 4n+2, \\ I_{\mathrm{f}}(i) &= [k_i \;\; k_{i+1}-2n-2] \text{ if } i \notin \{0, i^*\} \text{ and } \delta_i(\chi) \ge 2n+2, \\ I_{\mathrm{f}}(i^*) &= [k_{i^*} \;\; N-4n-2] \text{ if } \delta_{i^*}(\chi) \ge 4n+3. \end{aligned}$$

In the case of the backward time-varying similarity transformation, these intervals are replaced by

$$\begin{aligned} I_{\mathrm{b}}(0) &= [4n+1 \;\; k_1-1] \text{ if } \delta_0(\chi) \ge 4n+1, \\ I_{\mathrm{b}}(i) &= [k_i+2n \;\; k_{i+1}-1] \text{ if } i \notin \{0, i^*\} \text{ and } \delta_i(\chi) \ge 2n+1, \\ I_{\mathrm{b}}(i^*) &= [k_{i^*}+2n \;\; N-2n-1] \text{ if } \delta_{i^*}(\chi) \ge 4n+2. \end{aligned}$$

The dwell times enforced in the forward/backward intervals I_f/I_b are to ensure those intervals are not empty sets. Whenever a dwell time constraint fails, the corresponding interval for that index must be replaced by the empty set $\emptyset$.

Corollary 1: Suppose that Assumptions 1–2 hold and $I_\mathrm{f}(i) \neq \emptyset$, $I_\mathrm{b}(i) \neq \emptyset$, $\forall i \in [0 \;\; i^*]$. Then, for all $i \in [0 \;\; i^*]$, $\mathscr{P}_\mathrm{f}(k) \sim \mathscr{P}_{\varphi(k)}$, $\forall k \in I_\mathrm{f}(i)$ and $\mathscr{P}_\mathrm{b}(k) \sim \mathscr{P}_{\varphi(k)}$, $\forall k \in I_\mathrm{b}(i)$.

The time-varying coordinate transformations were also utilized in the eigensystem realization algorithm [29] which uses input-output data from multiple experiments.

IV. Estimation of the discrete-states and the switching sequence

In this section, we combine the results in Sections II–III to propose a realization algorithm that solves Problem 1. Though the forward and backward correction operators can be applied to all points in $(2n \;\; N-2n]$, it is more efficient to priorly apply the first order difference operator presented next. This is the case since if a point belongs to the stationary point set of $\mathscr{H}$ it also belongs to $I_f \cap I_b$ as we shall see next.

A. The Stationary Point Set of $\mathscr{H}$

The gradient of the Hankel matrix $\mathscr{H}(k)$ in the forward direction is defined by $\delta_{\mathscr{H}}(k) = \mathscr{H}(k+1) - \mathscr{H}(k)$ for all $k \in (2n \;\; N-2n]$. The stationary point set of $\mathscr{H}$ is then

$$\mathbb{Z}_{\mathscr{H}} := \{k \in (2n \;\; N-2n] : \delta_{\mathscr{H}}(k) = 0\}.$$

It is possible to relate $\mathbb{Z}_{\mathscr{H}}$ to $I_\mathrm{f}(i)$ and $I_\mathrm{b}(i)$, $0 \le i \le i^*$ when (1) satisfies the following (one-step) switch detectability condition and the discrete-states have no poles at zero.

Assumption 3: The SLS (1) satisfies

$$\begin{bmatrix} C_{\varphi(k)} & D_{\varphi(k)} \end{bmatrix} \neq \begin{bmatrix} C_{\varphi(l)} & D_{\varphi(l)} \end{bmatrix} \Longleftrightarrow \varphi(k) \neq \varphi(l),$$
$$\begin{bmatrix} B_{\varphi(k)} \\ D_{\varphi(k)} \end{bmatrix} \neq \begin{bmatrix} B_{\varphi(l)} \\ D_{\varphi(l)} \end{bmatrix} \Longleftrightarrow \varphi(k) \neq \varphi(l),$$

Assumption 4: For all $1 \le j \le \sigma$, $\mathscr{P}_j$ has no poles at 0.

Let

$$K(0) = I_\mathrm{f}(0) \cap I_\mathrm{b}(0) \text{ if } \delta_0(\chi) \ge 6n+2,$$
$$K(i) = I_\mathrm{f}(i) \cap I_\mathrm{b}(i) \text{ if } i \notin \{0, i^*\} \text{ and } \delta_i(\chi) \ge 4n+2,$$
$$K(i^*) = I_\mathrm{f}(i^*) \cap I_\mathrm{b}(i^*) \text{ if } \delta_{i^*}(\chi) \ge 6n+3$$

and denote the cardinality of a given set X by $|X|$, that is, the number of the elements in X. Thus, $|K(i)| = 1$ even if $K(i)$ has length zero, i.e., a singleton and $K(i) \subset I(i)$, $\forall 0 \le i \le i^*$. The following appeared in [18] as Lemma 5.1.

Lemma 2: Consider Algorithm 1. Suppose that Assumptions 1–2 and 3–4 with $D(k) \equiv 0$ hold. Then, $\mathbb{Z}_{\mathscr{H}} \cap I(i) = K(i)$ if $|K(i)| \ge 1$, $0 \le i \le i^*$.

B. A realization algorithm

For each $K(i) = [\alpha_i \;\; \beta_i]$ with $|K(i)| \ge 1$, set $c_i = (\alpha_i + \beta_i)/2$ if $|K(i)|$ is odd and $c_i = (\alpha_i + \beta_i - 1)/2$ if $|K(i)|$ is even. Let $\mathbb{I}_\nu = \{i \in [0 \;\; i^*] : |K(i)| \ge \nu\}$, $\nu \in \mathbb{N}$. From Corollary 1, $\mathscr{P}_\mathrm{f}(c_i) \sim \mathscr{P}_{\varphi(c_i)}$ and $\mathscr{P}_\mathrm{b}(c_i) \sim \mathscr{P}_{\varphi(c_i)}$ if $\nu \neq 0$. These relations generate the equivalance classes on the collection $I(i)$, $0 \le i \le i^*$ and from each equivalance class, a representative discrete-state may be selected. The number of such choices is bounded above by σ. If $\varphi(\mathbb{I}_\nu) = \mathbb{S}$ for some $\nu \in \mathbb{N}$, all the discrete-states in $\mathscr{P}$ are uniquely recovered up to σ similarity transformations. Obviously ν is a design variable but larger values are desired since they permit easier recovery. The equivalence classes can be retrieved by a clustering algorithm [30]. A simple feature for clustering is the sequence $\mathscr{M}_F(k) = \|\mathscr{H}(k)\|_F$ for $2n < k \le N-2n$. This choice is consistent with the definitions of $\mathscr{P}_\mathrm{f}(k)$ or $\mathscr{P}_\mathrm{b}(k)$ if $\nu \neq 0$ since on the interval $[\alpha_i \;\; \beta_i]$, $\mathscr{H}(k) = \mathscr{H}(k+1)$ and therefore $\mathscr{V}(k) = I_n$ and $\mathscr{W}(k) = I_n$ which is key in retrieving the discrete-state set. From Lemma 2, observe that

$$\begin{aligned} k_{i+1} &= \beta_i + 2n + 2, \; i \in [0 \;\; i^*), \\ k_i &= \alpha_i - 2n, \; i \in (0 \;\; i^*] \end{aligned} \tag{5}$$

and the recovery will be complete if $|K(i)| \ge 1$ for all $0 \le i \le i^*$. When $|K(i)| \ge 1$ does not hold for some $i \in [0 \;\; i^*]$, k_i and/or k_{i+1} can be detected by alternative iterative switch detection algorithms reported in [18]. Furthermore, the basis construction scheme outlined there nicely fits into the realization scheme of this paper. It must be applied when the realized SLS model is to be used for predicting outputs to prescribed inputs. An error analysis can be carried out analogously to [18] to show that Algorithm 2 is robust to amplitude bounded noise.

Let $\mathbb{Z}_{\mathscr{H},\varepsilon} = \{k \in (2n \;\; N-2n] : \|\delta_{\mathscr{H}}(k)\|_F < \varepsilon\}$. Since $N < \infty$, $i^* < \infty$, and the Markov parameters are noiseless, $\mathbb{Z}_{\mathscr{H}} = \mathbb{Z}_{\mathscr{H},\varepsilon_\mathbb{Z}}$ for some $\varepsilon_\mathbb{Z} > 0$ from $\mathbb{Z}_{\mathscr{H}} = \cap_{\varepsilon > 0} \mathbb{Z}_{\mathscr{H},\varepsilon}$. This fact is useful in the implementation of $\delta_{\mathscr{H}}(k) = 0$, which is essential for the retrieval of the stationary point set. We summarize the results derived in this section in the following theorem.

Theorem 1: Consider Algorithm 2 with the noiseless Markov parameters of the SLS (1). Suppose that Assumptions 1-2, 3–4 with $D(k) \equiv 0$ hold. Then, Algorithm 2 recovers $\mathscr{P}$ and χ if $|K(i)| \ge 1$ for all $0 \le i \le i^*$.

Algorithm 2: SLS realization from Markov parameters

Input: Markov parameters $h(k+i, k-j)$ for $0 \le i \le 2n$, $0 \le j \le 2n$, $2n < k \le N-2n$, a small $\varepsilon_\mathbb{Z} > 0$, a large ν, and $\hat{\mathscr{P}}(k)$, $2n < k \le N-2n$ from Algorithm 1

Compute $\delta_{\mathscr{H}}(k)$ for all $2n < k \le N-2n$
Initialize $\mathbb{Z}_{\mathscr{H},\varepsilon_\mathbb{Z}} = \phi$
while $k \in (2n^* \;\; N-2n^*]$ **do**
    if $\|\delta_{\mathscr{H}}(k)\|_F \le \varepsilon_\mathbb{Z}$ **then**
        $\mathbb{Z}_{\mathscr{H},\varepsilon_\mathbb{Z}} = \mathbb{Z}_{\mathscr{H},\varepsilon_\mathbb{Z}} \cup \{k\}$
    end
    $k \leftarrow k+1$
end
Determine points $\hat{c}_i \in \hat{K}(i)$, $i \in \mathbb{I}_\nu$
Choose σ points from $\{\hat{c}_i : i \in \mathbb{I}_\nu\}$ by clustering $\mathscr{M}_F(\hat{c}_i)$
Calculate $\hat{\mathscr{P}}_\mathrm{f}(\hat{c}_i)$ or $\hat{\mathscr{P}}_\mathrm{f}(\hat{c}_i)$, $1 \le i \le \sigma$ from Algorithm 1
Estimate $\hat{k}_i$ and $\hat{k}_{i+1}$ from (5) for all $0 \le i \le i^*$

Output: $\hat{\mathscr{P}}$ and $\hat{\chi}$

The discrete-states obtained from Algorithm 2 should undergo a basis correction to produce an identical input-output map to that of the original system, see for instance

Section 6 in [18] and reference [10] for basis correction recipes for SLS.

V. Numerical example

A. SLS Realization in a Noiseless Setup

Consider the MIMO-SLS adopted from [31]. The discrete-states are given by

$$A_1 = \begin{pmatrix} 0.4 & 0.1 & 0 \\ 0.8 & 0.4 & 0 \\ 0 & 0 & 0.8 \end{pmatrix}, B_1 = \begin{pmatrix} 1.5 & 0.9 \\ 1 & -1 \\ -1.5 & 2.3 \end{pmatrix},$$
$$C_1 = \begin{pmatrix} 0.8 & 1.1 & 2 \\ -1.3 & 0.7 & 1.7 \\ 1.5 & 0.7 & -0.9 \end{pmatrix}, D_1 = 0;$$

$$A_2 = \begin{pmatrix} 0.3 & 0.2 & 0 \\ 0.8 & 0.2 & 0 \\ 0 & 0 & -0.75 \end{pmatrix}, B_2 = B_1, C_2 = C_1, D_2 = 0.$$

Note that this SLS conforms to Assumptions 1, 3–4. A switching sequence that conforms with Assumption 2 was generated. It is plotted in Figure 1.

We generated the noiseless Markov parameters for $\{\mathscr{P}, \varphi\}$ and run Algorithm 1, thus yielding a topologically equivalent realization to (1) on $(2n \ \ N-2n]$. Next, we run Algorithm 2 with the Markov parameter sequence and the state-space matrices in $\hat{\mathscr{P}}(k)$ for $2n < k \leq N-2n$ returned by Algorithm 1. Clustering over the feature space $\mathscr{M}_F(k)$, $k \in \mathbb{Z}_{\mathscr{H},\varepsilon_{\mathbb{Z}}}$ with $\varepsilon_{\mathbb{Z}} = 10^{-5}$ using the DBSCAN algorithm [30] revealed an estimate of σ denoted by $\hat{\sigma}$. The feature for clustering $\mathscr{M}_F(k)$ and the histogram of clustering are plotted in Figure 2. The number of the submodels was correctly identified, i.e., $\hat{\sigma} = \sigma = 2$. Next, according to Step 10 of the algorithm we targeted the σ intervals characterized by $\hat{K}(i)$, $i \in \mathbb{I}_\nu$ where we picked $\nu = 30$. Recall that the corrected realizations over the mid-points of these intervals are the discrete-state estimates. The estimated eigenvalues retrieved from the estimated discrete states are found to be an exact match to the true ones since the algorithm was driven via noiseless Markov parameters. Note the perfect match between the two sets. Step 12 of Algorithm 2 allowed us to recover the entire switching signal since $|K(i)| \geq 1$ for all $0 \leq i \leq i^*$ for the SLS in the example. The estimated switching signal and the true one are plotted in Figure 3.

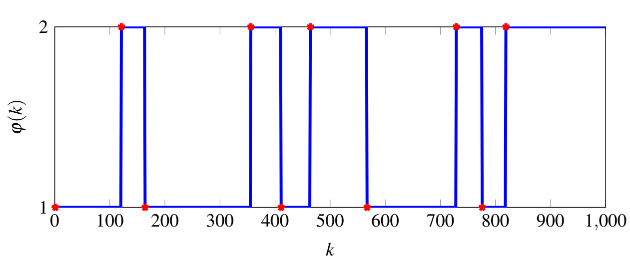
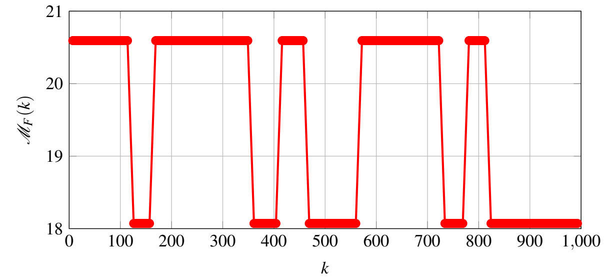
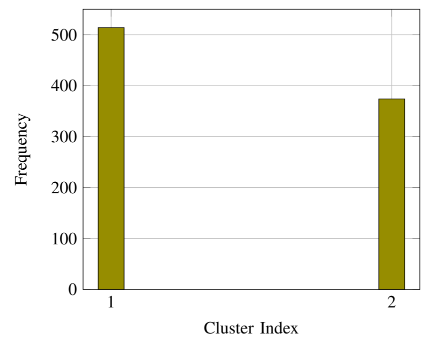


Fig. 1: The switching sequence.



(a) $\mathscr{M}_F(k)$, $k \in \mathbb{Z}_{\mathscr{H},\varepsilon_{\mathbb{Z}}}$.



(b) The histogram of the clusters.

Fig. 2: The retrieval of σ via the DBSCAN algorithm [30].

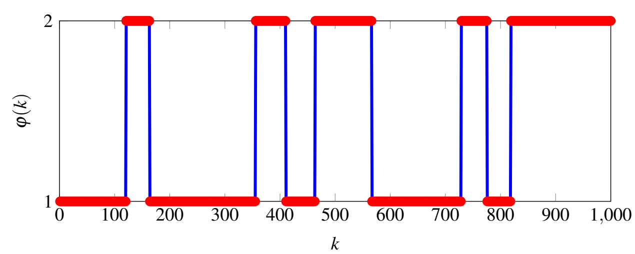


Fig. 3: The switching sequence estimate ('∘') and the true switching sequence ('-').

B. SLS Realization in a Noisy Setup

We numerically show that the proposed algorithm persists a good performance when driven with noise corrupted Markov parameters. We consider the following SISO-SLS adapted from [13]

$$A_1 = \begin{pmatrix} 0 & 0.8 \\ -0.8 & 0.5 \end{pmatrix}, B_1 = \begin{pmatrix} 0.4 \\ 0 \end{pmatrix}, C_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}, D_1 = 0;$$

$$A_2 = \begin{pmatrix} 0 & 0.5 \\ -0.5 & 0 \end{pmatrix}, B_2 = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}, C_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, D_2 = 0.$$

We drive Algorithm 1 with noisy Markov parameters belonging to this SLS. We assess the performance by employing two different performance metrics. First, we compute the relative errors $\|M_j - \hat{M}_j\|_F / \|M_j\|_F$ for $j = 1,2$ where $\|M_j\|_F^2 = \|A_j\|_F^2 + \|B_j\|_F^2 + \|C_j\|^2 + \|D_j\|_F^2$ and sum over $\mathbb{S}$. The result is denoted by $\delta_{\mathscr{P}}$. Second, a measure of fit for the switching sequence estimates is the percentage of the

TABLE I: The average performance of Algorithm 2 when driven with noisy Markov parameters.

Errors	$\delta_{\mathscr{P}}$			$\text{FIT}_{\varphi}(\%)$		
SNR (dB)	50dB	40dB	30dB	50dB	40dB	30dB
Mean	0.006	0.0.0174	0.0613	100	99.98	99.45

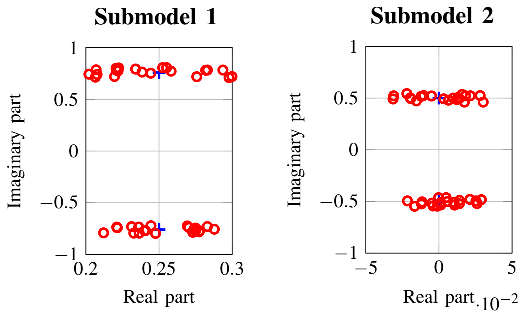


Fig. 4: The estimated ('∘') and the true ('+') submodel eigenvalues for different noise realizations and $\text{SNR} = 30db$.

correctly classified points. It is called FIT_{φ} and calculated via the formula

$$\text{FIT}_{\varphi} = \left(1 - \frac{1}{N}\sum_{k=1}^{N} \text{sign}\,|\hat{\varphi}(k) - \varphi(k)|\right) \times 100\%.$$

Table I displays the average performance of Algorithm 2 computed over 100 noise realizations for different signal-to-noise ratio (SNR) in terms of both performance metrics, $\delta_{\mathscr{P}}$ and FIT_{φ}. Lastly, the estimated eigenvalues are contrasted to the true ones in Figure 4.

VI. Summary

In this paper, we studied the problem of MIMO-SLS realization from the Markov parameters initiated in [18]. We proposed a simplified algorithm based on the forward and backward corrections of an LTV realization that is topologically equivalent to the SLS estimated from the Markov parameters.

References

[1] T. Schlegl, M. Buss, and G. Smith, "A hybrid systems approach toward modeling and dynamic simulation of dexterous manipulation," *IEEE/ASME Trans. Mechatronics*, vol. 8, pp. 352–361, 2003.

[2] J. Lee, S. Bohacek, J. P. Hespanha, and K. Obraczka, "Modelling communication networks with hybrid systems," *IEEE/ACM Trans. Networking*, vol. 15, pp. 630–643, 2007.

[3] G. S. Deaecto, M. Souza, and J. C. Geromel, "Discrete-time switched linear systems state feedback design with application to networked control," *IEEE Trans. Automatic Control*, vol. 60, pp. 877–881, 2014.

[4] R. Vidal and Y. Ma, "A unified algebraic approach to 2-d and 3-d motion segmentation and estimation," *Journal of Mathematical Imaging and Vision*, vol. 25, pp. 403–421, 2006.

[5] A. Garulli, S. Paoletti, and A. Vicino, "A survey on switched and piecewise affine system identification," *IFAC Proc. Volumes*, vol. 45, pp. 344–355, 2012.

[6] J. C. Geromel and P. Colaneri, "Stability and stabilization of discrete time switched systems," *Int. J. Control*, vol. 79, pp. 719–728, 2006.

[7] W. Chen and S. Mehrdad, "Observer design for linear switched control systems," in: Proc. *2004 ACC*, Boston, MA, June-July 2004, pages 5796–5801.

[8] I. V. Gosea, M. Petreczky, A. C. Antoulas, and C. Fiter, "Balanced truncation for linear switched systems," *Advances in Computational Mathematics*, vol. 44, pp. 1845–1886, 2018.

[9] M. Verhaegen and P. Dewilde, "Subspace model identification, Part 1. The output-error state-space model identification class of algorithm," *Int. J. Control*, vol. 56, pp. 1187–1210, 1992.

[10] F. Bencherki, S. Türkay, and H. Akçay, "Basis transform in switched linear system state-space models from input-output data," *Int. J. Adaptive Control and Signal Processing*, pp. 1–18, doi: 10.1002/acs.3677, 2023.

[11] K. Huang, A. Wagner, and Y. Ma, "Identification of hybrid linear time-invariant systems via subspace embedding and segmentation (SES)," in: Proc. *43rd IEEE CDC*, Paradise Island, Bahamas, December 2004, pages 3227–3234.

[12] R. Vidal, A. Chiuso, and S. Soatto, "Observability and identifiability of jump linear systems," in: Proc. *41st IEEE CDC*, Las Vegas, NV, December 2002, pages 3614–3619.

[13] V. Verdult and M. Verhaegen, "Subspace identification of piecewise linear systems," in: Proc. *43rd IEEE CDC*, Paradise Island, Bahamas, December 2004, pages 3838–3843.

[14] M. Petreczky, L. Bako, and J. H. Van Schuppen, "Realization theory of discrete-time linear switched systems," *Automatica*, vol. 49, pp. 3337–3344, 2013.

[15] M. Petreczky, R. Tóth, and G. Mercère, "Realization theory for LPV state-space representations with affine dependence," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4667–4674, 2017.

[16] M. Petreczky, R. Tóth, and G. Mercère, "LPV-ARX representations of LPV state-pace models with affine dependence," *Systems & Control Letters*, vol. 173, p. 105459, 2023.

[17] R. Tóth, H. S. Abbas, and H. Werner, "On the state-space realization of LPV input-output models: Practical approaches," *IEEE Trans. Control Systems Technology*, vol. 20, pp. 139–153, 2011.

[18] F. Bencherki, S. Türkay, and H. Akçay, "Realization of multi-input/multi-output switched linear systems from Markov parameters," *Nonlinear Analysis: Hybrid Systems*, vol. 48, p. 101311, 2023.

[19] M. Petreczky, R. Wisniewski, and J. Leth, "Balanced truncation for linear switched systems," *Nonlinear Analysis: Hybrid Systems*, vol. 10, pp. 4–20, 2013.

[20] I. V. Gosea, M. Petreczky, and A. C. Antoulas, "Data-driven model order reduction of linear switched systems in the Loewner framework," *SIAM J. Sci. Comput.*, vol. 40, pp. B572–B610, 2018.

[21] M. Rotulo, C. De Persis, and P. Tesi, "Online learning of data-driven controllers for unknown switched linear systems," *Automatica*, vol. 145, p. 110519, 2022.

[22] M. Rotulo, C. De Persis, and P. Tesi, "Online data-driven stabilization of switched linear systems," in: Proc. *2021 ECC*, Delft, Netherlands, June-July 2021, pp. 300–305.

[23] V. Breschi and S. Formentin, "Direct data-driven design of switching controllers," *Int. J. Robust and Nonlinear Control*, vol. 30, pp. 6042–6072, 2020.

[24] M. Philippe, R. Essick, G. E. Dullerud, and R. M. Jungers, "Stability of discrete-time switching systems with constrained switching sequences," *Automatica*, vol. 72, pp. 242–250, 2016.

[25] M. S. Branicky, "Multiple lyapunov functions and other analysis tools for switched and hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, pp. 475–482, 1998.

[26] D. Cheng, L. Guo, Y. Lin, and Y. Wang, "Stabilization of switched linear systems," *IEEE Transactions on Automatic Control*, vol. 50, pp. 661–666, 2005.

[27] F. Bencherki, S. Türkay, and H. Akçay, "Observer-based switched-linear system identification," *arXiv:2107.14571 [eess.SY]*, 2023.

[28] S. Shokoohi and L. M. Silverman, "Identification and model reduction of time-varying discrete-time systems," *Automatica*, vol. 23, pp. 509–521, 1987.

[29] M. Majji, J.-N. Juang, and J. L. Junkins, "Time-varying eigensystem realization algorithm," J. Guid., Contr., Dyn., vol. 33, pp. 13–28, 2010.

[30] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in: Proc. 2^{nd} *Int. Conf. Knowl. Disc. Data Min.*, August 1996, pages 226–231.

[31] M. G. Sefidmazgi, M. M. Kordmahalleh, A. Homaifar, A. Karimoddini, and E. Tunstel, "A bounded switching approach for identification of switched MIMO systems," in: Proc. *IEEE SMC*, October 2016, pages 4743–4748.