2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

On the notion of persistence of excitation for linear switched systems

Mihály Petreczky and Laurent Bako
Univ Lille Nord de France, F-59000 Lille, France
Ecole des Mines de Douai, IA, F-59500 Douai, France,
{mihaly.petreczky,laurent.bako}@mines-douai.fr.

Abstract— **The paper formulates the concept of persistence of excitation for discrete-time linear switched systems. In addition, the paper provides sufficient conditions for an input signal to be persistently exciting. Persistence of excitation is formulated as a property of the input signal, and it is not tied to any specific identification algorithm. The results of the paper rely on realization theory and on the notion of Markov-parameters for linear switched systems.**

I. INTRODUCTION

The paper formulates the concept of persistence of excitation for *discrete-time linear switched systems* (abbreviated by DTLSSs). DTLSSs are one of the simplest and best studied classes of hybrid systems, [22]. A DTLSS is a discrete-time switched system, such that the continuous sub-systems are linear. The switching signal is viewed as an external input, and all linear systems live on the same input-output- and state-space.

We define the notion of persistence of excitation for input signals as follows. Fix an input-output map and an input signal. We call the input signal persistently exciting for this input-output map, if the response of this map to that particular input determines the map uniquely. That is, the response of the system to a persistently exciting input is sufficient to predict the response of the system to **any** input.

Persistence of excitation is essential for system identification and adaptive control. Normally, the system of interest is tested only for one input sequence. In system identification this is due to practical limitations, and in adaptive control this is implied by the problem formulation. The objective is to find a model of the system based on the response to the chosen input signal. However, the knowledge of a model of the system implies that the response to **any** input signal is known. Hence, persistence of excitation of the input signal is a prerequisite for finding such a model.

Persistence of excitation is a joint property of the input and of the system. A particular input might be persistently exciting for a particular system and it might fail to be persistently exciting for another system. It is not even clear if every system admits a persistently exciting input.

In the existing literature, persistence of excitation is often defined as a specific property of the measurements which is **sufficient** for the correctness of **some** identification algorithm. In contrast, in this paper we propose a definition of persistence of excitation which is **necessary** for the correctness of **any** identification algorithm. Obviously, the two approaches are complementary. In fact, we hope that the results of this paper can serve as a starting point to derive persistence of excitation conditions for specific identification algorithms.

Contribution of the paper We define persistence of excitation for finite input sequences and persistence of excitation for infinite input sequences.

We will show the following. Assume that f is an input-output map which is realizable by a *reversible*[1] DTLSS. Then there exists a finite input sequence which is persistently exciting for the input-output map f. Moreover, we present a procedure for constructing such an input sequence.

Furthermore, we show that there exists a class of infinite input sequences, such that the following holds. Each sequence from this class is persistently exciting for all the input-output maps which are realizable by a *stable DTLSS*. The conditions which the input sequence must satisfy is that **(a)** each finite sequence of discrete modes occurs there infinitely often (i.e. the switching signal is rich enough) and **(b)** that the continuous input is a colored noise. Hence, this result is consistent with the classical result for linear systems.

It might be appealing to interpret the conditions above as ones which ensure that

- one stays in every discrete mode long enough and,
- the continuous input is persistently exciting in the classical sense.

One could then try to identify the linear subsystems separately and merge the results. However, this approach is in general incorrect. The reason for this is that there exists a broad class of input-output maps which can be realized by a linear switched system but not by a linear switched system, linear subsystems of which are minimal, [19]. The above scheme would not work for such systems. In fact, for such systems one has to test the systems response not only for each discrete mode, but for each combination of discrete modes.

The main idea behind the definition of persistence of excitation and the subsequent results is as follows. From realization theory for DTLSSs [19] we know that the knowledge of (finitely many) Markov-parameters is sufficient for computing a DTLSS state-space representation of the input-output map. Conversely, the Markov-parameters can always be computed from the matrices of a DTLSS state-space representation. We call an input sequence persistently exciting, if the Markov-parameters of the input-output map

[1]The formal definition of reversibility will be presented later

can be computed from the response of the map to that input. Combined with the previous remark, it means that one can compute a DTLSS representation of the input-output map from its response to a persistently exciting input. The latter means that the response to a persistently exciting input uniquely determines the input-output map.

Motivation of the system class The class of DTLSSs is the simplest and perhaps the best studied class of hybrid systems. In addition to its practical relevance, it also serves as a convenient starting point for theoretical investigations. In particular, any piecewise-affine hybrid system can be viewed as a feedback interconnection of a DTLSS with an event generating device. Hence, identification of a piecewise-affine system is related to the problem of closed-loop identification of a DTLSS. For the latter, it is indispensable to have a good notion of persistence of excitation for DTLSSs.

Related work Identification of hybrid systems is an active research area, with several significant contributions [8], [14], [15], [24], [11], [5], [12], [10], [8], [3], [2], [21], [6], [23], [16], [1]. While progress has been made in finding identification algorithms, the fundamental theoretical properties of these algorithms are only partially understood. Persistence of excitation of hybrid systems were already addressed in [25], [24], [23], [9]. However, the conditions of those papers are more method specific and their approach is quite different from the one we propose. For linear systems, persistence of excitation has thoroughly been investigated, see for example [13], [26] and the references therein.

Outline of the paper §II presents the formal definition of DTLSSs and it formulates the major system-theoretic concepts for this system class. §III presents a brief overview of realization theory for DTLSSs. §IV presents the main contribution of the paper.

Notation Denote by $\mathbb{N}$ the set of natural numbers including 0. The notation described below is standard in automata theory, see [4]. Consider a set X which will be called the *alphabet*. Denote by X^* the set of finite sequences of elements of X. Finite sequences of elements of X are referred to as *strings* or *words* over X. Each non-empty word w is of the form $w = a_1 a_2 \cdots a_k$ for some $a_1, a_2, \dots, a_k \in X$. The element a_i is called the *ith letter of* w, for $i = 1, \dots, k$ and k is called the *length of* w. We denote by ϵ the *empty sequence (word)*. The length of word w is denoted by $|w|$;note that $|\epsilon| = 0$. We denote by X^+ the set of non-empty words, i.e. $X^+ = X^* \setminus \{\epsilon\}$. We denote by wv the concatenation of word $w \in X^*$ with $v \in X^*$. For each $j = 1, \dots, m$, e_j is the jth unit vector of $\mathbb{R}^m$, i.e. $e_j = (\delta_{1,j}, \dots, \delta_{n,j})$, $\delta_{i,j}$ is the Kronecker symbol. For each $x \in \mathbb{R}^n$, $n > 0$ we denote the Euclidean norm of x by $||x||_2$.

II. LINEAR SWITCHED SYSTEMS

In this section we present the formal definition of DTLSSs along with a number of relevant system-theoretic concepts for DTLSSs. The presentation is based on [19], [17].

Definition 1: A discrete-time linear switched system (abbreviated by DTLSS), is a discrete-time control system of the form

$$\Sigma \begin{cases} x_{t+1} &= A_{q_t} x_t + B_{q_t} u_t \text{ and } x_0 = 0 \\ y_t &= C_{q_t} x_t. \end{cases} \tag{1}$$

Here $Q = \{1, \dots, D\}$ is the finite set of discrete modes, D is a positive integer, $q_t \in Q$ is the switching signal, $u_t \in \mathbb{R}$ is the continuous input, $y_t \in \mathbb{R}^p$ is the output and $A_q \in \mathbb{R}^{n \times n}$, $B_q \in \mathbb{R}^{n \times m}$, $C_q \in \mathbb{R}^{p \times n}$ are the matrices of the linear system in mode $q \in Q$.
Throughout the section, Σ *denotes a DTLSS of the form (1).* The *inputs of* Σ *are the continuous inputs* $\{u_t\}_{t=0}^{\infty}$ *and the switching signal* $\{q_t\}_{t=0}^{\infty}$. The state of the system at time t is x_t. Note *that any switching signal is admissible and that the initial state is assumed to be zero*. We use the following notation for the inputs of Σ.

Notation 1 (Hybrid inputs): Denote $\mathcal{U} = Q \times \mathbb{R}^m$.
We denote by $\mathcal{U}^*$ (resp. $\mathcal{U}^+$) the set of all finite (resp. non-empty and finite) sequences of elements of $\mathcal{U}$. A sequence

$$w = (q_0, u_0) \cdots (q_t, u_t) \in \mathcal{U}^+, \, t \geq 0 \tag{2}$$

describes the scenario, when the discrete mode q_i and the continuous input u_i are fed to Σ at time i, for $i = 0, \dots, t$.

Definition 2 (State and output): Consider a state $x_{init} \in \mathbb{R}^n$. For any $w \in \mathcal{U}^+$ of the form (2), denote by $x_\Sigma(x_{init}, w)$ *the state of* Σ at time $t + 1$, and denote by $y_\Sigma(x_{init}, w)$ the *output* of Σ at time t, if Σ is started from x_{init} and the inputs $\{u_i\}_{i=0}^{t}$ and the discrete modes $\{q_i\}_{i=0}^{t}$ are fed to the system.

Definition 3 (Input-output map): The map $y_\Sigma : \mathcal{U}^+ \to \mathbb{R}^p$, $\forall w \in \mathcal{U}^+ : y_\Sigma(w) = y(0, w)$, is called the input-output map of Σ.
That is, the input-output map of Σ maps each sequence $w \in \mathcal{U}^+$ to the output generated by Σ under the hybrid input w, if started from the zero initial state. The definition above implies that the input-output behavior of a DTLSS can be formalized as a map

$$f : \mathcal{U}^+ \to \mathbb{R}^p. \tag{3}$$

The value $f(w)$ for w of the form (2) represents the output of the underlying black-box system at time t, if the continuous inputs $\{u_i\}_{i=0}^{t}$ and the switching sequence $\{q_i\}_{i=0}^{t}$ are fed to the system.

Definition 4 (Realization): The DTLSS Σ is a *realization* of an input-output map f of the form (3), if f equals the input-output map of Σ, i.e. if $f = y_\Sigma$.
For the notions of observability, reachability and span-reachability of DTLSSs we refer the reader to [19], [22].

Definition 5 (Dimension): The dimension of Σ, denoted by $\dim \Sigma$, is the dimension n of its state-space.

Definition 6 (Minimality): Let f be an input-output map. Then Σ is *a minimal realization of* f, if Σ is a realization of f, and for any DTLSS $\hat{\Sigma}$ which is a realization of f, $\dim \Sigma \leq \dim \hat{\Sigma}$.

III. OVERVIEW OF REALIZATION THEORY

Below we present an overview of those results on realization theory of DTLSSs which are relevant for defining the notion of persistence of excitation. For more details on the topic see [19].

In the sequel, Σ *denotes a DTLSS of the form (1), and* f *denotes an input-output map* $f : \mathcal{U}^+ \to \mathbb{R}^p$. For our purposes the most important result is the one which states that a DTLSS realization of f can be computed from the Markov-parameters of f. The reason why this is so important is that we are going to use the notion of Markov-parameters to define the concept of persistence of excitation.

In order to present this result, we need to define the Markov-parameters of f formally.

Definition 7 (Markov-parameters): Denote $Q^{2,*} = \{w \in Q^* \mid |w| \geq 2\}$. Define the maps $S^f_j : Q^{2,*} \to \mathbb{R}^p$, $j = 1, \ldots, m$ as follows; for any $v = \sigma_1 \ldots \sigma_{|v|} \in Q^*$ with $\sigma_1, \ldots, \sigma_k \in Q$, and for any $q, q_0 \in Q$,

$$S^f_j(q_0 v q) = \begin{cases} f\big((q_0, e_j)(q, 0)\big) \text{ if } v = \epsilon \\ f\big((q_0, e_j)(\sigma_1, 0) \ldots (\sigma_{|v|}, 0)(q, 0)\big) \text{if } |v| \geq 1. \end{cases} \tag{4}$$

with $e_j \in \mathbb{R}^m$ is the vector with 1 as its jth entry and zero everywhere else. For each $w \in Q^*$, define the matrix $S^f(w)$ as follows

$$S^f(w) = \begin{bmatrix} S^f_1(w) & \ldots & S^f_m(w) \end{bmatrix} \in \mathbb{R}^{p \times m}.$$

With the notation above, the *Markov-parameter* $M^f(v)$ of f indexed by the word $v \in Q^*$ is the following $pD \times Dm$ matrix

$$M^f(v) = \begin{bmatrix} S^f(1v1), & \cdots, & S^f(Dv1) \\ S^f(1v2), & \cdots, & S^f(Dv2) \\ \vdots & \cdots & \vdots \\ S^f(1vD), & \cdots, & S^f(DvD) \end{bmatrix}. \tag{5}$$

Note that the values of the Markov-parameters can be obtained from the values of f. The matrices $S^f(w)$ can be viewed as *impulse responses* for the switching sequence w. The Markov-parameter $M^f(v)$ is then just the collection of the impulse responses for the switching sequences of the form $q_0 v q$ for all the possible values of $q, q_0 \in Q$. That is, we vary the first and the last discrete mode, and we assume that between the first and the last discrete modes the system switches according to v.

The interpretation of $S^f(w)$ will become more clear after we define the concept of a *generalized convolution representation*. For the latter, we need the following notation.

Notation 2 (Sub-word): Consider the sequence $v = q_0 \cdots q_t \in Q^+$, $q_0, \ldots, q_t \in Q$, $t \geq 0$. For each $j, k \in \{0, \ldots, t\}$, define the word $v_{j|k} \in Q^*$ as follows; if $j > k$, then $v_{j|k} = \epsilon$, if $j = k$, then $v_{j|j} = q_j$ and if $j < k$, then $v_{j|k} = q_j q_{j+1} \cdots q_k$. That is, $v_{j|k}$ is the sub-word of v formed by the letters from the jth to the kth letter.

Definition 8 (Convolution representation): The input-output map f has a *generalized convolution representation (abbreviated as **GCR**)*, if for all $w \in \mathcal{U}^+$ of the form (2), $f(w)$ can be expressed as follows.

$$f(w) = \sum_{k=0}^{t-1} S^f(q_k \cdot v_{k+1|t-1} \cdot q_t) u_k.$$

Note that if f has a **GCR**, then the Markov-parameters of f determine f uniquely. In fact, existence of a **GCR** means that the response of f to an input is a linear combination of the continuous inputs, and the Markov-parameters serve as coefficients of this linear combination. The existence of a **GCR** is a necessary condition for realizability by DTLSSs.

The concept of a **GCR** is a generalization of the known fact that the response y_t of a linear system to the inputs $u_0, \ldots, u_t$ is of the form

$$y_t = \sum_{k=0}^{t} M_{t-k} u_k,$$

where M_j is the jth Markov-parameter of the system. Note that in our case, M_0 is assumed to be zero, in order to avoid excessive notation. Recall that if the linear input-output map has a state-space representation (A, B, C), then the jth Markov-parameters satisfies $M_j = CA^{j-1}B$ for all $j \geq 1$.

In the same way, if f is realizable by a DTLSS, then the Markov-parameters of f can be expressed as products of the matrices of its DTLSS realization. In order to formulate this result more precisely, we need the following notation.

Notation 3: Consider the collection of $n \times n$ matrices A_σ, $\sigma \in Q$. For any $w \in Q^*$, the $n \times n$ matrix A_w is defined as follows. If $w = \epsilon$, then A_ϵ is the identity matrix. If $w = \sigma_1 \sigma_2 \cdots \sigma_k \in Q^*$, $\sigma_1, \cdots, \sigma_k \in Q$, $k > 0$, then

$$A_w = A_{\sigma_k} A_{\sigma_{k-1}} \cdots A_{\sigma_1}. \tag{6}$$

Lemma 1: The map f is realized by the DTLSS Σ if and only if f has a **GCR** and for all $v \in Q^*$, $q, q_0 \in Q$,

$$S^f(q_0 v q) = C_q A_v B_{q_0}. \tag{7}$$

Next, we define the concept of a Hankel-matrix. Similarly to the linear case, the entries of the Hankel-matrix are formed by the Markov parameters, and the finiteness of the rank of the Hankel-matrix will be the necessary and sufficient criterion for the existence of a finite-dimensional state-space representation.

For the definition of the Hankel-matrix of f, we will use lexicographical ordering on the set of sequences Q^*.

Remark 1 (Lexicographic ordering): Recall that $Q = \{1, \ldots, D\}$. We define a lexicographic ordering $\prec$ on Q^* as follows. For any $v, s \in Q^*$, $v \prec s$ if either $|v| < |s|$ or $0 < |v| = |s|$, $v \neq s$ and for some $l \in \{1, \ldots, |s|\}$, $v_l < s_l$ with the usual ordering of integers and $v_i = s_i$ for $i = 1, \ldots, l-1$. Here v_i and s_i denote the ith letter of v and s respectively. Note that $\prec$ is a complete ordering and $Q^* = \{v_1, v_2, \ldots\}$ with $v_1 \prec v_2 \prec \ldots$. Note that $v_1 = \epsilon$ and for all $i \in \mathbb{N}$, $q \in Q$, $v_i \prec v_i q$.

Definition 9 (Hankel-matrix): Consider the lexicographic ordering $\prec$ of Q^* from Remark 1. Define the Hankel-matrix H_f of f as the following infinite matrix

$$H_f = \begin{bmatrix} M^f(v_1 v_1) & M^f(v_2 v_1) & \cdots & M^f(v_k v_1) & \cdots \\ M^f(v_1 v_2) & M^f(v_2 v_2) & \cdots & M^f(v_k v_2) & \cdots \\ M^f(v_1 v_3) & M^f(v_2 v_3) & \cdots & M^f(v_k v_3) & \cdots \\ \vdots & \vdots & \cdots & \vdots & \cdots \end{bmatrix},$$

i.e. the $pD \times mD$ block of H_f in the block row i and block column j equals the Markov-parameter $M^f(v_j v_i)$ of f. The rank of H_f, denoted by $\mathrm{rank}\ H_f$, is the dimension of the linear span of its columns.

Recall from the definition of the Hankel-matrix for linear systems that it is of the form $H = (M_{k+l-1})_{k,l=1}^{\infty}$, where M_j is the jth Markov-parameter of the system. Hence, for $D = 1$, the definition above coincides with the linear case. The main result on realization theory of DTLSSs can be stated as follows.

Theorem 1 ([19]): 1) The map f has a realization by a DTLSS if and only if f has a **GCR** and $\mathrm{rank}\ H_f < +\infty$.

2) A minimal DTLSS realization of f can be constructed from H_f and any minimal DTLSS realization of f has dimension $\mathrm{rank}\ H_f$.

3) A DTLSS Σ is a minimal realization of f if and only if Σ is span-reachable, observable and it is a realization of f. Any two DTLSSs which are minimal realizations of f are isomorphic [2].

Note that Theorem 1 shows that the knowledge of the Markov-parameters is necessary and sufficient for finding a state-space representation of f. In fact, similarly to the continuous-time case [20], we can even show that the knowledge of *finitely many* Markov-parameters is sufficient. This will be done by formulating a *realization algorithm* for DTLSSs, which computes a DTLSS realization of f based on finitely many Markov-parameters of f.

In order to present the realization algorithm, we need the following notation.

Notation 4: Consider the lexicographic ordering $\prec$ of Q^* and recall that $Q^* = \{v_1, v_2, \ldots,\}$ where $v_1 \prec v_2 \cdots$. Denote by $\mathbf{N}(L)$ the number of sequences from Q^* of length at most L. It then follows that $|v_i| \le L$ if and only if $i \le \mathbf{N}(L)$.

Definition 10 ($H_{f,L,M}$ sub-matrices of H_f): For $L, K \in \mathbb{N}$ define the integers $I_L = \mathbf{N}(L)pD$ and $J_K = \mathbf{N}(K)mD$ Denote by $H_{f,L,K}$ the following upper-left $I_L \times J_K$ sub-matrix of H_f,

$$\begin{bmatrix} M^f(v_1v_1) & M^f(v_2v_1) & \cdots & M^f(v_{\mathbf{N}(K)}v_1) \\ M^f(v_1v_2) & M^f(v_2v_2) & \cdots & M^f(v_{\mathbf{N}(K)}v_2) \\ \vdots & \vdots & \cdots & \vdots \\ M^f(v_1v_{\mathbf{N}(L)}) & M^f(v_2v_{\mathbf{N}(L)}) & \cdots & M^f(v_{\mathbf{N}(K)}v_{\mathbf{N}(L)}) \end{bmatrix}.$$

Notice that the entries of $H_{f,L,K}$ are Markov-parameters indexed by words of length at most $L + K$, i.e. $H_{f,L,K}$ is uniquely determined by $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(L+K)}$.

The promised realization algorithm is Algorithm 1, which takes as input the matrix $H_{f,N,N+1}$ and produces a DTLSS. Note that the knowledge of $H_{f,N,N+1}$ is equivalent to the knowledge of the finite sequence $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2N+1)}$ of Markov-parameters. The Algorithm 1 is an extension of the well-known Kalman-Ho algorithm. Its correctness is stated below.

[2]see [19] for the definition of isomorphism between DTLSSs

Theorem 2: If $\mathrm{rank}\ H_{f,N,N} = \mathrm{rank}\ H_f$, then Algorithm 1 returns a minimal realization Σ_N of f. The condition $\mathrm{rank}\ H_{f,N,N} = \mathrm{rank}\ H_f$ holds for a given N, if there exists a DTLSS realization Σ of f such that $\dim \Sigma \le N + 1$.

The proof of Theorem 2 is completely analogous to its continuous-time counterpart [20]. *Theorem 2 implies that if f is realizable by a DTLSS, then a minimal DTLSS realization of f is computable from finitely many Markov-parameters*, using Algorithm 1. In fact, if f is realizable by a DTLSS of dimension n, then the *first* $\mathbf{N}(2n-1)$ *Markov-parameters* $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2n-1)}$ *uniquely determine* f.

Algorithm 1

Inputs: Hankel-matrix $H_{f,N,N+1}$.

Output: DTLSS Σ_N

1: Let $n = \mathrm{rank}\ H_{f,N,N+1}$. Choose a tuple of integers $(i_1, \ldots, i_n)$ such that the columns of $H_{f,N,N+1}$ indexed by $i_1, \ldots, i_n$ form a basis of $\mathrm{Im} H_{f,N,N+1}$. Let $\mathbf{O}$ be the $I_N \times n$ matrix formed by these linearly independent columns, i.e. the rth column of $\mathbf{O}$ equals the i_rth column of $H_{f,N,N+1}$. Let $\mathbf{R} \in \mathbb{R}^{n \times J_{N+1}}$ be the matrix, rth column of which is formed by coordinates of the rth column of $H_{f,N,N+1}$ with respect to the basis consisting of the columns $i_1, \ldots, i_n$ of $H_{f,N,N+1}$, for every $r = 1, \ldots, J_{N+1}$. It then follows that $H_{f,N,N+1} = \mathbf{OR}$ and $\mathrm{rank}\ \mathbf{R} = \mathrm{rank}\ \mathbf{O} = n$.

2: Define $\bar{\mathbf{R}} \in \mathbb{R}^{n \times J_N}$ as the matrix formed by the first J_N columns of $\mathbf{R}$.

3: For each $q \in Q$, let $\mathbf{R}_q \in \mathbb{R}^{n \times J_N}$ be such that for each $i = 1, \ldots J_N$, the ith column of $\mathbf{R}_q$ equals $r(i)$th column of $\mathbf{R}$. Here $r(i) \in \{1, \ldots, J_{N+1}\}$ is defined as follows. Consider the decomposition $i = (r-1)mD + z$ for some $z = 1, \ldots, mD$ and $r = 1, \ldots, \mathbf{N}(N)$. Consider the word $v_r q$ and notice that $|v_r q| \le N + 1$. Hence, $v_r q = v_d$ for some $d = 1, \ldots, \mathbf{N}(N+1)$. Then define $r(i)$ as $r(i) = (d-1)mD + z$.

4: Construct Σ_N of the form (1) such that

$$\begin{aligned} &\begin{bmatrix} B_1, \ldots, B_D \end{bmatrix} = \\ &\text{the first } mD \text{ columns of } \mathbf{R} \\ &\begin{bmatrix} C_1^T & C_2^T & \ldots & C_D^T \end{bmatrix}^T = \text{ the first } pD \text{ rows of } \mathbf{O} \\ &\forall q \in Q : A_q = \mathbf{R}_q \bar{\mathbf{R}}^+ \end{aligned}$$

where $\bar{\mathbf{R}}^+$ is the Moore-Penrose pseudoinverse of $\bar{\mathbf{R}}$.

5: Return Σ_N

The intuition behind Algorithm 1 is the following. The state-space of the DTLSS Σ_N returned by Algorithm 1 is an isomorphic copy of the space spanned by the columns of $H_{f,N,N}$. The isomorphism is determined by the matrix $\mathbf{R}$. The columns of B_q, $q \in Q$ are formed by first Dm columns of $H_{f,N,N}$. The rows of the matrices C_q, $q \in Q$ are formed by the first pD rows of $H_{f,N,N}$. Finally, the matrix A_q, $q \in Q$ is the matrix of the shift-like operator, which maps a block-column $\{M^f(v_j v_i)\}_{i=1}^{\mathbf{N}(L)}$ of $H_{f,N,N}$ to the block-column $\{M^f(v_j q v_i)\}_{i=1}^{\mathbf{N}(L)}$ of $H_{f,N,N+1}$.

IV. MAIN RESULTS OF THE PAPER

The main idea behind our definition of persistence of excitation is as follows. The measured time series is persistently exciting, if from this time-series we can reconstruct the Markov-parameters of the underlying system. Note that by Theorem 2, it is enough to reconstruct *finitely many* Markov-parameters. This means that our definition of persistence of excitation is also applicable to finite time series.

In order to present our main results, we will need some terminology.

Definition 11 (Output time-series): For any input-output map f and for any finite input sequence $w \in \mathcal{U}^+$ we denote by $\mathbf{O}(f, w)$ the output time series induced by f and w, i.e. if w is of the form (2), then $\mathbf{O}(f, w) = \{y_i\}_{i=0}^{t}$, such that $y_i = f((q_0, u_0) \cdots (q_i, u_i))$ for all $i = 0, 1, \ldots, t$.

Definition 12 (Persistence of excitation): The finite sequence $w \in \mathcal{U}^+$ is *persistently exciting for the input-output map* f, if it is possible to compute the Markov-parameters of f from the data $(w, \mathbf{O}(f, w))$.

Remark 2 (Interpretation): The input w is persistently exciting, if and only if there exists an algorithm which computes a DTLSS realization of f from the time series $(w, \boldsymbol{O}(f, w))$.

Indeed, assume that w is persistently exciting and that f admits a DTLSS realization of dimension at most n. Then the Markov-parameters $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2n-1)}$ can be computed from the data $(w, \mathbf{O}(f, w))$. By Theorem 2, one can use Algorithm 1 to compute a DTLSS realization of f from $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2n-1)}$.

Conversely, assume that there exists an algorithm which can correctly find a DTLSS realization of f from $(w, \mathbf{O}(f, w))$. Then by (7) all the Markov-parameters of f can be computed from that DTLSS realization. Hence, according to our definition, w is persistently exciting.

Next, we define persistence of excitation for infinite sequence of inputs. To this end, we need the following notation.

Notation 5: We denote by $\mathcal{U}^\omega$ the set of infinite sequences of hybrid inputs. That is, any element $w \in \mathcal{U}^\omega$ can be interpreted as a time-series $w = \{(q_t, u_t)\}_{t=0}^{\infty}$, $q_t \in Q$, $u_t \in \mathbb{R}^m$, $t \geq 0$. For each $N \in \mathbb{N}$, denote by w_N the sequence formed by the first N elements of w, i.e. $w_N = (q_0, u_0) \cdots (q_N, u_N)$.

Definition 13 (Asymptotic persistence of excitation): An infinite sequence of inputs $w \in \mathcal{U}^\omega$ is called *asymptotically persistently exciting for the input-output map* f, if the following holds. For every sufficiently large N, we can compute from $(w_N, \mathbf{O}(f, w_N))$ asymptotic estimates of the Markov-parameters of f. More precisely, for every $N \in \mathbb{N}$, we can compute from $(w_N, \mathbf{O}(f, w_N))$ some matrices $\{M_N^f(v)\}_{v \in Q^*}$ such that $\lim_{N \to \infty} M_N^f(v) = M^f(v)$ for all $v \in Q^*$. When clear from the context, we will use the term persistently exciting instead of asymptotically persistently exciting.

Remark 3 (Interpretation): The interpretation of asymptotic persistence of excitation is as follows. The infinite input sequence w *is asymptotically persistently exciting for f, if and only if there exists an asymptotically consistent algorithm for computing a DTLSS realization of f from its response to w.* By asymptotically consistent algorithm we mean the following. For each integer $N > 0$, the algorithm computes a DTLSS Σ_N from the finite data $(w_N, \mathbf{O}(f, w_N))$, and the DTLSSs Σ_N converge to a true DTLSS realization Σ of f, as $N \to \infty$. By convergence of Σ_N to Σ we mean convergence of the system matrices, i.e. the state-space dimension of Σ and Σ_N are the same, and if A_q^N, B_q^N, C_q^N, $q \in Q$ are the matrices of Σ_N and Σ is of the form (1), then $A_q = \lim_{N \to \infty} A_q^N$, $B_q = \lim_{N \to \infty} B_q^N$, $C_q = \lim_{N \to \infty} C_q^N$.

Indeed, assume that $w \in \mathcal{U}^\omega$ is asymptotically persistently exciting. Then for each N we can compute from the time-series $(w_N, \mathbf{O}(f, w_N))$ an approximation $\{M_N^f(v)\}_{v \in Q^*}$ of the Markov-parameters of f. Suppose that f is realizable by a DTLSS of dimension n and we know the indices $(i_1, \ldots, i_n)$ of those columns of $H_{f,n-1,n}$ which form a basis of the column space of $H_{f,n-1,n}$. Let $H_{f,n-1,n}^N$ be the matrix which is constructed in the same way as $H_{f,n-1,n}$, but with $M_N^f(v)$ instead of the Markov-parameters $M^f(v)$. Then each entry of $H_{f,n-1,n}^N$ converges to the corresponding entry of $H_{f,n-1,n}$. Modify Algorithm 1 by fixing the choice of columns to $(i_1, \ldots, i_n)$ in the first step. It is easy to see that the modified algorithm represents a continuous map from the input data (finite Hankel-matrix) to the output data (matrices of a DTLSS). For sufficiently large N, the columns of $H_{f,n-1,n}^N$ indexed by $(i_1, \ldots, i_n)$ also represent a basis of the column space of $H_{f,n-1,n}^N$. If we apply the modified Algorithm 1 to the sequence of matrices $H_{f,n-1,n}^N$, we obtain a sequence of DTLSSs Σ_N. Then the parameters of Σ_N converge to the parameters of the DTLSS Σ as $N \to \infty$, where Σ is the DTLSS which we obtain by applying Algorithm 1 to $H_{f,n-1,n}$. By Theorem 2, Σ is a realization of f.

Conversely, assume that there is an algorithm which from the sequence $(w_N, \mathbf{O}(w_N, f))$ computes a sequence of DTLSSs Σ_N such that Σ_N converge to a true DTLSS realization Σ of f, as $N \to \infty$. Then for each N, compute the Markov-parameter $M_N^f(v)$, $v \in Q^*$ of the input-output map of Σ_N, using (7). Notice that $M_N^f(v)$ depends continuously on the parameters of Σ_N, and the parameters of Σ_N converge to those of Σ. Hence, using the fact that Σ is a realization of f and (7), we obtain $\lim_{N \to \infty} M_N(v) = M(v)$. Hence, in this case w is asymptotically persistently exciting.

The discussion above implies that *if w is a finite or infinite input signal and w is persistently exciting for f, then Algorithm 1 can serve as an identification algorithm for computing a DTLSS realization of f.*

We will show that for every reversible DTLSS there exists **some** input which is persistently exciting. In addition, we present a class of inputs which are persistently exciting for any input-output map f realizable by a stable DTLSS.

A. Persistently exciting input for specific systems

Note that from (4) it follows that the Markov-parameters of f can be obtained from finitely many input-output data. However, the application of (4) implies evaluating the response of the system to different inputs, while started from a fixed initial state. In order to simulate this by evaluating the response of the system to one single input (which is then necessarily persistently exciting), one has to provide means to reset the system to its initial state. In order to be able to do so, we restrict attention to *reversible* DTLSSs.

Definition 14: A DTLSS Σ of the form (1) is *reversible*, if for every discrete mode $q \in Q$, the matrix A_q is invertible. Reversible DTLSSs arise naturally when sampling continuous-time linear switched systems.

Theorem 3: Consider an input-output map f. Assume that f has a realization by a reversible DTLSS. Then there exists an input $w \in \mathcal{U}^+$ such that w is persistently exciting for f.

Proof: [Sketch of the proof] The main idea behind the proof of Theorem 3 is as follows. If f admits a DTLSS realization of dimension n, then the finite sequence $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2n-1)}$ of Markov-parameters determine all the Markov-parameters of f uniquely. Hence, in order for a finite input w to be persistently exciting for f, it is sufficient that $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2n-1)}$ can be computed from the response $(w, \mathbf{O}(f,w))$.

Note that (4) implies that $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2n-1)}$ can be computed from $\{f(s) \mid s \in S\}$, where

$$\begin{aligned} &S = \{(q_0, e_j)(\sigma_1, 0)\dots(\sigma_{|v_i|}, 0)(q,0) \in \mathcal{U}^+ \mid q_0, q \in Q, \\ &v_i = \sigma_1 \cdots \sigma_{|v_i|}, \sigma_1, \dots, \sigma_{|v_i|} \in Q, \\ &j = 1, \dots, m, i = 1, \dots, \mathbf{N}(2n-1)\} \end{aligned}$$

Hence, if for each $s \in S$ there exists a prefix p of w such that $f(s) = f(p)$, then this w will be persistently exciting.

One way to construct such a w is to find for each $s \in S$ an input $s^{-1} \in \mathcal{U}^+$ such that

$$\forall v \in \mathcal{U}^+ : f(ss^{-1}v) = f(v).$$

That is, the input s^{-1} neutralizes the effect of the input s. We defer the construction of the input s^{-1} to the end of the proof. Assume for the moment being that such inputs s^{-1} exist. Let $S = \{s_1, \dots, s_d\}$ be an enumeration of S. Then it is easy to see that $f(s_1 s_1^{-1} s_2) = f(s_2)$, $f(s_1 s_1^{-1} s_2 s_2^{-1} s_3) = f(s_3)$, etc. Hence, if we define

$$w = s_1 s_1^{-1} \cdots s_{d-1} s_{d-1}^{-1} s_d,$$

then each $f(s)$, $s \in S$ can be obtained as a response of f to a suitable prefix of w. Hence, w is persistently exciting.

It is left to show that s^{-1} exists. Consider a reversible realization Σ of f. Then the controllable set and reachable set of Σ coincide by [7]. Hence, from any reachable state x of Σ, there exists an input $w(x)$ such that $w(x)$ drives Σ from x to zero, i.e. $x_\Sigma(x, w(x)) = 0$. For each $s \in S$, let $x(s) = x_\Sigma(0, s)$ and define $s^{-1} = w(x(s))$ as the input which drives $x(s)$ back to the initial zero state. ∎

The construction of the persistently exciting w from Theorem 3 requires the knowledge of a DTLSS realization of f. Below we present a subclass of input-output maps, for which the knowledge of a state-space representation is not required to construct a persistently exciting input.

Definition 15: Fix a map $.^{-1} : \mathcal{U} \ni \alpha \mapsto \alpha^{-1} \in \mathcal{U}$. A input-output map f is said to be *resettable with respect to the map* $.^{-1}$, if for all $\alpha \in \mathcal{U}$, $s, w \in \mathcal{U}^*$, $|sw| > 0$,

$$f(s\alpha\alpha^{-1}w) = f(sw).$$

Intuitively, f is resettable with respect to $.^{-1}$, if the effect of any input $\alpha = (q, u)$ on the initial state can be neutralized by the input α^{-1}.

Theorem 4: If f is resettable with respect to $.^{-1}$, then a persistently exciting input sequence w can be constructed for f without the knowledge of a DTLSS realization of f.

Proof: [Proof of Theorem 4] The proof differs from that of Theorem 3 only in the definition of s^{-1} for each $s \in S$. More precisely, for each $s = (q_0, u_0)\cdots(q_t, u_t) \in S$ define

$$s^{-1} = (q_t, u_t)^{-1}(q_{t-1}, u_{t-1})^{-1}\cdots(q_0, u_0)^{-1}$$

∎

B. Universal persistently exciting inputs

Next, we discuss classes of inputs which are persistently exciting for all input-output maps realizable by stable DTLSSs.

Definition 16 (Persistence of excitation condition): An infinite input $w = \{(q_t, u_t)\}_{t=0}^{\infty} \in \mathcal{U}^\omega$ satisfies PE condition, if for any word $v \in Q^+$ the limits below exist and satisfy the following conditions,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N} u_{t+j} u_t^T \chi(q_t q_{t+1} \cdots q_{t+|v|-1} = v) = 0,$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=j}^{N} u_{t-j} u_t^T \chi(q_{t-j} q_{t-j+1} \cdots q_{t-j+|v|-1} = v) = 0,$$

$$\mathcal{R} \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N} u_t u_t^T > 0,$$

$$\pi_v \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N} \chi(q_t q_{t+1} \cdots q_{t+|v|-1} = v) > 0,$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N} u_t u_t^T \chi(q_t q_{t+1} \cdots q_{t+|v|-1} = v) = \pi_v \mathcal{R}.$$

where χ is the indicator function, i.e. $\chi(A) = 1$ if A holds and $\chi(A) = 0$ otherwise. Note that by $\mathcal{R} > 0$ we mean that $\mathcal{R}$ is a strictly positive definite $m \times m$ matrix.

Remark 4: Note that if $w \in \mathcal{U}^\omega$ satisfies the conditions of Definition 16, then the switching signal is *rich enough*, i.e. any sequence of discrete modes occurs in the switching signal infinitely often. Relaxations of the requirement will be discussed at the end of this section.

Remark 5 (Relationship with stochastic processes): Fix a probability space $(\Omega, \mathcal{F}, P)$ and consider discrete-time stochastic processes $\mathbf{u}_t : \Omega \to \mathbb{R}^m$ and $\mathbf{q}_t : \Omega \to Q$ with values in $\mathbb{R}^m$ and Q respectively, such that the following conditions hold.

- The process $(\mathbf{u}_t, \mathbf{q}_t)$ is stationary and ergodic.
- The processes $\mathbf{u}_t$ and $\mathbf{q}_t$ are independent (i.e. the σ-algebras generated by $\{\mathbf{u}_t\}_{t=0}^{\infty}$ and by $\{\mathbf{q}_t\}_{t=0}^{\infty}$ are independent).

- The stochastic process $\mathbf{u}_t$ is zero-mean, $\mathbf{u}_t$ and $\mathbf{u}_s$ are uncorrelated for $s \neq t$, and $E[\mathbf{u}_t\mathbf{u}_t^T] = \mathcal{R} > 0$, with $E[\cdot]$ denoting the expectation operator.
- For each $v \in Q^+$, $\pi_v = P(\mathbf{q}_t \cdots \mathbf{q}_{t+|v|-1} = v) > 0$.

It then follows that almost all sample paths of $(\mathbf{u}_t, \mathbf{q}_t)$ satisfy the PE condition of Definition 16. That is, there exists a set $A \in \mathcal{F}$, such that $P(A) = 0$ and for all $\omega \in \Omega \setminus A$, the sequence $w = \{(q_t, u_t) = (\mathbf{q}_t(\omega), \mathbf{u}_t(\omega))\}_{t=0}^{\infty}$ satisfies the PE condition.

Remark 6: If $\mathbf{u}_t$ is a white-noise Gaussian process and if the variables $\mathbf{q}_t$ are uniformly distributed over Q (i.e. $P(\mathbf{q}_t = q) = \frac{1}{|Q|}$ for all $q \in Q$) and are independent from each other and from $\{\mathbf{u}_s\}_{s=0}^{\infty}$, then $\mathbf{u}_t$ and $\mathbf{q}_t$ satisfy the conditions of Remark 5 and hence almost any sample path of $\mathbf{u}_t$ and $\mathbf{q}_t$ satisfies the PE condition of Definition 16.

We will show that input sequences which satisfy the conditions of Definition 16 are asymptotically persistently exciting for a large class of input-output maps. The main idea behind the theorem is as follows. Consider a DTLSS Σ which is realization of f, and suppose we feed a stochastic input $\{\mathbf{q}_t, \mathbf{u}_t\}$ into Σ. Then the state $\mathbf{x}_t$ and the output response $\mathbf{y}_t$ of Σ will also be stochastic processes. Suppose that $\{\mathbf{q}_t, \mathbf{u}_t\}$ are stochastic processes which satisfy the conditions of Remark 5. It is easy to see that

$$\mathbf{y}_t = \sum_{k=0}^{t-1} C_{\mathbf{q}_t} A_{\mathbf{q}_{k+1}\cdots\mathbf{q}_{t-1}} B_{\mathbf{q}_k}\mathbf{u}_k$$

where $\mathbf{q}_{k+1}\cdots\mathbf{q}_{t-1}$ is interpreted as the empty sequence for $k+1 > t-1$. Hence for all $r, q \in Q$, $v \in Q^*$, $|rvq| = t+1$, with $v = \sigma_1 \cdots \sigma_{t-1}$, $\sigma_1, \ldots, \sigma_{t-1} \in Q$, and $\sigma_0 = r$,

$$\begin{aligned} &E[\mathbf{y}_t\mathbf{u}_0^T \chi(\mathbf{q}_0 \cdots \mathbf{q}_t = rvq)] = \\ &\sum_{k=0}^{t-1} C_q A_{\sigma_{k+1}\cdots\sigma_{t-1}} B_{\sigma_k} E[\mathbf{u}_k\mathbf{u}_0^T \chi(\mathbf{q}_0 \cdots \mathbf{q}_t = rvq)] = \\ &C_q A_v B_r \mathcal{R}\pi_{rvq} = S^f(rvq)\mathcal{R}\pi_{rvq}. \end{aligned} \tag{8}$$

Hence, if we know the expectations $E[\mathbf{y}_t\mathbf{u}_0^T\chi(\mathbf{q}_0 \cdots \mathbf{q}_t = rvq)]$ for all $r, q \in Q$, $v \in Q^*$, $|rvq| = t+1$, $t > 0$, then we can find all the Markov-parameters of f by the following formula

$$S^f(rvq) = E[\mathbf{y}_t\mathbf{u}_0^T\chi(\mathbf{q}_0 \cdots \mathbf{q}_{t+1} = rvq)]\mathcal{R}^{-1}\frac{1}{\pi_{rvq}} \tag{9}$$

Hence, the problem of estimating the Markov-parameters reduces to estimating the expectations

$$E[\mathbf{y}_t\mathbf{u}_0^T\chi(\mathbf{q}_0 \cdots \mathbf{q}_t = rvq)] \tag{10}$$

For practical purposes, the expectations in (10) have to be estimated from the time-series, and hence we obtain the following formula for estimating the Markov-parameters,

$$S^f(rvq) = \lim_{N\to\infty}\frac{1}{N}\sum_{i=0}^{N} y_{i+t}u_i^T\chi(q_i \cdots q_{i+t} = rvq) \tag{11}$$

where y_t, u_t, q_t denote the value at time t of a sample-path of $\mathbf{y}_t$, $\mathbf{u}_t$ and $\mathbf{q}_t$ respectively. Note that y_t is in fact the output of Σ at time t, if the input $\{u_i\}_{i=0}^{t}$ and the switching signal $\{q_i\}_{i=0}^{t}$ are fed to the system.

The problem with (10) is that the limit on the right-hand side of (11) may fail to converge to (10). A particular case when the limit in (11) converges to (10) is when the process $(\mathbf{y}_t, \mathbf{u}_t, \mathbf{q}_t)$ is ergodic. In that case, we can choose a sample-path (y_t, u_t, q_t) of $(\mathbf{y}_t, \mathbf{u}_t, \mathbf{q}_t)$ for which the limit in (11) equals the expectation (10), in fact 'almost all' sample paths will have this property. This means that we can choose a suitable deterministic input sequence $\{u_t\}_{t=0}^{\infty}$ and a switching signal $\{q_t\}_{t=0}^{\infty}$, such that for the resulting output $\{y_t\}_{t=0}^{\infty}$, the limit on the right-hand side of (11) equals the expectation (10). That is, in that case the input $w = (q_0, u_0) \cdots (q_t, u_t) \cdots$ is asymptotically persistently exciting. However, proving ergodicity of $\mathbf{y}_t$ is not easy. In addition, even if $\mathbf{y}_t$ is ergodic, the particular choice of the deterministic input w for which the right-hand side of (11) equals (10) might depend on the DTLSS itself.

For this reason, instead of using the concepts of ergodicity directly, we just show that for the input sequences w which satisfy the conditions of Definition 16, the corresponding output $\{y_t\}_{t=0}^{\infty}$ has the property that the limit (11) exists and it equals $S^f(rvq)\mathcal{R}\pi_{rvq}$, for any input-output map f which is realizable by a *l_1-stable DTLSS*. This strategy allows us to use elementary techniques, while not compromising the practical relevance of the result.

In order to present the main result of this section, we have to define the notion of l_1-stability of DTLSSs.

Definition 17 (Stability of DTLSSs): A DTLSS Σ of the form (1) is called l_1-stable, if for every $x \in \mathbb{R}^n$, the series $\sum_{v\in Q^*} ||A_v x||_2$ is convergent.

Remark 7 (Sufficient condition for stability): If for all $q \in Q$, $||A_q||_2 < \frac{1}{|Q|}$, where $||A_q||_2$ is the matrix norm of A_q induced by the standard Euclidean norm, then Σ is l_1-stable.

Remark 8 (Asymptotic stability): If Σ is l_1-stable, then it is asymptotically stable, in the sense that if $s_i \in Q^*$, $i > 0$ is a sequence of words such that $\lim_{i\to\infty}|s_i| = +\infty$, then $\lim_{i\to\infty} A_{s_i}x = 0$ for all $x \in \mathbb{R}^n$.

Intuitively it is clear why we have to restrict attention to stable systems. Recall that (4) allows us to compute the Markov-parameters of f from the responses of f to finitely many inputs. In order to obtain the response of f to several inputs from the response of f to one input, one has to find means to suppress the contribution of the current state of the system to future outputs. By assuming stability, we can make sure that the effect of the past state will asymptotically diminish in time. Hence, by waiting long enough, we can approximately recover the response of f to any input.

Theorem 5 (Main result): Assume that w satisfies the PE conditions of Definition 16. Assume that f is an arbitrary input-output map and assume that f admits an l_1-stable DTLSS realization. Then w is asymptotically persistently exciting for f.

The proof of Theorem 5 relies on the following technical result.

Theorem 6: Assume that Σ is a l_1-stable DTLSS of the form (1), and assume that w satisfies the PE conditions. Let $\{y_t\}_{t=0}^{\infty}$ and $\{x_t\}_{t=0}^{\infty}$ be the output and state responses of Σ to w, i.e. $y_t = y_{\Sigma}(w_t)$ and $x_t = x_{\Sigma}(0, w_t)$. Then for all $v, \beta \in Q^*$, $r, q \in Q$

$$\pi_{rvq\beta} A_v B_r \mathcal{R} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N} x_{t+|v|+1} u_t^T \chi(t, rvq\beta) \quad (12)$$

$$\pi_{rvq\beta} C_q A_v B_r \mathcal{R} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N} y_{t+|v|+1} u_t^T \chi(t, rvq\beta) \quad (13)$$

Here we used the following notation: for all $s \in Q^+$,

$$\chi(t,s) = \begin{cases} 1 & \text{if } s = q_t q_{t+1} \cdots q_{t+|s|-1} \\ 0 & \text{otherwise} \end{cases}$$

Informally, Theorem 6 implies that if f is realizable by a l_1-stable DTLSS, then the limit (11) equals (10). The proof of Theorem 6 can be found in [18].

Proof: [Proof of Theorem 5] For each t, denote by y_t the response of f to the first t elements of w, i.e. $y_t = f((q_0, u_0) \cdots (q_t, u_t))$. For each integer $N \in \mathbb{N}$ and for each word $v \in Q^*$, define the matrix $S_N(rvq)$ as

$$S_N(rvq) = (\frac{1}{N} \sum_{t=0}^{N} y_{t+|v|+1} u_t^T \chi(t, rvq)) \mathcal{R}^{-1} \frac{1}{\pi_{rvq}}$$

and define the matrix $M_N(v)$ by

$$\begin{bmatrix} S_N(1v1) & \cdots & S_N(Dv1) \\ \vdots & \vdots & \vdots \\ S_N(1vD) & \cdots & S_N(DvD) \end{bmatrix}$$

From Theorem 6 it follows that $\lim_{N\to\infty} S_N(rvq) = S^f(rvq)$ and hence $\lim_{N\to\infty} M_N(v) = M^f(v)$. Hence, w is indeed asymptotically persistently exciting. ■

Remark 9 (Relaxation of PE condition): Assume that we restrict attention to input-output maps which are realizable by a l_1-stable DTLSS of dimension at most n, and let f be such an input-output map. In this case, one can replace the assumption of Definition 16 that $\pi_v > 0$ for all $v \in Q^+$ by the condition that $\pi_s > 0$ for all $|s| \le 2n-1$ and still obtain asymptotically persistently exciting inputs for f.

Indeed, in this case Theorem 6 remains valid (the proof remains literally the same) and from the proof of Theorem 5 we get that for all $v \in Q^*, |v| \le 2n-1$, (11) holds. Hence, $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2n-1)}$ can asymptotically be estimated from $(w_N, \mathbf{O}(f, w_N))$. Since the modified Algorithm 1 from Remark 3 determines a continuous map from $\{M^f(v_i)\}_{i=1}^{\mathbf{N}(2n-1)}$ to the other Markov-parameters of f, w is asymptotically persistently exciting for f.

V. Conclusions

We defined persistence of excitation for input signals of linear switched systems. We showed existence of persistently exciting input sequences and we identified several classes of input signals which are persistently exciting.

Future work includes finding less restrictive conditions for persistence of excitation and extending the obtained results to other classes of hybrid systems.

References

[1] L. Bako. Identification of switched linear systems via sparse optimization. *Automatica*, 47:668–677, 2011.

[2] L. Bako, G. Mercère, and S. Lecoeuche. Online structured subspace identification with application to switched linear systems. *International Journal of Control*, 82:1496–1515, 2009.

[3] L. Bako, G. Mercère, R. Vidal, and S. Lecoeuche. Identification of switched linear state space models without minimum dwell time. In *IFAC Symposium on System Identification, Saint Malo, France*, 2009.

[4] Samuel Eilenberg. *Automata, Languages and Machines*. Academic Press, New York, London, 1974.

[5] G. Ferrari-Trecate, M. Musellu, D. Liberati, and M. Morari. A clustering technique for the identification of piecewise-affine systems. *Automatica*, 39:205–217, 2003.

[6] E.B. Fox. *Bayesian Nonparametric Learning of Complex Dynamical Phenomena*. Ph.D. thesis, MIT, Cambridge, MA, 2009.

[7] S.S. Ge, Zhendong Sun, and T.H. Lee. Reachability and controllability of switched linear discrete-time systems. *IEEE Trans. Automatic Control*, 46(9):1437 – 1441, 2001.

[8] Y. Hashambhoy and R. Vidal. Recursive identification of switched ARX models with unknown number of models and unknown orders. In *IEEE Conference on Decision and Control*, 2005.

[9] A. Hiskens. Identifiability of hybrid system models. In *Proceedings of the IEEE International Conference on Control Applications, Anchorage, AK*, 2000.

[10] A. Juloski. *Observer Design and Identification Methods for Hybrid Systems: Theory and Experiments*. PhD thesis, Eindhoven Universtity of Technology, 2004.

[11] A. Juloski, W.P.M.H. Heemels, G. Ferrari-Trecate, R. Vidal, S. Paoletto, and J. Niessen. Comparison of four procedures for the identification of hybrid systems. In *Hybrid Systems: Computation and Control*, volume 3414 of *LNCS*, pages 354–369. Springer-Verlag, Berlin, 2005.

[12] A. Juloski, S. Weiland, and M. Heemels. A bayesian approach to identification of hybrid systems. In *Proceedings of 43rd IEEE Conference on Decision and Control*, 2004.

[13] L. Ljung. *System Identification: Theory for the user (2nd Ed.)*. PTR Prentice Hall., Upper Saddle River, USA, 1999.

[14] Y. Ma and R. Vidal. A closed form solution to the identification of hybrid ARX models via the identification of algebraic varieties. In *Hybrid Systems: Computation and Control*, pages 449–465, 2005.

[15] S. Paoletti, A. Juloski, G. Ferrari-Trecate, and R. Vidal. Identification of hybrid systems: A tutorial. *European Journal of Control*, 13(2-3):242 – 260, 2007.

[16] S. Paoletti, J. Roll, A. Garulli, and A. Vicino. Input/ouput realization of piecewise affine state space models. In *46th IEEE Conf. on Dec. and Control*, 2007.

[17] M. Petreczky, L. Bako, and J.H. van Schuppen. Identifiability of discrete-time linear switched systems. In *Hybrid Systems: Computation and Control*, pages 141–150. ACM, 2010.

[18] Mihaly Petreczky and Laurent Bako. On the notion of persistence of excitation for linear sw itched systems. Technical Report arXiv:1103.1349v1, ArXiv, 2011.

[19] Mihaly Petreczky, Laurent Bako, and Jan H. van Schuppen. Realization theory for discrete-time linear switched systems. Technical Report arXiv:1103.1343v1, ArXiv, 2011.

[20] Mihaly Petreczky and Jan H. van Schuppen. Partial-realization of linear switched systems: A formal power series approach. *Automatica*, 47(10):2177–2184, 2011.

[21] J. Roll, A. Bemporad, and L. Ljung. Identification of piecewise affine systems via mixed-integer programming. *Automatica*, 40(1):37–50, 2004.

[22] Zhendong Sun and Shuzhi S. Ge. *Switched linear systems : control and design*. Springer, London, 2005.

[23] V. Verdult and M. Verhaegen. Subspace identification of piecewise linear systems. In *Proc.Conf. Decision and Control*, 2004.

[24] R. Vidal. Recursive identification of switched ARX systems. *Automatica*, 44(9):2274 – 2287, 2008.

[25] R. Vidal, A. Chiuso, and S. Sastry. Observability and identifiability of jump linear systems. In *Proc. IEEE Conf. Dec. and Control*, pages 3614 – 3619, 2002.

[26] Jan C. Willems, Paolo Rapisarda, Ivan Markovsky, and Bart L.M. De Moor. A note on persistency of excitation. *Systems & Control Letters*, 54(4):325 – 329, 2005.