Minimality and identifiability of discrete-time switched autoregressive exogenous systems

Petreczky Mihaly, Laurent Bako, S. Lecoeuche, K.M.D. Motchon

Received: 00 Month 0000 | Revised: 00 Month 0000 | Accepted: 00 Month 0000

DOI: xxx/xxxx

RESEARCH ARTICLE

Minimality and identifiability of discrete-time SARX systems

M. Petreczky*[1] | L. Bako[2] | S. Lecoeuche[3] | K.M.D. Motchon[4]

[1]CNRS, Centrale Lille, UMR 9189–CRIStAL–Centre de Recherche en Informatique, Signal et Automatique de Lille, F-59000 Lille, France

[2]Ecole Centrale Lyon, Laboratoire Ampère, 36 av. Guy de Collongue, 69130 Ecully, France

[3]IMT Lille Douai, Univ. Lille, Unité de Recherche Informatique Automatique, F-59000 Lille, France

[4]Université de Reims Champagne Ardenne, CReSTIC EA 3804, 51097 Reims, France

Correspondence
M. Petreczky, CNRS, Centrale Lille, UMR 9189–CRIStAL–Centre de Recherche en Informatique, Signal et Automatique de Lille
Email: mihaly.petreczky@centralelille.fr

Summary

The paper studies the problem of minimality and identifiability for Switched Auto-Regressive eXogenous (abbreviated by SARX) systems. We propose formal definitions of the concepts of identifiability and minimality for SARX models. Based on these formalizations, we derive conditions for minimality and identifiability of SARX systems. In particular, we show that polynomially parameterized SARX systems are generically identifiable.

KEYWORDS:
Switched ARX systems, minimality, identifiability, system identification, hybrid systems

1 | INTRODUCTION

System identification is the branch of control theory which is concerned with designing methods and algorithms for inferring parameterized mathematical models from input-output measurements. A fundamental criterion characterizing the quality of a model parameterization is that of identifiability. This refers to the formal question of whether a given parameterized model can, in principle, be uniquely determined from input-output data. More precisely, a parameterized model structure is a map from a certain parameter space to a set of dynamic systems. Such a parameterized model structure is said to be (structurally) identifiable, if no two different parameter vectors yield two models whose input-output behavior is the same. The concept of identifiability has a number of implications for the design of informative experiments, the development of parameter estimation algorithms, the analysis of identification methods and the significance of estimated models.

Contribution of the paper. The present paper deals with the problem of identifiability of switched ARX (*abbreviated as SARX*) systems. More precisely, we we introduce formal definitions identifiability for SARX systems and show that a particular notion of minimality, called *strong minimality*, is a sufficient condition for identifiability of SARX systems. We present conditions for checking strong minimality which are reminiscent of the well-know minimality conditions for ARX systems. We also show that minimality that SARX parameterizations are generically minimal and generically identifiable. Note that minimality and identifiability are properties of the structure of the model parameterization and not that of the data generated by the system. Based on our definitions, we derive checkable conditions guaranteeing these two properties. In addition to providing theoretical insights, the results of the paper allow us to check identifiability of SARX parameterizations, and to find identifiable parameterizations.

In this paper, we consider only single-input single-output (SISO) SARX systems. We conjecture that the extension of the result of this paper to the MIMO case is possible, but it is technically much more involved. This is consistent with the case of ARX systems, for which identifiability of the MIMO case is much more involved, see for example [1] and the references therein.

It is worth noting that identifiability and minimality SARX systems cannot be reduced to the corresponding properties of its ARX subsystems. It can be shown that a SARX system can be minimal (resp. identifiable), even if none of the ARX subsystems is minimal (resp. identifiable). That is, the relationship between identifiability and minimality of SARX systems and their ARX subsystems is not straightforward.

Motivation. SARX systems are popular in the hybrid systems community, due to their simplicity and modeling power. In particular, most of hybrid systems identification algorithms were developed for SARX systems [2,3,4,5,6,7]. Despite their popularity, identifiability and minimality of SARX systems are not yet completely understood.

Identifiability is an established topic in system identification, there is a consensus that identifiability is an essential condition for proving correctness of algorithms for system identification[8] and adaptive control[9]. Indeed, mathematical correctness of parametric system identification algorithms can be established only for identifiable parameterizations, since for non-identifiable parameterizations the corresponding system identification problem is ill-posed. In turn, minimality is often sufficient for achieving identifiability of input-output models.

The results of the paper are useful for the following purposes:

- They provide a strict mathematical proof that non-identifiable parameterizations are somewhat exceptional, in the sense that they are not generic. We believe that it is interesting in its own right.
- They provide means to come up with parameterizations which are guaranteed to be identifiable. In turn, for such parameterizations it might be possible to prove statistical consistency of hybrid system identification algorithms, since the underlying mathematical problem becomes well-posed.

Related work. Identification of hybrid systems is an active research topic [10,11,5,12,13,14,15,16,17,18,19,2,7]. Many of the major contributions are formulated only for SARX systems [4,20,21,22,5,3]. The relationship between SARX systems and state-space representations was addressed by Paoletti *et al*[3] and Weiland *et al*[23], and in this paper we use some of those results.

To the best of our knowledge, the results of the paper are new. The paper of Vidal[5] contains persistence of excitation conditions for SARX systems, which is related to identifiability. The main difference between the two concepts is that the former is a property of the data, while the latter is a property of the parameterization. Vidal[5] also proposes a definition of minimality of SARX systems which implies our definition. However, the two definitions are not equivalent.

The proofs of the main results of the paper are based on the results of[24] which analyzes minimality and identifiability of linear switched state-space representation. It should however be noted that the contributions of the paper are not trivial consequences of the work of[24]. This is due to the rich structure of SARX systems which allows us to derive stronger results than for general linear switched systems. For example, the relationship between minimality and identifiability is much more direct for SARX systems than for linear switched systems.

Finally, we note that a preliminary version of the material presented in the current paper appeared in the proceedings of the 16th IFAC Symposium on System Identification[25]. However, compared to the preliminary version, all detailed proofs of the SARX minimality and identifiability results are provided here. They are mainly based on some technical results that give important properties of state space representation which arise from SARX systems. These technical results can be also used for analysising other structural properties of SARX systems. Moreover, we present an algorithm for finding an identifiable parameterization of SARX systems, the latter was not included in[25]. Furthermore, the organization of the paper has been improved in order to ease the reading. A motivating example is also provided.

Outline. In Section 2, we define SARX systems and the corresponding system-theoretic concepts such as minimality and identifiability. In Section 3.1, we present sufficient conditions guaranteeing (strong) minimality of SARX systems. We also discuss the relationship between minimality of a SARX system and that of its subsystems. In Section 3.2, we discuss the relationship between minimality and identifiability. We show that minimality and identifiability are generic properties. Concluding remarks are provided in Section 4.

Notations. Denote by $T = \mathbb{N}$ the time-axis of natural numbers. The notation described below is standard in formal language and automata theory[26,27]. 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. We denote by ϵ the *empty sequence (word)*. The length of a word w is denoted by $|w|$, i.e. $|w| = k$ means that the length of w is k; notice 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 $v \in X^*$ and recall that $\epsilon w = w\epsilon = w$. Furthermore, we denote by I_d the $d \times d$ identity matrix and by $\mathbf{O}_{d\times l}$ the $d \times l$ zero matrix.

2 | DEFINITION OF SARX SYSTEMS AND PROBLEM FORMULATION

This section aims at introducing the concepts of minimality, strong minimality, and identifiability for discrete-time SARX systems. A formal description of this class of switched systems is recalled in Definition 1.

Definition 1 (SARX systems)**.** A SARX system $\mathbf{S}$ of type (n_y, n_u), where $0 < n_u \leq n_y$ are integers, is a collection $\mathbf{S} = \{h_q\}_{q\in Q}$, where Q is the finite set of discrete modes and for every $q \in Q$, h_q is a $p\times(n_y p+n_u m)$ matrices, where p is the output dimension and m is the input dimension of the system. We will call a SARX system a SISO SARX system if $p = m = 1$. The dimension of the SARX system $\mathbf{S}$ is the number $pn_y + n_u m$ and is denoted by $\dim \mathbf{S}$.

Assigning semantics to SARX systems defined above requires that we first formalize the concept of input-output behavior for SARX systems. For this let us introduce the following notion of hybrid inputs of SARX systems.

Definition 2 (Hybrid inputs of SARX systems)**.** The hybrid inputs of SARX system $\mathbf{S}$ in Definition 1 are the elements of $\mathcal{U}^* = Q \times \mathbb{R}^m$. For any $t \geq 0$, a sequence w of the form

$$w = (q_0, u_0)\cdots(q_t, u_t) \in \mathcal{U}^+, \tag{1}$$

where we recall that $\mathcal{U}^+$ denotes the set of non-empty finite sequences of elements of $\mathcal{U}^*$, describes the scenario when discrete mode $q_i \in Q$ and continuous input $u_i \in \mathbb{R}^m$ are fed to $\mathbf{S}$ at time i, for $i = 0, \dots, t$.

Notation 1. In the sequel, without loss of generality, we will set $Q = \{1, \dots, D\}$ and we will use the following decomposition for the matrices h_q:

$$h_q = \begin{bmatrix} h_q^1 & h_q^2 & \cdots & h_q^{n_u+n_y} \end{bmatrix},$$

where $h_q^i \in \mathbb{R}^{p\times p}$, $i = 1, \dots, n_y$, $h_q^j \in \mathbb{R}^{p\times m}$, $j = n_y + 1, \dots, n_u + n_y$.

In order to introduce the formal definition of SARX minimality, the following concepts of input-output map realization as well as equivalence of SARX systems are needed.

Definition 3 (Input-output map realization and equivalence of SARX systems)**.** The SARX system $\mathbf{S}$ is a realization of the input-output map $f : \mathcal{U}^+ \to \mathbb{R}^p$, if for all $w \in \mathcal{U}^+$ of the form (1), the outputs

$$\mathbf{y}_i = f((q_0, \mathbf{u}_0)\cdots(q_i, \mathbf{u}_i)), i = 0, \dots, t$$

of $\mathbf{S}$ satisfy the equation

$$\mathbf{y}_t = h_{q_t}\phi_t \tag{2}$$

where we define the regressors $\phi_t \in \mathbb{R}^{(n_y p+n_u m)}$ as

$$\phi_t = \begin{bmatrix} \mathbf{y}_{t-1}^T & \mathbf{y}_{t-2}^T & \cdots & \mathbf{y}_{t-n_y}^T & \mathbf{u}_{t-1}^T & \cdots \mathbf{u}_{t-n_u}^T \end{bmatrix}^T, \tag{3}$$

and for all $j < 0$, we set $\mathbf{y}_j = 0$ and $\mathbf{u}_j = 0$. Two SARXs are called *equivalent*, if they are realizations of the same input-output map.

For the input-output maps of interest $f : \mathcal{U}^+ \to \mathbb{R}^p$, the value $f(w)$, with w of the form (1), describes the output of the system in Definition 1 at time t, generated as a response of the system to the hybrid input w. Another important concept is that of minimality.

Definition 4 (Minimality of SARX systems)**.** A SARX system $\mathbf{S}$ is minimal, if there exists no equivalent SARX of dimension less than $\dim \mathbf{S}$.

Next, we define the concept of discrete-time *linear switched system*[28,29] (abbreviated by LSS) associated with a SARX. This will be used to define the concept of strong minimality, which will play a central role in the rest of the paper. We will use the notation and terminology from[29] for linear switched systems in state-space form, which we recall below.

Definition 5. A *linear switched system* (abbreviated by LSS) is a discrete-time system Σ represented by

$$\begin{aligned} x_{t+1} &= A_{q_t}x_t + B_{q_t}u_t,\ x_0 \text{ fixed} \\ y_t &= C_{q_t}x_t. \end{aligned} \tag{4}$$

where, $x_t \in \mathbb{R}^n$ is the continuous state at time $t \in T$, $u_t \in \mathbb{R}^m$ is the continuous input at time $t \in T$, $y_t \in \mathbb{R}^p$ is the continuous output at time $t \in T$, $q_t \in Q$ is the discrete mode (state) at time t, Q is the finite set of discrete modes, and $x_0 \in \mathbb{R}^n$ is the initial state of Σ. For each discrete mode $q \in Q$, the corresponding matrices are of the form $A_q \in \mathbb{R}^{n\times n}$, $B_q \in \mathbb{R}^{n\times m}$ and $C_q \in \mathbb{R}^{p\times n}$.

Notation 2. We will use $(n, Q, \{(A_q, B_q, C_q) \mid q \in Q\}, x_0)$ as a short-hand notation for LSSs of the form (4).

Definition 6 (Associated LSS of SARX)**.** Let $\mathbf{S} = \{h_q\}_{q\in Q}$ be a SARX system of type (n_y, n_u). The LSS $\Sigma_{\mathbf{S}} = (n, Q, \{(A_q, B_q, C_q) \mid q \in Q\}, x_0)$ associated with the SARX $\mathbf{S}$ is given by:

$$A_q = \begin{bmatrix} A_q^y & A_q^u \end{bmatrix}, \qquad C_q = h_q, \qquad B_q = \begin{bmatrix} \mathbf{O}_{pn_y\times m} \\ I_m \\ \mathbf{O}_{m(n_u-1)\times m} \end{bmatrix}, \quad x_0 = 0 \tag{5}$$

with

$$A_q^y = \begin{bmatrix} \begin{bmatrix} h_q^1 & h_q^2 & \dots & h_q^{n_y-1} \end{bmatrix} & h_q^{n_y} \\ I_{(n_y-1)p} & \mathbf{O}_{(n_y-1)p\times p} \\ \mathbf{O}_{m\times p(n_y-1)} & \mathbf{O}_{m\times p} \\ \mathbf{O}_{(n_u-1)m\times p(n_y-1)} & \mathbf{O}_{(n_u-1)m\times p} \end{bmatrix}, \qquad A_q^u = \begin{bmatrix} \begin{bmatrix} h_q^{n_y+1} & h_q^{n_y+2} & \dots & h_q^{n_y+n_u-1} \end{bmatrix} & h_q^{n_u+n_y} \\ \mathbf{O}_{(n_y-1)p\times(n_u-1)m} & \mathbf{O}_{(n_y-1)p\times m} \\ \mathbf{O}_{m\times m(n_u-1)} & \mathbf{O}_{m\times m} \\ I_{(n_u-1)m} & \mathbf{O}_{(n_u-1)m\times m} \end{bmatrix}, \tag{6}$$

and h_q decomposed as in Notation 1.

Similarly to the case of SARX systems, we can define the concept of a LSS being a realization of an input-output map $f : \mathcal{U}^+ \to \mathbb{R}^p$. Informally, a LSS is a realization of f, if the for any sequence of discrete modes and inputs $w \in \mathcal{U}$, the output response of the LSS from its initial state to this sequence of inputs and discrete equals $f(w)$. Formally, consider a state $x_{init} \in \mathbb{R}^n$. For any input sequence $w \in \mathcal{U}^*$, let $x_\Sigma(x_{init}, w)$ be the state of Σ reached from x_{init} under input w, i.e. $x_\Sigma(x_{init}, w)$ is defined recursively as follows; $x_\Sigma(x_{init}, \epsilon) = x_{init}$, and if $w = v(q, u)$ for some $(q, u) \in \mathcal{U}$, $v \in \mathcal{U}^*$, then $x_\Sigma(x_{init}, w) = A_q x_\Sigma(x_{init}, v) + B_q u$. If $w \in \mathcal{U}^+$, then denote by $y_\Sigma(x_{init}, w)$ the output response of Σ to w, from the state x_{init}, i.e. if $w = v(q, u)$ for some $(q, u) \in \mathcal{U}$, $v \in \mathcal{U}^*$, then $y_\Sigma(x_{init}, w) = C_q x_\Sigma(x_{init}, v)$. The function

$$y_\Sigma : \mathcal{U}^+ \to \mathbb{R}^p, \quad w \mapsto y_\Sigma(w) = y(x_0, w),$$

is called the *input-output map* of Σ. An input-output map $f : \mathcal{U}^+ \to \mathbb{R}^p$ is said to be *realized* by a LSS Σ of the form (4) if f equals the input-output map y_Σ of Σ. In this case Σ is said to be a realization of f. We say that two LSSs are *equivalent*, if their input-output maps are equal.

Lemma 1. The SARX system $\mathbf{S}$ is a realization of the input-output map f if and only if the associated LSS $\Sigma_{\mathbf{S}}$ is a realization of f.

Proof. See the paper of Weiland *et al*[23]. □

The following corollary of Lemma 1 allows us to relate the problem of minimality of SARX to that of LSSs. The latter has already been investigated by Petreczky *et al*[29]. Formally, the dimension of a LSS Σ of the form (4), denoted by $\dim \Sigma$, is defined as the dimension n of its continuous state-space. Let $f : \mathcal{U}^+ \to \mathbb{R}^p$ be an input-output map and let Σ be a LSS which is a realization of f. Then Σ is *a minimal realization of* f, if for any LSS realization $\hat{\Sigma}$ of f, $\dim \Sigma \leq \dim \hat{\Sigma}$. That is, a LSS realization is a minimal realization of f if it has the smallest dimension of state-space among all the LSS which are realizations of f. We will say that a LSS Σ is a *minimal*, if it is a minimal realization of its own input-output map y_Σ.

Corollary 1. Assume $\mathbf{S}$ is a SISO SARX. If the associated LSS $\Sigma_{\mathbf{S}}$ is minimal, then $\mathbf{S}$ is minimal.

Proof of Corollary 1. Assume that $\mathbf{S}$ is not minimal. Then there exists an equivalent $\mathbf{S}_m$ of type (n_y', n_u') such that $n_y' + n_u' < n_y + n_u$. But this implies that $\dim \Sigma_{\mathbf{S}_m} = n_y' + n_u' < n_y + n_u = \dim \Sigma_{\mathbf{S}}$, which contradicts to the minimality of $\Sigma_{\mathbf{S}}$. □

Remark 1. It can be noticed that none of the linear subsystems of $\Sigma_{\mathbf{S}}$ in (5)–(6) is minimal. Indeed, for each $q \in Q$, A_q contains a zero row, hence $\operatorname{rank} A_q < n_u + n_y$. This means that $\lambda = 0$ is an eigenvalue of A_q. By the PBH criterion, (C_q, A_q) is an observable pair if and only if the matrix $\begin{bmatrix} C_q^T, & \lambda I - A_q^T \end{bmatrix}$ has rank $n_y + n_u$ for all the eigenvalues of A_q. We will show that for $\lambda = 0$ this matrix cannot be of full row rank. To see this, for $\lambda = 0$ the matrix becomes $\begin{bmatrix} C_q^T, -A_q^T \end{bmatrix}$. But C_q equals the first row of A_q multiplied by -1. Hence, $[C_q^T, -A_q^T]$ will have the same rank as A_q and that is smaller than $n_u + n_y$. Thus, the linear subsystems $\Sigma_{\mathbf{S}}$ in (5)–(6) are not observable and consequently, they are not minimal. However, as we shall see later, they are generically minimal.

The result of Corollary 1 prompts us to propose the following definition.

Definition 7 (Strong minimality of SARX systems)**.** A SARX system $\mathbf{S}$ is called *strongly minimal*, if the corresponding LSS $\Sigma_{\mathbf{S}}$ is minimal.

Remark 2. By Corollary 1, strong minimality implies minimality. However, minimality does not imply strong minimality. Indeed, consider the SARX system $\mathbf{S}$ with discrete modes $Q = \{1, 2\}$ such that the ARX subsystem associated with mode 1 is $\mathbf{y}_t = -\mathbf{y}_{t-2} + \mathbf{u}_{t-1}$ and the ARX subsystem associated with mode 2 is $\mathbf{y}_t = -2\mathbf{y}_{t-2} + 2\mathbf{u}_{t-1}$. The two ARX subsystems are distinct, each of them is minimal, yet the associated LSS $\Sigma_{\mathbf{S}}$ is not minimal (in fact, it is not observable). The latter can be checked using the minimality conditions of Petreczky *et al*[24].

Remark 3. Minimality of the ARX subsystems is not necessary for strong minimality (and hence minimality) of the whole system. To see this, consider again the SARX system $\mathbf{S}$ with two discrete modes $Q = \{1, 2\}$ such that the ARX subsystem in mode 1 is of the form $\mathbf{y}_t = 8\mathbf{y}_{t-1} - 15\mathbf{y}_{t-2} + \mathbf{u}_{t-1} - 3\mathbf{u}_{t-2}$, and the ARX subsystem in mode 2 is of the form $\mathbf{y}_t = \mathbf{y}_{t-1} + 2\mathbf{y}_{t-2} + \mathbf{u}_{t-1} + \mathbf{u}_{t-2}$. The transfer function of the ARX in the first mode is $\frac{z-3}{z^2-8z+15} = \frac{1}{z-5}$ and the transfer function of the second ARX is $\frac{z+1}{z^2-z-2} = \frac{1}{z-2}$, hence neither of them is minimal. Yet, by using the conditions of Petreczky *et al*[24], it can be easily shown that the LSS $\Sigma_{\mathbf{S}}$ is minimal. Since strong minimality implies minimality of SARX systems, we get that $\mathbf{S}$ is minimal.

In order to be able to speak of identifiability, we need the notion of parameterization of SARX systems.

Notation 3. Denote by $SARX(n_y, n_u, m, p, Q)$ the set of all SARX systems of type (n_y, n_u) with input space $\mathbb{R}^m$, output space $\mathbb{R}^p$, and set of discrete modes Q.

Definition 8. (Parameterization of SARX systems) Assume that $\Theta \subseteq \mathbb{R}^d$ is the set of parameters. A *SARX parameterization* is a map

$$\mathbf{\Pi}^{\text{SARX}} : \Theta \to SARX(n_y, n_u, m, p, Q) \tag{7}$$

Example 1. Consider the discrete-time model of the intake manifold of a spark ignition engine as described in[30] and[31],

$$y_t = g(\phi_t, p_t, v_t),$$

where $\phi_t = \begin{bmatrix} y_{t-1} & y_{t-2} & u_{t-1} & u_{t-2} \end{bmatrix}^\top$, the output y_t is the normalized air charge, the input u_t is the opening of the throttle valve; p_t and v_t refer respectively to the pressure inside the intake manifold and the speed of the engine. Here, p and v are viewed as external signals which take values in some bounded intervals. We refer to[30] and[31] for more details. Inspired by the Linear Parameter Varying (LPV) parameterization of the intake model proposed in[30], one can consider approximating the intake manifold with a SISO SARX system of type $(2, 2)$ by viewing p and v as piecewise constant signals, each of which is allowed, for simplicity, to take only two possible values, $(p_1, p_2) = (0.3, 0.7)$, and $(v_1, v_2) = (2, 5)$ respectively. Indeed, since p and v are external excitation signals, one can select them in such a way that they are exactly piecewise constant with values prescribed as above. In that case, the SARX model considered here coincides with the LPV one.

Let us form a vector $\bar{p}_{(i,j)} = \begin{bmatrix} 10^{-3} & p_i & p_i^2 & v_j & v_j^2 \end{bmatrix}^\top$ and introduce the bijective map $\sigma : \{1, 2\} \times \{1, 2\} \to Q = \{1, 2, 3, 4\}$, $(i, j) \mapsto q = \sigma(i, j)$. where $p_1 = 0.3$, $p_2 = 0.7$, $v_1 = 2$, $v_2 = 5$. Then the LPV parameterization in[30] reduces to a SARX one $\mathbf{\Pi}^{\text{SARX}} : \Theta = \mathbb{R}^{20} \to SARX(2, 2, 1, 1, Q)$ with $\mathbf{\Pi}^{\text{SARX}}(\theta) = \left\{h_q(\theta)\right\}_{q \in Q}$, $h_q(\theta) = M(\theta)\bar{p}_{(i,j)}$ for $q = \sigma(i, j)$ and

$$M(\theta) = \begin{bmatrix} \theta_1 & \theta_5 & \cdots & \theta_{17} \\ \theta_2 & \theta_6 & \dots & \theta_{18} \\ \theta_3 & \theta_7 & \cdots & \theta_{19} \\ \theta_4 & \theta_8 & \dots & \theta_{20} \end{bmatrix} \in \mathbb{R}^{4 \times 5}$$

Here, $\bar{p}_{(i,j)}$ is a vector of the form $\bar{p}_{(i,j)} = \begin{bmatrix} 10^{-3} & p_i & p_i^2 & v_j & v_j^2 \end{bmatrix}^\top$ where $p_1 = 0.3$, $p_2 = 0.7$, $v_1 = 2$, $v_2 = 5$. $p_1 \in \{311.75; 737.25\}$ and $p_2 \in \{2132.5; 4877.5\}$ he input-output map of such a system is hence defined by

$$y_t = h_{q_t}^\top \phi_t$$

with $\phi_t = \begin{bmatrix} y_{t-1} & y_{t-2} & u_{t-1} & u_{t-2} \end{bmatrix}^\top$. This example illustrates how the behavior of some physical systems can be modeled by SARX systems.

Definition 9 (Identifiability of SARX parameterizations)**.** The parameterization $\mathbf{\Pi}^{\text{SARX}}$ is called *identifiable*, if for $\theta_1 \neq \theta_2 \in \Theta$, the corresponding SARX $\mathbf{\Pi}^{\text{SARX}}(\theta_1)$ and $\mathbf{\Pi}^{\text{SARX}}(\theta_2)$ are not equivalent.

The intuition behind the above definition is that if a parameterization is not identifiable, then there might exist different parameter values which yield the same observed behavior and hence they cannot be distinguished from each other by input-output experiments. Hence, the problem of identifying the parameters of a SARX models from a non identifiable parameterization is ill-posed. It would be tempting to try to reduce identifiability of SARX parametrizations to that of the parametrization of the corresponding ARX subsystems. This would then allow us to use existing theory on identifiability of ARX parametrizations. **Unfortunately, identifiability of a SARX parametrization does not imply the identifiability of the corresponding parametrization of ARX subsystems.** The example below demonstrates this point.

Example 2. Consider the SARX parametrization $\mathbf{\Pi}^{\text{SARX}}$ with $\Theta = \mathbb{R}^2$, and consider the parametrization $\mathbf{\Pi}^{\text{SARX}}((\theta_1, \theta_2)) = \{h_q(\theta_1, \theta_2)\}_{q \in Q}$, where

$$h_1 = \begin{bmatrix}(\theta_1 + \theta_2) & -\theta_1\theta_2 & 1 & -\theta_2\end{bmatrix},$$

and

$$h_2 = \begin{bmatrix}(2 + \theta_2) & -2\theta_2 & 1 & -\theta_2\end{bmatrix}.$$

Define the set $G = \{(\theta_1, \theta_2) \mid \theta_1 \neq 2\}$. Consider the restriction $\mathbf{\Pi}^{\text{SARX}}|_G$ of $\mathbf{\Pi}^{\text{SARX}}$ to G. Using Theorem 1 one can check that for any $(\theta_1, \theta_2) \in G$, the SARX system $\mathbf{\Pi}^{\text{SARX}}((\theta_1, \theta_2))$ is strongly minimal. Hence, the parametrization $\mathbf{\Pi}^{\text{SARX}}|_G$ is identifiable by Theorem 2. Identifiability of $\mathbf{\Pi}^{\text{SARX}}|_G$ can also be checked by considering the switching sequence 112 and input $\mathbf{u}_0 = 1, \mathbf{u}_t = 0$, $t > 0$ and noticing that then $\mathbf{y}_0$, $\mathbf{y}_1 = 1$, $\mathbf{y}_2 = \theta_1$, $\mathbf{y}_3 = 2\theta_1 + \theta_2\theta_1 - 2\theta_2$ from which $\theta_2 = \frac{(\mathbf{y}_3 - 2\theta_1)}{\theta_1 - 2}$. Hence, θ_1 and θ_2 can be determined from the outputs $\mathbf{y}_2$ and $\mathbf{y}_3$.
Note however, that for any (θ_1, θ_2), the ARX subsystems of $\mathbf{\Pi}^{\text{SARX}}(\theta_1, \theta_2)$ are not identifiable, since their dynamics does not depend on θ_2.

This implies that identifiability of SARX parametrizations has to be investigated separately. Recall now that identifiability of ARX parametrizations is closely related to their minimality. Hence, we start by investigating mnimality of SARX models.

3 | MAIN RESULTS

In this section we present the main results of the paper. First, in Section 3.1 we discuss minimality of SARX systems. In Section 3.2 we use the results of Section 3.1 to characterize identifiability of SARX systems. In the rest of the paper we consider only SISO SARX systems, i.e., we tacitly assume that $p = m = 1$ unless stated otherwise.

3.1 | Minimality conditions for switched ARX systems

In this section we will analyze minimality of SARX systems.We start by Lemmas 2, 3 and 4 that provide some simple and crucial properties of minimal SARX systems

Lemma 2. If the SISO SARX system $\mathbf{S}$ is minimal, then there must exist $q \in Q$ such that $h_q^{n_u+n_y} \neq 0$.

Proof. Assume the contrary, i.e. that $h_q^{n_u+n_y} = 0$ for all $q \in Q$. Define the vectors $\hat{h}_q \in \mathbb{R}^{n_y+n_u-1}$ by $\hat{h}_q = \begin{bmatrix}h_q^1 & \dots & h_q^{n_u+n_y-1}\end{bmatrix}$, $q \in Q$. Define the regressors $\hat{\phi}_t$ as

$$\hat{\phi}_t = \begin{bmatrix}\mathbf{y}_{t-1}^T & \dots, \mathbf{y}_{t-n_y}^T & \mathbf{u}_{t-1}^T & \dots & \mathbf{u}_{t-(n_u-1)}^T\end{bmatrix}^T,$$

where we used the convention that $\mathbf{y}_j = 0$ and $\mathbf{u}_j = 0$ for $j < 0$. It then follows that $\mathbf{y}_t = h_q\phi_t = \hat{h}_q\hat{\phi}_t$ for all $t \in T$. Hence, $\hat{\mathbf{S}} = (\{\hat{h}_q\}_{q \in Q})$ realizes the same input-output map as $\mathbf{S}$. But the dimension of $\hat{\mathbf{S}}$ is smaller than that of $\mathbf{S}$, which contradicts the minimality of $\mathbf{S}$. □

Lemma 3. A SISO ARX system is minimal according to Definition 4 if and only if the numerator and denominator of its transfer function are co-prime.

Proof. The proof follows from the classical linear theory, by observing that two ARX systems realize the same input-output map if and only if they have the same transfer function (modulo zero/pole cancellation).

Consider an ARX system $\mathbf{y}_t = h_{q_t}\phi_t$ and assume that it is minimal. If its transfer function admits a zero-pole cancellation, then the degrees of the numerator and denominator of the transfer function decrease by one. The latter means that the transfer function can be realized by an ARX of type $(n_y - 1, n_u - 1)$. The dimension of the latter is $n_y + n_u - 2$ and hence smaller than that of the original system, which was supposed to be minimal. Moreover, this new ARX system will realize the same input-output map as the original one.

Conversely, consider an ARX system $\mathbf{S}$ whose transfer function does not allow zero/pole cancellation. Let f be the input-output map of $\mathbf{S}$ and assume that the ARX system $\hat{\mathbf{S}}$ is a minimal realization of f. Then the transfer function $H_{\hat{\mathbf{S}}}(z)$ cannot allow a zero/pole cancellation and it must be equal to the transfer function $H_{\mathbf{S}}(z)$ of $\mathbf{S}$. Since neither $H_{\mathbf{S}}(z)$ nor $H_{\hat{\mathbf{S}}}(z)$ allow zero/pole cancellation, their equality implies the equality of the numerators and denominators respectively, viewed as polynomials. In particular, the corresponding coefficients are the same and hence the parameters of the two ARX systems are the same too. In particular, the dimensions of the two systems will be the same, and hence $\mathbf{S}$ is then a minimal realization of its input-output map. □

Remark 4. Recall that in the classical literature, a SISO ARX is said to be minimal if and only if the numerator and the denominator of its transfer function are co-prime polynomials. Consequently, Lemma 3 shows that our definition of minimality is consistent with the traditional one.

Lemma 4. If at least one of the ARX subsystems of a SISO SARX system is minimal, then the system is minimal.

Proof. Consider $\mathbf{S} = \{h_q\}_{q\in Q}$ and assume that for some $q_m \in Q$, the ARX $\mathbf{y}_t = h_{q_m}\phi_t$ is minimal. Assume that $\mathbf{S}$ is not minimal and hence there exists a SARX $\mathbf{S}_m = (\{\hat{h}_q\}_{q\in Q})$ such that $\dim \mathbf{S}_m \leq \dim \mathbf{S}$ and $\mathbf{S}_m$ realizes the same input-output map as $\mathbf{S}$. It then follows that the dimension of the ARX $\mathbf{y}_t = h_{q_m}\phi_t$ is larger than that of $\mathbf{y}_t = \hat{h}_{q_m}\phi_t$. It also follows that both $\mathbf{y}_t = \hat{h}_{q_m}\phi_t$ and $\mathbf{y}_t = h_{q_m}\phi_t$ realizes the same linear input-output map[1]. This contradicts the minimality of $\mathbf{y}_t = h_{q_m}\phi_t$. □

The definition 4 of minimality for SARX systems might seem ambiguous because it does not exclude explicitly the possibility of having two minimal SARX realizations of types (n_y, n_u) and $(\hat{n}_y, \hat{n}_u)$ respectively for the same input-output map with $(n_y, n_u) \neq (\hat{n}_y, \hat{n}_u)$. According to the lemma below, this is impossible at least in the SISO case.

Lemma 5. Assume that $\mathbf{S}_1$ and $\mathbf{S}_2$ are two minimal and equivalent SISO SARX systems such that $\mathbf{S}_1$ is of type (n_y, n_u) and $\mathbf{S}_2$ is of type $(\hat{n}_y, \hat{n}_u)$. Then $(n_y, n_u) = (\hat{n}_y, \hat{n}_u)$.

Proof. Pick any discrete state q and consider the transfer functions $H_i(z)$, $i = 1, 2$, of the ARX system in mode q associated with the SARX $\mathbf{S}_i$, $i = 1, 2$. Since $\mathbf{S}_1$ and $\mathbf{S}_2$ are equivalent, they produce the same response to any input if the discrete mode is kept to be q. Hence, the ARX systems corresponding to the mode q are also equivalent, i.e. $H_1(z)$ and $H_2(z)$ describe the same input-output behavior. This means that the transfer functions $H_1(z)$ and $H_2(z)$ are equal as rational expressions, after possibly performing zero/pole cancellation. The degrees of the numerators of $H_1(z)$ and $H_2(z)$ are respectively n_u and $\hat{n}_u$ and the degrees of the denominators of $H_1(z)$ and $H_2(z)$ are respectively n_y and $\hat{n}_y$. Performing zero/pole cancellation does not change the difference between the degree of the numerator and the degree of the denominator. Hence, we obtain that $n_y - n_u = \hat{n}_y - \hat{n}_u$ must hold. But since both $\mathbf{S}_1$ and $\mathbf{S}_2$ are minimal SARX realizations of the same input-output map, their dimensions must agree and hence $n_y + n_u = \hat{n}_y + \hat{n}_u$. It is easy to see that the only solution to the system of equations

$$\begin{cases} n_y - n_u = \hat{n}_y - \hat{n}_u, \\ n_y + n_u = \hat{n}_y + \hat{n}_u, \end{cases}$$

is $n_y = \hat{n}_y$ and $n_u = \hat{n}_u$. □

As we have seen in the previous section, strong minimality implies minimality. By [29,24], strong minimality and hence minimality, can be checked algorithmically. Indeed, strong minimality of a SARX system $\mathbf{S}$ means minimality of the associated LSS $\Sigma_{\mathbf{S}}$. The latter can be checked by checking if the rank of each of the finite span-reachability matrix $\mathcal{R}(\Sigma_{\mathbf{S}})$ of $\Sigma_{\mathbf{S}}$ and the finite observability matrix $\mathcal{O}(\Sigma_{\mathbf{S}})$ of $\Sigma_{\mathbf{S}}$ considered in Theorem 2 [29] equals the dimension of $\Sigma_{\mathbf{S}}$. We can also formulate sufficient conditions for minimality which do not involve computing LSSs.

[1] Namely, the map which maps input $\mathbf{u}_0, \ldots, \mathbf{u}_t$ to the output $\mathbf{y}_t = f((q_m, \mathbf{u}_0) \cdots (q_m, \mathbf{u}_t))$, where f is the input-output map of $\mathbf{S}$

Theorem 1 (Sufficient conditions for (strong) minimality)**.** Consider a SISO SARX system $\mathbf{S} = \{h_q\}_{q\in Q}$ of type (n_y, n_u). For all modes $q, \hat{q} \in Q$, define the polynomials

$$\chi_q(z) = z^{n_y} - \sum_{j=1}^{n_y} h_q^j z^{n_y-j}, \qquad \upsilon_q(z) = \sum_{j=1}^{n_y} h_q^j z^{n_y-j}, \qquad \phi_{\hat{q},q}(z) = \sum_{j=1}^{n_u} h_q^{j+n_y} \psi_{\hat{q},q,n_u-j}(z),$$

with $\psi_{\hat{q},q,j}(z)$ defined recursively for $j = 0, 1, 2 \dots$, as follows: $\psi_{\hat{q},q,0}(z) = 1$ and

$$\psi_{\hat{q},q,j+1}(z) = z\psi_{\hat{q},q,j}(z) + (h_q - h_{\hat{q}})d_j \tag{8}$$

where the vectors $d_j \in \mathbb{R}^{n_y+n_u}$ are defined as follows: $d_0 = e_1$ and when $d_j = (d_{j,1}, \dots, d_{j,n_y}, 0, \dots, 0)^T$ with $d_{j,1}, \dots, d_{j,n_y} \in \mathbb{R}$, $d_{j+1} = (h_q d_j, d_{j,1}, \dots, d_{j,n_y-1}, 0, \dots, 0)^T$. Then $\mathbf{S}$ is strongly minimal, if the following conditions hold:

(A) there exists discrete modes q_0 and q_1 such that the polynomials $\chi_{q_0}(z)$ and $\phi_{q_0,q_1}(z)$ are co-prime, and

(B) there exists discrete modes q_2 and q_3, such that $\upsilon_{q_3}(z)$ and $\chi_{q_2}(z)$ are co-prime, $h_{q_2}^{n_u+n_y} \neq 0$ and $h_{q_3}^{n_y} \neq \frac{h_{q_3}^{n_y+n_u}}{h_{q_2}^{n_y+n_u}} h_{q_2}^{n_y}$.

Proof. See Appendix B. □

Remark 5. Theorem 1 is analogous to the well-known result that if a SISO transfer has no zero-pole cancellation (i.e. its numerator and denominator are coprime) and its denominator is of degree n, then all its minimal realizations are of order n. Due to the presence of switching, the formulation of Theorem 1 is more involved. In addition, Theorem 1 does not imply the classical results, since condition **(B)** of the theorem is always false if there is only one discrete state.

In order to demonstrate the utility of Theorem 1, we present the following examples.

Example 3. Let us apply Theorem 1 to the SARX system $\mathbf{S}$ with two discrete modes $Q = \{1, 2\}$ such that the ARX subsystem in mode 1 is of the form $\mathbf{y}_t = 8\mathbf{y}_{t-1} - 15\mathbf{y}_{t-2} + \mathbf{u}_{t-1} - 3\mathbf{u}_{t-2}$, and the ARX subsystem in mode 2 is of the form $\mathbf{y}_t = \mathbf{y}_{t-1} + 2\mathbf{y}_{t-2} + \mathbf{u}_{t-1} + \mathbf{u}_{t-2}$. We obtain that $n_y = n_u = 2$, $h_1 = \begin{bmatrix}8 & -15 & 1 & -3\end{bmatrix}$ and $h_2 = \begin{bmatrix}1 & 2 & 1 & 1\end{bmatrix}$. Hence, $h_1^{n_y+n_u} = -3 \neq 0$, and $\chi_1(z) = z^2 - 8z + 15$, $\psi_{1,2,1}(z) = z - 7$ $\upsilon_2(z) = z + 2$ $\phi_{2,1}(z) = z - 6$. It is clear that the roots of $\chi_1(z)$ are 5 and 3 and hence $\upsilon_2(z)$ and $\chi_1(z)$ are co-prime and $\phi_{2,1}(z)$ and $\chi_1(z)$ are co-prime. Moreover, $h_2^{n_y} - \frac{h_2^{n_u+n_y} h_1^{n_y}}{h_1^{n_u+n_y}} = 2 - \frac{15}{3} = -3 \neq 0$. Hence, conditions **(A)** and **(B)** of Theorem 1 hold and thus $\mathbf{S}$ is (strongly) minimal.

Example 4. Consider the SARX system from Example 1, and choose the parameter vector $\underline{\theta} = \begin{bmatrix}\theta_1 & \dots & \theta_{20}\end{bmatrix}$ as

$$\begin{aligned}\underline{\theta} =& [0.0046, -0.0091, 0.0005, -0.0019, 0.4881, -0.9555, 0.0519, -0.1973, -0.4881, 0.9555, -0.0519,\\ & 0.1973, 6.4616, -12.6262, 0.6924, -2.6043, -1.2564, 2.6133, -0.0989, 0.5625]\end{aligned} \tag{9}$$

It then follows that with this parameters, the SARX becomes $\mathbf{S} = \{n_q\}_{q\in Q}$, $Q = \{(i, j) \mid i, j = 1, 2\}$ with $n_{(1,1)} = n_{(2,1)} = \begin{bmatrix}8 & -15 & 1 & -3\end{bmatrix}$ and $n_{(1,2)} = n_{(2,2)} = \begin{bmatrix}1 & 2 & 1 & 1\end{bmatrix}$. Note that the parameter vectors are the same as in Example 3. It follows that $\chi_{(1,1)}(z) = z^2 - 8z + 15$, $\phi_{(1,1),(1,2)}(z) = z - 6$. Hence, $\chi_{(1,1)}$ and $\phi_{(1,1),(1,2)}$ are co-prime and condition **(A)** of Theorem 1 holds for $q_0 = (1, 1)$ and $q_1 = (1, 2)$. Moreover, $\upsilon_{(1,2)}(z) = z + 2$ and $\chi_{(1,1)}(z) = z^2 - 8z + 15$ are also co-prime and $h_{(1,2)}^{n_y} = 2$, $h_{(1,2)}^{n_y+n_u} = 1$ and $h_{(1,1)}^{n_y} = -15$, $h_{(1,1)}^{n_u+n_y} = -3$. Hence, $h_{(1,2)}^{n_y} - \frac{h_{(1,2)}^{n_u+n_y} h_{(1,1)}^{n_y}}{h_{(1,1)}^{n_u+n_y}} = 2 - \frac{15}{3} = -3 \neq 0$, hence condition **(B)** of Theorem 1 holds for $q_2 = (1, 1)$ and $q_3 = (1, 2)$. That is, the SARX from Example 1 with the choice of parameters as in (9) is strongly minimal.

3.2 | Identifiability conditions for Switched ARX systems

In this section we study identifiability of SARX systems.

Theorem 2, which is one of the main results of the paper, describes the relationship between strong minimality and identifiability. More precisely, it show that strong minimality is sufficient for identifiability. In order to restrict attention to strongly minimal SARX systems, we introduce the following terminology.

Definition 10 (Minimality of SARX parametrizations)**.** The parametrization $\mathbf{\Pi}^{\text{SARX}}$ is called *minimal* (resp. *strongly minimal*), if for all $\theta \in \Theta$, $\mathbf{\Pi}^{\text{SARX}}(\theta)$ is minimal (resp. strongly minimal).

If a SARX parametrization is strongly minimal, then the corresponding LSSs parametrization will be minimal. Hence, we can apply the conditions and algorithms provided by Petreczky *et al*[24] for analyzing the identifiability of the latter parametrization. By Corollary 6 the identifiability of the latter parametrization is identifiability of the original SARX parametrization.

In fact, for the SISO case (i.e. when $p = m = 1$), we can derive even stronger results, by showing that *minimality is sufficient for identifiability*. To this end, we need the following definition.

Definition 11 (Injective SARX parametrizations)**.** A SARX parametrization $\mathbf{\Pi}^{\text{SARX}}$ is said to be *injective* if $\mathbf{\Pi}^{\text{SARX}}$ is an injective map.

An injective parametrization allows us to exclude the situation where two different parameter values lead to the same SARX system. The ARX parametrization $\mathbf{y}_t = \theta^2\mathbf{y}_{t-1} + \mathbf{u}_{t-1}$ with $\theta \in \mathbb{R}$ is not injective, since any θ and $-\theta$ always lead to the same ARX system.

Theorem 2. Assume that $p = m = 1$. If a SISO SARX parametrization $\mathbf{\Pi}^{\text{SARX}}$ is injective and strongly minimal, then $\mathbf{\Pi}^{\text{SARX}}$ is identifiable.

Proof. See Appendix C. □

Theorem 2 allows us to find an identifiable sub-parametrization of a SARX parametrization by checking finding a sub-parametrization which is strong minimal. One way to check strong minimality is by checking if the conditions of Theorem 1 are satisfied. This can easily be done, for parameterizations in which the coefficients of the SARX systems depend on the parameters in a polynomial way. To this end, we introduce the following terminology.

Definition 12 (Polynomial parametrization)**.** Let $K = (pn_y + mn_u)|Q|$. Then any SARX system of type (n_y, n_u) can be identified with a point in $\mathbb{R}^K$, by identifying the system with its parameters $\{h_q\}_{q\in Q}$. Thus, $SARX(n_y, n_u, m, p, Q)$ can be identified with the space $\mathbb{R}^K$. A parametrization $\mathbf{\Pi}^{\text{SARX}}$ is said to be *polynomial*, if Θ is an affine algebraic variety and $\mathbf{\Pi}^{\text{SARX}}$ is a polynomial map from Θ to $SARX(n_y, n_u, m, p, Q)$.

Let $\mathbf{\Pi}^{\text{SARX}}$ be a polynomial SISO SARX parametrization, i.e., $p = m = 1$. Below we present a procedure to find a subset $\hat{\Theta} \subseteq \Theta$ such that for each $\theta \in \hat{\Theta}$, the SARX system $\mathbf{\Pi}^{\text{SARX}}(\theta)$ satisfies the conditions of Theorem 1, hence it is strongly minimal, and as a consequence the parametrization $\mathbf{\Pi}^{\text{SARX}}|_{\hat{\Theta}} : \hat{\Theta} \ni \theta \mapsto \mathbf{\Pi}^{\text{SARX}}(\theta)$ is strongly minimal. To this end, we introduce the following notation. We will use the standard notation and terminology from commutative algebra, see[32]. In particular, we will need the notion of an ideal, generator of an ideal, Gröbner basi of an ideal, product of ideals from[32]. For each $\theta \in \Theta$, if $\mathbf{\Pi}^{\text{SARX}}(\theta) = \{n_q(\theta)\}_{q\in Q}$, then denote by $\chi_q(\theta)(z)$, $\upsilon_q(\theta)(z)$ and $\phi_{q,\hat{q}}(\theta)(z)$ the polynomials $\chi_q(z), \upsilon_q(z), \phi_{q,\hat{q}}(z)$, $\hat{q}, q \in Q$, defined in Theorem 1 for $n_q = n_q(\theta)$. Then, since $\mathbf{\Pi}^{\text{SARX}}$ is polynomial, the dependence of $n_q(\theta)$ and the coefficients of $\chi_q(\theta)(z)$, $\upsilon_q(\theta)(z)$ and $\phi_{q,\hat{q}}(\theta)(z)$ on θ is polynomial. That is, there exist polynomials $n_q^{f,i} \in \mathbb{R}[X_1, \dots, X_d]$, $i = 1, \dots, n_u + n_y$, in variables $X_1, \dots, X_d$, and polynomials $\chi_q^f(X_1, \dots, X_d, z), \upsilon_q^f(X_1, \dots, X_d, z), \phi_{q,\hat{q}}^f(X_1, \dots, X_d, z)$ in variables $X_1, \dots, X_d, z$ such that $n_q^i(\theta) = n_q^{f,i}(\theta)$, where $n_q^i(\theta)$ denotes the ith components of $n_q(\theta)$, and $i = 1, \dots, n_y + n_u$, $\chi_q(\theta)(z) = \chi_q^f(\theta, z)$, $\upsilon_q(\theta) = \upsilon_q^f(\theta, z)$ and $\phi_{q,\hat{q}}(\theta)(z) = \phi_{q,\hat{q}}^f(\theta, z)$, $\hat{q}, q \in Q$. In order to apply Theorem 1, it is necessary to have a sufficient conditions for co-primeness of two polynomials in z, coefficients of which are polynomial functions of θ. To this end, assume that $Q_i(X_1, \dots, X_d, z)$, $i = 1, 2$ are two polynomials. Consider the ideal $I(Q_1, Q_2)$ generated by the polynomials Q_1, Q_2 and consider the ideal $J(Q_1, Q_2) = I(Q_1, Q_2) \cap \mathbb{R}[X_1, \dots, X_d]$. The ideal $J(Q_1, Q_2)$ is finitely generated, and the set of its generators S can be calculated from the polynomial Q_1, Q_2 using standard algorithms from compute algebra, see[32] and the toolbox[33].

Lemma 6. If there exist $g \in S$ such that $g(\theta) \neq 0$, then the univariate polynomials $Q_i(\theta, z) \in \mathbb{R}[z]$, $i = 1, 2$ are co-prime.

Proof of Lemma 6. Indeed, since $g_i \in J(Q_1, Q_2) \subseteq I(Q_1, Q_2)$, $g(X_1, \dots, X_d) = Q_1(X_1, \dots, X_d, z)\alpha(X_1, \dots, X_d, z) + Q_1(X_1, \dots, X_d, z)\beta(X_1, \dots, X_d, z)$ for some polynomials α, β in $R[X_1, \dots, X_d]$. In particular, $g(\theta) = Q_1(\theta, z)\alpha(\theta, z) + Q_2(\theta, z)\beta(\theta, z)$ and since $g(\theta) \neq 0$, $Q_1(\theta, z)\frac{\alpha(\theta,z)}{g(\theta)} + Q_2(\theta, z)\frac{\alpha(\theta,z)}{g(\theta)} = 1$, which by Bezout's identity implies that $Q_1(\theta, z), Q_2(\theta, z)$ are co-prime. □

Lemma 6 implies that for any $\theta \in \{\theta \in \Theta \mid \exists g \in S : g(\theta) \neq 0\}$, the polynomials $Q_i(\theta, z) \in \mathbb{R}[z]$, $i = 1, 2$ are co-prime. We can apply Lemma 6 to find $\hat{\Theta} \subseteq \Theta$ such that for any $\theta \in \hat{\Theta}$, the SARX system $\mathbf{\Pi}^{\text{SARX}}(\theta)$ satisfies conditions **(A)** and **(B)** of Theorem 1. More precisely, we propose the following algorithm for finding a strongly minimal sub-parametrization of a parametrization.

Procedure 1 (Identifiable polynomial parametrization)**.**

1. For each $q, \hat{q} \in Q$, consider the ideal $I(\chi_q^f, \phi_{q,\hat{q}}^f)$ generated by the polynomials $\chi_q^f, \phi_{q,\hat{q}}^f$ and calculate the Gröbner basis $S_{A,q,\hat{q}} \subseteq \mathbb{R}[X_1, \dots, X_d]$ of the ideal $I(\chi_q^f, \phi_{q,\hat{q}}^f) \cap \mathbb{R}[X_1, \dots, X_d]$ using standard algorithms[32], implemented, for example, in the toolbox[33].
2. For each $q, \hat{q} \in Q$, consider $I(\chi_q^f, v_{\hat{q}}^f)$ generated by the polynomials $\chi_q^f, v_{\hat{q}}^f$ and calculate the Gröbner basis $S'_{B,q,\hat{q}} \subseteq \mathbb{R}[X_1, \dots, X_d]$ of the ideal $I(\chi_q^f, v_{\hat{q}}^f) \cap \mathbb{R}[X_1, \dots, X_d]$, using standard algorithms[32], for an implementation see[33]. Let $S_{B,q,\hat{q}} = \{P \cdot Q \mid P \in S'_{B,q,\hat{q}}\}$, where $Q = n_q^{f,n_u+n_y}(n_{\hat{q}}^{f,n_y} n_q^{f,n_u+n_y} - n_q^{f,n_y} n_{\hat{q}}^{n_y+n_u})$.
3. Let I_A be the ideal generated by $\bigcup_{q,\hat{q}\in Q} S_{A,q,\hat{q}}$, and let I_B be the ideal generated by $\bigcup_{q,\hat{q}\in Q} S_{B,q,\hat{q}}$ and let S be the Gröbner basis of the ideal $I_A \cdot I_B = \{P_1 \cdot P_2 \mid P_1 \in I_A, P_2 \in I_B\}$. Note that S can be computed from the Gröbner basis of I_A and I_B, which, in turn, can easily be computed from the finite sets $\bigcup_{q,\hat{q}\in Q} S_{A,q,\hat{q}}$ and $\bigcup_{q,\hat{q}\in Q} S_{B,q,\hat{q}}$ using a standard algorithm for computing Gröbner basis from a generator set of an ideal[32,33]. Define the parametrization: $\mathbf{\Pi}^{\text{SARX}}|_{\hat{\Theta}} : \hat{\Theta} \ni \theta \mapsto \mathbf{\Pi}^{\text{SARX}}(\theta)$, where
$$\hat{\Theta} = \{\theta \in \Theta \mid \exists P \in S : P(\theta) \neq 0\}.$$

Procedure 1 was implemented, the code is available in[34].

Lemma 7. The parametrization $\mathbf{\Pi}^{\text{SARX}}|_{\hat{\Theta}}$ calculated by Procedure 1 is strongly minimal and hence it is identifiable

Proof of Lemma 7. Assume that $\theta \in \hat{\Theta}$ and let $P \in S$ be such that $P(\theta) \neq 0$. Then, since $P \in I_A \cdot I_B$, $P = P_1 P_2$ for some $P_1 \in I_A$ and $P_2 \in I_B$, and since $P(\theta) \neq 0$, $P_1(\theta) \neq 0$ and $P_2(\theta) \neq 0$. Since $P_1 \in I_A$ and $P_1(\theta) \neq 0$, there must exist a polynomial $\hat{P}_1$ in the generator set $\bigcup_{q,\hat{q}\in Q} S_{A,q,\hat{q}}$ of I_A such that $\hat{P}_1(\theta) \neq 0$. In particular, $\hat{P}_1 \in S_{A,q,\hat{q}}$ for some $q, \hat{q} \in Q$. By applying Lemma 6 to $\chi_q^f, \phi_{q,\hat{q}}^f$ it follows that $\chi_q(\theta)(z) = \chi_q^f(\theta, z)$ and $\phi_{q,\hat{q}}^f(\theta, z) = \phi_{q,\hat{q}}(\theta)(z)$ are co-prime, hence for $q_1 = q$, $q_2 = \hat{q}$, condition **(A)** of Theorem 1 holds. Similarly, since $P_2 \in I_B$ and $P_2(\theta) \neq 0$, it follows that there exists a polynomial $\hat{P}_2$ such that $\hat{P}_2(\theta) \neq 0$ and for some $q_2, q_3 \in Q$, $\hat{P}_2 \in S_{B,q_2,q_3}$. The latter means that $\hat{P}_2 = P_3 \cdot n_{q_2}^{f,n_u+n_y}(n_{q_3}^{f,n_y} n_{q_2}^{f,n_u+n_y} - n_{q_2}^{f,n_y} n_{q_3}^{n_y+n_u})$ for some $P_3 \in S'_{B,q_2,q_3}$. From $\hat{P}_2(\theta) \neq 0$ it then follows that $P_3(\theta) \neq 0$ and $n_{q_3}^{f,n_y}(\theta) n_{q_2}^{f,n_u+n_y}(\theta) - n_{q_2}^{f,n_y}(\theta) n_{q_3}^{n_y+n_u}(\theta) \neq 0$ and $n_{q_2}^{f,n_u+n_y}(\theta) \neq 0$. From Lemma 6 it follows that $P_3(\theta) \neq 0$ implies that $v_{q_3}^f(\theta)(z)$ and $\chi_{q_2}^f(\theta)(z)$ are co-prime, hence condition **(B)** of Theorem 1 holds. That is, $\Theta(\theta)$ is strongly minimal. □

Example 5. Consider the parametrization from Example 1 and let us apply Procedure 1 to it. We re-parameterize this parametrization as follows: define $\phi : \mathbb{R}^2 \ni (\zeta_1, \zeta_2)^T \mapsto (\zeta_1, \zeta_1 + \zeta_2, 0, 0, \dots, 0, 1)^T \in \mathbb{R}^{20}$ and define the parametrization $\mathbf{\Pi}^{\bar{\text{SARX}}} : \mathbb{R}^2 \ni (\zeta_1, \zeta_2)^T \mapsto \mathbf{\Pi}^{\text{SARX}}(\phi(\zeta_1, \zeta_2))$, where $\mathbf{\Pi}^{\text{SARX}}$ is the parametrization from Example 1. It then follows that for $\underline{\zeta} = (\zeta_1, \zeta_2)$ are of the form $\mathbf{\Pi}^{\hat{\text{SARX}}}(\underline{\zeta}) = \{n_q(\underline{\zeta})\}_{q\in Q}$, $Q = \{(1,1), (1,2), (2,1), (2,2)\}$,

$$n_{(1,2)}(\underline{\zeta}) = \begin{bmatrix}0.001\zeta_1 & 0.001\zeta_1 + 0.001\zeta_2 & 0.0 & 16.0\end{bmatrix}^T,\ n_{(1,1)}(\underline{\zeta}) = \begin{bmatrix}0.001\zeta_1 & 0.001\zeta_1 + 0.001\zeta_2 & 0.0 & 4.0\end{bmatrix}^T$$
$$n_{(2,1)}(\underline{\zeta}) = \begin{bmatrix}0.001\zeta_1 & 0.001\zeta_1 + 0.001\zeta_2 & 0.0 & 4.0\end{bmatrix}^T,\ n_{(2,2)}(\underline{\zeta}) = \begin{bmatrix}0.001\zeta_1 & 0.001\zeta_1 + 0.001\zeta_2 & 0.0 & 16.0\end{bmatrix}^T$$

and the polynomials $\chi_q(\underline{\zeta}), v_q(\underline{\zeta}), \phi_{q_1,q_2}(\underline{\zeta})$,

$$\chi_q(\underline{\zeta})(z) = -0.001\zeta_1 z + z^2 - 0.001\zeta_1 - 0.001\zeta_2,\ v_q(\underline{\zeta})(z) = 0.001\zeta_1 z + 0.001\zeta_1 + 0.001\zeta_2,\ q \in Q$$
$$\phi_{q_1,q_2}(\underline{\zeta})(z) = 16z,\ (q_1, q_2) \in Z_1,\ \phi_{q_1,q_2}(\underline{\zeta})(z) = 4z,\ (q_1, q_2) \in Z_2,$$
$$Z_1 = \{((1,2),(2,2)), ((1,1),(2,1)), ((1,2),(2,2)), ((2,1),(1,1)), ((2,2),(1,2))\}$$
$$Z_2 = \{((1,1),(1,2)), ((1,1),(2,2)), ((1,2),(1,1)), ((2,1),(1,2)), ((2,1),(2,2)), ((2,2),(1,1))\}.$$

Then $S_{A,q,\hat{q}}$, $S'_{B,q,\hat{q}}$ and $S_{B,q,\hat{q}}$, $q, \hat{q} \in Q$ can all be calculated using[33]. In this case $S_{A,q,\hat{q}} = \{\zeta_1 + \zeta_2\}$, $q, \hat{q} \in Q$, $q \neq \hat{q}$ and $S_{B,q,\hat{q}} = \emptyset$, $(q, \hat{q}) \in Z_1$, $S_{B,q,\hat{q}} = \zeta_1^3 + 3\zeta_1^2\zeta_2 + 3\zeta_1\zeta_2^2 + \zeta_2^3$, $(q, \hat{q}) \in Z_2$. It then follows that I_A is generated by the Gröbner basis $\{\zeta_1 + \zeta_2\}$ and I_B is generated by the Gröbner basis $\{\zeta_1^3 + 3\zeta_1^2\zeta_2 + 3\zeta_1\zeta_2^2 + \zeta_2^3\}$. Hence, $S = \{\zeta_1^4 + 4\zeta_1^3\zeta_2 + 6\zeta_1^2\zeta_2^2 + 4\zeta_1\zeta_2^3 + \zeta_2^4\}$ It then follows that

$$\hat{\Theta} = \{\underline{z} = (\zeta_1, \zeta_2) \mid \zeta_1^4 + 4\zeta_1^3\zeta_2 + 6\zeta_1^2\zeta_2^2 + 4\zeta_1\zeta_2^3 + \zeta_2^4 \neq 0\} = \{\underline{z} = (\zeta_1, \zeta_2) \mid \zeta_1 \neq -\zeta_2\}$$

and the parametrization $\mathbf{\Pi}^{\bar{\text{SARX}}}|_{\hat{\Theta}} : \hat{\Theta} \ni \underline{\zeta} \mapsto \mathbf{\Pi}^{\bar{\text{SARX}}}(\underline{\zeta})$ is strongly minimal and identifiable.

We can also apply Procedure 1 to more complicated parametrizations, but the expressions for corresponding polynomials and the Gröbner bases are more involved. For example, define $\tilde{\phi} : \mathbb{R}^2 \ni (\zeta_1, \zeta_2)^T \mapsto (\zeta_1, \zeta_1 + \zeta_2, \kappa_3, \kappa_4, \dots, \kappa_{18}, \kappa_{19} + 1)^T \in \mathbb{R}^{20}$, $\kappa_i = 1$, if i is even, and $\kappa_i = (\zeta_1 - \zeta_2)$ if i is odd, and define the parametrization $\mathbf{\Pi}^{\tilde{\text{SARX}}} : \mathbb{R}^2 \ni (\zeta_1, \zeta_2)^T \mapsto \mathbf{\Pi}^{\text{SARX}}(\tilde{\phi}(\zeta_1, \zeta_2))$, where

$\mathbf{\Pi}^{\text{SARX}}$ is the parametrization from Example 1. The expressions for the Gröbner basis $S_{A,q,\hat{q}}$, $S_{B,q,\hat{q}}$, $q, \hat{q} \in Q$ is lengthy. However, using the implementation of [34] of Procedure 1, we obtain that I_A is generated by the polynomial 1, i.e., $I_A = \mathbb{R}[X_1, X_2]$, and I_B is generated by $\{\zeta_1^2, \zeta_1\zeta_2, \zeta_2^2\}$ and hence $I_A I_B = I_B$ and thus $S = \{\zeta_1^2, \zeta_1\zeta_2, \zeta_2^2\}$ and

$$\tilde{\Theta} = \{\underline{z} = (\zeta_1, \zeta_2) \mid \zeta_1^2 \neq 0, \text{ or } \zeta_1\zeta_2 \neq 0 \text{ or } \zeta_2^2 \neq 0\} = \{\underline{z} = (\zeta_1, \zeta_2) \mid \zeta_1 \neq 0 \text{ or } \zeta_2 \neq 0\}$$

and the parametrization $\mathbf{\Pi}^{\tilde{\text{SARX}}}|_{\tilde{\Theta}} : \tilde{\Theta} \ni \underline{\zeta} \mapsto \mathbf{\Pi}^{\tilde{\text{SARX}}}(\underline{\zeta})$ is strongly minimal and identifiable.

Remark 6 (Computional complexity)**.** Procedure 1 relies on computing Gröbner bases, and it is known that the computational complexity of the latter can be high. Hence, computational complexity of Procedure 1 might be an issue for applications. However, even for linear systems, identifiability analysis relies on symbolic algorithms, in particular, on algorithms based on calculation of Gröbner bases, and there the same problem arises [35]. For this reason, a detailed study of computational complexity of Procedure 1 cannot be handled within this paper.

3.3 | On the genericity of minimality and identifiability

In the previous sections we have established that strong minimality is sufficient for minimality and that it is also sufficient for identifiability. However, we have also demonstrated that for some minimal SARX systems, strong minimality does not hold. Hence, one may wonder how typical strong minimality is.

Below we will show that strong minimality is a generic property, i.e. it holds for almost all SARX systems, if $|Q| > 1$. This also means that identifiability is a generic property. In other words, strong minimality occurs very frequently. In order to formalize these results, we need the following terminology.

Definition 13 (Generic set)**.** A subset G of $\Theta \subset \mathbb{R}^d$ is *generic*, if G is non-empty and there exists a non-zero polynomial $P(X_1, \dots, X_d)$ in d variables such that $G = \{\theta \in \Theta \mid P(\theta) \neq 0\}$.

That is, a generic subset of Θ is a non-empty subset whose complement in Θ satisfies a polynomial equation.

Definition 14 (Generic identifiability and minimality of SARX parametrization)**.** The parametrization $\mathbf{\Pi}^{\text{SARX}}$ is said to be *generically identifiable* if there exists a generic subset G of Θ, such that the parametrization $\mathbf{\Pi}^{\text{SARX}}|_G : G \ni \theta \mapsto \mathbf{\Pi}^{\text{SARX}}(\theta)$ is identifiable. Similarly, $\mathbf{\Pi}^{\text{SARX}}$ is *generically minimal* (respectively *generically strongly minimal*), if there exists a generic subset G of Θ, such that the parametrization $\mathbf{\Pi}^{\text{SARX}}{}_{|G} : G \ni \theta \mapsto \mathbf{\Pi}^{\text{SARX}}(\theta)$ is minimal (respectively strongly minimal).

Intuitively, if a property is generic for a parametrization, then every member of the parametrization can be approximated with arbitrary accuracy by another member which has this property. Another interpretation is that if we randomly generate parameters, then the property will hold for the obtained random parametrization with probability one.

Example 6. Consider the parametrization $\mathbf{\Pi}^{\text{SARX}}$ from Example 2. The set G from Example 2 is generic. Hence, since the parametrization $\mathbf{\Pi}^{\text{SARX}}|_G$ is strongly minimal and identifiable, the parametrization $\mathbf{\Pi}^{\text{SARX}}$ is generically strongly minimal, generically minimal, and generically identifiable.

Theorem 3 (Generic minimality)**.** If $|Q| > 1$, $\mathbf{\Pi}^{\text{SARX}}$ is a polynomial parametrization and $\mathbf{\Pi}^{\text{SARX}}$ contains a strongly minimal SARX system, (i.e. for some $\theta \in \Theta$, $\mathbf{\Pi}^{\text{SARX}}(\theta)$ is strongly minimal), then $\mathbf{\Pi}^{\text{SARX}}$ is *generically strongly minimal*.

Proof. See Appendix D. □

Notice that Theorem 2 implies the following corollary result.

Corollary 2. Consider the SISO case, i.e. $p = m = 1$. If a SARX parametrization is injective, polynomial, and generically strongly minimal, then it is generically identifiable.

Proof. If $\mathbf{\Pi}^{\text{SARX}}$ is generically strongly minimal, then there exists a generic set $G \subseteq \Theta$ such that the parametrization $\mathbf{\Pi}^{\text{SARX}}|_G : G \ni \theta \mapsto \mathbf{\Pi}^{\text{SARX}}(\theta)$ is strongly minimal. Hence, by Theorem 2, $\mathbf{\Pi}^{\text{SARX}}|_G$ is identifiable. This means that $\mathbf{\Pi}^{\text{SARX}}$ is generically strongly identifiable. □

Corollary 2 and Theorem 3 yield the following result.

Corollary 3. Assume that $p = m = 1$. If a SISO SARX parametrization is polynomial and it contains a strongly minimal element, then it is generically identifiable.

The trivial SISO SARX parametrization $\mathbf{\Pi}_{\text{triv}}$ is the SARX parametrization defined as follows: $\Theta = \mathbb{R}^{|Q|(n_u+n_y)}$ and $\mathbf{\Pi}_{\text{triv}}$ is the identity map. From Corollaries 2 and 3, we obtain that

Corollary 4. The trivial parametrization is generically minimal and in the SISO case, it is generically identifiable.

Proof. By Remark 3, there exists a strongly minimal SARX system, i.e. $\mathbf{\Pi}_{\text{triv}}$ contains a strongly minimal element. Moreover, $\mathbf{\Pi}_{\text{triv}}$ is clearly injective and polynomial. We can therefore apply Corollary 3 to conclude. □

Example 7. From Example 4 it follows that the parametrization defined in Example 1 contains a strongly minimal element, hence it is generically strongly minimal and generically identifiable.

4 | CONCLUDING REMARKS

In this paper we have studied minimality and identifiability of linear SARX systems. Formal definitions of these two concepts have been introduced and discussed with respect to their standard characterizations for ARX systems. Sufficient and necessary conditions have been derived for minimality and identifiability of SARX systems. In particular, it has been shown that minimal SARX parametrizations are also identifiable.

ORCID

How to cite this article: M. Petreczky, L. Bako, S. Lecoeuche,and K.M.D. Motchon (2019), Minimality and identifiability of discrete-time SARX systems, *Int J Robust Nonlinear Control*, *2019;00:1–17*.

APPENDIX

A ON MINIMALITY AND INDENTIFIABILITY OF LSSs

In this section we recall from[24] the notion of identifiability for LSSs, and its relationship with minimality. In addition, we recall from[24] sufficient and necessary conditions for identifiability of LSSs. We start with defining the notion of parametrization of LSSs. To this end, we need the following notation.

Notation 4. Denote by $\Sigma(n, m, p, Q)$ the set of all LSSs with state-space dimension n, input space $\mathbb{R}^m$, output space $\mathbb{R}^p$, and set of discrete modes Q.

Definition 15 (Parametrization of LSSs)**.** Assume that $\Theta \subseteq \mathbb{R}^d$ is the set of parameters. A *parametrization of LSSs belonging to* $\Sigma(n, m, p, Q)$ is a map $\mathbf{\Pi}^{\text{LSS}} : \Theta \to \Sigma(n, m, p, Q)$. For each $\theta \in \Theta$, we denote $\mathbf{\Pi}^{\text{LSS}}(\theta)$ by

$$\Sigma(\theta) = (n, Q, \{(A_q(\theta), B_q(\theta), C_q(\theta)) \mid q \in Q\}, x_0(\theta)).$$

Next, we define structural identifiability of parametrizations.

Definition 16 (Structural identifiability of LSSs parametrizations)**.** A parametrization $\mathbf{\Pi}^{\text{LSS}} : \Theta \to \Sigma(n, m, p, Q)$ is *structurally identifiable*, if for any two distinct parameters $\theta_1 \neq \theta_2$, the input-output maps of the corresponding LSSs $\mathbf{\Pi}^{\text{LSS}}(\theta_1) = \Sigma(\theta_1)$ and $\mathbf{\Pi}^{\text{LSS}}(\theta_2) = \Sigma(\theta_2)$ are different, i.e. $y_{\Sigma(\theta_1)} \neq y_{\Sigma(\theta_2)}$.

The condition $y_{\Sigma(\theta_1)} \neq y_{\Sigma(\theta_2)}$ means that there exists a sequence of inputs and discrete modes $w \in \mathcal{U}^{*+}$, such that $y_{\Sigma(\theta_1)}(w) \neq y_{\Sigma(\theta_2)}(w)$. In other words, a parametrization is structurally identifiable, if for every two distinct parameters there exists an input and a switching signal, such that the corresponding outputs are different. This means that every parameter can be uniquely reconstructed from the input-output map of the corresponding LSS.

It is an intuitive fact that minimality is somehow a necessary condition for structural identifiability[24]. If we allow non-minimal parametrizations, then either the parametrization is not identifiable, or all the parameters occur in the minimal part of the systems, and hence we can replace the parametrization by a minimal one. For this reason, *we will restrict attention to minimal LSSs when*

studying identifiability. In turn, structural minimality of parametrizations allow a simple characterization of identifiability, due to the fact that minimal LSSs are unique up to isomorphism.

Definition 17 (Structural minimality of LSSs parametrization)**.** The parametrization $\mathbf{\Pi}^{\text{LSS}}$ is called structurally minimal, if for any parameter value $\theta \in \Theta$, $\Sigma(\theta)$ is a minimal LSS realization of its input-output map $y_{\Sigma(\theta)}$.

Hence, by Petreczky *et al*[29, Theorem 1], $\mathbf{\Pi}^{\text{LSS}}$ is structurally minimal if and only if for every parameter $\theta \in \Theta$, $\Sigma(\theta)$ is span-reachable and observable. Since the latter concepts admit rank characterizations, structural minimality is a property that can be checked algorithmically.

Theorem 4 below recalls a necessary and sufficient condition for structural identifiability of a *structurally minimal* parametrization established by Petreczky *et al*[24].

Theorem 4 (Identifiability of structural minimal parametrizations)**.** A structurally minimal parametrization $\mathbf{\Pi}^{\text{LSS}}$ is structurally identifiable, if and only if for any two distinct parameter values $\theta_1, \theta_2 \in \Theta$, $\theta_1 \neq \theta_2$, there exists no LSS isomorphism $S : \Sigma(\theta_1) \to \Sigma(\theta_2)$.

The following important corollary which is an immediate consequence of the Theorem 4 can be useful for checking identifiability of parametrizations.

Corollary 5. Assume that $\mathbf{\Pi}^{\text{LSS}}$ is a structurally minimal parametrization, and for each two parameter values $\theta_1, \theta_2 \in \Theta$, $\Sigma(\theta_1) = \Sigma(\theta_2)$ implies that $\theta_1 = \theta_2$. Here, equality of two systems means equality of the matrices of the linear subsystems for each discrete state $q \in Q$ and equality of the initial state. Then $\mathbf{\Pi}^{\text{LSS}}$ is structurally identifiable if and only if the assumption that $S : \Sigma(\theta_1) \to \Sigma(\theta_2)$ is an LSS isomorphism implies that S is the identity matrix.

B PROOF OF THEOREM 1

For the proof of Theorem 1, we will need a number of auxiliary results. Below, we consider $\mathbf{S} = \{h_q\}_{q \in Q}$. We denote by A_q the corresponding matrix of the LSS $\Sigma_{\mathbf{S}}$. We will denote by e_i the ith standard basis vector of $\mathbb{R}^{n_y+n_u}$.

Lemma 8. Let $X_1 = \text{Span}\{e_1, \dots, e_{n_y}\}$. It then follows that for any $q \in Q$,

1. $A_q e_j = h_q^j e_1 + e_{j+1}$ for all $j = 1, \dots, n_y + n_u$, $j \neq n_y$ and $A_q e_{n_y} = h_q^{n_y} e_1$.
2. The space X_1 is A_q invariant and $\text{Span}\{A_q^i e_1 \mid i = 0, \dots, n_y - 1\} = X_1$
3. $A_q^j e_{n_y+1} \in X_1 + e_{n_y+j+1}$, $j = 0, \dots, n_u - 1$, and $A_q^{n_u} e_{n_y+1} \in X_1$.
4. For any $\hat{q} \in Q$, $\psi_{\hat{q},q,j}(A_{\hat{q}})e_1 = A_q^j e_1$, where $\psi_{\hat{q},q,0}(z) = 1$ and
$$\psi_{\hat{q},q,j+1}(z) = z\psi_{\hat{q},q,j}(z) + (h_q - h_{\hat{q}})\psi_{\hat{q},q,j}(A_{\hat{q}})e_1,$$
where the vectors $d_j \in \mathbb{R}^{n_y+n_u}$ are defined as follows: $d_0 = e_1$ and if $d_j = (d_{j,1}, \dots, d_{j,n_y}, 0, \dots, 0)^T$, $d_{j,1}, \dots, d_{j,n_y} \in \mathbb{R}$, then
$$d_{j+1} = (h_q d_j, d_{j,1}, \dots, d_{j,n_y-1}, 0, \dots, 0)^T$$

Proof. Part 1 follows by a simple computation. Part 2 follows from Part 1 by taking into account that $A_q e_{n_y} \in \text{Span}\{e_1\}$. Part 3 follows from the definition of A_q by induction. Indeed, $A_q e_{n_y+1} = h_q^{n_y+1} e_1 + e_{n_y+2} \in X_1 + e_{n_y+2}$ and if $A_q^j e_{n_y+1} \in X_1 + e_{n_y+j+1}$, then $A_q^{j+1} e_{n_y+1} \in X_1 + A_q e_{n_y+j+1} \subseteq X_1 + h_q^{n_y+j+1} e_1 + e_{n_y+j+2} \subseteq X_1 + e_{n_y+j+2}$.

Finally, $\psi_{\hat{q},q,j}(A_{\hat{q}})e_1 = A_q^j e_1$ we will prove by induction. For $j = 0$, the equality is trivial. Notice that $A_q e_i = h_q^i e_1 + e_{i+1} = A_{\hat{q}} e_i + (h_q^i - h_{\hat{q}}^i)e_1$ for all $i = 1, \dots, n_y - 1$, and $A_q e_{n_y} = h_q^{n_y} e_1 = A_{\hat{q}} e_{n_y} + (h_q^{n_y} - h_{\hat{q}}^{n_y})e_1$. Hence, for any $x = \sum_{i=1}^{n_y} x_i e_i$, $A_q x = A_{\hat{q}} x + \sum_{i=1}^{n_y} x_i (h_q^i - h_{\hat{q}}^i)e_1 = A_{\hat{q}} x + ((h_q - h_{\hat{q}})x)e_1$. Hence, if $\psi_{\hat{q},q,j}(A_{\hat{q}})e_1 = A_q^j e_1$ holds, then

$$A_q^{j+1} e_1 = A_q \psi_{\hat{q},q,j}(A_{\hat{q}})e_1 = A_{\hat{q}} \psi_{\hat{q},q,j}(A_{\hat{q}})e_1 + ((h_q - h_{\hat{q}})\psi_{\hat{q},q,j}(A_{\hat{q}})e_1)e_1. \tag{B1}$$

Finally, notice that $d_j = A_q^j e_j$ for all $j = 0, \dots, n_y$. Indeed, $d_0 = e_1$ and if $d_j = \sum_{i=1}^{n_y} d_{j,i} e_i$, then $A_q d_j = (\sum_{i=1}^{n_y} d_{j,i} h_q^i)e_1 + \sum_{i=2}^{n_y} d_{j,i-1} e_i = d_{j+1}$. Hence, by replacing $\psi_{\hat{q},q,j}(A_{\hat{q}})e_1 = A_q^j e_1$ by d_j in (B1), we obtain that

$$A_q^{j+1} e_1 = \psi_{\hat{q},q,j+1}(A_{\hat{q}})e_1.$$

Hence, by induction we get the last statement of the lemma. □

Lemma 9. If $h_q^{n_y+n_u} \neq 0$, then $\{e_{n_y}^T A_q^j \mid j = 0, \dots, n_y + n_u - 1\}$ spans $\mathbb{R}^{1\times(n_y+n_u)}$. Moreover, $e_i^T = e_{n_y}^T A_q^{n_y-i}$, $i = 1, \dots, n_y$, and $e_{n_y+j}^T = e_{n_y}^T \chi_q(A_q)\gamma_{j,q}(A_q)$, $j = 1, \dots, n_u$, where $\chi_q(z) = z^{n_y} - \sum_{j=1}^{n_y} h_q^j z^{n_y-j}$ and the polynomial $\gamma_{j,q}(z)$, $j = 1, \dots, n_u$ is defined recursively as follows:

$$\gamma_{1,q}(z) = \frac{1}{h_q^{n_u+n_y}} z^{n_u-1}$$

and

$$\gamma_{i,q}(z) = \frac{1}{h_q^{n_u+n_y}}(z^{n_u-i} - \sum_{j=1}^{i-1} \gamma_{j,q}(z) h_q^{n_y+n_u-i+j}).$$

Proof. In this proof we will view A_q as a linear map $x^T \mapsto x^T A_q$, defined on the space of row vectors $x^T \in \mathbb{R}^{1\times(n_y+n_u)}$. From the structure of A_q it then follows that $e_j^T A_q = e_{j-1}^T$, for $j = \{2, \dots, n_y\} \cup \{n_y + 2, \dots, n_y + n_u\}$. Hence, $e_{n_y}^T A_q^j$, $j = 0, \dots, n_y - 1$ spans $X_1 = \text{Span}\{e_1^T, \dots, e_{n_y}^T\}$. Notice that $e_1^T A_q = \sum_{i=1}^{n_y+n_u-1} h_q^i e_i^T + h_q^{n_u+n_y} e_{n_y+n_u}^T$. Hence

$$x^T = \sum_{i=n_y+1}^{n_y+n_u} h_q^i e_i^T = e_1^T A_q - \sum_{i=1}^{n_y} h_q^i e_i^T \tag{B2}$$

belongs to the linear span of $e_{n_y}^T A_q^j$, $j = 0, \dots, n_y$.

We proceed to prove that $x^T A_q^j$, $j = 0, \dots, n_u - 1$ span $X_2 = \text{Span}\{e_{n_y+1}^T, \dots e_{n_y+n_u}^T\}$. From this the first statement of the lemma follows. Notice that $e_{n_y+1}^T A_q = 0$ and $e_{n_y+j}^T A_q = e_{n_y+j-1}^T$ for all $j = 2, \dots, n_u$. Hence,

$$x^T A_q^j = \sum_{i=1}^{n_u-j} h_q^{n_y+i+j} e_{n_y+i}^T. \tag{B3}$$

From (B3) and $h_q^{n_y+n_u} \neq 0$ it then follows that

$$e_{n_y+1}^T = \frac{1}{h_q^{n_u+n_y}} x^T A_q^{n_u-1} \tag{B4}$$

and if $e_{n_y+1}^T, \dots, e_{n_y+j}^T$ have already been obtained from the linear combinations of $x^T A_q^i$, $i = n_u - j, \dots, n_u - 1$, then

$$e_{n_y+j+1}^T = \frac{1}{h_q^{n_u+n_y}}(x^T A_q^{n_u-j-1} - \sum_{i=1}^{j} h_q^{n_y+n_u-j-1+i} e_{n_y+i}). \tag{B5}$$

Hence, $x^T A_q^j$, $j = 0, \dots, n_u - 1$ span X_2.

Finally, the statement $e_j^T = e_{n_y}^T A_q^{n_y-j}$, $j = 1, \dots, n_y$ follows from the definition of A_q. The statement that $e_{n_y+j}^T = e_{n_y}^T \chi(A_q)\gamma_{j,q}(A_q)$, $j = 1, \dots, n_u$ can be shown as follows. From (B2) it follows that $x^T = e_{n_y}^T \chi_q(A_q)$. From (B4) and (B5) it follows that $e_{n_y+j}^T = x^T \gamma_{j,q}(A_q)$ for all $j = 1, \dots, n_u$. Combining the above statements implies the second statement of the lemma. □

Lemma 10. Assume that $h_q^{n_u+n_y} \neq 0$. The characteristic polynomial of A_q coincides with its minimal polynomial and it equals

$$\chi_{A_q}(z) = z^{n_u}(z^{n_y} - \sum_{i=1}^{n_y} h_q^i z^{n_y-i}).$$

Proof. From Lemma 9 it follows that $\mathbb{R}^{1\times(n_u+n_y)}$ is a cyclic subspace with respect to the linear operator $\hat{A}_q : x^T \mapsto x^T A_q$ [2]. By Theorem 4-Chapter VII of the book of Gantmacher[36], it then follows that the minimal polynomial of the linear operator $\hat{A}_q$ equals its characteristic polynomial and it is of degree $n_y + n_u$. Note that in the standard basis $e_1^T, \dots, e_{n_u+n_y}^T$, the basis of the linear operator $\hat{A}_q$ is A_q^T. Hence, the minimal polynomial and characteristic polynomial of A_q^T coincide. But these polynomials are the same for the matrices A_q and A_q^T.

Moreover, from Lemma 9, it also follows that $e_{n_y}^T$ is the generating element of the cyclic space $\mathbb{R}^{1\times(n_y+n_u)}$. Hence, by Subsection 4.1-Chapter VII of the book of Gantmacher[36], a polynomial $\psi(z)$ is a minimal polynomial of $\hat{A}_q$, if $\psi(\hat{A}_q)e_{n_y}^T = e_{n_y}^T \psi(A_q) = 0$

[2]The definition of cyclic subspaces can be found in Section 4-Chapter VII of the book of Gantmacher[36]

and it has the smallest possible degree. By the discussion above, the degree of the minimal polynomial of $\hat{A}_q$ must be $n_y + n_u$. Hence, the minimal polynomial of $\hat{A}_q$ is the unique monic polynomial $\psi(z)$ of degree $n_y + n_u$, such that $e_{n_y}^T \psi(A_q) = 0$.

If we show that $e_{n_y}^T \chi_{A_q}(A_q) = 0$, then the statement of the lemma follows. To this end, notice that if $y^T \in X_2 = \text{Span}\{e_{n_y+1}^T, \dots, e_{n_y+n_u}^T\}$, then $y^T A_q^{n_u} = 0$. In addition, $e_{n_y}^T A_q^{n_y} = e_1^T A_q = \sum_{i=1}^{n_y} h_q^i e_i^T + x^T$, where $x^T = \sum_{i=n_y+1}^{n_u} h_q^i e_i^T \in X_2$. Hence, by taking into account the remark above and that $e_i^T = e_{n_y} A_q^{n_y - i}$, $i = 1, \dots, n_y$,

$$e_{n_y}^T A_q^{n_y+n_u} = \sum_{i=1}^{n_y} h_q^i e_i^T A_q^{n_u} + x^T A_q^{n_u} = \sum_{i=1}^{n_y} h_q^i e_{n_y}^T A_q^{n_u+n_y-i}.$$

The latter is exactly equivalent to $e_{n_y}^T \chi_{A_q}(A_q) = 0$. □

Now, to complete the proof of Theorem 1, we will show that if Part **(A)** holds, then $\Sigma_{\mathbf{S}}$ is reachable, and if Part **(B)** holds, then $\Sigma_{\mathbf{S}}$ is observable.

Proof of Part (A)

We will show that if the conditions of **(A)** hold, then $(A_{q_0}, A_{q_1}^{n_u} B_{q_1})$ is a controllable pair. By Sun and Ge[28] it then follows that the LSS $\Sigma_{\mathbf{S}}$ is reachable. From Lemma 8 it follows that $A_{q_1}^j B_{q_1} = A_{q_1}^j e_{n_y+1} = \sum_{r=1}^{j} h_{q_1}^{n_y+r} A_{q_1}^{j-r} e_1 + e_{n_y+j+1}$ for $j = 1, \dots, n_u - 1$ and hence

$$A_{q_1}^{n_u} B_{q_1} = \sum_{j=1}^{n_u} h_{q_1}^{n_y+j} A_{q_1}^{n_u-j} e_1.$$

From Lemma 8 it also follows that $\psi_{q_0,q_1,j}(A_{q_0}) e_1 = A_{q_1}^j e_1$ and hence the polynomial $\phi_{q_0,q_1}(z)$ satisfies

$$A_{q_1}^{n_u} B_q = \phi_{q_0,q_1}(A_{q_0}) e_1.$$

From Lemma 8, it follows that $\phi_{q_0,q_1}(A_{q_0}) e_1 \in X_1$ and X_1 is A_{q_1} invariant, where $X = \text{Span}\{e_1, \dots, e_{n_y}\}$. In addition, from the construction of A_{q_0} it follows that with respect to the basis $e_1, \dots, e_{n_y}$, the matrix representation of the restriction of A_{q_0} to X_1 is of the form

$$\hat{A}_{q_0} = \begin{bmatrix} \begin{bmatrix} h_{q_0}^1 & \dots & h_{q_0}^{n_y-1} \end{bmatrix} & h_{q_0}^{n_y} \\ I_{n_y-1} & \mathbf{O}_{(n_y-1)\times 1} \end{bmatrix}.$$

The above matrix is in companion form and it is known that its characteristic polynomial equals its minimal polynomial and it equals $\chi_{q_0}(z)$. That is, $\chi_{q_0}(z)$ is the minimal polynomial of the linear operator A_{q_0} restricted to X_1. Moreover, from Lemma 8, it follows that $A_{q_0}^j e_1$, $j = 0, \dots, n_y - 1$ generate the space X_1, i.e. X_1 is a cyclic subspace w.r.t. to A_{q_0}. Then by Subsection 4.1–Chapter VII of the book of Gantmacher[36], $\chi_{q_0}(z)$ is a minimal polynomial of e_1 with respect to A_{q_0}, i.e. $\chi_{q_0}(A_{q_0}) e_1 = 0$ and $\chi_{q_0}(z)$ has the smallest degree among all the polynomials $\psi(z)$ such that $\psi(A_{q_0}) e_1 = 0$.

Suppose now that $\chi_{q_0}(z)$ and $\phi_{q_0,q_1}(z)$ are coprime, but $(A_{q_0}, A_{q_1}^{n_u} B_{q_1}) = (A_{q_0}, \phi_q(A_{q_0}) e_1)$ is not a controllable pair. Then the vectors $A_{q_0}^j x$, $j = 0, \dots, n_y - 1$, $x = \phi_{q_0,q_1}(A_{q_0}) e_1$ are linearly dependent, i.e. there exists a non-zero polynomial $\kappa(z)$ of degree at most $n_y - 1$ such that $\kappa(A_{q_0}) x = 0$. By substituting $x = \phi_{q_0,q_1}(A_{q_0}) e_1$, we get $\kappa(A_{q_0}) \phi_{q_0,q_1}(A_{q_0}) e_1 = 0$. That is, for the polynomial $\phi(z) = \kappa(z) \phi_{q_0,q_1}(z)$, $\phi(A_{q_0}) e_1 = 0$. This implies by Gantmacher[36] that the minimal polynomial $\chi_{q_0}(z)$ divides $\phi(z) = \kappa(z) \phi_{q_0,q_1}(z)$. Since $\chi_{q_0}(z)$ and $\phi_{q_0,q_1}(z)$ are co-prime, then this is possible only if $\chi_{q_0}(z)$ divides $\kappa(z)$. But the degree of $\kappa(z)$ is strictly smaller than the degree of $\chi_{q_0}(z)$, hence $\kappa(z)$ cannot be divisible by $\chi_{q_0}(z)$. We arrived to a contradiction. That is, we can conclude that $(A_{q_0}, A_{q_1}^{n_u} B_{q_1})$ is a controllable pair.

Proof of Part (B)

We will show that (C_{q_3}, A_{q_2}) is an observable pair. By Sun and Ge[28], this is sufficient for observability of $\Sigma_{\mathbf{S}}$.

To this end, using the notation of Lemma 9 define the polynomial

$$\hat{\psi}(z) = \upsilon_{q_3}(z) + \sum_{j=1}^{n_u} h_{q_3}^{n_y+j} \gamma_{j,q_2}(z) \chi_{q_2}(z).$$

Then from Lemma 9 it follows that $C_{q_3} = e_{n_y}^T \hat{\psi}(A_{q_2})$. Assume that (C_{q_3}, A_{q_2}) is not an observable pair. Then $C_{q_3} A_{q_2}^j$, $j = 0, \dots, n_y - 1$ are linearly dependent. Hence, there exists a polynomial $\kappa(z)$ of degree less than n_y, such that $C_{q_3} \kappa(A_{q_2}) = 0$. Hence, we obtain that $e_{n_y}^T \hat{\psi}(A_{q_2}) \kappa(A_{q_2}) = 0$. In other words, the polynomial $P(z) = \hat{\psi}(z) \kappa(z)$ is an annihilating polynomial with respect to the operator $\hat{A}_{q_2} : x \mapsto x A_{q_2}$ [3] of $e_{n_y}^T$. Since by Lemma 9 $e_{n_y}^T A_{q_2}^j$, $j = 0, \dots, n_y + n_u$ generate the whole space, it then follows

[3] For the definition, see the book of Gantmacher[36].

that $P(z)$ is the annihilating polynomial of the whole space, i.e. $P(A_{q_2}) = 0$. It then follows that $P(z)$ is divisible by the minimal polynomial of $\hat{A}_{q_2}$ which coincides with that of A_{q_2}. From Lemma 10 it follows that the minimal polynomial of A_{q_2} is $z^{n_u}\chi_{q_2}(z)$.

We will argue that if the conditions of Part **(B)** hold, then $\hat{\psi}(z)$ and $z^{n_u}\chi_{q_2}(z)$ are co-prime. Indeed, if $\hat{\psi}(z)$ and $z^{n_u}\chi_{q_2}(z)$ are not co-prime, then there exists an irreducible polynomial $q(z)$ which divides both $\hat{\psi}(z)$ and $z^{n_u}\chi_{q_2}(z)$. If $q(z)$ is an irreducible polynomial which divides $z^{n_u}\chi_{q_2}(z)$, then it either equals z or it divides $\chi_{q_2}(z)$. If $q(z) = z$ and it divides $\hat{\psi}(z)$, then 0 is a root of $\hat{\psi}(z)$, i.e. $\hat{\psi}(0) = 0$. Notice that by induction it follows that for $j = 1, \dots, n_u - 1$, $\gamma_{j,q_2}(0) = 0$ and $\gamma_{n_u,q_2}(0) = \frac{1}{h_{q_2}^{n_u+n_y}}$. Hence, from the definition of $\hat{\psi}(z)$ it follows that $\hat{\psi}(0) = h_{q_3}^{n_y} + \frac{h_{q_3}^{n_u+n_y}}{h_{q_2}^{n_u+n_y}}\chi_{q_2}(0) = h_{q_3}^{n_y} - \frac{h_{q_3}^{n_u+n_y}}{h_{q_2}^{n_u+n_y}}h_{q_2}^{n_y}$. Hence, $\hat{\psi}(0) = 0$ implies that $h_{q_3}^{n_y} = \frac{h_{q_3}^{n_u+n_y}}{h_{q_2}^{n_u+n_y}}h_{q_2}^{n_y}$, which contradicts to the condition of **(B)**. If $q(z)$ divides $\chi_{q_2}(z)$ and it divides $\hat{\psi}(z)$, then it divides $\upsilon_{q_3}(z) = \hat{\psi}(z) - (\sum_{i=1}^{n_u} h_{q_3}^{n_u+i}\gamma_{i,q_2}(z))\chi_{q_2}(z)$. But this contradicts to the assumption that $\psi_{q_3}(z)$ and $\chi_{q_2}(z)$ are co-prime.

Hence, by the discussion above, $\hat{\psi}(z)$ and $z^{n_u}\chi_{q_2}(z)$ are coprime, so if $z^{n_u}\chi_{q_2}(z)$ divides $P(z)$, it then must divide $\kappa(z)$. But the degree of $\kappa(z)$ is strictly smaller than that of $z^{n_u}\chi_{q_2}(z)$, hence $z^{n_u}\chi_{q_2}(z)$ cannot divide $\kappa(z)$. We arrived to a contradiction. Hence, (C_{q_3}, A_{q_2}) must be an observable pair.

C PROOF OF THEOREM 2

In order to prove Theorem 2 we will relate identifiability analysis of SARX systems to that of the associated LSSs (see Section A for the definition of a parametrization and identifiability of LSSs). This is possible due to the following corollary of Lemma 1.

Corollary 6. A SARX parametrization $\mathbf{\Pi}^{\text{SARX}}$ of the form (7) is identifiable, if and only if the LSS parametrization $\mathbf{\Pi}_{sw} : \Theta \ni \theta \mapsto \Sigma_{\mathbf{\Pi}^{\text{SARX}}(\theta)}$ is identifiable. Here, $\Sigma_{\mathbf{\Pi}^{\text{SARX}}(\theta)}$ is the LSS of the form (5)–(6) obtained from the SARX $\mathbf{\Pi}^{\text{SARX}}(\theta)$.

Proof of Corollary 6. Consider two SARX systems $\mathbf{S}_i = \{h_q^i\}_{q\in Q}$, $i = 1, 2$ of type (n_y, n_u). Notice that each $\mathbf{S}_i$, $i = 1, 2$, realizes the same input-output map as the associated LSSs $\Sigma_{\mathbf{S}_i}$.

Assume that the parametrization $\mathbf{\Pi}^{\text{SARX}}$ is identifiable, but $\mathbf{\Pi}_{sw}$ is not identifiable. Then there exist two parameters $\theta_1, \theta_2 \in \Theta$, $\theta_1 \neq \theta_2$, such that $\mathbf{\Pi}_{sw}(\theta_1)$ and $\mathbf{\Pi}_{sw}(\theta_2)$ realize the same input-output map. Since $\mathbf{\Pi}_{sw}(\theta_i) = \Sigma_{\mathbf{\Pi}^{\text{SARX}}(\theta_i)}$, $i = 1, 2$, by the remark above it follows that $\mathbf{\Pi}^{\text{SARX}}(\theta_1)$ and $\mathbf{\Pi}^{\text{SARX}}(\theta_2)$ are equivalent. This contradicts the identifiability of $\mathbf{\Pi}^{\text{SARX}}$.

Conversely, assume that $\mathbf{\Pi}_{sw}$ is identifiable, but $\mathbf{\Pi}^{\text{SARX}}$ is not identifiable. Then there exists parameters $\theta_1, \theta_2 \in \Theta$, $\theta_1 \neq \theta_2$, such that $\mathbf{\Pi}^{\text{SARX}}(\theta_1)$ and $\mathbf{\Pi}^{\text{SARX}}(\theta_2)$ are equivalent. This means that $\Sigma_{\mathbf{\Pi}^{\text{SARX}}(\theta_1)} = \mathbf{\Pi}_{sw}(\theta_1)$ and $\Sigma_{\mathbf{\Pi}^{\text{SARX}}(\theta_2)} = \mathbf{\Pi}_{sw}(\theta_2)$ realize the same input-output map. But this contradicts the identifiability of $\mathbf{\Pi}_{sw}$. □

In order to prove Theorem 2, we will need the following result which is interesting on its own right.

Theorem 5. Consider two SISO SARX systems $\mathbf{S}_1 = \{h_q\}_{q\in Q}$ and $\mathbf{S}_2 = \{g_q\}_{q\in Q}$ of type (n_y, n_u) and assume that for some $q \in Q$, either $h_q^{n_y} \neq 0$ or $h_q^{n_y+n_u} \neq 0$. If there exists an isomorphism[4] between the associated LSSs $\Sigma_{\mathbf{S}_1}$ and $\Sigma_{\mathbf{S}_2}$, then this isomorphism is the identity map.

Theorem 5 implies, under some mild conditions, that the transformation of two different SARX systems to state-space representations cannot result in isomorphic systems.

Proof of Theorem 5. Assume that $\Sigma_{\mathbf{S}_1} = (n, Q, \{(A_q, B_q, C_q) \mid q \in Q\}, x_0)$ and $\Sigma_{\mathbf{S}_2} = (n', Q, \{(A'_q, B'_q, C'_q) \mid q \in Q\}, x'_0)$ with $n = n' = n = n_y + n_u$ and $p = m = 1$. Consider an isomorphism S between $\Sigma_{\mathbf{S}_1}$ and $\Sigma_{\mathbf{S}_2}$. Denote by e_i the ith standard unit vector of $\mathbb{R}^n$. Then $e_1^T, \dots, e_n^T$ form the standard basis in $\mathbb{R}^{1\times n}$. The proof depends on the following series of technical results.

Proposition 1.

$$e_1^T A_q = e_1^T S A_q \tag{C6}$$

Proof. From the construction of $\Sigma_{\mathbf{S}_i}$, $i = 1, 2$ it then follows that $C_q = e_1^T A_q$, $C'_q = e_1^T A'_q$. From the definition of isomorphism between LSSs, it follows that $C'_q S = C_q$, $q \in Q$. Hence, we obtain that

$$e_1^T A_q = e_1^T A'_q S.$$

[4] See the work of Petreczky *et al*[24] for the definition of isomorphism between LSSs

But $A_q^{'}\mathcal{S} = \mathcal{S}A_q$ by the definition of a LSS isomorphism, and hence we obtain the claim of the proposition. □

Proposition 2. The columns of A_q span the space

$$\text{Span}\{e_1, \dots, e_{n_u+n_y}\} \setminus \{e_{n_y+1}\}.$$

Proof of Proposition 2. Indeed, $A_q e_n = h_q^n e_1$, $A_q e_{n_y} = h_q^{n_y} e_1$, $A_q e_j = e_{j+1} + h_q^j e_1$ for all $j \in \{1, \dots, n-1\} \setminus \{n_y\}$. Hence, if either $h_q^{n_y} \neq 0$ or $h_q^n \neq 0$, then e_1 belongs to the column space of A_q, and hence $e_j = A_q e_{j-1} - h_q^{j-1} e_1$ belongs to the column space of A_q, for $j \in \{2, \dots, n\} \setminus \{n_y + 1\}$. □

Proposition 3. For any $i = 1, \dots, n_u + n_y$, if $e_i^T A_q = e_i^T \mathcal{S} A_q$, then $e_i^T = e_i^T \mathcal{S}$.

Proof of Proposition 3. Indeed, if $e_i^T A_q = e_i^T \mathcal{S} A_q$, then this implies that $(e_i^T - e_i^T \mathcal{S})A_q = 0$. By Proposition 2 this implies that $(e_i^T - e_i^T \mathcal{S})e_j = 0$ for all $j \in \{1, \dots, n_u + n_y\} \setminus \{n_y + 1\}$. Notice that from the construction of $\Sigma_{\mathbf{S}_1}$, $\Sigma_{\mathbf{S}_2}$ and the definition of a LSS morphism it follows that $e_{n_y+1} = B_q^{'} = \mathcal{S}B_q = \mathcal{S}e_{n_y+1}$. Hence, $(e_i^T - e_i^T \mathcal{S})e_{n_y+1} = 0$ and thus

$$(e_i^T - e_i^T \mathcal{S})e_j = 0, j = 1, \dots, n_y + n_u.$$

This is just an alternative way of formulating the conclusion of the proposition. □

Proposition 4. If $e_{j-1}^T \mathcal{S} = e_{j-1}^T$, then $e_j^T A_q = e_j^T \mathcal{S} A_q$ for all $j = \{2, \dots, n_y + n_u\} \setminus \{n_y + 1\}$.

Proof. Notice that $e_j^T A_q = e_{j-1}^T$, $e_j^T A_q^{'} = e_{j-1}^T$, $j = n_y, \dots, 2$, and $e_j^T A_q = e_{j-1}^T$, $e_j^T A_q^{'} = e_{j-1}^T$, for $j = n_y + n_u, \dots, n_y + 2$. Hence, by using $A_q^{'}\mathcal{S} = \mathcal{S}A_q$, we derive

$$e_{j-1}^T \mathcal{S} = e_j^T A_q^{'} \mathcal{S} = e_j^T \mathcal{S} A_q \tag{C7}$$

for all $j \in \{2, \dots, n_y\} \cup \{n_y + 2, \dots, n_y + n_u\}$. Since $e_{j-1}^T \mathcal{S} = e_{j-1}^T$, and $e_j^T A_q = e_{j-1}^T$ for all $j = \{2, \dots, n_u + n_y\} \setminus \{n_y + 1\}$, from (C7) we obtain the claim of the proposition. □

The rest of the proof of Theorem 5 proceeds as follows. We will prove that

$$e_j^T = \mathcal{S}e_j^T, j = 1, \dots, n_y + n_u, \tag{C8}$$

which is just another way of saying that $\mathcal{S}$ is the identity matrix. To this end, from (C6) and Proposition 3 it follows that (C8) holds for $j = 1$. Moreover, the $n_y + 1$th row of A_q and $A_q^{'}$ are both zero, hence, $0 = e_{n_y+1}^T A_q\ 0 = e_{n_y+1}^T A_q^{'}$ and thus $0 = e_{n_y+1} A_q^{'} \mathcal{S} = e_{n_y+1}^T \mathcal{S} A_q$. From this we get that $e_{n_y+1}^T A_q = e_{n_y+1}^T \mathcal{S} A_q$ and by Proposition 3 this implies that (C8) holds for $j = n_y + 1$. Notice that if (C8) holds for $j = k \in \{1, \dots, n_u + n_y - 1\} \setminus \{n_y\}$, then by Proposition 4, $e_{k+1}^T A_q = e_{k+1}^T \mathcal{S} A_q$. By Proposition 3, the latter implies that (C8) holds for $j = k + 1$. Hence, by induction we get that (C8) holds for all j. □

Proof of Theorem 2. We will show that the LSS parameterization $\mathbf{\Pi}_{sw} : \Theta \ni \theta \mapsto \Sigma_{\mathbf{\Pi}^{\text{SARX}}(\theta)}$ is identifiable. By Corollary 6 this is sufficient for identifiability of $\mathbf{\Pi}^{\text{SARX}}$.

Since $\mathbf{\Pi}^{\text{SARX}}$ is strongly minimal, the LSS parameterization $\mathbf{\Pi}_{sw}$ is minimal[24]. In order to show identifiability of $\mathbf{\Pi}_{sw}$, according to Petreczky *et al*[24, Corollary 1], it is enough to show that the only isomorphism between elements of $\mathbf{\Pi}_{sw}$ is the identity. Consider now two elements $\Sigma_i = \Sigma_{\mathbf{\Pi}^{\text{SARX}}(\theta_i)}$, $\theta_i \in \Theta$, $i = 1, 2$ of $\mathbf{\Pi}_{sw}$. Notice that $\mathbf{\Pi}^{\text{SARX}}(\theta_1)$ is minimal, since it is strongly minimal, and thus if $\mathbf{\Pi}^{\text{SARX}}(\theta_1) = \{h_q\}_{q \in Q}$, then by Lemma 2 $h_q^{n_u+n_y} \neq 0$. But then from Theorem 5 it follows that the only isomorphism between Σ_1 and Σ_2 is the identity map.

□

D PROOF OF THEOREM 3

Let $K = (pn_y + mn_u)|Q|$. Then any SARX system of type (n_y, n_u) can be identified with a point in $\mathbb{R}^K$, by identifying the system with the collection of its parameters $\{h_q\}_{q \in Q}$, $h_q \in \mathbb{R}^{p \times (pn_y + mn_u)}$.

First, we construct a polynomial $P_{min}(X_1, \dots, X_K)$, such that $P_{min}(\mathbf{S}) \neq 0$ if and only if $\mathbf{S}$ is strongly minimal. To this end, consider the LSS $\Sigma_{\mathbf{S}}$ and consider the observability and controllability matrices $O(\Sigma_{\mathbf{S}})$ and $\mathcal{R}(\Sigma_{\mathbf{S}})$ as defined in[29]. Define

$$P_{obs}(X_1, \dots, X_K) = \det(O(\Sigma_{\mathbf{S}})^T O(\Sigma_{\mathbf{S}})),$$

and

$$P_{contr}(X_1, \dots, X_K) = \det(\mathcal{R}(\Sigma_{\mathbf{S}})\mathcal{R}(\Sigma_{\mathbf{S}})^T),$$

with the notation $\det(\cdot)$ referring to determinant. Then P_{obs} and P_{contr} are polynomials in the entries of the matrices of $\Sigma_{\mathbf{S}}$. Moreover, by applying the results of Petreczky *et al*[29],

- $P_{obs}(\mathbf{S}) \neq 0 \Leftrightarrow \mathcal{O}(\Sigma_{\mathbf{S}})$ has full rank $\Leftrightarrow \Sigma_{\mathbf{S}}$ is observable
- $P_{contr}(\mathbf{S}) \neq 0 \Leftrightarrow \mathcal{R}(\Sigma_{\mathbf{S}})$ has full rank $\Leftrightarrow \Sigma_{\mathbf{S}}$ is reachable.

Define now $P_{min} = P_{contr}P_{obs}$. Then $P_{min}(\mathbf{S}) \neq 0$ if and only if $\Sigma_{\mathbf{S}}$ is both observable and reachable, i.e. if and only if $\Sigma_{\mathbf{S}}$ is minimal.

Finally, consider a polynomial parametrization $\mathbf{\Pi}^{\text{SARX}}$ such that $\mathbf{\Pi}^{\text{SARX}}$ contains a strongly minimal element. The fact that $\mathbf{\Pi}^{\text{SARX}}$ is a polynomial parametrization implies that there exists polynomials $\mathbf{\Pi}_i^{\text{SARX}}$ in variables $X_1, \dots, X_d$, $i = 1, \dots, K$ such that $\mathbf{\Pi}^{\text{SARX}}(\theta) = (\mathbf{\Pi}_1^{\text{SARX}}(\theta), \dots, \mathbf{\Pi}_K^{\text{SARX}}(\theta))$ for all $\theta \in \Theta$. Here we used the identification of a SARX system of type (n_y, n_u) with a point in $\mathbb{R}^K$. Consider the polynomial

$$Q_{min}(X_1, \dots, X_d) = P_{min}(\mathbf{\Pi}_1^{\text{SARX}}(X_1, \dots, X_d), \dots \dots, \mathbf{\Pi}_K^{\text{SARX}}(X_1, \dots, X_d)).$$

Notice that the set of parameters from Θ which do not yield a minimal SARX system all satisfy the equation $Q_{min}(\theta) = 0$. From the assumption that $\mathbf{\Pi}^{\text{SARX}}$ contains a strongly minimal element it follows that for some $\theta \in \Theta$, $Q_{min}(\theta) = P_{min}(\mathbf{\Pi}^{\text{SARX}}(\theta)) \neq 0$. Hence, the set $G = \{\theta \in \Theta \mid Q_{min}(\theta) \neq 0\}$ is a non-empty subset of Θ and it is clearly generic. That is, $\mathbf{\Pi}^{\text{SARX}}$ is generically strongly minimal, and hence minimal.

References

1. Hannan E. J., Deistler M.. *The Statistical Theory of Linear Systems*. USA: John Wiley & Sons, Inc.; 1987.
2. Paoletti S, Juloski A, Ferrari-Trecate G, Vidal R. Identification of Hybrid Systems: A Tutorial. *European Journal of Control.* 2007;13(2-3):242-260.
3. Paoletti S, Roll J, Garulli A, Vicino A. On the Input-Output Representation of Piecewise Affine State Space Models. *IEEE Transactions on Automatic Control.* 2010;55:60-73.
4. Ozay N, Sznaier C, Camps O. A sparsification approach to set membership identification of a class of affine hybrid systems. *IEEE Transaction on Automatic Control.* 2012;57:634-648.
5. Vidal R. Recursive Identification of Switched ARX Systems. *Automatica.* 2008;44(9):2274-2287.
6. Lauer Fabien, Bloch Gérard. Hybrid System Identification: Theory and Algorithms for Learning Switching Models. In: :77-101; 2019.
7. Garulli A, Paoletti S, Vicino A. A survey on switched and piecewise affine system identification. In: 16th IFAC Symposium on System Identification:344-355; 2012.
8. Ljung L.. *System Identification: Theory for the user (2nd Ed.).* PTR Prentice Hall., Upper Saddle River, USA; 1999.
9. Astrom Karl Johan, Wittenmark Bjorn. *Adaptive Control*. USA: Addison-Wesley Longman Publishing Co., Inc.; 2nd ed.1994.
10. Bako L, Mercére G, Vidal R, Lecoeuche S. Identification of switched linear state space models without minimum dwell time. In: 15th IFAC Symposium on System Identification:569-574; 2009.
11. Bako L, Mercére G, Lecoeuche S. Online structured subspace identification with application to switched linear systems. *International Journal of Control.* 2009;82:1496-1515.
12. Verdult V, Verhaegen M. Subspace Identification of Piecewise Linear Systems. In: 43rd IEEE Conference on Decision and Control:3838-3843; 2004.

13. Juloski ALj, Weiland S, Heemels WPMH. A Bayesian Approach to Identification of Hybrid Systems. *IEEE Transactions on Automatic Control.* 2005;50:1520-1533.

14. Roll J, Bemporad A, Ljung L. Identification of piecewise affine systems via mixed-integer programming. *Automatica.* 2004;40:37-50.

15. Ferrari-Trecate G., Muselli M., Liberati D., Morari M.. A clustering technique for the identification of piecewise affine systems. *Automatica.* 2003;39:205-217.

16. Nakada H, Takaba K, Katayama T. Identification of piecewise affine systems based on statistical clustering technique. *Automatica.* 2005;41:905-913.

17. Paoletti S, Roll J, Garulli A, Vicino A. Input/ouput realization of piecewise affine state space models. In: 46th IEEE Conference on Decision and Control:3164-3169; 2007.

18. Fliess M, Join C, Perruquetti W. Real-time estimation for switched linear systems. In: 47th IEEE Conference on Decision and Control:941-946; 2008.

19. Juloski A, Heemels WPMH, Ferrari-Trecate G, Vidal R, Paoletto S, Niessen J. Comparison of four procedures for the identification of hybrid systems. In: Morari M., Thiele L. (eds) Hybrid Systems: Computation and Control. HSCC 2005. Lecture Notes in Computer Science, vol. 3414: Springer-Verlag, Berlin; 2005.

20. Bako L. Identification of switched linear systems via sparse optimization. *Automatica.* 2011;47:668-677.

21. Ohlsson H, Ljung L. Identification of switched linear regression models using sum-of-norms regularization. *Automatica.* 2013;49:1045-1050.

22. Lauer F, Bloch G, Vidal R. A continuous optimization framework for hybrid system identification. *Automatica.* 2011;47:608-613.

23. Weiland S, Juloski ALj, Vet B. On the equivalence of switched affine models and switched ARX models. In: 45th IEEE Conference on Decision and Control:2614-2618; 2006.

24. Petreczky M, Bako L, Schuppen JH. Identifiability of discrete-time linear switched systems. In: 13th ACM international conference on Hybrid Systems: Computation and Control:141-150; 2010.

25. Petreczky M, Bako L, Lecoeuche S. Minimality and identifiability of sarx systems. In: 16th IFAC Symposium on System Identification:541-546; 2012.

26. Gécseg F, Peák I. *Algebraic theory of automata*. Akadémiai Kiadó, Budapest; 1972.

27. Eilenberg S. *Automata, Languages and Machines*. Academic Press, New York, London; 1974.

28. Sun Z, Ge SS. *Switched linear systems: control and design*. Springer, London; 2005.

29. Petreczky M, Bako L, Schuppen H. Realization theory of discrete-time linear switched systems. *Automatica.* 2013;49(11):3337-3344.

30. Tóth R., Abbas H.S., Werner H.. On the state-space realization of LPV Input-Output Models: Practical Approaches. *IEEE Transactions on Control Systems Technology.* 2012;20(1).

31. Kwiatkowski A.. *LPV Modeling and Application of LPV Controllers to SI Engines*. Verlag Dr. Hut; 2008.

32. Cox D., Little J., O'Shea D.. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Springer-Verlag; 2007.

33. SageMath package.

34. Petreczky M., Bako L., Lecoeuche S., Motchon K.M.D.. *Minimality and identifiability of discrete-time SARX systems.* : Arxiv report; 2020.

35. van den Hof J.. Structural identifiability of linear compartmental systems. *IEEE Transactions on Automatic Control.* 1998;43:800-818.

36. Gantmacher FR. *The theory of matrices*. AMS Chelsea Publishing; 2000.