Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 54-5 (2021) 61–66

On the Stability of Switched ARX Models, with an Application to Learning via Regression Trees

Vittorio De Iuliis * Francesco Smarra * Costanzo Manes * Alessandro D'Innocenzo *

* Dipartimento di Ingegneria e Scienze dell'Informazione, e Matematica, Università degli Studi dell'Aquila, Via Vetoio, 67100 Coppito (AQ), Italy (e-mail: vittorio.deiuliis@univaq.it).

Abstract: This work studies the stability properties of Switched AutoRegressive eXogenous (SARX) models subject to arbitrary switching sequences. We provide necessary and sufficient conditions for the arbitrary switching stability of multiple-input single-output SARX models under nonnegativity constraints, and sufficient-only conditions removing sign constraints. The conditions are equivalently formulated on state-space representations of SARX models, due to their influential use in designing control strategies. As an application of the aforementioned results, we propose a novel algorithm for the identification of switched models with stability guarantees via Regression Trees, a powerful machine learning technique.

Keywords: Stability of switched systems, System identification, Regression Trees

1. INTRODUCTION

Switched AutoRegressive eXogenous (SARX) models are a powerful and popular choice to describe the input-output behaviour of complex dynamical systems. They constitute a reference class for switched and hybrid system identification (see Lauer and Bloch (2018)), and the recent literature shows a considerable effort in addressing important theoretical problems such as minimality, realizability, identifiability (see Petreczky et al. (2020) and related works).

A SARX model is a regressive model in input-output form described by:

$$y(k) = \Theta_{\sigma(k)}\phi(k) \tag{1}$$

where y is the measured output, $\phi(k)$ is the vector of regressive input-output terms, and $\Theta_{\sigma(k)}$ is a matrix of parameters, which switches according to a signal σ. The identification problem of SARX models is solved estimating the mode-dependent matrix $\Theta_{\sigma(k)}$ from input-output measurements according to some technique. Some relevant problems in identification and realization of SARX models are widely described in the seminal survey Garulli et al. (2012). For analysis, prediction and control purposes, a SARX model can be represented in state-space form, as discussed in Weiland et al. (2006). Nevertheless, assessing some important properties (such as stability and stabilizability) is challenging: the presence of switching dynamics requires sophisticated and computationally hard mathematical tools.

In this work we provide novel conditions for the stability of SARX models. The work starts showing how the stability analysis is quite simplified resorting to the state-space realization of model (1). After some preliminary results (Section 2.1) on the input-output stability of multiple-input multiple-output (MIMO) SARX models, the analysis focuses on the multiple-input single-output (MISO) case, for which novel conditions for the arbitrary switching stability are introduced. Indeed, we first formulate necessary and sufficient conditions for the arbitrary switching stability of MISO SARX models under nonnegativity constraints, then we remove the constraints at the expense of introducing conservatism (Section 2.2). The conditions are effective, and their test is straightforward, as it vastly simplifies the problem of deciding whether the joint spectral radius of the switching family of dynamic matrices corresponding to the state-space form of (1) is less than one. We also illustrate how the proposed conditions are less conservative than the sufficient-only existence of copositive linear common Lyapunov functions (Section 2.3). To the best of our knowledge, the results introduced in this work are the first stability conditions explicitly formulated for SARX models, meaning that their particular structure is exploited to obtain a remarkable simplification with respect to the classical tools for analyizing the stability of general switched systems.

As an application of the aforementioned results, Section 3 presents an identification technique with stability guarantees which leverages Regression Trees (RT, a well known machine learning technique) to estimate SARX models by input-output data gathered on an unknown system. The technique is a refinement of a method previously introduced in Smarra et al. (2020) to train a predictive model and apply Model Predictive Control, which has been successfully adopted in a number of heterogeneous and complex real-life scenarios (see, e.g. Smarra et al. (2018); Bünning et al. (2020); De Iuliis et al. (2021)). While the method of Smarra et al. (2020) exhibited impressive experimental results, it tends to produce unstable models even if the original system is stable. The impact of this limitation, which is a common drawback of most

10.1016/j.ifacol.2021.08.475

SARX estimation techniques, is negligible in a closed-loop regime, but preserving stability during an estimation procedure is a matter of consistency in terms of accuracy, and a fundamental requirement both in classical system identification and in recent learning-based approaches (see e.g. Van Gestel et al. (2001); Umlauft et al. (2017)). In Section 4 we validate the proposed learning methodology on a bilinear model derived from a real building experimental setup: a comparison in terms of prediction accuracy discusses the benefits of the novel approach with respect to the original technique of Smarra et al. (2020), the *k-LinReg* switched regression method of Lauer (2013) and a family of nonlinear ARX models.

Notation. I_n is the $n \times n$ identity matrix. Inequalities among vectors and matrices of the same dimensions have to be understood componentwise, i.e. $M \le N$ if $m_{ij} \le n_{ij}$ for all i, j. In this sense, a nonnegative matrix $M \in \mathbb{R}_+^{m\times n}$ is also denoted by $M \ge 0$, where 0 is the matrix of appropriate dimensions whose entries are all zero. Similarly, $|M|$ denotes the componentwise absolute value of matrix M. $s(M)$ denotes the spectrum of a square matrix M, and $\rho(M)$ is its spectral radius. M is said to be *Schur-stable* if $s(M) \subset \{z \in \mathbb{C} : |z| < 1\}$ or, equivalently, if $\rho(M) < 1$. For a set (or family) of matrices $\mathcal{M} = \{M_1, \dots, M_h\}$ the joint spectral radius (JSR) of $\mathcal{M}$ is defined as

$$\rho^*(\mathcal{M}) = \rho^*(M_1, \dots, M_h) = \lim_{k\to\infty} \max_{\Pi \in \mathcal{M}^k} \|\Pi\|^{1/k}, \quad (2)$$

with $\mathcal{M}^k$ being the set of all products of length k (allowing for repetitions) of matrices in $\mathcal{M}$. Clearly, for a single matrix M, $\rho^*(M) = \rho(M)$, i.e. the JSR of M is its spectral radius. See Jungers (2009) for further details.

2. STABILITY OF SARX MODELS

The general multi-input multi-output form of a switched ARX model is described by the regressive model (1):

$$y(k) = \Theta_{\sigma(k)}\phi(k)$$

where $\phi(k) = [y(k-1)^\top \cdots\ y(k-\delta_y)^\top u(k)^\top \cdots u(k-\delta_u)^\top]^\top \in \mathbb{R}^{q\delta_y + p(\delta_u+1)}$ is the vector of input-output regressive terms, with input $u \in \mathbb{R}^p$ and output $y \in \mathbb{R}^q$. $\Theta_{\sigma(k)} \in \mathbb{R}^{q\times[q\delta_y + p(\delta_u+1)]}$ is the matrix of parameters, which switches according to $\sigma(k)$. Rewriting (1) as:

$$y(k) = [\Theta^y_{\sigma(k)}\Theta^u_{\sigma(k)}]x(k) + \Theta^{u_0}_{\sigma(k)}u(k) \quad (3)$$

having defined the state vector

$$x(k) = [y(k-1)^\top \cdots y(k-\delta_y)^\top u(k-1)^\top \cdots u(k-\delta_u)^\top]^\top$$

model (1) can be realized in state-space form as follows:

$$\begin{aligned} x(k+1) &= A_{\sigma(k)}x(k) + B_{\sigma(k)}u(k) \\ y(k) &= C_{\sigma(k)}x(k) + D_{\sigma(k)}u(k) \end{aligned} \quad (4)$$

where:

$$A_{\sigma(k)} = \begin{bmatrix} \Theta^y_{\sigma(k)} & & \Theta^u_{\sigma(k)} & \\ I_{q(\delta_y-1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_{p(\delta_u-1)} & 0 \end{bmatrix}, \quad B_{\sigma(k)} = \begin{bmatrix} \Theta^{u_0}_{\sigma(k)} \\ 0 \\ I_p \\ 0 \end{bmatrix}$$

$$C_{\sigma(k)} = \left[\Theta^y_{\sigma(k)}\ \Theta^u_{\sigma(k)}\right], \quad D_{\sigma(k)} = \Theta^{u_0}_{\sigma(k)}. \quad (5)$$

The 0 symbol is used, above and hereinafter, for a matrix of suitable dimensions consisting of all zeros. See Weiland et al. (2006) for details on SARX state-space realization.

The bounded-input bounded-output stability under arbitrary switching of (1) is equivalent to the internal arbitrary switching stability of its state-space realization (4), which amounts at solving the NP-hard problem '$\rho^*(\mathcal{A}) < 1$?', where $\rho^*(\mathcal{A})$ is the JSR (see *Notation*) of the family of dynamic matrices $\mathcal{A} = \{A_1, \dots, A_h\}$ selected by $\sigma(k)$ in $\{1, \dots, h\}$ at time k. Briefly, we write that the family $\mathcal{A}$ is stable under arbitrary switching meaning that $\rho^*(\mathcal{A}) < 1$. In what follows, we first show that the arbitrary switching stability of SARX models is only determined by the properties of $\Theta^y_{\sigma(k)}$, i.e. the autoregressive part of the whole model. Then, for MISO SARX models ($q = 1$), we provide straightforward conditions for the arbitrary switching stability in terms of the coefficients of $\Theta^y_{\sigma(k)}$.

2.1 Preliminary results on MIMO SARX models

First of all, let us introduce a block notation for A_j:

$$A_j = \begin{bmatrix} A^y_j & A^u_j \\ 0 & N \end{bmatrix}, \quad j = 1, \dots, h, \quad (6)$$

where:

$$A^y_j = \begin{bmatrix} \Theta^y_j & \\ I_{q(\delta_y-1)} & 0 \end{bmatrix}, \quad A^u_j = \begin{bmatrix} \Theta^u_j \\ 0 \end{bmatrix}, \quad N = \begin{bmatrix} 0 & 0 \\ I_p(\delta_u-1) & 0 \end{bmatrix}$$

Notice that $A^y_j \in \mathbb{R}^{q\delta_y\times q\delta_y}$ and $N \in \mathbb{R}^{p\delta_u\times p\delta_u}$ are square matrices. Let's start seeing that $\rho^*(\mathcal{A}) = \rho^*(\mathcal{A}^y)$, where $\mathcal{A}^y = \{A^y_1, \dots, A^y_h\}$. The following result is a special case of Proposition 1.5 in Jungers (2009).

Lemma 1. The JSR of a family of block triangular matrices coincides with the maximum among the JSRs of the families of (square) matrices on their diagonal blocks, i.e.

$$\rho^*\left(\left\{\begin{bmatrix} M_i & P_i \\ 0 & Q_i \end{bmatrix}\right\}\right) = \max\{\rho^*(\{M_i\}), \rho^*(\{Q_i\})\}. \quad (7)$$

In the previous result, $\{M_i\}$ is a short notation for the family of matrices $\{M_1, \dots, M_h\}$ (the same holds for $\{P_i\}, \{Q_i\}$). Now consider (6).

Proposition 2. The arbitrary switching stability of $\mathcal{A}$ is equivalent to the arbitrary switching stability of $\mathcal{A}^y$, since

$$\rho^*(\mathcal{A}) = \rho^*(\mathcal{A}^y). \quad (8)$$

Proof. The proposition is proved by simply noting that each A_j shares the common diagonal block N, which is a triangular nilpotent matrix with $\rho(N) = \rho^*(N) = 0$. Then, Lemma 1 applies to the family of block triangular matrices $\mathcal{A}$ yielding:

$$\begin{aligned} \rho^*(\mathcal{A}) &= \rho^*\left(\left\{\begin{bmatrix} A^y_j & A^u_j \\ 0 & N \end{bmatrix}\right\}\right) \\ &= \max\left\{\rho^*(\{A^y_j\}), \rho(N)\right\} = \rho^*(\mathcal{A}^y). \end{aligned} \quad (9)$$

□

Proposition 2 holds for general MIMO SARX models and tells us that the input-output stability of (1) only depends on its autoregressive part (a quite expected results, along the lines of a classical property of non-switching models).

2.2 Stability conditions for MISO SARX models

Now we restrict our analysis to MISO SARX models, characterized by the switching vectors of parameters:

$$\Theta_{\sigma(k)} = \left[\Theta^y_{\sigma(k)}\ \Theta^{u_0}_{\sigma(k)}\ \Theta^u_{\sigma(k)}\right] \in \mathbb{R}^{\delta_y + p(\delta_u+1)} \quad (10)$$

where, at each $j = 1, \ldots, h$:

$$\begin{aligned} \Theta_j^y &= \left[a_{j,1} \ \ldots \ a_{j,\delta_y}\right], \quad \Theta_j^{u_0} = \left[b_{j,1} \ \cdots \ b_{j,p}\right] \\ \Theta_j^u &= \left[\alpha_{j,1} \ \ldots \ \alpha_{j,p\delta_u}\right]. \end{aligned} \tag{11}$$

Due to Proposition 2, we know that the only component of $\Theta_{\sigma(k)}$ which determines the arbitrary switching stability of SARX models is $\Theta^y_{\sigma(k)}$. Indeed, in the MISO case the stability is equivalent to checking that the family $\mathcal{A}^y = \{A_1^y, \ldots, A_h^y\}$ of companion matrices

$$A_j^y = \begin{bmatrix} \Theta_j^y \\ I_{\delta_y - 1} \ 0 \end{bmatrix} = \begin{bmatrix} a_{j,1} & a_{j,2} & \cdots & a_{j,\delta_y-1} & a_{j,\delta_y} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{12}$$

is such that $\rho^*(\mathcal{A}^y) < 1$.

Nesterov and Protasov (2013), Sec. 6.1, illustrates some properties of nonstationary difference equations resorting to switching models in companion form with *nonnegative* scalars on the first row. In particular, the following result is a direct consequence of Nesterov and Protasov (2013), Sec. 6.1, Corollary 1 and Lemma 3, which has direct implications on deciding the stability of family $\mathcal{A}$.

Theorem 3. Consider a family $\mathcal{A}^y = \{A_1^y, \ldots, A_h^y\}$ of *nonnegative* matrices in companion form as in (12), i.e. $a_{j,i} \geq 0$ for all i, j. Then it follows that

$$\rho^*(\mathcal{A}^y) < 1 \iff \sum_{i=1}^{\delta_y} a_{j,i} < 1, \quad \forall j = 1, \ldots, h. \tag{13}$$

Remark 4. Nesterov and Protasov (2013) also show that the larger growth rate for a model that switches among companion matrices as in (12) is attained staying on the A_j^y with largest spectral radius, and the JSR is equal to the largest spectral radius among all the possible A_j^y matrices.

At this point, we are in order to present the following condition of arbitrary switching stability for MISO SARX models as in (1), under nonnegativity constraints.

Theorem 5. Consider the state-space realization (4) of a SARX model. Then, if $a_{j,i} \geq 0$ for all i, j, the model is stable under arbitrary switching if and only if

$$\sum_{i=1}^{\delta_y} a_{j,i} < 1, \quad \forall j = 1, \ldots, h. \tag{14}$$

Proof. The result readily follows by Proposition 2 and Theorem 3. The former ensures that the arbitrary switching stability of $\mathcal{A}$ is equivalent to that of the family of companion matrices $\mathcal{A}^y$; the latter gives necessary and sufficient conditions to have $\rho^*(\mathcal{A}^y) < 1$ under the nonnegativity constraints on the first row coefficients $a_{j,i}$. □

So far, we have a necessary and sufficient condition for the stability of MISO SARX models that holds under nonnegativity constraints on $\Theta^y_{\sigma(k)}$. One may wonder whether such constraints can be relaxed. The answer is affirmative, since the definition of joint spectral radius naturally yields the following result (recall that $|M|$ stands for the componentwise absolute value of M):

Lemma 6. Consider a family of arbitrary matrices $\mathcal{M} = \{M_1, \ldots, M_h\}$. Then, the family $|\mathcal{M}| = \{|M_1|, \ldots, |M_h|\}$ is such that $\rho^*(\mathcal{M}) \leq \rho^*(|\mathcal{M}|)$.

Proof. Recalling that, for any $M, N \in \mathbb{R}^{n \times n}$ it holds that:

$$MN \leq |MN| \leq |M||N| \tag{15}$$

by defining $P = MN$, $Q = |M||N|$, for any submultiplicative matrix norm $\|\cdot\|$ one has

$$\|P\| \leq \||P|\| \leq \|Q\| \tag{16}$$

(indeed, the first inequality is an equality for the induced matrix norms $\|\cdot\|_{1,\infty}$). Hence, from the definition of JSR (2) and properties (15)-(16), it follows that:

$$\begin{aligned} \rho^*(\mathcal{M}) &= \lim_{k\to\infty} \max_{\Pi \in \mathcal{M}^k} \|\Pi\|^{1/k} \\ &\leq \lim_{k\to\infty} \max_{\bar{\Pi} \in |\mathcal{M}|^k} \|\bar{\Pi}\|^{1/k} = \rho^*(|\mathcal{M}|) \end{aligned} \tag{17}$$

where $\mathcal{M}^k$ and $|\mathcal{M}|^k$ are the set of all products of length k of matrices in $\mathcal{M}$ and $|\mathcal{M}|$, respectively. □

Finally, we can relax the assumptions of Theorem 5 and introduce the following sufficient condition for the arbitrary switching stability of MISO SARX models.

Theorem 7. Consider the state-space realization (4) of a SARX model. Then the model is stable under arbitrary switching if

$$\sum_{i=1}^{\delta_y} |a_{j,i}| < 1, \quad \forall j = 1, \ldots, h. \tag{18}$$

Proof. Due to Proposition 2 and Lemma 6 we have that $\rho^*(\mathcal{A}) = \rho^*(\mathcal{A}^y) < \rho^*(|\mathcal{A}^y|)$. Noting that each $|A_j^y|$ in $|\mathcal{A}^y|$ is a *nonnegative* companion matrix as in (12) made of the absolute values of the $a_{j,i}$ parameters of Θ_j^y, the thesis directly follows from Theorem 3. □

Remark 8. A unified formulation of Theorems 5 and 7 can be given just requesting that

$$\|\Theta_j^y\|_1 = \sum_{i=1}^{\delta_y} |a_{j,i}| < 1, \quad \forall j = 1, \ldots, h. \tag{19}$$

The condition is necessary and sufficient if $\Theta_j^y \geq 0$ (i.e. $a_{j,i} \geq 0$ for all i, j), otherwise it is only sufficient.

For what concerns the extension of the stability conditions of Theorem 5 and 7 to MIMO SARX models, we note that it is nontrivial, as it requires a generalization of Theorem 3 to the vastly more difficult case of block-companion matrices. Preliminary results in this direction have been illustrated in De Iuliis et al. (2020a,b), under some combinatorial assumptions which hardly fit with an identification scheme. A generalization of such result to address the general MIMO case is still an open problem.

2.3 Comparison with linear common Lyapunov functions

Theorem 3 and its application to SARX models (Theorems 5 and 7) are quite appealing since they are less conservative than the sufficient-only existence of a linear co-positive common Lyapunov function (LcCLF), which is an effective (and easy to check) stability condition for nonnegative families of matrices.

For a family of nonnegative matrices $\mathcal{M} = \{M_1, \ldots, M_h\}$, the existence of a LcCLF is equivalent to the existence of $v > 0$ s.t. $v^\top M_j < v^\top, \forall j = 1, \ldots, h$, and implies the arbitrary switching stability ($\rho^*(\mathcal{M}) < 1$). Indeed, the condition is only sufficient also for the special case of companion-form nonnegative matrices. To see this, consider the simple example of $\mathcal{M} = \{M_1, M_2\}$ with:

$$M_1 = \begin{bmatrix} 0.4 & 0.5 \\ 1 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0.5 & 0.4 \\ 1 & 0 \end{bmatrix}. \tag{20}$$

No LcCLF exists for $\mathcal{M}$, whereas Theorem 3 easily proves that $\rho^*(\mathcal{M}) < 1$, since $0.4 + 0.5 < 1$. Indeed $\rho^*(\mathcal{M}) = \rho(M_1) = 0.9348$.

3. LEARNING STABLE SWITCHED MODELS VIA REGRESSION TREES

Theorems 5 and 7, in principle, can be directly adopted in any identification method to constrain the estimation of $\Theta_{\sigma(k)}$, when the stability of the system to be estimated is a known property, or can be deduced from experimental findings. In this section, we provide an application to a particular technique to produce switched ARX models via Regression Trees. The method, which is formalized in Smarra et al. (2020) with the aim of producing a predictive model over a fixed horizon, is here adopted in the MISO special case (i.e. $q = 1$) and with the standard one-step predictive horizon of the regressive models considered in this work. The technique does not produce an arbitrary switching model as (4): instead, as will be illustrated in detail below, the switching signal is constrained since it depends on the regressive input-output signal $x(k)$ and on an exogenous disturbance signal $d(k)$. Nevertheless, the results of Section 2 apply also to this case, especially when the estimated model produces a large number of sub-modes and switches frequently (as is the case when the method is used to estimate an unknown highly nonlinear system). We now proceed recalling the basics of the learning algorithm, and the differences introduced with respect to the original formulation of Smarra et al. (2020).

Let us consider a dataset $\mathcal{D} = \{(y(k), u(k), d(k))\}_{k=1}^{\ell}$ of ℓ samples collected from the measurements of a physical system without any knowledge on the model structure, respectively consisting of output $y(k) \in \mathbb{R}$, input $u(k) \in \mathbb{R}^p$ and disturbances $d(k) \in \mathbb{R}^w$. Starting from the historical data $\mathcal{D}$, we want to estimate a model as in (4):

$$\begin{aligned} x(k+1) &= A_{\sigma(k)}x(k) + B_{\sigma(k)}u(k) \\ y(k) &= C_{\sigma(k)}x(k) + D_{\sigma(k)}u(k), \end{aligned} \tag{21}$$

where $x(k) = [y(k-1) \cdots y(k-\delta_y) \;\; u(k-1)^\top \cdots u(k-\delta_u)^\top]^\top$, $\sigma : \mathbb{R}^{n+w} \to \mathcal{M} \subset \mathbb{N}$ is a switching signal that at each time k depends on $(x(k), d(k))$, i.e. $\sigma(k) = \sigma(x(k), d(k))$ with a slight abuse of notation, $n = \delta_y + p\delta_u$ with $\delta_y > 0, \delta_u \geq 0$. We remark that here we extend the method of Smarra et al. (2020) considering arbitrary δ_u.

The algorithm in Smarra et al. (2020) to derive the model (21) using RTs first requires the definition of the datasets $\mathcal{Y} \doteq \{y(k)\}_{k=1}^{\ell}$ and $\mathcal{X} \doteq \{(x(k), d(k))\}_{k=1}^{\ell}$ from $\mathcal{D}$, and consists of 2 main steps: the conditions derived in Section 2 will be applied to the second step, in order to estimate a switching model with stability guarantees.

First step. The standard Classification And Regression Trees (CART) algorithm (see the Appendix in Smarra et al. (2020) or the reference book Breiman (2017) for details) is applied using $\mathcal{Y}$ as the predictor data set (i.e. the variable we want to predict) and $\mathcal{X}$ as the features dataset (i.e. the variables we use to compute the prediction). The CART algorithm constructs as output a *binary tree* $\mathcal{T}$ that partitions $\mathcal{X}$ into non-intersecting subsets $\mathcal{X}_j$, each corresponding to a leaf of $\mathcal{T}$. Each leaf j is associated to a hyper-rectangle $R_j \subset \mathbb{R}^{n+w}$. Let $|\mathcal{T}|$ denote the number of leaves obtained from the partitioning, then the hyper-rectangular sets R_j, $j = 1, \ldots, |\mathcal{T}|$ form a partition of $\mathbb{R}^{n+w}$. Thus, each leaf j consists of a set of samples $\mathcal{X}_j = \{(x(k_1), d(k_1)), \ldots, (x(k_\epsilon), d(k_\epsilon))\} \subset \mathcal{X}$, at time instants $k_1, \ldots, k_\epsilon$ that are not necessarily adjacent and such that each sample belongs to R_j.

Second step. Consider the partition $\{\mathcal{X}_j\}_{j=1}^{|\mathcal{T}|}$ of $\mathcal{X}$ constructed by the CART algorithm in the first step: for each leaf j of $\mathcal{T}$ we use the samples

$$\mathcal{X}_j^u \doteq \{(x(k_1), u(k_1)), \ldots, (x(k_\epsilon), u(k_\epsilon))\}$$

associated to the time instants $k_1, \ldots, k_\epsilon$ to identify a state-space ARX model

$$\begin{aligned} x(k+1) &= A_j x(k) + B_j u(k) \\ y(k) &= C_j x(k) + D_j u(k), \end{aligned}$$

where matrices A_j, B_j, C_j, D_j are structured as in (5), (11), and are computed solving the following constrained least squares problem:

Problem 1.

$$\underset{\xi_j}{\text{minimize}} \quad \| \Lambda_j \xi_j - \lambda_j \|_2^2 \tag{22}$$

$$\text{where } \xi_j^\top = [\Theta_j^y \quad \Theta_j^u \quad \Theta_j^{u_0}] \tag{23}$$

$$\text{subject to } \Lambda_j^\top = \begin{bmatrix} x(k_1) & x(k_2) & \cdots & x(k_\epsilon) \\ u(k_1) & u(k_2) & \cdots & u(k_\epsilon) \end{bmatrix} \tag{24}$$

$$\lambda_j^\top = [\, y(k_1) \;\; y(k_2) \;\; \cdots \;\; y(k_\epsilon) \,] \tag{25}$$

$$\Gamma_{eq}\, \xi_j = \gamma_{eq}, \quad \Gamma_{ineq}\, \xi_j \leq \gamma_{ineq} \tag{26}$$

$$\|\Theta_j^y\|_1 < 1 \tag{27}$$

where the linear equality and inequality (26) can be useful in practice to constrain elements in ξ_j due to additional information on the plant's dynamics.

Differently from the algorithm of Smarra et al. (2020), and in order to learn a switching model with stability guarantees, we added constraint (27) in Problem 1 (see Remark 8): this enforces the coefficients of all the matrices A_j to satisfy the stability condition (18) of Theorem 7. The constraint $\Theta_j^y \geq 0$ (i.e., $a_{j,i} \geq 0$ for all i) can also be added to Problem 1 to enforce the necessary and sufficient condition (14) of Theorem 5. Notice that $\Theta_j^y \geq 0$ does not impose input-output positivity of the estimated model, as the sign constraint only concerns the autoregressive terms of Θ_j (no sign constraints are imposed on $\Theta_j^u, \Theta_j^{u_0}$).

Let ξ_j^* be the optimal solution of Problem 1 for leaf j: we can associate to each leaf $j = 1, \ldots, |\mathcal{T}|$ a predictive model $\hat{y}(k) = [x^\top(k) \;\; u^\top(k)]\, \xi_j^*$, from which the matrices A_j, B_j, C_j, D_j in equation (21) can be trivially derived.

The above learning algorithm is run off-line, on a training dataset. When the derived model is used on a validation dataset for prediction, or in a run-time procedure for control purposes, at each step k the state and disturbance measurements $(x(k), d(k))$ are used to identify the unique hyper-rectangular set such that $(x(k), d(k)) \in R_j$, corresponding to the leaf j of $\mathcal{T}$ and thus to the matrices A_j,

B_j, C_j, D_j. Therefore, the obtained predictive model can be written as in (21), with the switching signal defined as

$$\sigma(x(k), d(k)) = j \iff (x(k), d(k)) \in R_j. \quad (28)$$

We remark that, for analogy with Smarra et al. (2020), $u(k)$ does not contribute to the partitioning process of *First step* and therefore σ does not depend on $u(k)$: this choice was taken to reduce the computational complexity of the control algorithm. Nevertheless, also considering $u(k)$ in the partitioning process is trivial.

We finally observe that, while in model (4) the switching is an *arbitrary* signal, in model (21) the switching signal (28) *is not arbitrary* since it depends on $(x(k), d(k))$. As a consequence the stability conditions derived in Section 2, which are based on the JSR of the family of all (switching) matrices associated to the $|\mathcal{T}|$ leaves of $\mathcal{T}$ (each of dimension n), are in general just sufficient conditions when applied to model (21), and in particular the necessary and sufficient condition of Theorem 5 becomes just sufficient. Indeed, as shown in Kozyakin (2014), to derive necessary and sufficient stability conditions on model (21) one should first determine the digraph representing all admissible switchings among the leaves of $\mathcal{T}$ (which is already a tough task), and then compute the JSR of a family of $|\mathcal{T}|$ matrices, each of dimension $n|\mathcal{T}|$. In this paper we do not follow this approach for two reasons: first, it is computationally expensive since usually $|\mathcal{T}| \gg n$; second, unless strong constraints can be assumed on the disturbance signal, it can in general activate (almost) all switching sequences.

4. VALIDATION

In this section we compare in terms of prediction accuracy the effect of the stability conditions introduced by Theorems 5 and 7 on the learning procedure described in Section 3, with respect to the "free" approach of Smarra et al. (2020), where no stability constraints are employed in the estimation. We also compare such approaches with:

- the *k-LinReg* algorithm of Lauer (2013), suggested as a baseline reference method for switched regression by the recent monograph Lauer and Bloch (2018);
- nonlinear ARX models estimated using the *System Identification toolbox* of MATLAB, with the default choice of wavelet networks nonlinearity.

Notice that neither of the aforementioned techniques allows constraining the identification procedure to estimate models with stability guarantees. The comparison is provided on a Building Automation System (BAS) simulator which considers a bilinear building model developed at the Automatic Control Laboratory at ETH Zurich.

Case study description. The system is described by the following bilinear model:

$$x(k+1) = Ax(k) + (B_u + B_{xu}[x(k)] + B_{du}[d(k)])u(k) + B_d d(k) \quad (29)$$

where $x \in \mathbb{R}^{12}$ is the internal system state and includes the temperatures characterizing the room and in particular the zone temperature, $u \in \mathbb{R}^4$ is the control input vector consisting of the blind position, the gains due to electric lighting, the evaporative cooling usage factor and the heat from the radiator, and $d \in \mathbb{R}^7$, is the disturbance vector that includes weather conditions and internal gains. Due to space limitation, we refer the reader to Oldewurtel (2011); Smarra et al. (2020) for further details about the model.

We used model (29) to generate a dataset for a whole year with a sampling time of 5 minutes. To simulate such trajectories we exploited a dataset from *MeteoSwiss* of the weather conditions, collected hourly, as disturbance signal. The output selected to construct the dataset is the first state component (i.e., the room ambient temperature). 10 months of data are used to learn (train) the model, and 2 months for validation (January & May). For brevity, only the validation over the month of January is reported.

Prediction accuracy. In this section we compare, in terms of prediction accuracy, models obtained using k-LinReg, nonlinear ARX, and the three methods based on regression trees: the original method of Smarra et al. (2020), the one with stability guarantees of Theorem 7 obtained by constraint (27) and the one with stability guarantees of Theorem 5 obtained by (27) enforced with $\Theta_j^y \geq 0$. In all the aforementioned methods we considered δ_y ranging from 1 to 6 and, for the sake of comparison with Smarra et al. (2020), $\delta_u = 1$. For the same reason, no feedthrough term is estimated (i.e., $\Theta_j^{u_0} = 0$), which is coherent with the choice of $y(k) = x_1(k)$ in model (29). Validation results are shown in Figure 1 in terms of Normalized Root Mean Square Error (NRMSE), with the normalization operated with respect to the mean, i.e. $NRMSE[\%] = \frac{\sqrt{\frac{1}{N}\sum_{k=1}^{N}(y(k)-\hat{y}(k))^2}}{\frac{1}{N}\sum_{k=1}^{N} y(k)} \cdot 100$, where N is the number of samples in the dataset and $y, \hat{y}$ denote the true and predicted values. For the k-LinReg method we ran several simulations varying the number of modes between 1 and 500. Rather surprisingly, the best accuracy has been achieved in the case of only 2 modes, so the results here reported refer to this case. Figure 1 illustrates that all the RT-based approaches outperform k-LinReg and nonlinear ARX in terms of prediction accuracy. Figure 2 shows the values of $y(k) = x_1(k)$ predicted by each method on the validation dataset for a fixed autoregressive order $\delta_y = 6$.

Moreover, we analyzed the state-space switching models obtained with the RT technique of Smarra et al. (2020) and the enhanced ones proposed in this work, for the case of $\delta_y = 6$. All of them produced 215 discrete modes: note that such large number of discrete states does not generate any computational issue in the run-time prediction, since at each time k the selection of $\sigma(x(k), d(k)) = j \iff (x(k), d(k)) \in R_j$ as in equation (28) requires to explore a binary tree checking only $\lceil \log_2(215) \rceil = 8$ linear scalar inequality conditions like $x_i(k) > \nu_i$ or $d_i(k) > \nu_i$. The approach of Smarra et al. (2020) produced 32 unstable discrete modes (out of the total 215 modes) and a JSR in the interval $[1.9243, 2.8351]$, computed by MATLAB JSR toolbox (see Vankeerberghen et al. (2014)). The models with stability guarantees have obviously zero unstable modes. For the model with $\Theta_j^y \geq 0$, the JSR of the family $\{A_1, \ldots, A_{215}\}$ is easily computable as the largest spectral radius among those matrices (see Remark 4) and equals to 0.9976. It is quite interesting to see that the JSR toolbox would give no definitive answers about the stability of $\{A_1, \ldots, A_{215}\}$ for the model with no sign constraint on Θ_j^y even after several minutes of computations, as it returns the uncertainty bound $\rho^*(A_1, \ldots, A_{215}) \in [0.9975, 1.1536]$. Indeed, we know by Theorem 7 that the true value is less than one by construction.

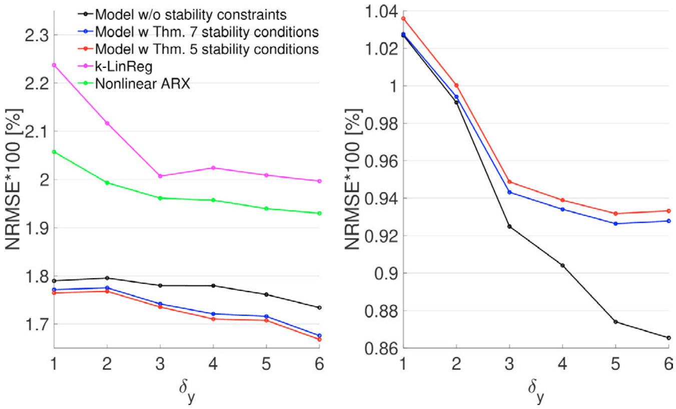

Fig. 1. NRMSE on the validation (left) and training (right) datasets considering different values of δ_y.

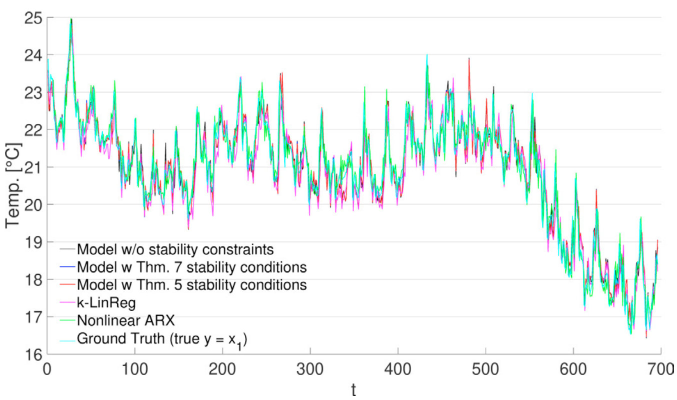

Fig. 2. Comparison of the values $\hat{y}(k) = \hat{x}_1(k)$ predicted by each method (with $\delta_y = 6$) on the validation dataset.

Discussion. Figure 1 shows that the best overall estimation, in terms of NRMSE on the validation dataset, is obtained by the methodologies with stability guarantees. Moreover, the addition of the nonnegativity constraint on Θ_j^y (i.e., exploiting Theorem 5) produces a slight increase in the predictive accuracy with respect to the stable model with no sign constraints. For what concerns the case study proposed in this work, we deal with a system whose input-output behavior is inherently nonnegative (and stable), as the observed variable is a temperature in a building and, even though expressed in Celsius, it is reasonably expected to only assume nonnegative values. In cases like this, the (partial) nonnegativity constraints in the estimation procedure do not introduce any sort of degradation. In fact, the model accuracies are even superior due to the combined stability and nonnegativity constraints.

Finally, a remark on the optimality of the solution. It could seem counterintuitive that by adding the stability constraint (27) (and possibly the nonnegativity constraint on Θ_j^y) to the QP Problem 1 we ended up with a better model with respect to the unconstrained formulation of Smarra et al. (2020), as shown by the validation in Figure 1 (left plot). To provide an answer to this point we tested the models' accuracies on the training dataset. Results are shown in the right plot of Figure 1, where it can be seen that the behaviour is reversed. This confirms that Problem 1 without added constraints finds the best solution within the training dataset: however, the constraints help avoiding overfitting and the stable models (of a stable system) provide better accuracy on the validation dataset.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the Italian Government under the PON Ricerca & Innovazione 2014-2020 (AIM1877124-Attività 1) and CIPE resolution n.135 (Dec. 21, 2012), project INCIPICT.

REFERENCES

Breiman, L. (2017). *Classification and regression trees.* Routledge.

Bünning, F., Huber, B., Heer, P., Aboudonia, A., and Lygeros, J. (2020). Experimental demonstration of data predictive control for energy optimization and thermal comfort in buildings. *Energy and Buildings*, 211, 109792.

De Iuliis, V., Di Girolamo, G.D., Smarra, F., and D'Innocenzo, A. (2021). A comparison of classical identification and learning-based techniques for cyber-physical systems. In *Proceedings of the 29th IEEE Mediterranean Conference on Control and Automation (to appear).*

De Iuliis, V., D'Innocenzo, A., Germani, A., and Manes, C. (2020a). Stability conditions for linear discrete-time switched systems in block companion form. *IET Control Theory & Applications*, 14(19), 3107–3115.

De Iuliis, V., D'Innocenzo, A., Germani, A., and Manes, C. (2020b). On the stability of discrete-time linear switched systems in block companion form. *IFAC-PapersOnLine*, 53(2), 2033–2038.

Garulli, A., Paoletti, S., and Vicino, A. (2012). A survey on switched and piecewise affine system identification. *IFAC Proceedings Volumes*, 45(16), 344–355.

Jungers, R. (2009). *The joint spectral radius: theory and applications*, volume 385. Springer.

Kozyakin, V. (2014). The Berger-Wang formula for the Markovian joint spectral radius. *Linear Algebra and its Applications*, 448, 315–328.

Lauer, F. (2013). Estimating the probability of success of a simple algorithm for switched linear regression. *Nonlinear Analysis: Hybrid Systems*, 8, 31–47.

Lauer, F. and Bloch, G. (2018). *Hybrid System Identification: Theory and Algorithms for Learning Switching Models.* Springer.

Nesterov, Y. and Protasov, V.Y. (2013). Optimizing the spectral radius. *SIAM Journal on Matrix Analysis and Applications*, 34(3), 999–1013.

Oldewurtel, F. (2011). *Stochastic model predictive control for energy efficient building climate control.* Ph.D. thesis, ETH Zurich.

Petreczky, M., Bako, L., Lecoeuche, S., and Motchon, K.M.D. (2020). Minimality and identifiability of discrete-time switched autoregressive exogenous systems. *International Journal of Robust and Nonlinear Control*, 30(15), 5871–5891.

Smarra, F., Di Girolamo, G.D., De Iuliis, V., Jain, A., Mangharam, R., and D'Innocenzo, A. (2020). Data-driven switching modeling for MPC using Regression Trees and Random Forests. *Nonlinear Analysis: Hybrid Systems*, 36, 100882.

Smarra, F., Jain, A., de Rubeis, T., Ambrosini, D., D'Innocenzo, A., and Mangharam, R. (2018). Data-driven model predictive control using random forests for building energy optimization and climate control. *Applied Energy*, 226, 1252–1272.

Umlauft, J., Lederer, A., and Hirche, S. (2017). Learning stable Gaussian process state space models. In *2017 American Control Conference (ACC)*, 1499–1504. IEEE.

Van Gestel, T., Suykens, J.A., Van Dooren, P., and De Moor, B. (2001). Identification of stable models in subspace identification by using regularization. *IEEE Transactions on Automatic Control*, 46(9), 1416–1420.

Vankeerberghen, G., Hendrickx, J., and Jungers, R.M. (2014). JSR: A toolbox to compute the joint spectral radius. In *Proceedings of the 17th International Conference on Hybrid systems: Computation and Control*, 151–156. ACM.

Weiland, S., Juloski, A.L., and Vet, B. (2006). On the equivalence of switched affine models and switched ARX models. In *Proceedings of the 45th IEEE Conference on Decision and Control*, 2614–2618.