Observability and Identifiability of Jump Linear Systems*

René Vidal[1] Alessandro Chiuso[2] Stefano Soatto[3]

Abstract

We analyze the observability of the continuous and discrete states of a class of linear hybrid systems. We derive rank conditions that the structural parameters of the model must satisfy in order for filtering and smoothing algorithms to operate correctly. We also study the identifiability of the model parameters by characterizing the set of models that produce the same output measurements. Finally, when the data are generated by a model in the class, we give conditions under which the true model can be identified.

1 Introduction

Hybrid systems are mathematical models of physical processes governed by differential equations that exhibit discontinuous behavior. Examples of such processes are ubiquitous in nature and in man-made devices, from action potentials in neurons to flight control systems in aircrafts. A particular but important class of hybrid systems is obtained by assuming that the dynamics between discrete events are *linear*. This class of systems is important not only because the analysis and design of linear control systems is well understood, but also because many real processes can be approximated arbitrarily well by models in this class. Previous work on hybrid systems concentrated on the areas of modeling, stability, controllability and verification. However (see Section 1.2), little attention has been devoted to the study of the *observability* of the continuous and discrete states of a hybrid system. Another important issue is whether the model itself can be inferred from data, *i.e.,* whether it is *identifiable*. While many identification algorithms have been proposed (see Section 1.2), most of them do not give conditions under which their solution is unique.

In this paper, we study the observability and identifiability of so-called jump linear systems (JLSs), *i.e.,* systems whose evolution is determined by a collection of linear models with *continuous state* $x_t \in \mathbb{R}^n$ connected by switches of a number of *discrete states* $\lambda_t \in \{1, 2, \ldots, N\}$. The evolution of the continuous state x_t is described by the linear system

$$\begin{aligned} x_{t+1} &= A(\lambda_t)x_t + v_t \\ y_t &= C(\lambda_t)x_t + w_t, \end{aligned} \tag{1}$$

where $A(k) \in \mathbb{R}^{n\times n}$ and $C(k) \in \mathbb{R}^{p\times n}$, for $k \in \{1, 2, \ldots, N\}$, $x_{t_0} \sim \mathcal{P}_0$, $v_t \sim \mathcal{N}(0, Q(\lambda_t))$ and $w_t \sim \mathcal{N}(0, R(\lambda_t))$. The evolution of the discrete state λ_t can be modeled, for instance, as an irreducible Markov chain governed by the transition map π, $P(\lambda_{t+1}|\lambda_t) = \pi_{t+1,t}$, or (a we do here) as a deterministic but unknown input that is piecewise constant and finite valued[1].

1.1 Contributions of this paper

In Section 2 we introduce the notion of observability for JLSs, and derive conditions under which the continuous and discrete states can be inferred from data. Unlike previous work, we derive simple and intuitive rank conditions that depend on the system parameters $A(\cdot), C(\cdot)$ and on the separation between jumps, and *not* on the noise or inference criterion. The rank conditions we derive can be thought of as an extension of the Popov-Belevic-Hautus rank test for linear systems [10] and are, to the best of our knowledge, novel.

In Section 3 we study two problems. The first one is concerned with the characterization of the set of models that produce the same outputs ("realizability"). We show that, in lack of assumptions on the model generating the data, there are infinitely many models that produce the same measurements and, therefore, a unique model cannot be inferred from data. We also show that the set of unidentifiable models is actually the entire set of possible models, unless proper conditions are imposed. Therefore, extreme care has to be exercised in the use of iterative identification techniques for JLSs, since they can in principle converge to any model if proper conditions are not enforced. We derive such conditions in Section 3.1. The second problem is concerned with the conditions under which the actual model that generated the data (the "true" model) can be recovered from the output ("identifiability"), under the assumption that the true model belongs to the class of JLSs. We address this problem in Section 3.2.

[1]Department of EECS, University of California, Berkeley CA 94720, USA, e-mail: rvidal@eecs.berkeley.edu.

[2]Dipartimento di Ingegneria dell'Informazione, Università di Padova, Italy, e-mail: chiuso@dei.unipd.it.

[3]Department of Computer Science, University of California, Los Angeles CA 90095, USA, e-mail: soatto@ucla.edu.

*The first author thanks Dr. Shankar Sastry for insights. Work funded by grants ONR N00014-00-1-0621, MURST "New Techniques for Identification and Adaptive Control of Industrial Systems", ASI 634, NSF-ECS 0200511 and NSF-STC CENS.

[1]Most of the literature on hybrid systems restricts the switching mechanism of the discrete state to depend on the value of the continuous state. While this is generally sensible in the study of stability, it is a significant restriction to impose in the context of filtering and identification. Our model is more general, as it imposes no restriction on the mechanism that governs the transitions between discrete states.

1.2 Relation to prior work

Filtering and identification of JLSs was an active area of research through the seventies and eighties: a review of the state of the art in 1982 can be found in [19]. That paper discusses sub-optimal algorithms for minimum mean-square error state estimation. In [18], the same author uses a finite-memory approximation to the maximum likelihood for simultaneously estimating the model parameters and the continuous and discrete states. The paper includes a condition for observability which, although tautological, is significant because it represents the first attempt to characterize the observability of hybrid systems.

The field was revived in the last few years, when many iterative algorithms were proposed, along with a few attempts to characterize observability. [16] gives an unusual and somewhat limited condition in terms of the existence of a discrete state trajectory. These conditions pertain to systems where the discrete state is controlled, rather than evolving out of its own dynamics. [17] gives conditions for a particular class of linear time-varying systems where the system matrix is a linear combination of a basis with time-varying coefficients. They relate to the conditions of [16], where the discrete state is controlled. [8] addresses observability and controllability for switched linear systems with known and periodic transitions. [2] proposes the notion of incremental observability of a hybrid system, which requires the solution of a mixed-integer linear program in order to be tested.

In a series of papers (see [7] and references therein) Krishnamurthy, Doucet et al. propose various forms of alternating minimization algorithms for estimating continuous and/or discrete states that are optimal in the maximum a-posteriori sense. Various forms of approximate inference for JLSs include [14, 12, 4]. In [12], approximate filtering and maximum likelihood parameter estimates are obtained via alternating minimization. The approximate filtering entails the assumption that the underlying conditional density is unimodal, similarly to [14], which implicitly approximates the posterior density with a Gaussian. Also [15] alternates between approximate filtering, which uses so-called Viterbi approximations (where the evolution of the conditional density is approximated by the trajectory of its mode; which subsumes the assumption that it is unimodal), and maximum likelihood parameter estimation. In [13] approximate filtering is performed by approximating the posterior density with a member of an ad-hoc parametric class. In [5] particle filtering is used for simultaneous filtering and identification. In [9] approximate filtering is applied to systems in which the discrete dynamics do not affect the evolution of the continuous states, but only the choice of the measured output. The textbook [1] uses the interacting multiple models scheme to provide an approximation of the two-step prediction using N Gaussian densities rather than N^2. [3] uses mixed-integer quadratic programming for identifying the parameters of piecewise affine models with known polyhedral partitions. The algorithm has polynomial complexity on the number of data.

2 Observability

Given a model $\Sigma = \{A(k), C(k); k = 1 \dots N\}$ of the type described by equation (1), we focus our attention on how to infer the state of the system and the system parameters from the output $\{y_t\}$. The simplest instance of this problem can be informally described as follows. Assume that we are given the model parameters $A(k), C(k)$ and that Σ evolves starting from an (unknown) initial condition (x_{t_0}, λ_{t_0}). Given the output on the interval $[t_0, t_0 + T]$, is it possible to reconstruct the sequence of continuous states $x_{t_0}, \dots, x_{t_0+T}$ as well as the sequence of discrete states $\lambda_{t_0}, \dots \lambda_{t_0+T}$ uniquely?

If the sequence of discrete states is known, then the output of the system between two consecutive jumps can be written explicitly in terms of the model parameters $A(\cdot), C(\cdot)$, and the initial value of the continuous state x_{t_0}. Thus the entire continuous state trajectory can be reconstructed from the initial value of the continuous state x_{t_0} and the discrete states $\lambda_{t_0}, \dots, \lambda_{t_0+T}$. More specifically, the output sequence is given by:

$$\begin{bmatrix} y_{t_0} \\ \vdots \\ y_{t_1-1} \\ y_{t_1} \\ \vdots \\ y_{t_2-1} \\ y_{t_2} \\ \vdots \end{bmatrix} = \begin{bmatrix} C(\lambda_{t_0}) x_{t_0} \\ \vdots \\ C(\lambda_{t_0}) A(\lambda_{t_0})^{t_1-t_0-1} x_{t_0} \\ C(\lambda_{t_1}) A(\lambda_{t_0})^{t_1-t_0} x_{t_0} \\ \vdots \\ C(\lambda_{t_1}) A(\lambda_{t_1})^{t_2-t_1-1} A(\lambda_{t_0})^{t_1-t_0} x_{t_0} \\ C(\lambda_{t_2}) A(\lambda_{t_1})^{t_2-t_1} A(\lambda_{t_0})^{t_1-t_0} x_{t_0} \\ \vdots \end{bmatrix} \tag{2}$$

$$= \begin{bmatrix} \mathcal{O}_{\tau_0}(\lambda_{t_0}) x_{t_0} \\ \mathcal{O}_{\tau_1}(\lambda_{t_1}) A(\lambda_{t_0})^{t_1-t_0} x_{t_0} \\ \mathcal{O}_{\tau_2}(\lambda_{t_2}) A(\lambda_{t_1})^{t_2-t_1} A(\lambda_{t_0})^{t_1-t_0} x_{t_0} \\ \vdots \end{bmatrix} \tag{3}$$

where $\{t_k, k \geq 1\}$ are the switching times, $\tau_k = t_{k+1} - t_k$ and $\mathcal{O}_{\tau_k}(\lambda_{t_k}) \in \mathbb{R}^{p\tau_k \times n}$, defined by the equations above, is the extended observability matrix of the pair $(A(\lambda_{t_k}), C(\lambda_{t_k}))$. We thus propose the following notions of indistinguishability and observability:

Definition 1 (Indistinguishability) *We say that the states $\{x_{t_0}, \lambda_{t_0}, \dots, \lambda_{t_0+T}\}$ and $\{\bar{x}_{t_0}, \bar{\lambda}_{t_0}, \dots, \bar{\lambda}_{t_0+T}\}$ are* **indistinguishable** *on the interval $[t_0, t_0 + T]$ if the corresponding outputs in free evolution $\{y_{t_0}, \dots, y_{t_0+T}\}$ and $\{\bar{y}_{t_0}, \dots, \bar{y}_{t_0+T}\}$ are equal. We denote the set of states which are indistinguishable from $\{x_{t_0}, \lambda_{t_0}, \dots, \lambda_{t_0+T}\}$ as $\mathcal{I}(x_{t_0}, \lambda_{t_0}, \dots, \lambda_{t_0+T})$.*

Definition 2 (Observability) *We say that a state* $\{x_{t_0}, \lambda_{t_0}, \ldots, \lambda_{t_0+T}\}$ *is observable on* $[t_0, t_0+T]$ *if* $\mathcal{I}(x_{t_0}, \lambda_{t_0}, \ldots, \lambda_{t_0+T}) = \{x_{t_0}, \lambda_{t_0}, \ldots, \lambda_{t_0+T}\}$. *When any admissible state is observable, we say that the model* Σ *is* **observable**.

2.1 Observability of the initial state

We first analyze the conditions under which we can determine $\lambda_{t_0} = \lambda_{t_0+1} = \cdots = \lambda_{t_1-1}$ and x_{t_0} uniquely, *i.e.,* before a switch occurs. We have that $\{x_{t_0}, \lambda_{t_0}, \ldots, \lambda_{t_1-1}\}$ is indistinguishable from $\{\bar{x}_{t_0}, \bar{\lambda}_{t_0}, \ldots, \bar{\lambda}_{t_1-1}\}$ if $\mathcal{O}_{\tau_0}(\lambda_{t_0})x_{t_0} = \mathcal{O}_{\tau_0}(\bar{\lambda}_{t_0})\bar{x}_{t_0}$, *i.e.,* if $\text{rank}([\mathcal{O}_{\tau_0}(\lambda_{t_0}) \ \ \mathcal{O}_{\tau_0}(\bar{\lambda}_{t_0})]) < 2n$. If either $p\tau_0 < 2n$ or $\text{rank}(\mathcal{O}_{\tau_0}(\lambda_{t_0})) < n$ or $\text{rank}(\mathcal{O}_{\tau_0}(\bar{\lambda}_{t_0})) < n$, the above rank condition is trivially satisfied, so we assume that $\tau_0 \geq \underline{\tau} \triangleq 2n$. We then have the following:

Lemma 1 (Observability of the initial state) *If* $\tau_0 \geq \underline{\tau} \triangleq 2n$, *then* $\{x_{t_0}, \lambda_{t_0}, \lambda_{t_0+1}, \ldots, \lambda_{t_0+\underline{\tau}-1}\}$ *is observable if and only if for all* $k \neq k' \in \{1, \ldots, N\}$ *we have* $\text{rank}([\mathcal{O}_{\underline{\tau}}(k) \ \ \mathcal{O}_{\underline{\tau}}(k')]) = 2n$.

Readers may notice that the *joint observability matrix* $\mathcal{O}_{2n}(k, k') \triangleq [\mathcal{O}_{2n}(k) \ \mathcal{O}_{2n}(k')]$ equals the observability matrix of the $2n$-dimensional system defined by:

$$A(k,k') = \begin{bmatrix} A(k) & 0 \\ 0 & A(k') \end{bmatrix} \quad C(k,k') = \begin{bmatrix} C(k) & C(k') \end{bmatrix}.$$

Therefore, we can define the *joint observability index* of systems k and k' as the minimum integer $\nu(k, k')$ such that the rank of the extended joint observability matrix $\mathcal{O}_j(k,k') \triangleq [\mathcal{O}_j(k) \ \mathcal{O}_j(k')]$ stops growing. Hence, we can rephrase Lemma 1 in terms of the largest joint observability index $\nu \triangleq \max_{k \neq k'}\{\nu(k,k')\} \leq 2n$ as follows:

Corollary 1 *If* $\tau_0 \geq \nu$, *then* $\{x_{t_0}, \lambda_{t_0}, \ldots, \lambda_{t_0+\nu-1}\}$ *is observable if and only if for all* $k \neq k' \in \{1, \ldots, N\}$ *we have* $\text{rank}([\mathcal{O}_\nu(k) \ \ \mathcal{O}_\nu(k')]) = 2n$.

Remark 1 (Observability subspaces) *Notice that the* rank-$2n$ *condition implies that each linear system* $(A(k), C(k))$ *must be observable, i.e.,* $\text{rank}(\mathcal{O}_\nu(k)) = n$ *for all* $k \in \{1, \ldots, N\}$. *In addition, the rank-$2n$ condition implies that the intersection of the observability subspaces of each pair of linear systems has to be empty. In fact, the set of unobservable states can be directly obtained from the intersection of the observability subspaces. One could therefore introduce a notion of distance between models using the angles between the observability spaces, similarly to [6].*

2.2 Observability of the switching times

Corollary 1 gives conditions for the observability of $\{x_{t_0}, \lambda_{t_0}, \lambda_{t_0+1}, \ldots, \lambda_{t_0+\nu-1}\}$. We are now interested in the observability of $\{x_{t_0}, \lambda_{t_0}, \lambda_{t_0+1}, \ldots, \lambda_{t_1-1}\}$. Since $\lambda_{t_1-1} = \cdots = \lambda_{t_0}$, we only need to concentrate on the conditions under which the first transition, t_1, can be uniquely determined. We will distinguish between two different aspects of the problem:

Detection of a switch: this refers to the problem of determining whether a jump has occurred or not, given the output of the system $\{y_t\}$ on the interval $[t_0, t_0+T]$. We will derive a rank condition that guarantees the detection of a switch, either at the same instant, or *post-mortem* (*i.e.,* after it has occurred).

Observation of a switch: this refers to the problem of uniquely determining the time instant t_k at which a jump occurs, given the output of the system $\{y_t\}$ on the interval $[t_0, t_0+T]$. We will derive additional conditions on the system parameters that guarantee that a switch is recovered uniquely when it is detected *post-mortem.*

Let us first derive conditions under which a switch is immediately reflected in the measurements. We have $y_{t_1} = C(\lambda_{t_1})A(\lambda_{t_0})^{t_1-t_0}x_{t_0}$ and want to determine whether it is possible that $y_{t_1} = C(\lambda_{t_0})A(\lambda_{t_0})^{t_1-t_0}x_{t_0}$. This happens for all x_{t_0} in the null-space of $(C(\lambda_{t_1}) - C(\lambda_{t_0}))\,A(\lambda_{t_0})^{t_1-t_0}$. Therefore:

Lemma 2 *If* $x_{t_0} \in \text{Null}\,(C(\lambda_{t_1}) - C(\lambda_{t_0}))\,A(\lambda_{t_0})^{t_1-t_0}$, *then* $\{x_{t_0}, \overbrace{\lambda_{t_0}, \ldots, \lambda_{t_0}}^{t_1-t_0 \text{ times}}, \lambda_{t_1}\}$ *and* $\{x_{t_0}, \overbrace{\lambda_{t_0}, \ldots, \lambda_{t_0}}^{t_1-t_0 \text{ times}}, \lambda_{t_0}\}$, *are indistinguishable in* $[t_0, t_1]$.

We conclude from Lemma 2 that a switch can be instantaneously detected for all $x_{t_0} \in \mathbb{R}^n$ if and only if $(C(k) - C(k'))\,A(k)$ is full rank for all $k \neq k' \in \{1, \ldots, N\}$. This condition, together with those of Corollary 1, enable us to uniquely recover x_{t_0}, λ_{t_0} and t_1 from the output $\{y_t, t \in [t_0, t_1]\}$ as follows. Let

$$\mathcal{Y}_i \triangleq \begin{bmatrix} y_{t_0}^T & \ldots & y_{t_0+i-1}^T \end{bmatrix}^T = \mathcal{O}_i(\lambda_{t_0})x_{t_0}. \tag{4}$$

If $i = \nu$, equation (4) has a solution for only one $\lambda_{t_0} \in \{1, \ldots, N\}$ which is given by

$$\lambda_{t_0} = \{k : \text{rank}([\mathcal{O}_\nu(k) \ \ \mathcal{Y}_\nu]) = n\}. \tag{5}$$

Given such a λ_{t_0}, equation (4) has a solution for any $i \in [\nu, \tau_0]$ and does not have a solution for $i = \tau_0 + 1$. Therefore t_1 can be uniquely recovered as:

$$t_1 = \min\{i : \text{rank}([\mathcal{O}_i(\lambda_{t_0}) \ \ \mathcal{Y}_i]) = n+1\} + t_0 - 1. \tag{6}$$

Once t_1 has been determined, we set $\tau_0 = t_1 - t_0$ and $x_{t_0} = \mathcal{O}_{\tau_0}(\lambda_{t_0})^\dagger \mathcal{Y}_{\tau_0}$. Then we repeat the process for the remaining jumps. The only difference is that x_{t_k}, $k \geq 1$, will be given. However, since λ_{t_0} is originally unknown, we still need to check the rank-$2n$ condition of Corollary 1 for any pair of extended observability matrices in order for x_{t_0} and λ_{t_0} to be uniquely recoverable. Therefore, we have:

Theorem 1 *If* $(C(k) - C(k'))\,A(k)$ *is full rank for all* $k \neq k' \in \{1, \ldots, N\}$ *and* $\tau_k \geq \nu$ *for all* $k \geq 0$, *then* $\{x_{t_0}, \lambda_{t_0}, \ldots, \lambda_{t_0+T}\}$ *is observable if and only if for all* $k \neq k' \in \{1, \ldots, N\}$ *we have* $\text{rank}([\mathcal{O}_\nu(k) \ \mathcal{O}_\nu(k')]) = 2n$.

Let us now consider the case in which a jump occurs at time t_1, but it cannot be detected at the same time. This happens, for example, when $C(\lambda_{t_1}) = C(\lambda_{t_0})$ or $p < n$. It can be easily verified that the output of the system in $[t_1, t_1 + j - 1]$, $y_{t_1+i} = C(\lambda_{t_1})A(\lambda_{t_1})^i A(\lambda_{t_0})^{t_1-t_0} x_{t_0}$, can also be obtained as $C(\lambda_{t_0})A(\lambda_{t_0})^{t_1+i-t_0} x_{t_0}$ if and only if x_{t_0} belongs to $\mathrm{Null}\left((C(\lambda_{t_1})A(\lambda_{t_1})^i - C(\lambda_{t_0})A(\lambda_{t_0})^i)A^{t_1-t_0}(\lambda_{t_0})\right)$ for all $i = 0, 1, \dots, j-1$. This condition is compactly expressed in the following lemma.

Lemma 3 (Detection of a switch) *If x_{t_0} belongs to $\mathrm{Null}\left((\mathcal{O}_i(\lambda_{t_1}) - \mathcal{O}_i(\lambda_{t_0}))A^{t_1-t_0}(\lambda_{t_0})\right)$ for all $i = 1, \dots, j$, then $\{x_{t_0}, \overbrace{\lambda_{t_0}, \dots, \lambda_{t_0}}^{t_1+j-t_0 \text{ times}}\}$ and $\{x_{t_0}, \overbrace{\lambda_{t_0}, \dots, \lambda_{t_0}}^{t_1-t_0 \text{ times}}, \overbrace{\lambda_{t_1}, \dots, \lambda_{t_1}}^{j \text{ times}}\}$ are indistinguishable on the interval $[t_0, t_1 + j - 1]$.*

Therefore, if $(\mathcal{O}_\nu(k) - \mathcal{O}_\nu(k'))\,A(k)$ is full rank for all $k \neq k' \in \{1, \dots, N\}$, then a switch can be detected for all $x_{t_0} \in \mathbb{R}^n$ (either immediately or after it has occurred). Even though this condition guarantees that a switch is detected, it does not guarantee that it can be uniquely recovered. For instance, imagine that a jump occurs at time t_1, but it is not detected until time $t_1 + j$. Since both t_1 and $j \leq \nu$ are unknown, in general we cannot tell *when* the jump occurred. However, if we assume that $t_2 \geq t_1 + j + \nu$, since $t_1 + j$ is known, under the assumptions of Corollary 1, we can determine $(A(\lambda_{t_0}), C(\lambda_{t_0}), x_{t_0})$ and $(A(\lambda_{t_1}), C(\lambda_{t_1}), x_{t_1+j})$ uniquely from the measurements on the intervals $[t_0, t_0 + \nu - 1]$ and $[t_1 + j, t_1 + j + \nu - 1]$, respectively. Since we must have $x_{t_1+j} = A(\lambda_{t_1})^j A(\lambda_{t_0})^{t_1-t_0} x_{t_0}$, in order for t_1 to be uniquely recoverable for all $x_{t_0} \in \mathbb{R}^n$, we need that $\mathrm{rank}\left((A(\lambda_{t_1})^j - A(\lambda_{t_1})^{j'} A(\lambda_{t_0})^{j-j'})A(\lambda_{t_0})^{t_1-t_0}\right) = n$ for all $0 \leq j' \leq j-1$. Since j is unknown we need to enforce this for all $j \leq \nu$. We have shown that:

Theorem 2 *If for all $k \geq 0$ we have $\tau_k \geq 2\nu$ and for all $k \neq k' \in \{1, \dots, N\}$, $j' = 0, \dots, j-1$ and $j \leq \nu$ we have $\mathrm{rank}([\mathcal{O}_\nu(k) \quad \mathcal{O}_\nu(k')]) = 2n$, $\mathrm{rank}((\mathcal{O}_\nu(k) - \mathcal{O}_\nu(k'))\,A(k)) = n$ and $\mathrm{rank}(A(k')^j - A(k')^{j'} A(k)^{j-j'}) = n$, then $\{x_{t_0}, \lambda_{t_0}, \lambda_{t_0+1}, \dots, \lambda_{t_0+T}\}$ is observable on $[t_0, t_0 + T]$.*

3 Identification

The study of identifiability aims at answering two crucial questions in modeling time series with JLSs.

The first question pertains to uniqueness: Given measurements $\{y_t\}_{t=t_0}^{t_0+T}$, generated by a model $\Sigma = \{A_i, C_i; i = 0 \dots N-1\}$, what is the set of models $\tilde{\Sigma} = \{\tilde{A}_i, \tilde{C}_i; i = 0 \dots \tilde{N} - 1\}$ that may have produced the same output sequence $\{y_t\}_{t=t_0}^{t_0+T}$? Clearly, this question is crucial in the design of inference algorithms: if the set $\tilde{\Sigma}$ is non-empty, then any inference algorithm can converge to any points of the set. In other words, all models in the set are indistinguishable. In order for inference algorithms to converge to a unique model, it is necessary to understand the structure of the set, and elect a representative for each class. This will need imposing conditions on the structure of the models. Failing that, any inference algorithm may give dramatically different answers when operating on the *same data* depending on initialization, or on numerical errors.

The second question pertains to identifiability of the "true" model: Assuming that the model that generates the data is actually a jump linear system, under what conditions can we recover *it* from the data? We address these two questions in order in the next two sections.

3.1 Realizability

The results we are about to present reveal that, given data alone, there are infinitely many systems that generate it, which differ in the trajectories of both discrete and continuous states, and model parameters. Even more surprisingly, given *any system*, one can always find arbitrary changes of basis of the *same* system that re-create the data. In other words, given any system, one can "simulate" the jumps. This is grim news for inference algorithms, because it means that – unless appropriate conditions are enforced – one can never recover a meaningful model from data.

Consider a vector $\mathcal{Y}_T$ collecting the measurements from t_0 to $t_0 + T < \infty$. We will now show that one can explain the data with one single model, or with two models and one switch, all the way to $T + 1$ models and T switches. In fact, we can always choose (infinitely many) models $\{A_0, C_0\}$ with extended observability matrix $\mathcal{O}_0$, and initial conditions $x_0 \in \mathbb{R}^n$ with n large enough, such that $\mathcal{Y}_T = \mathcal{O}_0 x_0$. However, we can also choose $\tau_0 \leq T/2$, two systems $\{A_1, C_1\}$, $\{A_2, C_2\}$ with observability matrices $\mathcal{O}_1$ and $\mathcal{O}_2$, two initial conditions $x_1, x_2 \in \mathbb{R}^{n_1}$, and n_1 large enough, such that:

$$\mathcal{Y}_T = \mathcal{O}_0 x_0 = \begin{bmatrix} \mathcal{O}_1 x_1 \\ \mathcal{O}_2 x_2 \end{bmatrix} \tag{7}$$

where $\{A_i, C_i\}_{i=1}^2$ are chosen with respect to a basis such that $x_2 = A_1^{\tau_0} x_1$. Similarly, one can split the time series into $3, \dots, T+1$ segments, and find $\{A_i, C_i\}$ for each one.

This shows that there is an inherent ambiguity between the number of models N and their complexity n: one cannot tell, from $\mathcal{Y}_T$, how many, and how complex, the models that generated the data are. Note that, even if the maximum number of models is known and equal to the true N, one can always generate the same data with an equal or smaller number of models $\tilde{N}$ and a larger dimension of their continuous state. Similarly, even if the maximum dimension of the continuous state is known and equal to the true one n, one can always generate the same data with models of smaller or equal state dimension $\tilde{n}$, and a larger number of models.

Therefore, imposing a limit on the dimension of the state space n *or* on the number of models N allowable in a given time interval T is not enough to make the inference task well-posed and we will need to constrain both of them simultaneously. If only one model is allowed, *i.e.,* if $N = 1$, then the order n of a minimal realization must satisfy $2n \leq T + 1$. If N models are allowed, and we are given a lower bound $\overline{\tau}$ on the minimum separation between consecutive switches, then we must have $2n+2n \leq \overline{\tau}$, hence $2nN+2n(N-1) \leq T+1$, in order to obtain a minimal realization for each one of the N linear systems and identify the $N-1$ switches.

From linear systems theory, we know that the realization of a linear system is obtained up to a change of basis of the state space. Since x_{t_0} is unknown, such a change of basis is arbitrary for the first linear system. However, the choice of basis for subsequent systems is not arbitrary, but depends on the choice of the initial continuous state. Therefore we have the following:

Theorem 3 *If we are given a lower bound $\overline{\tau}$ on the minimum separation between consecutive switches, then the set of models that generate the data $\{y_t\}_{t=t_0}^{t_0+T}$ is given by $\Sigma = \{M_i A_i M_i^{-1}, C_i M_i^{-1}; i = 0 \ldots N-1\}$, where $M_i \in GL(n)$ is such that $M_{i+1} x_{t_{i+1}} = M_i A_i^{t_i} x_{t_i}$, $i = 0, \ldots N-1$, $4n \leq \overline{\tau}$ and $4nN - 2n \leq T+1$.*

The intuition behind the theorem is as follows. Since the first switch occurs after $t_0 + \overline{\tau}$, we can use the first $\overline{\tau}$ measurements to identify the dimension of the state space n, the initial state (x_{t_0}, λ_{t_0}) and the system parameters $(A(\lambda_{t_0}), C(\lambda_{t_0}))$ up to a change of basis $M_0 \in GL(n)$. Then, assuming that t_1 can be uniquely recovered, one can use the measurements in $[t_1, t_1+\overline{\tau}]$ to obtain (x_{t_1}, λ_{t_1}), $(A(\lambda_{t_1}), C(\lambda_{t_1}))$ up to a change of basis $M_1 \in GL(n)$ that must satisfy $M_1 x_{t_1} = M_0 A(\lambda_{t_0}) x_{t_0}$ and so on.

Consider now a set of data $\mathcal{Y}_T$ and *any* (observable) system $\{A, C\}$, with continuous state space of dimension n, which generates an extended observability matrix $\mathcal{O}$. For simplicity, let T be a multiple of n, $T = kn$. Then, one can find k matrices $M_i \in GL(n)$, $i = 1 \ldots k$, such that the same system, with changes of basis M_i, generates the data. This means that one can choose an arbitrary model (A, C), of arbitrary order n, and generate k models which differ only by a change of basis while generating the data. Consider, in fact, the equation $\mathcal{Y}_T = \left[(\mathcal{O}x_1)^T \ldots (\mathcal{O}x_k)^T\right]$. If $\mathcal{O}$ has full rank (which is true provided that $\{A, C\}$ is minimal), then one can always choose $x_1, \ldots, x_k$ so that the above equation is satisfied. However, given an arbitrary choice of the state-space, one can always choose $M_2, \ldots, M_k \in GL(n)$ in the following way. Let $M_2 x_2 = A^{t_2-t_1} x_1$ and define $\mathcal{O}_2 \doteq \mathcal{O} M_2^{-1}$ and consequently A_2 and C_2. Similarly one defines M_3 so that $M_3 x_3 = A_2^{t_3-t_2} A^{t_2-t_1} x_1$ from which $\mathcal{O}_3 \doteq \mathcal{O} M_3^{-1}$ and so on. Therefore, given any (finite) dataset, one can pick an arbitrary model that will explain the data, provided that the number of models one can use (all constructed from the given one just by changes of basis) is large enough.

3.2 Identifiability

In this section we address the question of identifiability: under what conditions can we recover the "true" model from data? From the discussion of the previous sections, if the model can be uniquely identified (Theorems 2 and 3) and the "true" model belongs to the class of JLSs, then the model being identified has to be the true one.

Consider now the doubly infinite Hankel matrix

$$H = \begin{bmatrix} y_{t_0} & y_{t_0+1} & \cdots \\ y_{t_0+1} & y_{t_0+2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{8}$$

and let $H_{t|i|j} \in \mathbb{R}^{ip \times j}$ be a finite-dimensional Hankel sub-matrix with upper left entry y_t, lower left entry y_{t+i-1} and upper right entry y_{t+j-1}. If the dimension of the continuous state space n and the switching times t_k and t_{k+1} were known, then the continuous state of the system $X_k = [x_{t_k} \cdots x_{t_{k+1}-1}]$ and the model parameters (A_k, C_k) could be identified up to a change of basis $M_k \in GL(n)$ from the Hankel sub-matrix generated by measurements $\{y_t\}$ corresponding to that linear system, *i.e.,* from $H_{t_k|n+1|n}$. Such a computation can be done using, for example, a simplified version of the subspace identification algorithms described in [11]. The details of the computation are as follows (we use MATLAB notation):

$$H_{t_k|n+1|n} = U_k S_k V_k^T \tag{9}$$

$$A_k = U_k(1:pn, 1:n)^\dagger U_k(p+1:p(n+1), 1:n) \tag{10}$$

$$C_k = U_k(1:p, 1:n) \tag{11}$$

$$X_k = S_k(1:n, 1:n) V_k(:, 1:n)^T. \tag{12}$$

Since in practice n and t_k are unknown, we consider Hankel sub-matrices of the form $H_{t_k|i_k|j_k}$, with $i_k \geq \bar{n}+1$, $j_k \geq \bar{n}$ and $i_k + j_k \leq \overline{\tau} + 1$, where $\bar{n}$ is a given upper bound on n. Since $H_{t_k|i_k|j_k} = \mathcal{O}_{i_k}(\lambda_{t_k})[x_{t_k} \cdots x_{t_k+j_k-1}]$ and we are looking for a pair (A_k, C_k) that is observable, we have that $\text{rank}(H_{t_k|i_k|j_k}) = n$. Since t_0 is known, the dimension of the continuous state can be obtained as $n = \text{rank}(H_{t_0|i_0|j_0})$. Given n and t_0, we can obtain x_{t_0} and (A_0, C_0) from equations (9)-(12) up to a change of basis $M_0 \in GL(n)$.

We are now interested in identifying t_1, for which we consider a Hankel sub-matrix of the form $H_{t_0|n+1|j}$ with a variable number of columns $j \geq n$. If $n \leq j < t_1 - n$, $H_{t_0|n+1|j}$ is generated by measurements corresponding to one linear system, thus $\text{rank}(H_{t_0|n+1|j}) = n$. If $j \geq t_1 - n$, the matrix $H_{t_0|n|j}$ is generated by two linear systems. Since we have assumed that the switching times are detectable, we must have $\text{rank}(H_{t_0|n+1|j}) > n$ from some j on. Letting $j^* = \min\{j : \text{rank}(H_{t_0|n+1|j}) > n\}$, we can recover t_1 uniquely from j^* as follows. Compute x_{j*} and (A_1, C_1) up to a change of basis M_1 by applying (9)-(12) to $H_{j^*|n+1|n}$. We must have that $M_1 x_{j*} = M_0 A_1^{j^*-t_1} A_0^{t_1-t_0} x_{t_0}$. Under the assumptions of Theorem 2, there is a unique t_1 satisfying such an equation. Once such a t_1 has been computed, the identification process can be repeated starting with t_1 as we did with t_0.

Finally, we emphasize that the matrices $M_k \in GL(n)$ cannot be chosen arbitrarily, since they must satisfy the following constraint:

$$M_{k+1} x_{t_{k+1}} = M_k A_k^{t_{k+1}-t_k} x_{t_k}. \tag{13}$$

Thus, one can pick M_0 arbitrarily and then determine M_{k+1} from M_k, $x_{t_{k+1}}$ and x_{t_k} as follows:

$$M_{k+1} = [M_k x_{t_{k+1}} \ \ (M_k A_k^{t_{k+1}-t_k} x_{t_k})^\perp]^\dagger [x_{t_{k+1}} \ \ (x_{t_{k+1}})^\perp],$$

where $X^\perp \in \mathbb{R}^{n-1\times n}$ is the space orthogonal to $X \in \mathbb{R}^n$.

4 Conclusions

We have presented an analysis of the observability of the continuous and discrete states of a class of linear hybrid systems, as well as a characterization of the identifiability of the model parameters. The characterization of observability and identifiability given above sheds light on the geometry of the spaces spanned by the outputs of a JLS, regardless of noise. For a given system structure, which could be for instance a generative model, the conditions tell us whether one could use any of the filtering or identification algorithms proposed in the literature or, if the conditions are not satisfied, how one should modify the model or the inference procedure. We have also characterized the classes of unidentifiable models, *i.e.,* models that produce the same outputs, and derived conditions under which the "true" model can be identified from data.

An important issue that we did not addressed is concerned with characterizing the set of observationally equivalent models. In linear systems theory, this is done elegantly by the Kalman decomposition, which partitions the state space into orthogonal subspaces. Future work will address a characterization of the set of observationally equivalent models as well as a study of the observability and identifiability conditions in the presence of noise.

References

[1] Y. Bar-Shalom and X.-R. Li. *Estimation and Tracking: Principles, Techniques, and Software.* Artech House, Boston MA, 1993.

[2] A. Bemporad, G. Ferrari, and M. Morari. Observability and controllability of piecewise affine and hybrid systems. *IEEE Transactions on Automatic Control*, 45(10):1864–1876, October 2000.

[3] A. Bemporad, J. Roll, and L. Ljung. Identification of hybrid systems via mixed-integer programming. In *Proc. of IEEE Conf. on Decision and Control*, pages 786–792, 2001.

[4] M. Billio, A. Monfort, and C.P. Robert. Bayesian estimation of switching ARMA models. *Journal of Econometrics*, (93):229–255, 1999.

[5] A. Blake, B. North, and M. Isard. Learning multi-class dynamics. *Advances in Neural Information Processing Systems*, 11:389–395, 1999. MIT Press.

[6] K. De Cock and B. De Moor. Subspace angles and distances between ARMA models. In *Proc. of Mathematical Theory of Networks and Systems*, 2000.

[7] A. Doucet, A. Logothetis, and V. Krishnamurthy. Stochastic sampling algorithms for state estimation of jump Markov linear systems. *IEEE Transactions on Automatic Control*, 45(1):188–202, 2000.

[8] J. Ezzine and A. H. Haddad. Controllability and observability of hybrid systems. *International Journal of Control*, 49(6):2045–2055, 1989.

[9] Z. Ghahramani and G. E. Hinton. Variational learning for switching state-space models. *Neural Computation*, 12(4):963–996, 1998.

[10] T. Kailath. *Linear Systems.* Prentice Hall, 1980.

[11] B. De Moor, P. Van Overschee, and J. Suykens. Subspace algorithms for system identification and stochastic realization. Technical Report ESAT-SISTA Report 1990-28, Katholieke Universiteit Leuven, 1990.

[12] K. Murphy. Switching Kalman filters. Technical report, U. C. Berkeley, 1998.

[13] V. Pavlovic, B. J. Frey, and T. S. Huang. Time series classification using mixed state dynamic Bayesian networks. In *Proc. of the Intl. Conf. on Comp. Vision and Patt. Recog.*, 1999.

[14] V. Pavlovic, J. M. Rehg, T. J. Cham, and K. P. Murphy. A dynamic Bayesian network approach to figure tracking using learned dynamic models. In *Proc. of the Intl. Conf. on Comp. Vision*, pages 94–101, 1999.

[15] V. Pavlovic, J. M. Rehg, and J. MacCormick. Learning switching linear models of human motion. In *Neural Information Processing Systems, NIPS*, 2000.

[16] A. Sun, S. S. Ge, and T. H. Lee. Controllability and reachability criteria for switched linear systems. *Automatica*, 38:775–786, 2002.

[17] F. Szigeti. A differential algebraic condition for controllability and observability of time varying linear systems. In *Proc. of IEEE Conference on Decision and Control*, pages 3088–3090, 1992.

[18] J. K. Tugnait. Adaptive estimation and identification for discrete systems with Markov jump parameters. *IEEE Transactions on Automatic Control*, 27(5):1054–1065, 1982.

[19] J. K. Tugnait. Detection and estimation for abruptly changing systems. *Automatica*, 18(5):607–615, 1982.