Observability of Switched Linear Systems

Mohamed Babaali and Magnus Egerstedt

Electrical and Computer Engineering, Georgia Institute of Technology,
Atlanta, GA 30332, USA
{babaali,magnus}@ece.gatech.edu

Abstract. The observability of deterministic, discrete-time, switched, linear systems is considered. Depending on whether or not the modes are observed, and on whether the continuous state or the mode sequence is to be recovered, several observability concepts are defined, characterized through linear algebraic tests, and their decidability assessed.

1 Introduction

By switched linear systems (SLS), we refer to discrete-time systems that can be modeled as follows:

$$\begin{aligned} x_{k+1} &= A(\theta_k)x_k + B(\theta_k)u_k \\ y_k &= C(\theta_k)x_k, \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, and $y_k \in \mathbb{R}^p$ are the states, the inputs and the measurements, respectively. θ_k, which we refer to as the mode at time k, assumes its values in the set $\{1, \dots, s\}$, so that the system parameter matrices $A(\theta_k)$, $B(\theta_k)$, and $C(\theta_k)$ switch among s different known matrices. We assume that the mode sequence $\{\theta_k\}_{k=1}^{\infty}$, whether known or unknown, is arbitrary and independent of the initial state and inputs. In particular, we impose no constraints on the time separation between two consecutive switches, and we assume that the switches are not triggered by state space based events.

By observability, we mean the ability to infer the initial state x_1, and possibly a finite portion of the mode sequence (when unobserved), from a finite number of measurements $y_1, \dots, y_N$. While the concept of observability has a simple well-known characterization in classical linear systems, it has been associated with several notions in the SLS literature. Indeed, the fact that the mode sequence may or may not be observed, and, in the latter case, that one may or may not wish to recover it along with the state, makes for the need to consider several different problems, thus different definitions and characterizations. In this paper, our aim is to introduce and define several different concepts of observability in SLS's, to characterize them, and to assess their main properties, among which decidability is of special importance.

While we are concerned with the deterministic case, the first attempts at defining observability for SLS were motivated by stochastic optimal control and optimal estimation problems [7–9, 11], and are therefore of limited relevance to the present paper. In parallel with those works, the continuous-time case [5], and a special discrete-time class [6] were studied, both assuming observed modes. The recent surge of interest in hybrid systems expressed by the control and computer science communities has motivated renewed research efforts towards the SLS observability problem under unobserved modes. Particularly, the work in [3] concerned piece-wise affine systems, in which the future mode depends on the current state. In [2], two routes were proposed for recovering the modes: the first through a finite automaton formalism (which uses extra sources of information, beside the measurements, that we, in this paper, assume to be unavailable), and the second through failure detection techniques, which require the mode sequence to be slowly-varying. Finally, in [10], a systematic attempt at carefully defining observability for SLS under unobserved modes was made, from which some results in this paper draw inspiration. In the observed modes case, observability was characterized and its decidability proven in [1].

The outline of this paper is as follows. In Section 2, the autonomous case is studied. In Section 3, we discuss the non-autonomous case.

2 Autonomous Systems

We restrict our attention, for now, to autonomous systems of the form:

$$\begin{aligned} x_{k+1} &= A(\theta_k)x_k \\ y_k &= C(\theta_k)x_k, \end{aligned} \tag{2}$$

obtained simply by removing the $B(\theta_k)u_k$ term from (1). Before going any further, we need a few definitions. Since we are dealing with finite-time problems, we abandon the mode sequence notation and we define a path θ as a finite sequence (or string) of modes $\theta = \theta_1\theta_2 \dots \theta_N$, where N is the path length denoted by $|\theta|$. We also define Θ_N as the set of all paths of length N. Moreover, we denote by $\theta_{[i,j]}$ the infix of θ between i and j, i.e. $\theta_{[i,j]} = \theta_i\theta_{i+1}\dots\theta_j$, we use $\theta\theta'$ to denote the concatenation of θ with θ', and we let $\Phi(\theta) \triangleq A(\theta_N)\cdots A(\theta_1)$ denote the transition matrix of a path θ. By convention, we let $\theta_{[i,i-1]} = \epsilon$, the empty word, and $\Phi(\epsilon) = I$. We next define the observability matrix $\mathcal{O}(\theta)$ of a path θ as:

$$\mathcal{O}(\theta) \triangleq \begin{pmatrix} C(\theta_1) \\ \vdots \\ C(\theta_N)A(\theta_{N-1})\cdots A(\theta_1) \end{pmatrix}. \tag{3}$$

Finally, we define:

$$Y(\theta, x) \triangleq \mathcal{O}(\theta)x, \tag{4}$$

and we then get that if $x = x_1$ and $\theta = \theta_1\theta_2\dots\theta_N$ in (2), then $Y(\theta, x) = (y_1^T \ \dots \ y_N^T)^T$. Therefore, throughout the remainder of this section, we will use (4) to describe (2) in a more compact way.

In this section, we will define and characterize several concepts of observability for autonomous SLS (2). The observed modes case is considered in Section 2.1, and the unobserved modes case in Section 2.2.

2.1 The Observed Modes Case

We assume that both the mode path θ and $Y(\theta, x)$ are available, and we wish to recover the state x uniquely. More precisely, we study the existence of an integer N such that the map $(x, \theta) \mapsto (Y(\theta, x), \theta)$ is injective or, in other words, that

$$x \neq x' \Rightarrow Y(\theta, x) \neq Y(\theta, x') \tag{5}$$

for every path θ of length N. Since $Y(\theta, x) = \mathcal{O}(\theta)x$, this clearly amounts to requiring that $\mathcal{O}(\theta)$ be of full column rank (i.e. that $\rho(\mathcal{O}(\theta)) = n$, where $\rho(\cdot)$ denotes the rank function, and we then say θ is *observable*), in which case, letting $\mathcal{O}(\theta)^{\{1\}}$ be a $\{1\}$-inverse of $\mathcal{O}(\theta)$ (see Appendix A), we get

$$x = \mathcal{O}(\theta)^{\{1\}} Y(\theta, x). \tag{6}$$

The existence of such an integer N was defined in [1] as *pathwise observability*:

Definition 1 (Pathwise Observability (PWO)). *The SLS (2) is pathwise observable (PWO) if there exists an integer N such that every path θ of length N is observable, i.e. satisfies $\rho(\mathcal{O}(\theta)) = n$. We refer to the smallest such integer N as the index of PWO.* $\Diamond$

PWO was moreover shown to be decidable. More precisely, the index of pathwise observability was shown to be bounded above by numbers $\mathcal{N}(s, n)$ (actually, $\mathcal{N}(s, n, n)$ in [1]) depending only on s, the number of modes (or pairs) of the system, and n. Even though the numbers $\mathcal{N}(s, n)$ were not given explicitly, they were shown to be smaller than numbers $N(s, n)$ (again, $N(s, n, n)$ in [1]) that, moreover, satisfy (we let $\mathcal{R}(M)$ denote the row range space of a matrix M):

Theorem 1 *If θ is a path of length $N(s, n)$, then there exists a prefix θ^0 of θ (i.e. $\theta = \theta^0\theta^1$ for some θ^1) and a path θ' of arbitrary length such that*

$$\mathcal{R}(\mathcal{O}(\theta^0\theta')) \subset \mathcal{R}(\mathcal{O}(\theta^0)), \tag{7}$$

and thus $\rho(\mathcal{O}(\theta^0\theta')) = \rho(\mathcal{O}(\theta^0)) \leq \rho(\mathcal{O}(\theta))$. $\Diamond$

Even though this theorem was not explicitly stated in [1], its proof can easily be derived from that of [1, Theorem 1].

2.2 The Unobserved Modes Case: Mode Observability

In this section, we assume that only the continuous measurements $Y(\theta, x)$ are available, and we investigate the possibility to infer a prefix of the path θ (i.e. $\theta_{[1,N']}$ for some $N' < |\theta|$) from the successive measurements $Y(\theta, x)$ only. But first, noting that when $x = 0$, $Y(\theta, x) = 0$ for any path θ, we observe that it is *impossible* to distinguish between paths whenever $x = 0$. As it turns out, this happens in general for all states in a union of subspaces of $\mathbb{R}^n$. Moreover, this issue is closely related to false alarms in failure detection, as pointed out in [2]. We therefore have to consider the problem from a looser point of view, which leads us to the following definition, in which *a.e.* x stands for "for almost every x", by which we mean for all $x \in \mathbb{R}^n$ but a union of proper subspaces, thus for all x but a set of Lebesgue measure 0:

Definition 2 (Mode Observability (MO)). *The SLS (2) is MO at N if there exists an integer N' such that for all $\theta \in \Theta_{N+N'}$ and for a.e. $x \in \mathbb{R}^n$,*

$$\theta_{[1,N]} \neq \theta'_{[1,N]} \Rightarrow Y(\theta, x) \neq Y(\theta', x') \ \forall \ x' \in \mathbb{R}^n. \tag{8}$$

The index of MO at N is the smallest such N'. ♢

In other words, we require the possibility to recover the first N modes (i.e. $\theta_{[1,N]}$) uniquely whenever $N + N'$ measurements (i.e. $Y(\theta, x)$) are available, and for a.e. state x. To this end, we need a way to discern between the paths θ using the measurements $Y(\theta, x)$ they produce through $Y(\theta, x) = \mathcal{O}(\theta)x$. As we are about to show, the only way to achieve that without any information other than the available measurement $Y(\theta, x)$ is by taking advantage of the following inclusion, immediate from $Y(\theta, x) = \mathcal{O}(\theta)x$:

$$Y(\theta, x) \in \Re(\mathcal{O}(\theta)), \tag{9}$$

where $\Re(M)$ denotes the column range space of the matrix M. The question is then whether $\theta' \neq \theta \ \Rightarrow \ Y(\theta, x) \notin \Re(\mathcal{O}(\theta'))$, which would provide us with a simple procedure for recovering a path from the measurements (using the range inclusion test, see Appendix A):

$$\theta = \arg_{\theta' \in \Theta_N} \{Y(\theta, x) \in \Re(\mathcal{O}(\theta'))\} \tag{10}$$

The main issue lies in whether the test (10) has a unique solution. In order to analyze this, we introduce the concept of *discernibility*:

Definition 3 (Discernibility). *A path θ is discernible from another path θ' of the same length if*

$$\rho([\mathcal{O}(\theta)\mathcal{O}(\theta')]) > \rho(\mathcal{O}(\theta')), \tag{11}$$

where $[\mathcal{O}(\theta)\mathcal{O}(\theta')]$ denotes the horizontal concatenation of $\mathcal{O}(\theta)$ and $\mathcal{O}(\theta')$, and where the degree d of discernibility is defined as

$$d = \rho([\mathcal{O}(\theta)\mathcal{O}(\theta')]) - \rho(\mathcal{O}(\theta')). \tag{12}$$

We then say that θ is d-discernible from θ'. ♢

The following proposition is now in order:

Proposition 1 *$Y(\theta, x) \notin \Re(\mathcal{O}(\theta'))$ for almost any $x \in \mathbb{R}^n$ iff θ is discernible from θ'.* ♢

Proof: We have $Y(\theta, x) \in \Re(\mathcal{O}(\theta'))$ iff $Y(\theta, x)$ also lies in the the *output subspace of conflict of θ and θ'*, defined as:

$$C(\theta, \theta') \triangleq \Re(\mathcal{O}(\theta)) \cap \Re(\mathcal{O}(\theta')). \tag{13}$$

We therefore need to show that the dimension of the inverse image of $C(\theta, \theta')$ by $\mathcal{O}(\theta)$, $c(\theta, \theta') \triangleq \mathcal{O}(\theta)^{-1}(C(\theta, \theta'))$, which we refer to as the *input subspace of* *conflict of* θ *with* θ', is smaller than n (which implies that its Lebesgue measure is 0) if and only if θ is discernible from θ'. We have:

$$\dim(C(\theta,\theta')) = \rho(\mathcal{O}(\theta)) + \rho(\mathcal{O}(\theta')) - \rho([\mathcal{O}(\theta)\mathcal{O}(\theta')]). \tag{14}$$

Noting that $\dim(c(\theta,\theta')) = \dim(C(\theta,\theta')) + \dim\ker(\mathcal{O}(\theta))$, and then recalling that $\rho(\mathcal{O}(\theta)) + \dim\ker(\mathcal{O}(\theta)) = n$, we get

$$\dim(c(\theta,\theta')) = n - \rho([\mathcal{O}(\theta)\mathcal{O}(\theta')]) + \rho(\mathcal{O}(\theta')). \tag{15}$$

Therefore, by definition of discernibility, we see that $\dim(c(\theta,\theta')) < n$ if and only if θ is discernible from θ', in which case we moreover have

$$\dim(c(\theta,\theta')) = n - d, \tag{16}$$

where d is the degree of discernibility defined in (12). □

Remarks 1

- *Discernibility does not imply that either path is observable (see Example 1).*
- *From (16), the degree of discernibility appears as a measure of* separation *between the two paths* θ *and* θ'*: the larger the degree* d*, the smaller the dimension of the input subspace of conflict. It is clear that the maximum value for* d *is* n*, in which case the input subspace of conflict is trivial.*
- *Note that, as we have defined it, discernibility is not symmetric.*
- *If* $\theta_{[1,N-1]} = \theta'_{[1,N-1]}$*, i.e. if the two paths only differ by their last value, then their index of discernibility is bounded by* p*.* ♢

The last remark raises the question of whether an upper bound (p, the size of each measurement y_k) is imposed on the maximum degree of discernibility that can be guaranteed for all pairs of paths of a certain length. A related limitation was pointed out in [10], where p had to equal n, since the criterion considered for "switch detection" was similar to n-discernibility. It turns out that this limitation can be overcome, provided one can use further measurements in order to discern the paths, which leads us to the idea of *forward discernibility*:

Definition 4 (Forward Discernibility (FD)). *Given an integer* $d > 0$*, a path* θ *is forward* d*-discernible (*d*-FD) from another path* θ' *of the same length if there exists an integer* N_d *such that for any pair of paths* λ *and* λ' *of length* N_d*,* $\theta\lambda$ *and* $\theta'\lambda'$ *are discernible with degree at least* d*. The smallest such integer* N_d *is the index of* d*-FD of* θ *from* θ'*.* ♢

Proposition 2 $Y(\theta\lambda, x) \notin \Re(\mathcal{O}(\theta'\lambda'))$ *for all* $\lambda, \lambda' \in \Theta_{N'}$ *and for almost any* $x \in \mathbb{R}^n$ *iff* θ *is FD (i.e. 1-FD) from* θ' *with an index no larger than* N'*.* ♢

Proof: Clearly, the set $\{x \in \mathbb{R}^n \mid \exists \lambda, \lambda' \in \Theta_{N'},\ Y(\theta\lambda, x) \in \Re(\mathcal{O}(\theta'\lambda'))\}$ equals $\bigcup_{\lambda,\lambda'} c(\theta\lambda, \theta'\lambda')$, which, by Proposition 1 and by virtue of the fact that a finite union of null sets is a null set, has measure 0 iff θ is FD from θ'. □

We now turn to showing that forward discernibility is decidable. We first establish the following lemma, which indicates that the indexes of d-FD increase with d, which is not an obvious fact.

Lemma 1 *Let θ and θ' be two different paths of length N, and λ and λ' be any paths of length N'. The degree of discernibility of $\theta\lambda$ from $\theta'\lambda'$ is greater than or equal to the degree of discernibility of θ from θ'. In other words, the degree of discernibility is nondecreasing as the length increases.* $\Diamond$

Proof: It is easily shown, by elementary linear algebra, that

$$\rho([\mathcal{O}(\theta\lambda)\mathcal{O}(\theta'\lambda')]) - \rho([\mathcal{O}(\theta)\mathcal{O}(\theta')]) \geq \rho(\mathcal{O}(\theta\lambda)) - \rho(\mathcal{O}(\theta)). \tag{17}$$

In other words, the rank of the concatenation must increase by at least the increase of each path. $\square$

Theorem 2 (Decidability of Forward Discernibility) *FD is decidable for any degree, as the index of d-FD, where d is the maximum degree of FD between some pair of paths, is smaller than or equal to $N(s^2, 2n)$.* $\Diamond$

Proof: Fix θ and θ', and let λ and λ' be such that the degree of discernibility of $\theta\lambda$ from $\theta'\lambda'$ is minimal over all pairs of paths λ and λ' of length $N(s^2, 2n)$.

First, note that the matrices $[\mathcal{O}(\lambda)\mathcal{O}(\lambda')]$ are produced by the following set of s^2 pairs:

$$\begin{pmatrix} A(i) & 0 \\ 0 & A(j) \end{pmatrix} \qquad \begin{pmatrix} C(i) \; C(j) \end{pmatrix} \qquad (i,j) \in \{1, \ldots, s\}^2. \tag{18}$$

Therefore, by Theorem 1, there exist λ^0 and λ'^0, respective prefixes of λ and λ' of the same length, and two paths μ and μ' of the same, arbitrary, length, such that $\mathcal{R}([\mathcal{O}(\lambda^0\mu)\mathcal{O}(\lambda'^0\mu')]) \subset \mathcal{R}([\mathcal{O}(\lambda^0)\mathcal{O}(\lambda'^0)])$, which, by [1, Lemma 4] and upon some manipulation, implies that

$$\mathcal{R}([\mathcal{O}(\theta\lambda^0\mu)\mathcal{O}(\theta'\lambda'^0\mu')]) \subset \mathcal{R}([\mathcal{O}(\theta\lambda^0)\mathcal{O}(\theta'\lambda'^0)]). \tag{19}$$

By Lemma 1, Equation (19) implies that the degree of discernibility of $\theta\lambda^0\mu$ from $\theta'\lambda'^0\mu'$ is equal to that of $\theta\lambda^0$ from $\theta'\lambda'^0$, which, again by Lemma 1, is smaller than that of $\theta\lambda$ from $\theta'\lambda'$, which completes the proof since μ and μ' are of arbitrary length. $\square$

Before establishing the main result of this section characterizing mode observability, we need the following definition:

Definition 5 (Complete Forward Discernibility (CFD)). *Given an integer $d > 0$, a path θ is completely forward d-discernible (d-CFD) if it is d-FD from every other path θ' of the same length. The index of d-CFD of θ is the maximum index of d-FD, over all $\theta' \neq \theta$, of θ from θ'.* $\Diamond$

Theorem 3 *The SLS (2) is MO at N iff every path of length N is CFD, (i.e. 1-CFD). Moreover, the index of MO is the largest index of CFD, over all paths of length N.* $\Diamond$

Proof: Clearly, the set $\{x \in \mathbb{R}^n \mid \exists\theta,\theta' \in \Theta_N,\ \exists\lambda,\lambda' \in \Theta_{N'},\ Y(\theta\lambda, x) \in \Re(\mathcal{O}(\theta'\lambda'))\}$ equals $\bigcup_{\theta,\theta'}\bigcup_{\lambda,\lambda'} c(\theta\lambda, \theta'\lambda')$. Therefore, by Proposition 2, and by virtue of the fact that a finite union of null sets is a null set, it has measure 0 iff every θ is CFD with index at most N'. $\square$

We now complete our study of mode observability in autonomous systems by answering the following two questions:

– what effect does N have on MO? In other words, is MO at larger N stronger or weaker?
– Is MO decidable?

The following proposition answers the first question:

Proposition 3 *If a system is MO at N, then it is MO at any $M \leq N$.* $\diamondsuit$

Proof: Let N' be the index of MO at N. Then for every pair of paths θ, θ' of length $N + N'$ with a switch at or before N (i.e. such that $\theta_i \neq \theta'_i$ for some $i \leq N$), θ must be discernible from θ'. But, since $M \leq N$, this implies the same whenever a switch occurs at or before M, which implies MO at M with an index smaller than or equal to $N' + (N - M)$. $\square$

The converse is unfortunately not true, unless the A matrices are all invertible (a counterexample is given next):

Proposition 4 *If $A(1), \ldots, A(s)$ are all invertible, then MO at 1 implies MO at any positive integer N.* $\diamondsuit$

Proof: Let θ and θ' be two different paths of length N, and assume that the maximum index of FD over all pairs of different modes (i.e. paths of length 1) is N'. It suffices to show that θ is FD from θ' with index at most N'. Let λ and λ' be any two paths of length N'. It is easy to show that

$$c(\theta\lambda, \theta'\lambda') = \bigcap_{i=1}^{N} \phi(\theta_{[1,i-1]})^{-1} c(\theta_{[i,N]}\lambda, \theta'_{[i,N]}\lambda'). \tag{20}$$

Since $\theta \neq \theta'$, there exists $i \leq n$ such that $\theta_i \neq \theta'_i$. Note that $c(\theta_{[i,N]}\lambda, \theta'_{[i,N]}\lambda') = c(\theta_i\mu, \theta'_i\mu')$ with $\mu = \theta_{[i+1,N]}\lambda$, $\mu' = \theta'_{[i+1,N]}\lambda'$ and $|\mu| = |\mu'| \geq N'$. Therefore, since, by assumption, θ_i is FD from θ'_i with index at most N', $\dim(c(\theta_{[i,N]}\lambda, \theta'_{[i,N]}\lambda')) < n$, and since all the A matrices, and thus $\phi(\theta_{[1,i-1]})$, are invertible, and using (20), we finally get $\dim(c(\theta\lambda, \theta'\lambda')) < n$, which completes the proof. $\square$

Example 1. Consider

$$\begin{aligned} A(1) &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \; A(2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ C(1) &= \begin{pmatrix} 1 & 0 \end{pmatrix} \; C(2) = \begin{pmatrix} 1 & 0 \end{pmatrix}. \end{aligned} \tag{21}$$

The paths 1 and 2 (of length 1) are mutually FD with index 1, but the paths $1 \cdot 1$ and $1 \cdot 2$ are not, because they are not discernible and $A(1) = 0$, which prevents further measurements from increasing their discernibility. $\triangle$

And finally,

Theorem 4 *MO at any index N is decidable.* $\diamondsuit$

Proof: Since the number of paths of length N is finite, and since, by Theorem 2, FD is decidable, it follows that CFD is decidable, and thus that MO is decidable as well. □

Note that more precise versions of these last 3 results can be obtained, provided one extends the notion of *degree* to MO.

2.3 The Unobserved Modes Case: State Observability

In this section, we are concerned with whether the continuous state x only is recoverable. From the previous results, we know that if a system is PWO (with index N_{pwo}) and MO at N_{pwo} with index N_{mo}, then one can recover the state x uniquely from $Y(\theta, x)$ for all θ of length $N_{pwo} + N_{mo}$ and for almost any x. But if we do not need θ (which is the primary reason behind the "a.e."), is this still the best we can do? It turns out that we can do better, in that we can sometimes recover *all* states x uniquely, for all paths θ of a certain length. For now, we define our concept of *state observability*:

Definition 6 (State Observability (SO)). *The SLS (2) is SO if there exists an integer N (the smallest being the index) such that $\forall x \in \mathbb{R}^n$ and $\forall \theta \in \Theta_N$,*

$$x \neq x' \Rightarrow Y(\theta, x) \neq Y(\theta', x') \; \forall \; \theta' \in \Theta_N \tag{22}$$

In other words, a system is SO if any N consecutive measurements $Y(\theta, x)$ yield x uniquely without knowledge of θ, i.e. if the map $(x, \theta) \mapsto Y(\theta, x)$ is injective in its first coordinate. We first establish a sufficient condition.

Proposition 5 *If a system is PWO with index N_{pwo}, and if every path of length N_{pwo} is n-CFD, then it is SO.* ♢

Proof: Let N_{cfd} be the maximum index of n-CFD, and $N = N_{pwo} + N_{cfd}$. This implies that the dimension of the input subspaces of conflict of any two paths of length N satisfying $\theta_{[1,N_{PWO}]} \neq \theta'_{[1,N_{PWO}]}$ is 0: 0 is therefore the only state whose measurements $Y(\theta, x)$ do not yield $\theta_{[1,N_{PWO}]}$ unambiguously. We then have two cases:

- $x \neq 0$, in which case the range inclusion test yields $\theta_{[1,N_{PWO}]}$, which can then be used in $x = \mathcal{O}(\theta_{[1,N_{PWO}]})^{\{1\}} Y(\theta, x)$.
- $x = 0$, in which case $Y(\theta, x) = 0$. By pathwise observability, we then know that $x = 0$. □

Example 2. As an example, here is a system satisfying the conditions of Proposition 5:

$$\begin{aligned} A(1) &= \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \; A(2) = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} \\ C(1) &= \begin{pmatrix} 3 & 0 \end{pmatrix} \; C(2) = \begin{pmatrix} 2 & 0 \end{pmatrix}, \end{aligned} \tag{23}$$

Here, $N_{pwo} = 2$ and $N_{cfd} = 2$, and it is easy to check that the rank of $[\mathcal{O}(\theta\lambda)\mathcal{O}(\theta'\lambda')]$ equals 4 for any pair θ, θ' of different paths of length 2 and any pair λ, λ' of paths of length 2. △

It turns out that the conditions given in Proposition 5 are not necessary (we will later give a counterexample). In order to study SO further, we introduce the concept of *joint observability*:

Definition 7 (Joint Observability (JO)). *Two different paths θ and θ' of the same length are jointly observable (JO) if they are both observable, and if their left inverses agree on $C(\theta, \theta')$, i.e.*[1]

$$(\mathcal{O}(\theta)^{\{1\}} - \mathcal{O}(\theta')^{\{1\}})P_{C(\theta,\theta')} = 0, \tag{24}$$

or equivalently,

$$(\mathcal{O}(\theta) - \mathcal{O}(\theta'))P_{c(\theta,\theta')} = 0 \textit{ and } (\mathcal{O}(\theta) - \mathcal{O}(\theta'))P_{c(\theta',\theta)} = 0. \tag{25}$$

Note that, as opposed to discernibility, joint observability is symmetric. A direct consequence of this definition is:

Proposition 6 *θ and θ' are JO iff for all $x, x' \in \mathbb{R}^n$,*

$$x \neq x' \Rightarrow Y(\theta, x) \neq Y(\theta', x'). \tag{26}$$

We also need to define *forward joint observability*:

Definition 8 (Forward Joint Observability (FJO)). *Two different observable paths θ and θ' of the same length are forward jointly observable (FJO) if there exists an integer N such that for all λ and λ' of length N, $\theta\lambda$ and $\theta'\lambda'$ are JO. The index of FJO is the smallest such integer.* ♢

Before characterizing SO, we next show that FJO is decidable.

Theorem 5 *FJO is decidable, as the index of JO is bounded by $N(s^2, 2n)$.* ♢

Proof: Suppose that θ and θ' are observable and that there exist λ and λ' of length $N(s^2, 2n)$ such that $\theta\lambda$ and $\theta'\lambda'$ are not JO. Similarly as in the proof of Theorem 2, we can find λ^0 and λ'^0, respective prefixes of λ and λ' of the same length, and two paths μ and μ' of the same, arbitrary, length, such that

$$\mathcal{R}([\mathcal{O}(\theta\lambda^0\mu)\mathcal{O}(\theta'\lambda'^0\mu')]) \subset \mathcal{R}([\mathcal{O}(\theta\lambda^0)\mathcal{O}(\theta'\lambda'^0)]). \tag{27}$$

Now, since $\theta\lambda$ and $\theta'\lambda'$ are not JO, neither can be $\theta\lambda^0$ and $\theta'\lambda'^0$, since $Y(\theta\lambda, x) = Y(\theta'\lambda', x')$ implies $Y(\theta\lambda^0, x) = Y(\theta'\lambda'^0, x')$.

Moreover, by Lemma 1, Equation (27) implies that the degree of discernibility of $\theta\lambda^0\mu$ from $\theta'\lambda'^0\mu'$ equals that of $\theta\lambda^0$ from $\theta'\lambda'^0$, which furthermore implies that $c(\theta\lambda^0\mu, \theta'\lambda'^0\mu') = c(\theta\lambda^0, \theta'\lambda'^0)$, thus that $(\mathcal{O}(\theta\lambda^0\mu) - \mathcal{O}(\theta'\lambda'^0\mu'))P_{c(\theta\lambda^0\mu,\theta'\lambda'^0\mu')}$ equals $(\mathcal{O}(\theta\lambda^0\mu) - \mathcal{O}(\theta'\lambda'^0\mu'))P_{c(\theta\lambda^0,\theta'\lambda'^0)}$ and cannot equal zero since its submatrix $(\mathcal{O}(\theta\lambda^0) - \mathcal{O}(\theta'\lambda'^0))P_{c(\theta\lambda^0,\theta'\lambda'^0)}$ is not, because, as we have just shown, $\theta\lambda^0$ and $\theta'\lambda'^0$ are not JO. Therefore, $\theta\lambda^0\mu$ and $\theta'\lambda'^0\mu'$ are not JO, which completes the proof since μ and μ' are of arbitrary length. □

We now characterize SO:

[1] Given a subspace V, we let P_V denote the matrix of a linear projection on V.

Theorem 6 *The following are equivalent.*

1. *The SLS (2) is SO.*
2. *The SLS (2) is PWO with index N_{pwo}, and every pair of different paths of length N_{pwo} is FJO.*
3. *The SLS (2) is PWO, and every* minimally observable *path (i.e. a path with no observable prefix) is FJO with every other observable path of the same length.* ◊

Proof:

$2 \Rightarrow 1$: Let N_{fjo} be be the largest index of FJO over all pairs of paths of length N_{pwo}. Let us show that the system is SO with index at most $N = N_{pwo} + N_{fjo}$. Fix a path θ of length N, and suppose that θ' is such that $Y(\theta, x) = Y(\theta', x')$. Let $\theta_{[1,k]}$ be the minimally observable prefix of θ. First, if $\theta'_{[1,k]} = \theta_{[1,k]}$, then $x = x'$ by observability of $\theta_{[1,k]}$, since $Y(\theta, x) = Y(\theta', x')$ implies $Y(\theta_{[1,k]}, x) = Y(\theta'_{[1,k]}, x') = Y(\theta_{[1,k]}, x')$. On the other hand, if $\theta'_{[1,k]} \neq \theta_{[1,k]}$, then since $k \leq N_{pwo}$, it is easy to show that $\theta'_{[1,k]}$ and $\theta_{[1,k]}$ are FJO with index at most $N - k$. Proposition 6 then concludes that $x = x'$.

$3 \Rightarrow 2$: It is easily seen that the only pairs of paths of length N_{pwo} left to check for FJO are those sharing the same minimally observable prefix. Let θ and θ' be two paths of length N_{pwo}, and let $\theta'_{[1,k]} = \theta_{[1,k]}$ be their minimally observable prefix. $Y(\theta, x) = Y(\theta', x')$, which implies $Y(\theta_{[1,k]}, x) = Y(\theta'_{[1,k]}, x') = Y(\theta_{[1,k]}, x')$, implies that $x = x'$ by observability of $\theta_{[1,k]}$. θ and θ' are therefore JO, thus FJO.

$1 \Rightarrow 3$: Necessity of PWO to SO is obvious. Suppose that a minimally observable path is not FJO with another observable path, i.e. that there exist λ, λ' of arbitrary length such that $\theta\lambda$ and $\theta\lambda'$ are not JO, which, by proposition 6, implies the existence of $x \neq x'$ such that $Y(\theta\lambda, x) = Y(\theta'\lambda', x')$, which contradicts SO. □

The reason we give two characterizations is that their equivalence is not obvious, and because the second one is easier to check, since the number of minimally observable paths is in general smaller than $s^{N_{pwo}}$. Moreover, it is, in a sense, much tighter, since two paths can be non FJO only if they do not share the same minimally observable path. Finally,

Theorem 7 *SO is decidable.* ◊

Proof: PWO is decidable. Since FJO is decidable, and since there is a finite number of paths of length N_{pwo}, the first characterization of Theorem 6 concludes. □

We now give an example of an SO system that does not satisfy the requirements of Proposition 5:

Example 3. Let

$$\begin{array}{ll} A(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & A(2) = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \\ C(1) = \begin{pmatrix} 1 & 0 \end{pmatrix} & C(2) = \begin{pmatrix} 1 & 0 \end{pmatrix}, \end{array} \tag{28}$$

This system is PWO with index 2, and any paths of length 2 are FJO with index 1. For instance, letting $\theta = 11$, $\theta' = 22$, and $\lambda = \lambda' = 1$, one gets

$$\mathcal{O}(\theta\lambda) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \mathcal{O}(\theta'\lambda') = \begin{pmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 8 \end{pmatrix}, \tag{29}$$

hence that $\theta\lambda$ and $\theta'\lambda'$ are JO because the first columns of their observability matrices, which span $C(\theta, \theta')$, are equal. Thus if we measure $Y(\mu, x) = (\alpha\ \alpha\ \alpha)^T$, then the initial state can only be $(\alpha\ 0)^T$, regardless of the path μ. △

3 Non-Autonomous Systems

We now return to the general non-autonomous case, and recall our model:

$$\begin{aligned} x_{k+1} &= A(\theta_k)x_k + B(\theta_k)u_k \\ y_k &= C(\theta_k)x_k. \end{aligned} \tag{30}$$

Our aim here is to extend some of the previous analysis to the system in (30). We thus first define, for a path θ of length N:

$$\Gamma(\theta) \triangleq \begin{pmatrix} 0 & \cdots & 0 & 0 \\ C(\theta_2)B(\theta_1) & \cdots & 0 & 0 \\ C(\theta_3)A(\theta_2)B(\theta_1) & \cdots & \vdots & 0 \\ \vdots & \cdots & 0 & \vdots \\ C(\theta_N)\Phi(\theta_{[2,N]})B(\theta_1) & \cdots & C(\theta_N)B(\theta_{N-1}) & 0 \end{pmatrix},$$

which enables us to further define:

$$Y(\theta, x, U) \triangleq \mathcal{O}(\theta)x + \Gamma(\theta)U, \tag{31}$$

where U is a control vector in $\mathbb{R}^{mN}$. Again, if $x = x_1$, $\theta = \theta_1 \cdots \theta_N$, and $U = (u_1^T\ \dots\ u_N^T)^T$ in (30), then $Y(\theta, x, U) = (y_1^T\ \dots\ y_N^T)^T$, and we can concentrate on equation (31). Now, by the separation principle for linear time-varying systems, it turns out that we do not need to repeat the analysis of the known modes case, and pathwise observability remains necessary and sufficient for state observability in finite time. In this section, we will instead begin directly by taking a first look at mode observability.

Given θ and θ', our objective in the autonomous case has been, roughly speaking, to make the intersection $C(\theta, \theta')$ of $\Re(\mathcal{O}(\theta))$ with $\Re(\mathcal{O}(\theta'))$ as small as possible. Here, Equation (31) suggests that we should rather consider the intersection of the affine subspaces $\Re(\mathcal{O}(\theta)) + \Gamma(\theta)U$ and $\Re(\mathcal{O}(\theta')) + \Gamma(\theta')U$ ($V + v$, where V is a subspace and v a vector of $\mathbb{R}^n$, denotes the affine subspace $\{x + v \mid x \in V\}$), and study what effect U has on it. Recalling the following classic theorem,

Theorem 8 *The intersection of $V+v$ and $V'+v'$ is either empty or equal to $V \cap V' + w$ for some w, in which case it has the dimension of $V \cap V'$.* ♢

we realize that, while the $\Gamma(\theta)U$ terms cannot increase the degree of discernibility, they can achieve something impossible in the non-autonomous case: they can render the output affine subspaces of θ and θ', i.e. $\Re(\mathcal{O}(\theta)) + \Gamma(\theta)U$ and $\Re(\mathcal{O}(\theta')) + \Gamma(\theta')U$, totally disjoint, which motivates the following definition:

Definition 9 (Strong Mode Observability (SMO)). *The SLS (30) is strongly mode observable (SMO) at N if there exists an integer N' and a vector U such that for all $x \in \mathbb{R}^n$ and all $\theta \in \Theta_{N+N'}$,*

$$\theta_{[1,N]} \neq \theta'_{[1,N]} \Rightarrow Y(\theta, x, U) \neq Y(\theta', x', U) \ \forall \ x' \in \mathbb{R}^n \tag{32}$$

We refer to such a vector U as a discerning control. ♢

Note that the difference lies in the replacement of "a.e. x" with "$\forall$ x", which is a stronger statement. In order to characterize SMO, we unfortunately need a few more definitions:

Definition 10 (Controlled-Discernibility (CD)). *Two different paths θ and θ' of length N are controlled-discernible (CD) if*

$$(I-P)\big(\Gamma(\theta) - \Gamma(\theta')\big) \neq 0, \tag{33}$$

where P is the matrix of any projection on $\Re([O(\theta)\ \mathcal{O}(\theta')])$.

It can be verified that CD is well-defined, even though P is not unique. However, to fix the ideas, we let $P(\theta, \theta')$ be the matrix of the orthogonal projection on $\Re([O(\theta)\ \ \mathcal{O}(\theta')])$, throughout the remainder of this section. Furthermore, note that CD is also symmetric. We can now establish the following:

Proposition 7 *If θ and θ' are CD, then there exists a vector U such that*

$$\forall x \in \mathbb{R}^n,\ \ Y(\theta, x, U) \notin \Re(\mathcal{O}(\theta')) + \Gamma(\theta')U. \tag{34}$$

Even though $\Re(\mathcal{O}(\theta')) + \Gamma(\theta')U$ is an affine subspace, we can still use the range inclusion test, by testing whether $Y(\theta, x, U)) - \Gamma(\theta')U$ is in $\Re(\mathcal{O}(\theta'))$. The proof of Proposition 7 is as follows:

Proof: Let U satisfy $(I - P(\theta, \theta'))\big(\Gamma(\theta) - \Gamma(\theta')\big)U \neq 0$. Then, by elementary linear algebra, $\Re(\mathcal{O}(\theta)) + \Gamma(\theta)U$ and $\Re(\mathcal{O}(\theta')) + \Gamma(\theta')U$ are totally disjoint as affine subspaces of $\mathbb{R}^{pN}$, which completes the proof, since $Y(\theta, x, U) \in \Re(\mathcal{O}(\theta)) + \Gamma(\theta)U$. □

Finally, we define:

Definition 11 (Forward Controlled-Discernibility (FCD)). *Two different paths θ and θ' of length N are forward controlled-discernible (FCD) if there exists an integer N' such that $\theta\lambda$ and $\theta'\lambda'$ are controlled discernible for any pair of paths λ and λ' of length N'. The smallest such integer is the index of FCD.* ♢

Unfortunately, we do not know whether or not FCD is decidable. This is in part due to the fact that, as opposed to $\mathcal{O}(\theta)$, we know little about the structure of $\Gamma(\theta)$. Nevertheless, we can characterize SMO as follows:

Theorem 9 *The SLS (30) is SMO at N iff any two different paths θ and θ' of length N are FCD.* $\diamondsuit$

Proof: Suppose the system is SMO at N with index N'. It follows that there exists a control vector U such that for all $\theta, \theta' \in \Theta_N$, $\theta \neq \theta'$, and $\lambda, \lambda' \in \Theta_{N'}$, $\Re(\mathcal{O}(\theta\lambda)) + \Gamma(\theta\lambda)U$ and $\Re(\mathcal{O}(\theta'\lambda')) + \Gamma(\theta'\lambda')U$ are totally disjoint, which implies that $(I - P(\theta\lambda, \theta'\lambda'))\big(\Gamma(\theta\lambda) - \Gamma(\theta'\lambda')\big)U \neq 0$, thus that $(I - P(\theta\lambda, \theta'\lambda'))\big(\Gamma(\theta\lambda) - \Gamma(\theta'\lambda')\big) \neq 0$, hence FCD of θ and θ' with index at most N'.

Now, let N' be the maximum index of FCD, over all pairs of different paths of length N, and let us show that the system is SMO with index at most N'. We need to show the existence of a vector U in $\mathbb{R}^{m(N+N')}$ such that

$$(I - P(\theta\lambda, \theta'\lambda'))\big(\Gamma(\theta\lambda) - \Gamma(\theta'\lambda')\big)U \neq 0 \tag{35}$$

for all $\theta, \theta' \in \Theta_N$, $\theta \neq \theta'$, and $\lambda, \lambda' \in \Theta_{N'}$. Since every pair θ and θ' of different paths of length N is FCD with index at most N', we get

$$(I - P(\theta\lambda, \theta'\lambda'))\big(\Gamma(\theta\lambda) - \Gamma(\theta'\lambda')\big) \neq 0. \tag{36}$$

Therefore,

$$K = \bigcup_{\theta,\theta',\lambda,\lambda'} \ker\Big((I - P(\theta\lambda, \theta'\lambda'))\big(\Gamma(\theta\lambda) - \Gamma(\theta'\lambda')\big)\Big) \neq \mathbb{R}^{m(N+N')}, \tag{37}$$

since it is a finite union of proper subspaces of $\mathbb{R}^{m(N+N')}$, by (36). Any control vector $U \in \mathbb{R}^{m(N+N')} \backslash K$ will work in (35), and is therefore discerning. $\square$

It should be noted that the existence of a single discerning control U implies that "almost" any vector of the same length is discerning, as established by (37). Finally, we describe an SMO system in the next example.

Example 4. Let

$$A(1) = A(2) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad C(1) = C(2) = \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{38}$$

$$B(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad B(1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \tag{39}$$

Since the observability pairs $(A(1), C(1))$ and $(A(2), C(2))$ are equal, no two paths can be discernible, because all paths of the same length share the exact same observability matrix. However, this system is SMO at $N = 2$, with index $N' = 1$. To see this, it suffices to use Theorem 9 and to establish that $(I - P(\theta\lambda, \theta'\lambda'))\big(\Gamma(\theta\lambda) - \Gamma(\theta'\lambda')\big) \neq 0$ for any two different paths θ and θ' of length 2, and any pair of paths λ and λ' of length 1. $\triangle$

4 Conclusion

We have characterized several concepts of observability in switched linear systems through simple linear algebraic tests, and we have shown their decidability in the autonomous case. An assumption underlying all criteria studied was that the mode sequences were arbitrary, which is novel in the sense that most (if not all) previous work assumed constraints on the mode sequences, usually in the form of minimum "dwell times" between switches.

This paper is intended as an intermediate step towards a better understanding of the observability of switched systems. Indeed, some results need to be refined, some problems still need to be solved, and many extensions are in view. To mention a few, the decidability of *forward controlled-discernibility* (FCD), which seems to be a challenging problem, and the characterization and study of state observability in the non-autonomous case, still need to be addressed. Finally, the investigation of the application of the concept of *discernibility* to asymptotic observer design promises to be fruitful, and we leave it to a future endeavor.

References

1. M. Babaali and M. Egerstedt, "Pathwise observability and controllability are decidable," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, (Maui, HW), December 2003.
2. A. Balluchi, L. Benvenuti, M. D. Di Benedetto, and A. L. Sangiovanni-Vincentelli, "Design of observers for hybrid systems," vol. 2289 of *Lect. Notes Comp. Sc.*, pp. 76–89, 2002.
3. A. Bemporad, G. Ferrari-Trecate, and M. Morari, "Observability and controllability of piecewise affine and hybrid systems," *IEEE Transactions on Automatic Control*, vol. 45, pp. 1864–1876, October 2000.
4. S. L. Campbell and C. D. J. Meyer, *Generalized Inverses of Linear Transformations.* New York, NY: Dover, 1991.
5. J. Ezzine and A. H. Haddad, "Controllability and observability of hybrid systems," *Int. J. Control*, vol. 49, no. 6, pp. 2045–2055, 1989.
6. D. Goshen-Meskin and I. Y. Bar-Itzhack, "Observability analysis of piece-wise constant systems - part I: Theory," *IEEE Trans. Aerospace & Electronic Systems*, vol. 28, no. 4, pp. 1056–1067, 1992.
7. I. Hwang, H. Balakrishnan, and C. Tomlin, "Observability criteria and estimator design for stochastic linear hybrid systems," in *Proceedings of the IEE European Control Conference*, (Cambridge, UK), September 2003.
8. Y. Ji and H. Chizeck, "Controllability, observability and discrete-time jump linear quadratic control," *Int. J. Control*, vol. 48, no. 2, pp. 481–498, 1988.
9. M. Mariton, "Stochastic observability of linear systems with markovian jumps," in *Proceedings of the 25th IEEE Conference on Decision and Control*, (Athens, Greece), pp. 2208–2209, December 1986.
10. R. Vidal, A. Chiuso, and S. Soatto, "Observability and identifiability of jump linear systems," in *Proceedings of the 41st IEEE Conference on Decision and Control*, (Las Vegas, NV), pp. 3614–3619, December 2002.

11. P. D. West and A. H. Haddad, "On the observability of linear stochastic switching systems," in *Proceedings of the 1994 American Control Conf.*, (Baltimore, MD), pp. 1846–1847, 1994.

A Some Generalized Matrix Inversion Theory

We now present some definitions and results from matrix inversion theory (see, e.g., [4]). We first recall that $\mathcal{O}^{\{1\}}$ is a $\{1\}$-inverse of $\mathcal{O}$ if

$$\mathcal{O}\mathcal{O}^{\{1\}}\mathcal{O} = \mathcal{O}, \tag{40}$$

and that the (Moore-Penrose) pseudo-inverse of $\mathcal{O}$ is defined as

$$\mathcal{O}\mathcal{O}^{\dagger}\mathcal{O} = \mathcal{O},\ \mathcal{O}^{\dagger}\mathcal{O}\mathcal{O}^{\dagger} = \mathcal{O}^{\dagger},\ \mathcal{O}^{\dagger}\mathcal{O} = (\mathcal{O}^{\dagger}\mathcal{O})',\ \text{and}\ \mathcal{O}\mathcal{O}^{\dagger} = (\mathcal{O}\mathcal{O}^{\dagger})'. \tag{41}$$

Note that the pseudo-inverse $\mathcal{O}^{\dagger}$ of $\mathcal{O}$ always satisfies (40), and is therefore a $\{1\}$-inverse. If furthermore $\mathcal{O}$ has full column rank, then any $\{1\}$-inverse $\mathcal{O}^{\{1\}}$ of $\mathcal{O}$ is a left inverse of $\mathcal{O}$, in the sense that $\mathcal{O}^{\{1\}}\mathcal{O} = I$, the identity matrix. We next consider the following equation:

$$Y = \mathcal{O}x, \tag{42}$$

where $x \in \mathbb{R}^n$ and $Y \in \mathbb{R}^N$, and we examine the conditions on Y for (42) to have a solution in x, and how to compute that solution. Note that

$$\exists\ x \mid Y = \mathcal{O}x \Longleftrightarrow Y \in \Re(\mathcal{O}), \tag{43}$$

which is why we refer to the following test as the range inclusion test:

Proposition 8 (Range Inclusion Test) *If $\mathcal{O}^{\{1\}}$ is a $\{1\}$-inverse of $\mathcal{O}$, then*

$$Y \in \Re(\mathcal{O}) \Leftrightarrow (\mathcal{O}\mathcal{O}^{\{1\}} - I)Y = 0. \tag{44}$$

Proof:

$\Leftarrow$ Let $x = \mathcal{O}^{\{1\}}Y$. Then $Y = \mathcal{O}x$, which concludes the proof.
$\Rightarrow$ We have $Y \in \Re(\mathcal{O}) \Rightarrow \exists\ x$ s.t. $Y = \mathcal{O}x$. By definition of a left inverse, we have that $\mathcal{O}\mathcal{O}^{\{1\}}\mathcal{O}x = \mathcal{O}x$, which implies that $\mathcal{O}\mathcal{O}^{\{1\}}Y = Y$, which concludes the proof. □

In words, equation (42) has a solution if and only if $(\mathcal{O}\mathcal{O}^{\{1\}} - I)Y = 0$ holds for some $\{1\}$-inverse (it then holds for *any* $\{1\}$-inverse). Note that if (42) admits a solution, then $x = \mathcal{O}^{\{1\}}Y$ is a solution to (42) for any $\{1\}$-inverse $\mathcal{O}^{\{1\}}$ of $\mathcal{O}$.