Observability for Switched Linear Systems: Characterization and Observer Design

Aneel Tanwani, *Member, IEEE*, Hyungbo Shim, *Member, IEEE*, and Daniel Liberzon, *Fellow, IEEE*

***Abstract*—This paper presents a characterization of observability and an observer design method for switched linear systems with state jumps. A necessary and sufficient condition is presented for observability, globally in time, when the system evolves under predetermined mode transitions. Because this characterization depends upon the switching signal under consideration, the existence of singular switching signals is studied alongside developing a sufficient condition that guarantees uniform observability with respect to switching times. Furthermore, while taking state jumps into account, a relatively weaker characterization is given for determinability, the property that concerns with recovery of the original state at some time rather than at all times. Assuming determinability of the system, a hybrid observer is designed for the most general case to estimate the state of the system and it is shown that the estimation error decays exponentially. Since the individual modes of the switched system may not be observable, the proposed strategy for designing the observer is based upon a novel idea of accumulating the information from individual subsystems. Contrary to the usual approach, dwell-time between switchings is not necessary, but the proposed design does require persistent switching. For practical purposes, the calculations also take into account the time consumed in performing computations.**

***Index Terms*—Determinability, observability, observer design, switched linear systems.**

I. Introduction

THIS paper studies observability conditions and observer construction for switched linear systems described as

$$\dot{x}(t) = A_{\sigma(t)}x(t) + B_{\sigma(t)}u(t), \qquad t \neq \{t_q\} \tag{1a}$$

$$x(t_q) = E_{\sigma(t_q^-)}x\left(t_q^-\right) + F_{\sigma(t_q^-)}v_q, \qquad q \geq 1 \tag{1b}$$

$$y(t) = C_{\sigma(t)}x(t) + D_{\sigma(t)}u(t), \qquad t \geq t_0 \tag{1c}$$

where $x(t) \in \mathbb{R}^n$ is the state, $y(t) \in \mathbb{R}^{d_y}$ is the output, $v_q \in \mathbb{R}^{d_v}$ and $u(t) \in \mathbb{R}^{d_u}$ are the inputs, and $u(\cdot)$ is a locally bounded measurable function. For some index set $\mathcal{I}$, the switching signal $\sigma : \mathbb{R} \mapsto \mathcal{I}$ is a piecewise constant and right-continuous function that changes its value at switching times $\{t_q\}, q \in \mathbb{N}$. In our notation, if a function exhibits discontinuity at time instant t_q, we evaluate that function at t_q^- to represent its value prior to discontinuity, and at t_q to indicate its value right after the jump. It is assumed that there are a finite number of switching times in any finite time interval; thus, we rule out the Zeno phenomenon in our problem formulation. It is assumed throughout the paper that the signal $\sigma(\cdot)$ (and thus, the active mode and the switching times $\{t_q\}$ as well) is known over the interval of interest. The switching mode $\sigma(t)$ and the switching times $\{t_q\}$ may be governed by a supervisory logic controller, or considered as an external input, or determined internally depending on the system state (in the last case, it is assumed that the value of $\sigma(t)$ can be inferred from the measured output $y(t)$, or by using a sensor that detects mode transitions). Also, one may refer to, e.g., [4], [6], [19], [22], for the problem of estimation of the switching signal $\sigma(t)$.

In the past decade, the structural properties of switched systems have been investigated by many researchers and observability along with observer construction has been one of them, see for example [16], [18], and [22]. In switched systems, the observability can be studied from various perspectives. If we allow for the use of the differential operator in the observer, then it may be desirable to determine the continuous state of the system instantaneously from the measured output and inputs. This in turn requires each subsystem to be observable, and the problem becomes nontrivial when the switching signal is treated as an unknown discrete state and simultaneous recovery of the discrete and continuous state is desired. Some results on this problem are published in [2], [5], and [22].

On the other hand, if the mode transitions are represented by a known switching signal then, even though the individual subsystems are not observable, it is still possible to recover the initial state $x(t_0)$ by appropriately processing the measured signals over a time interval that involves multiple switching instants. This phenomenon is of particular interest for switched systems or systems with state jumps as the notion of instantaneous observability and observability over an interval coincide for non-switched linear time invariant systems. This variant of the observability problem in switched systems has been studied most notably by [3], [16], [24], and [25]. The authors in [8] and [9] have studied the similar problem for the systems that allow jumps in the states, but they do not consider the change in the dynamics that is introduced by switching to different matrices associated with the active mode.

The observer design has also received some attention in the literature [1], [4], [12], where the authors have assumed that each mode in the system is in fact observable admitting a state

Manuscript received August 03, 2011; revised March 31, 2012; accepted September 13, 2012. Date of publication October 11, 2012; date of current version March 20, 2013. This work was supported in part by a National Research Foundation Grant funded by the Korean Government (NRF-2011-220-D00043) and in part by the National Science Foundation under Grant ECCS-0821153. Recommended by Associate Editor P. Shi.

A. Tanwani is with the Institut National de Recherche en Informatique et Automatique (INRIA), Grenoble 38334, France (e-mail: atanwani@ieee.org).

H. Shim is with ASRI, Seoul National University, Seoul 151-744, Korea (e-mail: hshim@snu.ac.kr).

D. Liberzon is with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, IL 61801 USA (e-mail: liberzon@illinois.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2012.2224257

0018-9286/$31.00 © 2012 IEEE

observer, and have treated the switching as a source of perturbation effect. This approach not only has limited applicability but it also incurs the need of a common Lyapunov function for the switched error dynamics, or a fixed amount of dwell-time between switching instants, because it is intrinsically a stability problem of the error dynamics.

The main contribution of this paper is to present a characterization of observability and an observer design for the systems represented by (1), where the subsystems are no longer required to be observable. So the notion of observability adopted in this paper is related to [3], [24] in the sense that we also consider observability over an interval. However, the authors in [3] only present a coordinate-dependent sufficient condition that leads to the construction of an observer; and the work of [24] only focuses on a necessary and sufficient condition under which there exists a switching signal that makes it possible to recover $x(t_0)$, without any discussion on design of observers. This paper fills the void by constructing an asymptotic observer based on a necessary and sufficient condition for observability. To the best of authors' knowledge, the considered class of linear systems is the most general for this purpose in the literature.

In our approach, the switching signal is considered to be known and fixed, so that the trajectory of the system satisfies a time-varying linear differential equation with state jumps. Then for that particular trajectory, we answer the question whether it is possible to recover $x(t_0)$ from the knowledge of measured inputs and outputs. We present a necessary and sufficient condition for observability over an interval, which is independent of coordinate transformations. Since this condition depends upon the switching times, we study the denseness property of the set of switching signals with a fixed mode sequence such that system (1) satisfies the observability condition for each switching signal in that set. For the sake of completeness, conditions which can be verified independently of switching times, are derived as a corollary to the main result. Also, with a similar tool set, the notion of determinability (also called "final-state observability" in [15] and reconstructability in [16]), which is more in the spirit of recovering the current state based on the knowledge of inputs and outputs in the past, is developed. Moreover, a hybrid observer for system (1) is designed based on the proposed necessary and sufficient condition. Since the observers are useful for various engineering applications, their utility mainly lies in their online operation method. This thought is essentially rooted in the idea for observer construction adopted in this paper: the idea of combining the partial information available from each mode and processing this collected information at one instance of time to get the estimate of the state. For real-time implementation, the time required for processing this information is also taken into account in our design. We show that under mild assumptions, such an estimate converges to the actual state of the plant and the state estimation error satisfies an exponentially decaying bound.

More emphasis will be given to the case when the individual modes of system (1) are not observable (in the classical sense of linear time-invariant systems theory) because it is obvious that the system becomes immediately observable when it switches to an observable mode. In such cases, the switching signal plays a pivotal role as the observability of the switched system depends upon not only the mode sequence but also the switching times. In order to facilitate our understanding of this matter, let us begin with an example.

Example 1: Consider a switched system characterized by

$$A_a = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A_b = \begin{bmatrix} \epsilon & 1 \\ -1 & \epsilon \end{bmatrix}$$
$$C_a = [\,1 \quad 0\,], \qquad C_b = [\,0 \quad 0\,]$$

with $E_i = I$, $F_i = 0$, $B_i = 0$, $D_i = 0$ for $i \in \mathcal{I} := \{a, b\}$, and ϵ is a constant. It is noted that neither of the two pairs, (A_a, C_a) or (A_b, C_b), is observable. However, if the switching signal $\sigma(t)$ changes its value in the order $a \to b \to a$ at times t_1 and t_2, then we can recover the state. In fact, it turns out that at least two switchings are necessary and the switching sequence should contain the subsequence of modes (a, b, a). For instance, if the switching happens as $a \to b \to a$, then the output at time t_1^- and t_2 is: $y(t_1^-) = C_a x(t_1^-) = x_1(t_0)$, and $y(t_2) = C_a e^{A_b \tau} x(t_0) = e^{\epsilon\tau} \cos\tau \cdot x_1(t_0) + e^{\epsilon\tau} \sin\tau \cdot x_2(t_0)$, where $x(t_0) = [x_1(t_0), x_2(t_0)]^\top$ is the initial condition and $\tau = t_2 - t_1$. It is obvious that $x(t_0)$ can be recovered from two measurements $y(t_1^-)$ and $y(t_2)$ if $\tau \neq k\pi$ with $k \in \mathbb{N}$. On the other hand, any switching signal, whose duration for mode b is an integer multiple of π, is a "singular" switching signal (whose precise meaning will be given in Section II-A). ◁

Notation: For a square matrix A and a subspace $\mathcal{V}$, we denote by $\langle A|\mathcal{V}\rangle$ the smallest A-invariant subspace containing $\mathcal{V}$, and by $\langle \mathcal{V}|A\rangle$ the largest A-invariant subspace contained in $\mathcal{V}$. (See Property 7 in Appendix A for their computation.) For a matrix A, $\mathcal{R}(A)$ denotes the column space (range space) of A. The sum of two subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$ is defined as $\mathcal{V}_1 + \mathcal{V}_2 := \{v_1 + v_2 : v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2\}$. For a possibly non-invertible matrix A, the preimage of a subspace $\mathcal{V}$ under A is given by $A^{-1}\mathcal{V} = \{x : Ax \in \mathcal{V}\}$. Let $\ker A := A^{-1}\{0\}$; then it is seen that $A^{-1} \ker C = \ker(CA)$ for a matrix C. For convenience of notation, let $A^{-\top}\mathcal{V} := (A^\top)^{-1}\mathcal{V}$ where $A^\top$ is the transpose of A, and it is understood that $A_2^{-1}A_1^{-1}\mathcal{V} = A_2^{-1}(A_1^{-1}\mathcal{V})$. Also, we denote the products of matrices A_i as $\prod_{i=j}^{k} A_i := A_j A_{j+1} \cdots A_k$ when $j < k$, and $\prod_{i=j}^{k} A_i := A_j A_{j-1} \cdots A_k$ when $j > k$. The notation $\mathrm{col}(A_1, \ldots, A_k)$ means the vertical stack of matrices $A_1, \cdots, A_k$, that is, $[A_1^\top, \ldots, A_k^\top]^\top$.

Before going further, let us rename the switching sequence for convenience. For system (1), when the switching signal $\sigma(t)$takes the mode sequence $(q_1, q_2, q_3, \cdots)$, we rename them as increasing integers (1, 2, 3, $\cdots$), which is ever increasing even though the same mode is revisited; for convenience, this sequence is indexed by $q \in \mathbb{N}$. Moreover, it is often the case that the mode of the system changes without the state jump (1b), or the state jumps without switching to another mode. In the former case, we can simply take $E_q = I$ and $F_q = 0$, and in the latter case, we increase the mode index by one and take $A_q = A_{q+1}$, $B_q = B_{q+1}$ and so on. In this way, various situations fit into the description of (1) with increasing mode sequence. The switching time $t_q$is the instant when transition from mode q to mode $q + 1$ takes place.

II. Geometric Conditions for Observability

To make precise the notions of observability and determinability considered in this paper, let us introduce the formal definitions.

Definition 1: Let $(\sigma^i, u^i, v^i, y^i, x^i)$, for $i = 1, 2$, be the signals that satisfy (1) over an interval[1] $[t_0, T^+)$. We say that the system (1) is $[t_0, T^+)$-*observable* if the equality $(\sigma^1, u^1, v^1, y^1) = (\sigma^2, u^2, v^2, y^2)$ implies that $x^1(t_0) = x^2(t_0)$. Similarly, the system (1) is said to be $[t_0, T^+)$-*determinable* if the equality $(\sigma^1, u^1, v^1, y^1) = (\sigma^2, u^2, v^2, y^2)$ implies that $x^1(T) = x^2(T)$.

Since the initial state $x(t_0)$, the switching signal σ, and the inputs (u, v) uniquely determine $x(t)$ on $[t_0, T^+)$ by (1), observability is achieved if and only if the whole state trajectory $x(t)$ on $[t_0, T^+)$, is uniquely determined by the inputs, the output, and the switching signal. Obviously, observability implies determinability by forward integration of (1), but the converse is not true due to the possibility of non-invertible matrices E_σ. In case there is no jump map (1b), or each E_σ is invertible, observability and determinability are equivalent.

The use of Definition 1 leads to the following proposition whose proof is rather straightforward and can be found in [20].

Proposition 1: For a switching signal σ, the system (1) is $[t_0, T^+)$-observable (or, determinable) if, and only if, zero inputs and zero output on the interval $[t_0, T^+)$ imply that $x(t_0) = 0$ (or, $x(T) = 0$).

Because of Proposition 1, we are motivated to introduce the following homogeneous switched system, which has been obtained by setting the inputs (u, v) equal to zero in (1):

$$\dot{x}(t) = A_q x(t), \quad y(t) = C_q x(t), \quad t \in [t_{q-1}, t_q) \tag{2a}$$
$$x(t_q) = E_q x\left(t_q^-\right). \tag{2b}$$

If this homogeneous system is observable (or, determinable) with a given σ, then $y \equiv 0$ implies that $x(t_0) = 0$ (or, $x(T) = 0$), and in terms of the description of system (1), this means that zero inputs and zero output yield $x(t_0) = 0$ (or, $x(T) = 0$); hence, (1) is observable (or, determinable) because of Proposition 1. On the other hand, if system (1) is observable (or, determinable), then it is still observable (or, determinable) with zero inputs, which is described as system (2). Thus, the observability (or, determinability) of systems (1) and (2) are equivalent.

A. Necessary and Sufficient Condition for Observability

Based on the tool set developed in, e.g., [23], we present a characterization of the unobservable subspace for system (2) with a given switching signal. Towards this end, let $\mathcal{N}_q^m$ $(m \geq q)$ denote the set of states at $t = t_{q-1}$ for system (2) that generate identically zero output over $[t_{q-1}, t_{m-1}^+)$. Then, for fixed switching times, it is easily seen that $\mathcal{N}_q^m$ is actually a subspace due to linearity of (2), and we call $\mathcal{N}_q^m$ the *unobservable subspace for* $[t_{q-1}, t_{m-1}^+)$. It can be seen that system (2) is an LTI system between two consecutive switching times, so that its unobservable subspace on the interval $[t_{q-1}, t_q)$ is simply given by the largest A_q-invariant subspace contained in $\ker C_q$, i.e., $\langle \ker C_q | A_q \rangle = \ker G_q$ where

$$G_q := \operatorname{col}\left(C_q, C_q A_q, \cdots, C_q A_q^{n-1}\right).$$

So it is clear that $\mathcal{N}_q^q = \ker G_q$. Now, when the measured output is available over the interval $[t_{q-1}, t_{m-1}^+)$ that includes switchings at $t_q, t_{q+1}, \ldots, t_{m-1}$, more information about the state is obtained in general so that $\mathcal{N}_q^m$ gets smaller as the difference $m - q$ gets larger, and we claim that the subspace $\mathcal{N}_q^m$ can be computed recursively as follows:[2]

$$\begin{aligned} \mathcal{N}_m^m &= \ker G_m, \\ \mathcal{N}_q^m &= \ker G_q \cap e^{-A_q \tau_q} E_q^{-1} \mathcal{N}_{q+1}^m, \quad 1 \leq q \leq m-1 \end{aligned} \tag{3}$$

where $\tau_q := t_q - t_{q-1}$. The following theorem presents a necessary and sufficient condition for observability of the system (1) while proving the claim in the process.

Theorem 1: For system (2) with a switching signal $\sigma_{[t_0, t_{m-1}^+)}$, the unobservable subspace for $[t_0, t_{m-1}^+)$ at t_0 is given by $\mathcal{N}_1^m$ from (3). Therefore, system (1) is $[t_0, t_{m-1}^+)$-observable if, and only if

$$\mathcal{N}_1^m = \{0\}. \tag{4}$$

From (3), it is not difficult to arrive at the following formula for $\mathcal{N}_q^m$:

$$\mathcal{N}_q^m = \ker G_q \cap \left(\bigcap_{i=q+1}^{m} \prod_{j=q}^{i-1} e^{-A_j \tau_j} E_j^{-1} \ker G_i \right) \tag{5a}$$
$$= \ker G_q \cap \left(\bigcap_{i=q+1}^{m} \ker \left(G_i \prod_{l=i-1}^{q} E_l e^{A_l \tau_l} \right) \right) \tag{5b}$$

where $1 \leq q \leq m-1$. In order to inspect the observability of system (2), one can compute $\mathcal{N}_1^m$ using (3), (5a), or (5b). It is easily seen that $\mathcal{N}_1^{m_1} \supseteq \mathcal{N}_1^{m_2}$ if $m_1 \leq m_2$ (with arbitrary $\tau_i > 0$, $i \in \mathbb{N}$).

Proof of Theorem 1: Sufficiency. Using the result of Proposition 1, it suffices to show that the identically zero output of (2) implies $x(t_0) = 0$. Assume that $y \equiv 0$ on $[t_0, t_{m-1}^+)$. Then, it is immediate that $x(t_{m-1}) \in \mathcal{N}_m^m = \ker G_m$. We next apply the inductive argument to show that $x(t_{q-1}) \in \mathcal{N}_q^m$ for $1 \leq q \leq m-1$. Suppose that $x(t_q) \in \mathcal{N}_{q+1}^m$, then $x(t_{q-1}) \in e^{-A_q \tau_q} E_q^{-1} \mathcal{N}_{q+1}^m$ since $x(t)$ is the solution of (2). Zero output on the interval $[t_{q-1}, t_q)$ also implies that $x(t_{q-1}) \in \ker G_q$. Therefore,

$$x(t_{q-1}) \in \ker G_q \cap e^{-A_q \tau_q} E_q^{-1} \mathcal{N}_{q+1}^m$$

where the expression on right coincides with the definition of $\mathcal{N}_q^m$ in (3). This induction proves the claim that $\mathcal{N}_q^m$, the unobservable subspace for $[t_{q-1}, t_{m-1}^+)$, is given by (3). With $q = 1$, it is seen that $x(t_0) \in \mathcal{N}_1^m = \{0\}$, which proves the sufficiency.

[1]The notation $[t_0, T^+)$ is used to denote the interval $[t_0, T + \varepsilon)$, where $\varepsilon > 0$ is arbitrarily small. In fact, because of the right-continuity of the switching signal, the output $y(T)$ belongs to the next mode when T is the switching instant. Then, the point-wise measurement $y(T)$ is insufficient to contain the information from the new mode, and thus, it is imperative to consider the output signal over the interval $[t_0, T + \varepsilon)$ with $\varepsilon > 0$. This definition implicitly implies that the observability property does not change for sufficiently small ε (which is true, and becomes clear shortly).

[2]Although we use the matrix exponential in deriving conditions for observability, addressing the numerical implementation of such computations is beyond the scope of this paper. As an example, the reader may find [10] interesting for that purpose.

Necessity. Assuming that $\mathcal{N}_1^m \neq \{0\}$, we show that a nonzero initial state $x(t_0) \in \mathcal{N}_1^m$ yields the solution of (2) such that $y \equiv 0$ on $[t_0, t_{m-1}^+)$, which implies unobservability. First, we show the following implication:

$$x(t_{q-1}) \in \mathcal{N}_q^m \quad \Rightarrow \quad x(t_q) \in \mathcal{N}_{q+1}^m, \qquad 1 \leq q < m. \quad (6)$$

Indeed, assuming that $x(t_{q-1}) \in \mathcal{N}_q^m$ with $1 \leq q < m$, it follows that $x(t_q) = E_q e^{A_q \tau_q} x(t_{q-1})$, which further gives

$$\begin{aligned} x(t_q) \in & E_q e^{A_q \tau_q} \mathcal{N}_q^m = E_q e^{A_q \tau_q} \left(\ker G_q \cap e^{-A_q \tau_q} E_q^{-1} \mathcal{N}_{q+1}^m \right) \\ \subseteq & E_q \ker G_q \cap E_q E_q^{-1} \mathcal{N}_{q+1}^m \\ = & E_q \ker G_q \cap \mathcal{N}_{q+1}^m \cap \mathcal{R}(E_q) \subseteq \mathcal{N}_{q+1}^m \end{aligned}$$

by using (3) and Properties 2, 3, and 11 in Appendix A. Therefore, for $0 \leq q \leq m-1$, $x(t_q) \in \mathcal{N}_{q+1}^m \subseteq \ker G_{q+1}$, and the solution $x(t) = e^{A_{q+1}(t-t_q)} x(t_q)$ for $t \in [t_q, t_{q+1})$ satisfies that $y(t) = C_{q+1} x(t) = 0$ for $t \in [t_q, t_{q+1})$ due to A_{q+1}-invariance of $\ker G_{q+1}$. ∎

The observability condition (4) given in Theorem 1 is dependent upon a particular switching signal under consideration, and it is entirely possible that the system is observable for certain switching signals and unobservable for others (cf. Example 1). However, it would be more useful to know whether the observability property holds for a particular class of switching signals. Towards this end, we show that if there is a switching signal that satisfies (4), then the set of switching signals, with the same mode sequence, for which (4) does not hold, is nowhere dense.

To formalize this argument, consider the set $\mathcal{S}$ consisting of all switching signals σ (over a possibly different time domain) with a fixed mode sequence $(1, 2, \cdots, m)$ and switching times t_i such that $t_0 < t_1 < \cdots < t_{m-1}$. Then, for each $\sigma \in \mathcal{S}$, there is a corresponding vector $\tau = \mathrm{col}(\tau_1, \cdots, \tau_{m-1}) \in \mathcal{T} := \{\tau \in \mathbb{R}^{m-1} : \tau_i > 0\}$ with $\tau_i = t_i - t_{i-1}$ being the activation period for mode i under σ. We now introduce the metric $d(\cdot, \cdot)$ on the set $\mathcal{S}$ as follows:

$$d(\sigma^1, \sigma^2) := \|\tau^1 - \tau^2\|_1 = \sum_{i=1}^{m-1} \left| \tau_i^1 - \tau_i^2 \right|$$

for any $\sigma^1, \sigma^2 \in \mathcal{S}$, with $\tau^1, \tau^2 \in \mathcal{T}$.

Theorem 2: Let $\mathcal{S}^* := \{\sigma \in \mathcal{S} : \text{system } (1) \; is \; [t_0, t_{m-1}^+) - \text{observable with } \sigma\}$. If the set $\mathcal{S}^*$ is nonempty, then it is an open and dense subset of $\mathcal{S}$ under the topology induced by the metric $d(\cdot, \cdot)$.

The proof basically relies on the analyticity of the exponential map—an argument which is employed in several control theoretic results of this kind, see for example [15, Ch. 6], [16, Ch. 4]. The same idea is used in a formally worked out proof of Theorem 2 appearing in Appendix B. A consequence of Theorem 2 is that if there exists $\sigma' \in \mathcal{S} \setminus \mathcal{S}^*$, then we can find an element $\sigma'' \in \mathcal{S}^*$ by introducing arbitrarily small perturbations in the vector τ' that corresponds to σ'. We call such σ' a *singular* switching signal and the ones contained in the set $\mathcal{S}^*$ are called *regular* switching signals.

B. Conditions Independent of Switching Times

Existence of singular switching signals naturally raises the question whether, under certain conditions, observability holds uniformly with respect to switching times. In other words, it is desirable to know whether the observability could be verified for a given mode sequence independently of the switching times. For this, we again consider the sets $\mathcal{S}$ and $\mathcal{T}$ for a given mode sequence $(1, 2, \cdots, m)$.

Definition 2: The switched system (1) *is uniformly observable for all switching signals* $\sigma \in \mathcal{S}$ (i.e., $\mathcal{S}^* = \mathcal{S}$) if, and only if

$$\mathcal{V}_1^m := \cup_{\tau_1 > 0} \cup_{\tau_2 > 0} \cdots \cup_{\tau_{m-1} > 0} \mathcal{N}_1^m(\tau) = \{0\}. \quad (7)$$

By using the distributive property of intersection over union of sets, one can rewrite $\mathcal{V}_1^m$ by proceeding in the sequential manner as before

$$\begin{aligned} & \mathcal{V}_m^m := \ker G_m \\ & \mathcal{V}_q^m := \ker G_q \cap \left(\cup_{\tau_q > 0} e^{-A_q \tau_q} E_q^{-1} \mathcal{V}_{q+1}^m \right), \; 1 \leq q \leq m-1. \quad (8) \end{aligned}$$

However, in order to check condition (7) in practice, a difficulty arises due to the fact that $\mathcal{V}_1^m$ is not a subspace in general. (This is because the set $\cup_{\tau > 0} e^{A\tau} \mathcal{V}$ is not closed under addition of vectors, and hence not necessarily a subspace even though $\mathcal{V}$is a subspace.) To avoid this difficulty and obtain conditions for observability based on computing the dimension of a subspace, we seek a subspace containing $\mathcal{V}_q^m$, so that a sufficient condition is obtained for uniform observability with respect to switching times.

Corollary 1: Let $\overline{\mathcal{N}}_1^m$ be defined as follows:

$$\begin{aligned} & \overline{\mathcal{N}}_m^m := \ker G_m \\ & \overline{\mathcal{N}}_q^m := \left\langle A_q | \ker G_q \cap E_q^{-1} \overline{\mathcal{N}}_{q+1}^m \right\rangle, \quad 1 \leq q \leq m-1. \end{aligned}$$

Then, $\overline{\mathcal{N}}_1^m$ is a subspace that contains $\mathcal{V}_1^m$, and thus, the system (1) is uniformly observable for all $\sigma \in \mathcal{S}$ *if* $\overline{\mathcal{N}}_1^m = \{0\}$.

By construction, the subspace $\overline{\mathcal{N}}_1^m$ also contains $\mathcal{N}_1^m$ so that it serves a sufficient condition for (4) as well as for (7).

Proof: The proof is completed by showing that $\mathcal{V}_q^m \subseteq \overline{\mathcal{N}}_q^m$ for $1 \leq q \leq m$. First, note that $\mathcal{V}_m^m = \overline{\mathcal{N}}_m^m$. Assuming that $\mathcal{V}_{q+1}^m \subseteq \overline{\mathcal{N}}_{q+1}^m$ for $1 \leq q \leq m-1$, we now claim that $\mathcal{V}_q^m \subseteq \overline{\mathcal{N}}_q^m$. Indeed, by Properties 3, 9, and 11 in Appendix A, and the recursion (8), we obtain

$$\begin{aligned} \mathcal{V}_q^m &= \ker G_q \cap \left(\cup_{\tau_q > 0} e^{-A_q \tau_q} E_q^{-1} \mathcal{V}_{q+1}^m \right) \\ &= \cup_{\tau_q > 0} \, e^{-A_q \tau_q} \left(\ker G_q \cap E_q^{-1} \mathcal{V}_{q+1}^m \right) \\ &\subseteq \left\langle A_q | \ker G_q \cap E_q^{-1} \mathcal{V}_{q+1}^m \right\rangle \\ &\subseteq \left\langle A_q | \ker G_q \cap E_q^{-1} \overline{\mathcal{N}}_{q+1}^m \right\rangle = \overline{\mathcal{N}}_q^m, \quad 1 \leq q \leq m-1. \quad (9) \end{aligned}$$

Therefore, the condition $\overline{\mathcal{N}}_1^m = \{0\}$ implies (7). ∎

Since $\mathcal{V}_1^m$ is contained in $\overline{\mathcal{N}}_1^m$, it would be of interest to investigate how far the statement of Corollary 1 is from necessity, and the following lemma turns out to be helpful for that purpose (see Appendix B for its proof).

Lemma 1: For each $1 \leq q \leq m-1$, if $\mathcal{V}_{q+1}^m = \overline{\mathcal{N}}_{q+1}^m$, then $\overline{\mathcal{N}}_q^m$ is the smallest subspace containing the set $\mathcal{V}_q^m$.

The above discussion can be summarized as follows, which suggests when the condition in Corollary 1 becomes necessary.

Corollary 2: If each $\mathcal{V}_q^m$, for $1 \leq q \leq m-1$, is a subspace, then system (1) is uniformly observable with respect to the switching times if and only if $\overline{\mathcal{N}}_1^m = \{0\}$.

Example 2: The switched system considered in Example 1, with the mode sequence (a, b, a) and an arbitrary constant ϵ, provides an example of Corollary 2. It is seen that each $\mathcal{V}_q^m$ is a subspace as $\overline{\mathcal{N}}_3^3 = \mathcal{V}_3^3 = \text{span}\{\text{col}(0,1)\}$, $\overline{\mathcal{N}}_2^3 = \mathcal{V}_2^3 = \mathbb{R}^2$, and $\overline{\mathcal{N}}_1^3 = \mathcal{V}_1^3 = \text{span}\{\text{col}(0,1)\} \neq \{0\}$. Indeed, the switched system in Example 1 is not uniformly observable for the mode sequence (a, b, a), as seen by the existence of the singular switching signals.

Now, let us consider an additional subsystem, where $A_c := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $E_c = I_{2\times 2}$, and B_c, C_c, D_c, F_c are zero matrices with appropriate dimensions. For the switched system with the mode sequence (a, c, a), and $m = 3$, we get $\overline{\mathcal{N}}_3^3 = \text{span}\{\text{col}(0,1)\}$, $\overline{\mathcal{N}}_2^3 = \mathbb{R}^2$, and $\overline{\mathcal{N}}_1^3 = \text{span}\{\text{col}(0,1)\} \neq \{0\}$. The sufficient condition in Corollary 1 is violated but the resulting switched system is observable for all $\tau_1, \tau_2 > 0$ (which can be seen from Theorem 1). The source of this gap is the fact that the set $\mathcal{V}_2^3 = \cup_{\tau_2>0} e^{-A_c\tau_2}(\ker G_c \cap \ker G_a) = \cup_{\tau_2>0} e^{-A_c\tau_2}\text{span}\{\text{col}(0,1)\}$ is not a subspace. It is seen that the smallest subspace containing $\mathcal{V}_2^3$, in this case, is $\overline{\mathcal{N}}_2^3 = \mathbb{R}^2$. ◁

Having studied the uniform observability with respect to the switching times, we also discuss the existence of switching times for observability under a given mode sequence. This is to see whether the set $\mathcal{S}^*$ is empty or not when the mode sequence is given (or, the set $\mathcal{S}$ is given). Regarding this question, the idea of under-approximating $\mathcal{N}_1^m$ yields the following necessary condition.

Corollary 3: Let $\underline{\mathcal{N}}_1^m$ be defined as follows:

$$\underline{\mathcal{N}}_m^m := \ker G_m,$$
$$\underline{\mathcal{N}}_q^m := \left\langle \ker G_q \cap E_q^{-1}\underline{\mathcal{N}}_{q+1}^m | A_q \right\rangle, \qquad 1 \leq q \leq m-1.$$

Then, $\underline{\mathcal{N}}_1^m$ is a subspace contained in $\mathcal{N}_1^m(\tau)$ for all $\tau \in \mathcal{T}$, and thus, if there exists a vector $\tau \in \mathcal{T}$ such that $\mathcal{N}_1^m(\tau) = \{0\}$, that is, system (1) is $[t_0, t_{m-1}^+)$-observable (or, equivalently $\mathcal{S}^*$ is non-empty), then $\underline{\mathcal{N}}_1^m = \{0\}$.

Proof: For each $\tau \in \mathcal{T}$, the proof proceeds similar to Corollary 1. With $\mathcal{N}_m^m = \underline{\mathcal{N}}_m^m$, we assume that $\mathcal{N}_{q+1}^m \supseteq \underline{\mathcal{N}}_{q+1}^m$ for $1 \leq q \leq m-1$, and claim that $\mathcal{N}_q^m \supseteq \underline{\mathcal{N}}_q^m$. Again by Properties 3, 9, and 11 in Appendix A, and employing (3), we obtain

$$\begin{aligned} \mathcal{N}_q^m &= e^{-A_q\tau_q}\left(\ker G_q \cap E_q^{-1}\mathcal{N}_{q+1}^m\right) && \text{(10a)} \\ &\supseteq \left\langle \ker G_q \cap E_q^{-1}\mathcal{N}_{q+1}^m | A_q \right\rangle \\ &\supseteq \left\langle \ker G_q \cap E_q^{-1}\underline{\mathcal{N}}_{q+1}^m | A_q \right\rangle = \underline{\mathcal{N}}_q^m, \ 1 \leq q \leq m-1. && \text{(10b)} \end{aligned}$$

The condition $\underline{\mathcal{N}}_1^m = \{0\}$ is then implied by $\mathcal{N}_1^m(\tau) = \{0\}$. ■

As a matter of fact, it can be shown that[3]

$$\underline{\mathcal{N}}_1^m = \cap_{\tau_1>0} \cdots \cap_{\tau_{m-1}>0} \mathcal{N}_1^m(\tau). \qquad (11)$$

[3]Indeed, this follows from (10b) and from the claim that $\cap_{\tau>0} e^{-A\tau}\mathcal{V} = \langle \mathcal{V} | A \rangle$ for any subspace $\mathcal{V}$ and a matrix A, which is proved as follows. Since $e^{-A\tau}\mathcal{V} \supseteq \langle \mathcal{V}|A\rangle$ for all τ, we have that $\cap_{\tau>0} e^{-A\tau}\mathcal{V} \supseteq \langle \mathcal{V}|A\rangle$. On the other hand, since $(\cap_{\tau>0} e^{-A\tau}\mathcal{V})^\perp$ is a subspace containing $\cup_{\tau>0} e^{A^\top\tau}\mathcal{V}^\perp$, it holds that $(\cap_{\tau>0} e^{-A\tau}\mathcal{V})^\perp \supseteq \langle A^\top | \mathcal{V}^\perp \rangle$ because $\langle A^\top | \mathcal{V}^\perp \rangle$ is the smallest subspace containing $\cup_{\tau>0} e^{A^\top\tau}\mathcal{V}^\perp$ (see Lemma 1). Taking orthogonal complement, we get $\cap_{\tau>0} e^{-A\tau}\mathcal{V} \subseteq \langle \mathcal{V}|A\rangle$ (by Property 10 in Appendix A), which proves the claim.

Since the right-hand side of (11) can become $\{0\}$ even though $\mathcal{N}_1^m(\tau) \neq \{0\}$ for any $\tau \in \mathcal{T}$, it is clear that the condition $\underline{\mathcal{N}}_1^m = \{0\}$ is much weaker than requiring the existence of τ that satisfies $\mathcal{N}_1^m(\tau) = \{0\}$. However, it can be shown that[4] $\langle \mathcal{V}|A\rangle = \cap_{\tau>0} e^{-A\tau}\mathcal{V} = \cap_{i=1}^n e^{-As_i}\mathcal{V}$ for almost all $s_i > 0$, $1 \leq i \leq n$. Using this fact, if $\underline{\mathcal{N}}_1^m = \{0\}$, then a switching signal can be constructed where the mode sequence 1 through m is repeated at most n^m times such that system (1) is observable under this new switching signal for almost all switching times. This approach of constructing a switching signal, which makes the system (1) observable, coincides with the notion of observability adopted in [16], [24] and is illustrated in the following example.

Example 3: Suppose that, for the switched system of Example 1, the mode sequence (b, a) is given. Then, with $m = 2$, we obtain that $\underline{\mathcal{N}}_2^2 = \ker G_a = \text{span}\{\text{col}(0,1)\}$ and $\underline{\mathcal{N}}_1^2 = \langle \ker G_b \cap \ker G_a | A_b \rangle = \{0\}$. However, it is verified that $\mathcal{N}_1^2(\tau_1) = \ker[e^{\epsilon\tau_1}\cos\tau_1, e^{\epsilon\tau_1}\sin\tau_1] = \text{col}(\cos\tau, -\sin\tau) \neq \{0\}$, so that (4) does not hold for any $\tau_1 > 0$, showing that the system is not observable with $m = 2$ even though the necessary condition of Corollary 3 is satisfied. On the other hand, if the mode sequence is repeated at least once, so that the new mode sequence is (b, a, b, a), then, other than the case where mode b is not activated for a multiple of π time units, the system is observable under the new switching signal. ◁

Remark 1: By taking the orthogonal complements of $\mathcal{N}_q^m$, $\overline{\mathcal{N}}_q^m$ and $\underline{\mathcal{N}}_q^m$, respectively, we get the dual conditions for observability, using Properties 5, 6, 8, and 10 in Appendix A, as follows. System (1) is $[t_0, t_{m-1}^+)$-observable if and only if $\mathcal{P}_1^m = \mathbb{R}^n$ where

$$\mathcal{P}_1^m := (\mathcal{N}_1^m)^\perp = \mathcal{R}\left(G_1^\top\right) + \sum_{i=2}^{m}\prod_{j=1}^{i-1} e^{A_j^\top\tau_j}E_j^\top\mathcal{R}\left(G_j^\top\right).$$

Similarly, one can state Corollaries 1 and 3 in alternate forms. System (1) is uniformly observable if $\underline{\mathcal{P}}_1^m = \mathbb{R}^n$, where $\underline{\mathcal{P}}_1^m$ is computed as

$$\underline{\mathcal{P}}_m^m = \left(\overline{\mathcal{N}}_m^m\right)^\perp = \mathcal{R}\left(G_m^\top\right)$$
$$\underline{\mathcal{P}}_q^m = \left(\overline{\mathcal{N}}_q^m\right)^\perp = \left\langle \mathcal{R}\left(G_q^\top\right) + E_q^\top\underline{\mathcal{P}}_{q+1}^m | A_q^\top \right\rangle$$

for $1 \leq q \leq m-1$. Also, if system (1) is $[t_0, t_{m-1}^+)$-observable with a switching signal, then $\overline{\mathcal{P}}_1^m = \mathbb{R}^n$, where $\overline{\mathcal{P}}_1^m$ is defined sequentially, for $1 \leq q \leq m-1$, as

$$\overline{\mathcal{P}}_m^m = (\underline{\mathcal{N}}_m^m)^\perp = \mathcal{R}\left(G_m^\top\right)$$
$$\overline{\mathcal{P}}_q^m = \left(\underline{\mathcal{N}}_q^m\right)^\perp = \left\langle A_q^\top | \mathcal{R}\left(G_q^\top\right) + E_q^\top\overline{\mathcal{P}}_{q+1}^m \right\rangle.$$

◁

C. Necessary and Sufficient Conditions for Determinability

In order to study determinability of system (1) and arrive at a result parallel to Theorem 1, our first goal is to develop an object similar to $\mathcal{N}_q^m$. So, for system (2) with a given switching signal, let $\mathcal{Q}_q^m$ be the set of states at time t_{m-1} (or t_{m-1}^+) such that its corresponding solution $x(t)$ produces zero output on the

[4]One may refer to Corollary 6.2.4 and Proposition 6.2.11 in [15], where the result is given for $\mathcal{V} = \ker C$, but the same holds for any subspace $\mathcal{V}$.

interval $[t_{q-1}, t^+_{m-1})$. We call $\mathcal{Q}^m_q$ the *undeterminable subspace for* $[t_{q-1}, t^+_{m-1})$. Then, it can be shown that $\mathcal{Q}^m_q$ is computed recursively as follows:

$$\begin{aligned}\mathcal{Q}^q_q &:= \ker G_q \\ \mathcal{Q}^k_q &:= \ker G_k \cap E_{k-1} e^{A_{k-1}\tau_{k-1}} \mathcal{Q}^{k-1}_q, \quad q+1 \le k \le m. \end{aligned} \tag{12}$$

These sequential definitions lead to another equivalent expression for $\mathcal{Q}^m_q$:

$$\mathcal{Q}^m_q = \ker G_m \cap E_{m-1} \ker(G_{m-1}) \cap \left(\bigcap_{i=q}^{m-2} \prod_{l=m-1}^{i+1} E_l e^{A_l \tau_l} E_i \ker G_i \right). \tag{13}$$

In fact, the subspace $\Pi^{i+1}_{l=m-1} E_l e^{A_l\tau_l} E_i \ker G_i$ indicates the set of states obtained by propagating the unobservable states of the mode i (where $q \le i \le m-2$) to the time t_{m-1} under the dynamics of system (2). Intersection of these subspaces with $E_{m-1}\ker G_{m-1}$ and $\ker G_m$ shows that $\mathcal{Q}^m_q$ is the set of states that cannot be determined at time t_{m-1} from the zero output over the interval $[t_{q-1}, t^+_{m-1})$. Therefore, the determinability of system (2) can now be characterized in the following theorem.

Theorem 3: For system (2) and a given switching signal $\sigma_{[t_0, t^+_{m-1})}$, the undeterminable subspace for $[t_0, t^+_{m-1})$ at t_{m-1} is given by $\mathcal{Q}^m_1$ of (13). Therefore, system (1) is $[t_0, t^+_{m-1})$-determinable if and only if

$$\mathcal{Q}^m_1 = \{0\}. \tag{14}$$

The condition (14) is equivalent to (4) when all E_q matrices, $q = 1, \ldots, m-1$, are invertible because of the relation

$$\mathcal{Q}^m_1 = \prod_{l=m-1}^{1} E_l e^{A_l\tau_l} \mathcal{N}^m_1.$$

Example 4: If any of the jump maps E_q of (2), $q = 1, \ldots, m-1$, is a zero matrix, then (14) trivially holds regardless of whether (4) holds or not. This is intuitively clear because we can uniquely determine $x(t_{m-1}) = 0$ even if $x(t_0)$ cannot be determined. ◁

Recalling that $\mathcal{S}$ is the set of switching signals σ with mode sequence $(1, 2, \ldots, m)$ and $\tau \in \mathcal{T}$, the following two corollaries parallel Corollaries 1 and 3, and are given for completeness. Proofs are omitted but can be developed using the property that $\underline{\mathcal{Q}}^q_1 \subseteq \mathcal{Q}^q_1(\tau) \subseteq \cup_{\tau\in\mathcal{T}} \mathcal{Q}^q_1(\tau) \subseteq \overline{\mathcal{Q}}^q_1$.

Corollary 4: System (1) is uniformly determinable for all $\sigma \in \mathcal{S}$, i.e., $\mathcal{Q}^m_1(\tau) = \{0\}$ for all $\tau \in \mathcal{T}$, if $\overline{\mathcal{Q}}^m_1 = \{0\}$, where $\overline{\mathcal{Q}}^m_1$ is computed by

$$\begin{aligned}\overline{\mathcal{Q}}^1_1 &:= \ker G_1 \\ \overline{\mathcal{Q}}^q_1 &:= E_{q-1}\left\langle A_{q-1} | \overline{\mathcal{Q}}^{q-1}_1 \right\rangle \cap \ker G_q, \quad 2 \le q \le m.\end{aligned}$$

Corollary 5: If there exists a vector $\tau \in \mathcal{T}$ such that $\mathcal{Q}^m_1(\tau) = \{0\}$, i.e., system (1) is $[t_0, t^+_{m-1})$-determinable for some $\sigma \in \mathcal{S}$, then $\underline{\mathcal{Q}}^m_1 = \{0\}$, where $\underline{\mathcal{Q}}^m_1$ is computed by

$$\begin{aligned}\underline{\mathcal{Q}}^1_1 &:= \ker G_1 \\ \underline{\mathcal{Q}}^q_1 &:= E_{q-1}\left\langle \underline{\mathcal{Q}}^{q-1}_1 | A_{q-1} \right\rangle \cap \ker G_q, \quad 2 \le q \le m.\end{aligned}$$

Remark 2: An alternative dual characterization of determinability is possible by inspecting whether the complete state information is available while going forward in time. This is achieved in terms of the subspace $\mathcal{M}^m_q$, obtained by taking the orthogonal complement of $\mathcal{Q}^m_q$. Using Properties 5, 6, 8, and 10 in Appendix A, it follows from (13) that

$$\begin{aligned}\mathcal{M}^m_q &:= \left(\mathcal{Q}^m_q\right)^\perp \\ &= \sum_{i=q}^{m-2} \prod_{l=m-1}^{i+1} E_l^{-\top} e^{-A_l^\top \tau_l} E_i^{-\top} \mathcal{R}\left(G_i^\top\right) \\ &\quad + E_{m-1}^{-\top} \mathcal{R}\left(G_{m-1}^\top\right) + \mathcal{R}\left(G_m^\top\right).\end{aligned}$$

Note that, with $m > q$, the set-valued map $\prod_{j=q}^{m-1} e^{-A_j\tau_j} E_j^{-1} x$ pulls the state x at t_{m-1} back in time at t_{q-1}. This map was used in (5a) to pull $\mathcal{N}^m_m$ back at t_{q-1}. Since the dual of $\mathbb{R}^n$ is also $\mathbb{R}^n$, the adjoint of this map pushes the row vectors forward in time from t_{q-1} to t_{m-1}. In other words, $\mathcal{M}^m_q$ is the set of states at time instant $t = t_{m-1}$ that can be identified, modulo the unobservable subspace at t_{m-1}, from the information of y over the interval $[t_{q-1}, t^+_{m-1})$. Therefore, the dual statement to Theorem 1 for determinability is that the system (1) is $[t_0, t^+_{m-1})$-determinable if and only if

$$\mathcal{M}^m_1 = \mathbb{R}^n. \tag{15}$$

It is noted that a recursive expression for $\mathcal{M}^m_1$ is given by

$$\begin{aligned}\mathcal{M}^1_1 &= \mathcal{R}\left(G_1^\top\right) \\ \mathcal{M}^q_1 &= E_{q-1}^{-\top} e^{-A_{q-1}^\top \tau_{q-1}} \mathcal{M}^{q-1}_1 + \mathcal{R}\left(G_q^\top\right), \quad 2 \le q \le m\end{aligned}$$

and the dual statements of Corollaries 4 and 5, that are independent of switching times, are given as follows: system (1) is uniformly determinable for all $\sigma \in \mathcal{S}$ if $\underline{\mathcal{M}}^m_1 = \mathbb{R}^n$, where

$$\begin{aligned}\underline{\mathcal{M}}^1_1 &:= \left(\overline{\mathcal{Q}}^m_1\right)^\perp = \mathcal{R}\left(G_1^\top\right), \\ \underline{\mathcal{M}}^q_1 &:= \left(\overline{\mathcal{Q}}^q_1\right)^\perp = E_{q-1}^{-\top}\left\langle \underline{\mathcal{M}}^{q-1}_1 | A_{q-1}^\top \right\rangle + \mathcal{R}\left(G_q^\top\right)\end{aligned}$$

for $2 \le q \le m$. Similarly, if there exists a $\sigma \in \mathcal{S}$ such that system (1) is $[t_0, t^+_{m-1})$-determinable then $\overline{\mathcal{M}}^m_1 = \mathbb{R}^n$, where $\overline{\mathcal{M}}^m_1$ is computed as follows:

$$\begin{aligned}\overline{\mathcal{M}}^1_1 &:= \left(\underline{\mathcal{Q}}^m_1\right)^\perp = \mathcal{R}\left(G_1^\top\right), \\ \overline{\mathcal{M}}^q_1 &:= \left(\underline{\mathcal{Q}}^q_1\right)^\perp = E_{q-1}^{-\top}\left\langle A_{q-1}^\top | \overline{\mathcal{M}}^{q-1}_1 \right\rangle + \mathcal{R}\left(G_q^\top\right)\end{aligned}$$

for $2 \le q \le m$. ◁

III. Observer Design

In engineering practice, an observer is designed to provide an estimate of the actual state value at current time. In this regard, determinability (weaker than observability according to Definition 1) is a suitable notion for switched systems. Based on the conditions obtained for determinability in the previous section, an asymptotic observer is designed for system (1) in this section. By asymptotic observer, we mean that the estimate $\hat{x}(t)$ converges to the plant state $x(t)$ as $t \to \infty$.

A. Observer Overview

In order to construct an observer for system (1), we introduce the following assumptions.

Assumption 1:

1) The switching is persistent in the sense that there exists a constant $T_D > 0$ such that a switch occurs at least once in every time interval of length T_D; that is,
$$t_q - t_{q-1} \leq T_D, \qquad \forall q \in \mathbb{N}. \tag{16}$$
In addition, there are a bounded number of switchings in any finite time interval; i.e., there is a function $J_{\max}(\cdot)$ such that the number of switchings in any time interval of duration T is less than or equal to $J_{\max}(T)$.
2) The system is persistently determinable in the sense that there exists an $N \in \mathbb{N}$ such that
$$\dim \mathcal{M}^q_{q-N} = n, \qquad \forall q \geq N+1. \tag{17}$$
(The integer N is interpreted as the minimal number of switches required to gain determinability.)
3) There are constants b_A and b_E such that $\|A_q\| \leq b_A$ and $\max\{1, \|E_q\|\} \leq b_E$ for all $q \in \mathbb{N}$ (which is always the case when A_q and E_q belong to a finite set).

The observer we propose is a hybrid dynamical system of the form

$$\dot{\hat{x}}(t) = A_q\hat{x}(t) + B_q u(t), \quad t \in [t_{q-1}, t_q), \quad t \neq \hat{t}_k \tag{18a}$$
$$\hat{x}(t_q) = E_q\hat{x}\left(t_q^-\right) + F_q v_q, \quad q \geq 1 \tag{18b}$$
$$\hat{x}(\hat{t}_k) = \hat{x}\left(\hat{t}_k^-\right) - \xi_k, \quad k \geq 1 \tag{18c}$$

with an arbitrary initial state $\hat{x}(t_0) \in \mathbb{R}^n$. Here, $\hat{t}_k$ is the time for the kth estimation update (see Fig. 2), and we assume that $\hat{t}_k \neq t_q$ for any k and q because these updates (18b) and (18c) are executed sequentially on a digital processor. Fig. 1, together with Fig. 2, provides an illustration of how the proposed observer (18) is executed in practice. It is seen that the observer consists of a system copy and an estimate update law by a correction vector ξ_k. The vector ξ_k can be thought of as an approximation of the state estimation error and our goal is to design ξ_k such that $\hat{x}(t) \to x(t)$. In order to construct ξ_k, the signals available from the actual plant are gathered and stored over a time interval encompassing N switches. Because of Assumption 1.2, this stored information is rich enough to compute ξ_k that closely approximates the state estimation error. More specifically, we write ξ_k as a function of the observable components of individual modes, and the observable components are recovered by running the classical Luenberger observers for each of the past $N+1$ active modes. It is supposed that these observers process the stored data of the past switching intervals much faster than the real time scale. An estimate of the plant state is then obtained from the estimates of the observable components using an inversion logic. Computation of this estimate is performed on a digital processor while the observer (18) is also running (in parallel) synchronously with the plant. Unlike our conference paper [20], we no longer assume that this computation is performed instantaneously; we instead suppose that the computation is completed within a maximal computation time $T_C > 0$. Therefore, the estimate obtained for the plant

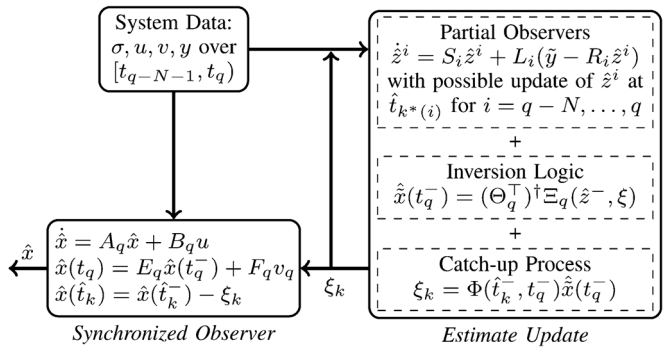

Fig. 1. Flow diagram of the proposed observer. The stored information is processed in the *Estimate Update* block to generate the updating signal ξ_k, which is passed on to the *Synchronized Observer* running in parallel to the plant at the same pace.

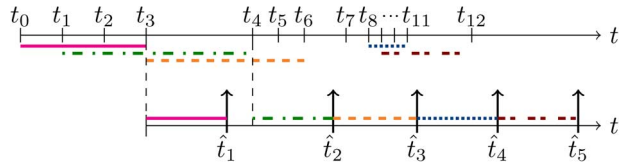

Fig. 2. Assuming that $N = 2$ in this figure, the computation of ξ_1 begins at time t_3. By processing the data between t_0 and t_3, the computation completes at $\hat{t}_1$ and the estimate $\hat{x}(\hat{t}_1)$ is updated by (18c). By assuming that T_C is the maximum computation time, $\hat{t}_1 \leq t_3 + T_C$. The computation for ξ_3 begins at $\hat{t}_2$ which is delayed because the previous computation for ξ_2, having started at t_4, does not finish when the new switchings occur at t_5 and t_6. While computing ξ_4, only the data between t_8 and t_{11} is processed.

state is not for the current time. In order to compensate for the computational delay, the *catch-up* process is introduced, with which the estimate ξ_k of the estimation error at the current time (denoted by $\hat{t}_k$) is obtained and used for updating (18c). For example, in Fig. 2, having gathered the information from the plant over the interval $[t_0, t_3)$, the computation for ξ_1 starts at t_3 and its value is available at $\hat{t}_1$, where $t_3 < \hat{t}_1 \leq t_3 + T_C$. After first N switches, the computation of ξ_k starts after every switch. However, in the case that another switch occurs before the on-going computation is complete, the request for the new computation is put in a waiting queue until the completion of the current computation. If several switches occur while a computation is being performed, then only the last $N+1$ active modes are considered during the next computation. See the information processed, marked with various types of lines, in Fig. 2.

Remark 3: One can always consider a nonswitched linear system with quadruple $\{A, B, C, D\}$ as a special case of switched system (1), where, for each $q \in \mathbb{N}$, $A_q = A$, $B_q = B$, $C_q = C$, $D_q = D$, $E_q = I$, and $F_q = 0$. Assumption 1.1 and Assumption 1.3 are trivially satisfied by introducing the pseudo-switches at some arbitrary times. Assumption 1.2 corresponds to the usual observability condition. The proposed observer scheme will be able to estimate the state of a nonswitched system. However, the state estimate is updated at discrete time instants only. ◁

B. Algorithm for Computing ξ_k

In the sequel, the above thought process is formalized by setting up a machinery to compute the correction vector ξ_k as indicated in Fig. 1. Based on these computations, a procedure for implementing the hybrid observer, according to the scheme shown

in Fig. 2, is outlined in Algorithm 1. It is then shown in Theorem 4 that the state estimate computed according to the parameter bounds given in Algorithm 1 indeed converges to the actual state of the system.

With $\tilde{x} := \hat{x} - x$, the error dynamics are described by

$$\dot{\tilde{x}}(t) = A_q \tilde{x}(t) \tag{19a}$$

$$\tilde{x}(t_q) = E_q \tilde{x}\left(t_q^-\right) \tag{19b}$$

$$\tilde{x}(\hat{t}_k) = \tilde{x}\left(\hat{t}_k^-\right) - \xi_k. \tag{19c}$$

The output error is defined as $\tilde{y}(t) := C_q\hat{x}(t) + D_q u(t) - y(t) = C_q\tilde{x}(t)$. Since estimating $\tilde{x}$ is equivalent to the estimation of x (obtained by subtracting $\tilde{x}$ from $\hat{x}$ of (18)), our design begins with the estimation of $\tilde{x}$ at time t_q^- (with $q \geq N+1$). In order to estimate the observable part of $\tilde{x}$ from each mode q, let us design the partial observers using the observability decomposition [23] and the dynamics in (19). Choose a matrix Z^q such that its columns are an orthonormal basis of $\mathcal{R}(G_q^\top)$, that is, $\mathcal{R}(Z^q) = \mathcal{R}(G_q^\top)$. Further, choose a matrix W^q such that its columns are an orthonormal basis of $\ker G_q$. From the construction, there are matrices $S_q \in \mathbb{R}^{r_q \times r_q}$ and $R_q \in \mathbb{R}^{d_y \times r_q}$, where $r_q = \operatorname{rank} G_q$, such that $Z^{q\top} A_q = S_q Z^{q\top}$ and $C_q = R_q Z^{q\top}$, and the pair (S_q, R_q) is observable. Let $z^q := Z^{q\top}\tilde{x}$ and $w^q := W^{q\top}\tilde{x}$, so that z^q (resp. w^q) denotes the observable (resp. unobservable) states of the mode q. Thus, for $z^q \in \mathbb{R}^{r_q}$, the error dynamics in (19) satisfy

$$\begin{aligned} \dot{z}^q(t) &= S_q z^q(t), \quad \tilde{y}(t) = R_q z^q(t), \quad t \in [t_{q-1}, t_q) \\ z^q(\hat{t}_k) &= z^q\left(\hat{t}_k^-\right) - Z^{q\top}\xi_k, \qquad \text{if} \quad \hat{t}_k \in (t_{q-1}, t_q) \end{aligned} \tag{20}$$

with the initial condition $z^q(t_{q-1}) = Z^{q\top}\tilde{x}(t_{q-1})$. Since z^q is observable over the interval $[t_{q-1}, t_q)$, a standard Luenberger observer, whose role is to estimate $z^q(t_q^-)$ at the end of the interval, is designed as

$$\begin{aligned} \dot{\hat{z}}^q(t) &= S_q \hat{z}^q(t) + L_q\left(\tilde{y}(t) - R_q\hat{z}^q(t)\right), \quad t \in [t_{q-1}, t_q) \\ \hat{z}^q(\hat{t}_k) &= \hat{z}^q\left(\hat{t}_k^-\right) - Z^{q\top}\xi_k, \qquad \text{if} \quad \hat{t}_k \in (t_{q-1}, t_q) \end{aligned} \tag{21}$$

with the initialization $\hat{z}^q(t_{q-1}) = 0$, where L_q is a matrix such that $(S_q - L_q R_q)$ is Hurwitz. Note that we have fixed the initial condition of the estimator to be zero for each interval, whose role becomes clear in (38).

Now let us define the state transition matrix $\Phi(s, r)$, $s > r$, that results in $\tilde{x}(s) = \Phi(s, r)\tilde{x}(r)$ along the dynamics (19a) and (19b) [but not (19c)]. For example, when $s = t_j^-$ and $r = t_i^-$ $(j > i)$ are switching instants, we have that

$$\Phi\left(t_j^-, t_i^-\right) = e^{A_j\tau_j} E_{j-1} e^{A_{j-1}\tau_{j-1}} E_{j-2} \cdots e^{A_{i+1}\tau_{i+1}} E_i =: \Psi_i^j \tag{22}$$

in which Ψ_i^j is defined for convenience. Note that Ψ_i^j is computed using the knowledge of the switching periods $\{\tau_{i+1}, \cdots, \tau_j\}$ which will be denoted simply by $\tau_{\{i+1,j\}}$, and note also that $\Psi_i^i := I$.

We now define a matrix Θ_i^q with $i \leq q$ whose columns form a basis of the subspace $\mathcal{R}(\Psi_i^q W^i)^\perp$; that is,

$$\mathcal{R}\left(\Theta_i^q\right) = \mathcal{R}\left(\Psi_i^q W^i\right)^\perp, \quad i = q-N, \cdots, q.$$

By construction, each column of Θ_i^q is orthogonal to the subspace $\ker G_i$ that has been transported from t_i^- to t_q^- along the error dynamics (19a) and (19b). This matrix Θ_i^q will be used for annihilating the unobservable component in the state estimate obtained from the mode i after being transported to the time t_q^-. As a convention, we take Θ_i^q to be a null matrix whenever $\mathcal{R}(\Psi_i^q W^i)^\perp = \{0\}$. Using the determinability of the system (Assumption 1.2), it will be shown later in the proof of Theorem 4 that the matrix

$$\Theta_q := \left[\Theta_q^q \vdots \cdots \vdots \Theta_{q-N}^q\right] \tag{23}$$

has rank n. Equivalently, $\Theta_q^\top$ has n independent columns and is left-invertible, so that $\left(\Theta_q^\top\right)^\dagger = \left(\Theta_q\Theta_q^\top\right)^{-1}\Theta_q$, where $\dagger$ denotes the left-pseudo-inverse. Introduce the notation

$$\begin{aligned} &\mathcal{K}_i^j := \left\{k \in \mathbb{N} : \hat{t}_k \in (t_i, t_j)\right\}, \quad \xi_{\{i,j\}} := \left\{\xi_k : k \in \mathcal{K}_i^j\right\}, \\ &\hat{z}_{\{i,j\}}^- := \left\{\hat{z}^i\left(t_i^-\right), \ldots, \hat{z}^j\left(t_j^-\right)\right\}. \end{aligned} \tag{24}$$

Let us also define the vector Ξ_q as as shown in the equation at the bottom of the page. The matrices M_i^q with $i = q-N, \cdots, q$ are defined such that M_i^q is a null matrix when Θ_i^q is null, and the following holds:

$$\begin{aligned} &\left[M_q^q, M_{q-1}^q, \cdots, M_{q-N}^q\right] := \left(\Theta_q^\top\right)^\dagger \\ &\times \operatorname{blockdiag}\left(\Theta_q^{q\top}\Psi_q^q, \Theta_{q-1}^{q\top}\Psi_{q-1}^q, \cdots, \Theta_{q-N}^{q\top}\Psi_{q-N}^q\right). \end{aligned} \tag{25}$$

Each non-empty M_i^q is an n by n matrix whose argument is $\tau_{\{q-N+1,q\}}$ in general (due to the inversion of $\Theta_q^\top$), while the argument of both Θ_i^q and Ψ_i^q is $\tau_{\{i+1,q\}}$.

$$\Xi_q\left(\hat{z}_{\{q-N,q\}}^-, \xi_{\{q-N,q\}}\right) := \begin{bmatrix} \Theta_q^{q\top}\Psi_q^q Z^q \hat{z}^q\left(t_q^-\right) \\ \Theta_{q-1}^{q\top}\left(\Psi_{q-1}^q Z^{q-1}\hat{z}^{q-1}\left(t_{q-1}^-\right) - \sum_{k\in\mathcal{K}_{q-1}^q}\Phi\left(t_q^-, \hat{t}_k^-\right)\xi_k\right) \\ \vdots \\ \Theta_{q-N}^{q\top}\left(\Psi_{q-N}^q Z^{q-N}\hat{z}^{q-N}\left(t_{q-N}^-\right) - \sum_{k\in\mathcal{K}_{q-N}^q}\Phi\left(t_q^-, \hat{t}_k^-\right)\xi_k\right) \end{bmatrix}$$

Finally, let $T_B := T_D + T_C$, where T_C is the upper bound on computation time, and define

$$\bar{b} := e^{b_A \cdot (2T_C + T_B)} \cdot b_E^{J_{\max}(2T_C) + J_{\max}(T_B)}. \tag{26}$$

Pick any number $\alpha \in (0,1)$ and compute the injection gain L_i such that

$$\bar{b} \left\| M_i^q \left(\tau_{\{q-N+1,q\}} \right) Z^i e^{(S_i - L_i R_i)\tau_i} Z^{i\top} \right\| \leq \frac{\alpha}{N+2}. \tag{27}$$

(One constructive way to compute such an L_i is from the *squashing lemma* [11, Lemma 1].) Using the information over the interval $[t_{q-N-1}, t_q)$, the error correction vector ξ_k in (18c) is now computed as

$$\xi_k = \Phi\left(\hat{t}_k^-, t_q^-\right) \hat{\tilde{x}}\left(t_q^-\right) \tag{28}$$

where

$$\hat{\tilde{x}}\left(t_q^-\right) = \left(\Theta_q^\top\right)^\dagger \Xi_q \left(\hat{z}^-_{\{q-N,q\}}, \xi_{\{q-N,q\}} \right). \tag{29}$$

Algorithm 1 summarizes these calculations for ξ_k and also illustrates how the schematics of Figs. 1 and 2 could be implemented. It comprises two processes running in parallel, *Synchronized Observer* and *Estimate Update*. Whenever the switching happens, the *Synchronized Observer* calls the *Estimate Update* if the latter is not already occupied with computation from previous switch. If the switch occurs while the Estimate Update is active, we wait for it to finish the previous computation and then look at the information from last $N+1$ active modes for the next update.

Algorithm 1: Implementation of the hybrid observer

Input: σ, u, v, y

Initialization: $\hat{x}(t_0) \in \mathbb{R}^n$, $q = N$, $k = 0$, `swCount` $= 0$, `Update` $=$ `idle`

1 **Synchronized Observer Loop**
2 Run the observer (18) synchronously to plant (1).
3 **if** *switching occurs* **then** increment `swCount`.
4 **if** `Update` $==$ `idle` *and* $q <$ `swCount` **then**
5 $q \leftarrow$ `swCount` and call **Estimate Update**.
6 **Loop end**

1 **Estimate Update**
2 `Update` $\leftarrow$ `active`
3 **for** $i = q - N$ **to** q **do**
4 Compute the gains L_i satisfying (27).
5 Obtain $\hat{z}^i(t_i^-)$ by running the individual observer (21) for the ith mode.
6 Compute $\hat{\tilde{x}}$ from (29).
7 Increment k and set $\hat{t}_k \leftarrow$ `CurrentTime`.
8 Set $\xi_k \leftarrow \Phi(\hat{t}_k^-, t_q^-)\hat{\tilde{x}}(t_q^-)$ and update $\hat{x}$ by (18c).
9 `Update` $\leftarrow$ `idle`
10 **end**

Remark 4: We remark that the idea of post-processing the stored information is really significant for switched systems with unobservable modes. While computing ξ_k from the observable components of last $N+1$ active modes, we first need to propagate these components under the dynamics of subsequent modes, which requires exact knowledge of the switching times. In addition, knowledge of the past switching times is also used when the observer gains are chosen in (27), with which it will be shown that the estimation error decreases by the desired factor α. ◁

C. *Analysis of Error Convergence*

The following theorem shows that the above implementation indeed guarantees the convergence of the state estimation error to zero.

Theorem 4: Under Assumption 1, consider the hybrid observer (18) in which the estimate update ξ_k is computed through (28) and introduced at $\hat{t}_k$, according to Algorithm 1. If the gains L_i, for each $i = q-N, \cdots, q$, are chosen so that (27) holds for any choice of $\alpha \in (0,1)$, then $\lim_{t\to\infty} |\hat{x}(t) - x(t)| = 0$. Furthermore, the estimation error satisfies the following exponential convergence rate:

$$|\tilde{x}(t)| \leq h(\alpha) e^{-\gamma \ln(\alpha^{-1})(t-t_0)} |\tilde{x}(t_0)|, \qquad \forall t \geq t_0 \tag{30}$$

where γ is a positive constant and the function $h : (0,1) \to \mathbb{R}$ has the property that $h(\alpha) \to \infty$ when $\alpha \to 0$.

In (30), the factor α^{-1} denotes the exponential decay rate in estimation error which can be increased by choosing the output injection gains appropriately, but it comes at the cost of poor transient response. This observation is consistent with the peaking phenomenon studied in [17].

We remark that if the computation time T_C is to be ignored, then the analysis becomes much simpler and for that case we refer to the conference version [20].

Proof of Theorem 4: From Assumption 1 and Algorithm 1, it can be seen that

$$\hat{t}_{k+1} - \hat{t}_k \leq T_D + T_C = T_B, \quad \forall k \geq 1.$$

Hence, it suffices to show that $\lim_{k\to\infty} |\tilde{x}(\hat{t}_k)| = 0$ because Assumptions 1.1 and 1.3 imply that

$$|\tilde{x}(t)| \leq e^{b_A T_B} \cdot b_E^{J_{\max}(T_B)} \left|\tilde{x}(\hat{t}_k)\right|, \quad \forall t \in [\hat{t}_k, \hat{t}_{k+1}). \tag{31}$$

In the remainder of this proof, an expression for $\tilde{x}(\hat{t}_k)$ is derived whose norm is shown to converge to zero. For this purpose, fix $k > N+3$ and suppose that the kth estimate update process for ξ_k completes at time $\hat{t}_k$ after having processed the data on the interval $[t_{q-N-1}, t_q)$ (where it is possible that there is another switch between t_q and $\hat{t}_k^-$).

The error at $t_q^-, \tilde{x}(t_q^-)$, can be written as

$$\tilde{x}\left(t_q^-\right) = \begin{bmatrix} Z^{q\top} \\ W^{q\top} \end{bmatrix}^{-1} \begin{bmatrix} z^q\left(t_q^-\right) \\ w^q\left(t_q^-\right) \end{bmatrix} = Z^q z^q\left(t_q^-\right) + W^q w^q\left(t_q^-\right). \tag{32}$$

The matrix Ψ_i^j with $j > i$, defined in (22), transports $\tilde{x}(t_i^-)$ to $\tilde{x}(t_j^-)$ along (19) by

$$\tilde{x}\left(t_j^-\right) = \Psi_i^j \tilde{x}\left(t_i^-\right) - \sum_{k \in \mathcal{K}_i^j} \Phi\left(t_j^-, \hat{t}_k^-\right) \xi_k. \tag{33}$$

We now have the following series of equivalent expressions for $\tilde{x}(t_q^-)$:

$$\begin{aligned}
\tilde{x}\left(t_q^-\right) &= Z^q z^q\left(t_q^-\right) + W^q w^q\left(t_q^-\right) \\
&= \Psi_{q-1}^q \left(Z^{q-1} z^{q-1}\left(t_{q-1}^-\right) + W^{q-1} w^{q-1}\left(t_{q-1}^-\right)\right) \\
&\quad - \sum_{k \in \mathcal{K}_{q-1}^q} \Phi\left(t_q^-, \hat{t}_k^-\right) \xi_k \\
&= \Psi_{q-2}^q \left(Z^{q-2} z^{q-2}\left(t_{q-2}^-\right) + W^{q-2} w^{q-2}\left(t_{q-2}^-\right)\right) \\
&\quad - \sum_{k \in \mathcal{K}_{q-2}^q} \Phi\left(t_q^-, \hat{t}_k^-\right) \xi_k \\
&\quad \vdots \\
&= \Psi_{q-N}^q \left(Z^{q-N} z^{q-N}\left(t_{q-N}^-\right) + W^{q-N} w^{q-N}\left(t_{q-N}^-\right)\right) \\
&\quad - \sum_{k \in \mathcal{K}_{q-N}^q} \Phi\left(t_q^-, \hat{t}_k^-\right) \xi_k.
\end{aligned} \tag{34}$$

To appreciate the implication of this equivalence, we first note that for each $q - N \le i \le q$, the term $\Psi_i^q Z^i z^i(t_i^-)$ transports the observable information of the ith mode from the interval $[t_{i-1}, t_i)$ to the time instant t_q^-. This observable information is corrupted by the unknown term $w^i(t_i^-)$, but since the information is being accumulated at t_q^- from modes $i = q-N, \cdots, q$, the idea is to combine the partial information from every intermediate mode to recover $\tilde{x}(t_q^-)$. This is where we use the notion of determinability. By Properties 1, 5, and 6 in Appendix A, and the fact that $\mathcal{R}(W^i)^\perp = (\ker G_i)^\perp = \mathcal{R}(G_i^\top)$ and $e^{-A_q^\top \tau_q} \mathcal{R}(G_q^\top) = \mathcal{R}(G_q^\top)$, it follows under Assumption 1.2 that

$$\begin{aligned}
&\mathcal{R}(W^q)^\perp + \mathcal{R}\left(\Psi_{q-1}^q W^{q-1}\right)^\perp + \cdots + \mathcal{R}\left(\Psi_{q-N}^q W^{q-N}\right)^\perp \\
&\quad = e^{-A_q^\top \tau_q} \Bigg(\mathcal{R}\left(G_q^\top\right) + E_{q-1}^{-\top} \mathcal{R}\left(G_{q-1}^\top\right) \\
&\qquad\qquad + \sum_{i=q-N}^{q-2} \prod_{l=q-1}^{i+1} E_l^{-\top} e^{-A_l^\top \tau_l} E_i^{-\top} \mathcal{R}\left(G_i^\top\right) \Bigg) \\
&\quad = e^{-A_q^\top \tau_q} \mathcal{M}_{q-N}^q = \mathbb{R}^n.
\end{aligned}$$

This equation shows that the matrix Θ_q defined in (23) has rank n, and is right-invertible. Keeping in mind that the range space of each Θ_i^q is orthogonal to $\mathcal{R}(\Psi_i^q W^i)$, each equality in (34) leads to the following relation:

$$\Theta_i^{q\top} \tilde{x}\left(t_q^-\right) = \Theta_i^{q\top} \left(\Psi_i^q Z^i z^i\left(t_i^-\right) - \sum_{k \in \mathcal{K}_i^q} \Phi\left(t_q^-, \hat{t}_k^-\right) \xi_k \right) \tag{35}$$

for $i = q - N, \cdots, q$. Stacking (35) from $i = q$ to $i = q - N$, and employing the left-inverse of $\Theta_q^\top$, we obtain that

$$\tilde{x}\left(t_q^-\right) = \left(\Theta_q^\top\right)^\dagger \Xi_q \left(z_{\{q-N,q\}}^-, \xi_{\{q-N,q\}} \right) \tag{36}$$

where $z_{\{q-N,q\}}^-$ denotes $\{z^{q-N}(t_{q-N}^-), \ldots, z^q(t_q^-)\}$. It is seen from (36) that, if we were able to estimate $z_{\{q-N,q\}}^-$ without error, then the plant state $x(t_q^-)$ would be exactly recovered by (36) because $x(t_q^-) = \hat{x}(t_q^-) - \tilde{x}(t_q^-)$ and both entities on the right side of the equation are known. However, since this is not the case, $z_{\{q-N,q\}}^-$ is replaced with its estimate $\hat{z}_{\{q-N,q\}}^-$ in (29), and $\hat{\tilde{x}}(t_q^-)$ is set as an estimate of $\tilde{x}(t_q^-)$, as done in (29).

Using the linearity of Ξ_q in its arguments, and substituting ξ_k from (28) in (19c), we get

$$\begin{aligned}
\tilde{x}(\hat{t}_k) &= \tilde{x}\left(\hat{t}_k^-\right) - \Phi\left(\hat{t}_k^-, t_q^-\right) \hat{\tilde{x}}\left(t_q^-\right) \\
&= \Phi\left(\hat{t}_k^-, t_q^-\right) \left(\Theta_q^\top\right)^\dagger \Big(\Xi_q \left(z_{\{q-N,q\}}^-, \xi_{\{q-N,q\}} \right) \\
&\quad - \Xi_q \left(\hat{z}_{\{q-N,q\}}^-, \xi_{\{q-N,q\}} \right) \Big) \\
&= -\Phi\left(\hat{t}_k^-, t_q^-\right) \left(\Theta_q^\top\right)^\dagger \Xi_q \left(\tilde{z}_{\{q-N,q\}}^-, 0 \right)
\end{aligned} \tag{37}$$

where $\tilde{z}_{\{q-N,q\}}^-$ denotes $\{\tilde{z}^{q-N}(t_{q-N}^-), \ldots, \tilde{z}^q(t_q^-)\}$ with $\tilde{z}^i(t_i^-) = \hat{z}^i(t_i^-) - z^i(t_i^-)$. It follows from (20) and (21) that $\dot{\tilde{z}}^i = (S_i - L_i R_i)\tilde{z}^i$ for $t \in [t_{i-1}, t_i)$ with

$$\tilde{z}^i(t_{i-1}) = \hat{z}^i(t_{i-1}) - z^i(t_{i-1}) = 0 - Z^{i\top} \tilde{x}(t_{i-1}) \tag{38}$$

which implies that

$$\tilde{z}^i\left(t_i^-\right) = e^{(S_i - L_i R_i)\tau_i} \tilde{z}^i(t_{i-1}) = -e^{(S_i - L_i R_i)\tau_i} Z^{i\top} \tilde{x}(t_{i-1}).$$

Plugging this expression in (37), and using the definition of M_i^q $(i = q - N, \ldots, q)$ from (25), we get

$$\begin{aligned}
\tilde{x}(\hat{t}_k) = \Phi\left(\hat{t}_k^-, t_q^-\right) \sum_{i=q-N}^{q} M_i^q \left(\tau_{\{q-N+1,q\}}\right) \\
\times Z^i e^{(S_i - L_i R_i)\tau_i} Z^{i\top} \tilde{x}(t_{i-1}).
\end{aligned} \tag{39}$$

For each $i = q - N - 1, \ldots, q - 1$, let $k^*(i) := \max\{k : \hat{t}_k < t_i\}$. Then it follows that

$$\tilde{x}(t_i) = \Phi\left(t_i, \hat{t}_{k^*(i)}\right) \tilde{x}\left(\hat{t}_{k^*(i)}\right).$$

From Fig. 2 and Algorithm 1, it is seen that

$$\hat{t}_k^- - t_q^- \le 2T_C \quad \text{and} \quad t_i - \hat{t}_{k^*(i)} \le T_B.$$

Thus, $\|\Phi(\hat{t}_k^-, t_q^-)\| \cdot \|\Phi(t_i, \hat{t}_{k^*(i)})\| \le \overline{b}, \forall i = q-N-1, \ldots, q-1$, where $\overline{b}$ is defined in (26).

Moreover, with k and q considered above, it can be seen that, for each $i = q - N - 1, \ldots, q - 1$, it holds that $k - N - 3 \le k^*(i) \le k - 2$ (since $k^*(q-1)$ either equals $k-2$ or $k-3$).

Then, from the selection of gains L_i's satisfying (27), it is seen that

$$\begin{aligned}\left|\tilde{x}(\hat{t}_k)\right| &\leq \frac{\alpha}{N+2}\sum_{i=q-N-1}^{q-1}\left|\tilde{x}\left(\hat{t}_{k^*(i)}\right)\right| \\ &\leq \alpha \max_{k-N-3\leq i\leq k-2}\left|\tilde{x}(\hat{t}_i)\right| \end{aligned} \tag{40}$$

where $0<\alpha<1$. Finally, applying the statement of Lemma 2, given in Appendix B, to (40) aids us in the completion of the proof as it shows that $|\tilde{x}(\hat{t}_k)|\to 0$ as $k\to\infty$.

In order to compute the exponential decay bound, note that (48) in the statement of Lemma 2, with $a_i=\tilde{x}(\hat{t}_i)$, leads to the following inequality for $i\geq 1$:

$$\left|\tilde{x}(\hat{t}_i)\right| \leq \frac{1}{\alpha}e^{-\frac{\ln(\alpha^{-1})}{(N+3)T_B}(\hat{t}_i-\hat{t}_1)}\max_{1\leq i\leq N+3}\left|\tilde{x}(\hat{t}_i)\right| \tag{41}$$

because $\hat{t}_{k+1}-\hat{t}_k\leq T_B$ and $0<\alpha<1$. And, since $e^{-(\ln(\alpha^{-1})/(N+3)T_B)(t-\hat{t}_i)}\geq\alpha$ for $t\geq\hat{t}_i$, it follows from (31) that for each $t\in[\hat{t}_i,\hat{t}_{i+1})$,

$$|\tilde{x}(t)|\leq e^{b_A T_B}\cdot b_E^{J_{\max}(T_B)}\cdot\frac{1}{\alpha}e^{-\frac{\ln(\alpha^{-1})}{(N+3)T_B}(t-\hat{t}_i)}\left|\tilde{x}(\hat{t}_i)\right|. \tag{42}$$

Combining (41) and (42), it holds for $t\geq\hat{t}_1$ that

$$|\tilde{x}(t)|\leq e^{b_A T_B}\cdot b_E^{J_{\max}(T_B)}\cdot\frac{1}{\alpha^2}e^{-\frac{\ln(\alpha^{-1})}{(N+3)T_B}(t-\hat{t}_1)}\max_{1\leq i\leq N+3}\left|\tilde{x}(\hat{t}_i)\right|. \tag{43}$$

On the other hand, since $t_{N+1}<\hat{t}_1\leq t_{N+1}+T_C$, an over-approximation of the error on the interval $[t_0,\hat{t}_{N+3}]$ is obtained, by ignoring the error updates, as

$$\max_{t\in[t_0,\hat{t}_{N+3}]}|\tilde{x}(t)|\leq\tilde{c}\cdot|\tilde{x}(t_0)| \tag{44}$$

where the constant $\tilde{c}:=e^{b_A((N+1)T_D+T_C+(N+2)T_B)}\cdot b_E^{N+1+J_{\max}(T_C+(N+2)T_B)}$. From (43) and (44), we arrive at the following, for $t\geq t_0$:

$$|\tilde{x}(t)|\leq e^{b_A T_B}\cdot b_E^{J_{\max}(T_B)}\cdot\frac{1}{\alpha^2}e^{-\frac{\ln(\alpha^{-1})}{(N+3)T_B}(t-\hat{t}_1)}\cdot\tilde{c}\cdot|\tilde{x}(t_0)| \tag{45}$$

in which it should be noted that the inequality holds for all $t\geq t_0$ because the right-hand side is greater than $\tilde{c}|\tilde{x}(t_0)|$ for $t\in[t_0,\hat{t}_1]$. Taking $\gamma=1/((N+3)T_B)$ and $h(\alpha)=(e^{b_A T_B}b_E^{J_{\max}(T_B)}\tilde{c})/\alpha^2\cdot e^{-\gamma\ln(\alpha^{-1})(t_0-\hat{t}_1)}$, the proof is completed. ∎

Example 5: We demonstrate the operation of the proposed observer for the switched system considered in Example 1 with $\epsilon>0$. We assume that each mode is activated for τ seconds and $\tau\neq\kappa\pi$ for any $\kappa\in\mathbb{N}$, so that, for each nonnegative integer m, mode a is active over $[2m\tau,(2m+1)\tau)$ (called *odd* interval henceforth), and mode $b$for $[(2m+1)\tau,(2m+2)\tau)$ (called *even* interval). We also use the notation q_o,k_o for odd positive integers and q_e,k_e for even positive integers. As mentioned earlier, the system is observable (and thus, determinable) over a time interval with the given switching signal if the mode sequence $a\to b\to a$ is contained in that interval. Hence, we pick $N=3$ in order to include both sequences (a,b,a,b) and (b,a,b,a), so that Assumption 1.2 holds. For simplicity, it is assumed that $0<T_C<\tau$ and that the computations always end at $\hat{t}_k=t_q+T_C$ with $q=k+N$ (because $\hat{t}_1=t_4+T_C$). With an arbitrary initial condition $\hat{x}(0)$, the observer to be implemented is

$$\left.\begin{aligned}\dot{\hat{x}}(t)&=A_a\hat{x}(t)\\ \hat{y}(t)&=C_a\hat{x}(t)\end{aligned}\right\},\quad t\in[2m\tau,(2m+1)\tau) \tag{46a}$$

$$\left.\begin{aligned}\dot{\hat{x}}(t)&=A_b\hat{x}(t)\\ \hat{y}(t)&=C_b\hat{x}(t)\end{aligned}\right\},\quad t\in[(2m+1)\tau,(2m+2)\tau) \tag{46b}$$

$$\hat{x}(\hat{t}_k)=\hat{x}\left(\hat{t}_k^-\right)-\xi_k,\quad \hat{t}_k=t_{k+3}+T_C,\quad k\in\mathbb{N}. \tag{46c}$$

In order to determine the value of ξ_k, we start off with the estimators for the observable part of each subsystem, denoted by z^q in (20). Note that mode a has a one-dimensional unobservable subspace whereas for mode b, the unobservable subspace is $\mathbb{R}^2$. Since mode a is active on every odd interval and mode b on every even interval, z^{q_o} represents the partial information obtained from mode a, and z^{q_e} is a null vector as no information is gathered from mode b. So the one-dimensional partial observer in (21) is implemented only for odd intervals. From mode a, we compute

$$G_{q_o}=\begin{bmatrix}1&0\\0&0\end{bmatrix},\quad W^{q_o}=\begin{bmatrix}0\\1\end{bmatrix},\quad Z^{q_o}=\begin{bmatrix}1\\0\end{bmatrix}$$

so that $S_{q_o}=0$ and $R_{q_o}=1$, which yields the observer in (21) as

$$\begin{aligned}\dot{\hat{z}}^{q_o}&=-l_{q_o}\hat{z}^{q_o}+l_{q_o}\tilde{y},\quad t\in[(q_o-1)\tau,q_o\tau)\\ \hat{z}^{q_o}(\hat{t}_k)&=\hat{z}^{q_o}\left(\hat{t}_k^-\right)-\xi_k^1,\quad \text{if } \hat{t}_k\in[(q_o-1)\tau,q_o\tau),\ k\in\mathbb{N}\end{aligned}$$

with the initial condition $\hat{z}^{q_o}((q_o-1)\tau)=0$, and $\tilde{y}$ being the difference between the measured output and the estimated output of (46). The notation ξ_k^1 denotes the first component of the vector ξ_k. The gain l_{q_o} will be chosen later by (47). From mode b, we get $W^{q_e}=I_{2\times 2}$, and $G_{q_e}=0_{2\times 2}$, so that Z^{q_e}, S_{q_e}, and R_{q_e} are null-matrices.

The next step is to use the value of $\hat{z}^{q_o}(t_{q_o}^-)$ to compute ξ_k, for each $k\in\mathbb{N}$. The matrices appearing in the computation of ξ_k are given as follows. For every $q_e>3$

$$\begin{aligned}\Psi_{q_e-3}^{q_e}&=e^{2\epsilon\tau}\begin{bmatrix}\cos 2\tau&\sin 2\tau\\-\sin 2\tau&\cos 2\tau\end{bmatrix}\Rightarrow\Theta_{q_e-3}^{q_e}=\begin{bmatrix}\cos 2\tau\\-\sin 2\tau\end{bmatrix}\\ \Psi_{q_e-2}^{q_e}&=e^{\epsilon\tau}\begin{bmatrix}\cos\tau&\sin\tau\\-\sin\tau&\cos\tau\end{bmatrix}\Rightarrow\Theta_{q_e-2}^{q_e}=\text{null}\\ \Psi_{q_e-1}^{q_e}&=e^{\epsilon\tau}\begin{bmatrix}\cos\tau&\sin\tau\\-\sin\tau&\cos\tau\end{bmatrix}\Rightarrow\Theta_{q_e-1}^{q_e}=\begin{bmatrix}\cos\tau\\-\sin\tau\end{bmatrix}\\ \Psi_{q_e}^{q_e}&=I_{2\times 2}\Rightarrow\Theta_{q_e}^{q_e}=\text{null}\end{aligned}$$

where, as a convention, we have taken $\Theta_i^{q_e}$ as a null matrix whenever $\mathcal{R}(\Psi_i^{q_e}W^i)^\perp=\{0\}$. Using the matrices $\Theta_j^{q_e}$, $j=q_e-3,\ldots,q_e$, we obtain for every $q_e>3$

$$\Theta_{q_e}=\begin{bmatrix}\Theta_{q_e-1}^{q_e}&\Theta_{q_e-3}^{q_e}\end{bmatrix}=\begin{bmatrix}\cos\tau&\cos 2\tau\\-\sin\tau&-\sin 2\tau\end{bmatrix}.$$

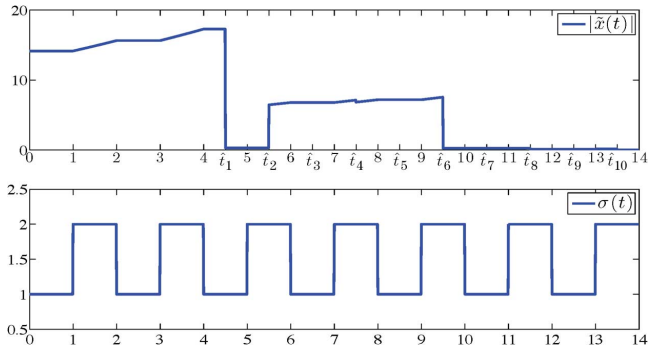

Fig. 3. Size of the state estimation error and the switching signal.

Since for every $q_e > 3$, $\Phi(\hat{t}^-_{k_o}, t^-_{q_e}) = I_{2\times 2}$ with $k_o = q_e - 3$, the error correction term ξ_{k_o} can be computed recursively by the formula

$$\xi_{k_o} = \hat{\tilde{x}}\left(t^-_{k_o+3}\right) = \hat{\tilde{x}}\left(t^-_{q_e}\right) = \Theta^{-\top}_{q_e}\mathrm{col}\left(\eta^1_{q_e}, \eta^2_{q_e}\right)$$

where $\eta^1_{q_e} := e^{\epsilon\tau}\hat{z}^{q_e-1}(t^-_{q_e-1}) - \Theta^{q_e^\top}_{q_e-1}e^{A_b(\tau-T_C)}\xi_{q_e-4}$, $\eta^2_{q_e} := e^{2\epsilon\tau}\hat{z}^{q_e-3}(t^-_{q_e-3}) - \Theta^{q_e}_{q_e-3}(e^{A_b(\tau-T_C)}\xi_{q_e-4} + e^{A_b\tau}\xi_{q_e-5} + e^{A_b(2\tau-T_C)}\xi_{q_e-6})$, and $\xi_k = 0$ for $k \le 0$.

Next, for every $q_o > 3$, we repeat the same calculations and obtain

$$\Theta_{q_o} = \begin{bmatrix}\Theta^{q_o}_{q_o} & \Theta^{q_o}_{q_o-2}\end{bmatrix} = \begin{bmatrix}1 & \cos\tau \\ 0 & -\sin\tau\end{bmatrix}.$$

Using $\Phi(\hat{t}^-_{k_e}, t^-_{q_e}) = e^{A_bT_C}$ with $k_e = q_o - 3$, we obtain

$$\xi_{k_e} = e^{A_bT_C}\hat{\tilde{x}}\left(t^-_{k_e+3}\right)$$

where

$$\hat{\tilde{x}}\left(t^-_{q_o}\right) = \Theta^{-\top}_{q_o}\mathrm{col}\left(\eta^1_{q_o}, \eta^2_{q_o}\right), \quad q_o > 3$$

and $\eta^1_{q_o} := \hat{z}^{q_o}(t^-_{q_o})$, $\eta^2_{q_o} := e^{\epsilon\tau}\hat{z}^{q_o-2}(t^-_{q_o-2}) - \Theta^{q_o^\top}_{q_o-2}(\xi_{q_o-1} + e^{A_b(\tau-T_C)}\xi_{q_o-2})$.

To compute the gain l_{q_o}, we note that $M^{q_e}_{q_e}$, $M^{q_e}_{q_e-2}$ are null matrices, and

$$M^{q_e}_{q_e-1} = e^{\epsilon\tau}\begin{bmatrix}\frac{\sin 2\tau}{\sin\tau} & 0 \\ \frac{\cos 2\tau}{\sin\tau} & 0\end{bmatrix} \text{ and } M^{q_e}_{q_e-3} = e^{2\epsilon\tau}\begin{bmatrix}-1 & 0 \\ -\frac{\cos\tau}{\sin\tau} & 0\end{bmatrix}.$$

Also, for $q_o > 3$, $M^{q_o}_{q_o-1}$ and $M^{q_o}_{q_o-3}$ are null matrices, and

$$M^{q_o}_{q_o} = \begin{bmatrix}1 & 0 \\ \frac{\cos\tau}{\sin\tau} & 0\end{bmatrix} \text{ and } M^{q_o}_{q_o-2} = \begin{bmatrix}0 & 0 \\ -\frac{e^{\epsilon\tau}}{\sin\tau} & 0\end{bmatrix}.$$

By taking l_{q_o} equal to l for each q_o, and computing the induced 2-norm of the matrix, it is seen that, $\max_{q-3\le j\le q, j:\mathrm{odd}, q>3}\|M^q_j Z^j e^{(S_j - lR_j)\tau} Z^{j\top}\| = e^{(2\epsilon-l)\tau}/|\sin\tau|$. Also, $\overline{b} = e^{\sqrt{1+\epsilon^2}(2T_C+\tau)}$. So, the lower bound for the gain l, is obtained as follows:

$$\overline{b}\frac{e^{(2\epsilon-l)\tau}}{|\sin\tau|} < \frac{1}{N+2} = \frac{1}{5} \quad \Rightarrow \quad l > 2\epsilon + \frac{1}{\tau}\ln\frac{5\overline{b}}{|\sin\tau|}. \quad (47)$$

Once again it can be seen that the singularity occurs when τ is an integer multiple of π. Moreover, if τ approaches this singularity, then the gain required for convergence gets arbitrarily large. This shows that even though the condition $\sin\tau \neq 0$ guarantees observability, it may cause some difficulty in practice if $\sin\tau \approx 0$. This also explains why the knowledge of the switching signal is required in general to compute the observer gains.

The results of simulations with $\tau = 1$, $T_C = 0.5\tau$, $\epsilon = 0.1$ and $l = 20$, are illustrated in Fig. 3. The error initially evolves according to the unstable system dynamics as no correction is applied till $t_4 + T_C$. The figure clearly shows the hybrid nature of the proposed observer, which is caused by the jump discontinuity in the error signal. The error grows between the error updates because the subsystem at mode b has unstable dynamics, but $\max_{k\le i\le k+N+2}|\tilde{x}(\hat{t}_i)|$ indeed gets smaller as k increases.◁

IV. Conclusion

This paper addressed the characterization of observability and determinability in switched linear systems with state jumps. It was shown that, for a fixed mode sequence, the set of switching signals over which these properties hold is either empty or dense under a certain metric topology. To study when the properties hold uniformly with respect to switching times, we derived separate sufficient and necessary conditions as corollaries to the main result. Later, using the property of determinability, an asymptotic observer was constructed that combines the partial information obtained from each mode to get an estimate of the state vector. For practical considerations, the proposed observer takes into account the time consumed in processing the information. Under the assumption of persistent switching, the error analysis shows that the estimate indeed converges to the actual state exponentially.

As an extension to the current work, it may be interesting to investigate how far these ideas carry over to nonlinear systems. The proposed method for observer design relies on the linearity of the system (1), in the sense that one can easily compute exact solutions of the linear system with zero inputs. In fact, it is seen in (34) that the transportation of the partially observable state information (represented by z), obtained at each mode, can be computed even with some unobservable information (by w). Since linearity guarantees that the observable information is not altered by this transportation process, the unobservable components are simply filtered out after the transportation. We emphasize that this idea may not be transparently applied to nonlinear systems, and may need a different approach as in [13] and [14]. However, the geometric approach adopted in this paper has been applied to the study of observability in another class of switched dynamical systems that comprise algebraic constraints [21].

Another interesting direction of research could be to make the observer design robust to uncertainties in the system. These uncertainties may manifest in two forms: perturbations in the model of the system, and perturbations in the signals used for constructing the state estimate. For example, in our work, the observer assumed exact knowledge of the output and the switching signal. It may happen that the output available to the observer is quantized, or the switching times are not known

exactly. Addressing such issues is nontrivial and requires further research.

APPENDIX

A. Some Useful Properties

Let $\mathcal{V}_1$, $\mathcal{V}_2$, and $\mathcal{V}$ be any linear subspaces, A be a (not necessarily invertible) $n \times n$ matrix, and B, C be matrices of suitable dimension. The following properties can be found in the literature such as [23], or developed with little effort.

1) $A\mathcal{R}(B) = \mathcal{R}(AB)$ and $A^{-1} \ker B = \ker(BA)$.
2) $A^{-1}A\mathcal{V} = \mathcal{V} + \ker A$, and $AA^{-1}\mathcal{V} = \mathcal{V} \cap \mathcal{R}(A)$.
3) $A^{-1}(\mathcal{V}_1 \cap \mathcal{V}_2) = A^{-1}\mathcal{V}_1 \cap A^{-1}\mathcal{V}_2$, and $A(\mathcal{V}_1 \cap \mathcal{V}_2) \subseteq A\mathcal{V}_1 \cap A\mathcal{V}_2$ (with equality if and only if $(\mathcal{V}_1+\mathcal{V}_2)\cap \ker A = \mathcal{V}_1 \cap \ker A + \mathcal{V}_2 \cap \ker A$, which holds, in particular, for any invertible A).
4) $A\mathcal{V}_1 + A\mathcal{V}_2 = A(\mathcal{V}_1 + \mathcal{V}_2)$, and $A^{-1}\mathcal{V}_1 + A^{-1}\mathcal{V}_2 \subseteq A^{-1}(\mathcal{V}_1 + \mathcal{V}_2)$ (with equality if and only if $(\mathcal{V}_1 + \mathcal{V}_2) \cap \mathcal{R}(A) = \mathcal{V}_1 \cap \mathcal{R}(A) + \mathcal{V}_2 \cap \mathcal{R}(A)$, which holds, in particular, for any invertible A).
5) $(\ker A)^{\perp} = \mathcal{R}(A^{\top})$.
6) $(A^{\top}\mathcal{V})^{\perp} = A^{-1}\mathcal{V}^{\perp}$ and $(A^{-1}\mathcal{V})^{\perp} = A^{\top}\mathcal{V}^{\perp}$.
7) $\langle A|\mathcal{V}\rangle = \mathcal{V} + A\mathcal{V} + A^2\mathcal{V} + \cdots + A^{n-1}\mathcal{V}$ and $\langle \mathcal{V}|A\rangle = \mathcal{V} \cap A^{-1}\mathcal{V} \cap A^{-2}\mathcal{V} \cap \cdots \cap A^{-(n-1)}\mathcal{V}$.
8) $\langle \mathcal{V}_1 \cap \mathcal{V}_2|A\rangle = \langle \mathcal{V}_1|A\rangle \cap \langle \mathcal{V}_2|A\rangle$ and $\langle A|\mathcal{V}_1 \cap \mathcal{V}_2\rangle \subset \langle A|\mathcal{V}_1\rangle \cap \langle A|\mathcal{V}_2\rangle$.
9) $e^{At}\mathcal{V} \subseteq \langle A|\mathcal{V}\rangle$ and $\langle \mathcal{V}|A\rangle \subseteq e^{At}\mathcal{V}$ for any t.
10) $\langle A|\mathcal{V}\rangle^{\perp} = \langle \mathcal{V}^{\perp}|A^{\top}\rangle$.
    Now, with $G := \mathrm{col}(C, CA, \ldots, CA^{n-1})$,
11) $e^{At} \ker G = \ker G$ and $e^{A^{\top}t}\mathcal{R}(G^{\top}) = \mathcal{R}(G^{\top})$ for all t.
12) $\langle \ker G|A\rangle = \ker G$ and $\langle A^{\top}|\mathcal{R}(G^{\top})\rangle = \mathcal{R}(G^{\top})$.

B. Proofs

Proof of Theorem 2: It is first shown that, for any $\sigma^* \in \mathcal{S}^*$, a neighborhood of σ^* is also contained in $\mathcal{S}^*$. Recalling the expression for $\mathcal{N}_1^m$ from (5b), introduce the following matrix:

$$W(\tau) := \mathrm{col}\left(G_1, G_2E_1e^{A_1\tau_1}, \cdots, G_m \prod_{l=m-1}^{1} E_l e^{A_l\tau_l}\right)$$

and let $\overline{W}(\tau) := W^{\top}(\tau)W(\tau)$. Note that $\mathcal{N}_1^m(\tau) = \ker W(\tau)$, so that $\mathcal{N}_1^m(\tau) = \{0\}$ if, and only if, $W(\tau)$ has full column rank, or equivalently $\psi(\tau) \neq 0$, where $\psi : \mathcal{T} \to \mathbb{R}$ denotes the determinant of the matrix $\overline{W}(\tau)$. Since $\overline{W}(\tau)$ comprises analytic functions of τ, the determinant $\psi(\tau)$ is also an analytic function. It is well-known that an analytic function is either identically zero, or the set comprising zeros of an analytic function has an empty interior [7, Ch. 4]. Therefore, the set $\mathcal{Z} := \{\tau \in \mathcal{T} : \psi(\tau) = 0\}$ has an empty interior (with respect to the topology induced by ℓ_1 norm), and is closed. Hence, the set $\mathcal{T} \setminus \mathcal{Z}$ is open and there exists an $\varepsilon > 0$ such that $\psi(\tau) \neq 0$ for each $\tau \in \mathcal{B}_{\varepsilon}(\tau^*) := \{\tau \in \mathcal{T} : \|\tau - \tau^*\|_1 < \varepsilon\}$, where τ^* is associated with σ^* and satisfies $\psi(\tau^*) \neq 0$ since $\sigma^* \in \mathcal{S}^*$. Now pick any $\sigma \in \mathcal{S}$ such that $d(\sigma, \sigma^*) < \varepsilon$. Then, the corresponding τ belongs to $\mathcal{B}_{\varepsilon}(\tau^*)$, which implies $\sigma \in \mathcal{S}^*$ showing that $\mathcal{S}^*$ is open.

Next, to show the denseness of $\mathcal{S}^*$, we pick $\sigma' \in \mathcal{S} \setminus \mathcal{S}^*$, and show that σ' is the limit point of $\mathcal{S}^*$. In this case, $\psi(\tau') = 0$, and $\tau' \in \mathcal{Z}$. Since $\mathcal{Z}$ has an empty interior, for every $\varepsilon > 0$, there exists $\tau'' \in \mathcal{B}_{\varepsilon}(\tau')$ such that $\psi(\tau'') \neq 0$. Let σ'' be the switching signal corresponding to τ'', then $\sigma'' \in \mathcal{S}^*$, proving that every neighborhood of σ', with respect to metric $d(\cdot,\cdot)$, has a non-empty intersection with $\mathcal{S}^*$. ∎

Proof of Lemma 1: From the inclusion relation of (9), it suffices to show that $\langle A_q|\mathcal{V}\rangle$ is the smallest subspace containing $\cup_{\tau_q>0} e^{-A_q\tau_q}\mathcal{V}$ where $\mathcal{V} := \ker G_q \cap E_q^{-1}\mathcal{V}_{q+1}^m = \ker G_q \cap E_q^{-1}\overline{\mathcal{N}}_{q+1}^m$. Let $\mathcal{W}$ denote the smallest subspace containing $\cup_{\tau_q>0} e^{-A_q\tau_q}\mathcal{V}$. Then, since $\langle A_q|\mathcal{V}\rangle$ is a subspace containing $\cup_{\tau_q>0} e^{-A_q\tau_q}\mathcal{V}$ by Property 9 in Appendix A, it follows that $\mathcal{W} \subseteq \langle A_q|\mathcal{V}\rangle$. Next, pick any $x \in \mathcal{W}^{\perp}$ and let V be a matrix such that $\mathcal{V} = \mathcal{R}(V)$. From the definition of $\mathcal{W}$, it follows that $x^{\top}e^{-A_q\tau_q}V = 0$ for all $\tau_q > 0$, but by continuity, it also holds for all $\tau_q \geq 0$. Repeated differentiation of both sides at $\tau_q = 0$ leads to $x^{\top}A_q^iV = 0$ for $i = 0, 1, \ldots, n-1$, or equivalently $x \in \langle A_q|\mathcal{V}\rangle^{\perp}$ by Property 7 in Appendix A. This shows that $\mathcal{W}^{\perp} \subseteq \langle A_q|\mathcal{V}\rangle^{\perp}$, and hence, $\mathcal{W} = \langle A_q|\mathcal{V}\rangle$. ∎

Lemma 2: Suppose that the sequence $\{a_k\}$ satisfies

$$|a_k| \leq \alpha \max_{k-N-3\leq i\leq k-2} |a_i|, \quad k > N+3$$

where $0 < \alpha < 1$. Then the following holds:

$$\max_{k\leq i\leq k+N+2} |a_i| \leq \alpha \max_{k-N-3\leq i\leq k-1} |a_i|, \quad k > N+3 \tag{48}$$

which implies that the maximum value of the sequence $\{a_k\}$ over a window of length $N+3$ is strictly decreasing and converging to zero, and thus, $\lim_{k\to\infty} a_k = 0$.

Proof of Lemma 2: By putting $|a_{k-1}|$ on the right-hand side, it is clear that

$$|a_k| \leq \alpha \max_{k-N-3\leq i\leq k-1} |a_i|. \tag{49}$$

Similarly, it follows that

$$\begin{aligned} |a_{k+1}| &\leq \alpha \max_{k-N-2\leq i\leq k} |a_i| \\ &\leq \alpha \max\left\{|a_{k-N-3}|, \max_{k-N-2\leq i\leq k-1} |a_i|, |a_k|\right\} \\ &\leq \alpha \max_{k-N-3\leq i\leq k-1} |a_i| \end{aligned}$$

where the last inequality follows from (49). By induction, this leads to

$$\max_{k\leq i\leq k+N+2} |a_i| \leq \alpha \max_{k-N-3\leq i\leq k-1} |a_i|.$$

∎

REFERENCES

[1] A. Alessandri and P. Coletta, "Switching observers for continuous-time and discrete-time linear systems," in *Proc. Amer. Control Conf.*, 2001, pp. 2516–2521.

[2] M. Babaali and G. J. Pappas, "Observability of switched linear systems in continuous time," in *Hyb. Syst: Comput. and Control*, M. Morari and L. Thiele, Eds. Berlin/Heidelberg, Germany: Springer, 2005, vol. 3414, Lecture Notes in Computer Science, pp. 103–117.

[3] A. Balluchi, L. Benvenuti, M. D. Di Benedetto, and A. L. Sangiovanni-Vincentelli, "Observability for hybrid systems," in *Proc. IEEE Conf. Decision Control*, 2003, pp. 1159–1164.

[4] A. Balluchi, L. Benvenuti, M. D. Di Benedetto, and A. L. Sangiovanni-Vincentelli, "Design of observers for hybrid systems," in *Hyb. Syst: Comput. and Control*, C. Tomlin and M. Greenstreet, Eds. Berlin/Heidelberg, Germany: Springer, 2002, vol. 2289, Lecture Notes in Computer Science, pp. 76–89.
[5] P. Collins and J. H. van Schuppen, "Observability of piecewise-affine hybrid systems," in *Hyb. Syst: Comput. and Control*, R. Alur and G. J. Pappas, Eds. Berlin/Heidelberg, Germany: Springer, 2004, vol. 2993, Lecture Notes in Computer Science, pp. 265–279.
[6] M. Fliess, C. Join, and W. Perruquetti, "Real-time estimation for switched linear systems," in *Proc. IEEE Conf. Decision Control*, 2008, pp. 941–946.
[7] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*, 2nd ed. Boston, MA: Birkhäuser, 2002.
[8] E. A. Medina and D. A. Lawrence, "Reachability and observability of linear impulsive systems," *Automatica*, vol. 44, no. 5, pp. 1304–1309, 2008.
[9] E. A. Medina and D. A. Lawrence, "State estimation for linear impulsive systems," in *Proc. Amer. Control Conf.*, 2009, pp. 1183–1188.
[10] C. Moler and C. van Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later," *SIAM Rev.*, vol. 45, no. 1, pp. 3–49, 2003.
[11] F. M. Pait and A. S. Morse, "A cyclic switching strategy for parameter adaptive control," *IEEE Trans. Autom. Control*, vol. 39, no. 6, pp. 1172–1183, 1994.
[12] S. Pettersson, "Designing switched observers for switched systems using multiple Lyapunov functions and dwell-time switching," in *Proc. IFAC Conf. Anal. Design of Hybrid Syst.*, 2006, pp. 18–23.
[13] H. Shim and A. Tanwani, "On a sufficient condition for observability of nonlinear switched systems and observer design strategy," in *Proc. Amer. Control Conf.*, 2011, pp. 1206–1211.
[14] H. Shim and A. Tanwani, "Hybrid-type observer design based on a sufficient condition for observability in switched nonlinear systems," *Int. J. Robust and Nonlinear Control*, 2012, Special Issue: High-Gain Observers in Nonlinear Feedback Control., to be published.
[15] E. D. Sontag, *Mathematical Control Theory: Finite Dimensional Deterministic Systems*, 2nd ed. New York: Springer, 1998.
[16] Z. Sun and S. S. Ge, *Switched Linear Systems: Control and Design*. London, U.K.: Springer-Verlag, 2005.
[17] H. Sussmann and P. V. Kokotovic, "The peaking phenomenon and the global stabilization of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 36, no. 4, pp. 424–440, Apr. 1991.
[18] A. Tanwani, "Invertibility and observability of switched systems with inputs and outputs" Ph.D. dissertation, Dept. Elect. Comp., Univ. of Illinois, Urbana-Champaign, USA, 2011 [Online]. Available: http://hdl.handle.net/2142/29755
[19] A. Tanwani and D. Liberzon, "Robust invertibility of switched linear systems," in *Proc. IEEE Conf. Decision Control*, 2011, pp. 441–446.
[20] A. Tanwani, H. Shim, and D. Liberzon, "Observability implies observer design for switched linear systems," in *Proc. ACM Conf. Hyb. Syst.: Comput. and Control*, 2011, pp. 3–12.
[21] A. Tanwani and S. Trenn, "On observability of switched differential algebraic equations," in *Proc. IEEE Conf. Decision Control*, 2010, pp. 5656–5661.
[22] R. Vidal, A. Chiuso, S. Soatto, and S. Sastry, "Observability of linear hybrid systems," in *Hyb. Sys: Comp. & Control*, O. Maler and A. Pnueli, Eds. Berlin, Germany: Springer, 2003, vol. 2623, Lecture Notes in Computer Science, pp. 526–539.
[23] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, 3rd ed. New York: Springer-Verlag, 1985.
[24] G. Xie and L. Wang, "Necessary and sufficient conditions for controllability and observability of switched impulsive control systems," *IEEE Trans. Autom. Control*, vol. 49, no. 6, pp. 960–966, Jun. 2004.
[25] S. Zhao and J. Sun, "A geometric approach for reachability and observability of linear switched impulsive systems," *Nonlinear Anal.*, vol. 72, pp. 4221–4229, 2010.

Aneel Tanwani (S'08–M'13) received the undergraduate degree in mechatronics engineering from National University of Sciences and Technology, Islamabad, Pakistan, in 2005, the D.E.A. degree in Automatique-Productique from the Institut National Polytechnique de Grenoble, Grenoble, France, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, under the supervision of Prof. Liberzon.

He is currently working as a Post-Doctoral Researcher at INRIA Rhône-Alpes, France. His research interests include switched and hybrid dynamical systems, nonlinear control, systems theory, and related applications.

Dr. Tanwani was a recipient of a Fulbright scholarship in 2006.

Hyungbo Shim (M'00) received the B.S., M.S., and Ph.D. degrees from Seoul National University, Seoul, Korea, in 1993, 1995, and 2000, respectively.

From 2000 to 2001, he was a Post-Doctoral Researcher at the Center for Control Engineering and Computation, University of California, Santa Barbara. In 2002, he joined the faculty of the Division of Electrical and Computer Engineering, Hanyang University, Seoul, Korea. Since 2003, he has been with the School of Electrical Engineering at Seoul National University, Korea. His research interests include analysis and control of nonlinear systems. He had been an Associate Editor for *European Journal of Control*, and is currently for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and *International Journal of Robust and Nonlinear Control*.

Daniel Liberzon (M'98–SM'04–F'13) was born in the former Soviet Union in 1973. He did his undergraduate studies in the Department of Mechanics and Mathematics at Moscow State University, Moscow, Russia, from 1989 to 1993 and received the Ph.D. degree in mathematics from Brandeis University, Waltham, MA, in 1998 (supervised by Prof. Roger W. Brockett of Harvard University).

Following a postdoctoral position in the Department of Electrical Engineering at Yale University from 1998 to 2000 (with Prof. A. Stephen Morse), he joined the University of Illinois at Urbana-Champaign, where he is now a Professor in the Electrical and Computer Engineering Department and the Coordinated Science Laboratory. His research interests include nonlinear control theory, switched and hybrid dynamical systems, control with limited information, and uncertain and stochastic systems. He is the author of the books *Switching in Systems and Control* (Birkhauser, 2003) and *Calculus of Variations and Optimal Control Theory: A Concise Introduction* (Princeton University Press, 2011).

Dr. Liberzon received several recognitions, including the 2002 IFAC Young Author Prize and the 2007 Donald P. Eckman Award. He delivered a plenary lecture at the 2008 American Control Conference. From 2007 to 2011, he has served as Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.