Determinability and state estimation for switched differential-algebraic equations

Aneel Tanwani, Stephan Trenn

Determinability and state estimation for switched differential-algebraic equations

Aneel Tanwani[a], Stephan Trenn[b]

[a]*LAAS–CNRS, University of Toulouse, 31400 Toulouse, France.*
[b]*Department of Mathematics, University of Kaiserslautern, 67663 Kaiserslautern, Germany.*

Abstract

The problem of state reconstruction and estimation is considered for a class of switched dynamical systems whose subsystems are modeled using linear differential-algebraic equations (DAEs). Since this system class imposes time-varying dynamic and static (in the form of algebraic constraints) relations on the evolution of state trajectories, an appropriate notion of observability is presented which accommodates these phenomena. Based on this notion, we first derive a formula for the reconstruction of the state of the system where we explicitly obtain an injective mapping from the output to the state. In practice, such a mapping may be difficult to realize numerically and hence a class of estimators is proposed which ensures that the state estimate converges asymptotically to the real state of the system.

1. Introduction

Switched differential-algebraic equations (DAEs) form an important class of switched systems, where the dynamics are not only discontinuous with respect to time, but also the state trajectories are constrained by certain algebraic equations which may also change as the system switches from one mode to another. Such dynamical models have found utility e.g. in the analysis of electrical power distribution (Gross et al., 2014) and of electrical circuits (Trenn, 2012). We consider switched DAEs of the following form

$$E_\sigma \dot{x} = A_\sigma x + B_\sigma u$$
$$y = C_\sigma x \tag{1}$$

where x, u, y denote the state (with dimension $n \in \mathbb{N}$), input (with dimension $\mathtt{u} \in \mathbb{N}$) and output (with dimension $\mathtt{y} \in \mathbb{N}$) of the system, respectively. The switching signal $\sigma : (t_0, \infty) \to \mathbb{N}$ is a piecewise constant, right-continuous function of time and in our notation it changes its value at time instants $t_1 < t_2 < \dots$ called the switching times. For a fixed σ, the triplet (x, u, y) is used to denote signals satisfying (1). We adopt the convention that over the interval $[t_p, t_{p+1})$ of length $\tau_p := t_{p+1} - t_p$, the active mode is defined by the quadruple $(E_p, A_p, B_p, C_p) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times \mathtt{u}} \times \mathbb{R}^{\mathtt{y} \times n}$, $p \in \mathbb{N}$. Over the interval (t_0, t_1), it is assumed that the system has some past which may be described by (E_0, A_0, B_0, C_0). The solution concept for switched DAEs is studied in (Trenn, 2009) and a brief discussion is also included in Section 3.1.

Dynamical system with discontinuous, or constrained trajectories have gathered a lot of interest in the control community, as they form an important class of hybrid, or discontinuous dynamical systems, see e.g. (Goebel et al., 2009). One

direction of research for these systems includes the study of structural properties that could be used for control design problems, and in this regard the problem of state reconstruction and estimation is of particular interest (Tanwani, 2011). In the current literature, one finds that the earlier work on observability and observers for switched/hybrid systems was aimed at using classical Luenberger observers for continuous dynamics and treat the switching or discontinuity as a perturbation that can be overcome by certain strong assumptions on system data. This line of work also requires that the underlying observability notion allows instantaneous recovery of the state from the measured output (Babaali and Pappas, 2005; Vidal et al., 2003). Modified Luenberger observers have also been used for constrained dynamical systems using similar observability concepts: see (Tanwani et al., 2014) when the constraint sets are convex (or, mildly nonconvex), and (Tanwani et al., 2016) when the constraints result in impacts.

However, for switched dynamical systems, several different notions of observability can be defined. In (Sun et al., 2002; Xie and Wang, 2003), a switched system comprising ordinary differential equations (ODEs) is called observable if there exists a switching signal that allows reconstruction of the initial state from the output measurements over an interval. This concept also appears in the observer construction (for continuous state and switching signal) proposed in (Balluchi et al., 2003). However, in our recent work, a more relaxed notion of observability has been proposed for switched ODEs (Tanwani et al., 2013) and switched DAEs (Tanwani and Trenn, 2012). The switching signal in this case is assumed to be known and fixed (i.e. playing the same role as the input u which is also assumed to be known and not influenced by the observer). By measuring the outputs and inputs over an interval, and using the data of subsystems activated during that interval, it is determined whether state reconstruction is possible or not. Several variants of this notion are

also collected in the survey (Petreczky et al., 2015). A state estimation algorithm based on these generalized observability concept can be found in (Shim and Tanwani, 2014; Tanwani et al., 2013, 2015) for switched ODEs. The main contribution of this paper is to address the observer design for switched DAEs which, apart from our preliminary work (Tanwani and Trenn, 2013), has not been addressed in this literature.

Already in the context of nonswitched systems several challenges arise in the study of DAEs. The difference basically arises due to the presence of algebraic equations (static relations) in the description of the system because of which the state trajectories can only evolve on the sets defined by the algebraic equations of the active mode. Observer designs have been studied for (nonswitched) DAEs since 1980's, e.g. (Dai, 1989; Fahmy and O'Reilly, 1989). Unlike ODEs, the observer design in DAEs requires additional structural assumptions and, furthermore, the order of the observer may depend on the design method. Because of these added generalities, observer design for nonswitched DAEs is still an active research field (Bobinyec et al., 2011; Darouach, 2012), and the recent survey articles summarize the development of this field (Berger and Reis, 2015; Bobinyec and Campbell, 2014).

In studying switched DAEs, our modeling framework allows for time-varying algebraic relations. The changes in algebraic constraints due to switching introduce jumps in the state of the system, and because of the possibility of a higher-index DAE, these jumps may get differentiated and generate impulsive solutions. The notion of observability studied in this paper takes into account the additional structure due to algebraic constraints, and the added information from the outputs in case there are impulses observed in the measurements. This observability concept is then used to construct a mapping from the output space to the state space, which allows us to theoretically reconstruct the state. The key element of constructing this mapping is to show how the structure of a linear DAE is exploited to recover the information about the state in individual subsystems. This structural decomposition is then combined with the expressions used for evolution of states in switched DAEs to accumulate all possible recoverable information from past measurements about the state at one time instant. The construction then yields a systematic procedure for writing the value of the state at a time instant in terms of outputs measured over an interval for which the system is observable.

The theoretical mappings constructed in the process are often not realizable in practice, but the derivation is used to describe a general class of state observers that generate asymptotically converging state estimates. The main result states that if the observable components of the individual subsystems can be estimated well-enough, and the required observability assumption persistently holds with time, then the estimates converge asymptotically. Our initial work on observers for switched DAEs (Tanwani and Trenn, 2013), and even the observers proposed for switched ODEs (Shim and Tanwani, 2014; Tanwani et al., 2013) can be seen as a special case of the general class of state estimators studied in this paper.

2. Contribution and Layout

This section provides a summary of all the technical results that are developed in this article, and a coherent view of how the different components are connected together to solve the state estimation problem for switched DAEs. The reader may also refer to this section as an index for finding the appropriate section for technical terms. The main contribution of this article is to provide a systematic procedure for designing observers for system class (1) which generate asymptotically convergent state estimates. The flow-diagram which shows the working of the proposed observer is given in Figure 1.
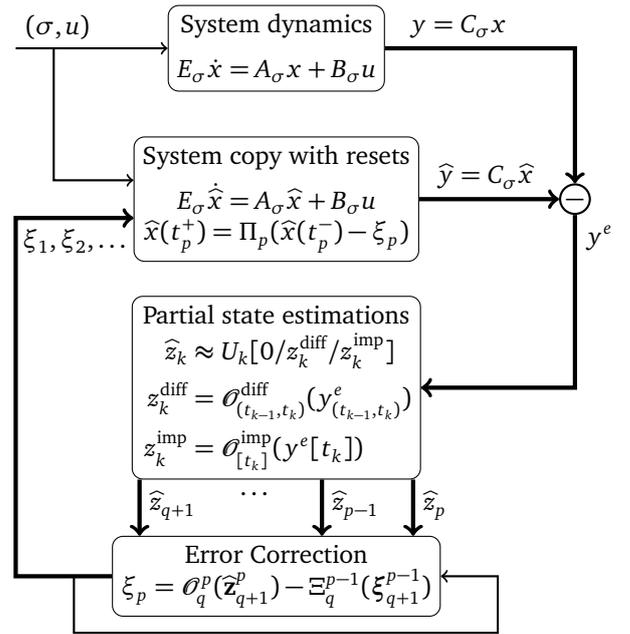

Figure 1: A schematic representation of the state estimator.

Our proposed observer basically comprises two layers. In the first layer, a system copy with variable $\widehat{x}$ (also the state estimate) is simulated with same time scale as the actual plant, and $\widehat{x}$ is reset at some switching times t_p by an error correction vector ξ_p. The second layer comprises an algorithm to compute ξ_p in very short (negligibly small) time. This article (in Sections 4–6) develops the theoretical tools and numerical certificates that gurantee efficient implementation of this algorithm. The central idea is that ξ_p closely approximates the estimation error prior to the reset times $e(t_p^-) := \widehat{x}(t_p^-) - x(t_p^-)$ by using the output measurements over the interval $(t_q, t_p]$, $q < p$. If this approximation can be achieved to a desired accuracy, and these sufficiently rich approximates are injected into the estimate's dynamics by resetting $\widehat{x}$ often enough, then our algorithms ensure that $\widehat{x}(t)$ converges to $x(t)$ as t tends to infinity. There are three primary mechanisms involved in achieving this desired goal:

- Determinability notion in Section 4 captures the property that, by measuring the output of (1) over an interval $(t_q, t_p]$, we must be able to recover the state of the system $x(t_p^+)$ exactly; and Theorem 1 provides conditions on system data to achieve this property.

- In Section 5, under the assumption that system (1) is determinable over the interval $(t_q, t_p]$, we build the theoretical engine to compute ξ_p. This involves looking at the observable components of the individual subsystems for the error dynamics denoted by z_k, $k = q+1, \cdots, p$. Using the structure of a DAE, we define the maps $\mathcal{O}_k^{\mathrm{diff}}$, $\mathcal{O}_k^{\mathrm{imp}}$, and decompose z_k into z_k^{cons}, $z_k^{\mathrm{diff}} := \mathcal{O}_k^{\mathrm{diff}}(y_{(t_{k-1}, t_k]}^e)$, $z_k^{\mathrm{imp}} := \mathcal{O}_k^{\mathrm{imp}}(y_{[t_k]}^e)$ based on whether they can be recovered from algebraic constraints, smooth output measurements, or impulsive output measurements, respectively. Theorem 2 then provides the construction of a matrix $\mathcal{O}_q^p$ such that

$$e(t_p^+) = \Pi_p \mathcal{O}_q^p \mathbf{z}_{q+1}^p,$$

where $\mathbf{z}_{q+1}^p := (z_{q+1}, \ldots, z_p)$, and Π_p is a projector depending on the system data (E_p, A_p).

- In Section 6, we show how to construct the estimate $\widehat{z}_k$ for the observable component of the error dynamics. These estimates are then used to define the vector ξ_p recursively (see Fig. 1 also). The final result, Theorem 3, proves the asymptotic convergence of $\widehat{x}(t)$ toward $x(t)$, as $t \to \infty$, in a rigorous manner.

3. Preliminaries

Before addressing the problem of observability, or state estimation, it helps to recall the solution framework associated with (1), and certain algebraic properties of the system matrices, that will form the foundation of the analysis carried out in this paper. In Section 3.1, we provide a motivation for the solution concept that is adopted for system class (1). Then, to analyze such solutions, certain properties of the matrix pairs (E_p, A_p), $p \in \mathbb{N}$, are described in Section 3.2.

3.1. Solution Concept for Switched DAEs

To discuss the solution concept for the switched DAEs described in (1), let us first consider a matrix pair (E, A) and the problem of finding a trajectory x which solves the following initial-trajectory problem (ITP):

$$x_{(-\infty,0)} = x_{(-\infty,0)}^0 \tag{2a}$$
$$(E\dot{x})_{[0,\infty)} = (Ax)_{[0,\infty)}, \tag{2b}$$

in some appropriate sense. In this equation, x^0 is some initial trajectory, and $f_\mathscr{I}$ denotes the restriction of f to an interval $\mathscr{I}$ and will be defined formally later. To avoid certain complications (such as nonuniquess of solutions), we limit our attention to *regular* matrix pairs (E, A), i.e. we assume that $\det(sE - A)$ is not the zero polynomial. A fundamental tool in studying (2) is the notion of *consistency space*, denoted $\mathfrak{C}_{(E,A)}$, and defined as

$$\mathfrak{C}_{(E,A)} := \left\{ x_0 \in \mathbb{R}^n \,|\, \exists\, x \in \mathscr{C}^1 : E\dot{x} = Ax \wedge x(0) = x_0 \right\}$$

where $\mathscr{C}^1$ is the space of differentiable functions $x : \mathbb{R} \to \mathbb{R}^n$. Intuitively speaking, $\mathfrak{C}_{(E,A)}$ corresponds to the set of initial

conditions which are consistent with the algebraic equations encoded in (2b), and thus a smooth solution evolving within $\mathfrak{C}_{(E,A)}$ can be obtained. An algebraic characterization of $\mathfrak{C}_{(E,A)}$ in terms of the matrices E, A will be given later in this section. At this point, it follows from the definition that if $x^0(\cdot)$ is absolutely continuous and $x^0(0^-) \in \mathfrak{C}_{(E,A)}$, then an absolutely continous trajectory x can be defined, such that (2) is satisfied, and furthermore $x(t) \in \mathfrak{C}_{(E,A)}$, for each $t \geq 0$. It is also shown in (Trenn, 2009) that such a solution is unique if, and only if, the matrix pair (E, A) is regular.

However, if $x^0(0^-) \notin \mathfrak{C}_{(E,A)}$, then it is not so obvious how the solution to (2) must be obtained. A valid solution x with $x(0^-) = x^0(0^-)$ must "jump" to $x(0^+) \in \mathfrak{C}_{(E,A)}$, and from there onwards, the solution can be extended uniquely over the interval $(0, \infty)$. In contrast to systems with ODEs which only comprise operations involving integration, the DAE (2b) allows the possibility of (higher-order) differentiation. For this reason, if a jump is introduced at time instant $t = 0$, then the notion of derivatives of jumps must also be introduced in the generalized solution concept associated with (2). This motivates us to enlarge the solution space from functions to *distributions*, because the derivative of a jump is then formally defined as the *Dirac impulse* (or Dirac delta). But the restriction to an interval as required in (2) cannot be defined for general distributions (Trenn, 2013, Lem. 5.1), and hence we consider the smaller space of *piecewise smooth distributions*, denoted by $\mathbb{D}_{\mathrm{pw}\mathscr{C}^\infty}$, for which (2) is indeed well defined. We refer the reader to (Trenn, 2009) for formal details, but for this paper, it suffices to recall that $x \in (\mathbb{D}_{\mathrm{pw}\mathscr{C}^\infty})^n$ is written as

$$x = x_\mathbb{D}^f + x[\cdot], \tag{3a}$$

where $x_\mathbb{D}^f$ denotes the distribution induced by the piecewise smooth function $x^f : \mathbb{R} \to \mathbb{R}^n$ and $x[\cdot]$ denotes the impulsive part of x given by

$$x[\cdot] = \sum_{k \in \mathbb{Z}} x[t_k] = \sum_{k \in \mathbb{Z}} \sum_{i=0}^{n_k} a_k^i \delta_{t_k}^{(i)}, \tag{3b}$$

where $\{\, t_k \in \mathbb{R} \mid k \in \mathbb{Z} \,\}$ is a strictly increasing set without finite accumulation and $\delta_{t_k}^{(i)}$ denotes the i-th derivative of the Dirac impulse with support at t_k. The $n_k + 1$ coefficients $a_k^0, a_k^1, \ldots, a_k^{n_k} \in \mathbb{R}^n$ parameterize the *impulsive part* of x at t_k, denoted by $x[t_k]$. Furthermore, the *left-* and *right-sided evaluations* $x(t^-)$, $x(t^+)$ are well defined for each $t \in \mathbb{R}$. Finally, a *distributional restriction* is well defined; in fact for x as in (3) the restriction to some interval $\mathscr{I} \subseteq \mathbb{R}$ is defined via

$$x_\mathscr{I} := (x_\mathscr{I}^f)_\mathbb{D} + \sum_{t_k \in \mathscr{I}} x[t_k], \tag{4}$$

where $x_\mathscr{I}^f(t) = x^f(t)$ for $t \in \mathscr{I}$ and $x_\mathscr{I}^f(t) = 0$ otherwise.

If in (2) it is assumed that $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathscr{C}^\infty})^n$, then we seek a solution x of the form (3), and in this particular case, the impulse times t_k either correspond to the impulse times of x^0, or the Dirac impulse that occurs at $t = 0$ due to the jump from an inconsistent initial condition to the consistency space.

3

When studying switched DAEs of the form (1), the consistency spaces for individual subsystems, in general, are different and affected by the presence of inputs as well. Hence, the switch from one subsystem to another may introduce jumps, and its higher order derivatives which we represent formally using distributions. One can draw an analogy between (2) and how the solution of a switched DAE (1) is extended beyond a switching time instant: The initial trajectory $x^0 \in \mathbb{D}_{\mathrm{pw}\mathscr{C}\infty}$ corresponds to some past over the interval (t_0, t_1), and then at t_1, the system first switches to a new subsystem described by a DAE, which would introduce a jump to the active consistency space, along with the possibility of Dirac impulses. In general, the state $x(t_p^-)$ may not be consistent with the consistency space of the subsystem (E_p, A_p, B_p, C_p) which is active over the interval $[t_p, t_{p+1})$, and the solution is extended by introducing a jump to the consistency space of the new subsystem, and introducing Dirac impulses at the switching times. The foregoing discussion motivates us to adopt the framework of piecewise smooth distributions to express solutions of a switched DAE.

Definition 1. *The triplet* (x, u, y) *is called a* (distributional) solution *of the switched DAE* (1) *on some interval* $\mathscr{I} \subseteq (t_0, \infty)$ *if, and only if,* $x \in (\mathbb{D}_{\mathrm{pw}\mathscr{C}\infty})^n$, $u \in (\mathbb{D}_{\mathrm{pw}\mathscr{C}\infty})^{\mathrm{u}}$, $y \in (\mathbb{D}_{\mathrm{pw}\mathscr{C}\infty})^{\mathrm{y}}$, *and* (1) *is satisfied within* $\mathbb{D}_{\mathrm{pw}\mathscr{C}\infty}$ *on* $\mathscr{I}$, *i.e.*

$$(E_\sigma \dot{x})_{\mathscr{I}} = (A_\sigma x + B_\sigma u)_{\mathscr{I}}, \quad y_{\mathscr{I}} = (C_\sigma x + D_\sigma u)_{\mathscr{I}}.$$

If $\mathscr{I} = (t_0, \infty)$ *we call* (x, u, y) *a* global *solution.*

In (Trenn, 2009) it is shown that (1) always has a distributional solution which is uniquely determined on (t_0, ∞) by $x(t_0^+)$ and the input u, provided *all matrix pairs* (E_k, A_k) *are regular*. Throughout this paper we always make this assumption.

Remark 2. *Under the regularity assumption and for a fixed input* u, *any solution of* (1) *on some interval* (s, t) *can be uniquely extended to a solution on* (s, ∞); *however, it is not always possible to extend this solution in the past as the following simple example shows:*

$$0 = x \text{ on } (t_0, t_1), \quad \text{and} \quad \dot{x} = 0 \text{ on } [t_1, \infty).$$

The only global solution is $x = 0$, *but* $x = c$ *for any constant* $c \in \mathbb{R}^n$ *is a solution on* $[t_1, \infty)$ *which for* $c \neq 0$ *cannot be extended on the whole interval* (t_0, ∞).

3.2. Properties of a matrix pair (E, A)

We now collect some important properties and definitions for a regular matrix pair (E, A), which can then of course be applied to each subsystem in (1). The properties, and the notations introduced in the process, are then used throughout the paper.

A very useful characterization of regularity is the following well-known result (see e.g. (Berger et al., 2012)).

Proposition 3 (Regularity and quasi-Weierstraß form). *A matrix pair* $(E, A) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ *is regular if, and only if, there exist invertible matrices* $S, T \in \mathbb{R}^{n \times n}$ *such that*

$$(SET, SAT) = \left(\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \right), \tag{5}$$

where $J \in \mathbb{R}^{n_1 \times n_1}$, $0 \leq n_1 \leq n$, *is some matrix and* $N \in \mathbb{R}^{n_2 \times n_2}$, $n_2 := n - n_1$, *is a nilpotent matrix.*

The most useful aspect of describing a DAE in quasi-Weierstraß form (5) is that it decomposes the differential and algebraic part. If we partition the state x of the system $E\dot{x} = Ax$ after applying coordinate transformation as $(v^\top, w^\top)^\top = Tx$ then (5) reads as follows: $\dot{v} = Jv$ and $N\dot{w} = w$. The smooth part of the solution is obtained by solving the ODE $\dot{v} = Jv$, and the only classical solution for the equation $N\dot{w} = w$ is $w = 0$. For the latter, the distributional response to an inconsistent initial value is given by

$$w[0] = -\sum_{i=0}^{n_2-2} N^{i+1} w(0^-) \delta_0^{(i)}.$$

We thus obtain an alternate meaning to the solution of a DAE, that is, the ODE component continues to evolve smoothly, whereas the pure algebraic component jumps to the origin with the possibility of Dirac impulses at the time the system is switched on.

To describe the system in the form (5), one can calculate the matrices S, T by constructing the so called Wong-sequences from the matrices E, A, see (Berger et al., 2012) for details. Based on these transformations, we define now the following important matrices.

Definition 4. *Consider the regular matrix pair* (E, A) *with corresponding quasi-Weierstraß form* (5). *The* consistency projector *of* (E, A) *is given by*

$$\Pi_{(E,A)} = T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T^{-1}.$$

Furthermore, let

$$A^{\mathrm{diff}} := T \begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} T^{-1}, \text{ and } E^{\mathrm{imp}} := T \begin{bmatrix} 0 & 0 \\ 0 & N \end{bmatrix} T^{-1}.$$

Finally, if also an output matrix C *is considered let*

$$C^{\mathrm{diff}} := C \Pi_{(E,A)}.$$

Note that none of the above matrices depends on the specific choice of the (non-unique) quasi-Weierstraß form (5).

To give an intuitive interpretation of the objects introduced in Definition 4, it is noted that the definition of the consistency projector yields $\mathrm{im}\, \Pi_{(E,A)} = \mathfrak{C}_{(E,A)}$, and the mapping $\Pi_{(E,A)}$ defines the jump rule onto the consistency space for inconsistent initial conditions, see Lemma 5. After the jump, the state evolves within the consistency space, and the matrices A^{diff} and C^{diff} are used to describe this evolution process, see Lemma 6. The impulsive part of the solution (2) resulting from differentiation of the jump at time $t = 0$, denoted by $x[0]$, is described by E^{imp}, see Lemma 8.

Lemma 5 ((Trenn, 2009, Thm. 4.2.8)). *Consider the ITP (2) with regular matrix pair (E, A) and with arbitrary initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathscr{C}^\infty})^n$. There exists a unique solution $x \in (\mathbb{D}_{\mathrm{pw}\mathscr{C}^\infty})^n$ and*

$$x(0^+) = \Pi_{(E,A)} x(0^-).$$

Lemma 6. *For any regular matrix pair (E, A) and output matrix C, the following implication holds for all continuously differentiable (x, y):*

$$\left. \begin{array}{c} E\dot{x} = Ax, \\ y = Cx \end{array} \right\} \Rightarrow \begin{cases} \dot{x} = A^{\mathrm{diff}}x, \\ y = C^{\mathrm{diff}}x \end{cases}$$

In particular, any classical solution x of $E\dot{x} = Ax$ satisfies

$$x(t) = e^{A^{\mathrm{diff}}(t-s)}x(s), \quad s, t \in \mathbb{R}.$$

Proof. The proof is a simple consequence of (Tanwani and Trenn, 2010, Lem. 3). □

Remark 7. *One may wonder why we replace the output matrix C by C^{diff}. As a motivation, consider a simple DAE with $E = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix} \right]$, $A = I_{2 \times 2}$, and $C = [0 \ 1]$. For this system, we have $A^{\mathrm{diff}} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix} \right]$, and $C^{\mathrm{diff}} = [0 \ 0]$. Hence, with C^{diff}, it is obvious that we cannot deduce any information about the state from the continuous flow and its corresponding output y. However, when using the original output equation the corresponding unobservable space of the ODE is not the whole space, i.e., some parts of the state can be recovered from the output. Although it is true that we can recover x_2 from the output, it is misleading because in the DAE, x_2 is always zero due to the algebraic constraint. Indeed, the ability to recover x_2 does not depend on the actual output matrix C.*

Lemma 8 ((Tanwani and Trenn, 2010, Cor. 5)). *Consider the ITP (2) with regular matrix pair (E, A) and the corresponding E^{imp} matrix from Definition 4. For the unique solution $x \in (\mathbb{D}_{\mathrm{pw}\mathscr{C}^\infty})^n$ of (2), it holds that*

$$x[0] = -\sum_{i=0}^{n-2} (E^{\mathrm{imp}})^{i+1} x(0^-) \delta_0^{(i)}, \tag{6}$$

where $\delta_0^{(i)}$ denotes the i-th (distributional) derivative of the Dirac-impulse δ_0 at $t = 0$.

4. Determinability Conditions

To address the problem of state estimation, we first need to study the appropriate notion of observability for switched DAEs. Several different notions of observability can be defined in the context of switched systems as proposed in a survey chapter (Petreczky et al., 2015). For the purpose of this paper, it suffices to recall that our approach towards studying observability has three distinctive features compared to the linear time-invariant ODE systems.

Final-state observability: For nonswitched dynamical systems described by linear ODEs, observability relates to recovering the unknown initial value of the state using the input and output measurements. For switched systems, due to the presence of non-invertible state reset maps (which is the case for switched DAEs), the backward flow in time may not be uniquely defined. Because of this reason, recovering the current state does not imply recovering the initial state (the converse however always holds). For asymptotic state estimation, it then only makes sense to talk about observability of the current-state, or final-state value at the end of an interval. This concept is closely related to the notion of final-state observability defined in (Sontag, 1998, Chapter 6).

Large-time observability: Typically, for continuous LTI systems, the interval over which the inputs and outputs are measured to recover the state can be arbitrarily small. However, for switched systems, the observation over larger intervals and the change in dynamics due to switching allow us to extract additional information about the state. This motivates us to study the problem of observability without requiring observability of the individual subsystems in the classical framework, and devise a framework to take different information from multiple modes to derive relaxed conditions. The concept of large-time observability has also been found useful in the study of nonlinear ODEs (Hespanha et al., 2005).

Algebraic structure: The state of a system comprising switched DAEs is constrained by certain algebraic equations in the model description. This is taken into account by the notions of R-observability or behavioral observability for unswitched DAEs (c.f. the recent survey Berger et al. (2016)). For the model-based state estimation, which is being studied in this paper, computing consistency spaces for individual subsystems from the algebraic equations provides additional information about the state. Moreover, one may observe the Dirac impulses in the output, which also result due to different algebraic equations allowed in the system class (1), and these impulsive measurements prove useful in state estimation. This is related to the notion of impulse observability or observability at infinity (c.f. Berger et al. (2016)), but here we only extract some partial information without assuming that the subsystem are impulse observable.

In the light of the above discussion, we now introduce the notion of *determinability*, that combines the properties of large-time and final-state observability. It is formally defined as follows:

Definition 9 (Determinability). *The switched DAE (1) is called (s, t)-determinable for $t_0 \leq s < t$ if for every pair of local solutions $(x, u, y), (\bar{x}, \bar{u}, \bar{y})$ of (1) on (s, ∞) with $u \equiv \bar{u}$, the following implication holds*

$$y_{(s,t]} = \bar{y}_{(s,t]} \implies x_{(t,\infty)} = \bar{x}_{(t,\infty)}. \tag{7}$$

Hence, for a fixed switching signal σ and given input, the system (1) is considered determinable over an interval if the output measurements over that interval allow us to reconstruct the value of the state in the future. Note that the left

expression of the implication (7) is equivalent to the conditions $y_{(s,t)} = \bar{y}_{(s,t)}$ and $y[t] = \bar{y}[t]$; in particular the knowledge of the distributional output $y[t]$ will play a major role in the forthcoming observer construction. Furthermore, the right expression of the implication (7) is, due to the regularity assumption, equivalent to the condition $x(t^+) = \bar{x}(t^+)$.

Remark 10. *In Definition 9, we only consider solutions that satisfy (1) over the interval (s, ∞) and such solutions are not required to satisfy the DAE prior to time s. The reason being, to characterize $(s, t]$-determinability, we are interested in computing the set of indistinguishable states at time t^+ using only the information from the input and output measurements over the interval $(s, t]$. To obtain such a characterization it is essential to allow for $x(s^-)$ to be arbitrary; Otherwise, if $x(s^-)$ is constrained as a solution of a DAE, then that information could possibly reduce the set of states reachable at time t^+, see Remark 2. In other words, the constrained past would influence the computation of the set of indistinguishable states which is not desired for our purposes.*

Based on (Tanwani and Trenn, 2012, Proposition 9), we can restrict our attention to zero-determinability. Although there is a slight difference in the definition used in this paper, and the one given in (Tanwani and Trenn, 2012), the proof of the following result goes through with the same arguments, and is thus not repeated here.

Proposition 11 ((Tanwani and Trenn, 2012, Proposition 9)). *The switched DAE (1) is $(s, t]$-determinable if, and only if, the following implication holds for any local solution (x, u, y) of (1) on (s, ∞) with $u \equiv 0$:*

$$y_{(s,t]} \equiv 0 \quad \Longrightarrow \quad x_{(t,\infty)} = 0.$$

Thus, in studying whether the state can be completely determined from the knowledge of external measurements, one can ignore the role of inputs and set them to zero. In essence, for studying determinability conditions, we can reduce our attention to the following system class:

$$\begin{aligned} E_\sigma \dot{e} &= A_\sigma e \\ y^e &= C_\sigma e. \end{aligned} \tag{8}$$

The reason for changing the notation for state variable from x to e in the homogenous system (8) is that later, in the observer design, we will be applying our notion of determinability, and the arguments that follow, to a certain homogenous system of error dynamics. Hence, it is helpful to work with this notation at this point.

The objective now is to compute the set of indistinguishable states for (8) in the sense of Definition 9. To do so, we first study the single switch case where we show how the structural decomposition in quasi-Weierstraß form (5) and the result of Lemmas 5, 6 and 8 allow us to extract the information about the state from the output measured on one switching interval. When studying the case of multiple switches in Section 4.2, we describe how this information is then accumulated at a single time instant to arrive at a characterization of determinability.

4.1. Observable Information from Single Switch

Consider the homogeneous switched DAE (8) on (t_0, ∞) and let $t_1 > t_0$ be the first switching instant after t_0. Let the active subsystem over the interval (t_0, t_1) be denoted by (E_0, A_0, C_0) and assume that the active mode after that is (E_1, A_1, C_1). We are interested in knowing what information can be deduced about $e(t_1^+)$ using the knowledge of the output y^e of (8) measured on $(t_0, t_1]$ and the system matrices. Invoking Lemma 5 we know that $e(t_1^+) = \Pi_1 e(t_1^-)$ and it suffices therefore to focus on $e(t_1^-)$ in the following.

Consistency space: Independently of the observed output it holds that $e(t_1^-) \in \mathfrak{C}_0$, where $\mathfrak{C}_0 := \mathfrak{C}_{(E_0, A_0)}$ denotes the consistency space of the DAE corresponding to the matrix pair (E_0, A_0).

Differential unobservable space: If $y^e_{(t_0, t_1)} \equiv 0$ then $(y^e)^{(i)}(t_1^-) = 0$ for all $i \in \mathbb{N}$, hence, invoking Lemma 6, we have $e(t_1^-) \in \ker O_0^{\text{diff}}$, where $\ker O_0^{\text{diff}}$ denotes the unobservable space of the ODE $\dot{e} = A_0^{\text{diff}} e$, $y^e = C_0^{\text{diff}} e$, i.e.[1]

$$O_0^{\text{diff}} := [C_0^{\text{diff}} / C_0^{\text{diff}} A_0^{\text{diff}} / \cdots / C_0^{\text{diff}} (A_0^{\text{diff}})^{n-1}]. \tag{9}$$

Impulse unobservable space: Finally, due to Lemma 8, from the equality $0 = y^e[t_1] = C_1 e[t_1]$ it follows that $e(t_1^-) \in \ker O_1^{\text{imp}}$, where

$$O_1^{\text{imp}} := [C_1 E_1^{\text{imp}} / C_1 (E_1^{\text{imp}})^2 / \cdots / C_1 (E_1^{\text{imp}})^{n-1}].$$

Altogether this provides the motivation to introduce the *locally unobservable subspace*:

$$\mathscr{W}_1 := \mathfrak{C}_0 \cap \ker O_0^{\text{diff}} \cap \ker O_1^{\text{imp}}. \tag{10}$$

Proposition 12. *The following equality of sets holds:*

$$\mathscr{W}_1 = \left\{ e(t_1^-) \mid (e, y^e) \text{ solves (8) and } y^e_{(t_0, t_1]} = 0 \right\}.$$

The proof appears in Appendix A. As a consequence of this result, we have

$$\Pi_1 \mathscr{W}_1 = \left\{ e(t_1^+) \mid (e, y^e) \text{ solves (8) and } y^e_{(t_0, t_1]} = 0 \right\}$$

and we call $\Pi_1 \mathscr{W}_1$ the locally undeterminable space.

4.2. Gathering Information over Multiple Switches

We are now interested in combining information from different modes to characterize how much knowledge about the state can be recovered if we observe measurements over intervals involving multiple switches. The discussion in the previous section applied to the interval $(t_{p-1}, t_p]$, yields that the measurements over an interval $(t_{p-1}, t_p]$, $p \geq 1$, allow us to deduce $e(t_p^-) \in \mathscr{W}_p$, where

$$\mathscr{W}_p := \mathfrak{C}_{p-1} \cap \ker O_{p-1}^{\text{diff}} \cap \ker O_p^{\text{imp}}$$

---

[1] By $[M_1/M_2/\ldots/M_k]$, we denote the matrix which is obtained by stacking the matrices $M_1, M_2, \ldots, M_k$ (with the same number of columns) over each other.

is the locally unobservable space at t_p. The objective is to gather the information from previous locally unobservable subspaces and deduce more knowledge about the value of the state at a given time. To do so, we introduce the following sequence of subspaces:

$$\begin{aligned}
\mathscr{Q}_q^{q+1} &:= \mathscr{W}_{q+1}, \\
\mathscr{Q}_q^p &:= \mathscr{W}_p \cap e^{A_{p-1}^{\mathrm{diff}} \tau_{p-1}} \Pi_{p-1} \mathscr{Q}_q^{p-1}, \ p > q+1.
\end{aligned} \tag{11}$$

The intuition behind this definition is as follows: Our aim is to let $\mathscr{Q}_q^p$ satisfy

$$\boxed{\mathscr{Q}_q^p = \left\{ e(t_p^-) \ \middle| \ (e, y_e) = (e, 0) \text{ solves (8) on } (t_q, t_p] \right\},} \tag{12}$$

in particular, $\Pi_p \mathscr{Q}_q^p$ is the $(t_q, t_p]$-*undeterminability space* of (8) in the sense that it is the subspace of points $e(t_p^+) = \Pi_p e(t_p^-)$ which cannot be determined from the knowledge of y^e on $(t_p, t_q]$. Indeed we have the following result.

Theorem 1 (Determinability characterization). *Consider the homogeneous switched DAE* (8) *with corresponding consistency projectors* Π_p *and subspaces* $\mathscr{Q}_q^p$, $p > q$, *recursively defined by* (11). *Then* (12) *holds for any* $p > q \geq 0$. *In particular, the switched DAE* (1) *is* $(t_q, t_p]$-*determinable, if and only if,*

$$\boxed{\Pi_p \mathscr{Q}_q^p = \{0\}.} \tag{13}$$

The proof is given in Appendix A. Note that the determinability characterization (13) is significantly weaker than the condition $\mathscr{Q}_q^p = \{0\}$ used in (Tanwani and Trenn, 2013). In fact, in the latter we aimed to recover $e(t_p^-)$, while for determinability it suffices to determine $e(t_p^+) = \Pi_p e(t_p^-)$ and any uncertainty within $\ker \Pi_p$ is irrelevant for determinability in the sense of Definition 9.

4.3. An illustrative example

To illustrate the above theoretical results, consider an academic example given by (1) with a periodic switching signal σ. The first switching time is $t_0 = 0$ and, for $k \in \mathbb{N}$,

$$\tau_{4k} = 1, \ \tau_{4k+1} = \pi/4, \ \tau_{4k+2} = 1, \ \tau_{4k+3} = 1.$$

For $p = 0, 1, 2, 3$, the subsystem which is active on the interval $[t_{4k+p}, t_{4k+p+1})$ is described by $(E_{4k+p}, A_{4k+p}, B_{4k+p}, C_{4k+p})$, $k \in \mathbb{N}$, defined as,

$$p = 0 : \left(I, \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, [\,1\ 0\ 0\ 0\,] \right),$$

$$p = 1 : \left(I, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & -1 & -1 & -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, [\,0\ 0\ 0\ 0\,] \right),$$

$$p = 2 : \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, [\,0\ 0\ 0\ 1\,] \right),$$

$$p = 3 : \left(I, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, [\,0\ 1\ 0\ 0\,] \right).$$

The subsystems indexed by $p = 0, 1, 3$ are actually ODEs and the one given by $p = 2$ is already in quasi-Weierstrass form (5). Hence, $T_p = S_p = I$, for each p. The matrices appearing in Definition 4, for $p = 0, 1, 3$, are

$$\Pi_p = I, \ A_p^{\mathrm{diff}} = A, \ C_p^{\mathrm{diff}} = C, \ E_p^{\mathrm{imp}} = 0,$$

and, for $p = 2$, are

$$\Pi_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ A_2^{\mathrm{diff}} = 0, \ C_2^{\mathrm{diff}} = 0, \ E_2^{\mathrm{imp}} = E.$$

The corresponding locally unobservable spaces are

$$\mathscr{W}_1 = \mathrm{im} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathscr{W}_2 = \mathrm{im} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathscr{W}_3 = \mathrm{im} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathscr{W}_4 = \mathrm{im} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

where $\mathrm{im}\, M$ denotes the image of the linear map induced by the matrix M (i.e., its column space). The $(t_q, t_p]$-undeterminable spaces $\mathscr{Q}_q^p$ with $\Pi_p Q_q^p = 0$ are as follows

$$Q_0^2 = \mathrm{im} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad Q_2^4 = \{0\}.$$

In particular, the switched DAE is determinable on each of the intervals $(t_{4k}, t_{4k+2}]$, $(t_{4k+2}, t_{4k+4}]$, $k \in \mathbb{N}$.

5. Reconstructing the State

Based on the determinability notions studied in the previous section, we are now interested in knowing how the state of the homogenous error system (8) can actually be reconstructed from the information obtained by (nonzero) output error measurements. We approach the problem of state reconstruction in the same manner as we derived the determinability conditions in the previous section. This approach involves two major steps:

(i) Reconstruct the observable components at each switching time from the local (in time) information.

(ii) Combine the information from each subsystem to find the whole state.

Once we have derived a method to theoretically compute the state of the error dynamics exactly, we can then take additional practical elements, such as estimation errors, into account.

5.1. Observable Component at a Switching Instant

The first step in reconstructing the state of the error dynamics is to be able to write down a systematic way of extracting the observable information around each switching time from locally active subsystems. Based on the discussion in the previous section, we know that for any solution (e, y^e) of (8) with $y_{(t_0, t_1)}^e = 0$ and $y^e[t_1] = 0$, the state $e(t_1^-)$ is contained in the locally unobservable space $\mathscr{W}_1$ given by

(10). For a general solution (e, y^e) with *nonzero* y^e the corresponding state vector $e(t_1^-)$ is not contained in $\mathscr{W}_1$.

In the following, we will decompose $e(t_1^-)$ according to the direct sum $\mathbb{R}^n = \mathscr{W}_1 \oplus \mathscr{W}_1^\perp$. For that, let us introduce orthonormal matrices W_1 and Z_1 such that $\operatorname{im} W_1 = \mathscr{W}_1$ and $\operatorname{im} Z_1 = \mathscr{W}_1^\perp$. Let d_{w_1} denote the dimension of $\mathscr{W}_1$, then there exist unique elements $w_1 \in \mathbb{R}^{d_{w_1}}$, and $z_1 \in \mathbb{R}^{n-d_{w_1}}$ such that

$$e(t_1^-) = W_1 w_1 + Z_1 z_1. \tag{14}$$

Because of orthonormality of the matrix Z_1, multiplying the above equation from left by $Z_1^\top$, we get $z_1 = Z_1^\top e(t_1^-)$. Similarly, multiplication from left by $W_1^\top$ gives $w_1 = W_1^\top e(t_1^-)$. It is noted that w_1 is the unobservable component of $e(t_1^-)$ and z_1 the component of the vector $e(t_1^-)$ that can be recovered from measuring $y^e_{(t_0, t_1)}$ and $y^e[t_1]$. In the sequel, we construct a mapping that relates the vector z_1 to the measurements of $y^e_{(t_0, t_1]}$.

We first make the observation that

$$\begin{aligned}
\mathscr{W}_1^\perp &= (\mathfrak{C}_0 \cap \ker O_0^{\mathrm{diff}} \cap \ker O_1^{\mathrm{imp}})^\perp \\
&= \mathfrak{C}_0^\perp + \operatorname{im}\left(O_0^{\mathrm{diff}\top}\right) + \operatorname{im}\left(O_1^{\mathrm{imp}\top}\right),
\end{aligned}$$

that is, $\mathscr{W}_1^\perp$ is a sum of three subspaces. Since z_1 corresponds to the projection of $e(t_1^-)$ onto the subspace $\mathscr{W}_1^\perp$, we further decompose the vector z_1 along each of the three constituent subspaces of $\mathscr{W}_1^\perp$, which are subsequently denoted by z_1^{cons}, z_1^{diff}, and z_1^{imp}. More formally, we let Z_0^{cons}, Z_0^{diff}, and Z_1^{imp} be matrices with orthonomal columns such that

$$\begin{aligned}
\operatorname{im} Z_0^{\mathrm{cons}} &= \mathfrak{C}_0^\perp, & z_1^{\mathrm{cons}} &:= Z_0^{\mathrm{cons}\top} e(t_1^-), \\
\operatorname{im} Z_0^{\mathrm{diff}} &= \operatorname{im}\left(O_0^{\mathrm{diff}\top}\right), & z_1^{\mathrm{diff}} &:= Z_0^{\mathrm{diff}\top} e(t_1^-), \\
\operatorname{im} Z_1^{\mathrm{imp}} &= \operatorname{im}\left(O_1^{\mathrm{imp}\top}\right), & z_1^{\mathrm{imp}} &:= Z_1^{\mathrm{imp}\top} e(t_1^-).
\end{aligned}$$

Note that the images of the matrices Z_0^{cons}, Z_0^{diff} and Z_1^{imp} might intersect non-trivially with each other. In this case, some part of the unknown error $e(t_1^-)$ can be determined from the consistency or classically differentiable part as well as from the impulsive information. From a mathematical point of view this redundancy can be eliminated by choosing a full column rank matrix U_1 such that

$$[Z_0^{\mathrm{cons}}, Z_0^{\mathrm{diff}}, Z_1^{\mathrm{imp}}] U_1 = Z_1, \tag{15}$$

and for brevity we let $\overline{Z}_1 := [Z_0^{\mathrm{cons}}, Z_0^{\mathrm{diff}}, Z_1^{\mathrm{imp}}]$. We thus obtain,

$$z_1 = U_1^\top \overline{Z}_1^\top e(t_1^-) = U_1^\top \left(z_1^{\mathrm{cons}} / z_1^{\mathrm{diff}} / z_1^{\mathrm{imp}}\right). \tag{16}$$

We now describe, how each component of the vector z_1 can be calculated.

The consistency information z_1^{cons}: By definition, it holds that $Z_0^{\mathrm{cons}\top} \mathfrak{C}_0 = \{0\}$ and hence $z_1^{\mathrm{cons}} = Z_0^{\mathrm{cons}\top} e(t_1^-) = 0$ because any solution of the homogenous DAE (8) evolves within the consistency space $\mathfrak{C}_0$ on the interval (t_0, t_1).

Mapping for the differentiable part z_1^{diff}: In order to define $z_1^{\mathrm{diff}} \in \mathbb{R}^{r_0}$, where $r_0 = \operatorname{rank} O_0^{\mathrm{diff}}$, we first introduce the function $\mathbf{z}_0^{\mathrm{diff}} : (t_0, t_1) \to \mathbb{R}^{r_0}, t \mapsto Z_0^{\mathrm{diff}\top} e(t)$, which represents the observable component of the subsystem (E_0, A_0, C_0) that can be recovered from the smooth output measurements y^e over the interval (t_0, t_1). It follows from Lemma 17 in Appendix B that the evolution of $\mathbf{z}_0^{\mathrm{diff}}$ is governed by an observable ODE

$$\begin{aligned}
\dot{\mathbf{z}}_0^{\mathrm{diff}} &= S_0^{\mathrm{diff}} \mathbf{z}_0^{\mathrm{diff}}, \\
y^e &= R_0^{\mathrm{diff}} \mathbf{z}_0^{\mathrm{diff}},
\end{aligned} \tag{17}$$

where $S_0^{\mathrm{diff}} := Z_0^{\mathrm{diff}\top} A_0^{\mathrm{diff}} Z_0^{\mathrm{diff}}$ and $R_0^{\mathrm{diff}} := C_0^{\mathrm{diff}} Z_0^{\mathrm{diff}}$. Because of the observability of the pair $(S_0^{\mathrm{diff}}, R_0^{\mathrm{diff}})$ in (17), there exists a (linear) map $\mathscr{O}_{(t_0, t_1)}^{\mathrm{diff}}$ such that

$$\mathbf{z}_0^{\mathrm{diff}} = \mathscr{O}_{(t_0, t_1)}^{\mathrm{diff}}(y^e_{(t_0, t_1)})$$

and we set

$$z_1^{\mathrm{diff}} = \mathbf{z}_0^{\mathrm{diff}}(t_1^-).$$

From the vast literature that exists on observability of classical linear time-invariant systems, one can find various ways for representing and approximating the map $\mathscr{O}_{(t_0, t_1)}^{\mathrm{diff}}$. For our purposes, this choice is not essential. We are just interested in knowing that there exist techniques which allow us to (approximately) recover z_1^{diff} and in the design of our estimators in Section 6, it will be specified what (approximation) properties are required.

Mapping for the impulsive part z_1^{imp}: A particular characteristic of the switched DAEs is that the different algebraic constraints for different modes may lead to Dirac impulses in the solution trajectories at switching times. If such impulses are observed in the output, then this information can be used to recover a certain part of the state $e(t_1^-)$. The impulsive part of the output at switching time t_1 can be represented as

$$y^e[t_1] = \sum_{i=0}^{n-2} \eta_1^i \delta_{t_1}^{(i)},$$

where due to Lemma 8 the coefficients η_1^i satisfy the relation $\boldsymbol{\eta}_1 = -O_1^{\mathrm{imp}} e(t_1^-)$, with $\boldsymbol{\eta}_1 := (\eta_1^0 / \cdots / \eta_1^{n-2}) \in \mathbb{R}^{(n-1)y}$. We now want to find a linear map $\mathscr{O}_{[t_1]}^{\mathrm{imp}}$ such that $z_1^{\mathrm{imp}} = \mathscr{O}_{[t_1]}^{\mathrm{imp}}(y^e[t_1])$. For that, we chose a matrix U_1^{imp} such that

$$-O_1^{\mathrm{imp}\top} U_1^{\mathrm{imp}} = Z_1^{\mathrm{imp}}, \tag{18}$$

then

$$z_1^{\mathrm{imp}} = Z_1^{\mathrm{imp}\top} e(t_1^-) = -U_1^{\mathrm{imp}\top} O_1^{\mathrm{imp}} e(t_1^-) = U_1^{\mathrm{imp}\top} \boldsymbol{\eta}_1, \tag{19}$$

hence $\mathscr{O}_{[t_1]}^{\mathrm{imp}}$ is given by:

$$y^e[t_1] = \sum_{i=0}^{n-2} \eta_1^i \delta_{t_1}^{(i)} \quad \mapsto \quad U_1^{\mathrm{imp}\top}(\eta_1^0 / \dots / \eta_1^{n-2}).$$

5.2. Construct Global Mapping from Local Mappings

In the previous subsection, we described how the information over an interval with one switch can be combined with the data of active subsystems over that interval to reconstruct the observable information from that data at the switching time. This can obviously be done at each switching time t_k, $k \in \mathbb{N}$, that is, by looking at the intervals $(t_{k-1}, t_k]$ and measuring $y^e_{(t_{k-1}, t_k)}$ and $y^e[t_k]$, one can repeat the procedure in Section 5.1 to compute z_k where

$$z_k = Z_k^\top e(t_k^-) \tag{20}$$

and Z_k is an orthonormal matrix with range space $\mathscr{W}_k^\perp$. Because of our notion of determinability, the next step is to be able to combine these local informations obtained at each switching time to reconstruct the entire state of the system (8). More specifically, assuming that (8) is (t_q, t_p)-determinable for some $q, p \in \mathbb{N}$, we are interested in finding a linear map $\mathscr{O}_q^p$ such that

$$e(t_p^+) = \Pi_p \mathscr{O}_q^p(\mathbf{z}_{q+1}^p),$$

where $\mathbf{z}_{q+1}^p := (z_{q+1}/z_{q+2}/\ldots/z_p)$. We next construct this map $\mathscr{O}_q^p$. A schematic representation of the development of this section is given in Figure 2.
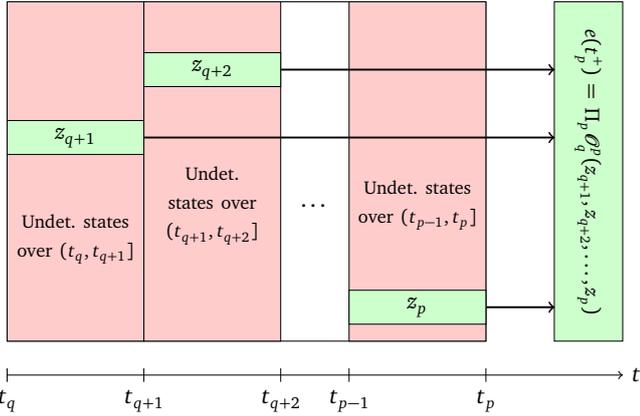

Figure 2: Accumulating local information z_k, $k = q, q+1, \ldots, p$ at time instant t_p to calculate $e(t_p^+)$.

For $k \in \mathbb{N}$ with $q < k \le p$ let P_q^k and Q_q^k be matrices with orthonormal columns such that

$$\operatorname{im} Q_q^k = \mathscr{Q}_q^k \quad \text{and} \quad \operatorname{im} P_q^k = (\mathscr{Q}_q^k)^\perp,$$

where $\mathscr{Q}_q^k$ is recursively defined by (11) and is the subspace of all points $e(t_k^-)$ which cannot be determined from $y^e_{(t_q, t_k)}$. There exist unique vectors φ_q^k, and χ_q^k such that

$$e(t_k^-) = P_q^k \varphi_q^k + Q_q^k \chi_q^k. \tag{21}$$

Because of orthonormality of the matrices P_q^k and Q_q^k, we have that $\varphi_q^k = P_q^{k\top} e(t_k^-)$ and $\chi_q^k = Q_q^{k\top} e(t_k^-)$. In particular, if (8) is (t_q, t_p)-determinable, we have $\mathscr{Q}_q^p \subseteq \ker \Pi_p$ and therefore

$$e(t_p^+) = \Pi_p P_q^p \varphi_q^p,$$

i.e. the state $e(t_p^+)$ can be recovered if we are able to find an expression for φ_q^k, $k = q+1, q+2, \ldots, p$. Note that $\varphi_q^{q+1} = z_{q+1}$ corresponds to the determinable information from the interval $(t_q, t_{q+1}]$, and we already discussed in 5.1 how it can be obtained. We now derive a recursive expression for φ_q^k, $k = q+2, q+3, \ldots, p$ in terms of φ_q^{k-1}.

For that we need to introduce a matrix Θ_q^{k-1} with orthonormal columns such that

$$\operatorname{im} \Theta_q^{k-1} = \left(e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} \mathscr{Q}_q^{k-1} \right)^\perp. \tag{22}$$

Note that then by definition

$$\operatorname{im} P_q^k = \mathscr{Q}_q^{k\perp} = \operatorname{im}[Z_k, \Theta_q^{k-1}],$$

in particular there is a matrix U_q^k such that

$$P_q^k = [Z_k, \Theta_q^{k-1}] U_q^k.$$

From

$$\begin{aligned} e(t_k^-) &= e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} e(t_{k-1}^-) \\ &= e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} \left(P_q^{k-1} \varphi_q^{k-1} + Q_q^{k-1} \chi_q^{k-1} \right), \end{aligned} \tag{23}$$

together with

$$\Theta_q^{k-1\top} e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} Q_q^{k-1} = 0.$$

we can conclude that

$$\begin{aligned} \varphi_q^k &= P_q^{k\top} e(t_k^-) = U_q^{k\top} \begin{bmatrix} Z_k^\top \\ \Theta_q^{k-1\top} \end{bmatrix} e(t_k^-) \\ &= U_q^{k\top} \begin{pmatrix} z_k \\ \Theta_q^{k-1\top} e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} P_q^{k-1} \varphi_q^{k-1} \end{pmatrix}, \end{aligned} \tag{24}$$

which is the desired recursion formula for φ_q^k for $k = q+2, q+3, \ldots, p$ with "initial value" $\varphi_q^{q+1} = z_{q+1}$.

We have thus arrived at the following result:

Theorem 2. *Consider the homogenous switched DAE (8) with corresponding A_p^{diff}, Π_p as in Definition 4, $\tau_p := t_{p+1} - t_p$, $p \in \mathbb{N}$, and assume (t_q, t_p)-determinability for some $0 \le q < p$. Furthermore, consider for $k = q+2, q+3, \ldots, p$ the matrices U_q^k, Θ_q^{k-1}, P_q^{k-1} as above and let*

$$F_q^k := U_q^{k\top} \begin{bmatrix} I \\ 0 \end{bmatrix}, \quad and \quad F_q^{q+1} := I,$$

$$G_q^k := U_q^{k\top} \begin{bmatrix} 0 \\ \Theta_q^{k-1\top} e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} P_q^{k-1} \end{bmatrix}$$

and with $\mathbf{z}_{q+1}^p = (z_{q+1}, z_{q+2}, \cdots, z_p)$ let [2]

$$\mathscr{O}_q^p(\mathbf{z}_{q+1}^p) := P_q^p \sum_{k=q+1}^{p} \left(\prod_{j=k+1}^{p} G_q^j \right) F_q^k z_k. \tag{25}$$

---

[2] For $q+1 \le k \le p-1$, the notation $\left(\prod_{j=k+1}^{p} G_q^j \right)$ denotes the product of matrices $G_q^p G_q^{p-1} \ldots G_q^{k+1}$ (note the decreasing order). By convention, when $p \le k$, the resulting product is set to identity.

Then

$$e(t_p^+) = \Pi_p \mathcal{O}_q^p(\mathbf{z}_{q+1}^p),$$

i.e., the linear map $\mathcal{O}_q^p$ describes how the state $e(t_p^+)$ can be recovered from the knowledge of locally observable parts z_{q+1}, z_{q+2}, ..., z_p, for which the construction was provided in Section 5.1.

Remark 13 (Dependence on switching signal). *While the reconstruction of the locally observable component z_k at the switching time t_k only depends on the two modes that are active prior and after the switch, i.e., $(E_{k-1}, A_{k-1}, C_{k-1})$ and (E_k, A_k, C_k), the overall reconstruction of the state, as in Theorem 2, additionally depends on the duration $\tau_k = t_{k+1} - t_k$ of each mode, because the matrix Θ_q^{k-1} used in the construction depends on $\tau_{q+1}, \ldots, \tau_{k-1}$. Hence the map $\mathcal{O}_q^p$ depends on $\tau_{q+1}, \tau_{q+2}, \ldots, \tau_p$ (but not on τ_q).*

5.3. Example

We revisit our example from Section 4.3 and recall that it is $(t_0, t_2]$ determinable, so that the above derivation can be used to recover the state $e(t_2^+)$ of the homogenous switched DAE (8) from the output y^e on $(t_0, t_2]$. For the reconstruction via (24), the following matrices are needed:

$$P_0^1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Theta_0^1 = \begin{bmatrix} 0 & 1 \\ -\sqrt{2}/2 & 0 \\ \sqrt{2}/2 & 0 \\ 0 & 0 \end{bmatrix}, \quad U_0^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\sqrt{2} & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$e^{A_1^{\text{diff}}\tau_1}\Pi_1 \approx \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 & 0 \\ -0.544 & -0.251 & -\sqrt{2}/2 & 0.456 \end{bmatrix}.$$

Furthermore, for estimating z_1^{diff} via (17), an observer for the following ODE has to be implemented:

$$\dot{\mathbf{z}}_0^{\text{diff}} = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}\mathbf{z}_0^{\text{diff}}, \quad y^e = \begin{bmatrix} 1 & 0 \end{bmatrix}\mathbf{z}_0^{\text{diff}}.$$

Finally, for estimating z_2^{imp} via (19) from the impulses in the output, we use the compression matrix $U_2^{\text{imp}} = [1\ 0\ 0]^\top$.

6. Design of State Estimators

In the previous section, we provided a method for reconstructing the state of the homogenous error dynamic system (without inputs) by constructing the mappings that exist between the output error and the observable components of the individual system. In practice, these mappings are not numerically robust, or physically realizable, and we are thus interested in obtaining estimates of the state trajectories through numerically robust algorithms. So, the purpose of this section is to design an observer for the system class (1) under the interval wise determinability assumption introduced in Section 4, which generates asymptotically convergent state estimates. The underlying structure of these estimators is based on the result of Theorem 2 developed in Section 5.

Our proposed observer for the system class (1) is given by (see also Figure 1):

$$E_\sigma \dot{\widehat{x}}_p = A_\sigma \widehat{x}_p + B_\sigma u, \text{ on } [t_p, t_{p+1}^+], \tag{26a}$$

$$\widehat{x}_p(t_p^-) = \widehat{x}_{p-1}(t_p^-) - \xi_p, \quad p \in \mathbb{N}, \tag{26b}$$

with arbitrary initial condition $\widehat{x}_0(t_0^-) \in \mathbb{R}^n$ and $[t_p, t_{p+1}^+] := [t_p, t_{p+1} + \varepsilon)$ for some arbitrarily small $\varepsilon > 0$. The desired estimate $\widehat{x}$ is defined as:

$$\widehat{x} := \sum_{p \in \mathbb{N}} (\widehat{x}_p)_{[t_p, t_{p+1})}.$$

It is seen that the observer consists of a system copy and unlike classical methods where the continuous dynamics of the estimate are driven by an error injection term, *the observer (26) updates the state estimate only at discrete switching instants by an error correction vector ξ_p*, which is determined by the difference between the system output $y = C_\sigma x$ and system copy output $\widehat{y} = C_\sigma \widehat{x}$. To give an intuitive interpretation of how to calculate ξ_p, note that, *under the assumption* that for some $p \in \mathbb{N}$, the correction term ξ_p satisfies

$$\Pi_p \xi_p = \Pi_p(\widehat{x}_{p-1}(t_p^-) - x(t_p^-)),$$

with x being a solution of (1), the equation (26b) gives $\widehat{x}_p(t_p^+) = \Pi_p \widehat{x}_p(t_p^-) = \Pi_p x(t_p^-) = x(t_p^+)$, and from there onwards the system copy (26) with $\xi_k = 0$, $k > p$ will follow exactly the original system, at least in theory. In reality, however, even after a perfect match at time t_p, the system copy will deviate again from the original system due to some uncertainties and we have to apply a correction $\xi_{\tilde{p}}$ at a later switching time $t_{\tilde{p}}$ again. It is not necessary (and may also not be possible) to apply this correction at every step. So our goal is to compute ξ_p repeatedly, for "sufficiently many" $p \in \mathbb{N}$, such that $\Pi_p \xi_p$ approximates the value of state estimation error at time t_p^+ closely as p gets large. Since the growth of error between the reset times can be upper bounded by the solution of a linear ODE, the resets (under determinability assumption) allow us to make the estimation error at switching times arbitrarily small, which eventually results in convergence of $\widehat{x}(t)$ toward $x(t)$ as t tends to infinity.

With this motivation, we introduce the state estimation error. Let $e_p := \widehat{x}_p - x$ denote the state estimation error on $[t_p, t_{p+1})$ and $e := \sum_p (e_p)_{[t_p, t_{p+1})} = \widehat{x} - x$, then

$$E_p \dot{e}_p = A_p e_p, \quad \text{on } [t_p, t_{p+1}), \tag{27a}$$

$$e_p(t_p^-) = e_{p-1}(t_p^-) - \xi_p. \tag{27b}$$

Note that equations (26a) and (27a) are both to be interpreted in the sense of distributions, in particular, the impulsive parts $x_p[t_p]$ and $e_p[t_p]$ are uniquely determined by (26) and (27), respectively. However, the error dynamics (27a) are homogenous and there are no impulses between two switches. As a result, the solution of (27a) for $t \in (t_p, t_{p+1})$ is described as

$$\begin{aligned}
e(t) = e_p(t) &= e^{A_p^{\text{diff}}(t-t_p)}\Pi_p e_p(t_p^-) \\
&= e^{A_p^{\text{diff}}(t-t_p)}\Pi_p\left(e(t_p^-) - \xi_p\right).
\end{aligned} \tag{28}$$

The output estimation error is

$$y^e = C_p \widehat{x}_p - y$$

on each open interval (t_p, t_{p+1}). The impulsive error $y^e[t_p]$ at the switching times is obtained by the difference between $y[t_p]$ and the output impulse resulting from (26a) *without* taking the correction ξ_p into account, that is,

$$y^e[t_p] := C_p \widehat{x}_{p-1}[t_p] - y[t_p]. \tag{29}$$

Note that, in general, $y^e[t_p] \neq C_p \widehat{x}[t_p] - y[t_p]$, because $\widehat{x}[t_p] = \widehat{x}_p[t_p] \neq \widehat{x}_{p-1}[t_p]$. In fact, the knowledge of $y^e[t_p]$ is based on the knowledge of $\widehat{x}_{p-1}$ which will be used to determine ξ_p, which in turn determines $C_p \widehat{x}_p[t_p]$. Furthermore, note that $y[t_p]$ as well as $\widehat{x}_{p-1}[t_p]$ depends on $u[t_p]$ and the derivates $u^{(i)}(t_p^+)$ immediately after time t_p (see (Trenn, 2012, Thm. 6.5.1)). This will render the observer slightly acausal, as the information immediately after t_p is used to set the value of $\widehat{x}_p(t_p^-)$. For DAEs, this is not surprising because the transfer functions are not necessarily proper and hence involve differentiation. However, this is not a serious problem from an implementation-point-of-view because the system copy (26) is not required to run synchronously to the original system. Another way to overcome this issue is to assume that the input is smooth at the switching instants (i.e. jumps in the inhomogeneity are induced by a switching B-matrix), then $u^{(i)}(t_p^+) = u^{(i)}(t_p^-)$.

Now that we have described the homogenous (input-free) dynamics for the state estimation error, and the corresponding output equation, we are interested in computing the vector ξ_p such that $\Pi_p \xi_p$ estimates $e_{p-1}(t_p^+) = \Pi_p e_{p-1}(t_p^-)$. We make use of the analysis carried out in Section 5 to implement the following basic idea in calculating ξ_p:

Step 1: Identify the observable component z_p of the individual subsystems for the error dynamics (27). For subsystem $p \in \mathbb{N}$, we let Z_p be an orthonormal matrix with range space $\mathscr{W}_p^\perp$, in particular, $z_p = Z_p^\top e(t_p^-)$.

Step 2: Under the assumption that for $p \in \mathbb{N}$, there exists a positive integer q such that $(t_q, t_p]$-determinability holds, we derive a linear function $\Xi_q^{p-1}(\cdot)$ such that any solution e of the error dynamics (27) satisfies

$$\Pi_p e(t_p^-) = \Pi_p \left(\mathcal{O}_q^p(\mathbf{z}_{q+1}^p) - \Xi_q^{p-1}(\xi_{q+1}^{p-1}) \right) \tag{30}$$

where

$$\xi_{q+1}^{p-1} = (\xi_{q+1}, \xi_{q+2}, \ldots, \xi_{p-1}).$$

and $\mathcal{O}_q^p$ is given by (25).

Step 3: Finally, the estimates $\widehat{z}_k$ for the observable components z_k are constructed at times t_k^-, $q + 1 \leq k \leq p$, and the error correction vector ξ_p is defined as

$$\xi_p := \mathcal{O}_q^p(\widehat{\mathbf{z}}_{q+1}^p) - \Xi_q^{p-1}(\xi_{q+1}^{p-1}), \tag{31}$$

where $\widehat{\mathbf{z}}_{q+1}^p = (\widehat{z}_{q+1}, \widehat{z}_{q+2}, \ldots, \widehat{z}_p)$.

The "error correction" block is basically a compact representation of the structure given in Figure 2, where additionally the effect of the state resets ξ_p are taken into account. In

Section 5, we have shown how to compute the map $\mathcal{O}_q^p$ using the output measurements for a homogeneous system. For computing the error correction vector ξ_p, we use the same operator but applied to the estimates of the observable components. Computations of the estimates $\widehat{z}_p$ of components of z_p, $p \in \mathbb{N}$, will be discussed in Section 6.1. Before that, we close this section by deriving the second missing component for computing ξ_p in (31), which is the map Ξ_q^{p-1} and it appears due to the state resets at previous switching instants (27b)[3]. From (30), it is clear that the computation of Ξ_q^{p-1} requires us to write $\Pi_p e(t_p^-)$ in terms of $z_{q+1}, \ldots, z_p$, and $\xi_{q+1}, \ldots, \xi_{p-1}$. Analogously to the derivation leading to Theorem 2, there exist unique vectors ψ_q^k and χ_q^k such that, cf. (21),

$$e(t_k^-) = P_q^k \psi_q^k + Q_q^k \chi_q^k, \quad k = q + 1, q + 2, \ldots, p,$$

where, as before, P_q^k and Q_q^k are orthonormal matrices whose columns span $(\mathcal{Q}_q^k)^\perp$ and $\mathcal{Q}_q^k$, respectively, and $\mathcal{Q}_q^k$ is given by (11). Invoking

$$e(t_k^-) = e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} \left(e(t_{k-1}^-) - \xi_{k-1} \right),$$

we arrive at the following recursion formula, cf. (24),

$$\psi_q^k = U_q^{k\top} \left(\Theta_q^{k-1\top} e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} \left(\substack{z_k \\ P_q^{k-1} \psi_q^{k-1} - \xi_{k-1}} \right) \right)$$

with "initial condition": $\psi_q^{q+1} = z_{q+1}$. We thus obtain the following generalization of Theorem 2:

Proposition 14. *Consider the error system* (27) *and assume* $(t_q, t_p]$-*determinability of* (1) *for some* $0 \leq q < p$. *Using the notation from Theorem 2 let*

$$H_q^k := U_q^{k\top} \begin{bmatrix} 0 \\ \Theta_q^{k-1\top} e^{A_{k-1}^{\mathrm{diff}} \tau_{k-1}} \Pi_{k-1} \end{bmatrix}$$

and

$$\Xi_q^{p-1}(\xi_{q+1}^{p-1}) := \sum_{k=q+1}^{p-1} \left(\prod_{k+2}^{p} H_q^j \right) \xi_k, \tag{32}$$

then

$$\Pi_p e(t_p^-) = \Pi_p \left(\mathcal{O}_q^p(\mathbf{z}_{q+1}^p) - \Xi_q^{p-1}(\xi_{q+1}^{p-1}) \right).$$

6.1. Observable Components and their Estimates

Before presenting the main convergence result, the last ingredient for computing state resets ξ_p for the estimator are the estimates of the observable components identified at each switching time. In Section 5.1, it was shown that these observable components z_k, $k \in \mathbb{N}$, comprise three subcomponents: z_k^{cons}, z_k^{diff}, and z_k^{imp}. Because of the algebraic constraints, we set $z_k^{\mathrm{cons}} = 0$, but for z_k^{diff} and z_k^{imp}, we compute

---

[3]This analysis can be skipped, if the state-resets are only applied at the end of a determinable interval $(t_q, t_p]$. However, in general we allow a reset of the estimator state at any switching time.

11

some appropriate estimates, denoted $\widehat{z}_k^{\mathrm{diff}}$ and $\widehat{z}_k^{\mathrm{imp}}$, respectively. Using these estimates of the individual components, we let $\widehat{z}_k$ denoted the estimate of z_k, and define it as

$$\widehat{z}_k := U_k^{\top}(z_k^{\mathrm{cons}}/\widehat{z}_k^{\mathrm{diff}}/\widehat{z}_k^{\mathrm{imp}})$$

with suitable compression matrix U_k defined analogously as in (15). In the next two subsections, we explain how the estimates $\widehat{z}_k^{\mathrm{diff}}$ and $\widehat{z}_k^{\mathrm{imp}}$ must be calculated for the convergence result proved in Section 6.2.

6.1.1. Estimate the smooth part z_k^{diff}

Based on the discussion in Section 5.1, it is possible to introduce a function $\mathbf{z}_{k-1}^{\mathrm{diff}}(\cdot) = Z_{k-1}^{\mathrm{diff}}{}^{\top} e(\cdot)$ on (t_{k-1}, t_k), $k \in \mathbb{N}$, and define an operator $\mathcal{O}_{(t_{k-1}, t_k)}^{\mathrm{diff}}$ such that $\mathbf{z}_{k-1}^{\mathrm{diff}} = \mathcal{O}_{(t_{k-1}, t_j)}^{\mathrm{diff}}(y_{(t_{pk1}, t_k)}^e)$ denotes the component of the state that can be recovered on (t_{k-1}, t_k) from the smooth part of the output. We are interested in computing the estimate of the vector $z_k^{\mathrm{diff}} = \mathbf{z}_{k-1}^{\mathrm{diff}}(t_k^-)$. The property that we require from the estimator is the following one:

(EP-1) For a given $\varepsilon_k > 0$, $k \in \mathbb{N}$, there is an estimator $\widehat{\mathcal{O}}_k^{\mathrm{diff}}$ such that

$$\widehat{z}_k^{\mathrm{diff}} = \widehat{\mathcal{O}}_k^{\mathrm{diff}}(y_{(t_{k-1}, t_k)}^e)$$

has the property that

$$|z_k^{\mathrm{diff}} - \widehat{z}_k^{\mathrm{diff}}| \leq \varepsilon_k |\mathbf{z}_{k-1}^{\mathrm{diff}}(t_{k-1}^+)|,$$

where $|\cdot|$ denotes the Euclidian norm.

In the literature, one can find many estimation techniques for linear systems of the form (17). One example of an estimator which satisfies this property is the classical *Luenberger observer*. Indeed, for $\mathbf{z}_{k-1}^{\mathrm{diff}}$ satisfying the equation (17) (defined on (t_{k-1}, t_k)), such an estimator is of the following form:

$$\dot{\widehat{\mathbf{z}}}_{k-1}^{\mathrm{diff}} = (S_{k-1}^{\mathrm{diff}} - L_{k-1}R_{k-1}^{\mathrm{diff}})\widehat{\mathbf{z}}_{k-1}^{\mathrm{diff}} + L_{k-1}y^e \text{ on } (t_{k-1}, t_k),$$

and we choose $\widehat{\mathbf{z}}_{k-1}^{\mathrm{diff}}(t_{k-1}^+) = 0$. Because $(S_{k-1}^{\mathrm{diff}}, R_{k-1}^{\mathrm{diff}})$ is an observable pair by construction, it follows from the *squashing lemma* (Pait and Morse, 1994, Lemma 1) that for a given $\varepsilon_k > 0$, and $\tau_{k-1} > 0$, there exists a matrix L_{k-1} such that $\|e^{(S_{k-1} - L_{k-1}R_{k-1})\tau_{k-1}}\| \leq \varepsilon_k$, where $\|\cdot\|$ denotes the induced matrix norm with respect to the Euclidian norm $|\cdot|$. By setting $\widehat{z}_k^{\mathrm{diff}} = \widehat{\mathbf{z}}_{k-1}^{\mathrm{diff}}(t_k^-)$, and looking at the dynamics for $\mathbf{z}_{k-1}^{\mathrm{diff}} - \widehat{\mathbf{z}}_{k-1}^{\mathrm{diff}}$, it follows that the desired estimate $|z_k^{\mathrm{diff}} - \widehat{z}_k^{\mathrm{diff}}| \leq \varepsilon_k |\mathbf{z}_k^{\mathrm{diff}}(t_{k-1}^+)|$ holds. By now, there exist many different estimation techniques for linear systems in the literature and the motivation for not fixing one particular estimation technique is to allow the possibility of several other estimators which have their own advantages.

6.1.2. Estimate the impulsive part z_k^{imp}

To estimate z_k^{imp} from the impulsive part of the output, we write the impulsive output of the error system (29) as:

$$y^e[t_k] = \sum_{i=0}^{n-2} \eta_k^i \delta_{t_k}^{(i)}.$$

Then, recalling (19),

$$z_k^{\mathrm{imp}} = U_k^{\mathrm{imp}^{\top}} \boldsymbol{\eta}_k,$$

where U_k^{imp} is a compression matrix defined analogously as in (18) and $\boldsymbol{\eta}_k := (\eta_k^0/\eta_k^1/\ldots/\eta_k^{n-2})$. Recall that

$$y^e[t_k] = C_k \widehat{x}_{k-1}[t_k] - y[t_k] =: C_k \sum_{i=0}^{n-2} \zeta_k^i \delta_{t_k}^{(i)} - \sum_{i=0}^{n-2} v_k^i \delta_{t_k}^{(i)},$$

i.e.,

$$\eta_k^i = C_k \zeta_k^i - v_k^i,$$

where v_k^i must be obtained via measuring the impulsive part of the system's output and ζ_k^i results from running the system copy (26); in particular, ζ_k^i depends on the impulsive part of the input $u[t_k]$ as well as on the derivatives $u^{(i)}(t_k^+)$, $i = 0, \ldots, n-2$, of the input immediately after the switch at t_k.

While the input may be known exactly (and hence ζ_k^i may be calculated analytically), obtaining v_k^i from measurements may prove to be a very difficult task using physical sensors, because Dirac impulses do not occur in reality. One possibility to approximately determine v_k^i is the following observation: Assume $\int$ denotes an ideal integrator, then $v_k^0 = (\int y)(t_k+) - (\int y)(t_k^-)$, and in general

$$v_k^i = \left(\underbrace{\int \int \cdots \int}_{i+1 \text{ times}} y\right)(t_k^+) - \left(\int \int \cdots \int y\right)(t_k^-).$$

If in reality, a Dirac impulse in $y[t_k]$ is "smeared out" on the interval $[t_k, t_k + \varepsilon]$ then one gets an estimate of v_k^0 by $v_k^0 \approx (\int y)(t_k + \varepsilon) - (\int y)(t_k)$, and analogously for v_k^i. The smaller ε is and the better the integrator is implemented, the closer the approximation is to the exact value v_k^i (and also η_k^i). Hence, we may approximate $y^e[t_k]$ as $\widehat{y}^e[t_k] \approx \sum_{i=0}^{n-2} \widehat{\eta}_k^i \delta_{t_k}^{(i)}$ and consequently

$$\widehat{z}_k^{\mathrm{imp}} := U_k^{\mathrm{imp}^{\top}} \widehat{\boldsymbol{\eta}}_k.$$

(EP-2) For a given $\varepsilon_k > 0$, $k \in \mathbb{N}$, one can obtain an approximation $\widehat{\boldsymbol{\eta}}_k$ of $\boldsymbol{\eta}_k$ such that

$$|\boldsymbol{\eta}_k - \widehat{\boldsymbol{\eta}}_k| \leq \varepsilon_k |\boldsymbol{\eta}_k|.$$

6.2. Convergence

We now show that the error correction vector ξ_p, $p \in \mathbb{N}$, computed from these estimates would make the state estimation error converge to zero, if the estimates of the observable components at each switching time are good enough, and a certain determinability assumption over intervals holds repeatedly. To formalize this result, let us introduce the following assumption:

(A-1) Assume that there exists a pair of subsequence in $\mathbb{N}$, non-decreasing and unbounded, denoted as $\{(q_i, p_i)\}_{i=1}^{\infty}$ with $q_i < p_i < p_{i+1}$ and such that (1) is $(t_{q_i}, t_{p_i}]$-determinable, i.e.

$$\mathcal{Q}_{q_i}^{p_i} \subseteq \ker \Pi_{p_i}, \quad i = 1, 2, 3, \cdots. \tag{33}$$

Assumption **(A-1)** basically requires that system (1) is persistently determinable, i.e., after any time instant, a determinable interval appears again eventually. At the end of these determinability intervals our observer resets the state estimate. Depending on the estimation accuracies formulated in **(EP-1)** and **(EP-2)** for individual components, the state resets make the overall estimation error sufficiently small. Afterwards the system copy runs without any continuous feedback, but its deviation from the original state is bounded by the systems dynamics. More formally, we can formulate the following qualitative convergence result:

Theorem 3. *Consider the switched system* (1) *satisfying Assumption* **(A-1)**. *For the impulsive observer* (26)*, let*

$$
\xi_p = \begin{cases} \mathscr{O}_{q_i}^{p_i}(\widehat{\mathbf{z}}_{q_i+1}^{p_i}) - \Xi_{q_i}^{p_i-1}(\xi_{q_i+1}^{p_i-1}) & \text{if } p = p_i, \ i \in \mathbb{N}, \\ 0, & \text{otherwise,} \end{cases}
$$

where the map $\mathscr{O}_{q_i}^{p_i}$ *is given by* (25)*, the estimates* $\widehat{\mathbf{z}}_{q_i+1}^{p_i} = (\widehat{z}_{q_i+1}, \widehat{z}_{q_i+2}, \ldots, \widehat{z}_{p_i})$ *are computed as in Section 6.1, the map* $\Xi_{q_i}^{p_i-1}$ *is given by* (32) *and* $\xi_{q_i+1}^{p_i-1} = (\xi_{q_i+1}, \xi_{q_i+2}, \ldots, \xi_{p_i-1})$.

For each $p \in \mathbb{N}$, *there exists* $\varepsilon_p > 0$ *such that, if* $\widehat{z}_p$ *satisfies the estimation properties* **(EP-1)** *and* **(EP-2)** *for the given* ε_p, *then it holds that*

$$
\boxed{\lim_{t \to \infty} |\widehat{x}(t^+) - x(t^+)| = 0.}
$$

The proof of Theorem 3 is constructive from design viewpoint and a quantitative bound on the ε_p is computed. Before proving this result in its generality, we highlight two special cases, whose convergence proofs form the basis for the general proof of Theorem 3.

Definition 15 (Interval-wise and sliding window observer). *Consider our general observer design as given in Theorem 3. We call this observer* interval-wise observer *if* $q_{i+1} = p_i$ *for all* $i \in \mathbb{N}$, *i.e. the determinability intervals cover the whole time axes without overlap. On the other hand, when there is the maximal possible overlap, i.e.* $p_{i+1} = p_i + 1$ *for all* $i \in \mathbb{N}$, *then we call our observer* sliding window observer.

Proof of Theorem 3. We first observe in general that, due to Proposition 14 and (31) for $p = p_i$,

$$
\begin{aligned}
e(t_{p_i}^+) &= \Pi_{p_i}(e(t_{p_i}^-) - \xi_{p_i}) \\
&= \Pi_{p_i} \left(\mathscr{O}_{q_i}^{p_i}(\mathbf{z}_{q_i+1}^{p_i}) - \Xi_{q_i}^{p_i-1}(\xi_{q_i+1}^{p_i-1}) \right. \\
&\quad \left. - \left(\mathscr{O}_{q_i}^{p_i}(\widehat{\mathbf{z}}_{q_i+1}^{p_i}) - \Xi_{q_i}^{p_i-1}(\xi_{q_i+1}^{p_i-1}) \right) \right) \\
&= \Pi_{p_i} \mathscr{O}_{q_i}^{p_i}(\mathbf{z}_{q_i+1}^{p_i} - \widehat{\mathbf{z}}_{q_i+1}^{p_i}).
\end{aligned}
$$

From Lemma 18 in Appendix B, for each $k \in \mathbb{N}$, there exists a constant $M_{k-1,k}^z > 0$ depending on $\tau_{k-1} = t_k - t_{k-1}$, $(E_{k-1}, A_{k-1}, C_{k-1})$ and (E_k, A_k, C_k) such that

$$
|z_k - \widehat{z}_k| \le \varepsilon_k M_{k-1,k}^z |e(t_{k-1}^+)|. \tag{34}
$$

Using the notation of Theorem 2, (25) and (34) yield

$$
|e(t_{p_i}^+)| \le \sum_{k=q_i+1}^{p_i} \varepsilon_k M_{k,p_i}^{\mathscr{O}} M_{k-1,k}^z |e(t_{k-1}^+)|, \tag{35}
$$

where

$$
M_{k,p_i}^{\mathscr{O}} := \left\| \Pi_{p_i} P_{q_i}^{p_i} \left(\prod_{j=k+1}^{p_i} G_{q_i}^j \right) F_{q_i}^k \right\|.
$$

Interval-wise observer. For this case, we first utilize the fact that $\xi_k = 0$ for $k = q_i + 1, q_i + 2, \ldots, p_i - 1$, and by invoking Lemma 6, we have for these k:

$$
e(t_k^+) = \left(\prod_{j=q_i}^{k-1} \Pi_{j+1} e^{A_j^{\text{diff}} \tau_j} \right) e(t_{q_i}^+).
$$

Substitution in (35) gives

$$
|e(t_{p_i}^+)| \le c_i |e(t_{q_i}^+)| = c_i |e(t_{p_{i-1}}^+)|.
$$

The constant c_i is defined as:

$$
c_i := \sum_{k=q_i+1}^{p_i} \varepsilon_k M_{k,p_i}^{\mathscr{O}} M_{k-1,k}^z M_{q_i,k-2}^{\text{diff}},
$$

where, for $k = q_i + 2, \ldots, p_i$, we let

$$
M_{q_i,k-2}^{\text{diff}} := \left\| \prod_{j=q_i}^{k-2} \Pi_{j+1} e^{A_j^{\text{diff}} \tau_j} \right\| \tag{36}
$$

and by convention, $M_{q_i,q_i-1}^{\text{diff}} = 1$. On each determinability interval $(p_{i-1}, p_i]$ we therefore can chose $\varepsilon_{p_{i-1}+1}, \varepsilon_{p_{i-1}+2}, \ldots \varepsilon_{p_i}$ sufficiently small such that $c_i \in (0, 1)$ and c_i is uniformly bounded away from 1. We can thus conclude that in this case

$$
e(t_{p_i}^+) \xrightarrow[i \to \infty]{} 0. \tag{37}
$$

Note that the ability to let i tend to infinity (and also $t_{p_i} \to \infty$) follows from assumption **(A-1)**.

Next, we have for $t \in (t_k, t_{k+1}) \subseteq (t_{p_i}, t_{p_{i+1}})$:

$$
Iche(t^+) = e^{A_k^{\text{diff}}(t-t_k)} \left(\prod_{j=p_i}^{k-1} \Pi_{j+1} e^{A_j^{\text{diff}} \tau_j} \right) e(t_{p_i}^+),
$$

which gives $|e(t^+)| \le \overline{M}_{p_i,p_{i+1}}^{\text{diff}} |e(t_{p_i}^+)|$, where

$$
\overline{M}_{p_i,p_{i+1}}^{\text{diff}} := \sup_{t \in (t_{p_i}, t_{p_{i+1}})} \left\| e^{A_{k(t)}^{\text{diff}}(t-t_{k(t)})} \left(\prod_{j=p_i}^{k(t)-1} \Pi_{j+1} e^{A_j^{\text{diff}} \tau_j} \right) \right\|,
$$

and $k(t) \in \{p_i, p_i + 1, \ldots, p_{i+1} - 1\}$ is such that $t \in (t_{k(t)}, t_{k(t)+1})$. We may now chose c_i small enough (by choosing $\varepsilon_{p_{i-1}+1}, \ldots, \varepsilon_{p_i}$ to be sufficiently small) such that $c_i \overline{M}_{p_i,p_{i+1}}^{\text{diff}}$ is uniformly bounded, say by $\overline{c} > 0$, then for all $t \in (t_{p_i}, t_{p_{i+1}})$ we have

$$
|e(t^+)| \le \overline{M}_{p_i,p_{i+1}}^{\text{diff}} c_i |e(t_{p_{i-1}}^+)| \le \overline{c} |e(t_{p_{i-1}}^+)|,
$$

and the convergence of $|e(t^+)|$ towards zero as $t \to \infty$ now follows from (37).

Sliding window observer. By Assumption **(A-1)**, the sequence $(q_i)_{i\in\mathbb{N}}$ is nondecreasing and unbounded, i.e., for sufficiently large i we have $q_i \geq p_1$. By using the relation $p_{i+1} = p_i + 1$, we can rewrite (35) as

$$|e(t_{p_i}^+)| \leq \sum_{k=q_i+1}^{p_i} \varepsilon_k M_{k,p_i}^{\mathscr{O}} M_{k-1,k}^z |e(t_{p_i-p_i+k-1}^+)|$$

for sufficiently large i. Let

$$c_i := (p_i - q_i) \max_{k=q_i+1,\ldots,p_i} \varepsilon_k M_{k,p_i}^{\mathscr{O}} M_{k-1,k}^z.$$

Since $(q_i)_{i\in\mathbb{M}}$ is non-decreasing and unbounded, for any fixed k, there are only finitely many indices i such that c_i depends on ε_k. Hence for sufficiently small ε_k we have $c_i \in (0,1)$ uniformly bounded away from 1 and

$$|e(t_{p_i}^+)| \leq \frac{c_i}{p_i - q_i} \sum_{j=1}^{p_i-q_i} |e(t_{p_{i-j}}^+)|.$$

Applying Lemma 19 from the Appendix B results in

$$|e(t_{p_i}^+)| \xrightarrow[i\to\infty]{} 0. \tag{38}$$

Next, for any $t \in (t_{p_i}, t_{p_i+1})$, we have $e(t^+) = e^{A_{p_i}^{\text{diff}}(t-t_{p_i})} e(t_{p_i}^+)$ and hence

$$|e(t^+)| \leq M_{p_i}^{\text{diff}} |e(t_{p_i}^+)| \leq M_{p_i}^{\text{diff}} c_i \frac{1}{p_i - q_i} \sum_{j=1}^{p_i-q_i} |e(t_{p_{i-j}}^+)|,$$

where $M_{p_i}^{\text{diff}} = \sup_{s\in(0,\tau_{p_i})} \left\| e^{A_{p_i}^{\text{diff}}s} \right\|$. For sufficiently small ε_k, we can ensure that $M_{p_i}^{\text{diff}} c_i$ is uniformly bounded, say by $\bar{c} > 0$. Furthermore, for any $\varepsilon > 0$ there exists an index i_ε such that $|e(t_p^+)| \leq \varepsilon$ for all $p \geq q_{i_\varepsilon}$, hence for all $t \in (t_{p_i}, t_{p_i+1})$ and all $i \geq i_\varepsilon$ we have

$$|e(t^+)| \leq \bar{c} \frac{1}{p_i - q_i} \sum_{j=1}^{p_i-q_i} |e(t_{p_{i-j}}^+)| \leq \bar{c}\varepsilon,$$

This shows convergence of $e(t^+)$ towards 0 for $t \to \infty$.

The general case. We now combine the proof ideas from the interval-wise and sliding window observer to also prove the general case. To this end, for a fixed $i \in \mathbb{N}$, introduce the function $h(i,\cdot) : \{q_i, \ldots, p_i - 1\} \to \mathbb{N}$ such that[4]

$$h(i,k) = \max\{j \,|\, p_j \leq k\}$$

and let the set J_i be the range of $h(i,\cdot)$, that is, $J_i := \{h(i,k), k = q_i, \ldots, p_i - 1\}$. For each $k = q_i, \ldots, p_i - 1$, it holds that, recalling (36),

$$|e(t_k^+)| \leq M_{p_{h(i,k)},k-1}^{\text{diff}} |e(t_{p_{h(i,k)}}^+)|$$

---

[4]For the interval-wise observer, $h(i,k) = p_{i-1} = q_i$. For sliding-window observer, where $p_{i-1} = p_i - 1$, we had $h(i,k) = k = p_i - p_i + k = p_{i-p_i+k}$.

and (35) becomes

$$|e(t_{p_i}^+)| \leq \sum_{k=q_i+1}^{p_i} \varepsilon_k M_{k,p_i}^{\mathscr{O}} M_{k-1,k}^z M_{p_{h(i,k-1)},k-2}^{\text{diff}} |e(t_{p_{h(i,k-1)}}^+)|.$$

Let $|J_i|$ denote the cardinality of J_i, and let

$$c_i := |J_i| \max_{k=q_i+1,\ldots,p_i} \varepsilon_k M_{k,p_i}^{\mathscr{O}} M_{k-1,k}^z M_{p_{h(i,k-1)},k-2},$$

then by choosing ε_k sufficiently small, we again have $c_i \in (0,1)$ uniformly bounded away from 1, and

$$|e(t_{p_i}^+)| \leq \frac{c_i}{|J_i|} \sum_{j=1}^{|J_i|} |e(t_{p_{i-j}}^+)|.$$

Once again, it follows from Lemma 19 in Appendix B that $|e(t_{p_i}^+)| \to 0$ as $i \to \infty$. To show that, $|e(t^+)|$ converges to zero for $t \in (t_{p_i}, t_{p_{i+1}})$, one can follow exactly the same arguments as in the case of interval-wise observer to concluded that $e(t^+) \to 0$ as $t \to \infty$. $\qquad\square$

Remark 16 (Convergence of impulsive part $\hat{x}[t_p] - x[t_p]$). *It was already observed that $e[t_p] = -\sum_{i=0}^{n-2}(E_p^{\text{imp}})^{i+1}(e(t_p^-) - \xi_p)\delta_{t_p}^{(i)}$. From Theorem 3, we have $e(t_p^-)$ converging to zero for large $p \in \mathbb{N}$. Since ξ_p is by construction an estimate of $e(t_p^-)$ which gets closer and closer to the real value of $e(t_p^-)$ as p gets large, it follows that the coefficients multiplying Dirac impulses get smaller with time.*

7. Simulations

For the simulation of a sliding window observer, we refer to our conference paper (Tanwani and Trenn, 2013). The simulations for the interval-wise observer are now presented for the example considered in Section 4.3. As a known input to the system, we chose $u(t) = 1 + \sin(t)$. The corresponding output of the switched system (1) with initial condition $x(t_0) = [1, 3/2, 2, 5/2]^\top$ is shown in Figure 3.
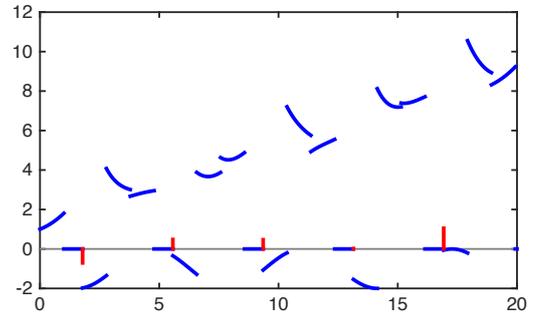

Figure 3: Output $y = y_{\mathbb{D}}^f + y[\cdot]$ of (1) with regular part y^f (blue) and indication of Dirac impulses in $y[\cdot]$ (red).

For our observer design, we estimate z_p^{diff} using a classical Luenberger observer with $L_0 = [1/4, 3/8]^\top$, $L_3 = [3/8, 1/4]^\top$, and we add ten percent measurement noise in
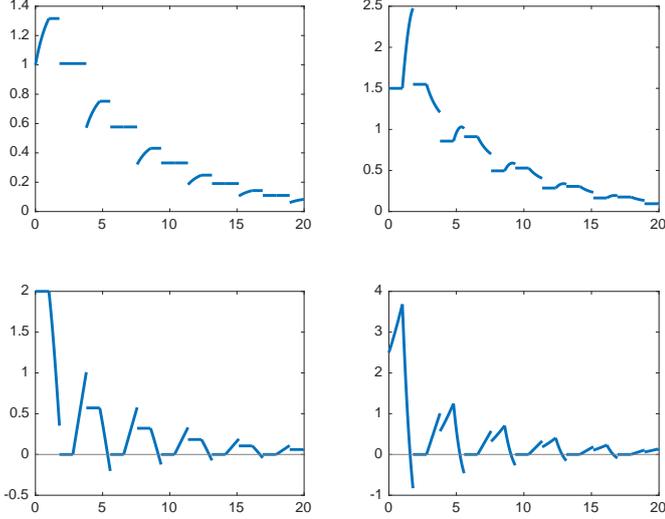
Figure 4: Estimation errors (without Dirac impulses), $x_1 - \widehat{x}_1$ upper left figure, $x_2 - \widehat{x}_2$ upper right figure, $x_3 - \widehat{x}_3$ lower left figure, $x_4 - \widehat{x}_4$ lower right figure.

the impulsive part of $y[\cdot]$. The system copy starts with zero initial value. The resulting estimation errors (without the Dirac impulses) are shown in Figure 4. Clearly, the estimation error converges (slowly) to zero. Note that the convergence can be accelerated significantly by using a more aggressive gain matrix in Luenberger observers. However, there exists a lower bound on the gain matrices which is determined by the length of the determinability intervals.

8. Conclusions

The paper considered the problem of state estimation in switched linear DAEs. The notion of determinability studied in this paper relates to the reconstruction of the state value at some time by processing outputs and inputs over an interval. This does not necessarily require observability of the initial state, or the individual subsystems. The geometric characterization of determinability is then used for synthesis of a class of state estimators. In contrast to classical estimation techniques which require continuous output injection, in our approach the estimator is reset at some discrete time instants after processing the external measurements over an interval.

For future work, we are interested in developing state estimators which only require the property of detectability from system. Our preliminary results on detectability of switched DAEs have appeared in (Tanwani and Trenn, 2015). Similar concepts have been used for studying the notion of controllability in switched DAEs (Küsters et al., 2015) and a duality result is also available (Küsters and Trenn, 2015). It would be interesting to investigate if such ideas can be used for designing stabilizing controllers.

Appendix A. Detailed Proofs

Proof of Proposition 12. The above discussion already shows "$\supseteq$". To show the converse subspace inclusion, let $e_1 \in \mathscr{W}_1$.

We need to show that there exists a solution e of (8) such that $e(t_1^-) = e_1$ and the resulting output is such that $y_{(t_0, t_1]}^e \equiv 0$. For that let $e_0 := e^{-A_0^{\mathrm{diff}}(t_1 - t_0)} e_1$. Since $e_1 \in \mathfrak{C}_0$ and $\mathfrak{C}_0$ is A_0^{diff}-invariant, it follows that $e_0 \in \mathfrak{C}_0$. Lemma 6 now yields that (8) with initial condition $e(t_0^+) = e_0$ has a solution given by $e(t) = e^{A_0^{\mathrm{diff}}(t - t_0)} e_0$ on (t_0, t_1), and in particular, $e(t_1^-) = e_1$. Furthermore, $e_1 \in \ker O_0^{\mathrm{diff}}$ and A_0^{diff}-invariance of $\ker O_0^{\mathrm{diff}}$ implies that $e(t) \in \ker O_0^{\mathrm{diff}}$ for $t \in (t_0, t_1)$, i.e., $y_{(t_0, t_1)}^e = 0$. Finally, $e(t_1^-) = e_1 \in \ker O_0^{\mathrm{imp}}$, hence $y^e[t_1] = 0$ by Lemma 8. $\qquad \square$

Proof of Theorem 1. We first prove (12).

($\supseteq$). To prove this inclusion, we consider the solution e that satisfies (8) on $(t_q, t_p]$ and the corresponding output y^e is such that $y_{(t_q, t_p]}^e = 0$. From Proposition 12, we have $e(t_{q+1}^-) \in \mathscr{W}_{q+1} = \mathscr{Q}_q^{q+1}$. Using an induction based argument, it is shown that $e(t_k^-) \in \mathscr{Q}_q^k$ for all $k = q+1, q+2, \ldots, p$. Assume that the desired relation holds for some $k \geq q+1$. As $e(t_{k+1}^-) = e^{A_k^{\mathrm{diff}} \tau_k} \Pi_k e(t_k^-)$, it follows that $e(t_{k+1}^-) \in e^{A_k^{\mathrm{diff}} \tau_k} \Pi_k \mathscr{Q}_q^k$. Also, the application of Proposition 12 on the interval $(t_k, t_{k+1}]$ yields that $e(t_{k+1}^-) \in \mathscr{W}_{k+1}$. Hence we have shown that $e(t_{k+1}^-) \in \mathscr{W}_{k+1} \cap e^{A_k^{\mathrm{diff}} \tau_k} \Pi_k \mathscr{Q}_q^k = \mathscr{Q}_q^{k+1}$.

($\subseteq$). For each $e_p \in \mathscr{Q}_q^p$, we want to construct a solution e of (8) over the interval $(t_q, t_p]$ such that $e(t_p^-) = e_p$ and $e(t_k^-) \in \mathscr{W}_k$, for $k = q+1, q+2, \ldots, p$. From Proposition 12, it would then follow that $y_{(t_{k-1}, t_k]}^e = 0$ and the claim is shown. To construct the desired solution, let $e_p \in \mathscr{Q}_q^p$ and chose $e_k \in \mathscr{Q}_q^k$ for $k = p-1, p-2, \ldots, q+1$ so that

$$e^{A_k^{\mathrm{diff}} \tau_k} \Pi_k e_k = e_{k+1},$$

which is possible due to the definition of $\mathscr{Q}_q^{k+1}$. Finally let $e_q^+ := e^{-A_q^{\mathrm{diff}} \tau_q} e_{q+1}$. By construction, $e_{q+1} \in \mathfrak{C}_q$ and hence $e_q^+ \in \mathfrak{C}_q$ as well. Hence there exists a local[5] solution e of (8) on $(t_q, t_p]$ with $e(t_q^+) = e_q^+$. Furthermore, Lemmas 5 and 6 yield $e(t_{k+1}^-) = e^{A_k^{\mathrm{diff}} \tau_k} \Pi_k e(t_k^-)$ for all $k = q+2, q+3, \ldots, p-1$ as well as $e(t_{q+1}^-) = e^{A_q^{\mathrm{diff}} \tau_q} e(t_q^+) = e_{q+1}$. Inductively, we can now conclude that $e(t_k^-) = e_k \in \mathscr{Q}_q^k \subseteq \mathscr{W}_k$ for all $k = q+1, q+2, \ldots p$ which concludes this proof step.

Finally, due to Lemma 5, the condition (13) is a characterization for determinability of (8) and Proposition 11 yields the same determinability characterization for the inhomogeneous switched DAE (1). $\qquad \square$

Appendix B. Lemmas Used in Derivations

Lemma 17. *Consider the ODE*

$$\dot{x} = Ax, \quad y = Cx$$

---

[5]Note that there may not exist a global solution with the initial condition $e(t_q^+) = e_q^+$, as already shown by example in Remark 2.

for some $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{y \times n}$, and chose a matrix Z with orthonormal columns such that

$$\text{im}\, Z = \text{im}[C/CA/CA^2/\dots/CA^{n-1}]^\top.$$

Then for any solution x of $\dot{x} = Ax$ we have that $z = Z^\top x$ is a solution of the observable ODE

$$\dot{z} = Z^\top AZz, \quad y = CZz.$$

Proof. This is a simple consequence from the well known Kalman observability decomposition. □

Lemma 18. *Consider the switched DAE* (1) *together with the impulsive observer* (26) *and corresponding error dynamics* (27). *For $k \in \mathbb{N}$ let z_k and $\widehat{z}_k$ be given as in Section 5.1 and Section 6.1. If for some $\varepsilon_k > 0$, the estimation properties* **(EP-1)** *and* **(EP-2)** *hold, then there exists a constant $M^z_{k-1,k} > 0$ depending on $\tau_{k-1} = t_k - t_{k-1}$, $(E_{k-1}, A_{k-1}, C_{k-1})$ and (E_k, A_k, C_k) such that*

$$|z_k - \widehat{z}_k| \le \varepsilon_k M^z_{k-1,k} |e(t^+_{k-1})|. \tag{B.1}$$

Proof. Since $z_k - \widehat{z}_k = U_k^\top (0/z_k^{\text{diff}} - \widehat{z}_k^{\text{diff}}/z_k^{\text{imp}} - \widehat{z}_k^{\text{imp}})$, it suffices to consider the differences $z_k^{\text{diff}} - \widehat{z}_k^{\text{diff}}$ and $z_k^{\text{imp}} - \widehat{z}_k^{\text{imp}}$ individually.

Bound for the difference $|z_k^{\text{diff}} - \widehat{z}_k^{\text{diff}}|$. Invoking **(EP-1)** and the definition of $\mathbf{z}^{\text{diff}}_{k-1}$ implies

$$\left| z_k^{\text{diff}} - \widehat{z}_k^{\text{diff}} \right| \le \varepsilon_k \left| \mathbf{z}^{\text{diff}}(t^+_{k-1}) \right| = \varepsilon_k \left| Z^{\text{diff}}_{k-1}{}^\top e(t^+_{k-1}) \right|$$
$$\le \varepsilon_k \left\| Z^{\text{diff}}_{k-1}{}^\top \right\| \left| e(t^+_{k-1}) \right|,$$

which is the desired bound.

Bound for the difference $|z_k^{\text{imp}} - \widehat{z}_k^{\text{imp}}|$. We first observe that

$$y^e[t_k] = C_k(\widehat{x}_{k-1}[t_k] - x[t_k]) = C_k(\widehat{x}_{k-1}[t_k] - \widehat{x}_k[t_k] - e[t_k]).$$

From Lemma 8, it is easily seen that,

$$\widehat{x}_{k-1}[t_k] - \widehat{x}_k[t_k] = -\sum_{i=0}^{n-2} (E_k^{\text{imp}})^{i+1} \xi_k \delta^{(i)}$$

and

$$e[t_k] = -\sum_{i=0}^{n-2} (E_k^{\text{imp}})^{i+1} (e(t^-_k) - \xi_k) \delta^{(i)}_{t_k}.$$

Hence, $y^e[t_k] = \sum_{i=0}^{n-2} \eta_k^i \delta^{(i)}_{t_k}$ with

$$\eta_k^i = -C_k(E_k^{\text{imp}})^{i+1} e(t^-_k) = -C_k(E_k^{\text{imp}})^{i+1} e^{A^{\text{diff}}_{k-1}\tau_{k-1}} e(t^+_{k-1}),$$

where we invoked Lemma 6. Combining this with **(EP-2)**, we obtain the desired bound:

$$\left| z_k^{\text{imp}} - \widehat{z}_k^{\text{imp}} \right| = \left| U_k^{\text{imp}\top}(\boldsymbol{\eta}_k - \widehat{\boldsymbol{\eta}}_k) \right| \le \varepsilon_k \left\| U_k^{\text{imp}\top} \right\| |\boldsymbol{\eta}_k|$$
$$\le \varepsilon_k M^{\text{imp}}_{k-1,k} \left| e(t^+_{k-1}) \right|,$$

where $M^{\text{imp}}_{k-1,k} := \left\| U_k^{\text{imp}\top} \right\| \left\| O_k^{\text{imp}} e^{A^{\text{diff}}_{k-1}\tau_{k-1}} \right\|$. Altogether, we have (B.1) with

$$M^z_{k-1,k} = \left\| U_k^\top \right\| \left| \begin{pmatrix} \|Z^{\text{diff}}_{k-1}\| \\ M^{\text{imp}}_{k-1,k} \end{pmatrix} \right|.$$

□

Lemma 19. *Consider a sequence $(a_k)_{k \in \mathbb{N}}$ in $\mathbb{R}^m$, $m \in \mathbb{N}$, where*

$$|a_k| \le \frac{\alpha}{n_k} \sum_{j=1}^{n_k} |a_{k-j}| \tag{B.2}$$

for some $\alpha \in (0,1)$, and the positive integers $n_k \le k$ are such that the sequence $(k - n_k)_{k \in \mathbb{N}}$ is nondecreasing and unbounded. It then holds that $\lim_{k \to \infty} |a_k| = 0$.

Proof. We first extract a subsequence $(a_{k_i})_{i \in \mathbb{N}}$ such that $k_i \le k_{i+1} - n_{k_{i+1}}$ which is possible due to assumptions on $(k - n_k)_{k \in \mathbb{N}}$. We will show that

$$\max_{k_i \le j < k_{i+1}} |a_j| \le \alpha \max_{k_{i-1} \le j < k_i} |a_j| \tag{B.3}$$

which implies that $|a_k|$ converges to zero as $k \to \infty$.

From (B.2), it follows that $|a_{k_i}| \le \alpha \max_{k_i - n_{k_i} \le j < k_i} |a_j|$, and knowing that $k_{i-1} \le k_i - n_{k_i}$, we have

$$|a_{k_i}| \le \alpha \max_{k_{i-1} \le j < k_i} |a_j|.$$

We will now show inductively that for all $\ell \in \mathbb{N}$

$$|a_{k_i + \ell}| \le \alpha \max_{k_{i-1} \le j < k_i} |a_j|,$$

from which (B.3) follows.

For this, it is first observed that, for all $\ell > 0$,

$$k_i + \ell - n_{k_i + \ell} \ge k_i - n_{k_i} \ge k_{i-1},$$

where the first inequality is due to the nondecreasing assumption on $(k - n_k)_{k \in \mathbb{N}}$, and the second inequality results from the definition of $(a_{k_i})_{i \in \mathbb{N}}$. To invoke the induction argument, we assume

$$|a_j| \le \max_{k_{i-1} \le j < k_i} |a_j| \quad \text{for } k_{i-1} \le j \le k_i + \ell - 1$$

and arrive at

$$|a_{k_i + \ell}| \le \alpha \max_{k_i + \ell - n_{k_i + \ell} \le j \le k_i + \ell - 1} |\alpha_j|$$
$$\le \alpha \max_{k_{i-1} \le j \le k_i + \ell - 1} |\alpha_j|$$
$$\le \alpha \max_{k_{i-1} \le j \le k_i} |\alpha_j|.$$

□

References

Babaali, M., Pappas, G. J., 2005. Observability of switched linear systems in continuous time. In: Hybrid Systems: Computation and Control. Vol. 3414 of LNCS. Springer, Berlin, pp. 103–117.

Balluchi, A., Benvenuti, L., Di Benedetto, M. D., Sangiovanni-Vincentelli, A., 2003. Observability for hybrid systems. In: Proc. 42nd IEEE Conf. Decis. Control, Hawaii, USA. Vol. 2. pp. 1159–1164.

Berger, T., Ilchmann, A., Trenn, S., 2012. The quasi-Weierstraß form for regular matrix pencils. Linear Algebra Appl. 436 (10), 4052–4069.

Berger, T., Reis, T., 2015. Observers and dynamic controllers for linear differential-algebraic systems, submitted for publication.

Berger, T., Reis, T., Trenn, S., 2016. Observability of linear differential-algebraic systems. In: Ilchmann, A., Reis, T. (Eds.), Surveys in Differential-Algebraic Equations IV. Differential-Algebraic Equations Forum. Springer-Verlag, Berlin-Heidelberg, to appear.

Bobinyec, K., Campbell, S. L., Kunkel, P, 2011. Full order observers for linear DAEs. In: Proc. 50th IEEE Conf. Decis. Control and European Control Conference ECC 2011, Orlando, USA. pp. 4011 – 4016.

Bobinyec, K. S., Campbell, S. L., 2014. Linear differential algebraic equations and observers. In: Ilchmann, A., Reis, T. (Eds.), Surveys in Differential-Algebraic Equations II. Differential-Algebraic Equations Forum. Springer-Verlag, Berlin-Heidelberg, pp. 1–67.

Dai, L., 1989. Singular Control Systems. No. 118 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin.

Darouach, M., 2012. On the functional observers for linear descriptor systems. Syst. Control Lett. 61, 427 – 434.

Fahmy, M. M., O'Reilly, J., 1989. Observers for descriptor systems. Int. J. Control 49, 2013–2028.

Goebel, R., Sanfelice, R. G., Teel, A. R., 2009. Hybrid dynamical systems. IEEE Control Systems Magazine 29 (2), 28–93.

Gross, T. B., Trenn, S., Wirsen, A., 2014. Topological solvability and index characterizations for a common DAE power system model. In: Proc. 2014 IEEE Conf. Control Applications (CCA). IEEE, pp. 9–14.

Hespanha, J., Liberzon, D., Angeli, D., Sontag, E., 2005. Nonlinear norm-observability notions and stability of switched systems. ieeetac 50 (2), 154–168.

Küsters, F., Ruppert, M. G.-M., Trenn, S., 2015. Controllability of switched differential-algebraic equations. Syst. Control Lett. 78 (0), 32 – 39.

Küsters, F., Trenn, S., 2015. Duality of switched DAEs, submitted for publication.

Pait, F., Morse, S., 1994. A cyclic switching strategy for parameter-adaptive control. IEEE Trans. Autom. Control 39 (6), 1172 – 1183.

Petreczky, M., Tanwani, A., Trenn, S., 2015. Observability of switched linear systems. In: Djemai, M., Defoort, M. (Eds.), Hybrid Dynamical Systems. Vol. 457 of Lecture Notes in Control and Information Sciences. Springer-Verlag, pp. 205–240.

Shim, H., Tanwani, A., 2014. Hybrid-type observer design based on a sufficient condition for observability in switched nonlinear systems. Int. J. Robust & Nonlinear Control: Special Issue on High Gain Observers and Nonlinear Output Feedback Control 24 (6), 1064 – 1089.

Sontag, E. D., 1998. Mathematical Control Theory: Deterministic Finite Dimensional Systems, 2nd Edition. Springer-Verlag, New York.

Sun, Z., Ge, S. S., Lee, T. H., 2002. Controllability and reachability criteria for switched linear systems. Automatica 38, 775–786.

Tanwani, A., 2011. Invertibility and observability of switched systems with inputs and outputs. Ph.D. thesis, University of Illinois at Urbana-Champaign, IL, USA, available online: http://hdl.handle.net/2142/29755.

Tanwani, A., Brogliato, B., Prieur, C., 2014. Stability and observer design for multivalued Lur'e systems with non-monotone, time-varying nonlinearities and state jumps. SIAM J. Control & Optim. 52 (6), 3639–3672.

Tanwani, A., Brogliato, B., Prieur, C., 2016. Observer design for unilaterally constrained Lagrangian systems: A passivity-based approach. IEEE Trans. Autom. Control 61 (9), 2386–2401.

Tanwani, A., Shim, H., Liberzon, D., 2013. Observability for switched linear systems: Characterization and observer design. IEEE Trans. Autom. Control 58 (4), 891–904.

Tanwani, A., Shim, H., Liberzon, D., 2015. Hybrid Dynamical Systems. Vol. 457 of Lecture Notes in Control and Information Sciences. Springer, Ch. Observer Design for Switched Linear Systems with State Jumps, pp. 179 – 203.

Tanwani, A., Trenn, S., 2010. On observability of switched differential-algebraic equations. In: Proc. 49th IEEE Conf. Decis. Control, Atlanta, USA. pp. 5656–5661.

Tanwani, A., Trenn, S., 2012. Observability of switched differential-algebraic equations for general switching signals. In: Proc. 51st IEEE Conf. Decis. Control, Maui, USA. pp. 2648–2653.

Tanwani, A., Trenn, S., 2013. An observer for switched differential-algebraic equations based on geometric characterization of observability. In: Proc. 52nd IEEE Conf. Decis. Control, Florence, Italy. pp. 5981–5986.

Tanwani, A., Trenn, S., 2015. On detectability of switched linear differential-algebraic equations. In: Proc. 54th IEEE Conf. Decis. Control, Osaka, Japan. pp. 2957–2962.

Trenn, S., 2009. Distributional differential algebraic equations. Ph.D. thesis, Institut für Mathematik, Technische Universität Ilmenau, Universitätsverlag Ilmenau, Germany. URL http://www.db-thueringen.de/servlets/DocumentServlet?id=13581

Trenn, S., 2012. Switched differential algebraic equations. In: Vasca, F., Iannelli, L. (Eds.), Dynamics and Control of Switched Electronic Systems - Advanced Perspectives for Modeling, Simulation and Control of Power Converters. Springer-Verlag, London, Ch. 6, pp. 189–216.

Trenn, S., 2013. Solution concepts for linear DAEs: a survey. In: Ilchmann, A., Reis, T. (Eds.), Surveys in Differential-Algebraic Equations I. Differential-Algebraic Equations Forum. Springer-Verlag, Berlin-Heidelberg, pp. 137–172.

Vidal, R., Chiuso, A., Soatto, S., Sastry, S., 2003. Observability of linear hybrid systems. In: Hybrid Systems: Computation and Control. Vol. 2623 of Lecture Notes in Computer Science. Springer, Berlin, pp. 526–539.

Xie, G., Wang, L., 2003. Controllability and stabilizability of switched linear systems. Syst. Control Lett. 48 (2), 135–155.