

Detectability and Observer Design for Switched Differential Algebraic Equations

Aneel Tanwani^a, Stephan Trenn^b

^a*Team MAC (Decision and Optimization), LAAS – CNRS, Toulouse, France*

^b*Jan C. Willems Center for Systems and Control, University of Groningen, Netherlands.*

Abstract

This paper studies detectability for switched linear differential-algebraic equations (DAEs) and its application to the synthesis of observers, which generate asymptotically converging state estimates. Equating detectability to asymptotic stability of zero-output-constrained state trajectories, and building on our work on interval-wise observability, we propose the notion of interval-wise detectability: If the output of the system is constrained to be identically zero over an interval, then the norm of the corresponding state trajectories scales down by a certain factor at the end of that interval. Conditions are provided under which the interval-wise detectability leads to asymptotic stability of zero-output-constrained state trajectories. An application is demonstrated in designing state estimators. Decomposing the state into observable and unobservable components, we show that if the observable component of the system is reset appropriately and persistently, then the estimation error converges to zero asymptotically under the interval-wise detectability assumption.

Keywords: Switched systems, differential-algebraic equations, detectability, observer design, state estimation, asymptotic convergence

1. Introduction

The growing application of switched systems in modeling and analysis has contributed toward immense research in the area of dynamical systems which combine discrete and continuous dynamics. Different classes of switched systems can be introduced based on the models associated with the switching signal, or the particular characteristics of the individual subsystems. In this regard, this article studies the problem of *detectability* for switched systems where the subsystems are described by differential-algebraic equations (DAEs) and the switching signal is assumed to be known a priori.

Switched DAEs arise naturally when the system dynamics undergo sudden structural changes (switches) and the dynamics of each mode are algebraically constrained. A typical example are electrical circuits with switches where the constraints are induced by Kirchhoff's laws. Our previous works on structural properties of switched DAEs has addressed the problem of observability (Tanwani and Trenn, 2012) and the observer design (Tanwani and Trenn, 2013, 2017a) under the stronger assumption of determinability (which in the nonswitched case is equivalent to observability and roughly speaking means that the state at the end of the observation interval can be determined to any given accuracy). Building on this line of work, this article proposes the notion of detectability for switched linear DAEs and its application in the observer design.

Roughly speaking, the property of detectability combines the notions of observability and stability, that is, a dynamical system is called detectable if the state trajectories, which correspond to same input and output, converge asymptotically to the same value. Seen as a generalization of the observability property for classical linear systems, detectability is characterized by asymptotic stability of the unobservable modes for linear time-invariant systems. For nonlinear systems, while there are different notions for observability (Sontag, 1998), the notion of output-to-state stability (OSS) provides one possible framework (Sontag and Wang, 1997) which has also been used in observer design (Astolfi and Praly, 2003). These techniques are generalized for the switched systems as well. The work of (De Santis et al., 2009) proposes detectability conditions for switched linear system in terms of the stability of a reduced order switched system. The notion of OSS has been studied for switched nonlinear systems in (Muller and Liberzon, 2012), where the focus is on characterizing a class of switching signals under which the growth of the state trajectory is bounded by some increasing function of the output norm.

The viewpoint adopted in this article builds on the observability notions studied in (Tanwani et al., 2013) and (Tanwani and Trenn, 2012, 2017a) where we use the output information from different modes active over an interval to recover the value of the state, either at the start of the interval (observability), or at the end of the interval (determinability). Due to this generalized notion, we have to introduce unobservable dynamics over an interval which not only depend on the unobservable dynamics of

Email addresses: aneel.tanwani@laas.fr (Aneel Tanwani), s.trenn@rug.nl (Stephan Trenn)

individual subsystems but also their activation times. The detectability notion proposed in this article thus relates to the stability of the unobservable dynamics over an interval (and not the individual subsystems). Inspired by the fact that detectability is a sufficient condition for designing state observers for linear systems, we use these ideas to propose an observer design algorithm for switched DAEs.

The contribution of this paper lies in proposing this novel notion of detectability for switched systems (see Section 3 for the formal definition) and design state estimators for systems satisfying the detectability assumption in an appropriate sense. This work builds on our two conference papers: geometric conditions for detectability of switched DAEs were studied in (Tanwani and Trenn, 2015) and the preliminary design of the observer was proposed in (Tanwani and Trenn, 2017b). Using the presentation of the later article as a template, this paper provides additional details, rigorous proofs of the results, and the simulation results which were not a part of the conference paper. To the best of our knowledge, these results are also new for the case of switched ordinary differential equations (switched ODEs), as the previous works have only dealt with observable switched systems (Tanwani et al., 2013). It turns out that an observer for the detectable case has to work fundamentally different to our observer proposed for the determinable case. We illustrate this by the following simple example.

Example 1. Consider the switched ODE on the interval $[0, 3)$ given by

$$\begin{array}{l|l} \dot{x}_1(t) = 0 & \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = 0 & \dot{x}_2(t) = 0 \\ \dot{x}_3(t) = 0 & \dot{x}_3(t) = x_2(t) - x_3(t) \\ y(t) = x_1(t) & y(t) = 0 \\ t \in [0, 1) \cup [2, 3), & t \in [1, 2). \end{array}$$

Based on our previous work (Tanwani et al., 2013; Tanwani and Trenn, 2017a), if we restrict our attention to the interval $[0, 3)$, then $y(t) \equiv 0$ on this interval implies $x_1(t) \equiv x_2(t) \equiv 0$, and hence (x_1, x_2) is observable. Also, the identically zero output would imply that the magnitude of x_3 decreases, which is the notion of detectability we adopt in this paper (see Section 3). It is possible to design an impulsive estimator with states $\hat{x}_1, \hat{x}_2, \hat{x}_3$ which copies the system dynamics over the interval $[0, 3)$, and at $t = 3$ we reset the estimator state as

$$\begin{pmatrix} \hat{x}_1(3) \\ \hat{x}_2(3) \end{pmatrix} := \mathcal{O}(y_{[0,3)})$$

for some map \mathcal{O} , so that, if $e = x - \hat{x}$ denotes the state estimation error, we have

$$\left| \begin{pmatrix} e_1(3) \\ e_2(3) \end{pmatrix} \right| \leq \alpha \begin{pmatrix} e_1(0) \\ e_2(0) \end{pmatrix}$$

for some desired $\alpha \in (0, 1)$. However, for the unobservable

error e_3 , we get

$$\begin{aligned} \dot{e}_3(t) &= 0, & t \in [0, 1) \cup [2, 3) \\ \dot{e}_3(t) &= e_2(t) - e_3(t), & t \in [1, 2) \end{aligned} \quad (1)$$

and hence

$$e_3(3) = e^{-1}e_3(0) + (1 - e^{-1})e_2(0).$$

Thus, for a large initial value $e_2(0)$, the final error $e_3(3)$ may be significantly larger than $e_3(0)$ and therefore a direct application of our previous presented observer to detectable systems will not work. The underlying problem for this example is that it is not enough to have a good estimate of the observable states at the end of the considered interval, but the estimate must be available already when the observable states influence the unobservable states.

The remainder of the paper is structured as follows: In Section 2 we formally introduce the system class of switched DAEs and also highlight the importance of taking induced Dirac impulses into account. Afterwards we introduce in Section 3 the notion of detectability. In particular, we introduce the notion of uniform interval-detectability, which is fundamental for our observer design, which we present in Section 4. The key result in Section 4 is Theorem 13, which shows how the ideal correction term decreases the estimation error. Convergence of the observer for non-ideal correction terms is shown in Theorem 15 in Section 5.

2. Preliminaries

2.1. Switched DAEs

We consider switched linear DAEs of the form

$$\begin{aligned} E_\sigma \dot{x} &= A_\sigma x + B_\sigma u \\ y &= C_\sigma x + D_\sigma u \end{aligned} \quad (2)$$

where x, u, y denote the state (with dimension $n \in \mathbb{N}$), input (with dimension $u \in \mathbb{N}$) and output (with dimension $y \in \mathbb{N}$) of the system, respectively. The switching signal $\sigma : [0, \infty) \rightarrow \mathbb{N}$ is a piecewise constant, right-continuous function of time and in our notation it changes its value at time instants $0 < t_1 < t_2 < \dots$ called the switching times. We adopt the convention that over the interval $[t_k, t_{k+1})$ of length $\tau_k := t_{k+1} - t_k$, the active mode is defined by the quintuple $(E_k, A_k, B_k, C_k, D_k) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times u} \times \mathbb{R}^{y \times n} \times \mathbb{R}^{y \times u}$, $k \in \mathbb{N}$ and $t_0 := 0$. If $E_k = I$ for all $k \in \mathbb{N}$ we call (2) a switched ODE. In general, E_k is not assumed to be invertible, which means that in addition to differential equations the state x has to satisfy certain algebraic constraints. At a switching instant the algebraic constraints before the switch and the algebraic constraints after the switch do not match in general, i.e. the state variable has to jump in order to satisfy the algebraic constraints after the switch. These induced jumps are a first major difference to switched ODEs (which

do not exhibit jumps unless one imposes some additional external jump rules). The second major difference is the possible presence of Dirac impulses in the state variable x in response to a state jump. The following example shows this effect and also the role of the Dirac impulses in the state estimation problem.

Example 2. Consider the switched DAE given by, $i \in \mathbb{N}$,

$$\begin{array}{l|l} t \in [2i, 2i+1) & t \in [2i+1, 2i+2) \\ \hline \dot{x}_1 = x_3 & 0 = x_1 \\ \dot{x}_2 = 0 & \dot{x}_1 = x_2 \\ \dot{x}_3 = 0 & \dot{x}_3 = 0 \\ \dot{x}_4 = x_3 - x_4 & \dot{x}_4 = x_3 - x_4 \\ y = 0 & y = x_2 \end{array}$$

The dynamics for x_3 and x_4 are actually non-switched and it is therefore obvious that the overall switched system can only be detectable when it is possible to determine x_3 from the output. On the intervals $[2i, 2i+1)$, $i \in \mathbb{N}$, the output is zero by definition and on the open intervals $(2i+1, 2i+2)$, $i \in \mathbb{N}$, it holds that $x_1 = 0$, hence $y = x_2 = \dot{x}_1 = 0$, i.e. the output is zero almost everywhere. Consequently, we are not able to deduce anything about x_3 from the output if we do not take into account what the output is doing at the switching times. So what is x_2 doing at the switching times $t = 2i+1$, $i \in \mathbb{N}$? The state x_1 jumps from $x_1((2i+1)^-)$ to $x_1((2i+1)^+) = 0$, hence x_2 contains the derivative of this jump! The derivative of a jump is only well defined in a distributional (generalized functions) solution framework; in this framework x_2 contains a Dirac impulse with magnitude $-x_1((2i+1)^-)$ and this Dirac impulse is visible at the output. Consider now the switching time $t = 3$ then we can deduce from the Dirac impulse of the output at $t = 3$ the value $x_1(3^-)$. We know that $x_1 = 0$ on $(1, 2)$, $\dot{x}_1 = x_3$ on $[2, 3)$ and x_3 is constant, hence $x_1(3^-) = x_1(2^-) + (3-2)x_3 = x_3$. In particular, if we observe a zero output (including zero Dirac impulses) we can conclude that $x_3 = 0$, $(x_1, x_2) = (0, 0)$ on $(1, \infty)$ and $x_4(t) \rightarrow 0$ as $t \rightarrow \infty$. In summary, the above switched DAE is detectable but it is not possible to estimate the state without taking into account the Dirac impulses in the output.

The above example shows that an observer design which does not utilize the information from possible Dirac impulses in the output will not work for general switched DAEs. We will therefore recall now the distributional solution framework for (2) as introduced in Trenn (2009).

2.2. Distributional solution framework

Let \mathbb{D} denote the space of distributions in the sense of Schwartz (1950, 1951), i.e. $D \in \mathbb{D}$ if, and only if, $D : \mathcal{C}_0^\infty \rightarrow \mathbb{R}$ is linear and continuous, where \mathcal{C}_0^∞ is the space of test functions consisting of smooth functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ with compact support and equipped with a suitable topology. Any locally integrable function $f : \mathbb{R} \rightarrow \mathbb{R}$

induces a distribution $f_{\mathbb{D}} \in \mathbb{D}$ given by

$$f_{\mathbb{D}}(\varphi) := \int_{\mathbb{R}} f \varphi.$$

For differentiable f it is easily seen via integration by parts that

$$(f')_{\mathbb{D}}(\varphi) = -f_{\mathbb{D}}(\varphi'),$$

which motivates the definition of the derivative of a general distribution $D \in \mathbb{D}$:

$$D'(\varphi) := -D(\varphi').$$

For some interval $\mathcal{I} \subseteq \mathbb{R}$ let $\mathbb{1}_{\mathcal{I}}$ be the indicator function of \mathcal{I} , i.e. $\mathbb{1}_{\mathcal{I}}(t) = 1$ for $t \in \mathcal{I}$ and $\mathbb{1}_{\mathcal{I}}(t) = 0$ otherwise. Then the Dirac impulse δ can be defined as the distributional derivative of the unit jump (or Heaviside step function), i.e.

$$\delta := ((\mathbb{1}_{[0, \infty)})_{\mathbb{D}})'$$

Note that $\delta(\varphi) = \varphi(0)$ for any test function $\varphi \in \mathcal{C}_0^\infty$. The Dirac impulse at $t \in \mathbb{R}$, denoted by δ_t is the distributional derivative of $\mathbb{1}_{[t, \infty)}$. As shown in (Trenn, 2009) it is not possible to use the space \mathbb{D} directly as the underlying solution space for the switched DAE (2). Instead, the smaller space of piecewise-smooth distributions $\mathbb{D}_{\text{pwc}^\infty}$ will be used as underlying solution space for (2), where

$$\mathbb{D}_{\text{pwc}^\infty} := \left\{ D = f_{\mathbb{D}} + \sum_{t \in T} D_t \left| \begin{array}{l} f \in \mathcal{C}_{\text{pw}}^\infty, T \subseteq \mathbb{R} \text{ discrete,} \\ D_t \in \text{span}\{\delta_t, \delta_t', \delta_t'', \dots\} \end{array} \right. \right\},$$

i.e. a piecewise-smooth distribution is the sum of a piecewise-smooth function and Dirac impulses (and their derivatives) at isolated points in time. Any piecewise smooth distribution $D = f_{\mathbb{D}} + \sum_{t \in T} D_t \in \mathbb{D}_{\text{pwc}^\infty}$ can be evaluate at time $t \in \mathbb{R}$ in three different ways:

$$D(t^+) := f(t^+) := \lim_{\varepsilon \searrow 0} f(t + \varepsilon),$$

$$D(t^-) := f(t^-) := \lim_{\varepsilon \searrow 0} f(t - \varepsilon),$$

$$D[t] := \begin{cases} D_t, & t \in T \\ 0, & \text{otherwise} \end{cases}$$

where we denote by $D[t]$ the impulsive part of D at time t . Note that for any $t \in T$

$$D[t] = \sum_{j=0}^{n_t} \alpha_j^t \delta_t^{(j)}$$

for some finite $n_t \in \mathbb{N}$ and $\alpha_0^t, \alpha_1^t, \dots, \alpha_{n_t}^t \in \mathbb{R}$. Furthermore, the product of a piecewise-smooth function with a piecewise-smooth distribution is well defined, in particular, (2) can be evaluated for piecewise-smooth distributions. It is also possible to define the restriction to intervals for piecewise-smooth distributions, in particular, for any interval $[a, b] \subseteq \mathbb{R}$ we have

$$D_{[a, b]} = \mathbb{1}_{[a, b]} D.$$

Lemma 3 (cf. Trenn (2009)). *Consider the switched DAE (2) and assume that each matrix pair (E_p, A_p) is regular, i.e. $\det(sE_p - A_p)$ is not the zero polynomial. Then for every $u \in \mathbb{D}_{\text{pwC}\infty}^u$, any $x_0 \in \mathbb{R}^n$ and any interval $[a, b) \subseteq [0, \infty)$ there exists $x \in \mathbb{D}_{\text{pwC}\infty}^n$ uniquely defined on $[a, b)$ such that $x(a^-) = x_0$ and (2) holds as an equation of piecewise-smooth distributions restricted to $[a, b)$.*

This motivates the following solution definition of (2).

Definition 4 (Solution of switched DAE). *A tuple (x, u, y) (or just x when u and y are clear) is called a solution of (2) on an interval \mathcal{I} if $x \in \mathbb{D}_{\text{pwC}\infty}^n$, $u \in \mathbb{D}_{\text{pwC}\infty}^u$, $y \in \mathbb{D}_{\text{pwC}\infty}^y$ and (2) restricted to \mathcal{I} holds in the distributional sense. If $\mathcal{I} = [0, \infty)$ we omit “on the interval $[0, \infty)$ ” in the following.*

3. Detectability Notions

Roughly speaking, in classical literature on nonswitched systems, a dynamical system is called detectable if, for a fixed input and an observed output, the trajectories starting from every pair of indistinguishable initial states converge to a common trajectory asymptotically. The formal definition for switched DAEs is as follows:

Definition 5. *The switched DAE (2) is called detectable for a given switching signal σ , if there exists a class \mathcal{KL} function¹ $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ such that, for any two distributional solutions (x_1, u, y) , (x_2, u, y) of (2) we have*

$$|x_1(t^+) - x_2(t^+)| \leq \beta(|x_1(0^-) - x_2(0^-)|, t), \quad \forall t \geq 0. \quad (3)$$

Because of linearity the definition can be simplified to the case that $u = 0$ and $y = 0$, in particular, convergence to zero has only to be checked for the homogeneous system and the initial states in

$$\mathcal{N}^\sigma := \left\{ x^0 \in \mathbb{R}^n \mid \begin{array}{l} (x, u = 0, y = 0) \text{ solves (2)} \\ \wedge x(0^-) = x^0 \end{array} \right\}, \quad (4)$$

or in other words, detectability is the same as asymptotic stability of the switched DAE (2) with $u = 0$ and $y = 0$.

Remark 6. *In contrast to previous works on stability of switched DAEs (Liberzon and Trenn, 2009, 2012) we do not require impulse-freeness of solutions for asymptotic stability. The reason is that the presence of Dirac impulses may actually help to make certain states observable (c.f. Example 2), hence the exclusion of Dirac impulses may exclude an important class of problems where Dirac impulses are needed for observability (or detectability). It should also be noted that the magnitude of the Dirac impulses is always proportional to the state value prior to the*

time the Dirac impulse occurs (c.f. the explicit expression (A.3) in the Appendix), i.e. when the state converges to zero as $t \rightarrow \infty$ the magnitude of the Dirac impulses also converges to zero (under an additional mild boundedness assumption on (E_k, A_k) as $k \rightarrow \infty$).

Computation of the set \mathcal{N}^σ in general depends on all switching times and the data of all subsystems. For certain applications, such as state estimation which we discuss later, it may be desirable to work with system data available on finite intervals only, and in that case, Definition 5 may not be suitable. To overcome this problem, we consider the system behavior on finite intervals, and introduce the notion of interval-detectability:

Definition 7 (Interval-detectability). *The switched DAE (2) is called $[t_p, t_q)$ -detectable for a given switching signal σ , if there exists a class \mathcal{KL} function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ with*

$$\beta(r, t_q - t_p) < r, \quad \forall r > 0 \quad (5a)$$

and for any local solution $(x, u = 0, y = 0)$ of (2) on $[t_p, t_q)$ we have

$$|x(t^+)| \leq \beta(|x(t_p^-)|, t - t_p), \quad \forall t \in [t_p, t_q). \quad (5b)$$

One should be aware, that a solution on some interval is not always a part of a solution on a larger interval. Consequently, detectability does not always imply interval-detectability: The switched system $0 = x$ on $[t_0, t_1)$ and $\dot{x} = 0$ on $[t_1, \infty)$ with zero output is obviously detectable (because zero is the only global solution), but it is not interval-detectable on $[t_1, s)$ for any $s > t_1$ because on $[t_1, s)$ there are nonzero solutions which do not converge towards zero.

Furthermore, we would like to emphasize that the interval $[t_p, t_q)$ in general contains multiple switches, i.e. it is *not assumed* that the individual modes of the switched systems are detectable. We need some uniformity assumption to conclude that interval-detectability on each interval of a partition of $[t_0, \infty)$ implies detectability:

Assumption 1 (Uniform interval-detectability). *Consider the switched system (2) with switching signal σ and switching times t_k , $k \in \mathbb{N}$. Assume that there exists a strictly increasing sequence $(q_i)_{i=0}^\infty$ with $q_0 > 0 =: q_{-1}$ such that for $p_i := q_{i-1}$ the system is $[t_{p_i}, t_{q_i})$ -detectable with \mathcal{KL} -function β_i for which additionally it holds that*

$$\beta_i(r, t_{q_i} - t_{p_i}) \leq \alpha r, \quad \forall r > 0, \forall i \in \mathbb{N}, \quad (6a)$$

$$\beta_i(r, 0) \leq M r, \quad \forall r > 0, \forall i \in \mathbb{N}, \quad (6b)$$

for some uniform $\alpha \in (0, 1)$ and $M \geq 1$.

We can now show the following result:

Proposition 8. *If the switched system (2) is uniformly interval-detectable in the sense of Assumption 1 then (2) is detectable.*

¹A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is called a class \mathcal{KL} function, if 1) for each $t \geq 0$, $\beta(\cdot, t)$ is continuous, strictly increasing, with $\beta(0, t) = 0$; 2) for each $r \geq 0$, $\beta(r, \cdot)$ is decreasing and converging to zero as $t \rightarrow \infty$.

Proof. Let for $i \in \mathbb{N}$

$$\widehat{\beta}_i(r, t - t_{p_i}) = Mr - (t - t_{p_i}) \frac{Mr(1 - \alpha)}{t_{q_i} - t_{p_i}},$$

i.e. for each $r > 0$ the function $\widehat{\beta}_i(r, \cdot)$ is linear on $[t_{p_i}, t_{q_i})$ and decreasing from Mr towards αMr . Now let

$$\beta(r, t) := \max \left\{ \beta_i(\alpha^i r, t - t_{p_i}), \widehat{\beta}_i(\alpha^i r, t - t_{p_i}) \right\},$$

where $i \in \mathbb{N}$ is such that $t \in [t_{p_i}, t_{q_i})$. Clearly, for fixed t , $\beta(\cdot, t)$ is continuous and strictly increasing. From (6a) and $M \geq 1$ it follows that

$$\beta(r, t_{q_i}^-) = \max \left\{ \beta_i(\alpha^i r, t_{q_i} - t_{p_i}), M\alpha^{i+1}r \right\} = M\alpha^{i+1}r$$

and, invoking (6b),

$$\beta(r, t_{p_i}) = \max \left\{ \beta_i(\alpha^i r, 0), M\alpha^i r \right\} = M\alpha^i r.$$

Because $q_i = p_{i+1}$, continuity of $\beta(r, \cdot)$ with fixed $r > 0$ follows. Furthermore, on each interval $[t_{p_i}, t_{q_i})$ the function $\beta(r, \cdot)$ is strictly decreasing as a maximum of two strictly decreasing functions. Additionally, $\beta(r, t_{p_i}) = M\alpha^i r$ with $\alpha \in (0, 1)$ implies that $\beta(r, t)$ converges to zero as $t \rightarrow \infty$. So β is a \mathcal{KL} -function and it remains to be shown that $|x(t^+)| \leq \beta(|x(t_0^-)|, t)$ for any solution $(x, u = 0, y = 0)$ of (2). First observe, that by (6a) and continuity of β_i it follows that

$$|x(t_{p_{i+1}}^-)| = |x(t_{q_i}^-)| \leq \beta_i(|x(t_{p_i}^-)|, t_{q_i} - t_{p_i}) \leq \alpha |x(t_{p_i}^-)|,$$

hence $|x(t_{p_i}^-)| \leq \alpha^i |x(0^-)|$. Therefore,

$$\begin{aligned} |x(t^+)| &\leq \beta_i(|x(t_{p_i}^-)|, t - t_{p_i}) \leq \beta_i(\alpha^i |x(0^-)|, t - t_{p_i}) \\ &\leq \beta(|x(0^-)|, t). \end{aligned}$$

□

The uniformity conditions (6a) and (6b) are both crucial, see (Tanwani and Trenn, 2017b, Example 2) for counterexamples.

Example 9 (Example 1 revisited). *Consider the system in Example 1 with periodic switching where the mode sequence and activation times defined for the interval $[0, 3)$ are repeated on the interval $[3i, 3i + 3)$, $i \in \mathbb{N}$. It can be verified that the resulting system is uniformly interval-detectable, and hence detectable by Proposition 8. To see this, we consider the sequence $q_i = 3i + 3$ and let*

$$\beta(r, s) := r e^{2-s}.$$

The function β satisfies the inequalities (6a) and (6b), with $\alpha = e^{-1}$ and $M = e^2$, respectively. The constraint $y \equiv 0$ yields $x_1 = x_2 \equiv 0$, and it can be verified that

$$|x_3(t)| \leq e^{2-(t-3i)} x_3(3i), \quad t \in [3i, 3i + 3).$$

Remark 10. *Proposition 8 can actually be seen as a statement about asymptotic stability of switched systems and when it is possible to conclude asymptotic stability from some stability notion on finite intervals. Furthermore, the statement carries over to the nonlinear case without much change, because in the proof we did not exploit the special (linear) form of the switched system (2).*

4. Observer design

Our observer design is an extension of the algorithm proposed in (Tanwani and Trenn, 2017a) for the determinable case (in particular, the interval-wise observer design), i.e. we propose an impulsive observer which consists of a system copy and a correction term which updates the state of the system copy at the end of the detectability interval.

More formally, under the uniform interval-detectability (Assumption 1) with detectability intervals $[t_{p_i}, t_{q_i})$, $i \in \mathbb{N}$, the state estimator is given by $\widehat{x} := \sum_{i \in \mathbb{N}} (\widehat{x}_i)_{[t_{p_i}, t_{q_i})}$ with

$$\left. \begin{aligned} E_\sigma \dot{\widehat{x}}_i &= A_\sigma \widehat{x}_i + B_\sigma u, \\ \widehat{y} &= C_\sigma \widehat{x}_i + D_\sigma u, \end{aligned} \right\} \text{ on } [t_{p_i}, t_{q_i}), \quad (7)$$

$$\widehat{x}_{i+1}(t_{q_i}^-) = \widehat{x}_i(t_{q_i}^-) - \xi_i.$$

where $\xi_i \in \mathbb{R}^n$ is a state estimation correction obtained from the available data on the interval $[t_{p_i}, t_{q_i})$ applied at the end of the corresponding interval. Similar to the technique adopted in (Tanwani and Trenn, 2017a), the correction term ξ_i is obtained by collecting the local observability data for each mode. However, these local data is combined in a fundamentally different way compared to (Tanwani and Trenn, 2017a), because ξ_i is obtained by composing the local observability data backward in time first and then propagating this forward in time under the error dynamics, c.f. Example 1.

In particular, a much more complicated algorithm is needed to obtain the correction term at the end of the interval. In fact, it consists of the three following steps which have to be carried out on each of the detectability intervals $[t_{p_i}, t_{q_i})$:

- (i) Collect local observability data for each mode synchronous to the system dynamics from the measured input and output over the interval $[t_{p_i}, t_{q_i})$.
- (ii) Propagate back the collected information to obtain an estimation correction ξ_i^{left} at the beginning of the detectability interval.
- (iii) Propagate forward the correction term ξ_i^{left} to obtain the actual estimation correction ξ_i at the end of the interval.

We will now explain each of the steps in detail, for that we drop the index i and just consider the generic detectability interval $[t_p, t_q)$ for some $q > p \geq 0$. It is helpful to also introduce the estimation error $e := \widehat{x} - x$ (which we don't know, because x is not known) and the corresponding output mismatch $y^e := \widehat{y} - y$ (which we know). It is easily seen that the error is governed by the following homogeneous switched DAE on $[t_p, t_q)$:

$$E_\sigma \dot{e} = A_\sigma e, \quad y^e = C_\sigma e \quad (8)$$

and the idea of the observer is to estimate the error signal e from the measured output mismatch y^e .

Remark 11. A key feature of our observer is the consideration of the homogeneous error dynamics (8) not only in the analysis but also in the implementation of our observer. In particular, it is not necessary to store the input and output values over a (possibly long) time interval to carry out Steps (ii) and (iii), see also Remark 14. This approach is only possible because the observer consists of a system copy without a continuous update of the state estimation based on an output error injection; instead our observer is an impulsive observer in the sense that only at isolated time points the state estimation is updated discontinuously. Another reason not to use continuous updates of the state estimations via output error injection is the problem that the observable subspace is not necessarily aligned with the original coordinates. While for the original Example 1 it would be possible to continuously update the estimation of x_1 already on the first interval; this update is not possible if we slightly change the example such that the output on the intervals $[0, 1)$ and $[2, 3)$ takes the form $y = x_1 + x_2$. It is easily seen that also with this output the switched system is detectable; however, now it is unclear how the local (one-dimensional) observability information available on the interval $[0, 1)$ can be injected continuously to update the state variable in a meaningful way.

4.1. Collecting local observability data for each mode

For each mode k with $p \leq k \leq q - 1$ consider the local unobservable space:

$$\mathcal{W}_k := \left\{ e_0 \in \mathbb{R}^n \left| \begin{array}{l} e(t_k^-) = e_0, \text{ where } (e, y^e = 0) \\ \text{solves (8) on } [t_k, t_{k+1}) \end{array} \right. \right\} \quad (9)$$

Defining Π_k , O_k^{diff} and O_k^{imp} in terms of (E_k, A_k) as in the Appendix A, it can be shown (c.f. (Tanwani and Trenn, 2013, 2017a)) that

$$\mathcal{W}_k = \Pi_k^{-1} \ker O_k^{\text{diff}} \cap \ker O_k^{\text{imp}}.$$

Note that in general Π_k is not invertible and Π_k^{-1} stands for the set-valued preimage.

Remark 12 (Different definitions of local unobservable space). In (Tanwani and Trenn, 2012, 2013, 2017a) slightly different definitions of the local unobservable spaces are used. The difference is based on the different solution interval; in the previous works this interval was (t_{k-1}, t_{k+1}) or $(t_{k-1}, t_k]$, while here the interval is $[t_k, t_{k+1})$. As a consequence \mathcal{W}_k here only depends on the system's parameters of mode k and not on parameters of two modes. The different definitions are motivated by the overall observability notion studied. In our first work we defined the local unobservable space in such a way that all information around a single switching time is utilized, in particular, the local unobservable space for a system with a single switch matched the overall unobservable space. Our later works focused on observer design and determinability (i.e. the ability to determine the state value at the end of

the observation interval), for this reason it made sense to consider as local information the continuous output measurement before the current switching time and the Dirac impulses instantaneously induced at that switching time. Here our focus is on observable components of the state at the beginning of the observed interval (cf. Example 1) which motivated the definition (9).

If the output mismatch y^e is nonzero then the value of e in (8) prior to the switching time t_k can be decomposed as

$$e(t_k^-) = W_k w_k + Z_k z_k,$$

where $\text{im } W_k = \mathcal{W}_k$ and $\text{im } Z_k = \mathcal{W}_k^\perp$ and W_k, Z_k are orthonormal matrices. In particular, $z_k = Z_k^\top e(t_k^-)$ is the observable part of the error $e(t_k^-)$ based on the knowledge on the interval $[t_k, t_{k+1})$. It is possible to write the observable part z_k in terms of y^e :

$$z_k = \mathfrak{D}_k(y_{[t_k, t_{k+1})}^e) \quad (10)$$

with some operator \mathcal{O}_k which evaluates the impulsive part $y^e[t_k]$ as well as the smooth part $y_{(t_k, t_{k+1})}^e$ (possibly depending on the derivatives of y^e). The construction of this “ideal” observability operator \mathcal{O}_k is provided in Appendix B.1. One may also refer to (Tanwani and Trenn, 2017a, Section 5) for a detailed treatment. In practice, only an approximation $\hat{\mathcal{O}}_k$ of \mathcal{O}_k will be available, this will be discussed in Section 5.

4.2. Combining local information backwards in time

Next we want to combine the observable information $z_p, z_{p+1}, \dots, z_{q-1}$ collected on the interval $[t_p, t_q)$ via (10), to arrive at an expression for $e(t_p^-)$. To do so, we first quantify the information that can be extracted from the output over an interval $[t_k, t_q)$ by introducing the subspace

$$\mathcal{N}_k^q := \left\{ e_0 \in \mathbb{R}^n \left| \begin{array}{l} e(t_k^-) = e_0, \text{ where } (e, y^e = 0) \\ \text{solves (8) on } [t_k, t_q) \end{array} \right. \right\} \quad (11)$$

which can be recursively calculated (backwards in time, i.e. for $k = q - 1, q - 2, \dots, p$), see (B.7) in the Appendix. We then decompose the state estimation error just before the interval $[t_k, t_q)$ accordingly:

$$e(t_k^-) = M_k^q \mu_k + N_k^q \nu_k \quad (12)$$

for some vectors μ_k and ν_k of appropriate dimension; here, M_k^q and N_k^q are the matrices with orthonormal columns such that

$$\text{im } N_k^q = \mathcal{N}_k^q \quad \text{and} \quad \text{im } M_k^q = (\mathcal{N}_k^q)^\perp.$$

As shown in Appendix B.2, there exists a matrix \mathcal{F}_k^q given in terms of $M_{k+1}^q, N_{k+1}^q, (E_k, A_k)$ and the duration time $\tau_k = t_{k+1} - t_k$ such that for $p \leq k \leq q - 2$

$$\mu_k = \mathcal{F}_k^q \begin{pmatrix} z_k \\ \mu_{k+1} \end{pmatrix}.$$

and $\mu_{q-1} = z_{q-1}$. Note that by construction, for all $p \leq k \leq q-1$

$$e(t_k^-) - M_k^q \mu_k \in \mathcal{N}_k^q.$$

Now the ideal estimation error correction is

$$\begin{aligned} \xi^{\text{left}} &:= M_p^q \mu_p \\ &= M_p^q \mathcal{F}_p^q \left(\mathcal{F}_{p+1}^q \left(\mathcal{F}_{p+2}^q \left(\dots \mathcal{F}_{q-2}^q \left(\begin{matrix} z_p \\ z_{p+1} \\ \vdots \\ z_{q-2} \\ z_{q-1} \end{matrix} \right) \right) \right) \right) \right) \quad (13) \\ &=: \mathcal{O}_p^{q-1} \mathbf{z}_p^{q-1}, \end{aligned}$$

where $\mathbf{z}_p^{q-1} = (z_p/z_{p+1}/\dots/z_{q-1})$; here the notation $(\dots/\dots/\dots)$ stands for a vector (or matrix) resulting from stacking all entries over each other. In fact, by construction the following is true:

$$\boxed{e(t_p^-) - \xi^{\text{left}} \in \mathcal{N}_p^q \text{ and } \xi^{\text{left}} \in \mathcal{N}_p^{q\perp},}$$

i.e. we are able to obtain the orthogonal projection of $e(t_p^-)$ onto \mathcal{N}_p^q without actually knowing $e(t_p^-)$.

4.3. Propagating correction term forward in time

For the detectability interval $[t_p, t_q]$, let ξ^{left} be given as above, then let

$$\boxed{\xi := \Phi_p^q \xi^{\text{left}},}$$

where $\Phi_p^p = I$ and Φ_p^{k+1} , $k = p, p+1, \dots, q-1$ is recursively given by

$$\Phi_p^{k+1} = \mathbf{e}^{A_k^{\text{diff}} \tau_k} \Pi_k \Phi_p^k \quad (14)$$

with Π_k and A_k^{diff} are given as in Definition 17 in the Appendix. In fact, as a consequence from Lemmas 18 and 20 in the Appendix, Φ_p^q is the transition matrix of the homogeneous error DAE (8) from $e(t_p^-)$ to $e(t_q^-)$. We then have the following result:

Theorem 13. *Consider the switched DAE (2) which is detectable on $[t_p, t_q]$ with corresponding \mathcal{KL} -function β . Let (\hat{x}, \hat{y}) be the solution of the system copy*

$$\begin{aligned} E_\sigma \dot{\hat{x}} &= A_\sigma \hat{x} + B_\sigma u, \\ \hat{y} &= C_\sigma \hat{x} + D_\sigma u \end{aligned} \quad (15)$$

on $[t_p, t_q]$. Based on the output mismatch $y^e = \hat{y} - y$ let

$$\xi = \Phi_p^q \xi^{\text{left}} = \Phi_p^q \mathcal{O}_p^{q-1} \mathbf{z}_p^{q-1}$$

where Φ_p^q is given by (14), \mathcal{O}_p^{q-1} is given (13) and $\mathbf{z}_p^{q-1} = (z_p/z_{p+1}/\dots/z_{q-1})$ with $z_k = \mathfrak{D}_k(y_{[t_k, t_{k+1}]})$, $k = p, p+1, \dots, q-1$ is given by (10). Then

$$\begin{aligned} |\hat{x}(t_q^-) - \xi - x(t_q^-)| &\leq \beta(|\hat{x}(t_p^-) - x(t_p^-)|, t_q - t_p) \\ &< |\hat{x}(t_p^-) - x(t_p^-)|, \end{aligned}$$

i.e. the correction term ξ indeed reduces the estimation error at the end of the interval in comparison to the estimation error at the beginning of the interval.

Proof. Let \hat{x}^{new} be the (virtual) solution of the system copy (15) with corrected initial value $\hat{x}^{\text{new}}(t_p^-) = \hat{x}(t_p^-) - \xi^{\text{left}}$. By construction $\hat{x}^{\text{new}}(t_p^-) - x(t_p^-) = e(t_p^-) - \xi^{\text{left}} \in \mathcal{N}_p^q$ hence $y = \hat{y}^{\text{new}}$ on $[t_p, t_q]$ and therefore, for all $t \in [t_p, t_q]$,

$$|\hat{x}^{\text{new}}(t^+) - x(t^+)| \leq \beta(|\hat{x}^{\text{new}}(t_p^-) - x(t_p^-)|, t - t_p).$$

Note that $e = \hat{x} - x$ as well as $e^{\text{new}} := \hat{x}^{\text{new}} - x$ are solutions of the homogenous error DAE (8), in particular

$$e(t_q^-) - e^{\text{new}}(t_q^-) = \Phi_p^q (e(t_p^-) - e^{\text{new}}(t_p^-)) = \Phi_p^q \xi^{\text{left}} = \xi,$$

or, in other words,

$$\hat{x}(t_q^-) - \xi = \hat{x}^{\text{new}}(t_q^-)$$

Finally, by construction $\xi^{\text{left}} \in \mathcal{N}_p^{q\perp}$ and therefore, by Pythagoras' Theorem,

$$\begin{aligned} |\hat{x}(t_p^-) - x(t_p^-)|^2 &= |\underbrace{\hat{x}^{\text{new}}(t_p^-) - x(t_p^-)}_{\in \mathcal{N}_p^q} + \xi^{\text{left}}|^2 \\ &= |\hat{x}^{\text{new}}(t_p^-) - x(t_p^-)|^2 + |\xi^{\text{left}}|^2 \\ &\geq |\hat{x}^{\text{new}}(t_p^-) - x(t_p^-)|^2. \end{aligned}$$

Altogether we have:

$$\begin{aligned} |\hat{x}(t_q^-) - \xi - x(t_q^-)| &= |\hat{x}^{\text{new}}(t_q^-) - x(t_q^-)| \\ &\leq \beta(|\hat{x}^{\text{new}}(t_p^-) - x(t_p^-)|, t_q - t_p) \\ &\leq \beta(|\hat{x}(t_p^-) - x(t_p^-)|, t_q - t_p) \end{aligned}$$

which is the desired estimate. \square

Remark 14. *The proof of Theorem 13 reveals that by applying the correction ξ^{left} at the beginning of the interval the output of the system copy is then identical to the output of the original system. However, for the observer design it is not necessary to rerun the system copy (in particular storing the whole input signal over the interval $[t_p, t_q]$), because we just propagate ξ^{left} via the homogenous error dynamics (8) which is independent of the in- and output. This actually allows us to calculate the error correction for an arbitrary future; this fact can be utilized to deal with time delays due to computation times, see the discussion in Section 6.*

5. Estimation errors and asymptotic convergence

In theory, it is possible to determine the observable part exactly from the output, however, in practice one can only get approximations. Nevertheless, these approximations may be as accurate as desired (e.g. by choosing appropriate gains in a Luenberger observer). Similar as in (Tanwani and Trenn, 2017a) we therefore make the following assumption about the ability to approximate the observable part to any given accuracy:

Assumption 2. For each mode k of the switched DAE (2) and a given $\varepsilon_k > 0$, there exists an estimator $\widehat{z}_k = \mathfrak{D}_k(y_{[t_k, t_{k+1}]})$ such that

$$|\widehat{z}_k - z_k| \leq \varepsilon_k |z_k|, \quad (16)$$

where $z_k = \mathfrak{D}_k(y_{[t_k, t_{k+1}]})$ is the ideal estimator of the observable part on $[t_k, t_{k+1})$ as given in Section 4.1.

Under Assumption 2, the state estimation correction in (7) for the interval $[t_{p_i}, t_{q_i})$ is given by

$$\xi_i := \Phi_{p_i}^{q_i} \mathcal{O}_{p_i}^{q_i-1} \widehat{\mathbf{z}}_{p_i}^{q_i-1}, \quad (17)$$

where $\widehat{\mathbf{z}}_{p_i}^{q_i-1} = (\widehat{z}_{p_i} / \widehat{z}_{p_i+1} / \cdots / z_{q_i-1})$.

As detailed in B.1, the observable component z_k of the estimation error $e = \widehat{x} - x$ on the interval $[t_k, t_{k+1})$ is composed of the two components z_k^{diff} and z_k^{imp} , where the former is obtained from the continuous output mismatch y^e on (t_k, t_{k+1}) and the latter is obtained from the impulsive mismatch $y^e[t_k]$. The estimation of z_k^{diff} can be reduced to the classical state estimation problem for non-switched linear ODEs and there are many methods to do that. The only non-standard aspect here is that we have to obtain the state-estimation at the beginning of the interval (t_k, t_{k+1}) and not (as usual) at the end of the interval. This does not pose any serious problems, as we can use a standard Luenberger observer on the interval (t_k, t_{k+1}) to get an estimate at the end of the interval and then propagate this estimate back in time. Since the (homogeneous) ODE dynamics are known as well as the length of the interval, we can ensure the desired estimation accuracy at the beginning of the interval by increasing the accuracy of the estimate at the end of the interval.² Another (more sophisticated) way of obtaining such estimates is by the use of “back-and-forth observer” as presented in (Shim et al., 2012), however, this requires the storage of the output over the whole interval (t_k, t_{k+1}) .

The estimation accuracy for z_k^{imp} is actually concerned with the measurement accuracy of the impulsive part $y^e[t_k]$, i.e. on how well Dirac impulses and their derivatives can be measured in practice, see (Tanwani and Trenn, 2017a) for details.

Assumption 2, together with Assumption 1, provide all the ingredients we need for obtaining converging state estimates.

Theorem 15. Consider the switched DAE (2) satisfying the uniform local detectability Assumption 1, and the local estimation accuracy Assumption 2. For the α given in (6a), choose ε_k , $k \in \mathbb{N}$, (depending on α) such that

$$c_i \varepsilon_i^{\max} \leq \widehat{\alpha} - \alpha \quad (18)$$

²In fact, consider an observable LTI system $\dot{z} = Az$, $y = Cz$ over the interval $[0, T]$, with the estimator $\hat{z} = (A - LC)\hat{z} + Ly$, $\hat{z}(0) = 0$. For every $\epsilon > 0$, there exists L such that $|\hat{z}(T) - z(T)| \leq \epsilon |z(0)|$. Choose $\hat{z}_0^* = e^{-AT} \hat{z}(T)$, and $\epsilon \leq \delta / \|e^{-AT}\|$ for some desired $\delta > 0$, then it is easily verified that $|z(0) - z_0^*| \leq \delta |z(0)|$.

for some $\widehat{\alpha} \in (\alpha, 1)$, where

$$c_i := \|\Phi_{p_i}^{q_i} \mathcal{O}_{p_i}^{q_i-1}\| \left\| \begin{bmatrix} Z_{p_i}^\top \\ Z_{p_i+1}^\top \Phi_{p_i}^{p_i+1} \\ \vdots \\ Z_{q_i-1}^\top \Phi_{p_i}^{q_i-1} \end{bmatrix} \right\|$$

and

$$\varepsilon_i^{\max} := \max \{\varepsilon_k \mid p_i \leq k \leq q_i - 1\}.$$

Then the observer given by the system copies (7), with error corrections ξ_i in (17) and the estimate \widehat{z}_k chosen to satisfy (16) for ε_k specified in (18), results in

$$\widehat{x}(t^+) \rightarrow x(t^+) \text{ as } t \rightarrow \infty,$$

i.e. the observer achieves asymptotic estimation of the state.

Proof. From Theorem 13 we know that for each detectable interval $[t_{p_i}, t_{q_i})$ the ideal correction $\xi_i^{\text{ideal}} = \Phi_p^q \mathcal{O}_p^{q-1} \mathbf{z}_p^{q-1}$ ensures

$$\begin{aligned} |\widehat{x}_i(t_{q_i}^-) - \xi_i^{\text{ideal}} - x(t_{q_i}^-)| &\leq \beta_i (|\widehat{x}_i(t_{p_i}^-) - x(t_{p_i})|, t_{q_i} - t_{p_i}) \\ &\leq \alpha |\widehat{x}_i(t_{p_i}^-) - x(t_{p_i})|. \end{aligned}$$

Hence for the actual correction term $\xi_i = \Phi_p^q \mathcal{O}_p^{q-1} \widehat{\mathbf{z}}_p^{q-1}$, we have

$$|\widehat{x}_{i+1}(t_{q_i}^-) - x(t_{q_i}^-)| \leq \alpha |\widehat{x}_i(t_{p_i}^-) - x(t_{p_i})| + |\xi_i^{\text{ideal}} - \xi_i|.$$

By assumption, $|\widehat{z}_k - z_k| \leq \varepsilon_k |z_k|$ and since $z_k = Z_k^\top e(t_k^-) = Z_k^\top \Phi_{p_i}^k e(t_{p_i}^-)$ for any k with $p_i \leq k \leq q_i - 1$, we have

$$\begin{aligned} |\xi_i^{\text{ideal}} - \xi_i| &= |\Phi_{p_i}^{q_i} \mathcal{O}_{p_i}^{q_i-1} (\widehat{\mathbf{z}}_{p_i}^{q_i-1} - \mathbf{z}_{p_i}^{q_i-1})| \\ &\leq \|\Phi_{p_i}^{q_i} \mathcal{O}_{p_i}^{q_i-1}\| \varepsilon_i^{\max} \left\| \begin{bmatrix} Z_{p_i}^\top \\ Z_{p_i+1}^\top \Phi_{p_i}^{p_i+1} \\ \vdots \\ Z_{q_i-1}^\top \Phi_{p_i}^{q_i-1} \end{bmatrix} \right\| |e(t_{p_i}^-)| \\ &= c_i \varepsilon_i^{\max} |\widehat{x}_i(t_{p_i}^-) - x(t_{p_i}^-)|. \end{aligned}$$

Altogether we have

$$|\widehat{x}_{i+1}(t_{q_i}^-) - x(t_{q_i}^-)| \leq \widehat{\alpha} |\widehat{x}_i(t_{p_i}^-) - x(t_{p_i}^-)|,$$

i.e. on each detectability interval $[t_{p_i}, t_{q_i})$ the estimation error $\widehat{x} - x$ decreases uniformly by a factor $\widehat{\alpha}$ and the same proof technique as in Proposition 8 shows asymptotic convergence. \square

6. Observer implementation

We have implemented the observer in Matlab and show in the next section the simulation results for the academic Examples 1 and 2. Before presenting the simulation results, we would like to discuss some implementation issues, in particular, which calculations can be carried out offline

Algorithm 1: Observer for detectable switched DAEs

Data: Modes $(E_k, A_k, B_k, C_k, D_k)$, $k = 0, 1, 2, \dots$
switching times t_k , $k = 1, 2, 3, \dots$, $t_0 := 0$
update-time indicies q_i , $i = 0, 1, 2, \dots$, $q_{-1} := 0$
access to input u and output y
Result: State-estimation \hat{x}
INITIALIZATION (OFFLINE):
forall modes k **do**
 Calculate $\Pi_k, A_k^{\text{diff}}, E_k^{\text{imp}}, C_k^{\text{diff}}$ as in Def. 17
 Calculate $e^{A_k^{\text{diff}} \tau_k} \Pi_k$ and $e^{-A_k^{\text{diff}} \tau_k}$
 Calculate $W_k, Z_k, Z_k^{\text{diff}}, Z_k^{\text{imp}}, U_k^{\text{obs}}$ as in App. B.1
 Calculate $S_k^{\text{diff}}, R_k^{\text{diff}}, U_k^{\text{imp}}$ as in (B.3), (B.5)
 Choose L_k s. t. $S_k - R_k L_k$ is “sufficiently” Hurwitz
 (in view of Assumption 2)
end
RUN OBSERVER ON DETECTABLE INTERVALS $[p_i, q_i]$:
forall $i \in \mathbb{N}$ **do**
 $p := q_{i-1}$, $q := q_i$
 GET LOCAL ESTIMATION DATA (ONLINE):
 forall $k = p, p+1, \dots, q-1$ **do**
 Run system copy (7) with input u
 Run Luenberger observer with gain L_k for
 (B.3) on (t_k, t_{k+1}) with output-injection $y^e = \hat{y} - y$
 → estimation of z_k^{diff} on (t_k, t_{k+1})
 Estimate impulse differences $\hat{y}[t_k] - y[t_k]$
 → estimation of η_k
 Calculate $\hat{z}_k^{\text{diff}}, \hat{z}_k^{\text{imp}}$ according to (B.4), (B.6)
 Calculate \hat{z}_k via (B.2)
 end
 COMBINE LOCAL INFORMATION BACKWARDS (OFFLINE)
 $\hat{\mu}_{q-1} := \hat{z}_{q-1}$
 forall $k = q-2, q-3, \dots, p$ (*backwards*) **do**
 Calculate subspaces \mathcal{N}_k^q recursively via (B.7)
 Choose Θ_k^q, M_k^q, U_k^q via (B.8), (B.9), (B.10)
 Calculate $\hat{\mu}_k$ from $\hat{z}_k, \hat{\mu}_{k+1}$ via (B.11)
 end
 $\xi^{\text{left}} := M_p^q \hat{\mu}_p$
 PROPAGATE CORRECTION FORWARD (OFFLINE)
 $\xi_i \leftarrow \xi^{\text{left}}$
 forall $k = p+1, p+2, \dots, q-1$ (*forward*) **do**
 Calculate $\xi_i \leftarrow e^{A_k^{\text{diff}} \tau_k} \Pi_k \xi_i$
 end
 Update state-estimation $\hat{x}(t_{q_i}^-) \leftarrow \hat{x}(t_{q_i}^-) - \xi_i$
end

and how to treat the necessary computation times. The overall structure of the observer is given in Algorithm 1.

In the initialization phase, certain matrices and subspaces are calculated for each individual subsystem; in particular, a decomposition into unobservable and observable states is carried out. It should be noted that, in practice and in our setup, these calculations have to be carried out only for finitely many modes. In fact, it suffices to carry out the calculation for all modes occurring in the next detectability interval (and these calculations can be done in parallel to running the system copy and collecting the corresponding measurements). For a suitable choice of the Luenberger gain it is necessary to know (at least

some bounds of) the values c_i in (18) and consequently (bounds on) the necessary estimation accuracies ε_k as in Assumption 2. In case of a periodic switching signal one may adapt the Luenberger gain in each iteration until the estimation error is sufficiently small such that convergence of the state-estimator is guaranteed.

In the online (or synchronous) phase of the observer, a system copy must be simulated (driven by the current input) and its output must be compared with the actual output of the system. Although the output of the original system may contain Dirac impulses at arbitrary times, for the observer only the Dirac impulses *at the switching times* are compared with the predicted Dirac impulses of the system copy at the switching times. All other Dirac impulses are induced by discontinuities in the input (and are independent of the current state), hence (at least in theory) they are identical for the system copy and the original system and do not provide any additional information for the state estimation problem. Hence the impulse measurement needs only be active around the switching times. For estimating the observable part between the switching times, one could either run a classical observer (for the system (B.3) with the desired output $y^e = \hat{y} - y$) synchronously to the system without the need to store the measures output difference y^e ; however, as the dimension of the output is usually low, it may also be feasible to store the whole trajectory y^e and carry out some more sophisticated estimation procedure offline.

Finally, after the local observability data is obtained, it must be combined in a suitable way to obtain the impulsive update ξ for the state estimation \hat{x} . Although all the involved matrices can be computed offline, the actual calculations can only be carried out after the last estimate \hat{z}_{q_i-1} is obtained, hence some unavoidable processing time $\Delta > 0$ is required to compute ξ . However, the effect of the processing time can be entirely compensated as follows: For a generic detectability interval $[t_p, t_q]$, assume that an upper bound $\Delta > 0$ is known for the time required to calculate ξ . Furthermore, we assume that $[t_p, t_q - \Delta]$ is still a detectable interval in the sense of Definition 7 (this is always the case for sufficiently small Δ). In particular, ξ^{left} is an arbitrarily good estimate of the projection of $e(t_p^-)$ on the unobservable space \mathcal{N}_p^q . Now, we just propagate forward ξ^{left} with the matrix Φ_p^q to get a good estimate of $e(t_q^-)$, and we can update \hat{x} at the correct time. The key observation is that once we have obtained ξ^{left} , we can freely choose the update time (i.e. how far we propagate forward the error correction) without losing any accuracy.

7. Simulations

7.1. Simulation of Example 1

Consider the switched ODE given in Example 1 with the periodic switching signal. It was already shown in Example 9 that this system is uniformly detectable, and under periodicity assumption, it is persistently detectable.

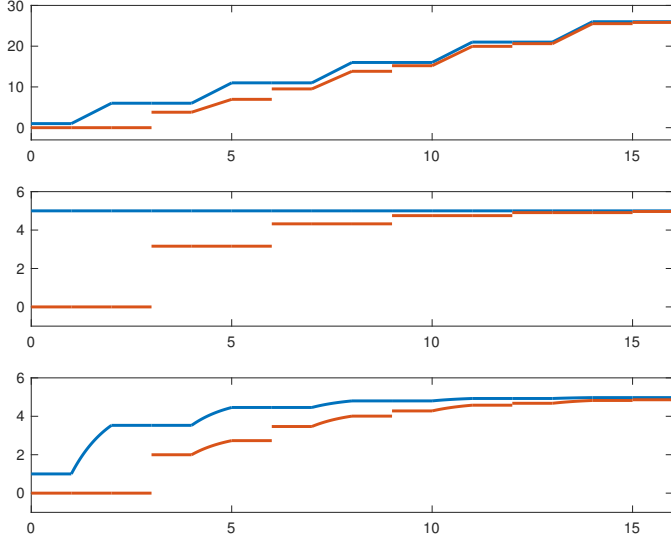


Figure 1: State estimation, x_1 (blue) and \hat{x}_1 (red) top figure; x_2 (blue) and \hat{x}_2 (red) middle figure; x_3 (red) and \hat{x}_3 (red) bottom figure.

To implement the proposed observer, we run the system copy (7) on the intervals $[3i, 3i + 3)$, $i \in \mathbb{N}$, and apply the correction term at $t_{q_i} = 3i + 3$. The correction terms are obtained by

$$\xi_i = e^{A_1 \tau_{3i+1}} M_{3i}^{3i+3} \hat{\mu}_{3i}, \quad i \in \mathbb{N},$$

where $\tau_{3i+1} = 1$, for each i , and $\hat{\mu}_{3i}$ is computed from the estimates of the observable states of individual subsystems: \hat{z}_{3i} and \hat{z}_{3i+2} . We recall that z_{3i+1} is an empty vector because the output over the interval $[3i + 1, 3i + 2)$ is zero for each i and nothing can be deduced about the state. The values of \hat{z}_{3i} and \hat{z}_{3i+2} are obtained by running a Luenberger observer (with gain $L = 1$) for the x_1 -dynamics over the intervals $[3i, 3i + 1)$ and $[3i + 2, 3i + 3)$, respectively. We use the later to first compute

$$\hat{\mu}_{3i+1} = \mathcal{F}_{3i+1}^{3i+3} \begin{pmatrix} \hat{z}_{3i+1} \\ \hat{z}_{3i+2} \end{pmatrix} = \Theta_{3i+1}^\top e^{-A_{3i+1} \tau_{3i+1}} Z_{3i+2} \hat{z}_{3i+2}$$

where $\Theta_{3i+1}^\top = (-1/\sqrt{2}, -1/\sqrt{2}, 0)$, and $Z_2 = (1/0/0)$. This leads to

$$\hat{\mu}_{3i} = \mathcal{F}_{3i}^{3i+3} \begin{pmatrix} \hat{z}_{3i} \\ \hat{\mu}_{3i+1} \end{pmatrix} = \begin{pmatrix} \Theta_{3i}^\top \hat{z}_{3i} \\ \Theta_{3i}^\top M_{3i+1} \hat{\mu}_{3i+1} \end{pmatrix}$$

where $\Theta_{3i+1}^\top = (1, 1, 0)$ and $M_{3i+1}^{3i+3^\top} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

The results of the simulation are reported in Figure 1.

It is observed that whenever a correction is applied at $t_{q_i} = 3i + 3$, the estimation error decreases and the convergence to zero is achieved asymptotically.

7.2. Simulation of Example 2

We now implement our observer on the system given in Example 2 where one of the subsystem is a DAE. As already discussed above the presence and evaluation of the

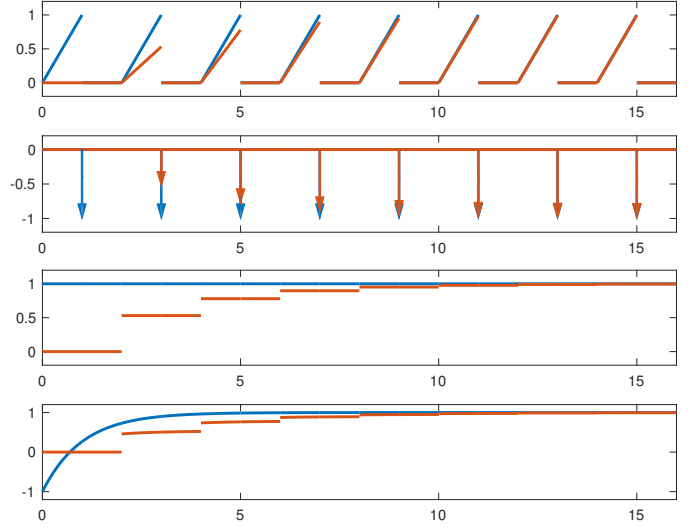


Figure 2: State estimation, x_1 (blue) and \hat{x}_1 (red) in top figure; x_2 (blue) and \hat{x}_2 (red) with Dirac impulses (shown as arrows) in second figure from top (the Dirac impulses are also visible in the output); x_3 (blue) and \hat{x}_3 (red) in third figure; x_4 (blue) and \hat{x}_4 (red) in bottom figure.

occurring Dirac impulses in the output are crucial for the state estimation.

The system is detectable on the intervals $[2i, 2i + 2)$, for $i \in \mathbb{N}$ and so we run the system copy (7) on the intervals $[2i, 2i + 2)$, and apply the correction term at $t_{q_i} = 2i + 2$. The correction term, for each $i \in \mathbb{N}$, is obtained by

$$\xi_i = e^{A_{2i+1}^{\text{diff}} \tau_{2i+1}} \Pi_{2i+1} e^{A_{2i}^{\text{diff}} \tau_{2i}} \Pi_{2i} M_{2i}^{2i+2} \hat{\mu}_{2i},$$

where we set $\tau_{2i} = \tau_{2i+1} = 1$, and

$$A_{2i}^{\text{diff}} = A_{2i} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \quad A_{2i+1}^{\text{diff}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix},$$

$$\Pi_{2i} = I_{4 \times 4}, \quad \Pi_{2i+1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The term $\hat{\mu}_{2i} \in \mathbb{R}$ is computed from the estimates of the observable states of individual subsystems: \hat{z}_{2i} and \hat{z}_{2i+1} . Because $y = 0$ on $[2i, 2i + 1)$, and the corresponding subsystem is an ODE, we set \hat{z}_{2i} to be the empty vector (nothing can be concluded from the output). Also, due to the structure of the second subsystem, the only observable information is due to impulses in the output, so that $\hat{z}_{2i+1} = \hat{z}_{2i+1}^{\text{imp}} \in \mathbb{R}$ because

$$O_{2i+1}^{\text{diff}} = 0_{4 \times 4}, \quad O_{2i+1}^{\text{imp}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

To compute $\hat{\mu}_{2i}$, we have $\hat{\mu}_{2i+1} = \hat{z}_{2i+1}$, and

$$\hat{\mu}_{2i} = \mathcal{F}_{2i}^{2i+2} \hat{\mu}_{2i+1} = \Theta_{2i}^\top e^{-A_{2i}^{\text{diff}} \tau_{2i}} Z_{2i+1} \hat{z}_{2i+1}$$

where $\Theta_{2i}^\top = \left(\frac{-1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}, 0 \right)$, and $Z_{2i+1}^\top = (1, 0, 0, 0)$. Finally, we compute $M_{2i}^{2i+2^\top} = \left[\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0 \right]$ and use it along with $\hat{\mu}_{2i}$ to compute the correction terms ξ_{2i+2} .

Since in the simulation both the system and the system copy are simulated, we would be able to obtain $\hat{z}_{2i+1}^{\text{imp}}$ without estimation error (i.e. $\varepsilon_{2i+1} = 0$ in Assumption 2) and then already after one correction we would have a perfect state-estimation. Therefore, we introduced some artificial random noise when “measuring” $y[t_{2i+1}]$, so that (16) is only satisfied with $\varepsilon_{2i+1} = 0.1$.

The results of the simulation are reported in Figure 2 where we see that the estimation error converges to zero in each of the state components. In particular, we see impulsive corrections for state component x_2 .

8. Conclusion

We have studied the notion of detectability for switched DAEs which allows us to consider the problem of observer design under relaxed assumptions on system dynamics compared to the existing works. A novel estimation algorithm is proposed which relies on propagating backward and forward the correction terms obtained by processing the measured outputs and inputs. Rigorous convergence analysis of estimation error for the proposed algorithm is carried out and the results are illustrated by studying two academic examples with simulations.

Appendix

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

A very useful characterization of regularity is the following well-known result.

Proposition 16 (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*

$$(SET, SAT) = \left(\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \right), \quad (\text{A.1})$$

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. \triangleleft

We call (A.1) a *quasi-Weierstrass form* of (E, A) following Berger et al. (2012); therein it also shown how to easily obtain (A.1) via the Wong-sequences Wong (1974).

Definition 17. *Consider the regular matrix pair (E, A) with corresponding quasi-Weierstraß form (A.1). The consistency projector of (E, A) is given by*

$$\Pi = T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T^{-1}.$$

Furthermore, let

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

Finally, if also an output matrix C is considered let

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

To see the utility of the matrices introduced in Definition 17, consider the problem of finding a trajectory x which solves the initial-trajectory problem (ITP)

$$x_{(-\infty, 0)} = x_{(-\infty, 0)}^0 \quad (\text{A.2a})$$

$$(E\dot{x})_{[0, \infty)} = (Ax)_{[0, \infty)}, \quad (\text{A.2b})$$

in some appropriate sense.

Lemma 18 (Role of consistency projector, (Trenn, 2009, Thm. 4.2.8)). *Consider the ITP (A.2) with regular matrix pair (E, A) and with arbitrary initial trajectory $x^0 \in (\mathbb{D}_{\text{pwc}^\infty})^n$. There exists a unique solution $x \in (\mathbb{D}_{\text{pwc}^\infty})^n$ and*

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

Lemma 19 ((Tanwani and Trenn, 2010, Cor. 5)). *Consider the ITP (A.2) with regular matrix pair (E, A) and the corresponding E^{imp} matrix. For the unique solution $x \in (\mathbb{D}_{\text{pwc}^\infty})^n$, it holds that*

$$x[0] = - \sum_{j=0}^{n-2} (E^{\text{imp}})^{j+1} x(0^-) \delta_0^{(j)}, \quad (\text{A.3})$$

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

Lemma 20. *For any regular matrix pair (E, A) and output matrix C , the following implication holds for all continuously differentiable (x, y) :*

$$\left. \begin{aligned} E\dot{x} &= Ax, \\ y &= Cx \end{aligned} \right\} \Rightarrow \begin{cases} \dot{x} = A^{\text{diff}}x, \\ y = C^{\text{diff}}x \end{cases}$$

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

$$x(t) = e^{A^{\text{diff}}t}x(0), \quad t \in \mathbb{R}. \quad \triangleleft$$

B. Output-to-State Mappings

B.1. Observable component of a subsystem

The local unobservable space (9) is given by

$$\mathcal{W}_k = \Pi_k^{-1} \ker O_k^{\text{diff}} \cap \ker O_k^{\text{imp}},$$

where

$$\begin{aligned} O_k^{\text{diff}} &:= [C_k^{\text{diff}}/C_k^{\text{diff}}A_k^{\text{diff}}/\dots/C_k^{\text{diff}}(A_k^{\text{diff}})^{n-1}], \\ O_k^{\text{imp}} &:= [C_kE_k^{\text{imp}}/C_k(E_k^{\text{imp}})^2/\dots/C_k(E_k^{\text{imp}})^{n-1}]. \end{aligned} \quad (\text{B.1})$$

In other words, $\ker O_k^{\text{diff}}$ denotes the unobservable space of the ODE $\dot{e} = A_k^{\text{diff}}e$, $y^e = C_k^{\text{diff}}e$, and $\ker O_k^{\text{imp}}$ denotes the impulse unobservable space in the sense that $y^e[t_k] = 0$ implies $e(t_k^-) \in \ker O_k^{\text{imp}}$.

We may now write

$$e(t_k^-) = W_k w_k + Z_k z_k,$$

where $\text{im } W_k = \mathcal{W}_k$ and $\text{im } Z_k = \mathcal{W}_k^\perp$ and W_k, Z_k are orthonormal matrices. Here z_k determines the projection of $e(t_k^-)$ onto the subspace \mathcal{W}_k^\perp . The latter can further be decomposed as

$$\mathcal{W}_k^\perp = \text{im}(O_k^{\text{diff}} \Pi_k)^\top + \text{im } O_k^{\text{imp}\top}$$

Let Z_k^{diff} , and Z_k^{imp} be the orthonormal matrices such that

$$\begin{aligned} \text{im } Z_k^{\text{diff}} &= \text{im} \left(O_k^{\text{diff}\top} \right), & z_k^{\text{diff}} &:= Z_k^{\text{diff}\top} \Pi_k e(t_k^-) \\ & & &= Z_k^{\text{diff}\top} e(t_k^+), \\ \text{im } Z_k^{\text{imp}} &= \text{im} \left(O_k^{\text{imp}\top} \right), & z_k^{\text{imp}} &:= Z_k^{\text{imp}\top} e(t_k^-). \end{aligned}$$

The motivation for introducing the components z_k^{diff} and z_k^{imp} is that they can be estimated using the output measurements on the interval $[t_k, t_{k+1})$. To express the vector z_k in terms of these components, we introduce the matrix U_k^{obs} such that

$$Z_k = \begin{bmatrix} \Pi_k^\top Z_k^{\text{diff}} & Z_k^{\text{imp}} \end{bmatrix} U_k^{\text{obs}}.$$

Such a matrix U_k^{obs} always exists because

$$\begin{aligned} \text{im } Z_k &= \mathcal{W}_k^\perp = (\Pi_k^{-1}(\ker O_k^{\text{diff}}))^\perp + (\ker O_k^{\text{imp}})^\perp \\ &= \Pi_k^\top \text{im } Z_k^{\text{diff}} + \text{im } Z_k^{\text{imp}} \\ &= \text{im} \begin{bmatrix} \Pi_k^\top Z_k^{\text{diff}} & Z_k^{\text{imp}} \end{bmatrix}. \end{aligned}$$

It then follows that

$$\begin{aligned} z_k &= Z_k^\top e(t_k^-) = U_k^{\text{obs}\top} \begin{bmatrix} Z_k^{\text{diff}\top} \Pi_k \\ Z_k^{\text{imp}\top} \end{bmatrix} e(t_k^-) \\ &= U_k^{\text{obs}\top} \begin{bmatrix} Z_k^{\text{diff}\top} e(t_k^+) \\ Z_k^{\text{imp}\top} e(t_k^-) \end{bmatrix} = U_k^{\text{obs}\top} \begin{bmatrix} z_k^{\text{diff}} \\ z_k^{\text{imp}} \end{bmatrix}. \end{aligned}$$

If only estimates $\widehat{z}_k^{\text{diff}}$ and $\widehat{z}_k^{\text{imp}}$ of z_k^{diff} and z_k^{imp} are available, we therefore obtain an estimate of z_k as follows:

$$\widehat{z}_k = U_k^{\text{obs}\top} \begin{bmatrix} \widehat{z}_k^{\text{diff}} \\ \widehat{z}_k^{\text{imp}} \end{bmatrix}. \quad (\text{B.2})$$

Next, we specify how to write z_k^{diff} and z_k^{imp} in terms of the output measured over the interval $[t_k, t_{k+1})$.

Mapping for the differentiable part z_k^{diff} : In order to define $z_k^{\text{diff}} \in \mathbb{R}^{r_k}$, where $r_k = \text{rk } O_k^{\text{diff}}$, we first introduce the function $\mathbf{z}_k^{\text{diff}} : (t_k, t_{k+1}) \rightarrow \mathbb{R}^{r_k}, t \mapsto Z_k^{\text{diff}\top} e(t)$, which represents the observable component of the subsystem (E_k, A_k, C_k) that can be recovered from the smooth output measurements y^e over the interval (t_k, t_{k+1}) . It follows (cf. (Tanwani and Trenn, 2017a, Lem. 17)) that the evolution of $\mathbf{z}_k^{\text{diff}}$ is governed by an observable ODE

$$\begin{aligned} \dot{\mathbf{z}}_k^{\text{diff}} &= S_k^{\text{diff}} \mathbf{z}_k^{\text{diff}}, \\ y^e &= R_k^{\text{diff}} \mathbf{z}_k^{\text{diff}}, \end{aligned} \quad (\text{B.3})$$

where $S_k^{\text{diff}} := Z_k^{\text{diff}\top} A_k^{\text{diff}} Z_k^{\text{diff}}$ and $R_k^{\text{diff}} := C_k^{\text{diff}} Z_k^{\text{diff}}$. Because of the observability of the pair $(S_k^{\text{diff}}, R_k^{\text{diff}})$ in (B.3), there exists a (linear) operator $\mathfrak{D}_{(t_k, t_{k+1})}^{\text{diff}}$ such that

$$\mathbf{z}_k^{\text{diff}} = \mathfrak{D}_{(t_k, t_{k+1})}^{\text{diff}} (y_{(t_k, t_{k+1})}^e)$$

and we set

$$z_k^{\text{diff}} = \mathbf{z}_k^{\text{diff}}(t_k^+).$$

Note that for an estimation $\widehat{\mathbf{z}}_k^{\text{diff}}$ of $\mathbf{z}_k^{\text{diff}}$ obtained by a standard observer (e.g. the Luenberger observer) the evaluation at the beginning of the observation interval is not meaningful (because this value is not affected by the output injection). However, a good estimate can easily be obtained by propagating back the final estimate with the known homogeneous system dynamics; i.e.

$$\widehat{z}_k^{\text{diff}} = e^{-S_k^{\text{diff}} \tau_k} \widehat{\mathbf{z}}_k^{\text{diff}}(t_{k+1}^-). \quad (\text{B.4})$$

Mapping for the impulsive part z_k^{imp} : The impulsive part of the output at switching time t_k can be represented as

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

where due to Lemma 19 the coefficients η_k^j satisfy the relation $\boldsymbol{\eta}_k = -O_k^{\text{imp}} e(t_k^-)$, with $\boldsymbol{\eta}_k := (\eta_k^0 / \dots / \eta_k^{n-2}) \in \mathbb{R}^{(n-1)y}$. We chose a matrix U_k^{imp} such that

$$-O_k^{\text{imp}\top} U_k^{\text{imp}} = Z_k^{\text{imp}}, \quad (\text{B.5})$$

then

$$z_k^{\text{imp}} = Z_k^{\text{imp}\top} e(t_k^-) = -U_k^{\text{imp}\top} O_k^{\text{imp}} e(t_k^-) = U_k^{\text{imp}\top} \boldsymbol{\eta}_k. \quad (\text{B.6})$$

B.2. Observable component over an interval

For $q-1 \geq k \geq p$, the $[t_k, t_q)$ -unobservable subspace (11) can be computed recursively as follows

$$\mathcal{N}_{q-1}^q = \mathcal{W}_{q-1} \quad (\text{B.7a})$$

$$\mathcal{N}_k^q = \mathcal{W}_k \cap \Pi_k^{-1} e^{-A_k^{\text{diff}} \tau_k} \mathcal{N}_{k+1}^q, \quad k \leq q-2. \quad (\text{B.7b})$$

The objective is to compute the observable part $\mu_k = M_k^{q\top} e(t_k^-)$ in (12) recursively for $k = q-1, q-2, \dots, p$. We choose $\mu_{q-1} = z_{q-1}$. By construction, we know that

$$\begin{aligned} \text{im } M_k^q &= \mathcal{M}_k^q = (\mathcal{N}_k^q)^\perp = \left(\mathcal{W}_k \cap \Pi_k^{-1} (e^{-A_k^{\text{diff}} \tau_k} \mathcal{N}_{k+1}^q) \right)^\perp \\ &= \mathcal{W}_k^\perp + \Pi_k^\top (e^{-A_k^{\text{diff}} \tau_k} \mathcal{N}_{k+1}^q)^\perp, \quad k \leq q-2. \end{aligned}$$

Recalling that $\text{im } Z_k = (\mathcal{W}_k)^\perp$, and introducing the matrix Θ_k^q , for $k = p, p+1, \dots, q-2$, such that

$$\text{im } \Theta_k^q = (e^{-A_k^{\text{diff}} \tau_k} \mathcal{N}_{k+1}^q)^\perp \quad (\text{B.8})$$

we obtain

$$\text{im } M_k^q = \text{im} [Z_k, \Pi_k^\top \Theta_k^q]. \quad (\text{B.9})$$

Hence there exists a matrix U_k^q such that

$$M_k^q = [Z_k, \Pi_k^\top \Theta_k^q] U_k^q. \quad (\text{B.10})$$

Noting that

$$\begin{aligned} \Pi_k e(t_k^-) &= e(t_k^+) = \mathbf{e}^{-A_k^{\text{diff}} \tau_k} e(t_{k+1}^-) \\ &= \mathbf{e}^{-A_k^{\text{diff}} \tau_k} (M_{k+1}^q \mu_{k+1} + N_{k+1}^q \nu_{k+1}) \end{aligned}$$

and multiplication on both sides from left by $\Theta_k^{q\top}$ gives

$$\begin{aligned} \Theta_k^{q\top} \Pi_k e(t_k^-) &= \Theta_k^{q\top} \mathbf{e}^{-A_k^{\text{diff}} \tau_k} M_{k+1}^q \mu_{k+1} \\ &\quad + \underbrace{\Theta_k^{q\top} \mathbf{e}^{-A_k^{\text{diff}} \tau_k} N_{k+1}^q}_{=0} \nu_{k+1}. \end{aligned}$$

This allows us to compute μ_k , $k = q - 2, q - 3, \dots, p$, as follows:

$$\begin{aligned} \mu_k &= M_k^{q\top} e(t_k^-) = U_k^{q\top} \begin{bmatrix} Z_k^\top \\ \Theta_k^{q\top} \Pi_k \end{bmatrix} e(t_k^-) \\ &= U_k^{q\top} \left(\Theta_k^{q\top} \mathbf{e}^{-A_k^{\text{diff}} \tau_k} M_{k+1}^q \mu_{k+1} \right) =: \mathcal{F}_k^q \begin{pmatrix} z_k \\ \mu_{k+1} \end{pmatrix}. \quad (\text{B.11}) \end{aligned}$$

References

- Astolfi, A., Praly, L., 2003. Global complete observability and output-to-state stability imply the existence of a globally convergent observer. In: Proc. 42nd IEEE Conf. Decis. Control, Hawaii, USA. pp. 1562–1567.
- Berger, T., Ilchmann, A., Trenn, S., 2012. The quasi-Weierstraß form for regular matrix pencils. *Linear Algebra Appl.* 436 (10), 4052–4069.
- De Santis, E., Di Benedetto, M. D., Pola, G., 2009. A structural approach to detectability for a class of hybrid systems. *Automatica* 45 (5), 1202–1206.
- Liberzon, D., Trenn, S., December 2009. On stability of linear switched differential algebraic equations. In: Proc. IEEE 48th Conf. on Decision and Control. pp. 2156–2161.
- Liberzon, D., Trenn, S., May 2012. Switched nonlinear differential algebraic equations: Solution theory, Lyapunov functions, and stability. *Automatica* 48 (5), 954–963.
- Muller, M., Liberzon, D., 2012. Input/output-to-state stability and state-norm estimators for switched nonlinear systems. *Automatica* 48 (9), 2029–2039.
- Schwartz, L., 1950, 1951. *Théorie des Distributions I,II*. No. IX,X in Publications de l’institut de mathématique de l’Université de Strasbourg. Hermann, Paris.
- Shim, H., Tanwani, A., Ping, Z., 2012. Back-and-forth operation of state observers and norm estimation of estimation error. In: Proc. 51st IEEE Conf. Decis. Control, Maui, USA. pp. 3221 – 3226.
- Sontag, E., Wang, Y., 1997. Output-to-state stability and detectability of nonlinear systems. *Syst. Control Lett.* 29 (5), 279–290.
- Sontag, E. D., 1998. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, 2nd Edition. Springer-Verlag, New York.
- 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., 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.
- Tanwani, A., Trenn, S., 2017a. Determinability and state estimation for switched differential-algebraic equations. *Automatica* 76, 17–31.
- Tanwani, A., Trenn, S., 2017b. Observer design for detectable switched differential-algebraic equations. In: Proceedings of the 20th IFAC World Congress. pp. 2953–2958.
- 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>
- Wong, K.-T., 1974. The eigenvalue problem $\lambda T x + S x$. *J. Diff. Eqns.* 16, 270–280.