An Observer for Switched Differential-Algebraic Equations Based on Geometric Characterization of Observability

Aneel Tanwani and Stephan Trenn

***Abstract*— Based on our previous work dealing with geometric characterization of observability for switched differential-algebraic equations (switched DAEs), we propose an observer design for switched DAEs that generates an asymptotically convergent state estimate. Without assuming the observability of individual modes, the central idea in constructing the observer is to filter out the maximal information from the output of each of the active subsystems and combine it with the previously extracted information to obtain a good estimate of the state after a certain time has passed. In general, observability only holds when impulses in the output are taken into account, hence our observer incorporates the knowledge of impulses in the output. This is a distinguishing feature of our observer design compared to observers for switched ordinary differential equations.**

I. Introduction

In this paper, we propose an observer for a class of switched systems where the dynamical subsystems are modeled as *differential-algebraic equations* (DAEs):

$$\begin{aligned} E_p\dot{x} &= A_p x + B_p u, \quad \text{over} \quad [t_{p-1}, t_p) \\ y(t) &= C_p x(t), \quad t \in [t_{p-1}, t_p), \end{aligned} \tag{1}$$

where $x : \mathbb{R} \to \mathbb{R}^n$, $u : \mathbb{R} \to \mathbb{R}^{d_u}$, and $y : \mathbb{R} \to \mathbb{R}^{d_y}$ denote the state, input, and output respectively; and $E_p, A_p \in \mathbb{R}^{n\times n}$, $B_p \in \mathbb{R}^{n\times d_u}$, $C_p \in \mathbb{R}^{d_y\times n}$, for $p \in \mathbb{N}$. The description (1) is not a restriction of generality as we do, of course, allow $(E_p, A_p, B_p, C_p) = (E_q, A_q, B_q, C_q)$ for $p \neq q$. In order to use the piecewise-smooth distributional solution framework and to avoid technical difficulties in general, it is assumed that the inputs u are piecewise smooth and that there is no accumulation of switching times. The forthcoming observer design will not rely on observability of individual subsystems in the classical sense, however we will assume observability conditions in line of our recent observability characterization for switched DAEs [15]. In particular, the knowledge of the switching times and the active mode is assumed.

The main motivation for studying this problem is of theoretical nature, however switched DAEs (1) occur naturally when modeling e.g. electrical circuits with switches or sudden component faults. Observers are necessary to monitor the inner states of a large system where only some external signals are available. A possible future application might be the use of observers in electrical grids to monitor the energy flows through the transmission lines and prevent overloading.

Observability and observer design are classical problems in systems theory and the earliest solution of these problems for linear time invariant systems date back to 1960's. Since then the problem has been well studied for different kinds of dynamical systems. In the context of (nonswitched) DAEs, observer design methods were already studied in 1980's, e.g. [5], [7]. In contrast to ordinary differential equations (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 [4], [6].

During the past decade, however, the focus has shifted towards the study of observability and observer design for nonsmooth dynamical systems since they generalize a large number of physical and digitally-interfaced models, e.g. [2], [13], [19]. Out of several existing formalisms for modeling nonsmooth behaviors, switched systems form an important subclass which comprise a family of subsystems and a switching rule that determines the active subsystem [8]. The presence of a switching signal brings an extra dimension to the problem of observability for such systems. Observability and observer design for switched linear ODEs with unknown switching signal (or discrete state) were studied by [1], [19]. Assuming that the individual subsystems are observable, algorithms are proposed for computing the continuous as well as the discrete state. However, if the switching signal is known, then without requiring the observability of individual subsystems, the results on recovering the continuous state appear in [10], [21]; but the observer construction remains unaddressed. Based on the latter viewpoint, a unified approach towards observability and observers in a more general framework is studied in the recent papers [9], [11], [12]. In contrast to the classical approach, observers with state jumps have been employed in [12] to compensate for the lack of complete information about the state at each time instant.

The idea of our observer design for switched DAEs is heavily influenced by the approach in [11], however there are two major differences: 1) Switched DAEs exhibit jumps in the state given by non-invertible jump maps (the approach in [11] is only valid for invertible jump maps) and 2) Switched DAEs might even produce Dirac impulses in the output and the information from these Dirac impulses is in general necessary for observability; hence the observer must take the presence of Dirac impulses into account.

II. Preliminaries

A. Properties and Definitions for Regular Matrix Pairs

In the following, we collect important properties and definitions for matrix pairs (E, A). We only consider *regular*

Aneel Tanwani is with the INRIA research centre, Grenoble, France, email: aneel.tanwani@inria.fr

Stephan Trenn is with the Technomathematics group, University of Kaiserslautern, Germany, email: trenn@mathematik.uni-kl.de

matrix pairs, i.e. for which the polynomial $\det(sE-A)$ is not the zero polynomial. A very useful characterization of regularity is the following well-known result.

Proposition 1 (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

$$(SET, SAT) = \left(\left[\begin{smallmatrix} I & 0 \\ 0 & N \end{smallmatrix}\right], \left[\begin{smallmatrix} J & 0 \\ 0 & I \end{smallmatrix}\right]\right), \tag{2}$$

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

In view of [3], we call the decomposition (2) *quasi-Weierstrass form*. An easy way to calculate the transformation matrices S and T for (2) is to use the following so-called *Wong sequences* [20], [3]:

$$\begin{aligned} \mathcal{V}_0 &:= \mathbb{R}^n, & \mathcal{V}_{i+1} &:= A^{-1}(E\mathcal{V}_i), & i &= 0,1,\cdots \\ \mathcal{W}_0 &:= \{0\}, & \mathcal{W}_{i+1} &:= E^{-1}(A\mathcal{W}_i), & i &= 0,1,\cdots \end{aligned}$$

The Wong sequences are nested and become stationary after finitely many steps. The limiting subspaces are defined as follows:

$$\mathcal{V}^* := \bigcap_i \mathcal{V}_i, \quad \mathcal{W}^* := \bigcup_i \mathcal{W}_i.$$

For any full rank matrices V,W with $\operatorname{im} V = \mathcal{V}^*$ and $\operatorname{im} W = \mathcal{W}^*$, the matrices $T := [V,W]$ and $S := [EV, AW]^{-1}$ are invertible and (2) holds.

Based on the Wong-sequences we define the following "projectors".

Definition 2 (Consistency, differential and impulse projectors): Consider the regular matrix pair (E,A) with corresponding quasi-Weierstrass form (2). The *consistency projector* of (E,A) is given by

$$\Pi_{(E,A)} = T\left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right]T^{-1},$$

the *differential and impulse projectors* are given by

$$\Pi^{\text{diff}}_{(E,A)} = T\left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right]S, \quad \Pi^{\text{imp}}_{(E,A)} = T\left[\begin{smallmatrix} 0 & 0 \\ 0 & I \end{smallmatrix}\right]S,$$

where the block sizes correspond to the ones in (2). $\lhd$

Note that only the consistency projector is a projector in the usual sense (i.e. $\Pi_{(E,A)}$ is an idempotent matrix); whereas $\Pi^{\text{diff}}_{(E,A)}$ and $\Pi^{\text{imp}}_{(E,A)}$ are not projectors because, in general, $\Pi^{\text{diff}}_{(E,A)}\Pi^{\text{diff}}_{(E,A)} \neq \Pi^{\text{diff}}_{(E,A)}$ and the same holds for $\Pi^{\text{imp}}_{(E,A)}$. Let

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

be the *consistency space* of the DAE $E\dot{x} = Ax$, where $\mathcal{C}^1$ is the space of differentiable functions $x : \mathbb{R} \to \mathbb{R}^n$. Then the following observations hold [3]:

1) All solutions $x \in \mathcal{C}^1$ of $E\dot{x} = Ax$ evolve within $\mathfrak{C}_{(E,A)}$,
2) $\mathfrak{C}_{(E,A)} = \mathcal{V}^*$, i.e. the first Wong-sequence converges to the consistency space,
3) $\operatorname{im}\Pi_{(E,A)} = \mathcal{V}^* = \mathfrak{C}_{(E,A)}$, hence the consistency projector maps onto the consistency space.

For understanding the role of the consistency projector and for studying impulsive solutions, we consider the space of *piecewise-smooth distributions* $\mathbb{D}_{\text{pw}\mathcal{C}^\infty}$ from [17] as the solution space; that is, we seek a solution $x \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$ to the following initial-trajectory problem (ITP):

$$\begin{aligned} x_{(-\infty,0)} &= x^0_{(-\infty,0)} \\ (E\dot{x})_{[0,\infty)} &= (Ax+Bu)_{[0,\infty)}, \end{aligned} \tag{3}$$

where $x^0 \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$ is some initial trajectory, and $f_{\mathcal{I}}$ denotes the restriction of a piecewise-smooth distribution f to an interval $\mathcal{I}$. In [16], [17] it is shown that the ITP (3) has a unique solution for any initial trajectory if, and only if, the matrix pair (E,A) is regular. In particular, the following result concerning the consistency projector holds.

Lemma 3 (Role of consistency projector, [16, Thm. 4.2.8]): Consider the ITP homogenous (3) (i.e. $Bu = 0$) with regular matrix pair (E,A) and with arbitrary initial trajectory $x^0 \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$. Let $\Pi_{(E,A)}$ be the consistency projector of (E,A), then the unique solution $x \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$ satisfies

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

Note furthermore, that it can be shown that the unique solution of (3) restricted to $[0,\infty)$ only depends on $x^0(0^-)$, hence in the following we can replace (3) by

$$(E\dot{x})_{[0,\infty)} = (Ax+Bu)_{[0,\infty)}, \quad x(0^-) = x_0 \in \mathbb{R}^n. \tag{4}$$

The following lemma motivates the name of the differential projector.

Lemma 4 ([14, Lem. 3]): Consider the DAE $E\dot{x} = Ax$ with regular matrix pair (E,A). Then any solution $x \in \mathcal{C}^1$ of $E\dot{x} = Ax$ fulfills

$$\dot{x} = \Pi^{\text{diff}}_{(E,A)}Ax =: A^{\text{diff}}x. \quad \lhd$$

Finally, the role of the impulsive projector becomes clear when expressing the impulsive part, denoted by $x[0]$, of the distributional solution x of the ITP (3).

Lemma 5 ([14, Cor. 5]): Consider the ITP (3) with regular matrix pair (E,A) and corresponding impulse and consistency projectors $\Pi^{\text{imp}}_{(E,A)}$, $\Pi_{(E,A)}$. Let $E^{\text{imp}} := \Pi^{\text{imp}}_{(E,A)}E$ then, for the unique solution $x \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$,

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

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

Remark 6: The actual formula for the impulses given in [14] is $x[0] = -\sum_{i=0}^{n-2}(E^{\text{imp}})^{i+1}(I - \Pi_{(E,A)})\delta_0^{(i)}x(0-)$, however it is easily seen that $E^{\text{imp}}(I - \Pi_{(E,A)}) = E^{\text{imp}}$, where E^{imp} is a nilpotent matrix of index smaller than or equal to $n_2 \le n$. $\lhd$

III. Observability Conditions

Our observer is built on the notion of *determinability* considered in [15, Definition 8] and here we recall some tools that are used in deriving the conditions for determinability and later in observer construction.

Using the notation $[M_1/M_2]{:=}\left[\begin{smallmatrix}M_1\\M_2\end{smallmatrix}\right]$, we define for each $p{>}0$

$$\begin{aligned}\Pi_p &:= \Pi_{(E_p,A_p)}\\ \mathfrak{C}_p &:= \mathfrak{C}_{(E_p,A_p)},\\ O_p^{\text{diff}} &:= [C_p\Pi_p/C_pA_p^{\text{diff}}/\cdots/C_p(A_p^{\text{diff}})^{n-1}],\\ O_p^{\text{imp}} &:= [C_pE_p^{\text{imp}}/C_p(E_p^{\text{imp}})^2/\cdots/C_p(E_p^{\text{imp}})^{n-1}].\end{aligned}$$

In view of Lemma 4, O_p^{diff} is the Kalman observability matrix of the ODE

$$\begin{aligned}\dot{x} &= A_p^{\text{diff}}x\\ y &= C_px = C_p\Pi_px\end{aligned}$$

taking into account that x only evolves within the consistency space (yielding $\Pi_px = x$) as well as $\Pi_pA^{\text{diff}} = A^{\text{diff}}$. Similarly as in [15] we can define the local unobservable space $\mathcal{W}_p$ as follows

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

while taking into account the measurements (u,y) over the interval (t_{p-1},t_p) and the impulsive information $y[t_p]$ only.

The following sequence of subspaces is central for the observer construction:

$$\begin{aligned}\mathcal{Q}_p^p &:= \mathcal{W}_p,\\ \mathcal{Q}_p^{p+k} &:= \mathcal{W}_{p+k} \cap e^{A_{p+k}^{\text{diff}}\tau_{p+k}}\Pi_{p+k}\mathcal{Q}_p^{p+k-1},\ k>0,\end{aligned} \tag{5}$$

where $\tau_p := t_p - t_{p-1}$, for $p \in \mathbb{N}$. The intuition behind this sequence of subspaces is as follows: If we measure (u,y) over the interval (t_{p-1},t_p) and $y[t_p]$, then we can determine $x(t_p^-)$ modulo the subspace $\mathcal{Q}_p^p = \mathcal{W}_p$. Similarly, by measuring (u,y) over the interval (t_{p-1},t_{p+k}) and $y[t_{p+k}]$, we can recover $x(t_{p+k}^-)$ modulo $\mathcal{Q}_p^{p+k}$, for $k \in \mathbb{N}$.

For the observer design the orthogonal complement of the above sequence is also needed, i.e. $\mathcal{P}_p^p := \mathcal{Q}_p^{p\perp} = \mathcal{W}_p^\perp$ and

$$\begin{aligned}\mathcal{P}_p^{p+k} &:= \mathcal{Q}_p^{p+k\perp}\\ &= \mathcal{W}_{p+k}^\perp + \Pi_{p+k}^{-\top}e^{-A_{p+k}^{\text{diff}\top}\tau_{p+k}}\mathcal{P}_p^{p+k-1},\ k>0.\end{aligned} \tag{6}$$

Theorem 7 (Determinability Characterization): Consider the switched DAE (1) with zero input. Then $\mathcal{Q}_q^p$ for some $p \geq q \geq 1$ characterizes the unobservable space in the following sense:

$$y_{(t_{q-1},t_p)} \equiv 0 \wedge y[t_p] = 0 \quad \Leftrightarrow \quad x(t_p^-) \in \mathcal{Q}_q^p.$$

In particular, if there exists $p \geq q$ such that $\mathcal{Q}_q^p = \{0\}$ the state $x(t_p^-)$ (and hence the complete future trajectory) can be determined from the knowledge of (u,y) over the interval (t_{q-1},t_p) and $y[t_p]$.

The proof uses the same arguments as the proof of [15, Thm. 15] and is therefore omitted.

IV. Observer Design

Assumption 8: The following assumptions are imposed on the system data for our proposed observer design:

1) Each switching interval has a finite maximum length; that is, there exist $D>0$ such that
$$t_{p+1} - t_p < D, \qquad \forall\, p \in \mathbb{N}. \tag{7}$$
2) The system is persistently determinable in the sense that there exists an $N \in \mathbb{N}$ such that, $\forall\, p > N$,
$$\dim \mathcal{Q}_{p-N}^p = 0 \ (\Leftrightarrow\ \dim \mathcal{P}_{p-N}^p = n). \tag{8}$$
(The integer N is interpreted as the minimal number of switches required to gain determinability.)
3) The induced matrix norms $\|A_p^{\text{diff}}\|$ and $\|\Pi_p\|$ are uniformly bounded for all $p \in \mathbb{N}$ (which is always the case when A_p and Π_p belong to a set of finite elements).

We propose the following observer for the switched DAE (1):

$$\boxed{\begin{aligned}E_p\dot{\hat{x}}_p &= A_p\hat{x}_p + B_pu, \text{ on } [t_{p-1},t_p),\\ \hat{x}_p(t_{p-1}^-) &= \hat{x}_{p-1}(t_{p-1}^-) - \xi_{p-1}\end{aligned}} \qquad \begin{aligned}&(9a)\\ &(9b)\end{aligned}$$

where the initial condition $\hat{x}_0(t_0^-) \in \mathbb{R}^n$ is arbitrarily chosen and the overall estimation is $\hat{x} := \sum_{p\in\mathbb{N}}(\hat{x}_p)_{[t_{p-1},t_p)}$. The error correction vector ξ_p is defined as:

$$\xi_p = \begin{cases}\mathcal{L}_p(y_{(t_{p-N-1},t_p]}, u_{(t_{p-N-1},t_p]}),\ p>N,\\ 0, \qquad\qquad\qquad\qquad\qquad 0\leq p\leq N,\end{cases}$$

where the operator $\mathcal{L}_p$ will be designed in the sequel.

Remark 9: For nonswitched systems, Assumption 8.1 is not a restriction since one can always add "dummy" switches (with similar system matrices) to satisfy (7). However, note that there is no continuous error injection term in observer (9) and the estimate is updated only at the switching instants. That also explains the necessity of (7) because we need to update the estimate repeatedly to compensate for the error propagating between two switches. ◁

Let $\tilde{x}_p := \hat{x}_p - x$ denote the state estimation error on $[t_{p-1},t_p)$ and $\tilde{x} := \sum_p(\tilde{x}_p)_{[t_{p-1},t_p)} = \hat{x} - x$, then

$$E_p\dot{\tilde{x}}_p = A_p\tilde{x}_p, \quad \text{ on } [t_{p-1},t_p), \tag{10a}$$
$$\tilde{x}_p(t_{p-1}^-) = \tilde{x}_{p-1}(t_{p-1}^-) - \xi_{p-1}. \tag{10b}$$

Note that equations (9a) and (10a) are both to be interpreted in the sense of distributions. However, the error dynamics (10a) are homogenous and there are no impulses between two switches. As a result, the solution of (10a) on the open interval (t_{p-1},t_p) is given by $\tilde{x}_p(t) = e^{A_p^{\text{diff}}(t-t_{p-1})}\Pi_p\tilde{x}_p(t_{p-1}^-)$.

Let the output estimation error be $\tilde{y} = C_p\hat{x}_p - y$ on each open interval (t_{p-1},t_p). The impulsive error $\tilde{y}[t_p]$ at the switching times is obtained by the difference between $y[t_p]$ and the output impulse resulting from (9a) *without* taking the correction ξ_p into account i.e. (invoking Lemma 5):

$$\begin{aligned}\tilde{y}[t_p] &:= -\sum_{i=0}^{n-2} C_{p+1}(E_{p+1}^{\text{imp}})^{i+1}\hat{x}_p(t_p^-)\delta_{t_p}^{(i)} - y[t_p]\\ &= -\sum_{i=0}^{n-2} C_{p+1}(E_{p+1}^{\text{imp}})^{i+1}\tilde{x}(t_p^-)\delta_{t_p}^{(i)},\end{aligned} \tag{11}$$

in particular, we have $\tilde{y}[t_p] = C_{p+1}\tilde{x}[t_p] = -\sum_{i=0}^{n-2} C_{p+1}(E_{p+1}^{\text{imp}})^{i+1}(\tilde{x}(t_p^-) - \xi_p)\delta_{t_p}^{(i)}$. Note that we need to be able to measure the impulsive part of y at t_p which depends on u and its derivates immediately after time t_p (c.f. [18, Thm. 6.5.1]). This will render the observer slightly acausal, as the information immediately after t_p

is used to estimate $\tilde{x}_p(t_p^-)$. However, this is not a serious problem from an implementation-point-of-view as (10a) does not depend on u.

In the remainder of this section, we develop a machinery to compute ξ_p. It is noted that ξ_p approximates the value of $\tilde{x}(t_p^-)$, and thus the basic idea in deriving an expression for ξ_p is to first write $\tilde{x}(t_p^-)$ in terms of the *known* quantities.

A. Local estimation around a switch

For each $p \in \mathbb{N}$, we are interested in decomposing the (unknown) error vector $\tilde{x}(t_p^-)$ along $\mathcal{W}_p$ and $\mathcal{W}_p^\perp$. For that, let us introduce the orthonormal matrices W_p and Z_p such that $\mathcal{R}(W_p) = \mathcal{W}_p$ and $\mathcal{R}(Z_p) = \mathcal{W}_p^\perp$, where $\mathcal{R}(M)$ denotes the range space of the columns of a matrix M. It then follows that $[Z_p, W_p]^{-1} = [Z_p, W_p]^\top$. Now define, $z_p := Z_p^\top \tilde{x}(t_p^-)$ and $w_p := W_p^\top \tilde{x}(t_p^-)$. Thus, we have

$$\tilde{x}(t_p^-) = Z_p z_p + W_p w_p. \tag{12}$$

Note that z_p denotes the component of the error vector $\tilde{x}(t_p^-)$ that can be recovered from measuring $(u_{(t_{p-1},t_p)}, y_{(t_{p-1},t_p)})$ and $y[t_p]$. Hence, we are interested in obtaining a good estimate of z_p. Since $\mathcal{W}_p^\perp = \left(\mathfrak{C}_p \cap \ker O_p^{\text{diff}} \cap \ker O_{p+1}^{\text{imp}}\right)^\perp = \mathfrak{C}_p^\perp + \mathcal{R}\left({O_p^{\text{diff}}}^\top\right) + \mathcal{R}\left({O_{p+1}^{\text{imp}}}^\top\right)$ is a sum of three subspaces and z_p is the projection of $\tilde{x}(t_p^-)$ along the subspace $\mathcal{W}_p^\perp$, we further decompose the vector z_p along each of the three constituent subspaces. Towards this end, let Z_p^{cons}, Z_p^{diff}, Z_p^{imp} be the matrices whose columns form an orthonormal basis of the subspaces $\mathfrak{C}_p^\perp$, $\mathcal{R}\left({O_p^{\text{diff}}}^\top\right)$, and $\mathcal{R}\left({O_{p+1}^{\text{imp}}}^\top\right)$, respectively. Define $z_p^{\text{cons}} := {Z_p^{\text{cons}}}^\top \tilde{x}(t_p^-)$, $z_p^{\text{diff}}(\cdot) := {Z_p^{\text{diff}}}^\top \tilde{x}(\cdot)$, $z_p^{\text{imp}} := {Z_p^{\text{imp}}}^\top \tilde{x}(t_p^-)$. Note that $[Z_p^{\text{cons}}, Z_p^{\text{diff}}]$ has full column rank, however the image of Z_p^{imp} might non-trivially intersect with the image of $[Z_p^{\text{cons}}, Z_p^{\text{diff}}]$. In this case, some part of the unknown error $\tilde{x}(t_p^-)$ 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_p such that

$$[Z_p^{\text{cons}},\ Z_p^{\text{diff}},\ Z_p^{\text{imp}}]\, U_p = Z_p, \tag{13}$$

and for brevity we let $\overline{Z_p} := [Z_p^{\text{cons}},\ Z_p^{\text{diff}},\ Z_p^{\text{imp}}]$.

We thus obtain,

$$z_p = Z_p^\top \tilde{x} = U_p^\top \left(z_p^{\text{cons}} / z_p^{\text{diff}}(t_p^-) / z_p^{\text{imp}}\right). \tag{14}$$

a) The consistency information z_p^{cons}: In the above expression, $z_p^{\text{cons}} = {Z_p^{\text{cons}}}^\top \tilde{x}(t_p^-) = 0$ because any solution of the homogenous DAE (10a) evolves within the consistency space $\mathfrak{C}_p$ and ${Z_p^{\text{cons}}}^\top \mathfrak{C} = \{0\}$ by definition.

b) Recover the differentiable part $z_p^{\text{diff}}(\cdot)$: The observable part $z_p^{\text{diff}}(\cdot)$ can in theory be determined exactly from the output error $\tilde{y}(\cdot)$ on the interval (t_{p-1}, t_p). However, in practice the values of z_p^{diff} will be approximated by a standard Luenberger observer based on the Kalman decomposition of $(A_p^{\text{diff}}, C_p \Pi_p)$. In fact, choose matrices $S_p \in \mathbb{R}^{r_p \times r_p}$ and $R_p \in \mathbb{R}^{d_y \times r_p}$, where $r_p = \operatorname{rank} O_p^{\text{diff}}$, such that ${Z_p^{\text{diff}}}^\top A_p^{\text{diff}} = S_p {Z_p^{\text{diff}}}^\top$ and $C_p \Pi_p = R_p {Z_p^{\text{diff}}}^\top$. Then (S_p, R_p) is an observable pair in the classical sense. For the interval (t_{p-1}, t_p), the use of Lemma 4 yields

$$\begin{aligned} \dot{z}_p^{\text{diff}} &= {Z_p^{\text{diff}}}^\top A_p^{\text{diff}} \tilde{x} = S_p z_p^{\text{diff}}, \\ \tilde{y} &= C_p \Pi_p \tilde{x} = R_p z_p^{\text{diff}}. \end{aligned} \tag{15}$$

Since z_p^{diff} is observable over the interval (t_{p-1}, t_p), a standard Luenberger observer is designed as

$$\dot{\hat{z}}_p^{\text{diff}} = S_p \hat{z}_p^{\text{diff}} + L_p(\tilde{y} - R_p \hat{z}_p^{\text{diff}}), \quad t \in (t_{p-1}, t_p), \tag{16a}$$

$$\hat{z}_p^{\text{diff}}(t_{p-1}^+) = 0, \tag{16b}$$

whose role is to estimate z_p^{diff} especially at the end of the interval. In our forthcoming main result we will have to assume that L_p is chosen such that the difference $\hat{z}_p^{\text{diff}}(t_p^-) - z_p^{\text{diff}}(t_p^-)$ is sufficiently small.

c) Recover the impulsive part z_p^{imp}: When comparing the observed impulses in the output y at t_p with the impulses predicted by the system copy (9) via the formula (11), then it is possible to recover a certain part of the error $\tilde{x}(t_p^-)$. In fact, let

$$\tilde{y}[t_p] = \sum_{i=0}^{n-2} \eta_p^i \delta_{t_p}^{(i)},$$

then (11) implies that for $\eta_p = (\eta_p^0 / \cdots / \eta_p^{n-2})$, we have the relation $\eta_p = O_{p+1}^{\text{imp}} \tilde{x}(t_p^-)$. If U_p^{imp} is a matrix such that ${O_{p+1}^{\text{imp}}}^\top U_p^{\text{imp}} = Z_p^{\text{imp}}$, then

$${U_p^{\text{imp}}}^\top \eta_p = {U_p^{\text{imp}}}^\top O_{p+1}^{\text{imp}} \tilde{x}(t_p^-) = {Z_p^{\text{imp}}}^\top \tilde{x}(t_p^-) = z_p^{\text{imp}}.$$

Altogether, we now let $\hat{z}_p$ be defined as follows:

$$\hat{z}_p = U_p^\top \left(0 / \hat{z}_p^{\text{diff}}(t_p^-) / {U_p^{\text{imp}}}^\top \eta_p\right). \tag{17}$$

B. Merging the local information

For $p, q \in \mathbb{N}$ with $p \geq q$ let P_q^p and Q_q^p be matrices such that its columns are an orthonormal basis of $\mathcal{P}_q^p$ and $\mathcal{Q}_q^p$, respectively. The corresponding projections of $\tilde{x}(t_p^-)$ onto these subspaces are defined by letting $\varphi_q^p := {P_q^p}^\top \tilde{x}(t_p^-)$ and $\chi_q^p := {Q_q^p}^\top \tilde{x}(t_p^-)$. Thus, it is seen that in addition to (12), another way of expressing $\tilde{x}(t_p^-)$ is:

$$\tilde{x}(t_p^-) = P_q^p \varphi_q^p + Q_q^p \chi_p^q. \tag{18}$$

Furthermore, let Θ_q^p be a matrix whose columns form the basis of the subspace $\mathcal{R}(e^{A_{p+1}^{\text{diff}} \tau_{p+1}} \Pi_{p+1} Q_q^p)^\perp$; that is,

$${\Theta_q^p}^\top e^{A_{p+1}^{\text{diff}} \tau_{p+1}} \Pi_{p+1} Q_q^p = 0.$$

The key idea of the observer design is to combine the observable information φ_q^{p-1}, $p > q$, for $\tilde{x}(t_{p-1}^-)$ obtained from $(u_{(t_{q-1},t_{p-1})}, y_{(t_{q-1},t_{p-1})})$ and $y[t_{p-1}]$ with the locally observable information z_p for $\tilde{x}(t_p^-)$ obtained from $(u_{(t_{p-1},t_p)}, y_{(t_{p-1},t_p)})$ and $y[t_p]$ to accumulate more information φ_q^p about $\tilde{x}(t_p^-)$. For that, the following relationship between $\tilde{x}(t_p^-)$ and φ_q^{p-1}, $q < p$, is crucial:

$$\begin{aligned} \tilde{x}(t_p^-) &= e^{A_p^{\text{diff}} \tau_p} \Pi_p (\tilde{x}(t_{p-1}^-) - \xi_{p-1}) \\ &= e^{A_p^{\text{diff}} \tau_p} \Pi_p \left(P_q^{p-1} \varphi_q^{p-1} + Q_q^{p-1} \chi_q^{p-1} - \xi_{p-1}\right). \end{aligned} \tag{19}$$

Combining this with (12) we obtain

$$\begin{bmatrix} Z_p^\top \\ \Theta_q^{p-1\top} \end{bmatrix} \tilde{x}(t_p^-) = \begin{pmatrix} z_p \\ \Theta_q^{p-1\top} e^{A_p^{\text{diff}}\tau_p} \Pi_p \left(P_q^{p-1}\varphi_q^{p-1} - \xi_{p-1}\right) \end{pmatrix},$$

hence we can obtain more information about $\tilde{x}(t_p^-)$ by combining z_p and φ_q^{p-1} accordingly. In fact, from $\varphi_q^p = P_q^{p\top}\tilde{x}(t_p^-)$ it now follows that

$$\varphi_q^p = U_q^{p\top} \begin{bmatrix} Z_p^\top \\ \Theta_q^{p-1\top} \end{bmatrix} \tilde{x}(t_p^-) \tag{20}$$

$$= U_q^{p\top} \begin{pmatrix} z_p \\ \Theta_q^{p-1\top}\left(e^{A_p^{\text{diff}}\tau_p} \Pi_p \left(P_q^{p-1}\varphi_q^{p-1} - \xi_{p-1}\right)\right) \end{pmatrix}, \tag{21}$$

where U_q^p is a full column rank matrix such that

$$[Z_p, \Theta_q^{p-1}] U_q^p = P_q^p.$$

This matrix always exists because from the definition of $\mathcal{P}_q^p$ and Z_p it follows that

$$\mathcal{R}(P_q^p) = \mathcal{R}([Z_p, \Theta_q^{p-1}]),$$

Note that (20) expresses the vector φ_q^p recursively in terms of φ_q^{p-1}. Recall that $\mathcal{P}_{p-N}^{p-N} = \mathcal{W}_{p-N}{}^\perp = \mathcal{R}(Z_{p-N})$, hence we can assume $P_{p-N}^{p-N} = Z_{p-N}$ and we have the "initial value" for the recursion (20) given by $\varphi_{p-N}^{p-N} = z_{p-N}$.

If we know $z_p, z_{p-1}, \ldots, z_{p-N}$ exactly then the above recursion formula would allow us to reconstruct $\tilde{x}(t_p)$ after N steps and we would choose $\xi_p = P_{p-N}^p \varphi_{p-N}^p = \tilde{x}(t_p^-)$. The error dynamics (10) would then jump to zero and remain zero after t_p, i.e. our observer would have recovered the state exactly. Since we only know the approximation $\hat{z}_p$ of z_p we can only get an approximation of φ_{p-N}^p and the error dynamics will not jump to zero. That is why the above recursion formula has to be repeated at each switching time, making the error smaller and smaller.

C. Summary of observer design

Altogether we have derived the following algorithm for calculating the jump corrections ξ_p in (9) at the p-th switching time t_p:

1) Calculate the matrices Π_p, A_p^{diff}, E_{p+1}^{imp}, e.g. via the Wong-sequences, and the corresponding local unobservable space $\mathcal{W}_p$.
2) Run the observer (16) on the interval (t_{p-1}, t_p) to obtain $\hat{z}_p^{\text{diff}}$ using the difference between the output $\hat{y}$ of the system copy (9) and the real output y.
3) Measure the impulsive part $y[t_p]$ in the output at time t_p and calculate the approximation $\hat{z}_p$ via (17).
4) For $k = N-1, \ldots, 0$, calculate the matrices P_{p-N}^{p-k} and Θ_{p-N}^{p-k-1}.
5) For $k = N-1, \ldots, 0$, calculate the approximation $\hat{\varphi}_{p-N}^{p-k}$ of φ_{p-N}^{p-k} via the following recursion formula:

$$\boxed{\begin{aligned} \hat{\varphi}_{p-N}^{p-N} &= \hat{z}_{p-N} \\ \hat{\varphi}_{p-N}^{p-k} &= F_{p-N}^{p-k} Z_{p-k} \hat{z}_{p-k} \\ &\quad + G_{p-N}^{p-k}\left(P_{p-N}^{p-k-1}\hat{\varphi}_{p-N}^{p-k-1} - \xi_{p-k-1}\right), \end{aligned}}$$

where

$$\begin{aligned} &[F_{p-N}^{p-k}, G_{p-N}^{p-k}] \\ &\quad := U_{p-N}^{p-k\top} \begin{bmatrix} Z_{p-k}^\top & 0 \\ 0 & \Theta_{q-N}^{p-k-1\top} e^{A_{p-k}^{\text{diff}}\tau_{p-k}} \Pi_{p-k} \end{bmatrix}. \end{aligned}$$

6) If $p > N$ let $\boxed{\xi_p = P_{p-N}^p \hat{\varphi}_{p-N}^p.}$

V. Error Convergence

In order to state the criteria for choosing the gain matrix that guarantees the convergence of the state estimation error to zero, we introduce the matrix

$$\Lambda_p := \text{block diag}\left(0, e^{(S_p - L_p R_p)\tau_p}, 0\right) \tag{22}$$

where the zero blocks correspond to the sizes of z_p^{cons} and z_p^{imp} in (17). Due to the observability of (S_p, R_p) the norm of Λ_p can be made arbitrarily small by choosing L_p accordingly. In order to make precise statements about the "smallness" of Λ_p we need to define the following matrices for $p > N$, $k = N-2, \ldots, 0$ and $i = 0, \ldots, N-k-1$

$$V_{p-N,p-N}^{p-N+1} := G_{p-N}^{p-N+1} \tag{23a}$$

$$V_{p-N,p-N+1}^{p-N+1} := F_{p-N}^{p-N+1} \tag{23b}$$

$$V_{p-N,p-N+i}^{p-k} := G_{p-N}^{p-k} P_{p-N}^{p-k-1} V_{p-N,p-N+i}^{p-k-1} \tag{23c}$$

$$V_{p-N,p-k}^{p-k} := F_{p-N}^{p-k}. \tag{23d}$$

The main result on observer convergence now follows:

Theorem 10: Consider the observer (9) under Assumption 8, with ξ_p, $p > N$, given as in Section IV-C. If, for each $k = 0, \ldots, N$, the output injection matrices L_{p-k} are chosen to make the norm of Λ_{p-k} small enough such that

$$\|P_{p-N}^p V_{p-N,p-k}^p Z_{p-k} U_{p-k}^\top \Lambda_{p-k} \overline{Z_{p-k}}^\top \Pi_{p-k}\| < \frac{1}{N+1},$$

then it holds that $\boxed{\lim_{t\to\infty} \|\hat{x}(t^+) - x(t^+)\| = 0}$ and $\hat{x}[t] - x[t] \to 0$ in the distributional sense as $t \to \infty$.
The proof has been omitted due to space limitations.

VI. Simulations

We illustrate the observer design and its effectiveness with an example. In system (1), for $k \geq 1$, let

$$(E_{2k-1}, A_{2k-1}) =$$

$$\left(\begin{bmatrix} 1 & -2 & 0 & 0 & 0 & 0 & 0 \\ -2 & 5 & 0 & 0 & 0 & 0 & 0 \\ -2 & 2 & 1 & 0 & 0 & 0 & 0 \\ 3 & -5 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0.08 & -0.16 & 0 & 0 & 0 & 0 & 0 \\ -0.2 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ -0.24 & 0.24 & 0.12 & 0 & 0 & 0 & 0 \\ 2 & -4 & 0 & 1 & 0 & 0 & 0 \\ 3 & -5 & -1 & 0 & 1 & 0 & 0 \\ -0.15 & 0.45 & 0 & 0 & 0 & 0.15 & 0.3 \\ 2 & -3 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \right)$$

$$(E_{2k}, A_{2k}) =$$

$$\left(\begin{bmatrix} 1 & -2 & 0 & 0 & 0 & 0 & 0 \\ -2 & 5 & 0 & 0 & 0 & 0 & 0 \\ -2 & 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -3 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & -3 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 5 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 & 0 & 0 \\ -0.18 & 0.18 & 0.09 & 0 & 0 & 0 & 0 \\ 2 & -4 & 0 & 1 & 0 & 0 & 0 \\ 1 & -3 & 0 & 0 & 1 & 0 & 0 \\ 0.4 & -0.6 & 0 & 0 & 0 & 0 & 0.2 \\ -1 & 3 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \right)$$

the matrix $B_k = 0$, and

$$C_{2k-1} = \begin{bmatrix} 3 & -6 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad C_{2k} = \begin{bmatrix} 2 & -4 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

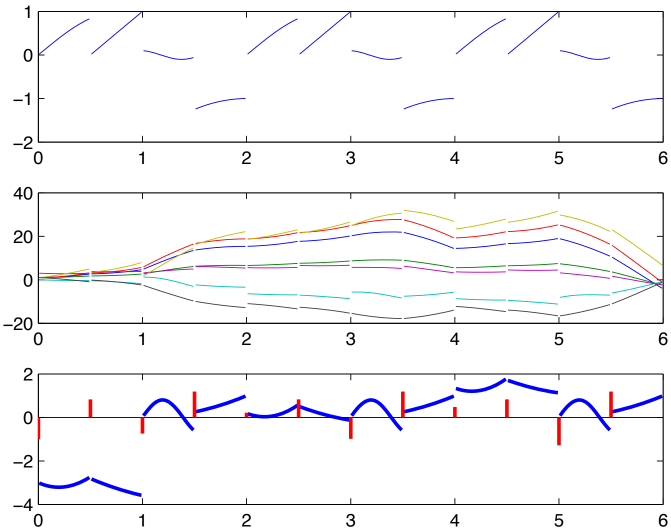

Fig. 1. The input (top), the seven states (middle) and the output (bottom) of the original system over the time interval $[0, 10]$. The impulses in the output are illustrated by red vertical lines (the height corresponds to the strength of the Dirac impulse), Dirac impulses in the states are not shown.

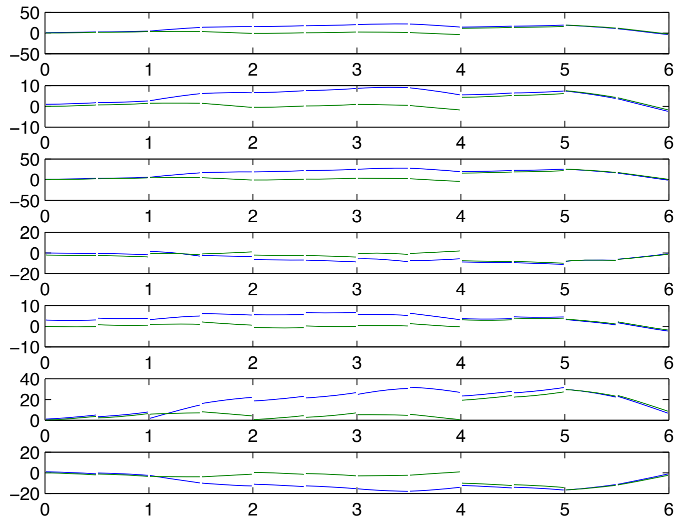

Fig. 2. The original seven states (x_1 at the top to x_7 at the bottom) in blue and its estimates in green. The Dirac impulses in the state are not shown.

The example is purely academic but has some special features:

1) Each mode is unobservable.
2) The switched DAE is unobservable in the sense of [15], i.e. $x(0^-)$ can not be determined.
3) After the switching sequence $1 \to 2 \to 3$ the system is determinable. Because of the repeated mode sequence, we also have, for each $p \in \mathbb{N}$, $\mathcal{Q}^p_{p-N} = \{0\}$ with $N = 3$.
4) In order to determine the current state, the information about the Dirac-impulses present in the output must be used.

We apply a discontinuous input to the original system leading to additional Dirac impulses in the output between the switching times (see Figure 1). However, the system copy (9) produces the same Dirac impulses so that these Dirac impulses do not appear in the difference $\hat{y} - y$.

The estimation of the seven states via our proposed observer is shown in Figure 2. It is clearly seen that, on the first four intervals, the observer does not improve the estimation of the states as there is not enough information available to improve the estimation. At the switching time $t = 4$ the observer can correct the estimation for the first time, which is clearly visible in the figure.

References

[1] M. Babaali and George J. Pappas. Observability of switched linear systems in continuous time. In *Hybrid Systems: Computation and Control*, volume 3414 of *LNCS*, pages 103–117. Springer, Berlin, 2005.

[2] A. Balluchi, L. Benvenuti, M.D. Di Benedetto, and A.L. Sangiovanni-Vincentelli. Observability for hybrid systems. In *Proc. 42rd IEEE Conf. Decis. Control, Hawaii, USA*, volume 2, pages 1159 – 1164, 2003.

[3] T. Berger, A. Ilchmann, and S. Trenn. The quasi-Weierstraß form for regular matrix pencils. *Lin. Alg. Appl.*, 436(10):4052–4069, 2012.

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

[5] L. Dai. *Singular Control Systems*. Lecture Notes in Control and Information Sciences, vol 118, Springer-Verlag, Berlin, 1989.

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

[7] M. M. Fahmy and J. O'Reilly. Observers for descriptor systems. *Int. J. Control*, 49:2013–2028, 1989.

[8] D. Liberzon. *Switching in Systems and Control*. Systems and Control: Foundations and Applications. Birkhäuser, Boston, 2003.

[9] H. Shim and A. Tanwani. 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*, 2012. DOI: 10.1002/rnc.2901, To appear.

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

[11] A. Tanwani, H. Shim, and D. Liberzon. Observability implies observer design for switched linear systems. In *Proc. ACM Conf. Hybrid Systems: Computation and Control*, pages 3 – 12, 2011. Also see the erratum: `www.inrialpes.fr/bipop/people/atanwani/errHSCC.pdf`

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

[13] A. Tanwani. *Invertibility and Observability of Switched Systems with Inputs and Outputs*. PhD thesis, University of Illinois at Urbana-Champaign, IL, USA, 2011. Available online: `http://hdl.handle.net/2142/29755`.

[14] A. Tanwani and S. Trenn. On observability of switched differential-algebraic equations. In *Proc. 49th IEEE Conf. Decis. Control, Atlanta, USA*, pages 5656–5661, 2010.

[15] A. Tanwani and S. Trenn. Observability of switched differential-algebraic equations for general switching signals. In *Proc. 51st IEEE Conf. Decis. Control, Maui, Hawaii*, pages 2648 – 2653, 2012.

[16] S. Trenn. *Distributional differential algebraic equations*. PhD thesis, Institut für Mathematik, Technische Universität Ilmenau, Universitätsverlag Ilmenau, Ilmenau, Germany, 2009.

[17] S. Trenn. Regularity of distributional differential algebraic equations. *Math. Control Signals Syst.*, 21(3):229–264, 2009.

[18] S. Trenn. Switched differential algebraic equations. In F. Vasca and L. Iannelli, eds., *Dynamics and Control of Switched Electronic Systems - Advanced Perspectives for Modeling, Simulation and Control of Power Converters*, chapter 6, pages 189–216. Springer-Verlag, London, 2012.

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

[20] K. T. Wong. The eigenvalue problem $\lambda T x + S x$. *J. Diff. Eqns.*, 16:270–280, 1974.

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