DAE Panel

DOI:10.xxx/xxx-x
Submitted: xx.xx.20xx
Accepted: xx.xx.20xx
Published: xx.xx.20xx

Model reduction for switched differential-algebraic equations with known switching signal

MD. SUMON HOSSAIN , AND STEPHAN TRENN

Abstract: Building on our recently proposed model reduction methods for switched ordinary linear systems we propose a comprehensive model reduction method for linear switched differential-algebraic equations (DAEs). In contrast to most other available model reduction methods for switched systems we consider the switching signal as a given time-variance of the system. This allows us to exploit certain linear subspaces in the reduction process and also provide in general significantly smaller reduced models compared to methods which consider arbitrary switching signals. Model reduction for switched DAEs has some unique features which makes a generalization of the available methods nontrivial; in particular, the presence of jumps and Dirac impulses in response to switches have to be carefully treated. Furthermore, due the algebraic constraints, the reachability subspaces cannot be the full space, hence a straightforward application of balanced truncation is not possible (because the corresponding reachability Gramians will be structurally non-invertible). We resolve this problem by first apply an exact model reduction which reduces the switched DAE to a switched ordinary systems with jumps and carefully keep track of the impulsive effects. As a second step we then apply a midpoint balanced truncation approach to further reduce the switched system. In addition to the challenge to appropriately take into account the Dirac impulses, another novel challenge was the occurrence of input-dependent state-jumps. We propose to deal with input-dependent jumps by combining certain discrete-time reachability Gramians with continuous time reachability Gramians. We provide corresponding Matlab implementations of the proposed algorithms and illustrate their effectiveness with some academic examples.

Keywords: balanced truncation, reduced realization, descriptor systems, reachability/observability Gramians, time-varying systems.

AMS Subject Classification (2020): 34A09, 65D99, 34A38, 93C15, 93B20, 93-08

1 Introduction

Differential-algebraic equations (DAEs) have become an important tool for modelling and simulation of constrained dynamical systems and are also known as descriptor systems or singular systems. DAEs naturally occur when modelling linear electrical circuits, simple mechanical systems or, in general, linear systems with additional linear algebraic constraints; they have been used for modelling a vast variety of problem, e.g. in economics [29, 28], demography [6], mechanical systems [5, 14], multibody dynamics [13, 36], electrical networks [10, 1, 31, 12], fluid mechanics [17, 27], chemical engineering [11, 32] and they are also particularly important in the simulation and design of very large scale integrated circuits. The theory of linear DAEs is well developed and mature cf. [6, 47, 7, 9, 8, 10, 33, 23, 15, 16, 34, 24, 5] and the references therein.

Switched DAEs arise when the system changes suddenly and the switching between the systems is induced by faults or an external switching rule, cf. [25, 44]. Switches or component faults induce jumps in certain state variables, and it is common to define additional jump-maps based on physical arguments. It turns out that general switched DAEs can have not only jumps in the solutions but also Dirac impulses and/or their derivatives. This makes it necessary to enlarge the underlying solution space to the space of distributions (also known as generalized functions). However, it can be shown that the classical Schwartz distribution space [37] is not suitable as a solution space for switched DAEs (the reason is that this space is too large and doesn't e.g. allow the restriction to intervals [41]). This problem was resolved in [40] where the space of *piecewise-smooth distributions* as an underlying solution space was proposed. Based on this piecewise-smooth distributional solution space several modelling and control results related to switched DAE have been obtained so far, see e.g., [26, 46, 42, 43, 18, 4, 38, 30].

Here we are interested in the problem of model reduction for linear switched DAEs of the following form:

$$\begin{aligned} E_\sigma \dot{x} &= A_\sigma x + B_\sigma u, \quad x(t_0^-) = \mathscr{X}_0 \subseteq \mathbb{R}^n, \\ y &= C_\sigma x + D_\sigma u, \end{aligned} \tag{1}$$

where x, u, y denote respectively, the state, input and output; $\sigma : [t_0, t_f) \to \mathscr{Q} := \{0, 1, \dots, \mathtt{m}\}$ is the switching signal defined on the interval of interest $[t_0, t_f)$ with switching times $s_k \in (t_0, t_f)$, $k \in \mathscr{Q}$, and is defined as

$$\sigma(t) = k, \quad t \in [s_k, s_{k+1}), \quad k = 0, 1, \dots, \mathtt{m}, \tag{2}$$

where $s_0 := t_0$ and $s_{\mathtt{m}+1} := t_f$. Furthermore, we define the mode duration of mode k as $\tau_k := s_{k+1} - s_k$. The active mode $k \in \mathscr{Q}$ is characterized by the matrix quintuple $(E_k, A_k, B_k, C_k, D_k) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n} \times \mathbb{R}^{p\times m}$. In order to guarantee existence and uniqueness of solution we will assume throughout this paper that the matrix pairs (E_k, A_k) are regular (see Section 2.1 for the formal definition and related consequences). If $E_k = I$ for all $k \in \mathscr{Q}$, system (1) is called a switched ordinary differential equation (ODE).

While model reduction for non-switched DAEs have been studied rather extensively, see the survey [2] and the references therein, there seems to be almost no research yet on model reduction for switched DAEs. The only exception seems to be [35], which however only considers a homogenous system without inputs and outputs.

Our approach extends and utilizes our previous approaches on reduced realization for switched ODEs [22] and midpoint balanced truncation for switched ODEs [21]. In particular, the first step in the reduction of (1) is finding an equivalent switched ODE with jumps and Dirac impulses without altering the input-output behavior, which is summarized in Algorithm 1. For this step there are no further assumption required, in particular, the DAEs can have arbitrary indeces and impulsive solutions. The second step is then applying model reduction based on balancing certain Gramians on the outcome of Algorithm 1. In contrast to the first step, there are certain assumptions which need to be satisfied. A first assumption is that the involved reachability and observability Gramians are non-singular, which we conjecture is the case in almost all practical situation because the reduction carried out in the first steps removes unreachable and unobservable states. Another assumption is a certain structure-decoupling property in the sense of Assumption 3, where the states (and the corresponding flow and jump maps) are decoupled into states which effect the impulses in the output and those which don't. Since arbitrary small errors in an Dirac impulse have an infinite L_2-norm, we only apply a further model reduction on those states which do not effect the output impulses. With this decoupling it is then possible to apply the midterm balanced truncation approach from [21], however, the presence of input-dependent state-jumps complicates the setting, because these jumps effect the reachability spaces and also have an quantitative effect on the reachability Gramians. How exactly a discrete input effects the continuous reachability Gramian is not really clear. Here, we first observe that the effect of the discrete input (effecting the state-jump) can be decoupled from the effect of the continuous input and a discrete-time reachability Gramian can be defined. We then propose to define the overall (midpoint) Gramian as the weighted sum of the continuous reachability Gramian and the discrete-time reachability Gramian. Since in the context of switched DAEs, the discrete input is given by the values of the continuous input (and its derivatives) at the jump time the weight between the two Gramians reflects how we measure the control energy needed to adjust specific values for the input (and its derivatives) at a specific time in relation to the overall integral-cost of the continuous input. At the moment we cannot provide any further theoretical justification on how to choose this weight optimally; furthermore, as for the switched ODE case we are not able to proof any error bounds but instead see our main contribution to propose some intuitively meaningful model reduction method for switched DAEs at all.

Notation: For two matrices P,Q with the same number of columns, we define $[P/Q] := \left[\begin{smallmatrix} P \\ Q \end{smallmatrix}\right]$; this notation generalizes in an obvious way to the case when more than two matrices are placed over each other. For a subspace $\mathscr{V} \in \mathbb{R}^n$ and a matrix $A \in \mathbb{R}^{n\times n}$ the smallest A-invariant subspace containing $\mathscr{V}$ is denoted by $\langle A \mid \mathscr{V}\rangle$ and the largest A-invariant subspace contained in $\mathscr{V}$ is denoted as $\langle \mathscr{V} \mid A\rangle$. In particular, when $\mathscr{V} = \operatorname{im} B$ for some matrix $B \in \mathbb{R}^{n\times m}$, then $\langle A \mid \mathscr{V}\rangle = \operatorname{im}[B, AB, \ldots, A^{n-1}B]$ and if $\mathscr{V} = \ker C$ for some matrix $C \in \mathbb{R}^{p\times n}$, then $\langle \mathscr{V} \mid A\rangle = \ker[C/CA/\cdots/CA^{n-1}]$.

2 Preliminaries

2.1 Regular matrix pairs

In this subsection we recall important properties of matrix pairs (E,A), in particular, associated projectors and selectors, which allow a convenient decomposition of the solutions of the corresponding DAE $E\dot{x} = Ax + Bu$. First recall, that a matrix pair (E,A) is called *regular* if the matrices E and A are square matrices of the same size and the polynomial $\det(sE - A)$ is not the zero polynomial. In the following we will also call the switched DAE (1) regular, if each mode's matrix pair (E_k, A_k) is regular.

For every regular matrix pair, there exists a unique decomposition of $\mathbb{R}^n = \mathscr{V}^* \oplus \mathscr{W}^*$ such that the linear operators E and A with respect to this decomposition become block diagonal

$$(E,A) \simeq \left(\left[\begin{smallmatrix} I & 0 \\ 0 & N \end{smallmatrix}\right], \left[\begin{smallmatrix} J & 0 \\ 0 & I \end{smallmatrix}\right]\right), \tag{3}$$

where N is a nilpotent linear map on $\mathscr{W}^*$. We will call the decomposition (3) the quasi-Weierstrass form (QWF) and in [3] it is shown, that the subspaces $\mathscr{V}^*$ and $\mathscr{W}^*$ can be calculated via the Wong sequences [49]:

$$\mathscr{V}_0 := \mathbb{R}^n, \quad \mathscr{V}_{k+1} := A^{-1}(E\mathscr{V}_k), \qquad \mathscr{W}_0 := \{0\}, \quad \mathscr{W}_{k+1} := E^{-1}(A\mathscr{W}_k),$$

where $A^{-1}(\cdot)$ and $E^{-1}(\cdot)$ stand for the pre-image (E and A are not assumed to be invertible). It is easily seen, that the two subspace sequences are nested and become stationary after at most n steps; in fact, they become stationary after the same number of steps, say $\nu \in \mathbb{N}$, which coincides with the nilpotency index of N in the QWF and is called the *index* of the matrix pair (E,A) (or the corresponding DAE). The *Wong limits* are defined as

$$\mathscr{V}^* = \bigcap_{k\in\mathbb{N}} \mathscr{V}_k = \mathscr{V}_\nu, \qquad \mathscr{W}^* = \bigcup_{k\in\mathbb{N}} = \mathscr{W}_\nu.$$

In addition to $\mathbb{R}^n = \mathscr{V}^* \oplus \mathscr{W}^*$, it can also be shown that $\mathbb{R}^n = E\mathscr{V}^* \oplus A\mathscr{W}^*$. By choosing basis matrices $V \in \mathbb{R}^{n\times n_J}$ and $W \in \mathbb{R}^{n\times n_N}$ of $\mathscr{V}^*$ and $\mathscr{W}^*$, resp., the QWF can explicitly be obtained by (SET, SAT), where

$$T = [V, W] \quad \text{and} \quad S = [EV, AW]^{-1}.$$

Define the consistency projector[1] as

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

and the differential/impulsive selector as

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

Note that in all three cases the definition does not depend on the specific choice of bases matrices V and W of $\mathscr{V}^*$ and $\mathscr{W}^*$. Furthermore, the consistency projector is a projector onto $\mathscr{V}^*$ along $\mathscr{W}^*$ (but in general it is not an orthogonal projector, because $\mathscr{V}^*$ and $\mathscr{W}^*$ are not orthogonal to each other), while the differential and impulsive selectors are not projectors in general.

Based on the above selectors we define the following matrices:

$$\begin{aligned} A^{\mathrm{diff}} &:= \Pi^{\mathrm{diff}}_{(E,A)} A, & B^{\mathrm{diff}} &:= \Pi^{\mathrm{diff}}_{(E,A)} B, \\ E^{\mathrm{imp}} &:= \Pi^{\mathrm{imp}}_{(E,A)}, & B^{\mathrm{imp}} &:= \Pi^{\mathrm{imp}}_{(E,A)} B. \end{aligned}$$

For a regular switched DAE (1) we define $\Pi_k := \Pi_{(E_k,A_k)}$, $\Pi^{\mathrm{diff/imp}}_k := \Pi^{\mathrm{diff/imp}}_{(E_k,A_k)}$ and A^{diff}_k, B^{diff}_k, E^{imp}_k, B^{imp}_k accordingly. Furthermore, if the corresponding matrix pair (E,A) is clear, we drop the index (E,A) from the consistency projector and differential/impulsive selectors.

2.2 Piecewise-smooth distributions

In addition to jumps, solutions of switched DAEs may also contain Dirac impulses (and their derivatives) [44], hence a distributional solution framework needs to be considered to study (1). We first recall the basic definitions of distributions (generalized functions) in the sense of Schwartz [37]. Smooth functions $\varphi : \mathbb{R} \to \mathbb{R}$ with bounded support are called *test-functions* and the linear space containing all such test-functions is denoted by $\mathscr{C}_0^\infty$. This linear space can be equipped with a certain topology, such that a sequence $(\varphi_k)_{k\in\mathbb{N}}$ of test-functions converges to zero, if, and only if, there exists a compact subset of $\mathbb{R}$ which contains all supports of the φ_k-s and furthermore, for each $i \in N$ the sequence $\varphi_k^{(i)}$ converges uniformly to 0 as $k \to \infty$. The space of *distributions* $\mathbb{D}$ is then defined as the dual space of $\mathscr{C}_0^\infty$, i.e.

$$\mathbb{D} := \{\, D : \mathscr{C}_0^\infty \to \mathbb{R} \mid D \text{ is linear and continuous} \,\},$$

where continuity can be checked by checking that $D(\varphi_k) \to 0$ for all sequences $(\varphi_k)_{k\in\mathbb{N}}$, which converge to zero in the above sense.

For every locally integrable function $f : \mathbb{R} \to \mathbb{R}$, the map

$$f_{\mathbb{D}} : \mathscr{C}_0^\infty \to \mathbb{R}, \quad \varphi \mapsto \int_{\mathbb{R}} f\varphi$$

defines a distribution and is called a regular distribution induced by f. The map $f \mapsto f_{\mathbb{D}}$ is an injective homomorphism[2], which justifies to call distributions “generalized functions”. For any distribution a derivative exists and is given by

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

This definition is consistent with the classical derivative, i.e. for a differentiable function $f : \mathbb{R} \to \mathbb{R}$, it holds that

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

which is a simple consequence from the integration by parts. The space of distributions contains many elements which are not induced by locally integrable functions, the most famous example is the Dirac

[1] The reason for this name becomes clear in the following subsection, when inconsistent initial values are discussed

[2] Locally integrable functions which only differ on a set of measure zero are identified with each other.

impulse (Dirac delta, delta "function") at $t \in \mathbb{R}$, given by

$$\delta_t(\varphi) := \varphi(t).$$

It is easily seen that the Dirac impulse at $t \in \mathbb{R}$ is the (distributional) derivative of the Heaviside step function $\mathbb{1}_{[t,\infty)}$ (with jump from zero to one at $t \in \mathbb{R}$), i.e.

$$\delta_t = ((\mathbb{1}_{[t,\infty)})_{\mathbb{D}})'.$$

In order to consider the switched DAE (1) with distributional solutions, it is necessary to consider restrictions of distributions to intervals of the form $[s_k, s_{k+1})$. However, it was shown in [41], that it is impossible to define such a distributional restriction for general distributions. To resolve this problem, the space of piecewise-smooth distributions is considered in [40], which is defined as

$$\mathbb{D}_{\mathrm{pw}\mathscr{C}^\infty} := \left\{ \alpha_{\mathbb{D}} + \sum_{\tau \in T} D_\tau \;\middle|\; \begin{array}{l} \alpha \in \mathscr{C}^\infty_{\mathrm{pw}}, T \subseteq \mathbb{R} \text{ is discrete}, \\ \forall \tau \in T:\; D_\tau \in \operatorname{span}\{\delta_\tau, \delta_\tau', \delta_\tau'', \dots\} \end{array} \right\},$$

where $\mathscr{C}^\infty_{\mathrm{pw}}$ is the space of piecewise-smooth functions and a set $T \subseteq \mathbb{R}$ is called discrete if each intersection with a bounded interval only contains finitely many points. This space has many nice properties, in particular, it is closed under differentiation, it is possible to define left, right and impulse evaluation at any $t \in \mathbb{R}$, where for $D = \alpha_{\mathbb{D}} + \sum_{\tau \in T} D_\tau$,

$$D(t^-) := \alpha(t^-), \quad D(t^+) := \alpha(t^+)$$

and

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

Furthermore, it is possible to define a restriction to any interval $\mathscr{I} \subseteq \mathbb{R}$ as follows:

$$D_{\mathscr{I}} := (\alpha_{\mathscr{I}})_{\mathbb{D}} + \sum_{\tau \in T \cap \mathscr{I}} D_\tau,$$

where $\alpha_{\mathscr{I}}(t) := \alpha(t)$ for $t \in \mathscr{I}$ and zero otherwise.

These properties of $\mathbb{D}_{\mathrm{pw}\mathscr{C}^\infty}$ allow to consider (1) with distributional solutions, i.e. $x \in \mathbb{D}^n_{\mathrm{pw}\mathscr{C}^\infty}$, $u \in \mathbb{D}^m_{\mathrm{pw}\mathscr{C}^\infty}$ and consequently $y \in \mathbb{D}^p_{\mathrm{pw}\mathscr{C}^\infty}$. However, in the following we restrict our attention to impulse free inputs, i.e. the input is a regular distribution induced by a piecewise-smooth function.

2.3 Inconsistent initial values and solution decomposition

Due to the algebraic constraints, not all initial values are feasible for DAEs of the form

$$E\dot{x} = Ax + Bu,$$

however in [40] it was shown that for regular DAEs the corresponding initial trajectory problem (ITP)

$$\begin{aligned} x_{(-\infty,t_0)} &= x^0_{(-\infty,t_0)} \\ (E\dot{x})_{[t_0,\infty)} &= (Ax + Bu)_{[t_0,\infty)} \end{aligned} \tag{4}$$

considered in the space of piecewise-smooth distributions has a unique solution $x \in \mathbb{D}^n_{\mathrm{pw}\mathscr{C}^\infty}$ for any arbitrary past trajectory $x^0 \in \mathbb{D}^n_{\mathrm{pw}\mathscr{C}^\infty}$ and any input $u \in \mathbb{D}^m_{\mathrm{pw}\mathscr{C}^\infty}$. It was also shown there that the solution on $[t_0, \infty)$ only depends on $x(t_0^-)$, hence, in the following we will abbreviate the ITP (4) as a classical initial value problem

$$E\dot{x} = Ax + Bu, \quad x(t_0^-) = x_0 \in \mathbb{R}^n. \tag{5}$$

Furthermore, in the context of switched DAEs of the form (1), we assume that in the past the ODE

$$\dot{x} = 0, \quad x(t_0^-) = x_0,$$

was active; in particular, we formally extend the switching signal to $-\infty$ with the value -1 and let $E_{-1} := I$, $A_{-1} := 0$, $B_{-1} = 0$, $C_{-1} = 0$, $D_{-1} = 0$.

Lemma 1 (cf. [48]). *Consider the (inconsistent) initial value problem* (5) *with regular matrix pair* (E,A) *with corresponding consistency projector* Π *and corresponding matrices* A^{diff}, B^{diff}, E^{imp}, B^{imp}*. Then* (x,u) *is a (distributional) solution of* (5) *on* $[t_0,t_f)$*, if and only if* $x = x^{\text{diff}} + x^{\text{imp}}$ *where* x^{diff} *and* x^{imp} *are the (distributional) solutions of the following two (inconsistent) initial value problems on* $[t_0,t_f)$

$$\begin{aligned} \dot{x}^{\text{diff}} &= A^{\text{diff}} x^{\text{diff}} + B^{\text{diff}} u, & x^{\text{diff}}(t_0^-) &= \Pi x_0, \\ E^{\text{imp}} \dot{x}^{\text{imp}} &= x^{\text{imp}} + B^{\text{imp}} u, & x^{\text{imp}}(t_0^-) &= (I-\Pi) x_0. \end{aligned}$$

Proof. Necessity. Let $x^{\text{diff}} := \Pi x$ and $x^{\text{imp}} := (I-\Pi)x$, then the claim follows from multiplying the DAE with Π^{diff} or Π^{imp}, respectively, and observing that $\Pi^{\text{diff}} E = \Pi$ and $\Pi^{\text{imp}} A = (I-\Pi)$.
Sufficiency. First observe that the solutions x^{diff} and x^{imp} satisfy $\Pi x^{\text{diff}} = x^{\text{diff}}$ and $(I-\Pi)x^{\text{imp}} = x^{\text{imp}}$. It then follows that

$$\begin{aligned} E\dot{x} &= E\Pi\dot{x}^{\text{diff}} + E(I-\Pi)\dot{x}^{\text{imp}} = S^{-1}T^{-1}(\dot{x}^{\text{diff}} + E^{\text{imp}}\dot{x}^{\text{imp}}) \\ &= S^{-1}T^{-1}(A^{\text{diff}}\Pi + (I-\Pi))x + S^{-1}T^{-1}(B^{\text{diff}} + B^{\text{imp}})u \end{aligned}$$

and the claim follows from verifying that $S^{-1}T^{-1}(A^{\text{diff}}\Pi + (I-\Pi)) = A$ and $S^{-1}T^{-1}(B^{\text{diff}} + B^{\text{imp}}) = B$. □

Note that for an impulse-free input, any solution x^{diff} in Lemma 1 satisfies $x^{\text{diff}}(t^-) = x^{\text{diff}}(t^+)$ for all $t \geq t_0$, in particular, the initial condition can then be replaced by $x(t_0^+) = \Pi x_0$.

Lemma 2. *The solution of* $E^{\text{imp}}\dot{x}^{\text{imp}} = x^{\text{imp}} + B^{\text{imp}}u$, $x^{\text{imp}}(t_0^-) = (I-\Pi)x_0$ *considered on* $[t_0,t_f)$ *and for an impulse-free input* u *is given by*

$$\begin{aligned} x^{\text{imp}} &= \mathbf{B}^{\text{imp}} \mathbf{U}^{\nu}, \quad \text{on } (t_0,t_f), \\ x^{\text{imp}}[t_0] &= -\sum_{i=0}^{\nu-2} (E^{\text{imp}})^{i+1} (x_0 - \mathbf{B}^{\text{imp}} \mathbf{U}^{\nu}(t_0^+)) \delta_{t_0}^{(i)}, \end{aligned}$$

where $\nu \in \mathbb{N}$ *is the nilpotency index of* E^{imp} *and*

$$\mathbf{U}^{\nu} := \left[u^\top, \dot{u}^\top, \cdots, u^{(\nu-1)^\top}\right]^\top, \quad \mathbf{B}^{\text{imp}} := -\left[B^{\text{imp}}, E^{\text{imp}}B^{\text{imp}}, \ldots, (E^{\text{imp}})^{\nu-1}B^{\text{imp}}\right].$$

Proof. This is a simple consequence from Theorems 6.4.4 and 6.5.1 in [44], taking into account that $E^{\text{imp}}(I-\Pi) = 0$. □

Note that $\mathbf{U}^{\nu}$ may be impulsive despite u being impulse-free, because u is not assumed to be continuous and jumps in u will lead to Dirac impulses in $\mathbf{U}^{\nu}$ if $\nu > 1$.

3 Reduced realization

3.1 Equivalent switched ODEs with jumps and impulses

For the regular switched DAE (1) with switching signal (2) we consider the following surrogate switched ODE with jumps and Dirac impulses:

$$\boxed{\begin{aligned} \dot z &= A_k^{\text{diff}} z + B_k^{\text{diff}} u, \quad \text{on } (s_k, s_{k+1}), \quad z(t_0^-) = x_0, \\ z(s_k^+) &= \Pi_k \left[z(s_k^-) + \mathbf{B}_{k-1}^{\text{imp}} \mathbf{U}^{\nu_{k-1}}(s_k^-) \right], \; k \geq 0 \\ w &= C_k z + D_k u + \mathbf{D}_k^{\text{imp}} \mathbf{U}^{\nu_k}, \quad \text{on } (s_k, s_{k+1}), \end{aligned}} \tag{6a}$$

$$\boxed{w[s_k] = -C_k \sum_{i=0}^{\nu_k - 2} (E_k^{\text{imp}})^{i+1} \left[z(s_k^-) + \mathbf{B}_{k-1}^{\text{imp}} \mathbf{U}^{\nu_{k-1}}(s_k^-) - \mathbf{B}_k^{\text{imp}} \mathbf{U}^{\nu_k}(s_k^+) \right] \delta_{s_k}^{(i)},} \tag{6b}$$

where $\nu_k \in \mathbb{N}$ is the nilpotency index of E_k^{imp}; $\mathbf{U}^{\nu_k}$ and $\mathbf{B}_k^{\text{imp}}$ are given as in Lemma 2, $\mathbf{D}_k^{\text{imp}} := C_k \mathbf{B}_k^{\text{imp}}$ and $\mathbf{B}_{-1}^{\text{imp}} := 0$.

Theorem 3. *Consider an arbitrary impulse-free input $u \in \mathbb{D}^m_{\text{pw}\mathscr{C}^\infty}$ with the corresponding (distributional) outputs $y \in \mathbb{D}^p_{\text{pw}\mathscr{C}^\infty}$ of the regular switched DAE (1) and $w \in \mathbb{D}^p_{\text{pw}\mathscr{C}^\infty}$ of the switched ODE (6). Then $y = w$ on $[t_0, t_f)$.*

Proof. Let x denote the solution of the switched DAE (1) and z the solution of the ODE (6) for the same (impulse-free) input u. The proof can be divided into three steps as follows.

Step 1: We show that $z = \Pi_k x =: x_k^{\text{diff}}$ on (s_k, s_{k+1}).

From Lemma 1 it follows that it suffices to show that $z(s_k^+) = x_k^{\text{diff}}(s_k^+)$. For $k = 0$ this is trivially satisfied, because $z(s_0^+) = z(t_0^+) = \Pi_0 x_0 = x_0^{\text{diff}}(t_0^-) = x_0^{\text{diff}}(s_0^+)$. Continuing inductively, assume now $z(s_k^+) = x_k^{\text{diff}}(s_k^+)$ for some $k \geq 0$. Then in view of Lemma 1, we have $z(s_{k+1}^-) = x_k^{\text{diff}}(s_{k+1}^-)$ and together with Lemma 2 we have $x(s_{k+1}^-) = z(s_{k+1}^-) + \mathbf{B}_k^{\text{imp}} \mathbf{U}^{\nu_k}(s_{k+1}^-)$. Furthermore, $x_{k+1}^{\text{diff}}(s_{k+1}^+) = x_{k+1}^{\text{diff}}(s_{k+1}^-) = \Pi_{k+1} x(s_{k+1}^-) = z(s_{k+1}^+)$.

Step 2: The corresponding outputs on (s_k, s_{k+1}) are equal.

The output equation on (s_k, s_{k+1}) is given by

$$y = C_k x + D_k u = C_k \Pi_k x + C_k (I - \Pi_k) x + D_k u = C_k z + C_k x_k^{\text{imp}} + D_k u,$$

where $x_k^{\text{imp}} := (I - \Pi_k) x = \mathbf{B}_k^{\text{imp}} \mathbf{U}^{\nu_k}$ according to Lemma 2; from which $y = w$ on (s_k, s_{k+1}) follows.

Step 3: The impulse parts of the outputs are equal.

According to Lemma 2 we have

$$x[s_k] = x_k^{\text{imp}}[s_k] = -\sum_{i=0}^{\nu_k - 2} (E_k^{\text{imp}})^{i+1} \left(x(s_k^-) - \mathbf{B}_k^{\text{imp}} \mathbf{U}^{\nu_k}(s_k^+) \right) \delta_{s_k}^{(i)}$$

as well as

$$x(s_k^-) = x_{k-1}^{\text{diff}}(s_k^-) + x_{k-1}^{\text{imp}}(s_k^-) = z(s_k^-) + \mathbf{B}_{k-1}^{\text{imp}} \mathbf{U}^{\nu_{k-1}}(s_k^-).$$

From this it is clear that $y[s_k] = C_k x[s_k] = w[s_k]$ and the proof is complete. □

Remark 4. *As already highlighted after Lemma 2, the term $\mathbf{U}^{\nu_k}$ appearing in the equation for the output w in (6a) may also contain Dirac impulses away from the switching times, if the input u (or one of its derivatives) is discontinuous. In particular, not all impulses in w are explicitly given by (6b). However, this "inconvinience" can easily be resolved by adding extra artificial switching times at all discontinuities of u (with identical coefficient matrices before and after the switch).*

A consequence of the equivalent input-output behavior between the switched DAE (1) and the switched ODE with jumps and impulses (6) is that a reduced order model of (1) can be obtained by considering a reduced order model of (6) instead. In particular, we are seeking a reduced switched ODE with jumps and impulses of the form

$$\left.\begin{aligned} \dot{\widehat{z}} &= \widehat{A}_k^{\text{diff}} z + \widehat{B}_k^{\text{diff}} u, \quad \text{on } (s_k, s_{k+1}), \quad \widehat{z}(t_0^-) = \widehat{x}_0 \in \widehat{\mathscr{X}_0}, \\ \widehat{z}(s_k^+) &= \widehat{\Pi}_k \widehat{z}(s_k^-) + \widehat{J}_k^{\nu} \mathbf{U}^{\nu_{k-1}}(s_k^-),\ k \geq 0 \\ \widehat{w} &= \widehat{C}_k \widehat{z} + D_k u + \mathbf{D}_k^{\text{imp}} \mathbf{U}^{\nu_k}, \quad \text{on } (s_k, s_{k+1}), \end{aligned}\right\} \tag{7a}$$

$$\widehat{w}[s_k] = \sum_{i=0}^{\nu_k-2} \widehat{C}_k^i \widehat{z}(s_k^-) \delta_{s_k}^{(i)} + \sum_{i=0}^{\nu_k-2} \left(\mathbf{D}_{k,i}^{\text{imp}-} \mathbf{U}^{\nu_{k-1}}(s_k^-) - \mathbf{D}_{k,i}^{\text{imp}+} \mathbf{U}^{\nu_k}(s_k^+) \right) \delta_{s_k}^{(i)}, \tag{7b}$$

where $\mathbf{D}_{k,i}^{\text{imp}-} := -C_k (E^{\text{imp}})_k^{i+1} \mathbf{B}_{k-1}^{\text{imp}}$ and $\mathbf{D}_{k,i}^{\text{imp}+} := -C_k (E^{\text{imp}})_k^{i+1} \mathbf{B}_k^{\text{imp}}$ are not affected by the reduction. Note that each mode can have different state-dimension $\widehat{n}_k$, so formally we would have to consider mode-dependent state variables $\widehat{z}_k : (s_k, s_{k+1}) \to \mathbb{R}^{\widehat{n}_k}$. However, to simplify notation, we just write $\widehat{z}(\cdot)$ instead of $\widehat{z}_k(\cdot)$, where we have to keep in mind that the dimension of $\widehat{z}(t)$ depends on t.

In our previous work [21] we have proposed a model reduction method for switched ODEs with jumps, however there are three obstacles which prevents a direct application of these methods to (6):

1) The solutions of (6a) considered on (s_k, s_{k+1}) evolve within the subspace $\operatorname{im} \Pi_k$ which is a strict subspace of $\mathbb{R}^n$ unless the switched DAE is already a switched ODE (without jumps). In particular, the corresponding reachability Gramian can never have full rank, which was an assumption to carry out the midpoint reduced balanced truncation method.
2) The jump rule in (6a) depends on the input, which was not considered in our previous work.
3) The additional impulsive output (6b) needs to be taken into account appropriately.

Note that the additional feed-through terms in the output equations of (6a) and (6b) involving derivatives of the input, do not pose any technical difficulties, because these terms remain unaffected by any model reduction procedure (note the missing "hats" in all D-terms of (7)).

There are different approaches to deal with the above mentioned challenges. To deal with the first challenge, we can simply apply our already proposed method to obtained a reduced realization for a switched ODE with jumps; however, the other two challenges still needs to be addressed. A simple way to address the other two challenges in general is simply to assume that the switched DAE is such that the jump rule is independent from the input and that the Dirac impulses in the output do only depend directly on the input. These assumption can be formalized as follows:

Assumption 1 (Input-independent jumps)**.** *Consider the regular switched DAE (1) on $[t_0, t_f)$ with switching signal (2) and corresponding consistency projectors Π_k and extended input matrices $\mathbf{B}_k^{\text{imp}}$ as in Lemma 2. Assume that for all $k \geq 1$ that*

$$\Pi_k \mathbf{B}_{k-1}^{\text{imp}} = 0.$$

Note that Assumption 1 is satisfied if $\Pi_k \Pi_{k-1}^{\text{imp}} = 0$, however, this is in general a more conservative assumption.

Assumption 2 (Only input induced output impulses)**.** *Consider the regular switched DAE* (1) *on* $[t_0, t_f)$ *with switching signal* (2) *and corresponding consistency projectors* Π_k*. Assume that for all* $k \geq 1$ *and all* $i \geq 1$,

$$C_k (E_k^{\mathrm{imp}})^i \Pi_{k-1} = 0.$$

Note that Assumption 2 is satisfied if $C_k E_k^{\mathrm{imp}} = 0$, however, this is in general stronger than necessary.

Under these assumption and utilizing the fact that every solution of (6a) satisfies $z = \Pi_k z$ on (s_k, s_{k+1}) we have the following simplified version of (6).

Corollary 5. *Consider the regular switched DAE* (1) *with switching signal* (2) *satisfying Assumptions 1 and 2, then for any impulse-free input the output of* (1) *equals the output of*

$$\begin{aligned}
\dot{z} &= A_k^{\mathrm{diff}} z + B_k^{\mathrm{diff}} u, \quad \textit{on } (s_k, s_{k+1}), \quad z(t_0^-) = x_0 \\
z(s_k^+) &= \Pi_k z(s_k^-), \quad k \geq 0, \\
w &= C_k z + D_k u + \mathbf{D}_k^{\mathrm{imp}} \mathbf{U}^{\nu_k}, \quad \textit{on } (s_k, s_{k+1}), \\
w[s_k] &= \sum_{i=0}^{\nu_k - 2} \left(\mathbf{D}_{k,i}^{\mathrm{imp}-} \mathbf{U}^{\nu_{k-1}}(s_k^-) - \mathbf{D}_{k,i}^{\mathrm{imp}+} \mathbf{U}^{\nu_k}(s_k^+) \right) \delta_{s_k}^{(i)}.
\end{aligned}$$

Consequently, the reduced realization method for switched ODEs with jumps [22] can be applied without alteration to obtain a reduced realization with mode-dependent state dimensions. Afterwards, the midpoint balanced truncation method [21] can directly be applied to further reduce the state-dimension while introducing a small output error.

3.2 Reduced realization for switched ODEs with input dependent jumps and impulsive outputs

In this section, we want to generalize the existing reduced realization approach for switched ODE with jumps to switched systems of the form (6) without making any further assumptions. In order to streamline the notation, we consider in the following general switched ODEs with jumps and impulsive output (not necessarily induced by a switched DAE) governed by a switching signal of the form (2):

$$\left.\begin{aligned}
\dot{x} &= A_k x + B_k u, \quad \text{on } (s_k, s_{k+1}), \quad x(t_0^-) = x_0 \in \mathscr{X}_0, \\
x(s_k^+) &= J_k^x x(s_k^-) + J_k^v v_k, \quad k \geq 0, \\
y &= C_k x, \quad \text{on } (s_k, s_{k+1}),
\end{aligned}\right\} \tag{8a}$$

$$y[s_k] = \sum_{i=0}^{\rho_k} C_k^i x(s_k^-) \delta_{s_k}^{(i)}, \tag{8b}$$

Note that compared to (6) we do not consider the feed-through terms in the output, because they do not play a role in obtaining a reduced realization. Furthermore, we view $v_k \in \mathbb{R}^{m_k}$, $k = 0, 1, \ldots$ as discrete additional inputs which are independent from the smooth input u. While formally $v_k = \mathbf{U}^{\nu_k}(s_k^-)$ in (6a) depends on the smooth input u (and its derivatives), the values $u^{(i)}(s_k^-)$ determining v_k can be chosen completely independently from the values of u on $(s_k, s_{k+1} - \varepsilon)$ for any $\varepsilon > 0$. From a model reduction viewpoint it is therefore justified to consider these effects of the input on the output as independent from each other. In contrast to (6) we do allow a mode-dependent state-dimension, i.e. $A_k \in \mathbb{R}^{n_k \times n_k}$, $B_k \in \mathbb{R}^{n_k \times m}$, $J_k^x \in \mathbb{R}^{n_k \times n_{k-1}}$, $J_k^v \in \mathbb{R}^{n_k \times m_k}$, $C_k, C_k^i \in \mathbb{R}^{p \times n_k}$ and formally the state has to be consider mode-wise as $x_k : (s_k, s_{k+1}) \to \mathbb{R}^{n_k}$, but with some mild abuse of notation we just write x instead of x_k.

In the following we will remove unreachable and unobservable states and it turns out, that the approach proposed in [22] via the weak Kalman-decomposition and extended/restricted reachable/unobservable spaces can be adapted in a quite straightforward way to the situation of an input dependent jump map and impulsive outputs. First we define the (exact) reachable and unobservable subspaces of (8) as follows.

Definition 6 (cf. [19, Defs. 4.4&4.6])**.** *The reachable subspace of* (8) *on* $[t_0,t)$ *for* $t \in (t_0,t_f]$ *is*

$$\mathscr{R}^{\sigma}_{[t_0,t)} := \left\{ x(t^-) \;\middle|\; x(\cdot) \text{ is solution of (8) for some } u(\cdot), v_0, v_1, \ldots \right\};$$

the unobservable subspace of (8) *on* (t,t_f) *for* $t \in [t_0,t_f)$ *is*

$$\mathscr{U}^{\sigma}_{(t,t_f)} := \left\{ x(t^+) \;\middle|\; \begin{array}{l} x(\cdot) \text{ solve (8) on } (t,t_f) \text{ with arbitrary } x(t^+) \text{ for } u(\cdot) = 0, v_0 = v_1 = \ldots = v_m = 0 \\ \text{and has output } y = 0 \text{ on } (t,t_f) \end{array} \right\}$$

In order to calculate the reachable and unobservable subspaces at $t = s_k$ recursively we define the following sequence of subspaces:

$$\begin{aligned} &\mathscr{M}^{\sigma}_{-1} := \mathscr{X}_0, \\ &\mathscr{M}^{\sigma}_{k} := \mathscr{R}_k + e^{A_k\tau_k}(J^x_k \mathscr{M}^{\sigma}_{k-1} + \operatorname{im} J^v_k), \quad k = 0,1,\ldots \mathtt{m}, \end{aligned}$$

where $\mathscr{R}_k := \langle A_k \mid \operatorname{im} B_k \rangle$ is the classical reachable subspace and

$$\begin{aligned} &\mathscr{N}^{\sigma}_{\mathtt{m}} := \mathscr{U}_m, \\ &\mathscr{N}^{\sigma}_{k} := \mathscr{U}_k \cap e^{-A_k\tau_k}(((J^x_k)^{-1}\mathscr{N}^{\sigma}_{k+1}) \cap \mathscr{U}^{\text{imp}}_{k+1}), \quad k = \mathtt{m}-1,\ldots,0, \end{aligned}$$

where $\mathscr{U}_k := \langle \ker C_k \mid A_k \rangle$ is the classical unobservable space of the pair (C_k, A_k) and $\mathscr{U}^{\text{imp}}_k := \ker[C^0_k / C^1_k / \ldots / C^{p_k}_k]$ is the impulse unobservable space (cf. [39]).

Lemma 7. *Consider the switched system* (8)*, then, for all* $k = 0,1,\ldots,m$

$$\mathscr{M}^{\sigma}_k = \mathscr{R}^{\sigma}_{[t_0,s_{k+1})} \quad \text{and} \quad \mathscr{N}^{\sigma}_k = \mathscr{U}^{\sigma}_{(s_k,t_f)}.$$

Proof. The proof is analogously as the one of [20, Lems. 10 & 13] and therefore omitted. □

Following the approach of [20] we define a sequence of extended reachable / restricted unobservable subspace which are also A_k-invariant and therefore allow a weak Kalman decomposition. The underlying idea is to utilize the fact that for any subspace $\mathscr{V} \in \mathbb{R}^n$ and any matrix $A \in \mathbb{R}^{n\times n}$

$$\langle \mathscr{V} \mid A \rangle \subseteq e^{At}\mathscr{V} \subseteq \langle A \mid \mathscr{V} \rangle, \quad \forall t \in \mathbb{R}.$$

This leads to the consideration of the following sequences of subspaces (which do not depend on the mode durations anymore):

$$\begin{aligned} &\overline{\mathscr{R}}_{-1} := \mathscr{X}_0, \\ &\overline{\mathscr{R}}_{k} := \mathscr{R}_k + \langle A_k \mid J^x_k \overline{\mathscr{R}}_{k-1} + \operatorname{im} J^v_k \rangle, \quad k = 0,\ldots,\mathtt{m} \end{aligned} \tag{9}$$

and

$$\begin{aligned} &\underline{\mathscr{U}}_{\mathtt{m}} := \mathscr{U}_{\mathtt{m}}, \\ &\underline{\mathscr{U}}_{k} := \mathscr{U}_k \cap \langle ((J^x_{k+1})^{-1}\underline{\mathscr{U}}_{k+1}) \cap \mathscr{U}^{\text{imp}}_{k+1}) \mid A_k \rangle, \qquad k = \mathtt{m}-1,\ldots,0. \end{aligned} \tag{10}$$

By definition, the extended reachable subspaces $\overline{\mathscr{R}}_k$ and the restricted unobservable subspaces $\underline{\mathscr{U}}_k$ are A_k-invariant; furthermore, it is easily seen that for $t \in (s_k, s_{k+1})$

$$\overline{\mathscr{R}}_k \supseteq \mathscr{R}^{\sigma}_{[t_0,t)} \supseteq \mathscr{R}_k \quad \text{and} \quad \underline{\mathscr{U}}_k \subseteq \mathscr{U}^{\sigma}_{(t,t_f)} \subseteq \mathscr{U}_k.$$

Now for each mode a weak Kalman decomposition as in [20] can be obtained, which then can be used to define (mode-dependent) left- and right-projectors which remove unreachable and unobservable states.

In fact, for each mode k choose invertible matrix $Q_k = [P_k^1, P_k^2, P_k^3, P_k^4]$ such that

$$\operatorname{im} P_k^1 = \overline{\mathscr{R}}_k \cap \underline{\mathscr{U}}_k, \quad \operatorname{im}[P_k^1, P_k^2] = \overline{\mathscr{R}}_k, \quad \operatorname{im}[P_k^1, P_k^3] = \underline{\mathscr{U}}_k. \tag{11}$$

Then the left projector is $V_k := P_k^2$ and the corresponding right projector W_k is the second block row of Q_k^{-1}, in particular, $W_k V_k = I$. Finally, to also reduce the dimension of the initial state, we can remove those states, which do not contribute to the reduced state after the initial jump and also do not contribute in an impulsive output; this can be achieved by choosing a basis matrix V_{-1} such that

$$\mathscr{X}_0 = \operatorname{im} V_{-1} \oplus (\mathscr{X}_0 \cap \ker W_0 J_0^x \cap \ker \mathscr{U}_0^{\text{imp}}). \tag{12}$$

In fact, let W_{-1} be a basis matrix of $(\mathscr{X}_0 \cap \ker W_0 J_0^x \cap \ker \mathscr{U}_0^{\text{imp}})$, then there is for all $x_0 \in \mathscr{X}_0$ unique pair $(\widehat{x}_0, \bar{x}_0)$ of appropriate size, such that $x_0 = [V_{-1}, W_{-1}] \binom{\widehat{x}_0}{\bar{x}_0}$; let's denote the matrix representation of the corresponding map $x_0 \mapsto \widehat{x}_0$ by $\Pi^{\mathscr{X}_0}$.

Theorem 8 (Reduced realization of (8))**.** *Consider the switched system (8) with corresponding extended reachable subspaces $\overline{\mathscr{R}}_k$ and restricted unobservable subspaces $\underline{\mathscr{U}}_k$ together with the induced weak Kalman-decomposition; let W_k and V_k denote the corresponding left- and right-projectors; choose V_{-1} according to (12) and with corresponding initial value reduction matrix $\Pi^{\mathscr{X}_0}$. Define the reduced system*

$$\begin{aligned}
\dot{\widehat{x}} &= \widehat{A}_k \widehat{x} + \widehat{B}_k u, \quad \text{on } (s_k, s_{k+1}), \quad \widehat{x}(t_0^-) = \widehat{x}_0 := \Pi^{\mathscr{X}_0} x_0 \in \widehat{\mathscr{X}_0},\\
\widehat{x}(s_k^+) &= \widehat{J}_k^x \widehat{x}(s_k^-) + \widehat{J}_k^v v_k, \quad k \geq 0,\\
\widehat{y} &= \widehat{C}_k \widehat{x}, \quad \text{on } (s_k, s_{k+1}), \quad \widehat{y}[s_k] = \sum_{i=0}^{\rho_k} \widehat{C}_k^i \widehat{x}(s_k^-) \delta_{s_k}^{(i)},
\end{aligned} \tag{13}$$

with $\widehat{A}_k := W_k A_k V_k$, $\widehat{B}_k := W_k B_k$, $\widehat{C}_k := C_k V_k$, $\widehat{C}_k^i := C_k^i V_{k-1}$, $\widehat{J}_k^x := W_k J_k^x V_{k-1}$, $\widehat{J}_k^v := W_k J_k^v$, $\widehat{\mathscr{X}_0} := \mathbb{R}^{\widehat{n}_{-1}}$, where $\widehat{n}_{-1} = \operatorname{rank} V_{-1}$. Then for all impulse-free inputs u, all $v_0, v_1, \ldots, v_m$ and all initial values $x_0 \in \mathscr{X}_0$ the (impulsive) outputs y of (8) and $\widehat{y}$ of (13) are equal.

Proof. Step 1: We show that $\widehat{x} := W_k x$ solves (13) on (s_k, s_{k+1}).

Consider a solution x of (8) for some initial value $x_0 \in \mathscr{X}_0$ and let $\widehat{x} := W_k x$. The satisfaction of the differential equation $\dot{\widehat{x}} = \widehat{A}_k \widehat{x} + \widehat{B} u$ on (s_k, s_{k+1}) follows with the same arguments as in the first part of the proof of [22, Lem. 16] and is therefore omitted. It remains to be shown that $\widehat{x}$ satisfies the initial jump rule for some suitable $\widehat{x}_0$ and all subsequent jumps at the switches. In view of (12), we can choose $\widehat{x}_0$ such that $x_0 = V_{-1}\widehat{x}_0 + \widetilde{x}_0$, where $\widetilde{x}_0 \in \mathscr{X}_0 \cap \ker W_0 J_0^x \cap \ker \mathscr{U}_0^{\text{imp}} \subseteq \ker W_0 J_0^x$. We then have

$$\widehat{x}(t_0^+) = W_0 x(t_0^+) = W_0(J_0^x x_0 + J_0^v v_0) = W_0 J_0 (V_{-1}\widehat{x}_0 + \widetilde{x}_0) + W_0 J_0^v v_0 = \widehat{J}_0^x \widehat{x}_0 + \widehat{J}_0^v v_0$$

as desired. It remains to be shown that for $k \geq 1$ we have $\widehat{x}(s_k^+) = \widehat{J}_k^x \widehat{x}(s_k^-) + \widehat{J}_k^v v_k$. Assuming inductively, that $\widehat{x}$ solves (13) on $[t_0, s_k)$ already, we have

$$\widehat{x}(s_k^+) = W_k x(s_k^+) = W_k(J_k^x x(s_k^-) + J_k^v v_k) = W_k J_k^x x(s_k^-) + \widehat{J}_k^v v_k.$$

Hence it remains to be shown that $W_k J_k^x x(s_k^-) = \widehat{J}_k^x \widehat{x}(s_k^-) = W_{k-1} J_k^x V_{k-1} W_{k-1} x(s_k^-)$. This can be shown following the same arguments as in second part of the proof of [22, Lem. 16] and is therefore omitted.

Step 2: We show that $\widehat{y} = y$.

From $\underline{\mathscr{U}}_k \subseteq \mathscr{U}_k \subseteq \ker C_k$ and because (using the notation (11)) $\underline{\mathscr{U}}_k = \operatorname{im}[P_k^1, P_k^3]$ we have $C_k Q_k = [0, C_k P_k^2, 0, *]$. Furthermore, for all $t \in (s_k, s_{k+1})$ we have $x(t) \in \mathscr{R}^{\sigma}_{[t_0,t)} \subseteq \overline{\mathscr{R}}_k = \operatorname{im}[P_k^1, P_k^2]$, hence

$Q_k^{-1}x(t) = \begin{pmatrix} * \\ W_k x(t) \\ 0 \\ 0 \end{pmatrix}$. Consequently,

$$C_k x(t) = C_k Q_k Q_k^{-1} x(t) = C_k V_k W_k x(t) = \widehat{C}_k \widehat{x}(t)$$

and it remains to be shown that $y[s_k] = \widehat{y}[s_k]$. For $k=0$, we have, according to (12), $x(t_0^-) = V_{-1}\widehat{x}_0 + \tilde{x}_0$, where $\tilde{x}_0 \in \mathscr{X}_0 \cap \ker W_0 J_0^x \cap \ker \mathscr{U}_0^{\mathrm{imp}} \subseteq \mathscr{U}_0^{\mathrm{imp}} \subseteq \ker C_0^i$, hence $C_k^i x(t_0^-) = C_k^i V_{-1}\widehat{x}_0 + C_k^i \tilde{x}_0 = C_k^i \widehat{x}(t_0^-)$, which shows that $y[t_0] = \widehat{y}[t_0]$. For $k \geq 1$ we have, similar as above, $x(s_k^-) \in \overline{\mathscr{R}}_{k-1}$ and hence $Q_{k-1}^{-1}x(s_k^-) = \begin{pmatrix} * \\ W_{k-1}x(t) \\ 0 \\ 0 \end{pmatrix}$. Furthermore, $\underline{\mathscr{U}}_{k-1} \subseteq \mathscr{U}_k^{\mathrm{imp}} \subseteq \ker C_k^i$ and hence $C_k^i Q_{k-1} = [0, C_k^i V_{k-1}, 0, *]$ and, therefore,

$$C_k^i x(s_k^-) = C_k^i Q_{k-1}^{-1} Q_{k-1} x(s_k^-) = C_k^i V_{k-1} W_{k-1} x(s_k^-) = \widehat{C}_k^i \widehat{x}(s_k^-).$$

This concludes the proof. □

We conclude this section by summarizing the reduction process of a regular switched DAE (1) while preserving the input-output behavior in Algorithm 1 and a Matlab implementation can be found at [45]. This algorithm is illustrated with the following academic example.

Algorithm 1: Algorithm to obtain reduced realization of a switched DAE

Data: $\mathscr{X}_0$, regular modes (E_k, A_k, B_k, C_k), $k = 0, 1, \ldots, \mathtt{m}$.

Result: Reduced system (7) given by $\widehat{\mathscr{X}_0}$, $\widehat{A}_k^{\mathrm{diff}}$, $\widehat{B}_k^{\mathrm{diff}}$, $\widehat{C}_k$, $\widehat{C}_k^0, \ldots, \widehat{C}_k^{\nu_k-2}$, $\widehat{\Pi}_k$, $\widehat{J}_k^{\nu}$, $\mathbf{D}_k^{\mathrm{imp}}$, $\mathbf{D}_{k,0}^{\mathrm{imp}-}$, $\mathbf{D}_{k,1}^{\mathrm{imp}-}, \ldots, \mathbf{D}_{k,\nu_k-2}^{\mathrm{imp}-}$, $\mathbf{D}_{k,0}^{\mathrm{imp}+}$, $\mathbf{D}_{k,1}^{\mathrm{imp}+}, \ldots, \mathbf{D}_{k,\nu_k-2}^{\mathrm{imp}+}$, $k = 0, \ldots, \mathtt{m}$.

for $k = 0, 1, \ldots, m$ **do**

Calculate matrices: Π_k, A_k^{diff}, B_k^{diff}, E_k^{imp}, B_k^{imp}, ν_k via the Wong-limits as in Section 2.1.

Calculate matrices: $\mathbf{B}_k^{\mathrm{imp}} := -\left[B_k^{\mathrm{imp}}, \ldots, (E_k^{\mathrm{imp}})^{\nu_k-1} B_k^{\mathrm{imp}}\right]$, $\mathbf{D}_k^{\mathrm{imp}} := C_k \mathbf{B}_k^{\mathrm{imp}}$, and for $i = 0, 1, \ldots, \nu_k - 2$, $C_k^i := -C_k (E_k^{\mathrm{imp}})^{i+1}$, $\mathbf{D}_{k,i}^{\mathrm{imp}+} := C_k^i \mathbf{B}_k^{\mathrm{imp}}$, $\mathbf{D}_{k,i}^{\mathrm{imp}-} := C_k^i \mathbf{B}_{k-1}^{\mathrm{imp}}$, where $\mathbf{B}_{-1}^{\mathrm{imp}} := 0$.

Calculate (basis matrices of) subspaces: $\mathscr{R}_k := \operatorname{im}\left[B_k^{\mathrm{diff}}, A_k^{\mathrm{diff}} B_k^{\mathrm{diff}}, \ldots, (A_k^{\mathrm{diff}})^{n_k-1} B_k^{\mathrm{diff}}\right]$, $\mathscr{U}_k := \ker\left[C_k / C_k A_k^{\mathrm{diff}} / \ldots / C_k (A_k^{\mathrm{diff}})^{n_k-1}\right]$, $\mathscr{U}_k^{\mathrm{imp}} := \ker\left[C_k E_k^{\mathrm{imp}} / \ldots / C_k (E_k^{\mathrm{imp}})^{\nu_k-1}\right]$.

end

Compute the sequence of extended reachability subspaces $\overline{\mathscr{R}}_0, \overline{\mathscr{R}}_1, \ldots, \overline{\mathscr{R}}_{\mathtt{m}}$ via

$$\overline{\mathscr{R}}_0 := \mathscr{R}_0 + \langle A_0^{\mathrm{diff}} | \Pi_0 \mathscr{X}_0 \rangle, \qquad \overline{\mathscr{R}}_k := \mathscr{R}_k + \langle A_k^{\mathrm{diff}} \mid \Pi_k \overline{\mathscr{R}}_{k-1} + \operatorname{im} \Pi_k \mathbf{B}_{k-1}^{\mathrm{imp}} \rangle, \quad k = 1, \ldots, \mathtt{m}.$$

Compute the sequence of restricted unobservability subspaces $\underline{\mathscr{U}}_{\mathtt{m}}, \underline{\mathscr{U}}_{\mathtt{m}-1}, \ldots, \underline{\mathscr{U}}_0$ via

$$\underline{\mathscr{U}}_{\mathtt{m}} := \mathscr{U}_{\mathtt{m}}, \qquad \underline{\mathscr{U}}_k := \mathscr{U}_k \cap \langle (\Pi_{k+1}^{-1} \underline{\mathscr{U}}_{k+1} \cap \mathscr{U}_{k+1}^{\mathrm{imp}}) \mid A_k^{\mathrm{diff}} \rangle, \quad k = \mathtt{m} - 1, \ldots, 0.$$

for $k = 0, \ldots, m$ **do**

Obtain left-/right-projector W_k and V_k via the weak Kalman decomposition based on $\overline{\mathscr{R}}_k$ and $\underline{\mathscr{U}}_k$.

Compute reduced ODE matrices: $\widehat{A}_k^{\mathrm{diff}} := W_k A_k^{\mathrm{diff}} V_k$, $\widehat{B}_k^{\mathrm{diff}} := W_k B_k^{\mathrm{diff}}$, $\widehat{C}_k := C_k V_k$.

end

Choose basis matrices V_{-1}, W_{-1} s.t. $\mathscr{X}_0 = \operatorname{im}[V_{-1}, W_{-1}]$ and $\operatorname{im} W_{-1} = \mathscr{X}_0 \cap \ker W_0 \Pi_0 \cap \mathscr{U}_0^{\mathrm{imp}}$.

$\widehat{\mathscr{X}_0} := \mathbb{R}^{\operatorname{rank} V_{-1}}$ and calculate $\Pi^{\mathscr{X}_0} := [I, 0]([V_{-1}, W_{-1}]^\top [V_{-1}, W_{-1}])^{-1} [V_{-1}, W_{-1}]^{-1}$.

for $k = 0, \ldots, m$ **do**

Calculate reduced impulse output matrices: $\widehat{C}_k^i := C_k^i V_{k-1}$.

Calculate reduced jump maps: $\widehat{\Pi}_k := W_k \Pi_k V_{k-1}$ and $\widehat{J}_k^{\nu} := W_k \Pi_k \mathbf{B}_{k-1}^{\mathrm{imp}}$.

end

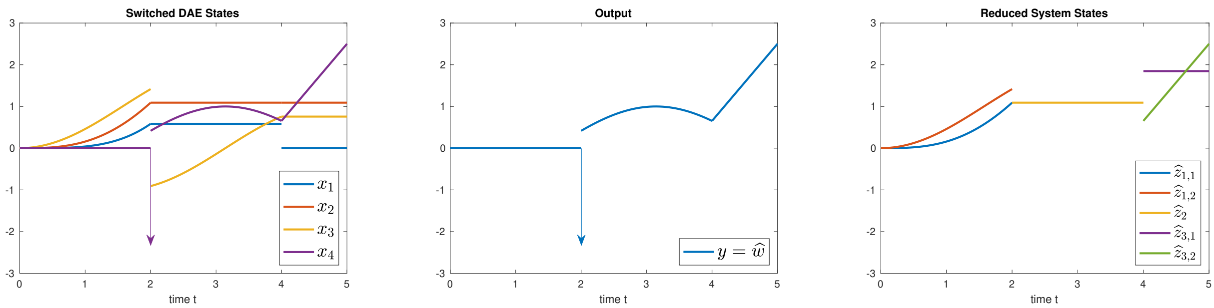

Figure 1. Numerical simulation of Example 9; Diracs are shown as arrows, where the length corresponds to the magnitude. Left: Evolution of the four states of the switched DAE (1). Middle: Output of switched DAE (1) and reduced system (7). Right: Evolution of states (which have mode-dependent dimensions) of reduced system (7). The Matlab code to produce these simulations is available at [45].

Example 9. *Consider the switched DAE* (1) *with the modes* $(E_k, A_k, B_k, C_k, D_k)$, $k = 0,1,2$, *given by*

$$i=0: \quad \left(\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1 \end{bmatrix}, \begin{bmatrix} 0&1&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0\\0\\1\\0 \end{bmatrix}, \begin{bmatrix} 0&0&0&0 \end{bmatrix}, 0 \right),$$
$$i=1: \quad \left(\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&1&0 \end{bmatrix}, \begin{bmatrix} 0&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&1 \end{bmatrix}, \begin{bmatrix} 0\\0\\1\\0 \end{bmatrix}, \begin{bmatrix} 0&0&0&1 \end{bmatrix}, 0 \right),$$
$$i=2: \quad \left(\begin{bmatrix} 0&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1 \end{bmatrix}, \begin{bmatrix} 1&0&0&0\\0&0&0&0\\0&0&0&0\\0&1&1&0 \end{bmatrix}, \begin{bmatrix} 0\\0\\0\\0 \end{bmatrix}, \begin{bmatrix} 0&0&0&1 \end{bmatrix}, 0 \right).$$

It is easily seen that the index ν_k *of* (E_k, A_k) *is given by* $\nu_0 = 0$ *(this mode is an ODE),* $\nu_1 = 2$ *(the pair* (E_1, A_1) *is already in QWF with* $N_1 = \left[\begin{smallmatrix} 0&0\\1&0 \end{smallmatrix}\right]$*) and* $\nu_2 = 1$ *(the pair* (E_2, A_2) *is also already in a (permuted) QWF with* $N_2 = 0$*). Furthermore, the matrices* $(\Pi_k, A_k^{\mathrm{diff}}, B_k^{\mathrm{diff}}, E_k^{\mathrm{imp}}, B_k^{\mathrm{imp}})$ *are given by*

$$i=0: \quad \left(\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1 \end{bmatrix}, \begin{bmatrix} 0&1&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0\\0\\1\\0 \end{bmatrix}, \begin{bmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0\\0\\0\\0 \end{bmatrix} \right),$$
$$i=1: \quad \left(\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0\\0\\0\\0 \end{bmatrix}, \begin{bmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&1&0 \end{bmatrix}, \begin{bmatrix} 0\\0\\1\\0 \end{bmatrix} \right),$$
$$i=2: \quad \left(\begin{bmatrix} 0&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1 \end{bmatrix}, \begin{bmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&0\\0&1&1&0 \end{bmatrix}, \begin{bmatrix} 0\\0\\0\\0 \end{bmatrix}, \begin{bmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0\\0\\0\\0 \end{bmatrix} \right).$$

The corresponding feedthrough terms are then

$$\mathbf{D}_0^{\mathrm{imp}} = 0_{1\times 0}, \quad \mathbf{D}_1^{\mathrm{imp}} = \begin{bmatrix} 0 & -1 \end{bmatrix}, \quad \mathbf{D}_2^{\mathrm{imp}} = 0_{1\times 1}, \quad \mathbf{D}_{1,0}^{\mathrm{imp}+} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \mathbf{D}_{1,0}^{\mathrm{imp}-} = 0_{1\times 0}.$$

The solution of the corresponding switched system (1) *on the time interval* $[t_0, t_f] := [0, 5]$ *with switching times* $s_1 = 2$, $s_2 = 4$ *and input* $u(t) = \sin(t)$ *is illustrated in the left part of Figure 1. It can clearly be seen that the state trajectories contain jumps (at both switching instants) as well as a Dirac impulse (at the first switching instant). The corresponding output is shown in the middle of Figure 1 and it contains as well jumps and a Dirac impulse.*

With $\mathscr{X}_0 = \{0\}$, *the extended reachable spaces can be calculated as*

$$\overline{\mathscr{R}}_0 = \operatorname{im} \begin{bmatrix} 1&0&0\\0&1&0\\0&0&1\\0&0&0 \end{bmatrix}, \quad \overline{\mathscr{R}}_1 = \operatorname{im} \begin{bmatrix} 1&0\\0&1\\0&0\\0&0 \end{bmatrix}, \quad \overline{\mathscr{R}}_2 = \operatorname{im} \begin{bmatrix} 0&0&0\\1&0&0\\0&1&0\\0&0&1 \end{bmatrix},$$

and the restricted unobservable spaces are

$$\underline{\mathscr{U}}_0 = \operatorname{im}\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \quad \underline{\mathscr{U}}_1 = \operatorname{im}\begin{bmatrix}1&0\\0&0\\0&1\\0&0\end{bmatrix}, \quad \underline{\mathscr{U}}_2 = \operatorname{im}\begin{bmatrix}1&0\\0&-1\\0&1\\0&0\end{bmatrix}.$$

Via the corresponding weak Kalman decomposition, we obtain the left- and right-projectors given by

$$\begin{aligned} W_0 &= \begin{bmatrix}0&1&0&0\\0&0&1&0\end{bmatrix}, & W_1 &= \begin{bmatrix}0&1&0&0\end{bmatrix}, & W_2 &= \begin{bmatrix}0&1&1&0\\0&0&0&1\end{bmatrix} \\ V_0 &= \begin{bmatrix}0&0\\1&0\\0&1\\0&0\end{bmatrix}, & V_1 &= \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, & V_2 &= \begin{bmatrix}0&0\\1&0\\0&0\\0&1\end{bmatrix}. \end{aligned}$$

Together with $V_{-1} = 0_{4\times 0}$, this yields the following reduced switched system with jumps and Dirac impulses.
- On $[s_0, s_1)$ with state-space dimension $\widehat{n}_0 = 2$:

$$\begin{aligned} \dot{\widehat{z}} &= \begin{bmatrix}0&1\\0&0\end{bmatrix}\widehat{z} + \begin{bmatrix}0\\1\end{bmatrix}u, & \widehat{z}(s_0^+) &= \begin{bmatrix}0\\0\end{bmatrix}, \\ \widehat{w} &= 0, & \widehat{w}[s_0] &= 0. \end{aligned}$$

- On $[s_1, s_2)$ with state-space dimension $\widehat{n}_1 = 1$:

$$\begin{aligned} \dot{\widehat{z}} &= 0, & \widehat{z}(s_1^+) &= \begin{bmatrix}1&0\end{bmatrix}\widehat{z}(s_1^-), \\ \widehat{w} &= \begin{bmatrix}0&-1\end{bmatrix}\begin{bmatrix}u\\\dot{u}\end{bmatrix}, & \widehat{w}[s_1] &= \left(\begin{bmatrix}0&-1\end{bmatrix}\widehat{z}(s_1^-) - \begin{bmatrix}1&0\end{bmatrix}\begin{bmatrix}u(s_1^+)\\\dot{u}(s_1^+)\end{bmatrix}\right)\delta_{s_1}. \end{aligned}$$

- On $[s_2, s_3)$ with state-space dimension $\widehat{n}_2 = 2$:

$$\begin{aligned} \dot{\widehat{z}} &= \begin{bmatrix}0&0\\1&0\end{bmatrix}, & \widehat{z}(s_2^+) &= \begin{bmatrix}1\\0\end{bmatrix}z(s_2^-) + \begin{bmatrix}-1&0\\0&-1\end{bmatrix}\begin{bmatrix}u(s_2^-)\\\dot{u}(s_2^-)\end{bmatrix}, \\ \widehat{w} &= \begin{bmatrix}0&1\end{bmatrix}\widehat{z}, & \widehat{w}[s_2] &= 0. \end{aligned}$$

The evolution of the reduced states are shown in the right part of Figure 1. Note that these states do never contain Dirac impulses, because the Dirac-impulses are explicitely calculated in the corresponding output; the latter is then indeed identical to the one from the switched DAE shown in the middle of Figure 1.

4 Midpoint balanced truncation

4.1 Model reduction in the presence of Dirac impulses

In this section we also consider the switched system (8) but assume that all unobservable and unreachable states have already been removed (e.g. via the method proposed by Theorem 8). If after this reduction, there are no input-induced jumps and no state-induced output impulses (i.e. Assumptions 1 and 2 are satisfied in the context of switched DAEs) then Gramians can be defined which provide on the one hand the minimal input energy $\int_{t_0}^{t} u(\tau)^\top u(\tau)\mathrm{d}\tau$ which is needed to reach a given state from zero on the time interval $[t_0, t)$ and, on the other hand, provide the output energy $\int_{t}^{t_f} y(\tau)^\top y(\tau)\mathrm{d}\tau$ which any given initial state at time t produces on (t, t_f) for a zero input. Based on this energy interpretation of the Gramians, a coordinate transformation can be constructed which results in equal reachability and observability Gramians, which means that states can be identified which are simultaneously difficult to reach and difficult to observe; these states are then prime candidates to be removed to obtain a reduced system without significantly alter the input-output behavior.

However, this intuitive approach is not valid anymore in case there are input dependent jumps or state-dependent output impulses. For the latter this is obvious, because a Dirac impulse in the output results in an infinite output energy (because the L_2-norm of a Dirac impulse is infinite). In particular, any state-approximation which results in a different impulsive output response has infinite L_2-error. This means that model reduction is restricted to a subspace of the state space which is contained in $\ker \mathscr{U}_k^{\text{imp}}$. More precisely, the state space of mode $k-1$ must be decomposable as $\mathbb{R}^{n_{k-1}} = \mathscr{X}_{k-1}^{\overline{\text{imp}}} \oplus \mathscr{X}_{k-1}^{\text{imp}}$, such that $\mathscr{X}_{k-1}^{\overline{\text{imp}}}$ is invariant under the dynamics of mode $k-1$ and such that $\mathscr{X}_{k-1}^{\text{imp}} \subseteq \ker \mathscr{U}_k^{\text{imp}}$. This decomposition can be obtained by a classical Kalman-observability decomposition of the pair $(A_{k-1}, [C_k^0/C_k^1/\ldots/C_k^{\rho_k}])$. For each mode k, we therefore have to restrict the model-reduction on the subspace $\mathscr{X}_k^{\overline{\text{imp}}}$. Furthermore, we need to assume that the dynamics on the subspace $\mathscr{X}_k^{\text{imp}}$ are initialized with an exact initial value, hence the component of the jump map J_k^x mapping onto $\mathscr{X}_k^{\text{imp}}$ must be independent from the previous reduction space $\mathscr{X}_{k-1}^{\overline{\text{imp}}}$, i.e. $J_k^x \mathscr{X}_{k-1}^{\overline{\text{imp}}} \subseteq \mathscr{X}_k^{\overline{\text{imp}}}$. Finally, we assume that the dynamics on $\mathscr{X}_k^{\overline{\text{imp}}}$ and $\mathscr{X}_k^{\text{imp}}$ can be completely decoupled (a sufficient condition for this decoupling is that in the corresponding Kalman observability decomposition the set of eigenvalues of the block-diagonal matrices are disjoint), including the state-jump rules. Altogether, we make the following assumptions to handle state-dependent output impulses:

Assumption 3. *Consider the switched system* (8)*. Assume there is a mode dependent coordinate transformation* $T_k = [T_k^{\overline{\text{imp}}}, T_k^{\text{imp}}]$ *for* $k = -1, 0, 1, \ldots, m$*, such that for all* $k = 0, 1, \ldots, m$

$$T_k^{-1} A_k T_k = \begin{bmatrix} A_k^{\overline{\text{imp}}} & 0 \\ 0 & A_k^{\text{imp}} \end{bmatrix}, \quad T_k^{-1} J_k^x T_{k-1} \quad = \begin{bmatrix} J_k^{x^{\overline{\text{imp}}}} & 0 \\ 0 & J_k^{x^{\text{imp}}} \end{bmatrix}, \quad C_k^i T_{k-1}^{\overline{\text{imp}}} = 0, \quad i = 0, 1, \ldots, \rho_k.$$

Lemma 10. *Consider* (8) *satisfying Assumption* 3*. Let* $\begin{bmatrix} x^{\overline{\text{imp}}} \\ x^{\text{imp}} \end{bmatrix} = T_k^{-1} x$ *on* (s_k, s_{k+1})*, then* (8) *has the same input-output behavior as*

$$\begin{aligned} \dot{x}^{\overline{\text{imp}}} &= A_k^{\overline{\text{imp}}} x^{\overline{\text{imp}}} + B_k^{\overline{\text{imp}}} u, \quad \text{on } (s_k, s_{k+1}), \quad x^{\overline{\text{imp}}}(t_0^-) = x_0^{\overline{\text{imp}}} \in \mathscr{X}_0^{\overline{\text{imp}}}, \\ \dot{x}^{\text{imp}} &= A_k^{\text{imp}} x^{\text{imp}} + B_k^{\text{imp}} u, \quad \text{on } (s_k, s_{k+1}), \quad x^{\text{imp}}(t_0^-) = x_0^{\text{imp}} \in \mathscr{X}_0^{\text{imp}}, \\ x^{\overline{\text{imp}}}(s_k^+) &= J_k^{x^{\overline{\text{imp}}}} x^{\overline{\text{imp}}}(s_k^-) + J_k^{v^{\overline{\text{imp}}}} v_k, \\ x^{\text{imp}}(s_k^+) &= J_k^{x^{\text{imp}}} x^{\text{imp}}(s_k^-) + J_k^{v^{\text{imp}}} v_k, \\ y = y^{\overline{\text{imp}}} + y^{\text{imp}} &= C_k^{\overline{\text{imp}}} x^{\overline{\text{imp}}} + C_k^{\text{imp}} x^{\text{imp}}, \quad y[s_k] = \sum_{i=0}^{\rho_k} C_k^{\text{imp},i} x^{\text{imp}}(s_k^-) \delta_{s_k}^{(i)}, \end{aligned} \tag{14}$$

where $A_k^{\overline{\text{imp}}}$, A_k^{imp}, $J_k^{x^{\overline{\text{imp}}}}$ *and* $J_k^{x^{\text{imp}}}$ *are given according to Assumption* 3, $\begin{bmatrix} B_k^{\overline{\text{imp}}} \\ B_k^{\text{imp}} \end{bmatrix} = T_k^{-1} B_k$, $\begin{bmatrix} J_k^{v^{\overline{\text{imp}}}} \\ J_k^{v^{\text{imp}}} \end{bmatrix} = T_k^{-1} J_k^v$, $[C_k^{\overline{\text{imp}}}, C_k^{\text{imp}}] = C_k T_k$, $C_k^{\text{imp},i} = C_k^i T_{k-1}^{\text{imp}}$, $\begin{pmatrix} x_0^{\overline{\text{imp}}} \\ x_0^{\text{imp}} \end{pmatrix} := T_{-1}^{-1} x_0$, $\mathscr{X}_0^{\overline{\text{imp}}} := [I, 0] T_{-1}^{-1} \mathscr{X}_0$, $\mathscr{X}_0^{\text{imp}} := [0, I] T_{-1}^{-1} \mathscr{X}_0$. *In particular, any (approximative) model reduction applied on* $x^{\overline{\text{imp}}}$ *does not effect the impulsive output.*

Proof. This follows straightforwardly by plugging $x = T_k \begin{bmatrix} x^{\overline{\text{imp}}} \\ x^{\text{imp}} \end{bmatrix}$ into (8) while taking into account Assumption 3. $\square$

Remark 11. *If* (8) *is induced by the switched DAE* (1)*, then Assumption* 2 *implies Assumption* 3*, because the former implies that the output does not contain state-induced impulses at all and hence* $C_k^i = 0$ *can be assumed in* (8)*, which means that* $C_k^{i,\text{imp}} = 0$ *in* (14)*, which of course is not a necessary consequence of Assumption* 3*.*

4.2 Model reduction for switched ODEs with input dependent jumps

As argued above, we should not attempt to apply model reduction on the state component x^{imp} in (14), because even small mismatches will be infinitely "amplified" by the Dirac impulses. Therefore, we can in the following restrict our attention to the $\overline{x^{\text{imp}}}$-part of (14) and how to carry out model reduction for the corresponding switched ODE with input dependent jumps. The corresponding switched ODE has exactly the same form as (8a) and with the corresponding adjustment of notation we therefore simply continue to study (8a). We first show that we can decouple the effect of the continuous and discrete input as follows.

Lemma 12. *The trajectory x is a solution of (8a) if, and only if, $x = x_u + x_v$, where x_u is a solution of (8a) where $v_k = 0$, $x_u(t_0^-) = 0$ and x_v is a solution of (8a) with $u = 0$, $x_v(t_0^-) = x(t_0^-)$.*

Proof. Let $x = x_u + x_v$ for solutions x_u and x_v of (8a) with only one type of input. Then

$$\dot{x} = \dot{x}_u + \dot{x}_v = A_k x_u + B_k u + A_k x_v = A_k x + B_k u$$

and

$$x(s_k^+) = x_u(s_k^+) + x_v(s_k^+) = J_k^x x_u(s_k^-) + J_k^x x_v(s_k^-) + J_k^v v_k = J_k^x x(s_k^-) + J_k^v v_k.$$

This shows that x is indeed a solution of (8a). Conversely, let x be a solution of (8a) and let x_u be a solution of (8a) with vanishing discrete inputs, then similarly as above it can be shown that $x_v := x - x_u$ solves (8a) with vanishing continuous input. This concludes the proof. □

Based on the decoupling of the continuous and discrete input we propose the following model reduction approach for switched systems of the form (8a):

(i) Calculate midpoint reachability Gramians and observability Gramians as in [21] for (8a) with $v_k = 0$ and $x_0 = 0$.
(ii) Define suitable discrete reachability Gramians for (8a) with $u = 0$.
(iii) Define the overall reachability Gramian as weighted sum of the midpoint and discrete reachability Gramian.
(iv) Apply balanced truncation based on the overall reachability Gramian and the midpoint observability Gramian as in [21].

For defining the discrete reachability Gramian of (8a) with $u = 0$, we consider the following *midpoint* discrete time system

$$x_{k+1}^m = A_k^m x_k^m + B_k^m v_k, \quad x_0^m = x_0 \in \mathscr{X}_0, \tag{15}$$

where $A_k^m := e^{A_k \tau_k/2} J_k^x e^{A_{k-1}\tau_{k-1}/2} \in \mathbb{R}^{n_k \times n_{k-1}}$ (with $A_{-1} := 0$) and $B_k^m := e^{A_k \tau_k/2} J_k^v$. It is easily seen that for every discrete input sequence $(v_k)_{k\in\mathbb{N}}$ and any corresponding solution x of (8a) with $u = 0$ the midpoint sequence $x_k^m := x(\frac{s_k + s_{k+1}}{2})$ solves (15). In other words, (15) captures the reachability properties of the system (8a) at the *midpoints* with respect to the discrete input v_k. We can now define the discrete-time midpoint reachability Gramian recursively as follows:

$$\mathscr{P}_{-1}^m := \gamma X_0 X_0^\top, \quad \mathscr{P}_k^m = A_k^m \mathscr{P}_{k-1}^m A_k^{m\top} + B_k^m B_k^{m\top}, \tag{16}$$

where X_0 is an orthonormal matrix such that $\operatorname{im} X_0 = \mathscr{X}_0$ and $\gamma > 0$ is a parameter to relate the (inverse) "control-cost" in the past to achieve a nonzero initial state with the future control costs. Note that here (8a) was obtained as a result of Algorithm 9, hence $\mathscr{X}_0 = \mathbb{R}^{n_{-1}}$, i.e. $\mathscr{P}_0^m = \gamma I$ in case $n_{-1} = \dim \mathscr{X}_0 > 0$ or $\mathscr{P}_0^m = 0$ otherwise.

We conjecture that similar as in the continuous time case (cf. [21, Thm. 5]) the discrete-time reachability Gramian as defined in (16) is tightly related to the required input energy in (15) to reach a final state from the origin. However, a formal proof is outside the scope of this manuscript.

Note that considering a discrete version of the observability Gramian is not necessary, because the observability Gramian is independently defined from the input signal, hence the presence of an additional discrete input in the jump map has no effect on the continuous observability Gramian.

As mentioned above we now define the overall reachability Gramian $\mathscr{P}_k^\lambda$ as the weighted sum of the midpoint reachability Gramian $\overline{\mathscr{P}}_k^\sigma$ (using the notation from [21]) and the discrete-time midpoint reachability Gramian $\mathscr{P}_k^m$, i.e.

$$\mathscr{P}_k^\lambda := \overline{\mathscr{P}}_k^\sigma + \lambda \mathscr{P}_k^m, \tag{17}$$

for some $\lambda > 0$. For small values of $\lambda > 0$ the costs for effecting the jump via the input is small compared to the cost of the continuous inputs, while for large $\lambda > 0$ assigning specific values for v_k is "expensive" compared to the cost of specifying a continuous input on an interval. The choice of λ is related to the earlier discussion that the discrete inputs v_k are not independent from the continuous input u when considering a switched DAE, because v_k is related to the derivatives of u at the switching times.

Once the reachability Gramian $\mathscr{P}_k^\lambda$ has been defined, left and right projectors can be defined via a balanced truncation approach identical to the approach in [21], provided each of the Gramians $\mathscr{P}_k^\lambda$ and $\overline{\mathscr{Q}}_k$ (the midpoint observability Gramian for (8a)) is invertible, which is usually the case after a reduced realization is obtained via Algorithm 1. Note that the discrete reachability Gramian is often not invertible (especially for small k), however it is added to the usually invertible midpoint reachability Gramians and since both matrices are symmetric and positive (semi-)definite, the weighted sum will be invertible. The midpoint balanced truncation method for a switched ODE with input dependent jumps (8a) is summarized in Algorithm 2 and a Matlab implementation is available at [45].

Algorithm 2: Midpoint balanced truncation for linear switched system with input dependent jumps.

Data: $\mathscr{X}_0$, modes of (8a) given by A_k, B_k, J_k^x, J_k^v, C_k, $k = 0, 1, \dots, \mathtt{m}$,
initial, switching and final times: $t_0 = s_0, s_1, \dots, s_\mathtt{m}, s_{\mathtt{m}+1} = t_f$.

Parameters: $\lambda > 0$, $\gamma > 0$ (if $\mathscr{X}_0 \neq \{0\}$), desired approximation thresholds $\varepsilon_k > 0$ or desired reduction size $r_k \leq n_k$.

Result: Reduced system of (8a) given by $\widehat{\mathscr{X}_0}$, $\widehat{A}_k$, $\widehat{B}_k$, $\widehat{J}_k^x$, $\widehat{J}_k^v$, $\widehat{C}_k$, $k = 0, \dots, \mathtt{m}$.

Compute the sequence of midpoint reachability and observability Gramians $\overline{\mathscr{P}}_0^\sigma, \overline{\mathscr{P}}_1^\sigma, \dots, \overline{\mathscr{P}}_\mathtt{m}^\sigma$; $\overline{\mathscr{Q}}^\sigma{}_\mathtt{m}, \overline{\mathscr{Q}}_{\mathtt{m}-1}^\sigma, \dots, \overline{\mathscr{Q}}_0^\sigma$ according to [21, Alg. 2].

Let $F_{-1} := I$; $\mathscr{P}_{-1}^m := \gamma X_0 X_0^\top$ where X_0 is an orthonormal basis matrix of $\mathscr{X}_0$.

for $k = 0, 1, \dots, m$;
do
- Compute discrete-time dynamics: $F_k^m := e^{A_k(s_{k+1}-s_k)/2}$, $A_k^m := F_k^m J_k^x F_{k-1}^m$, $B_k^m := F_k^m J_k^v$.
- Compute discrete-time reachability Gramian: $\mathscr{P}_k^m := A_k^m \mathscr{P}_{k-1}^m A_k^{m\top} + B_k^m B_k^{m\top}$.
- Combine discrete-time with continuous-time Gramian: $\mathscr{P}_k^\lambda := \lambda \mathscr{P}_k^m + \overline{\mathscr{P}}_k^\sigma$.

end

Calculate left- and right projectors Π_k^l and Π_k^r via [21, Alg. 1] for (A_k, B_k, C_k, J_k^x) and midpoint Gramians $\mathscr{P}_k^\lambda$ and $\overline{\mathscr{Q}}_k^\sigma$.

Let $\Pi_{-1}^r := I$.

for $k = 0, 1, \dots, m$;
do
- Compute reduced system matrices: $\widehat{A}_k := \Pi_k^l A_k \Pi_k^r$, $\widehat{B}_k := \Pi_k^l B_k$, $\widehat{C}_k := C_k \Pi_k^r$.
- Compute reduced jump matrices: $\widehat{J}_k^x := \Pi_k^l J_k^x \Pi_{k-1}^r$, $\widehat{J}_k^v := \Pi_k^l J_k^v$.

end

The effectiveness of Algorithm 2 is illustrated with the following medium size academic example.

Example 13. *We consider the switched ODE with jumps (8a) with switching times $s_0 = t_0 = 0$, $s_1 = 2$, $s_2 = 4$, $s_3 = t_f = 6$ and with mode-dependent state dimensions $n_0 = 50$, $n_1 = 60$, $n_2 = 40$. The subspace* *of possible initial states is assumed to be $\mathscr{X}_0 = \mathbb{R}^5$ (i.e. $n_{-1} = 5$). The continuous input u is assumed to be one-dimensional, whereas the discrete input is two dimensional and given by $v_k = (u(s_k), \dot{u}(s_k))$. The initial values and all entries of the coefficient and jump matrices are chosen randomly with a normal distribution (mean zero and variance one), apart from the A_k-matrices, which are constructed by first randomly creating a diagonal matrix with values uniformly distributed in the interval $[-0.1, 0.1]$ and then a similarity transformation is applied with an orthogonalized random coordinate transformation (in particular, the modes are not asymptotically stable, nevertheless the growth rates are moderate). Algorithm 2 is applied with paramteters $\lambda = 1$ and $\gamma = 0.1$ and as reduction threshold for the singular values of the product of the midpoint Gramians we choose $\varepsilon_k = 10^{-3}$ for each mode k. Algorithm 2 results in a reduced switched ODE with jumps with reduced state-dimensions $\widehat{n}_0 = 8$, $\widehat{n}_1 = 10$, $\widehat{n}_2 = 6$ and the run time on a standard office laptop is about 30ms. For the input $u(t) = \cos(t)$ the correspondent outputs of the original system and of the reduced system together with the relative error are shown in Figure 2. Clearly, the output of the reduced system approximates the original output very well (including the jumps at the switching times): the relative error is between 0.1% (for the first mode) and 1% (for the last mode).*

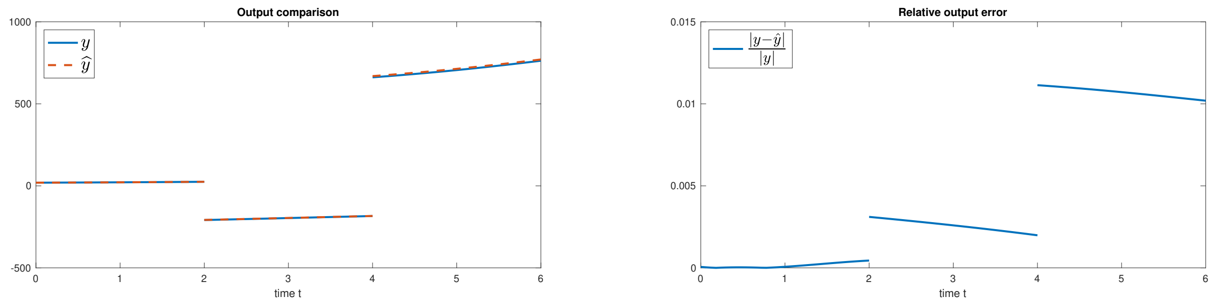

Figure 2. Output comparison (left part) and relative output error (right part) for Example 13. The Matlab scripts to produce these simulations are available at [45].

4.3 Overall midpoint balanced truncation for switched DAEs

In the original notation of (1) the midpoint balanced truncation method for switched DAEs is now summarized as follows, see also Figure 3.

Step 1 (first part of Algorithm 1): For the regular switched DAE (1) obtain an equivalent switched ODE with jumps and impulses in the form (6) via the Wong sequences as described in Section 2.1.

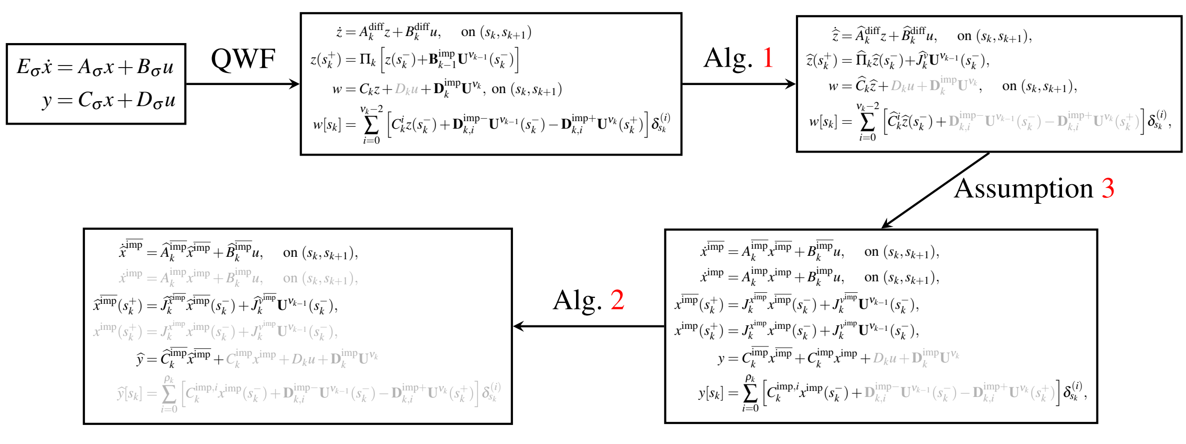

Figure 3. Overview of steps necessary for carrying out model reduction for switched DAEs. Parts which remain unchanged in a step (and its subsequent steps) are displayed in gray.

Step 2 (second part of Algorithm 1): Reduce (6) via Algorithm 1 to obtain (7).

Step 3: Check Assumption 3 for (7). In particular, try to find a family of $\widehat{A}_k^{\mathrm{diff}}$-invariant subspaces $\mathscr{X}_k^{\overline{\mathrm{imp}}}$ and $\mathscr{X}_k^{\mathrm{imp}}$ such that

- $\mathscr{X}_k^{\overline{\mathrm{imp}}} \oplus \mathscr{X}_k^{\mathrm{imp}} = \mathbb{R}^{\widehat{n}_k}$,
- $\mathscr{X}_k^{\overline{\mathrm{imp}}} \subseteq \ker[\widehat{C}_k^0 / \widehat{C}_k^1 / \widehat{C}_k^1 / \dots / \widehat{C}_k^{\rho_k}]$,
- $\widehat{\Pi}_k \mathscr{X}_{k-1}^{\overline{\mathrm{imp}}} \subseteq \mathscr{X}_{k-1}^{\overline{\mathrm{imp}}}$ and $\widehat{\Pi}_k \mathscr{X}_{k-1}^{\mathrm{imp}} \subseteq \mathscr{X}_{k-1}^{\mathrm{imp}}$.

If such a family of subspaces exists, apply a corresponding coordinate transformation T_k to obtain (14).

Step 4 (Algorithm 2): Calculate midpoint Gramians $\overline{\mathscr{P}}_k^{\sigma}$, $\overline{\mathscr{Q}}_k^{\sigma}$ for the switched ODE (obtained from (14))

$$
\begin{aligned}
\dot{x}^{\overline{\mathrm{imp}}} &= A_k^{\overline{\mathrm{imp}}} x^{\overline{\mathrm{imp}}} + B_k^{\overline{\mathrm{imp}}} u, \quad \text{on } (s_k, s_{k+1}), \quad x^{\overline{\mathrm{imp}}}(t_0^-) = 0, \\
x^{\overline{\mathrm{imp}}}(s_k^+) &= J_k^{x^{\overline{\mathrm{imp}}}} x^{\overline{\mathrm{imp}}}(s_k^-), \\
y^{\overline{\mathrm{imp}}} &= C_k^{\overline{\mathrm{imp}}} x^{\overline{\mathrm{imp}}},
\end{aligned}
\tag{18}
$$

according to [21, Alg. 2]. Furthermore, consider the discrete-time midpoint dynamical system (15) obtained from (14) and calculate the corresponding discrete-time Gramian $\mathscr{P}_k^m$. Finally, choose $\lambda > 0$ and let $\mathscr{P}_k^{\lambda} := \overline{\mathscr{P}}_k + \lambda \mathscr{P}_k^m$. In order to be able to continue, it has to be assumed that $\mathscr{P}_k^{\lambda}$ and $\overline{\mathscr{Q}}_k$ are invertible for all k. Finally, calculate the left- and right reduction projectors Π_k^l and Π_k^r based on classical balanced truncation, see e.g. [21, Alg. 1]. The (mode-dependent) reduction size r_k can either be prespecified or can implicitely be given by providing a (mode-dependent) threshold ε_k for the singular values of the Gramians. The overall reduced switched system is then given by

$$
\begin{aligned}
\dot{\widehat{x}}^{\overline{\mathrm{imp}}} &= \widehat{A}_k^{\overline{\mathrm{imp}}} \widehat{x}^{\overline{\mathrm{imp}}} + \widehat{B}_k^{\overline{\mathrm{imp}}} u, \quad \text{on } (s_k, s_{k+1}), \quad \widehat{x}^{\overline{\mathrm{imp}}}(t_0^-) = x_0^{\overline{\mathrm{imp}}} \in \mathscr{X}_0^{\overline{\mathrm{imp}}}, \\
\dot{x}^{\mathrm{imp}} &= A_k^{\mathrm{imp}} x^{\mathrm{imp}} + B_k^{\mathrm{imp}} u, \quad \text{on } (s_k, s_{k+1}), \quad x^{\mathrm{imp}}(t_0^-) = x_0^{\mathrm{imp}} \in \mathscr{X}_0^{\mathrm{imp}}, \\
\widehat{x}^{\overline{\mathrm{imp}}}(s_k^+) &= \widehat{J}_k^{x^{\overline{\mathrm{imp}}}} \widehat{x}^{\overline{\mathrm{imp}}}(s_k^-) + \widehat{J}_k^{v^{\overline{\mathrm{imp}}}} \mathbf{U}^{\nu_{k-1}}(s_k^-), \\
x^{\mathrm{imp}}(s_k^+) &= J_k^{x^{\mathrm{imp}}} x^{\mathrm{imp}}(s_k^-) + J_k^{v^{\mathrm{imp}}} \mathbf{U}^{\nu_{k-1}}(s_k^-), \\
\widehat{y} &= \widehat{C}_k^{\overline{\mathrm{imp}}} \widehat{x}^{\overline{\mathrm{imp}}} + C_k^{\mathrm{imp}} x^{\mathrm{imp}} + D_k u + \mathbf{D}_k^{\mathrm{imp}} \mathbf{U}^{\nu_k}, \\
\widehat{y}[s_k] &= \sum_{i=0}^{\rho_k} C_k^{\mathrm{imp},i} x^{\mathrm{imp}}(s_k^-) \delta_{s_k}^{(i)} + \sum_{i=0}^{\nu_k - 2} \Big(\mathbf{D}_{k,i}^{\mathrm{imp}-} \mathbf{U}^{\nu_{k-1}}(s_k^-) - \mathbf{D}_{k,i}^{\mathrm{imp}+} \mathbf{U}^{\nu_k}(s_k^+) \Big) \delta_{s_k}^{(i)},
\end{aligned}
$$

where $\widehat{A}_k^{\overline{\mathrm{imp}}} := \Pi_k^l A_k^{\overline{\mathrm{imp}}} \Pi_k^r$, $\widehat{B}_k^{\overline{\mathrm{imp}}} := \Pi_k^l B_k^{\overline{\mathrm{imp}}}$, $\widehat{J}_k^{x^{\overline{\mathrm{imp}}}} = \Pi_k^l J_k^{x^{\overline{\mathrm{imp}}}} \Pi_{k-1}^r$, $\widehat{J}_k^{v^{\overline{\mathrm{imp}}}} = \Pi_k^l J_k^{v^{\overline{\mathrm{imp}}}}$, where $\Pi_{-1}^r := I$.

5 Numerical considerations and conclusions

While the provided academic examples illustrate that the proposed model reduction approach for switched DAEs in principle is working, there are the following major challenges for using the proposed approach in a more realistic large scale scenario:

I Precise rank decisions required for Algorithm 1.
II Assumption 3 is not constructive.
III Large-scale matrix-exponentials are required for Algorithm 2.
IV Switching signal must be known a-priori.

We will discuss each point in more details in the following.

Challenge I

Algorithm 1 heavily relies on calculating certain subspaces and corresponding coordinate transformations. The Matlab implementation of Algorithm 1 provided at [45] resolves this issue by carrying out all calculations with exact arithmetics (via the symbolic toolbox). This severely limits the size of the systems for which Algorithm 1 can be used and hence contradicts the original intended usage in model reduction. However, an accurate identification of the relevant subspaces is crucial to adequately capture the jumping behaviors as well as the impulsive outputs. Nevertheless, we believe that there are two approaches to resolve this problem: The first (obvious) approach is the development of numerically robust methods to carry out the used subspace operations (e.g. preimages, subspace addition, subspace intersection, ...) e.g. via utilizing suitable singular value decompositions. The second approach would be to utilize the sparsity which is usually present in the coefficient matrices; it then may still possible to continue using exact arithmetics for much larger systems. It may also be possible to combine both approaches. In general, we believe that for specific large-scale problems it is always necessary to develop some tailor-made algorithm exploiting domain-specific knowledge about the structure of the problem; we see our main contribution to provide the starting point for such a specific implementation.

Challenge II

Assumption 3 can in general not be easily checked because suitable coordinate transformations T_k must be found. While parts of the coordinate transformation can be obtained by constructing certain invariant subspaces within other subspaces (for which algorithms exists), there is no general method to obtain the complete coordinate transformation. However, the underlying reason for requiring Assumption 3 was the limitation of not changing the impulsive output. However, if an approximation of the output Dirac impulses is permitted in the application domain considered, Assumption 3 can be relaxed or even completely omitted, but it is not straightforward how to quantify an error in the impulsive part and how the (midpoint) balanced truncation approach needs to be adjusted. Similar as in Challenge I it may be necessary to exploit domain-specific knowledge to resolve this issue.

Challenge III

The calculation of the full size matrix exponential which is used in the calculation of the midpoint reachability and observability Gramian as well as in the discrete-time reachability Gramian is not feasible for very large scale systems (because this is equivalent to solving a corresponding large scale linear ODE); furthermore, the calculation of the midpoint Gramians also involve the solution of a Lyapunov equation. This is however only an issue for very large scale systems, for system orders of up to one thousand the standard Matlab implementation `expm` and `lyap` provide solutions in less then a second. There is however another issue that the numerically calculated Gramians are not always positive definite, which then results in a failure to obtain the corresponding Cholesky decomposition, this problem already occurs for rather small model orders and in the provided Matlab implementation [45] this is resolved by adding a small multiple of the identity matrix to the Gramians before carrying out the Cholesky decomposition.

Challenge IV

For the model-reduction approach it is required to know the whole switching signal in advance, which in many application may be an unrealistic assumption, in particular, when the switching signal itself is seen as an input signal or is the result of faults. This is a structural limitation of our approach, because by considering the switching signal as a given time-variance the overall system is still linear and hence model reduction can utilize certain subspaces and linear methods. In the context of model reduction for ordinary switched systems other approaches have been discussed in our previous works [22, 21], however, we are not aware of any model reduction approaches for switched DAEs where the switching signal is considered an input or unknown. From our previous work on model reduction for switched ordinary system it is clear that the obtained reduced model which does not take into account the switching signal will in general be much bigger than the reduced model for a specific switching signal. Furthermore, our approach only needs exact knowledge of the mode sequence and not of the switching times; in fact, Algorithm 1 does not depend at all on the switching times, whereas Algorithm 2 depends on the switching times (in fact on the switching durations), but small variations in the switching times will only result in small variations in the Gramians and hence the reduced models are not expected to change much if the switching times are not exactly known.

Conclusions

Switched DAEs naturally occur when modelling large networks with changing topology, hence obtaining a reduced model capturing the essential input-output behaviour is highly relevant. To our best knowledge we are the first to propose a model reduction method for switched DAEs which fully takes into account the jumps and Dirac impulses induced by the switches. While the two main ingredients (Algorithms 1 and 2) have already been reported in similar form in our previous work on model reduction for switched ordinary systems, the utilizations in the context of switched DAEs was nontrivial; in particular, the presence of input-depend jumps required the consideration of an additional discrete-time reachability Gramian. While Algorithm 1 is provably correct with no output error, we haven't been able to obtain provable error bounds for Algorithm 2, instead we rely on the intuition that the midpoint reachability and observability Gramians encode which states are difficult to reach/observe. It is the topic of future research to investigate whether there are provable error bounds. Furthermore, resolving the above mentioned numerical challenges also need to be addressed when our method is utilized for specific large scale applications.

Funding

This work was partially supported by the NWO Vidi-grant 639.032.733 and was mainly carried out while the first author was at the University of Groningen.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Underlying and related material

The Matlab scripts implementing Algorithms 1 and 2 and which were used to produce the simulations are openly available at [45].

Author contributions

Both authors contributed equally to all aspects of the manuscript.

Competing interests

The authors declare that they have no competing interests.

References

[1] U. M. Ascher, L. R. Petzold. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Philadelphia: SIAM Publications, 1998.

[2] P. Benner, T. Stykel. "Model order reduction for differential-algebraic equations: a survey". *Surveys in Differential-Algebraic Equations IV*. Ed. by A. Ilchmann, T. Reis. Cham: Springer-Verlag, 2017:107–160. DOI: 10.1007/978-3-319-46618-7_3.

[3] T. Berger, A. Ilchmann, S. Trenn. The quasi-Weierstraß form for regular matrix pencils. *Linear Algebra Appl.* 2012;436(10):4052–4069. DOI: 10.1016/j.laa.2009.12.036.

[4] T. Berger, T. Reis, S. Trenn. "Observability of linear differential-algebraic systems: a survey". Surveys in Differential-Algebraic Equations IV. Ed. by A. Ilchmann, T. Reis. Differential-Algebraic Equations Forum. Berlin-Heidelberg: Springer-Verlag, 2017:161–219. DOI: 10.1007/978-3-319-46618-7_4.

[5] K. E. Brenan, S. L. Campbell, L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. Amsterdam: North-Holland, 1989. DOI: 10.1137/1.9781611971224.

[6] S. L. Campbell. *Singular Systems of Differential Equations I*. New York: Pitman, 1980.

[7] S. L. Campbell. *Singular Systems of Differential Equations II*. New York: Pitman, 1982.

[8] J. D. Cobb. Controllability, observability and duality in singular systems. *IEEE Trans. Autom. Control*. 1984;29:1076–1082. DOI: 10.1109/TAC.1984.1103451.

[9] J. D. Cobb. Descriptor variable systems and optimal state regulation. *IEEE Trans. Autom. Control*. 1983;28:601–611. DOI: 10.1109/TAC.1983.1103283.

[10] L. Dai. *Singular Control Systems*. Lecture Notes in Control and Information Sciences 118. Berlin: Springer-Verlag, 1989. DOI: 10.1007/BFb0002475.

[11] M. Diehl, H. G. Bock, J. P. Schlöder, R. Findeisen, Z. Nagy, F. Allgöwer. Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. *Journal of Process Control*. 2002;12:577–585. DOI: 0.1016/S0959-1524(01)00023-3.

[12] B. Dziurla, R. W. Newcomb. Nonregular semistate systems: examples and input-output pairing. Proc. 26th IEEE Conf. Decis. Control. Vol. 26. New York: IEEE Press, 1987:1125–1126. DOI: 10.1109/CDC.1987.272580.

[13] E. Eich-Soellner, C. Führer. *Numerical Methods in Multibody Dynamics*. Stuttgart: Teubner, 1998. DOI: 10.1007/978-3-663-09828-7.

[14] C. Führer, B. J. Leimkuhler. Numerical solution of differential-algebraic equations for constrained mechanical motion. *Numer. Math.* 1991;59:55–69. DOI: 10.1007/BF01385770.

[15] F. R. Gantmacher. *The Theory of Matrices (Vol. I & II)*. New York: Chelsea, 1959.

[16] A. H. W. T. Geerts. Solvability conditions, consistency and weak consistency for linear differential-algebraic equations and time-invariant linear systems: The general case. *Linear Algebra Appl.* 1993;181:111–130. DOI: 10.1016/0024-3795(93)90027-L.

[17] P. M. Gresho. Incompressible fluid dynamics: Some fundamental formulation issues. *Annu. Rev. Fluid Mech.* 1991;23:413–453. DOI: 10.1146/annurev.fl.23.010191.002213.

[18] T. B. Gross, S. Trenn, A. Wirsen. Solvability and stability of a power system DAE model. *Syst. Control Lett.* 2016;29:12–17. DOI: 10.1016/j.sysconle.2016.08.003.

[19] M. S. Hossain. "Reduced realizations and model reduction for switched linear systems: a time-varying approach". PhD thesis. Groningen, Netherlands: University of Groningen, 2022. DOI: 10.33612/diss.222461251.

[20] M. S. Hossain, S. Trenn. A weak Kalman decomposition approach for reduced realizations of switched linear systems. IFAC-PapersOnLine. Vol. 55. 20. MATHMOD 2022, Vienna, Austria. July 27, 2022:157–162. DOI: 10.1016/j.ifacol.2022.09.088.

[21] M. S. Hossain, S. Trenn. Midpoint based balanced truncation for switched linear systems with known switching signal. *IEEE Trans. Autom. Control*. 2024;69(1). DOI: 10.1109/TAC.2023.3269721.

[22] M. S. Hossain, S. Trenn. Reduced realization for switched linear systems with known mode sequence. *Automatica*. 2023;154(111065):1–9. DOI: 10.1016/j.automatica.2023.111065.

[23] P. Kunkel, V. Mehrmann. *Differential-Algebraic Equations. Analysis and Numerical Solution*. Zürich, Switzerland: EMS Publishing House, 2006. DOI: 10.4171/017.

[24] F. L. Lewis. A survey of linear singular systems. *IEEE Proc. Circuits, Systems and Signal Processing*. 1986;5(1):3–36. DOI: 10.1007/BF01600184.

[25] D. Liberzon. *Switching in Systems and Control*. Systems and Control: Foundations and Applications. Boston: Birkhäuser, 2003. DOI: 10.1007/978-1-4612-0017-8.

[26] D. Liberzon, S. Trenn. Switched nonlinear differential algebraic equations: solution theory, Lyapunov functions, and stability. *Automatica*. 2012;48(5):954–963. DOI: 10.1016/j.automatica.2012.02.041.

[27] P. Lötstedt, L. R. Petzold. Numerical solution of nonlinear differential equations with algebraic constraints I: Convergence results for backward differentiation formulas. *Math. Comp.* 1986;46(174):491–516. DOI: 10.1090/S0025-5718-1986-0829621-X.

[28] D. G. Luenberger. Dynamic equations in descriptor form. *IEEE Trans. Autom. Control*. 1977;22:312–321. DOI: 10.1109/TAC.1977.1101502.

[29] D. G. Luenberger, A. Arbel. Singular dynamic Leontief systems. *Econometrica*. 1977;45:991–995. DOI: 10.2307/1912686.

[30] E. Mostacciuolo, S. Trenn, F. Vasca. Averaging for switched DAEs: convergence, partial averaging and stability. *Automatica*. 2017;82:145–157. DOI: 10.1016/j.automatica.2017.04.036.

[31] R. W. Newcomb. The semistate description of nonlinear time-variable circuits. *IEEE Trans. Circuits Syst.* 1981;CAS-28:62–71. DOI: 10.1109/TCS.1981.1084908.

[32] C. C. Pantelides. The consistent initialization of differential-algebraic systems. *SIAM J. Sci. Stat. Comput.* 1988;9:213–231. DOI: 10.1137/0909014.

[33] P. J. Rabier, W. C. Rheinboldt. "Theoretical and numerical analysis of differential-algebraic equations". Handbook of Numerical Analysis. Ed. by P. G. Ciarlet, J. L. Lions. Vol. VIII. Amsterdam, The Netherlands: Elsevier Science, 2002:183–537. DOI: 10.1016/S1570-8659(02)08004-3.

[34] H. H. Rosenbrock. *State Space and Multivariable Theory*. New York, NY: John Wiley and Sons Inc., 1970.

[35] S. S. K. Sajja, M. Corless, E. Zeheb, R. N. Shorten. On dimensionality reduction and the stability of a class of switched descriptor systems. *Automatica*. 2013;49(6):1855–1860. DOI: 10.1016/j.automatica.2013.02.056.

[36] W. Schiehlen. Multibody system dynamics: Roots and perspectives. *Multibody System Dynamics*. 1997;1:149–188. DOI: 10.1023/A:1009745432698.

[37] L. Schwartz. *Théorie des Distributions I,II*. Publications de l'institut de mathématique de l'Universite de Strasbourg IX,X. Paris: Hermann, 1950, 1951.

[38] A. Tanwani, S. Trenn. Detectability and observer design for switched differential algebraic equations. *Automatica*. 2019;99:289–300. DOI: 10.1016/j.automatica.2018.10.043.

[39] A. Tanwani, S. Trenn. On observability of switched differential-algebraic equations. Proc. 49th IEEE Conf. Decis. Control, Atlanta, USA. 2010:5656–5661. DOI: 10.1109/CDC.2010.5717685.

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

[41] S. Trenn. Distributional restriction impossible to define. *Examples and Counterexamples*. Nov. 30, 2021;1(100023). open access:1–4. DOI: 10.1016/j.exco.2021.100023.

[42] S. Trenn. "Solution concepts for linear DAEs: a survey". Surveys in Differential-Algebraic Equations I. Ed. by A. Ilchmann, T. Reis. Differential-Algebraic Equations Forum. Berlin-Heidelberg: Springer-Verlag, 2013:137–172. DOI: 10.1007/978-3-642-34928-7_4.

[43] S. Trenn. "Stability of switched DAEs". Hybrid Systems with Constraints. Ed. by J. Daafouz, S. Tarbouriech, M. Sigalotti. Automation - Control and Industrial Engineering Series. London: Wiley, 2013:57–83. DOI: 10.1002/9781118639856.ch3.

[44] S. Trenn. "Switched differential algebraic equations". Dynamics and Control of Switched Electronic Systems - Advanced Perspectives for Modeling, Simulation and Control of Power Converters. Ed. by F. Vasca, L. Iannelli. London: Springer-Verlag, 2012. Chap. 6:189–216. DOI: 10.1007/978-1-4471-2885-4_6.

[45] S. Trenn, M. S. Hossain. *MRswDAE - model reduction for switched DAEs - matlab scripts*. 2023. DOI: 10.5281/zenodo.8133789.

[46] S. Trenn, F. R. Wirth. Linear switched DAEs: Lyapunov exponent, converse Lyapunov theorem, and Barabanov norm. Proc. 51st IEEE Conf. Decis. Control, Maui, USA. 2012:2666–2671. DOI: 10.1109/CDC.2012.6426245.

[47] G. C. Verghese, B. C. Levy, T. Kailath. A generalized state-space for singular systems. *IEEE Trans. Autom. Control*. 1981;26(4):811–831. DOI: 10.1109/TAC.1981.1102763.

[48] P. Wijnbergen, S. Trenn. Optimal control of DAEs with unconstrained terminal costs. Proc. 60th IEEE Conf. Decis. Control, Austin, USA. 2021:5275–5280. DOI: 10.1109/CDC45484.2021.9682950.

[49] K.-T. Wong. The eigenvalue problem $\lambda Tx + Sx$. *J. Diff. Eqns.* 1974;16:270–280. DOI: 10.1016/0022-0396(74)90014-X.

Md. Sumon Hossain
Dept. of Mathematics & Physics
North South University
Dhaka-1229, Bangladesh
sumon.hossain@northsouth.edu

Stephan Trenn
Bernoulli Institute for Math, CS and AI
University of Groningen
Nijenborgh 9
9747 AG Groningen, Netherlands
s.trenn@rug.nl