Impulse-free linear quadratic optimal control of switched differential algebraic equations

Paul Wijnbergen, Stephan Trenn*

KTH Royal Institute for Technology, School of Electrical Engineering and Computing Science, Stockholm
Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence
University of Groningen, Groningen, the Netherlands

This work was supported by the NWO Vidi-grant 639.032.733 while the first author was at the University of Groningen.

Abstract

In this paper the finite horizon linear quadratic regulator (LQR) problem for switched linear differential algebraic equations is studied. It is shown that for switched DAEs with a switching signal that induces locally finitely many switches the problem can be solved by solving several LQR problems for non-switched DAE recursively. First, it is shown how to solve the non-switched problems for index-1 DAEs followed by an extension of the results to higher index DAEs. The resulting optimal control can be computed based on the solution of a Riccati differential equation expressed in terms of the differential system matrices. The paper concludes with the extension of the results to the LQR problem for general switched DAEs.

Keywords: Switched systems, Differential Algebraic Equations, Impulse-controllability, Geometric control

1. Introduction

In this paper, we aim to find necessary and sufficient conditions for the existence of an input that solves the finite horizon linear quadratic regulator problem for switched differential-algebraic equations.

Problem 1. [LQR for switched DAEs] Find an input u that minimizes

$$J(x_0, u, t_0) = \int_{t_0}^{t_f} \|y(t)\|^2 \,\mathrm{d}t + x(t_f^-)Px(t_f^-), \tag{1}$$

$$\begin{aligned} \text{s.t.} \quad E_\sigma \dot{x} &= A_\sigma x + B_\sigma u, && \text{(2a)}\\ y &= C_\sigma x + D_\sigma u, && \text{(2b)}\\ x(t_0^-) &= x_0, && \text{(2c)}\\ x(t_f^-) &\in \mathcal{V}^{\text{end}}, && \text{(2d)} \end{aligned}$$

where $\sigma : [t_0, t_f) \to \mathbb{N}$ is a given piecewise constant switching signal, x is the state, the matrices $E_p, A_p \in \mathbb{R}^{n\times n}$ form a regular matrix pair, (i.e., $\det(sE - A)$ is not indentically zero), $B_p \in \mathbb{R}^{n\times m}$, $C_p \in \mathbb{R}^{q\times n}$ and $D_p \in \mathbb{R}^{q\times m}$, $p \in \mathbb{N}$, $P = P^\top \in \mathbb{R}^{n\times n}$ is some symmetric positive semi-definite matrix and $\mathcal{V}^{\text{end}} \subseteq \mathbb{R}^n$ is some subspace.

Switched differential algebraic equations (swDAEs) of the form (2a) arise naturally when modeling physical systems with certain algebraic constraints on the state variables; examples of applications of non-switched DAEs in electrical circuits (with distributional solutions) can be found, e.g., in [1]. For non-switched DAEs, these constraints are often eliminated such that the system is described by ordinary differential equations. However, in the case of switched systems, the elimination process of the constraints is in general different for each individual mode and therefore there does in general not exist a description as a switched ODE with a common state variable. This problem can be overcome by studying switched DAEs directly.

In the context of linear systems the linear quadratic regulator (LQR) problem on both the finite and infinite horizon has been studied extensively, see [2, 3, 4, 5, 6] for results on ODEs and [7, 8, 9, 10, 11, 12, 13, 14, 15, 16] for DAEs. Most recent studies regarding the optimal control problem for DAEs focus on finding solutions based on the Lure inequality or an extension of the Kalman-Yakubovich-Popov lemma [17, 18, 19]; further results have been opbtained in the context of model predictive control [20, 21, 22, 23]. For switched differential algebraic equations it seems that so far only qualitative properties such as controllability, stabilizability [24, 25, 26, 27, 28, 29, 30, 31], and observability have been studied [32, 33, 34, 35, 36, 37, 38]. To the best of the authors knowledge quantitative properties such as optimal control have not been studied for switched DAEs. This paper aims to close this gap in

*Corresponding author
Email addresses: wijnberg@kth.se (Paul Wijnbergen), s.trenn@rug.nl (Stephan Trenn)

the literature.

As trajectories of switched DAEs generally exhibit jumps (or even impulses), which may exclude classical solutions from existence, the *piece-wise smooth distributional solution framework* introduced in [39] is adopted, i.e. $(x,u) \in \mathbb{D}_{\text{pw}\mathcal{C}^\infty}^{n+m}$, where $\mathbb{D}_{\text{pw}\mathcal{C}^\infty}$ denotes the space of piece-wise smooth distributions. In particular, Problem 1 is considered in a piece-wise smooth distributional setup. Since within this setup, the integral over the norm squared of a Dirac impulse is not well defined, it follows directly that in order to have finite cost, the output (2b) needs to be impulse-free. Focusing on solutions that result in an impulse-free output, we denote the output as a piece-wise continuous function, whereas it is actually a distribution.

We consider Problem 1 under the assumption that the switching signal does not induce chattering behavior, i.e., we assume that it induces locally finitely many switches. Since the switching signal could still induce infinitely many switches, which is troublesome for solving the problem in finitely many steps, we consider the bounded interval $[t_0, t_f)$. In this interval thus only finitely many switches are present.

For many applications, it is of interest to extend an optimal solution in an impulse-free way on the interval $[t_f, \infty)$. This is the case for example in choosing suitable terminal costs if the LQR problem is to be solved on a receding horizon. To allow for such extensions, we impose the subspace endpoint constraint (2d) to the state at t_f^-. As we will show, this subspace endpoint constraint fits naturally in the LQR problem for switched DAEs as there exists a solution to Problem 1 if and only if the initial value x_0 is contained in a certain subspace.

The remainder of the paper is structured as follows. First mathematical notation and preliminaries are introduced in Section 2. Then the approach to solving Problem 1 is formulated in Section 3 and the main result is presented. In Section 4 necessary and sufficient conditions for solvability of Problem 1 for non-switched DAEs of index-1 presented and it is shown how to generalize these results to arbitrary index-DAEs in Section 5.

2. Mathematical preliminaries

In this section we recall some notation and properties related to the non-switched DAE

$$E\dot{x} = Ax + Bu. \tag{3}$$

In the following, we call a matrix pair (E,A) and the associated DAE (3) *regular* iff the polynomial $\det(sE - A)$ is not the zero polynomial. Recall the following result on the *quasi-Weierstrass form*[40].

Proposition 1. *A matrix pair $(E,A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ is regular if, and only if, there exist invertible matrices $S, T \in \mathbb{R}^{n\times n}$ such that*

$$(SET, SAT) = \left(\left[\begin{smallmatrix} I & 0 \\ 0 & N \end{smallmatrix}\right], \left[\begin{smallmatrix} J & 0 \\ 0 & I \end{smallmatrix}\right]\right), \tag{4}$$

where $J \in \mathbb{R}^{n_1 \times n_1}$, $0 \leqslant n_1 \leqslant n$, is some matrix and $N \in \mathbb{R}^{n_2\times n_2}$, $n_2 := n - n_1$, is a nil-potent matrix of order $\nu \in \mathbb{N}$. In particular, ν is referred to as the index of (3).

The matrices S and T can be calculated by using the so-called *Wong sequences* [40, 41]:

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

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

$$\mathcal{V}^* := \bigcap_i \mathcal{V}_i, \qquad \mathcal{W}^* := \bigcup \mathcal{W}_i.$$

For any full rank matrices V, W with $\operatorname{im} V = \mathcal{V}^*$ and $\operatorname{im} W = \mathcal{W}^*$, the matrices $T := [V, W]$ and $S := [EV, AW]^{-1}$ are invertible and (4) holds. Based on the Wong sequences we define the following projector and selectors.

Definition 2. Consider the regular matrix pair (E,A) with corresponding quasi-Weierstrass form (4). The *consistency projector* of (E,A) is given by

$$\Pi_{(E,A)} := T \left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right] T^{-1}, \tag{5}$$

the *differential* and *impulse selector* are given by

$$\Pi_{(E,A)}^{\text{diff}} := T \left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right] S, \; \Pi_{(E,A)}^{\text{imp}} := T \left[\begin{smallmatrix} 0 & 0 \\ 0 & I \end{smallmatrix}\right] S. \tag{6}$$

In all three cases, the block structure corresponds to the block structure of the quasi-Weierstrass form. Furthermore, we define

$$\begin{aligned} A^{\text{diff}} &:= \Pi_{(E,A)}^{\text{diff}} A, \qquad E^{\text{imp}} := \Pi_{(E,A)}^{\text{imp}} E, \\ B^{\text{diff}} &:= \Pi_{(E,A)}^{\text{diff}} B, \qquad B^{\text{imp}} := \Pi_{(E,A)}^{\text{imp}} B. \end{aligned}$$

A classical solution, i.e., a differentiable or locally integrable solution, to (3) in terms of these matrices yields $x = x^{\text{diff}} + x^{\text{imp}}$,where x^{diff} and x^{imp} satisfy

$$\dot{x}^{\text{diff}} = A^{\text{diff}} x^{\text{diff}} + B^{\text{diff}} u, \quad x^{\text{diff}}(t_0^-) = \Pi x_0, \tag{7}$$

$$x^{\text{imp}} = -\sum_{i=0}^{\nu-1} (E^{\text{imp}})^i B^{\text{imp}} u^{(i)}. \tag{8}$$

Observe that for index-1 systems we find $x^{\text{imp}} = -B^{\text{imp}}u$. Note that all the above-defined matrices do not depend on the choice of transformation matrices S and T; they are uniquely determined by the original matrix pair (E,A).

The switched DAE (2a) usually will not have classical solutions. Due to the switching between modes $x^{\text{imp}}(t_i^-) \neq x^{\text{imp}}(t_i^+)$ in general. Consequently, the state can contain jumps or even Dirac impulses. We therefore utilize the piecewise-smooth distributional framework as introduced in [39], i.e., x and u are vectors of piecewise-smooth distributions given by

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

where $\mathcal{C}^\infty_{\mathrm{pw}}$ denotes the space of piecewise-smooth functions, $f_{\mathbb{D}}$ denotes the regular distribution induced by f, δ_t denotes the Dirac impulse with support $\{t\}$ and δ_t' denotes distributional derivative of δ_t. For $D = f_{\mathbb{D}} + \sum_{t\in T} D_t \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ three types of "evaluation at time t" are defined: left side evaluation $D(t^-) := f(t^-)$, right side evaluation $D(t^+) := f(t^+)$ and the impulsive part $D[t] := D_t$ if $t \in T$ and $D[t] = 0$ otherwise.

It can be shown (cf. [42]) that the space $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ can be equipped with a multiplication, in particular, the multiplication of a piecewise-constant function with a piecewise-smooth distribution is well defined and the switched DAE (2a) can be interpreted as an equation within the space of piecewise-smooth distributions. Specifically, restrictions of x and u to intervals, are well defined. Given the notation $x_{\mathcal{I}}$ for the restriction of x to the interval $\mathcal{I} \subseteq \mathbb{R}$, it is shown in [39] that the *initial trajectory problem* (ITP)

$$x_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}, \tag{9a}$$

$$(E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)} + (Bu)_{[t_0,\infty)}, \tag{9b}$$

has a unique solution for any initial trajectory if, and only if, the matrix pair (E, A) is regular. Note that it can be shown that the solution of (9) on $[t_0, \infty)$ is uniquely determined by $x(t_0^-)$, hence it is justified to replace (9a) by $x(t_0^-) = x_0$ for some $x_0 \in \mathbb{R}^n$.

The impulsive part of a solution of (9) is given by

$$x[t_0] = -\sum_{i=0}^{\nu-1} \left(E^{\mathrm{imp}}\right)^{i+1} \left(x(t_0^-) - x(t_0^+)\right) \delta^{(i)}. \tag{10}$$

For a single mode, the concept of impulse-controllable space is defined as follows.

Definition 3. The impulse-controllable space for (3) is given by

$$\mathcal{C}^{\mathrm{imp}}_{(E,A,B)} := \left\{ x_0 \;\middle|\; \begin{array}{l} \exists (x,u) \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n+m} \text{ solving (9)} \\ \text{s.t. } x(t_0^-) = x_0 \text{ and } (x,u)[t_0] = 0. \end{array} \right\}.$$

Furthermore, the DAE is called impulse-controllable if all initial values are impulse-controllable, i.e., $\mathcal{C}^{\mathrm{imp}} = \mathbb{R}^n$.

It can be shown (see e.g. [31, Lem. 13]), that

$$\mathcal{C}^{\mathrm{imp}}_{(E,A,B)} = \operatorname{im} \Pi_{(E,A)} + \langle E^{\mathrm{imp}}, B^{\mathrm{imp}} \rangle + \ker E, \tag{11}$$

where

$$\langle E^{\mathrm{imp}}, B^{\mathrm{imp}} \rangle := \operatorname{im}[B^{\mathrm{imp}}, E^{\mathrm{imp}} B^{\mathrm{imp}}, \ldots, (E^{\mathrm{imp}})^{n-1} B^{\mathrm{imp}}].$$

Lemma 4 ([43, Prop. 3])**.** *The regular DAE* (3) *is impulse controllable if and only if*

i) $\operatorname{im} E + A \ker E + \operatorname{im} B = \mathbb{R}^n$,

ii) There exists a matrix L such that the closed loop with feedback $u = Lx$ results in an index-1 matrix pair $(E, A + BL)$; the latter can be characterized by $\operatorname{im} E + (A + BL) \ker E = \mathbb{R}^n$.

We conclude this section with an explicit definition of a solution to the switched DAE (2a).

Definition 5. A distribution $(x, u) \in \mathbb{D}^{(n+m)}_{\mathrm{pw}\mathcal{C}^\infty}$ is called a solution to the switched DAE (2a) for a given right continuous switching signal σ with switching times $t_0, t_1, ...,$, if (x, u) considered on each interval $[t_k, t_{k+1})$ is a local (distributional) solution to ITP (9) on $[t_k, t_{k+1})$ with $E = E_{\sigma(t_k)}$, $A = A_{\sigma(t_k)}$ and $B = B_{\sigma(t_k)}$, where the initial condition $x(t_k^-)$ is either given by (2c) or by the final value of the solution from the previous interval.

Since by assumption each matrix pair (E_p, A_p) is regular, it follows that each local ITP is uniquely solvable and hence the overall switched DAE is uniquely solvable for any given input and any given initial value x_0.

3. Problem Formulation and approach

As mentioned in the introduction, we consider Problem 1 in a distributional setup. As such, the aim is to find a distribution $u \in \mathbb{D}^m_{\mathrm{pw}\mathcal{C}^\infty}$ that minimizes a quadratic cost functional subject to a switched differential-algebraic equation. In order to utilize the distributional solution framework and to avoid technical difficulties in general, we only consider systems with a switching signal from the following class

$$\mathcal{S} := \left\{ \sigma : \mathbb{R} \to \mathbb{N} \;\middle|\; \begin{array}{l} \sigma \text{ is right continuous with a} \\ \text{locally finite number of jumps} \end{array} \right\},$$

i.e., we exclude an accumulation of switching times (see [39]). Since a bounded interval $[t_0, t_f)$ is considered in Problem 1, the switching signal thus induces $\mathtt{n} \in \mathbb{N}$ switches on this interval, each occurring at t_k, where $k \in \{1, 2, ..., \mathtt{n}\}$. The switching signal is assumed to be known a priory; in particular, solvability and the solution of Problem 1 depends on the specific switching signal. By appropriately relabeling the matrices we can therefore assume without loss of generality that

$$\sigma(t) = k, \qquad \text{for } t_k \leqslant t < t_{k+1}, \tag{12}$$

where $t_{\mathtt{n}+1} := t_f$.

The switching signal is thus *not* regarded as a control input. Consequently, a switched differential algebraic equation of the form (2a) with a switching signal $\sigma \in \mathcal{S}$ can be regarded as (piecewise-constant) time-varying linear systems. Such systems have a linear solutions space where the sum of solutions is also a solution. Furthermore, the subspace endpoint constraint (2d) is also a linear constraint and hence the sum of solutions satisfying (2d) will also satisfy (2d). Together with the fact that the cost functional (1) is quadratic in the state and input all ingredients are present to prove several important properties of Problem 1. Namely if there exists an input that solves Problem 1 the optimal cost is quadratic in the initial value and the optimal input is linear in the state, i.e., it is a feedback.

Lemma 6. *If there exists an input $u \in (\mathbb{D}_{pw\mathcal{C}^\infty})^m$ that solves Problem 1 then $u(t^+) = F(t)x(t^-)$ for some $F : \mathbb{R} \to \mathbb{R}^{m\times n}$.*

The (technical) proof can be found in the Appendix. Note that we only consider piecewise-smooth solutions, hence $x(t^-) \neq x(t^+)$ for only finitely many $t \in [t_0, t_f)$ and hence we can assume that the input is right-continuous and we can simply write $u(t) = F(t)x(t)$ in the following.

Corollary 7. *If there exists an input that solves Problem 1 then the optimal cost $J(x_0, u, t_0)$ is quadratic in x_0, i.e.,*

$$J(x_0, u, t_0) = x_0^\top K(t_0)x_0,$$

for some $K : \mathbb{R} \to \mathbb{R}^{n\times n}$.

Proof. In the proof of Lemma 6 it was shown that the map $x_0 \mapsto V(x_0, t_0)$ satisfies the parallelogram equality (A.2). Hence it is a (semi-) norm induced by an inner product. Consequently, there exists a (positive semi-definite) matrix $K(t_0) \in \mathbb{R}^{n\times n}$ such that $V(x_0, t_0) = x_0^\top K(t_0)x_0$.

The result of Lemma 6 also leads to the observation that the space of initial values for which Problem 1 is solvable must be a subspace.

Definition 8. The set of initial values for which Problem 1 is solvable on $[t_0, t_f)$ is given by

$$\mathcal{V}^{\text{init}}_{t_0} := \left\{ x_0 \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \exists\, u \text{ that solves Problem 1 on} \\ [t_0, t_f) \text{ satisfying } x(t_0^-) = x_0 \end{array} \right\}.$$

Corollary 9. *The set $\mathcal{V}^{\text{init}}_{t_0}$ is a subspace.*

Proof. Suppose that $x_0, y_0 \in \mathcal{V}^{\text{init}}_{t_0}$. Since the inputs u_{x_0} and u_{y_0}that solve Problem 1 for $x(t_0^-) = x_0$ and $x(t_0^-) = y_0$ are feedbacks, it follows that $\alpha u_{x_0} + \beta u_{y_0}$ is the optimal input that solves Problem 1 for $x(t_0^-) = z_0 = \alpha x_0 + \beta y_0$. Consequently, $z_0 \in \mathcal{V}^{\text{init}}_{t_0}$ and thus $\mathcal{V}^{\text{init}}_{t_0}$ is a subspace.

Let $\mathcal{V}^{\text{init}}_{t_i}$ be the subspace of initial values for which there exists a solution to Problem 1 on the interval $[t_i, t_f)$ with terminal subspace $\mathcal{V}^{\text{end}}$ and terminal cost matrix P. Furthermore, let the optimal cost matrix be given by $K_i(t_i)$, that is, the solution to Problem 1 yields an optimal cost $J(x_i, u, t_i) = x_i^\top K(t_i)x_i$. Then the following lemma is a reformulation of the Bellman principle of optimality.

Lemma 10. *Problem 1 with initial value x_0, terminal cost matrix P and terminal subspace $\mathcal{V}^{\text{end}}$ has a solution on $[t_0, t_f)$ if and only if Problem 1 on the interval $[t_0, t_i)$ with initial value x_0, terminal cost matrix $K_i(t_i)$ and terminal subspace $\mathcal{V}^{\text{init}}_{t_i}$ has a solution.*

Proof. The statement follows directly from the Bellman principle of optimality [44].

As a consequence of Lemma 10, it follows that if we can characterize $\mathcal{V}^{\text{init}}_{t_i}$ and we are able to compute the corresponding cost matrix $K_i(t)$ and corresponding optimal control, we can reduce the problem of solving Problem 1 on the interval $[t_0, t_f)$ to solving Problem 1 on the interval $[t_0, t_i)$. Moreover, by choosing $t_i = t_{\mathtt{n}}$, Problem 1 on the interval $[t_{\mathtt{n}}, t_f)$ reduces to an optimal control problem subject to a non-switched DAE. By applying Lemma 10 recursively and choosing each t_i to be a switching time, it follows that we can solve Problem 1 by solving $\mathtt{n}$ optimal control problems for non-switched DAEs, each defined on the interval $[t_{i-1}, t_i)$, $i \in \{\mathtt{n}, \mathtt{n}-1, ..., 1\}$. Therefore, we will first focus on the following problem.

Problem 2. Find an input $u \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^m$ that minimizes

$$J(x_0, u, t_0) = \int_{t_0}^{t_f} \|y(t)\|^2 \, \mathrm{d}t + x(t_f^-)Px(t_f^-), \tag{13}$$

$$\text{s.t.} \quad E\dot{x} = Ax + Bu, \tag{14a}$$

$$y = Cx + Du, \tag{14b}$$

$$x(t_0^-) = x_0 \in \mathbb{R}^n, \tag{14c}$$

$$x(t_f^-) \in \mathcal{V}^{\text{end}}, \tag{14d}$$

on the interval $[t_0, t_f)$, $x \in \mathbb{D}^n_{\text{pw}\mathcal{C}^\infty}$ is the state, $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{q\times n}$ and $D \in \mathbb{R}^{q\times m}$, $P = P^\top \in \mathbb{R}^{n\times n}$ is some symmetric positive semi-definite matrix and $\mathcal{V}^{\text{end}} \subseteq \mathbb{R}^n$ is some subspace.

We are able to solve Problem 2 without assuming index-1 or impulse controllability, but we will first focus on how to solve Problem 2 for the simpler index-1 case. Then, we will show how to rewrite Problem 2 into an optimal control problem for an index-1 DAE (Problem 3) and how the solvability conditions and the optimal control for the latter provides a solution for the former.

A phenomenon already well known for ODE optimal control problems is that the cost for the input (given by D) needs to be non-singular to avoid impulsive optimal controls. We will make a similar assumption here as well, which reads as follows:

$$\operatorname{rank}\,(D - CB^{\text{imp}}) = m. \tag{15}$$

The main result for the index-1 case of Problem 2 is then given by the forthcoming Theorem 21, which shows that Problem 2 is solvable if, and only if, the initial value x_0 is an element of a subspace $\mathcal{V}^{\text{init}}$ which is defined in terms of the given final subspace $\mathcal{V}^{\text{end}}$. Furthermore, an explicit solution for the optimal control is provided.

Now utilizing Lemma 10 we can define recursively a sequence of subspaces $\mathcal{V}_{t_i}^{\text{init}}$ as the subspace of feasible initial values for Problem 2 for mode i considered on the time interval $[t_i, t_{i+1})$ with final subspace $\mathcal{V}_{t_{i+1}}^{\text{init}}$ (where $\mathcal{V}_{t_{\mathtt{n}+1}}^{\text{init}} := \mathcal{V}^{\text{end}}$). Together with the non-singular input-cost assumption

$$\operatorname{rank}\,(D_p - C_p B_p^{\text{imp}}) = m. \tag{16}$$

for $p \in \{0, 1, ..., \mathtt{n}\}$, we can conclude that Problem 1 with regular, index-1, matrix pairs (E_p, A_p) is solvable if, and only if, $x_0 \in \mathcal{V}_{t_0}^{\text{init}}$. Furthermore, we can explicitly provide the optimal control:

Theorem 11. *Consider the regular, index-1, switched DAE* (2) *satisfying* (16) *for which Problem 1 is solvable, i.e.* $x_0 \in \mathcal{V}_{t_0}^{\text{init}}$. *Then the optimal input is given by*

$$u(t) = -R_{\sigma(t)}^{-1}(B_{\sigma(t)}^{\text{diff}\top} K(t) + S_{\sigma(t)}^\top)\Pi_{\sigma(t)} x(t),$$

where $R_\sigma = (D_\sigma - C_\sigma B_\sigma^{\text{imp}})^\top (D_\sigma - C_\sigma B_\sigma^{\text{imp}})$, $S_\sigma = (D_\sigma - C_\sigma B_\sigma^{\text{imp}})^\top C_\sigma$ *and* Π_i *is a projector resulting from the Wong sequence based on* (E_i, A_i). *Finally,* $K(t)$ *is given by the solution of*

$$\begin{aligned}\dot{K} = &-A_i^{\text{diff}\top} K - K A_i^{\text{diff}} \\ &+ (S_i + K^\top B_i^{\text{diff}}) R_i^{-1} (B_i^{\text{diff}\top} K + S_i^\top) - Q_i,\end{aligned}$$

on $[t_i, t_{i+1})$, *where* $Q_i = C_i^\top C_i$ *and boundary conditions*

$$\begin{aligned}K(t_{i+1}^-) &= \Psi_i^\top K(t_{i+1}^+)\Psi_i, \qquad i \in \{0, 1, ..., \mathtt{n}-1\}, \\ K(t_{\mathtt{n}+1}^-) &= \Psi_{\mathtt{n}}^\top P \Psi_{\mathtt{n}}.\end{aligned}$$

where $\Psi_i = (I - B_i^{\text{imp}} N_i)\Pi_i$, *for some* N_i *that satisfies* $\begin{bmatrix} I & 0 & N_i \end{bmatrix} \ker \mathcal{H}_i = 0$, *with*

$$\mathcal{H}_i = \begin{bmatrix} B_i^{\text{imp}\top} P B_i^{\text{imp}} & B_i^{\text{imp}\top}(I - \Pi_{\mathcal{V}_i^{\text{end}}})^\top \\ (I - \Pi_{\mathcal{V}_i^{\text{end}}}) B_i^{\text{imp}} & 0 \\ -\Pi_i^\top P B_i^{\text{imp}} & -\Pi_i^\top (I - \Pi_{\mathcal{V}_i^{\text{end}}})^\top \end{bmatrix}^\top. \tag{17}$$

and $\Pi_{\mathcal{V}_i^{\text{end}}}$ *is a projector onto the subspace* $\mathcal{V}_i^{\text{end}} = \mathcal{V}_{t_{i+1}}^{\text{init}}$. *Finally, the optimal cost is given by*

$$\min_u J(x_0, u, t_0) = x_0^\top K(t_0) x_0.$$

In the next section we will show how to arrive at the result of Theorem 11 in the case of a non-switched DAE, i.e., the main result for Problem 2. Then we will show how this result can be utilized to obtain necessary and sufficient conditions for solvability of Problem 1. Finally, we will show how these results should be modified to treat arbitrary index (switched) DAEs.

4. Optimal control for non-switched index-1 DAEs

As mentioned previously, we will consider first the optimal control problem for non-switched DAEs, i.e., Problem 2[1]. Furthermore, we will first consider Problem 2 subject to an index-1 DAE. As such, the state can be decomposed as

$$x = x^{\text{diff}} + x^{\text{imp}} = x^{\text{diff}} - B^{\text{imp}} u, \tag{18}$$

where the differential state component satisfies

$$\dot{x}^{\text{diff}} = A^{\text{diff}} x^{\text{diff}} + B^{\text{diff}} u, \qquad x^{\text{diff}}(t_0^-) = \Pi x_0. \tag{19}$$

respectively. As a consequence, we can state the following result which follows from Lemma 6.

Corollary 12. *If there exists an input* $u \in (\mathbb{D}_{pw\mathcal{C}^\infty})^m$ *that solves Problem 2 where the DAE* (14) *is of index-1, then* $u(t) = F(t)x^{\text{diff}}(t)$ *for some* $F : \mathbb{R} \to \mathbb{R}^{m \times n}$.

Proof. After decomposing the state as (18) and considering the ODE dynamics (19) the proof is analogous to the proof of Lemma 6.

4.1. Terminal cost

Decomposing the state as in (18) allows us to express the terminal cost as a quadratic function of the differential state $x^{\text{diff}}(t_f^-)$ and the input. Consequently, an input u with a value $u(t_f^-)$ that minimizes the terminal cost with respect to the resulting $x^{\text{diff}}(t_f^-)$ can be chosen. However, as the terminal cost penalizes the value of u at t_f^- from the left and this value needs to be well-defined, the input u needs to be continuous on at least $[t_f - \varepsilon, t_f)$ for some $\varepsilon > 0$. Therefore altering a solution (x^{diff}, u) such that the output has a desired value at t_f^- will in general influence the running cost. As a result, we can not optimize the running cost and the terminal cost independently of each other. However, the following result shows that the value of the optimal input $u(t_f^-)$ minimizes the terminal cost with respect to the value $x^{\text{diff}}(t_f^-) \in \operatorname{im}\Pi$.

Lemma 13. *Let* u *be an input that solves Problem 2 and let* x^{diff} *be the corresponding optimal trajectory. Denote* $u(t_f^-) = \psi^* \in \mathbb{R}^m$ *and* $x^{\text{diff}}(t_f^-) = \zeta^* \in \operatorname{im}\Pi$. *Then* ψ^* *is a minimizer of the following problem.*

$$\begin{aligned}\min_{\psi \in \mathbb{R}^m} \quad & (\zeta^* - B^{\text{imp}}\psi)^\top P (\zeta^* - B^{\text{imp}}\psi), \\ s.t. \quad & \zeta^* - B^{\text{imp}}\psi \in \mathcal{V}^{\text{end}}.\end{aligned} \tag{20}$$

The proof can be found in the Appendix.

[1] We have already studied this problem with $\mathcal{V}^{\text{end}} = \mathbb{R}^n$ in [45]; but the consideration of a general subspace $\mathcal{V}^{\text{end}}$ increases the difficulty significantly and is crucial for utilizing the result in the context of switched DAEs.

Lemma 14. *For a given $\zeta \in \left(\mathcal{V}^{\mathrm{end}} + \operatorname{im} B^{\mathrm{imp}}\right) \cap \operatorname{im}\Pi$ the vector $\psi \in \mathbb{R}^m$ solves*

$$\begin{aligned}\min_{\psi\in\mathbb{R}^m} \quad & (\zeta - B^{\mathrm{imp}}\psi)^\top P(\zeta - B^{\mathrm{imp}}\psi), \\ \text{s.t.} \quad & \zeta - B^{\mathrm{imp}}\psi \in \mathcal{V}^{\mathrm{end}},\end{aligned} \tag{21}$$

if and only if $\zeta = \left[\begin{smallmatrix} 0 & 0 & \Pi \end{smallmatrix}\right] h$ and $\psi = \left[\begin{smallmatrix} I & 0 & 0 \end{smallmatrix}\right] h$ for some $h \in \ker\mathcal{H}$, where

$$\mathcal{H} := \begin{bmatrix} B^{\mathrm{imp}\top}PB^{\mathrm{imp}} & B^{\mathrm{imp}\top}(I-\Pi_{\mathcal{V}^{\mathrm{end}}})^\top \\ (I-\Pi_{\mathcal{V}^{\mathrm{end}}})B^{\mathrm{imp}} & 0 \\ -\Pi^\top PB^{\mathrm{imp}} & -\Pi^\top(I-\Pi_{\mathcal{V}^{\mathrm{end}}})^\top \end{bmatrix}^\top \tag{22}$$

and $\Pi_{\mathcal{V}^{\mathrm{end}}}$ is any projector onto $\mathcal{V}^{\mathrm{end}}$.

The proof can be found in the Appendix.

Given the result of Lemma 14, we can compute which states $\zeta \in \operatorname{im}\Pi$ are possibly an endpoint of an optimal trajectory. Moreover, for each potential endpoint $\zeta \in \operatorname{im}\Pi$ we can compute a value of ψ that solves (21). Consequently, for a given optimal solution (x^{diff}, u) where $x^{\mathrm{diff}}(t_f^-) = \zeta$, we are able to express the terminal cost of this solution in terms of $x^{\mathrm{diff}}(t_f^-)$ only.

Corollary 15. *If there exists an input u that solves Problem 2 then the optimal terminal cost satisfies*

$$x(t_f^-)^\top Px(t_f^-) = x^{\mathrm{diff}}(t_f^-)^\top\Psi^\top P\Psi x^{\mathrm{diff}}(t_f^-),$$

where $\Psi = (I - B^{\mathrm{imp}}N)$, for any N satisfying

$$\begin{bmatrix} I & 0 & -N\Pi \end{bmatrix} \ker\mathcal{H} = 0, \tag{23}$$

where $\mathcal{H}$ is given by (22)

Proof. Since (x, u) is solving Problem 2 it follows from Lemma 13 that $\psi = u(t_f^-)$ minimizes (20) for $\zeta = x^{\mathrm{diff}}(t_f^-)$. By Corollary 12 the optimal input is linear in x^{diff}, i.e., $u = Nx^{\mathrm{diff}}$ for some linear map N. Hence by Lemma 14, N satisfies $\left[\begin{smallmatrix} I & 0 & -N\Pi \end{smallmatrix}\right] h = 0$ for any $h \in \ker\mathcal{H}$, i.e. (23) actually has a solution. Furthermore, for any other $\bar{N}$ which satisfies (23) it follows that $\bar{N}\zeta = \left[\begin{smallmatrix} 0 & 0 & \bar{N}\Pi \end{smallmatrix}\right] h = \left[\begin{smallmatrix} I & 0 & 0 \end{smallmatrix}\right] h = \left[\begin{smallmatrix} 0 & 0 & N\Pi \end{smallmatrix}\right] h = N\zeta$, hence the effective optimal feedback does not depend on the specific choice of N satisfying (23).

Although the minimum of the objective function in (21) is uniquely given for a particular $x^{\mathrm{diff}} \in \mathbb{R}^n$, a minimizer $u \in \mathbb{R}^m$ is not necessarily unique. However, the following result can still be concluded regarding an optimal input.

Corollary 16. *If an input u solves Problem 2 then the optimal feedback satisfies $u(t_f^-) = Nx^{\mathrm{diff}}(t_f^-)$ for some N satisfying* (23).

4.2. Running cost

We will now turn our attention to the running cost and the optimal control given on the half-open interval $[t_0, t_f)$. To that extent, we will write

$$\begin{aligned}\|y(t)\|^2 &= \begin{bmatrix} x \\ u \end{bmatrix}^\top \begin{bmatrix} C \\ D \end{bmatrix} \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}, \\ &= \begin{bmatrix} x^{\mathrm{diff}} \\ u \end{bmatrix}^\top \begin{bmatrix} C^\top \\ \bar{D}^\top \end{bmatrix} \begin{bmatrix} C & \bar{D} \end{bmatrix} \begin{bmatrix} x^{\mathrm{diff}} \\ u \end{bmatrix},\end{aligned}$$

where $\hat{D} = D - CB^{\mathrm{imp}}$. Then, after defining $Q = C^\top C$, $S = \hat{D}^\top C$ and $R = \hat{D}^\top\hat{D}$, we can rewrite the cost functional as

$$J(x_0, u, t_0) = \int_{t_0}^{t_f} \left[\begin{smallmatrix} x^{\mathrm{diff}} \\ u \end{smallmatrix}\right]^\top \left[\begin{smallmatrix} Q & S^\top \\ S & R \end{smallmatrix}\right] \left[\begin{smallmatrix} x^{\mathrm{diff}} \\ u \end{smallmatrix}\right] \mathrm{d}t + x(t_f^-)^\top Px(t_f^-).$$

Lemma 17. *Assume that the matrices C and D satisfy* (15). *Then $\hat{D} := D - CB^{\mathrm{imp}}$ has full column rank and $\hat{D}^\top\hat{D}$ is positive definite.*

Proof. Since $D - CB^{\mathrm{imp}}$ has full column rank it follows directly that $\hat{D}^\top\hat{D}$ is invertible.

Remark 18. As already mentioned in the introduction, the assumption (15) can be regarded as the differential-algebraic version of the assumption that $D^\top D$ is positive definite, which is commonly made in the LQR problem for ordinary differential equations. The assumption that $D^\top D$ is positive definite is usually made to penalize every input action in the cost. As the solution x of a DAE has a component that is directly determined by the input, the cost functional can penalize the input also indirectly via penalizing the corresponding state component. Hence penalizing all input actions is equivalent to the condition (15).

Lemma 19. *If an input $u \in (\mathbb{D}_{pw\mathcal{C}^\infty})^m$ solves Problem 2 then*

$$u(t) = -R^{-1}\left(B^{\mathrm{diff}\top}K(t) + S^\top\right)x^{\mathrm{diff}}(t), \tag{24}$$

where K solves

$$\begin{aligned}\dot{K} = &-A^{\mathrm{diff}\top}K - KA^{\mathrm{diff}} - Q \\ &+ (S + K^\top B^{\mathrm{diff}})R^{-1}(B^{\mathrm{diff}\top}K + S^\top),\end{aligned} \tag{25}$$

with terminal condition $K(t_f^-) = \Psi^\top P\Psi$.

Proof. For any symmetric-matrix-valued continuously differentiable function $K(t)$ defined on $[t_0, t_f)$ we can write the cost-functional as

$$\begin{aligned}&J(x^{\mathrm{diff}}, u, t_0) - x_0^\top K(t_0)x_0 \\ &\quad= \int_{t_0}^{t_f} \left[\begin{smallmatrix} x^{\mathrm{diff}} \\ u \end{smallmatrix}\right]^\top \left[\begin{smallmatrix} Q & S^\top \\ S & R \end{smallmatrix}\right] \left[\begin{smallmatrix} x^{\mathrm{diff}} \\ u \end{smallmatrix}\right] \\ &\qquad + \tfrac{\mathrm{d}}{\mathrm{d}t}(x^{\mathrm{diff}}(t)^\top K(t)x^{\mathrm{diff}}(t))\,\mathrm{d}t \\ &\qquad + x^{\mathrm{diff}}(t_f^-)^\top\left(\Psi^\top P\Psi - K(t_f^-)\right)x^{\mathrm{diff}}(t_f^-).\end{aligned}$$

Taking the two integrals together and computing the second integral using the differential equation and the completion of the squares formula, we obtain (omitting the dependence on t):

$$\begin{aligned}x^{\mathrm{diff}\top}Qx^{\mathrm{diff}}+&2x^{\mathrm{diff}\top}S^\top u+u^\top Ru+\frac{d}{dt}x^{\mathrm{diff}\top}Kx^{\mathrm{diff}}\\ =&x^{\mathrm{diff}\top}(Q+A^{\mathrm{diff}\top}K+KA^{\mathrm{diff}}+\dot{K})x^{\mathrm{diff}\top}\\ &+2u^\top(B^{\mathrm{diff}}K^\top+S)x^{\mathrm{diff}}+u^\top Ru\\ =&x^{\mathrm{diff}\top}KB^{\mathrm{diff}\top}R^{-1}B^{\mathrm{diff}}Kx^{\mathrm{diff}}\\ &+2u^\top(B^{\mathrm{diff}}K^\top+S)x^{\mathrm{diff}}+u^\top Ru\\ &+x^{\mathrm{diff}\top}Wx^{\mathrm{diff}}\\ =&\|Ru+(B^{\mathrm{diff}\top}K+S^\top)x^{\mathrm{diff}}\|^2\\ &+x^{\mathrm{diff}\top}Wx^{\mathrm{diff}},\end{aligned}$$

where

$$\begin{aligned}W:=&\dot{K}+A^{\mathrm{diff}\top}K+KA^{\mathrm{diff}}\\ &-(S+K^\top B^{\mathrm{diff}})R^{-1}(B^{\mathrm{diff}\top}K+S^\top)+Q.\end{aligned}$$

Consequently, we can rewrite the cost in Problem 2 as

$$\begin{aligned}J(x_0,u,t_0)=&x_0^\top K(t_0^-)x_0\\ &+\int_{t_0}^{t_f}\|Ru(t)+(B^{\mathrm{diff}\top}K(t)+S^\top)x^{\mathrm{diff}}(t)\|^2\\ &+x^{\mathrm{diff}}(t)^\top W(t)x^{\mathrm{diff}}(t)\,\mathrm{d}t\\ &+x^{\mathrm{diff}}(t_f^-)^\top\left(\Psi^\top P\Psi-K(t_f^-)\right)x^{\mathrm{diff}}(t_f^-),\end{aligned}$$

where $\Psi=(I-B^{\mathrm{imp}}N)\Pi$ for some N satisfying (23).

Under the assumption (15), it follows from the literature on solutions on the Riccati differential equation (cf. Theorem 10.7 in [46]) that a function K satisfying $K(t_f^-)=\Psi^\top P\Psi$ such that $W=0$ can always be chosen. Hence by choosing $K(t)$ such that $W=0$ and $K(t_f^-)=\Psi^\top P\Psi$ we obtain that the cost $J(x^{\mathrm{diff}},u)$ can be expressed as

$$\begin{aligned}&J(x_0,u,t_0)-x_0^\top K(t_0^-)x_0\\ &=\int_{t_0}^{t_f}\|Ru(t)+(B^{\mathrm{diff}\top}K(t)+S^\top)x^{\mathrm{diff}}(t)\|^2\,\mathrm{d}t.\quad(26)\end{aligned}$$

Clearly without the constraint $x^{\mathrm{diff}}(t_f^-)-B^{\mathrm{imp}}u(t_f^-)\in\mathcal{V}^{\mathrm{end}}$ it follows that $J(x_0,u)$ is minimized if the input is given by

$$u=-R^{-1}(B^{\mathrm{diff}\top}K+S^\top)x^{\mathrm{diff}}.$$

Next, we will show that for the problem with the constraint $\inf J(x_0,u,t_0)=x_0^\top K(t_0^-)x_0$.

Case 1: $x^{\mathrm{diff}}(t_f^-)\in\mathcal{V}^{\mathrm{end}}+\operatorname{im}B^{\mathrm{imp}}$
Suppose applying the input (24) to the initial condition x_0 results in $x^{\mathrm{diff}}(t_f^-)\in\mathcal{V}^{\mathrm{end}}+\operatorname{im}B^{\mathrm{imp}}$. Then consider the input $u_\delta=u+\bar{u}_\delta$ where u_δ is defined as

$$u_\delta=\begin{cases}0, & t_0\leqslant t<t_f-\delta,\\ \phi(t), & t_f-\delta\leqslant t<t_f-\frac{\delta}{2},\\ Nx^{\mathrm{diff}}(t_f^-)-u(t_f^-), & t_f-\frac{\delta}{2}\leqslant t<t_f,\end{cases}$$

for some N satisfying $[\,I\ 0\ -N\,]\ker\mathcal{H}$, is constant and $\phi(t)$ is chosen in such a way that the corresponding solution x_δ^{diff} satisfies $x_\delta^{\mathrm{diff}}(t_f^-)=x^{\mathrm{diff}}(t_f^-)$ (which is always possible, cf. Lemma 28 in the Appendix).

Note that

$$\begin{aligned}x_\delta^{\mathrm{diff}}(t_f^-)-B^{\mathrm{imp}}u_\delta(t_f)&=x^{\mathrm{diff}}(t_f^-)-B^{\mathrm{imp}}Nx_\delta^{\mathrm{diff}}(t_f^-)\\ &=x^{\mathrm{diff}}(t_f^-)-B^{\mathrm{imp}}Nx^{\mathrm{diff}}(t_f^-)\in\mathcal{V}^{\mathrm{end}}\end{aligned}$$

and thus u_δ is a feasible input.

It follows from (26) that for every $\varepsilon>0$ there exists a solution $(x_\delta^{\mathrm{diff}},u_\delta)$ such that

$$J(x_\delta^{\mathrm{diff}},u_\delta,t_0)=x_0^\top K(t_0^-)x_0+\varepsilon$$

and thus we can conclude

$$\inf J(x^{\mathrm{diff}},u,t_0)=x_0^\top K(t_0^-)x_0.$$

However, the infimum is attained if and only if the input is given by (24). Hence the result follows.

Case 2: $x^{\mathrm{diff}}(t_f^-)\notin\mathcal{V}^{\mathrm{end}}+\operatorname{im}B^{\mathrm{imp}}$
Suppose applying the input (24) to the initial condition x_0 results in $x^{\mathrm{diff}}(t_f^-)\in\mathcal{V}^{\mathrm{end}}+\operatorname{im}B^{\mathrm{imp}}$. We will prove in this case that there does not exist an optimal control. For the sake of contradiction, assume that the optimal control is given by $\tilde{u}\in(\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m$. Then the solution $(\tilde{x}^{\mathrm{diff}},\tilde{u})$ must satisfy $\tilde{x}^{\mathrm{diff}}(t_f^-)\in\mathcal{V}^{\mathrm{end}}+\operatorname{im}B^{\mathrm{imp}}$, as the solution is feasible.

Let $y_0\in\mathbb{R}^n$, $y_0\neq x_0$ be an initial value such that the solution (y^{diff},u) with $y^{\mathrm{diff}}(t_0^-)=y_0$ and u given by the feedback (24) satisfies $y^{\mathrm{diff}}(t_f^-)=q$. Recall, that the optimal control is a feedback. As a consequence of the linearity of the optimal control in the state, it must hold that $v=u-\tilde{u}$ is the optimal control for the initial value $z_0=y_0-x_0$. However, by linearity of solutions, the solution (z^{diff},v) satisfies

$$\begin{aligned}z^{\mathrm{diff}}(t^-)&=e^{A^{\mathrm{diff}}t-t_t}z_0+\int_{t_0}^t e^{A^{\mathrm{diff}}(t-\tau)}B^{\mathrm{diff}}v(\tau)\,\mathrm{d}\tau\\ &=e^{A^{\mathrm{diff}}t-t_t}x_0+\int_{t_0}^t e^{A^{\mathrm{diff}}(t-\tau)}B^{\mathrm{diff}}\tilde{u}(\tau)\,\mathrm{d}\tau\\ &\quad-e^{A^{\mathrm{diff}}t-t_t}y_0-\int_{t_0}^t e^{A^{\mathrm{diff}}(t-\tau)}B^{\mathrm{diff}}u(\tau)\,\mathrm{d}\tau\\ &=\tilde{x}^{\mathrm{diff}}(t^-)-y^{\mathrm{diff}}(t^-)\end{aligned}$$

and consequently $z^{\mathrm{diff}}(t_f^-)=0$. However, this implies that $z_0=0$, as a feedback, cannot control an initial condition to zero unless it is zero. Hence we can conclude that $x_0=y_0$, which yields a contradiction. Hence there does not exist an optimal control for x_0.

4.3. Combining the results

Thus far we have only been concerned with necessary conditions for solvability of Problem 2. The reason that the conditions in Corollary 19 are not sufficient in general is that a feedback of the form (24) does not necessarily ensure that all the constraints are satisfied. A solution (x^{diff}, u) with u given by (24) and $x^{\text{diff}}(t_0^-) = x_0 \in \operatorname{im}\Pi$ does not necessarily satisfy

$$x^{\text{diff}}(t_f^-) - B^{\text{imp}}u(t_f^-) \in \mathcal{V}^{\text{end}},$$

nor

$$x(t_f^-)^\top P x(t_f^-) = x^{\text{diff}}(t_f^-)\Psi^\top P \Psi x^{\text{diff}}(t_f^-),$$

for any N for which $\left[\begin{smallmatrix} I & 0 & -N \end{smallmatrix}\right] \ker \mathcal{H} = 0$. Both these conditions can be rewritten equivalently as

$$(I - \Pi_{\mathcal{V}^{\text{end}}})(I - B^{\text{imp}}\Lambda)x^{\text{diff}}(t_f^-) = 0 \tag{27}$$

and

$$(I - B^{\text{imp}}\Lambda)^\top P((I - B^{\text{imp}}\Lambda) - \Psi^\top P\Psi)x^{\text{diff}} = 0, \tag{28}$$

where we have written

$$\Lambda := -R^{-1}\left(B^{\text{diff}\top}\Psi^\top P\Psi + S^\top\right),$$

for convenience. However, it follows straightforwardly that if a solution (x^{diff}, u) with $x^{\text{diff}}(t_0^-) = x_0$ and u satisfying (24) is such that (27) and (28) are satisfied the input is optimal. To prove this, we will first introduce the backward state-transition matrix, defined similarly to [2] or [47] and which also appears in [8].

Definition 20. The backwards state transition matrix for the closed loop time-varying differential equation

$$\dot{x}^{\text{diff}} = \left(A^{\text{diff}} - B^{\text{diff}}R^{-1}(B^{\text{diff}\top}K + S^\top)\right)x^{\text{diff}},$$

is given by $\Omega(t, t_f)$, where K is a solution to (25) with terminal condition $K(t_f^-) = \Psi^\top P\Psi$. In particular, the state satisfies $x^{\text{diff}}(t) = \Omega(t, t_f)x^{\text{diff}}(t_f^-)$.

Theorem 21. *Problem 2 is solvable if and only if*

$$x_0 \in \mathcal{V}^{\text{init}} := \Omega(t_0, t_f)\ker \Xi\Pi, \tag{29}$$

with

$$\Xi = \begin{bmatrix} (I - \Pi_{\mathcal{V}^{\text{end}}})(I - B^{\text{imp}}\Lambda) \\ (I - B^{\text{imp}}\Lambda)^\top P(I - B^{\text{imp}}\Lambda) - \Psi^\top P\Psi \end{bmatrix}$$

where $\Omega(t_0, t_f)$ is the backward state transition matrix as defined in Definition 20 and the optimal control is given by

$$u(t) = -R^{-1}(B^{\text{diff}\top}K(t) + S^\top)x^{\text{diff}}(t), \tag{30}$$

where K is a solution to (25) with terminal condition $K(t_f^-) = \Psi^\top P\Psi$. Finally, the optimal cost is given by

$$J^*(x_0, u, t_f) = x^{\text{diff}}(t_0^-)K(t_0)x^{\text{diff}}(t_0^-)$$

and is quadratic in $x^{\text{diff}}(t_0^-)$.

5. LQR for Higher-index DAEs

In the previous section, Problem 2 has been considered where the DAE was assumed to be of index-1. This assumption allowed us to decompose the state into a component that solves an ODE and a feed-trough term depending directly on the input. Furthermore, the solution (x, u) was impulse-free regardless of the initial value as long as the input was impulse-free. This decomposition can not be made anymore if a higher index DAE is considered. As a result of the higher index of the DAE, the state will also depend on the derivatives of the input u and the state will not necessarily be impulse-free if the input is impulse-free.

In fact, there exists an input that results in an impulse-free solution (x, u) satisfying $x(t_0^-) = x_0$ if and only if the initial value is contained in the impulse-controllable space $\mathcal{C}^{\text{imp}}$. For such initial values, we will show in the following a particular impulse-controllable DAE can be considered equivalently instead of (14). Specifically, after applying a preliminary feedback, an index-1 DAE can be considered.

For initial values $x_0 \notin \mathcal{C}^{\text{imp}}$, i.e., initial values that are not contained in the impulse-controllable space, a solution x satisfying $x(t_0^-) = x_0 \notin \mathcal{C}^{\text{imp}}$ a Dirac impulse will occur inevitably, i.e., regardless of the choice of input. However, an optimal control might still exist for these initial values, as long as the corresponding Dirac impulses are not visible in the output. Combining these observations lead to the following result.

Lemma 22. *Consider the DAE* (14) *and assume it is of arbitrary index. There exists an impulse-free input $u \in (\mathbb{D}_{pw\mathcal{C}^\infty})^m$ such that for the solution (x, u) satisfying $x(t_0^-) = x_0$ of* (14) *the output is impulse-free at t_0, i.e., $y[t_0] = Cx[t_0] + Du[t_0] = 0$, if and only if $x_0 \in \mathcal{C}^{\text{imp}} + \mathcal{O}^{\text{imp}}$ where $\mathcal{O}^{\text{imp}}$ is the impulse-unobservable space defined as*

$$\mathcal{O}^{\text{imp}} := \ker \begin{bmatrix} CE^{\text{imp}} \\ C(E^{\text{imp}})^2 \\ \vdots \\ C(E^{\text{imp}})^{\nu-1} \end{bmatrix} \tag{31}$$

and ν is the index of nilpotency of E^{imp}.

Proof. The proof can be found in the Appendix.

As the condition $x_0 \in \mathcal{C}^{\text{imp}} + \mathcal{O}^{\text{imp}}$ is necessary and sufficient for the existence of an impulse-free output, it is a necessary condition for the existence of an impulse-free input that minimizes (13), subject to (14). However, it suffices to only consider initial values contained in $\mathcal{C}^{\text{imp}}$. The main reason for this is that initial values $x_0 \in (\mathcal{C}^{\text{imp}} + \mathcal{O}^{\text{imp}}) \setminus \mathcal{C}^{\text{imp}}$ will only produce a Dirac impulse at t_0 if a zero input is applied. This impulse will occur in the unobservable space, the output will remain zero. Hence it follows that $u = 0$ is trivially the optimal input.

Corollary 23. *Consider the DAE* (14). *For any $x_0 \in \mathcal{C}^{\text{imp}} + \mathcal{O}^{\text{imp}}$ a solution (x, u) with $x(t_0^-) = x_0$ satisfies $y(t) = \bar{y}(t)$ where $\bar{y}(t)$ is the output corresponding to the solution $(\bar{x}, u)$ with $\bar{x}(t_0^-) = Wx_0$ where W is an orthogonal projector onto $\mathcal{C}^{\text{imp}}$.*

Hence in the remainder of the paper, we will consider initial values contained in the impulse-controllable space of (14). However, instead of considering (14), which is not impulse-controllable and of higher index, we can consider an auxiliary impulse-controllable DAE. For there exists an impulse-controllable DAE which has the same input-output behavior as (14) for initial values $x_0 \in \mathcal{C}^{\text{imp}}$. Because this auxiliary DAE is impulse-controllable, it is much more convenient to consider in the analysis than (14) itself.

Lemma 24. *Let $\mathcal{C}^{\text{imp}}$ be the impulse-controllable space of (14). A distribution (x, u) satisfying $x(t_0^-)$ with $x_0 \in \mathcal{C}^{\text{imp}}$, solves (14) if and only if it solves*

$$EW\dot{x} = Ax + Bv, \tag{32}$$

$$y = Cx + Dv, \tag{33}$$

where W is an orthogonal projector onto $\mathcal{C}^{\text{imp}}$. Moreover the pair (EW, A) is regular and (32) is an impulse-controllable DAE.

Proof. The proof can be found in the Appendix.

The auxiliary DAE (32) is much easier to analyze with respect to the optimal control problem as for impulse-controllable DAEs there exists a feedback that reduces the index to 1, *cf.* Lemma 4. Let $u = Lx + v$ be such a feedback. After applying this feedback we obtain

$$\Sigma^{\text{aux}} : \begin{cases} EW\dot{x} &= (A + BL)x + Bv, \\ \quad y &= (C + DL)x + Dv, \end{cases} \tag{34}$$

which is of index-1. For index-1 DAEs the results have already been established and the following result shows that these results can be carried over to (34). As such, to solve Problem 2 subject to a higher index DAE, it suffices to find an optimal input v that solves the following auxiliary Problem.

Problem 3. Consider the DAE (14), let W be a projector onto $\mathcal{C}^{\text{imp}}$ corresponding to (14). Find an input $v \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^m$ that solves the following problem:

$$\min \quad J(x_0, v, t_0) = \int_{t_0}^{t_f} \|\bar{y}(t)\|^2 \, \mathrm{d}t + x(t_f^-)Px(t_f^-), \tag{35}$$

$$\text{s.t.} \quad EW\dot{x} = (A + BL)x + Bv, \tag{36}$$

$$\bar{y} = (C + DL)x + Dv, \tag{37}$$

$$x(t_0^-) = x_0 \in \mathbb{R}^n, \tag{38}$$

$$x(t_f^-) \in \mathcal{V}^{\text{end}}, \tag{39}$$

on the interval $[t_0, t_f)$, where L is a matrix, such that $(EW, A + BL)$ is of index-1.

Lemma 25. *Let $\mathcal{C}^{\text{imp}}$ be the impulse-controllable space corresponding to (14). There exists an input $u \in \mathbb{D}_{pw\mathcal{C}^\infty}$ that solves Problem 2 subject to $x_0 \in \mathcal{C}^{\text{imp}}$ if and only if there exists an input $v \in (\mathbb{D}_{pw\mathcal{C}^\infty})^m$ that solves Problem 3 subject to $x_0 \in \mathcal{C}^{\text{imp}}$.*

Furthermore, the optimal input that solves Problem 2 subject to (14) satisfies $u = Lx + v$, where v is the optimal input that solves Problem 3.

Proof. As $x_0 \in \mathcal{C}^{\text{imp}}$ it follows form Lemma 24 that the solution (x, u) solves (14) if and only if it solves (32). Hence we will consider solutions of (32). Applying a feedback to (32) can be regarded as a change of coordinates

$$\begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} I & 0 \\ L & I \end{bmatrix} \begin{bmatrix} \bar{x} \\ v \end{bmatrix}. \tag{40}$$

Writing (32) as

$$\begin{bmatrix} EW & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{u} \end{bmatrix} = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix},$$

enables us to write

$$\begin{aligned} \begin{bmatrix} EW & 0 \end{bmatrix} \begin{bmatrix} \dot{\bar{x}} \\ \dot{v} \end{bmatrix} &= \begin{bmatrix} EW & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{u} \end{bmatrix} \\ &= \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} I & 0 \\ L & I \end{bmatrix} \begin{bmatrix} \bar{x} \\ v \end{bmatrix} \\ &= \begin{bmatrix} (A + BL) & B \end{bmatrix} \begin{bmatrix} \bar{x} \\ v \end{bmatrix}. \end{aligned}$$

Hence (x, u) solves (32) if and only if, $(\bar{x}, v)$ satisfying (40) solves (34). Furthermore, it follows naturally from that if $u = Lx + v$ that $J(x_{,0} u) = \bar{J}(\bar{x}_0, v)$.

Given a method to compute the optimal input to Problem 2, it remains to characterize the space for which the problem can be solved. This space can easily be computed based on the computation of the optimal input for Problem 3.

Lemma 26. *Let $\overline{\mathcal{V}}^{\text{init}}$ be the space of initial values for which Problem 3 can be solved. Then the space of initial values for which Problem 2 can be solved is given by*

$$\mathcal{V}^{\text{init}} = \overline{\mathcal{V}}^{\text{init}} \cap (\mathcal{C}^{\text{imp}} + \mathcal{O}^{\text{imp}}), \tag{41}$$

where $\mathcal{C}^{\text{imp}}$ and $\mathcal{O}^{\text{imp}}$ be the impulse-controllable space and impulse-observable space corresponding to (14).

6. LQR for general switched DAEs

Given the results regarding Problem 2 where the DAE is assumed to be of arbitrary index, the results for Problem 1 where each mode of (2a) is of arbitrary index follow straightforwardly. A summarizing algorithm is presented in Algorithm 1.

We illustrate the overall procedure with the following illustrative (academic) example.

Example 27. Consider the switched DAE given by

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u, & 0 \leqslant t < 1, \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \dot{x} &= -x + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u, & 1 \leqslant t < 2, \\ \dot{x} &= \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u, & 2 \leqslant t < 3, \end{aligned}$$

Algorithm 1 LQR with subspace constraint

Input : $E_i, A_i, B_i, C_i, D_i, t_i (i = 0, 1, \dots, \mathtt{n}), P, \mathcal{V}^{\text{end}}$

Set $\mathcal{V}^{\text{init}}_{t_{\mathtt{n}+1}} := \mathcal{V}^{\text{end}}$ and $K_{\mathtt{n}}(t^-_{\mathtt{n}+1}) := P$

for $i = \mathtt{n}, \mathtt{n}-1, ..., 0$ **do**

Step 1: Preconditioning

Compute $\mathcal{C}^{\text{imp}}_i$ via (11) and $\mathcal{O}_i$ via (31)

Choose any projector W_i onto $\mathcal{C}^{\text{imp}}_i$

Utilizing Lemmas 4 and 24, choose L_i such that $(EW_i, A_i + B_i L_i)$ is of index 1

Define

$$(\bar{E}_i, \bar{A}_i, \bar{B}_i, \bar{C}_i, \bar{D}_i) := (E_i W_i, A_i + B_i L_i, B_i, C_i + D_i L_i, D_i)$$

Compute $\bar{\Pi}_i$, $\bar{A}^{\text{diff}}_i$, $\bar{B}^{\text{diff}}_i$,$\bar{B}^{\text{imp}}_i$ via the Wong sequences of $(\bar{E}_i, \bar{A}_i)$

Step 2: Solve Riccati equations

Set $\mathcal{V}^{\text{end}}_{t_{i+1}} = \mathcal{V}^{\text{init}}_{t_{i+1}}$

Solve

$$\dot{K}_i = -\bar{A}^{\text{diff}\top}_i K_i - K_i \bar{A}^{\text{diff}}_i + (\bar{S}_i + K_i^\top \bar{B}^{\text{diff}}_i)\bar{R}_i^{-1}(\bar{B}^{\text{diff}\top}_i K_i + \bar{S}_i^\top) - \bar{Q}_i,$$

on $[t_i, t_{i+1})$, with $\bar{R}_i := (\bar{D}_i - \bar{C}_i \bar{B}^{\text{imp}}_i)^\top (\bar{D}_i - \bar{C}_i \bar{B}^{\text{imp}}_i)$, $\bar{S}_i := (\bar{D}_i - \bar{C}_i \bar{B}^{\text{imp}}_i)^\top \bar{C}^\top$, $\bar{Q}_i := \bar{C}_i^\top \bar{C}_i$ and boundary condition

$$K_i(t^-_{i+1}) = \Psi_i^\top K_{i+1}(t^+_{i+1})\Psi_i, \qquad \text{if } i \neq \mathtt{n}$$

where $\Psi_i = (I - \bar{B}^{\text{imp}}_i N_i)\bar{\Pi}_i$, for some N_i that satisfies $\begin{bmatrix} I & 0 & N_i\bar{\Pi}_i \end{bmatrix} \ker \mathcal{H}_i = 0$, with

$$\mathcal{H}_i = \begin{bmatrix} \bar{B}^{\text{imp}\top}_i K_i(t^-_{i+1})\bar{B}^{\text{imp}}_i & \bar{B}^{\text{imp}\top}_i (I - \Pi_{\mathcal{V}^{\text{end}}_{t_{i+1}}})^\top \\ (I - \Pi_{\mathcal{V}^{\text{end}}_{t_{i+1}}})\bar{B}^{\text{imp}}_i & 0 \\ -\bar{\Pi}_i^\top P \bar{B}^{\text{imp}}_i & -\bar{\Pi}_i^\top (I - \Pi_{\mathcal{V}^{\text{end}}_{t_{i+1}}})^\top \end{bmatrix}^\top$$

and $\Pi_{\mathcal{V}^{\text{end}}_{t_{i+1}}}$ is a projector onto the subspace $\mathcal{V}^{\text{end}}_{t_{i+1}}$

Step 3: Compute subspace $\mathcal{V}^{\text{init}}_{t_i}$

Compute $\Omega_i(t_i, t_{i+1})$ (see Def. 20) for the system

$$\dot{x}^{\text{diff}} = \left(\bar{A}^{\text{diff}}_i - \bar{B}^{\text{diff}}_i \bar{R}_i^{-1}(\bar{B}^{\text{diff}\top}_i K_i + \bar{S}_i^\top)\right) x^{\text{diff}}.$$

Compute $\overline{\mathcal{V}}^{\text{init}}_{t_i} = \Omega_i(t_i, t_{i+1}) \ker \Xi_i \bar{\Pi}_i$, with

$$\Xi_i = \begin{bmatrix} (I - \Pi_{\mathcal{V}^{\text{end}}_{t_{i+1}}})(I - \bar{B}^{\text{imp}}\Lambda_i) \\ (I - \bar{B}^{\text{imp}}\Lambda_i)^\top P (I - \bar{B}^{\text{imp}}\Lambda_i) - \Psi_i^\top P \Psi_i \end{bmatrix}$$

where $\Lambda := -\bar{R}_i^{-1}\left(\bar{B}^{\text{diff}\top}_i \Psi_i^\top K(t^-_{i+1})\Psi_i + \bar{S}_i^\top\right)$.

Compute $\mathcal{V}^{\text{init}}_{t_i} = \overline{\mathcal{V}}^{\text{init}}_{t_i} \cap (\mathcal{C}^{\text{imp}}_i + \mathcal{O}^{\text{imp}}_i)$.

end for

Step 4: Compute optimal control

Compute

$$u(t) = -\bar{R}^{-1}_{\sigma(t)}\left(\bar{B}^{\text{diff}\top}_{\sigma(t)} K_{\sigma(t)}(t) + \bar{S}^\top_{\sigma(t)}\right)\bar{\Pi}_{\sigma(t)} x(t),$$

together with the output

$$y = x + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u.$$

The cost functional to be minimized is thus given by

$$J(x_0, u, 0) = \int_0^3 \|y(t)\| \, \mathrm{d}t,$$

subject to $x(3^-) \in \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\} := \mathcal{V}^{\text{end}}$. In this particular problem, the terminal cost matrix is given by $P = 0$.

Note that the mode active on $1 \leqslant t < 2$ is impulse-controllable, but not index-1. To that extent a preliminary index-reducing feedback given by

$$u(t) = \begin{cases} \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} x(t) + v(t), & 1 \leqslant t < 2, \\ v(t), & \text{otherwise}, \end{cases}$$

is applied, resulting in

$$\begin{aligned}
(\bar{E}_0, \bar{A}_0, \bar{B}_0, \bar{C}_0, \bar{D}_0) &= \left(I, \begin{bmatrix} 1&1&0 \\ 1&0&1 \\ 1&1&2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, I, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right), \\
(\bar{E}_1, \bar{A}_1, \bar{B}_1, \bar{C}_1, \bar{D}_1) &= \left(\begin{bmatrix} 1&0&0 \\ 0&0&1 \\ 0&0&0 \end{bmatrix}, \begin{bmatrix} -1&1&1 \\ 0&0&0 \\ 0&1&1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1&1&1 \\ 0&1&0 \\ 0&0&1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right), \\
(\bar{E}_2, \bar{A}_2, \bar{B}_2, C_2, D_2) &= \left(I, \begin{bmatrix} -1&0&0 \\ 0&0&0 \\ 0&0&0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, I, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right),
\end{aligned}$$

The optimal feedback matrix on each interval $[t_i, t_{i+1})$, $i \in \{0, 1, 2\}$ is computed after solving

$$\begin{aligned}
\dot{K}_i &= -\bar{A}^{\text{diff}\top}_i K_i - K_i \bar{A}^{\text{diff}}_i + (\bar{S}_i + K_i^\top \bar{B}^{\text{diff}}_i)\bar{R}_i^{-1}(\bar{B}^{\text{diff}\top}_i K + \bar{S}_i^\top) - \bar{Q}_i, \\
K_i(t^-_{i+1}) &= \Psi_i^\top K_{i+1}(t^+_{i+1})\Psi_i,
\end{aligned}$$

where $\Psi_i = (I - B^{\text{imp}}_i N_i)\bar{\Pi}_i$ for some N_i which satisfies $\begin{bmatrix} I & 0 & -N_i\bar{\Pi} \end{bmatrix} \ker \mathcal{H}_i = 0$. and $K_2(3^-) = 0$. The computation yields

$$\begin{aligned}
K_1(2^-) &= \begin{bmatrix} 0&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix}, \; K_0(1^-) = \begin{bmatrix} 0.39 & 0 & 0.38 \\ 0 & 0 & 0 \\ 0.38 & 0 & 2.40 \end{bmatrix}, \\
K_0(0^+) &= \begin{bmatrix} 0.21 & -0.03 & 0.07 \\ -0.03 & 0.03 & -0.19 \\ 0.07 & -0.19 & 1.59 \end{bmatrix}.
\end{aligned}$$

After computing the backward state transition matrices Ξ_i it follows that

$$\begin{aligned}
\mathcal{V}^{\text{init}}_2 &= \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}, \quad \mathcal{V}^{\text{init}}_1 = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0.54 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}, \\
\mathcal{V}^{\text{init}}_0 &= \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0.49 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0.06 \end{bmatrix}\right\}.
\end{aligned}$$

Given the solution K_i we can compute the optimal input and optimal state trajectory, which are shown in Figure 1 and 2, respectively. As can be seen, both the optimal input and the optimal trajectory are piecewise continuous and contain jumps. ◇

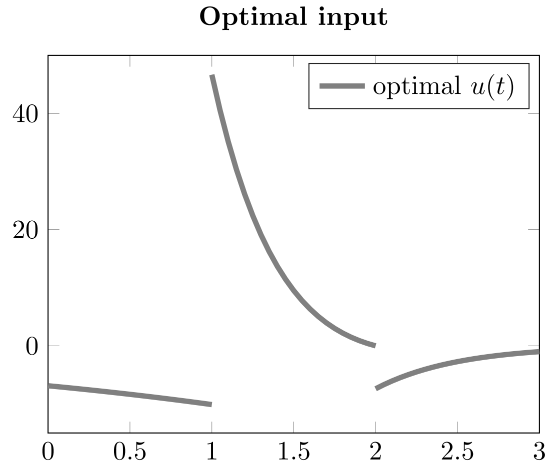

Figure 1: The optimal input $u(t)$ that solves Problem 1

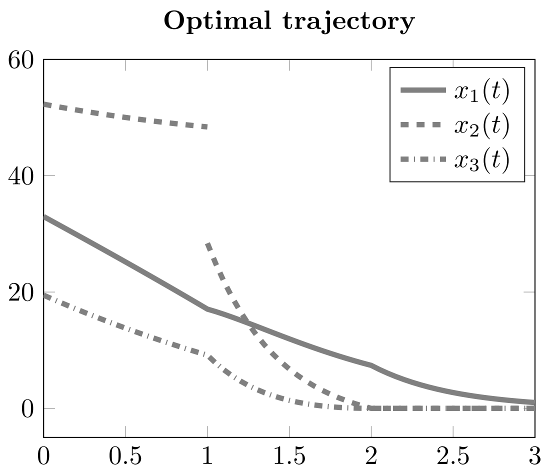

Figure 2: The corresponding optimal trajectory $x(t)$ that solves Problem 1 with initial value $x_0 = [\, 32.98 \; 52.30 \; 19.46 \,]^\top$.

7. Conclusion

In this paper, the finite horizon LQR problem for switched linear differential-algebraic equations has been studied. It was shown that for switched DAEs with a switching signal that induces locally finitely many switches the problem can be solved by solving LQR problems for non-switched DAE repeatedly. First, it was shown how to solve the non-switched problems for index-1 DAEs followed by an extension of the results to higher index DAEs. The resulting optimal control can be computed based on the solution of a Riccati differential equation expressed in terms of the differential system matrices. Although these differential systems matrices do not depend on a particular coordinate transformation, it remains a future direction of research to express the results in terms of the original system matrices,

Another natural direction of future research is to explore the admission of impulsive inputs. The authors suspect however that the results in this direction would not be much different than the ones already obtained in this paper and the results on singular optimal control obtained by Willems et all.

Appendix A. Proofs

Proof of Lemma 6

First we will show that the map $x_0 \mapsto u$ is linear, where $x(t_0^-) = x_0$ and u solves Problem 1; in particular, we will show that λu is the optimal control for any initial value λx_0 and that for any initial values $x_0, z_0 \in \mathbb{R}^n$ for which optimal inputs u_x, u_z exists, the input $u_x + u_z$ is optimal for the initial value $x_0 + z_0$.

To that extent, let $V(x_0, t)$ be the value function as

$$V(x_0, t) = \inf_u J(x_0, u, t_0) \tag{A.1}$$

Applying the input λu to an initial condition λx_0 results in a trajectory λx, due to the linearity of solutions of the switched DAE. This means that $J(\lambda x_0, \lambda u) = \lambda^2 J(x_0, u)$ for any $\lambda \in \mathbb{R}$ and we can conclude that

$$\lambda^2 V(x_0, t_0) = \lambda^2 J(x_0, u) = J(\lambda x_0, \lambda u) = V(\lambda x_0, t_0).$$

Hence we can conclude that if u is the optimal input for x_0 that λu is the optimal input for λx_0. In the following, we will prove if u_x and u_z are the optimal inputs for x_0 and z_0 respectively, that $u_x + u_z$ is the optimal input for $x_0 + z_0$. To do so, it will be first shown that

$$V(x_0 + z_0, t_0) + V(x_0 - z_0, t_0) = 2V(x_0, t_0) + 2V(z_0, t_0). \tag{A.2}$$

Observe that

$$\begin{aligned} &\|C_\sigma(x+z) + D_\sigma(u_x+u_z)\|^2 + \|C_\sigma(x-z) + D_\sigma(u_x-u_z)\|^2 \\ &\qquad = 2\|C_\sigma x + D_\sigma u_x\|^2 + 2\|C_\sigma z + D_\sigma u_z\|^2, \end{aligned}$$

and

$$\begin{aligned} &(x(t_f^-) + z(t_f^-))^\top P (x(t_f^-) + z(t_f^-)) \\ &\quad + (x(t_f^-) - z(t_f^-))^\top P (x(t_f^-) - z(t_f^-)) \\ &\quad = x(t_f^-)^\top P x(t_f^-) + z(t_f^-)^\top P z(t_f^-), \end{aligned}$$

from which we can conclude that

$$\begin{aligned} &J(x_0 + z_0, u_x + u_z, t_0) + J(x_0 - z_0, u_x - u_z, t_0) \\ &\qquad = 2J(x_0, u_x, t_0) + 2J(z_0, u_z, t_0). \end{aligned}$$

Hence for all input u_x and u_z (and thus not necessarily the optimal ones) we obtain

$$\begin{aligned} &V(x_0 + z_0, t_0) + V(x_0 - z_0, t_0) \\ &\qquad \leqslant J(x_0 + z_0, u_x + u_z) + J(x_0 - z_0, u_x - u_z) \\ &\qquad = 2J(x_0, u_x) + 2J(z_0, u_z), \end{aligned}$$

which means that $V(x_0+z_0,t_0)+V(x_0-z_0,t_0) \leqslant 2V(x_0,t_0)+2V(z_0,t_0)$. Conversely

$$\begin{aligned}2V(x_0,t_0)&+2V(z_0,t_0)\\ &\leqslant 2J(x_0,u_x,t_0)+2J(z_0,u_z,t_0)\\ &= J(x_0+z_0,u_x+u_z,t_0)+J(x_0-z_0,u_x-u_z,t_0),\end{aligned}$$

from which we can conclude that $2V(x_0,t_0)+2V(z_0,t_0) \leqslant V(x_0+z_0,t_0)+V(x_0-z_0,t_0)$ and therefore the equality $V(x_0+z_0,t_0)+V(x_0-z_0,t_0) = 2V(x_0,t_0)+2V(z_0,t_0)$ follows. Furthermore, if u_x is the optimal input for x_0 and u_z is the optimal input for z_0 then

$$\begin{aligned}V(x_0-z_0,t_0)&+V(x_0+z_0,t_0)\\ &= 2V(x_0,t_0)+2V(z_0,t_0)\\ &= 2J(x_0,u_x,t_0)+2J(z_0,u_z,t_0)\\ &= J(x_0+z_0,u_x+u_z,t_0)\\ &\quad + J(x_0-z_0,u_x-u_z,t_0).\end{aligned}$$

Since $V(x_0+z_0,t_0) \leqslant J(x_0+z_0,u_x+w_z)$ and similarly $V(x_0-z_0,t_0) \leqslant J(x_0-z_0,u_x-u_z)$, it follows that

$$\begin{aligned}0 &\leqslant J(x_0+z_0,u_x+w_z)-V(x_0+z_0,t_0)\\ &= V(x_0-z_0,t_0)-J(x_0-z_0,u_x-u_z) \leqslant 0,\end{aligned}$$

and thus

$$V(x_0+z_0,t_0) = J(x_0+z_0,u_x+u_z),$$

which shows that u_x+u_z is optimal for x_0+z_0.

Hence there exists a linear map between the optimal trajectory and the optimal input. In particular, the map $x(t_0^-) \mapsto u(t_0^+)$ is linear, i.e., there exists a matrix $F(t_0) \in \mathbb{R}^{m\times n}$ such that $u(t_0^+) = F(t_0)x(t_0^-)$.

From the dynamic programming principle [44, 48] it follows that $u_{[\tau,t_f)}$ is the optimal control for the cost function in Problem 1 considered on the interval $[\tau,t_f)$ for any $\tau \in [t_0,t_f)$, hence by replacing the initial time t_0 in the above argumentation by $\tau \in [t_0,t_f)$ we can conclude that for every $\tau \in [t_0,t_f)$ a matrix $F(\tau) \in \mathbb{R}^{m\times n}$ exists such that the optimal control satisfies $u(\tau^+) = F(\tau)x(\tau^-)$.

Proof of Lemma 13

Before proving Lemma 13 we need the following technical lemma:

Lemma 28. *Consider the ODE* (7) *on the interval* $[0,s]$ *and with zero initial condition. Then for any* $\alpha \in \mathbb{R}^m$, *there exists* $\phi : [0,s/2] \to \mathbb{R}^m$ *such that the input*

$$u(t) = \begin{cases}\phi(t), & t \in [0,s/2)\\ \alpha, & t \in [s/2,s)\end{cases}$$

has a corresponding solution x *with* $x(s^-) = 0$.

Proof. Let $x_1 := -e^{-A^{\mathrm{diff}}s/2}\int_0^{s/2} e^{A^{\mathrm{diff}}(s/2-\tau)}B^{\mathrm{diff}}\alpha \mathrm{d}\tau$, then applying $u(t)=\alpha$ on $[s/2,s)$ with initial value x_1 will result in a solution which reaches zero at $t=s$. Furthermore, by definition $e^{A^{\mathrm{diff}}s/2}x_1$ is reachable, and since the reachable space is A^{diff}-invariant, it follows that also x_1 is reachable from zero, which guarantees the existence of ϕ as claimed.

In order to prove Lemma 13, assume now that u solves Problem 2 for some fixed x_0. Let x^{diff} be the corresponding optimal trajectory on $[t_0,t_f)$. Denote $u(t_f^-) = \psi \in \mathbb{R}^m$ and $x^{\mathrm{diff}}(t_f^-) = \zeta \in \operatorname{im}\Pi$. Seeking a contradiction, assume there exists a value w for which $\zeta - B^{\mathrm{imp}}w \in \mathcal{V}^{\mathrm{end}}$ and

$$\begin{aligned}(\zeta - B^{\mathrm{imp}}w)^\top P(\zeta - B^{\mathrm{imp}}w)&\\ &= (\zeta - B^{\mathrm{imp}}\psi)^\top P(\zeta - B^{\mathrm{imp}}\psi) - M,\end{aligned}$$

for some $M>0$. Consider the solution (x_s,u_s) of (14) where $u_s = u+\bar{u}_s$ and $\bar{u}_s$ is defined as

$$\bar{u}_s = \left\{\begin{array}{ccl} 0, & t_0 & \leqslant t < t_f - s,\\ \phi(t), & t_f - s & \leqslant t < t_f - \frac{s}{2},\\ w-\psi & t_f-\frac{s}{2}, & \leqslant t < t_f\end{array}\right.$$

where $\phi(t)$ is chosen in such a way that $x_s^{\mathrm{diff}}(t_f^-) = x^{\mathrm{diff}}(t_f^-)$, which is always possible according to Lemma 28. Furthermore, for any $\varepsilon>0$ there exists a sufficiently small $s>0$ such that the output y_s resulting from the solution $(x_s^{\mathrm{diff}},u_s)$ satisfies

$$\begin{aligned}\int_{t_0}^{t_f}\|y_s(t)\|^2\,\mathrm{d}t &= \int_{t_0}^{t_f-s}\|y_s(t)\|^2\,\mathrm{d}t + \int_{t_f-s}^{t_f}\|y_s(t)\|^2\,\mathrm{d}t\\ &= \int_{t_0}^{t_f-s}\|y(t)\|^2\,\mathrm{d}t + \int_{t_f-s}^{t_f}\|y_s(t)\|^2\,\mathrm{d}t\\ &\leqslant \int_{t_0}^{t_f}\|y(t)\|^2\,\mathrm{d}t + \varepsilon,\end{aligned}$$

As $u_s(t_f^-) = u(t_f^-)+\bar{u}_s(t_f^-) = w$ we find that $x_s^{\mathrm{diff}}(t_f^-) - B^{\mathrm{imp}}u_s(t_f^-) \in \mathcal{V}^{\mathrm{end}}$ and

$$J(x_0,u_s) = J(x_0,u)+\varepsilon - M.$$

Hence for $\varepsilon < M$ there exists an s such that the solution $(x_s^{\mathrm{diff}},u_s)$ satisfies $J(x_0,u_s) < J(x_0,u)$, which contradicts the optimality of (x^{diff},u). Therefore the result follows.

Proof of Lemma 14

Note that the terminal cost function

$$(\zeta - B^{\mathrm{imp}}\psi)^\top P(\zeta - B^{\mathrm{imp}}\psi), \tag{A.3}$$

for a given $\zeta \in \operatorname{im}\Pi$ is a convex function of $\psi \in \mathbb{R}^m$. Furthermore $\psi \in \mathbb{R}^m$ minimizes (A.3) if and only if ψ minimizes

$$\tfrac{1}{2}\psi^\top B^{\mathrm{imp}\top}PB^{\mathrm{imp}}\psi - \zeta^\top\Pi^\top PB^{\mathrm{imp}}\psi,$$

where here and in the following we replace ζ by $\Pi\zeta$ to enforce that $\zeta = \Pi\zeta \in \operatorname{im}\Pi$. The constraint $\Pi\zeta - B^{\mathrm{imp}}\psi \in \mathcal{V}^{\mathrm{end}}$ is satisfied if and only if $(I-\Pi_{\mathcal{V}^{\mathrm{end}}})(\Pi\zeta - B^{\mathrm{imp}}\psi) = 0$, where $\Pi_{\mathcal{V}^{\mathrm{end}}}$ is a projector onto $\mathcal{V}^{\mathrm{end}}$. This condition can be written equivalently as

$$(I - \Pi_{\mathcal{V}^{\mathrm{end}}})B^{\mathrm{imp}}\psi = (I - \Pi_{\mathcal{V}^{\mathrm{end}}})\Pi\zeta.$$

As this constraint is a convex function and P is positive semi-definite, it follows that this optimization problem is a convex problem. Hence any local minimizer is a global minimizer. The first-order necessary conditions are thus also sufficient. Hence ψ is a minimizer that satisfies the constraints if and only if there exists a Lagrange multiplier λ such that

$$\begin{bmatrix} B^{\mathrm{imp}\top}PB^{\mathrm{imp}} & B^{\mathrm{imp}\top}(I-\Pi_{\mathcal{V}^{\mathrm{end}}})^\top \\ (I-\Pi_{\mathcal{V}^{\mathrm{end}}})B^{\mathrm{imp}} & 0 \end{bmatrix} \begin{bmatrix} \psi \\ \lambda \end{bmatrix} = \begin{bmatrix} B^{\mathrm{imp}\top}P \\ (I-\Pi_{\mathcal{V}^{\mathrm{end}}}) \end{bmatrix} \Pi\zeta.$$

This can equivalently be written as $\mathcal{H}\xi = 0$ where

$$\mathcal{H} := \begin{bmatrix} B^{\mathrm{imp}\top}PB^{\mathrm{imp}} & B^{\mathrm{imp}\top}(I-\Pi_{\mathcal{V}^{\mathrm{end}}})^\top \\ (I-\Pi_{\mathcal{V}^{\mathrm{end}}})B^{\mathrm{imp}} & 0 \\ -\Pi^\top PB^{\mathrm{imp}} & -\Pi^\top(I-\Pi_{\mathcal{V}^{\mathrm{end}}})^\top \end{bmatrix}^\top \tag{A.4}$$

and $\xi^\top = [\,\psi^\top\ \lambda^\top\ \zeta^\top\,]^\top$. Since $\zeta \in \operatorname{im}\Pi$ and hence $\zeta = \Pi\zeta$ the result follows.

Proof of Lemma 22

($\Rightarrow$) Suppose that there exists an impulse-free input such that $y[t] = 0$. Then since the input u is impulse-free, i.e., $u[t] = 0$, it follows that $y[t] = Cx[t] + Du[t] = Cx[t]$. Consequently, the output is impulse-free for a given impulse-free input if and only if $x[t] \in \ker C$. In the case $u[t] = 0$ then it follows from the solution formula (10) and observing that $E^{\mathrm{imp}} = E^{\mathrm{imp}}(I-\Pi)$ that

$$Cx[t] = -C\sum_{i=0}^{\nu-1}(E^{\mathrm{imp}})^{i+1}(I-\Pi)(x_0 - x(t_0^+))\delta^{(i)} = 0.$$

Consequently $(I-\Pi)(x_0 - x(t_0^+)) \in \ker C(E^{\mathrm{imp}})^i$, for $i \in \{1, 2, ..., \nu-1\}$. Hence we can conclude that $(I-\Pi)(x_0 - x(t_0^+)) \in \mathcal{O}^{\mathrm{imp}}$. Since $(I-\Pi)x(t_0^+) \in \mathcal{C}^{\mathrm{imp}}$ it follows that $(I-\Pi)x_0 \in \mathcal{O}^{\mathrm{imp}} + \mathcal{C}^{\mathrm{imp}}$. Finally, since that $\operatorname{im}\Pi \subseteq \mathcal{C}^{\mathrm{imp}}$ we can conclude that

$$x_0 = \Pi x_0 + (I-\Pi)x_0 \in \mathcal{C}^{\mathrm{imp}} + \mathcal{O}^{\mathrm{imp}},$$

which proves the desired result.

($\Leftarrow$). Let $x_0 = p_0 + q_0$ for some $p_0 \in \mathcal{C}^{\mathrm{imp}}$ and $q_0 \in \mathcal{O}^{\mathrm{imp}}$. Then by definition of $\mathcal{C}^{\mathrm{imp}}$ there exists an impulse-free input u such that (p, u) satisfying $p(t_0^-) = p_0$ is impulse-free, i.e., $p[t] = 0$ for all $t \geqslant t_0$. As $E^{\mathrm{imp}}(I-\Pi) = E^{\mathrm{imp}}$ the solution $(q, 0)$ with $q(t_0^-) = q_0$ will satisfy

$$\begin{aligned} Cq[t_0] &= -C\sum_{i=0}^{\nu-1}\left(E^{\mathrm{imp}}\right)^{i+1}(I-\Pi)\,q_0\delta^{(i)} \\ &= -C\sum_{i=0}^{\nu-1}\left(E^{\mathrm{imp}}\right)^{i+1} q_0\delta^{(i)} = 0. \end{aligned}$$

Hence the solution $(q, 0)$ with $q(t_0^-) = q_0$ will only generate a Dirac impulse at t_0, which will not appear in the output y. By linearity of solutions, (x, u) with $x(t_0^-) = x_0$ will satisfy $x(t) = p(t) + q(t)$ and hence

$$y[t] = Cx[t] + Du[t] = C(p[t] + q[t]) = Cq[t] = 0.$$

Hence u is an impulse-free input such that (x, u) with $x(t_0^-) = x_0$ ensures $y[t] = 0$.

Proof of Lemma 24

In order to prove the statement, we have to prove that $E\dot{x} = EW\dot{x}$. As $x(t_0^-) = x_0 \in \mathcal{C}^{\mathrm{imp}}$, it follows that $(I-W)x(t) = 0$ on $[t_0, \infty)$. Consequently $(I-W)\dot{x}(t) = 0$ on (t_0, ∞). Therefore $E\dot{x} = EW\dot{x}$ on (t_0, ∞). It remains to show that $E\dot{x}[t_0] = EW\dot{x}[t_0]$. However as $(I-W)x(t_0^-) = 0$, it follows that

$$E\dot{x}[t_0] = E(W\dot{x} + (I-W)\dot{x}[t_0]) = EW\dot{x}[t_0].$$

Conversely, assume that x solves $EW\dot{x} = Ax + Bu$. Then since $(I-W)x(t) \in \ker E$ and hence $(I-W)\dot{x}(t) \in \ker E$. Consequently, $EW\dot{x} = E(W + (I-W))\dot{x} = E\dot{x}$. Furthermore, since $x_0 \in \operatorname{im}W$, and $x(t_0^-) = x_0 = Wx_0$, it follows that $EW\dot{x}[t_0] = E(W\dot{x}[t_0] + (I-W)\dot{x}[t_0]) = E\dot{x}[t_0]$.

Now it remains to prove regularity of (EW, A) and impulse-controllability of $EW\dot{x} = Ax + Bu$. Let $\mathcal{C}_W^{\mathrm{imp}}$ be the impulse-controllable space of this auxiliary DAE. It follows from the above that $\mathcal{C}^{\mathrm{imp}} \subseteq \mathcal{C}_W^{\mathrm{imp}}$. However, since $\ker EW \subseteq \mathcal{C}_W^{\mathrm{imp}}$ it follows that

$$(\mathcal{C}^{\mathrm{imp}})^\perp \subseteq \ker EW + (\mathcal{C}^{\mathrm{imp}})^\perp = \ker EW \subseteq \mathcal{C}_W^{\mathrm{imp}},$$

and thus $\mathbb{R}^n \subseteq \mathcal{C}_W^{\mathrm{imp}}$. Hence the auxiliary DAE is impulse controllable.

Finally, for $x \in \mathcal{C}^{\mathrm{imp}}$ it follows that $EWx = Ex$ and hence $(\lambda EW - A)x = (\lambda E - A)x \neq 0$. Next observe that

$$\begin{aligned} \mathbb{R}^n &= \operatorname{im}EW + A\ker EW + \operatorname{im}B \\ &\subseteq \operatorname{im}E + A\ker E + \operatorname{im}B + A(\mathcal{C}^{\mathrm{imp}})^\perp \\ &\subseteq \mathcal{C}^{\mathrm{imp}} + A(\mathcal{C}^{\mathrm{imp}})^\perp \end{aligned}$$

and hence $Ax \neq 0$ for $x \in (\mathcal{C}^{\mathrm{imp}})^\perp$. Hence also for $x \in (\mathcal{C}^{\mathrm{imp}})^\perp$ it follows that

$$(\lambda EW - A)x = Ax \neq 0.$$

Thus we can conclude that (EW, A) is regular.

References

[1] Javier Tolsa and Miquel Salichs. Analysis of linear networks with inconsistent initial conditions. *IEEE Trans. Circuits Syst.*, 40(12):885 – 894, Dec 1993.

[2] Rudolf E. Kalman. Contributions to the theory of optimal control. *Bol. Soc. Matem. Mexico*, II. Ser. 5:102–119, 1960.

[3] Rudolf E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *ASME Trans., Part D*, 80:95–108, 1961.

[4] Rudolf E. Kalman. When is a linear control system optimal? *Trans. ASME J. Basic Eng.*, 86D:51–60, 1964.

[5] W. Murray Wonham. Optimal stationary control of a linear system with state-dependent noise. *SIAM J. Cont.*, 5(3):486–500, 1967.

[6] Jan C. Willems. Least squares optimal control and the algebraic Riccati equation. *IEEE Trans. Autom. Control*, 16:621–634, 1971.

[7] J. Daniel Cobb. Descriptor variable systems and optimal state regulation. *IEEE Trans. Autom. Control*, 28:601–611, 1983.

[8] Douglas J. Bender and Alan J. Laub. The linear-quadratic optimal regulator for descriptor systems. In *Proc. 24th IEEE Conf. Decis. Control, Ft. Lauderdale, FL*, pages 957–962, 1985.

[9] D. Bender and A. Laub. The linear quadratic optimal regulator problem for descriptor systems. *IEEE Trans. Autom. Control*, 32:672–688, 1987.

[10] Volker Mehrmann. *The Linear Quadratic Control Problem: Theory and Numerical Algorithms*. Habilitationsschrift, Universität Bielefeld, Bielefeld, FRG, 1987.

[11] Volker Mehrmann. Existence, uniqueness and stability of solutions to singular, linear-quadratic control problems. *Linear Algebra Appl.*, pages 291–331, 1989.

[12] Volker Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*. Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, 1991.

[13] Peter Kunkel and Volker Mehrmann. The linear quadratic control problem for linear descriptor systems with variable coefficients. *Math. Control Signals Syst.*, 10:247–264, 1997.

[14] Matthias Gerdts. A survey on optimal control problems with differential-algebraic equations. In *Surveys in Differential-Algebraic Equations II*, pages 103–161. Springer, 2015.

[15] Matthias Gerdts. *Optimal control of ODEs and DAEs*. Walter de Gruyter, 2012.

[16] Achim Ilchmann, Leslie Leben, Jonas Witschel, and Karl Worthmann. Optimal control of differential-algebraic equations from an ordinary differential equation perspective. *Optimal Control Applications and Methods*, 40(2):351–366, 2019.

[17] Timo Reis and Matthias Voigt. Linear-quadratic infinite time horizon optimal control for differential-algebraic equations - a new algebraic criterion. In *Proceedings of MTNS-2012*, 2012.

[18] Matthias Voigt. *On Linear-Quadratic Optimal Control and Robustness of Differential-Algebraic Systems*. PhD thesis, Otto-von-Guericke-Universität Magdeburg, publ. by Logos Verlag Berlin, Germany, 2015.

[19] Timo Reis, Olaf Rendel, and Matthias Voigt. The Kalman-Yakubovich-Popov inequality for differential-algebraic systems. Hamburger Beiträge zur Angewandten Mathematik 2014-27, Fachbereich Mathematik, Universität Hamburg, 2014. submitted for publication.

[20] Carlos E Garcia, David M Prett, and Manfred Morari. Model predictive control: Theory and practice - a survey. *Automatica*, 25(3):335–348, 1989.

[21] Manfred Morari and Jay H Lee. Model predictive control: past, present and future. *Computers and Chemical Engineering*, 23(4-5):667–682, 1999.

[22] Achim Ilchmann, Jonas Witschel, and Karl Worthmann. Model predictive control for linear differential-algebraic equations. *IFAC-PapersOnLine*, 51(20):98–103, 2018.

[23] Achim Ilchmann, Jonas Witschel, and Karl Worthmann. Model predictive control for singular differential-algebraic equations. *International Journal of Control*, pages 1–10, 2021.

[24] Daniel Liberzon and Stephan Trenn. On stability of linear switched differential algebraic equations. In *Proc. IEEE 48th Conf. on Decision and Control*, pages 2156–2161, December 2009.

[25] Daniel Liberzon and Stephan Trenn. Switched nonlinear differential algebraic equations: Solution theory, Lyapunov functions, and stability. *Automatica*, 48(5):954–963, May 2012.

[26] Andrii Mironchenko, Fabian R. Wirth, and Kai Wulff. Stabilization of switched linear differential-algebraic equations via time-dependent switching signals. In *Proc. 52nd IEEE Conf. Decis. Control, Florence, Italy*, pages 5975–5980, 2013.

[27] Andrii Mironchenko, Fabian Wirth, and Kai Wulff. Stabilization of switched linear differential algebraic equations and periodic switching. *IEEE Transactions on Automatic Control*, 60(8):2102–2113, 2015.

[28] Ferdinand Küsters, Markus G.-M. Ruppert, and Stephan Trenn. Controllability of switched differential-algebraic equations. *Syst. Control Lett.*, 78(0):32 – 39, 2015.

[29] Paul Wijnbergen and Stephan Trenn. Impulse controllability of switched differential-algebraic equations. In *2020 European Control Conference (ECC)*, pages 1561–1566. IEEE, 2020.

[30] Paul Wijnbergen and Stephan Trenn. Impulse-free interval-stabilization of switched differential algebraic equations. *Systems & Control Letters*, 149:104870, 2021.

[31] Paul Wijnbergen and Stephan Trenn. Impulse-controllability of system classes of switched differential algebraic equations. *Mathematics of Control, Signals, and Systems*, pages 1–30, 2023. to appear.

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

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

[34] Aneel Tanwani and Stephan Trenn. Determinability and state estimation for switched differential–algebraic equations. *Automatica*, 76:17–31, 2017.

[35] Aneel Tanwani and Stephan Trenn. Detectability and observer design for switched differential algebraic equations. *Automatica*, 99:289–300, 2019.

[36] Ferdinand Küsters, Stephan Trenn, and Andreas Wirsen. Switch-observer for switched linear systems. In *Proc. 56th IEEE Conf. Decis. Control, Melbourne, Australia*, 2017. to appear.

[37] Ferdinand Küsters and Stephan Trenn. Switch observability for switched linear systems. *Automatica*, 87:121–127, 2018.

[38] Ferdinand Küsters. *Switch observability for differential-algebraic systems*. PhD thesis, Department of Mathematics, University of Kaiserslautern, 2018.

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

[40] Thomas Berger, Achim Ilchmann, and Stephan Trenn. The quasi-Weierstraß form for regular matrix pencils. *Linear Algebra Appl.*, 436(10):4052–4069, 2012.

[41] Kai-Tak Wong. The eigenvalue problem $\lambda Tx + Sx$. *J. Diff. Eqns.*, 16:270–280, 1974.

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

[43] J. Daniel Cobb. Controllability, observability and duality in singular systems. *IEEE Trans. Autom. Control*, 29:1076–1082, 1984.

[44] R. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.

[45] Paul Wijnbergen and Stephan Trenn. Optimal control of daes with unconstrained terminal costs. In *Proc. 60th IEEE Conf. Decision and Control (CDC 2021)*, pages 5275–5280, 2021.

[46] Harry L. Trentelman, Anton A. Stoorvogel, and Malo L. J. Hautus. *Control Theory for Linear Systems*. Communications and Control Engineering. Springer-Verlag, London, 2001.

[47] D. R. Vaughan. A negative exponential solution for the matrix Riccati equation. *IEEE Trans. Autom. Control*, 14:72–75, 1969.
[48] Dimitri P Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena scientific Belmont, MA, 1995.