On discontinuous DAE: concept of solutions, numerical simulations, and a case study well-posedness

Alexandre Rocca[1], Vincent Acary[1], and Bernard Brogliato[1]

(1) Univ. Grenoble-Alpes, INRIA, CNRS, Grenoble INP,
LJK, 38000 Grenoble, France

Abstract. In this work, we study differential algebraic equations with constraints defined in a piecewise manner using a conditional statement. Such models classically appear in systems where constraints can evolve in a very small time frame compared to the observed time scale. The use of conditional statements or hybrid automata are a powerful way to describe such systems and are, in general, well suited to simulation with event driven numerical schemes. However, such methods may fail to efficiently simulate sliding motions and events accumulation. In contrast, the representation of such systems using differential inclusions and the methods from nonsmooth dynamics are often closer to the physical theory, but their solutions may be less intuitive or harder to interpret. Associated event-capturing numerical methods have been extensively used in mechanical modelling with success and then extended to other fields such as electronics and system biology. In a similar manner to the previous application of nonsmooth methods to the simulation of piecewise linear ODEs, we want to apply event-capturing numerical schemes to piecewise linear DAEs. In this paper, we explore three concepts of solutions for switching index-2 DAEs. To this aim, we first study in-depth the well-posedness of an example of index-2 planar dynamical system with a switching constraint, using a set-valued operators relaxation of the constraint. Then, for the same example, we give an analysis of the time-capturing implicit Euler scheme solutions and conjecture an improved numerical scheme to tackle the observed problems. In a second part, we propose three relaxations of switching constraints to obtain three concepts of solutions to switching DAEs, inspired by concepts of solutions for switching ODEs. We illustrate their properties and the possible effects on the well-posedness of switching DAEs through simple examples.

Keywords: Hybrid Systems, Switching DAE, Nonsmooth Dynamical Systems, Linear Complementarity System, Time-stepping scheme, Euler method

1 Introduction

The aim of this work is to study discontinuous differential algebraic equations (DAE), also known as hybrid DAE or switching DAE in the literature. They

can be defined, for instance, as continuous-time dynamical systems with some algebraic constraints switching with respect to the state variables. A very general framework of discontinuous DAE may be given by

$$F(t, \mathbf{y}, \dot{\mathbf{y}}) = 0 \tag{1}$$

where $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is a discontinuous, possibly set-valued function, with respect to its second and third arguments. Since it is very difficult to deal with this too general setting, a first specialisation is

$$\begin{cases} \dot{\mathbf{x}} = f(t, \mathbf{x}, \mathbf{z}) \\ 0 = g(t, \mathbf{x}, \mathbf{z}) \end{cases} \tag{2}$$

where $g : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ is a discontinuous, possibly set-valued function, with respect to its second and third arguments.

Such discontinuous (hybrid or switching) DAE systems are used in numerous fields from electronics [1], mechanics [8] to chemical process engineering [25]. They are especially used in model-based design through the use of languages like MODELICA as in [13]. In such languages, one can build a model defined by a finite number of dynamical systems whose switching is dictated by conditional statements over the time or the state variables. In general, after the compilation, models expressed in such languages results in the definition of so-called flat, multi-mode or hybrid, DAEs such as the ones studied in [12] and [7] that are particular cases of discontinuous DAEs (2). These discontinuous DAEs are then simulated using event-based methods. However, few works focus on the study and analysis of concept of solutions for such discontinuous DAEs, resulting in unclear simulating behaviours at mode change, outside of some restricted context.

1.1 Piecewise linear DAEs with a unique vector field.

The framework in (2) is still relatively too general to be able to provide us with useful insights on the concept of solutions. This is the reason why we focus our attention in this article on piecewise linear DAEs, with a constant vector field, that we wish to study from the point of view of nonsmooth dynamics [8], and event-capturing time integration methods [2].

Let us define a piecewise linear DAE with a unique vector field.

Definition 1 (Piecewise linear DAE with a unique vector field). *Let n, n_1, n_2 be three integers such that $n_1 + n_2 = n$. Let $\mathbf{x} : \mathbb{R} \to \mathbb{R}^{n_1}$ be a function of time t called the differential variable. Let $\mathbf{z} : \mathbb{R} \to \mathbb{R}^{n_2}$ be a function of time t called the algebraic variable. Let $\mathrm{A} \in \mathbb{R}^{n_1 \times n_1}$, $\mathrm{B} \in \mathbb{R}^{n_1 \times n_2}$ be two constant matrices, and $\mathrm{B}_i \in \mathbb{R}^{n_1 \times n_2}$ $\mathrm{C}_i \in \mathbb{R}^{n_2 \times n_1}$, a family of matrices indexed by i. A piecewise linear DAE with a unique vector field is defined by*

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathrm{A}\mathbf{x}(t) + \mathrm{B}\mathbf{z}(t) + \mathbf{b} \\ \quad 0 = \mathbf{g}_i(\mathbf{x}(t), \mathbf{z}(t)) = \mathrm{C}_i\mathbf{x}(t) + \mathrm{D}_i\mathbf{z}(t) + \mathbf{q}_i \\ \qquad \forall (\mathbf{x}(t), \mathbf{z}(t)) \in \mathcal{X}_i \,, \end{cases} \tag{3}$$

where the sets $\mathcal{X}_i = \{(\mathbf{x},\mathbf{z}) \in \mathbb{R}^n \mid \mathbf{h}_i(\mathbf{x},\mathbf{z}) = H_i\mathbf{x}(t) + F_i\mathbf{z}(t) + \mathbf{p}_i > 0\} \subset \mathbb{R}^n$ *define a partition of* $\mathbb{R}^n$ *such that:*

- $\bigcup_i \overline{\mathcal{X}}_i = \mathbb{R}^n$,
- $\mathrm{int}(\mathcal{X}_i) \neq \emptyset\,, \quad \forall i$,
- *for* $i \neq j$, $\mathcal{X}_i \cap \mathcal{X}_j = \emptyset$.

Using step-functions[1], we can build in a similar fashion to [3] a generalised constraint:

$$\mathbf{g}(\mathbf{x},\mathbf{z}) = \sum_i \left(\prod_{j\neq i} (1 - \mathrm{s}^+(\mathbf{h}_j(\mathbf{x},\mathbf{z}))) \right) \mathrm{s}^+(\mathbf{h}_i(\mathbf{x},\mathbf{z}))\mathbf{g}_i(\mathbf{x},\mathbf{z}) = \ 0\,, \tag{4}$$

where $\mathrm{s}^+(\mathbf{y}) = 0$ if $\mathbf{y} < 0$ and $\mathrm{s}^+(\mathbf{y}) = 1$ if $\mathbf{y} > 0$. The behaviour at $\mathbf{y} = 0$ is not stipulated yet, as it will depend on later relaxations: either using a multivalued step function relaxation or using a convex relaxation of the switching constraint, as it is explained in Section 4. In particular, in the context of piecewise ODE, the work of [3] shows that methods for nonsmooth dynamics can be efficiently applied using such a transformation. Then, depending on the concept of solutions applied on the switching surfaces (using convexification as in [11], or using multivalued functions as in [5]), the resulting solutions may differ. Here, we study the extension of such concepts of solutions when applied to switching constraints instead of switching ODE.

1.2 Related work

Various authors already studied the field of hybrid DAE. For example, DAE including complementarity constraints are a subset of differential variational inequalities (DVI). DVIs are defined and studied in [19]. In particular, they analyse the well-posedness of index-one and mixed index between 1 and 2 DVIs. In a similar manner, the authors of [1] study switching DAEs in the context of switching electrical systems. They show that such systems can be expressed as Mixed Complementarity Systems[2] (MCS), [1, Chapter 4], and efficiently simulated with associated event-capturing numerical schemes. However, none of these works study solutions to discontinuous switching DAEs given in the formalism of Definition 1. Apart from variational inequalities, or complementarity systems formalisms, Matrosov [15] proposes a concept of solutions, which is inspired from the Filippov concept of solution for discontinuous ODEs [11], for discontinuous DAEs (2). Let us consider the discontinuous constraints and a differential part $g(\cdot)$, $f(\cdot) \in \mathcal{C}(\mathbb{R} \times \mathbb{R}^{n_1} \setminus S \times \mathbb{R}^{n_2})$, with S the set of discontinuities assumed as a finite union of smooth hyper-surfaces in $\mathbb{R}^{n_1}$ of co-dimension larger or equal to 1. Let us assume some function $\mathbf{z}(t)$ is given, and define F_0 the set of vector

[1] or using sign functions.
[2] Let us remark that MCS are a particular case of DVI.

fields that are solution in $(t, \mathbf{x}, \mathbf{z})$ as:

$$F_0(t,\mathbf{x},\mathbf{z}) = \left\{ \mathbf{y} \in \mathbb{R}^{n_1} \text{ such that } \begin{array}{r} \exists \mathbf{y}_k \to \mathbf{y}, \exists t_k \to t, \exists \mathbf{x}_k \to \mathbf{x}, \exists \mathbf{z}_k \to \mathbf{z} \\ \mathbf{y}_k = f(t_k,\mathbf{x}_k,\mathbf{z}_k) \\ 0 = g(t_k,\mathbf{x}_k,\mathbf{z}_k) \end{array} \right\}, \quad (5)$$

where $\mathbf{x}_k \notin S$ for all k. Then, one defines a concept of solution for discontinuous DAEs as an extension of Filippov concept of solution for discontinuous ODEs: $(\mathbf{x}(t), \mathbf{z}(t))$ is a solution of (2) if $\mathbf{x}(t)$ is absolutely continuous in t, $\mathbf{z}(t)$ is continuous in t everywhere, and if $(\mathbf{x}(t), \mathbf{z}(t))$ is a solution of the differential inclusion:

$$\dot{\mathbf{x}}(t) \in \overline{\text{co}}(F_0(t,\mathbf{x}(t),\mathbf{z}(t))) \,. \quad (6)$$

Matrosov gives sufficient conditions for existence of such solutions in [15], and sufficient conditions for uniqueness in [16].

Merhmann et al. [12,17] provide a study of well-posedness of hybrid DAE structured as hybrid automata. In addition, a numerical implementation of sliding modes for DAE systems is provided to avoid chattering when switching occurs. To this aim, they propose a concept of sliding solutions defined by the convex hull of the vector fields obtained by index reductions at the right and left limit of the discontinuity. Furthermore, the work in [17] needs explicit transition functions from one mode (DAE) to another, in addition to consistent reset conditions. Trenn [26] defines solutions of hybrid DAE with exogenous switching. In particular, he introduces a notion of distributional solutions which can also be used to efficiently solve inconsistent initial conditions of classical DAE as an exogenous switching at $t = 0$. Camlibel et al. [9] extend results of well-posedness of differential inclusions with maximal monotone right-hand sides to differential algebraic inclusions $\mathrm{P}\dot{\mathbf{x}} \in -\mathcal{F}(\mathbf{x})$ with a maximal monotone operator $\mathcal{F}(\cdot)$ and a singular matrix P. This work considers nonsmooth DAEs in the form of:

$$\begin{cases} \mathrm{E}\dot{\mathbf{y}}(t) = \mathrm{A}\mathbf{y}(t) + \mathrm{B}\boldsymbol{\lambda}(t) + \mathbf{b} \\ \mathbf{w}(t) = \mathrm{C}\mathbf{y}(t) + \mathrm{D}\boldsymbol{\lambda}(t) + \mathbf{q} \\ \mathbf{w}(t) \in \mathcal{M}(-\boldsymbol{\lambda}(t)), \end{cases} \quad (7)$$

with $\mathcal{M}(\cdot)$ a maximal monotone operator and singular matrix E. Then, assuming some passivity conditions on the Weierstrass-Kronecker form of (7), sufficient conditions for the well-posedness of (7) are given.

It is important to note that this formalism is the closest one to the example studied in Sections 2 and 3, with the notable difference that in our case the operator $\mathcal{M}(\cdot)$ is not maximal monotone, but hypo-monotone. Stechlinski et al. [25,6] and Khan [14] define from the Clarke Jacobian a notion of generalised differential index and an associated index reduction procedure in the context of nonsmooth DAE, with Lipschitz continuous constraints. Current implementation and theory are limited to semi-explicit index-1 nonsmooth DAE. Finally, let us cite another work for index reduction of hybrid DAE by Benveniste et al. [7]. This work uses non-standard analysis to construct well-defined transitions from one

mode to another in the context of hybrid DAE, even in the presence of varying index. In particular, this work pairs well with [17], which needs the knowledge of transition and re-initialisation maps when switching from one mode to another.

1.3 Outline

As we mentioned in the previous section, the case of index-1 discontinuous DAEs has been well studied in the literature. This is the reason why we focus our attention on index-2 problems where the switching surface depends, only, on the state variables. Let us give a definition of the object of interest in this article.

Definition 2 (Single-Switching Index-2 Piecewise linear DAE with a unique vector field). *Let us consider a piecewise linear DAE with a unique vector field as in Definition 1. When considering a single switching surface and index-2 DAEs in each mode,* (3) *can be reduced to:*

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathrm{A}\mathbf{x}(t) + \mathrm{B}\mathbf{z}(t) + \mathbf{b} \\ \quad 0 = \mathbf{g}_1(\mathbf{x}(t)) = C_1\mathbf{x}(t) + \mathbf{q}_1 \quad \text{if } \mathbf{x}(t) \in \mathcal{X}_1 \\ \quad 0 = \mathbf{g}_2(\mathbf{x}(t)) = C_2\mathbf{x}(t) + \mathbf{q}_2 \quad \text{if } \mathbf{x}(t) \in \mathcal{X}_2\,, \end{cases} \tag{8}$$

with $\mathcal{X}_1 = \{\mathbf{x} \in \mathbb{R}^{n_1} \mid h_1(\mathbf{x}) = -h(\mathbf{x}) = -\mathrm{H}\mathbf{x}(t) - \mathbf{p} > 0\} \subset \mathbb{R}^{n_1}$ *and* $\mathcal{X}_2 = \{\mathbf{x} \in \mathbb{R}^{n_1} \mid h_2(\mathbf{x}) = h(\mathbf{x}) = \mathrm{H}\mathbf{x}(t) + \mathbf{p} > 0\} \subset \mathbb{R}^{n_1}$. *In addition,* (8) *is of differential index 2 in each mode and it follows that the matrices* $\mathrm{C}_i\mathrm{B}$ *are non-singular for all* $i \in \{1, 2\}$.

Throughout the paper, a working example of (8) will be detailed to give some insights on possible concepts of solutions and to large varieties of solutions we can obtain. Along this paper, we may call "left-hand constraint", the constraint $\mathrm{C}_1\mathbf{x}(t) + \mathbf{q}_1 = 0$ that is active if $\mathbf{x} \in \mathcal{X}_1$ (or equivalently $h(\mathbf{x}) < 0$). Similarly, we may call "right-hand constraint", the constraint $\mathrm{C}_2\mathbf{x}(t) + \mathbf{q}_2 = 0$ that is active if $\mathbf{x} \in \mathcal{X}_2$ (or equivalently $h(\mathbf{x}) > 0$). To study the solutions of systems in the form of (8), we construct a relaxation of these two constraints along the switching surface $h(\mathbf{x}, \mathbf{z}) = 0$ by "filling-in the gap" (see [18]). One way to construct such relaxed constraint in $\mathbf{h}(\mathbf{x}, \mathbf{z}) = 0$ is by considering the convex hull of the left and right limit of $\mathbf{g}(\mathbf{x}, \mathbf{z})$ when $\mathbf{h}(\cdot) < 0$ and $\mathbf{h}(\cdot) > 0$ respectively. We could also consider multi-valued step functions in (4) in a similar fashion to [5] for discontinuous ODEs. For the working example of Section 2, we consider the convexification of the constraints along the switching surface.

As we have seen, most works consider either a high index hybrid DAE framework with event-driven numerical methods and explicit transition functions, or index-1 DAE with nonsmooth constraints aiming to rewrite the system as a differential inclusion[3] into a Lipschitz function, or a maximal monotone operator. In this paper, we are looking to make a bridge between the state-dependent switching DAE formalism, and the nonsmooth DAE formalism studied by DVI [19] and MCS [1]. With this in mind, in Section 2 we first study in details the

[3] A differential algebraic inclusion in the case of [9] state-dependent switching.

well-posedness of nonsmooth DAEs obtained by a relaxation of the constraints on a simple working example. In Section 3, we analyse how the classical nonsmooth numerical methods perform in this context. We also propose some modification to the numerical scheme to overcome issues observed in this context. Part of the results presented in the Sections 2 and 3, were already briefly presented without proofs in [21]. In Section 4, we study various relaxation methods to fill the graph of the constraint by either a continuous extension of the left and right constraints, a relaxation by multi-valued step functions, or finally a relaxation by convexification. We also discuss their effect on the existence of solutions for index-2 hybrid DAEs through another illustrative example. Conclusions end the article in Section 5. Proofs and definitions are given in Appendix.

2 Analysis of a discontinuous DAE Example

Let us consider the following switching DAE:

$$\begin{cases} \dot{x}_1(t) = 1 + \mathrm{B}_1 z(t) \\ \dot{x}_2(t) = \mathrm{B}_2 z(t) \end{cases} \tag{9a}$$

$$\texttt{if } x_1 < 0 \texttt{ then :} \quad 0 = -1 - x_1(t) - x_2(t)\,, \tag{9b}$$

$$\texttt{if } x_1 > 0 \texttt{ then :} \quad 0 = 1 + x_1(t) - x_2(t) \tag{9c}$$

which is a particular case of (8) with $\mathbf{x} = (x_1, x_2)^{\mathrm{T}}$, $\mathrm{A} = 0$, $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$, $\mathrm{C}_1 = (-1, -1)$, $\mathrm{C}_2 = (1, -1)$, $\mathrm{H} = (1, 0)$, $\mathbf{b} = (1, 0)^{\mathrm{T}}$, $(q_1, q_2, p) = (-1, 1, 0)$. In $x_1 = 0$, the system does not have continuous solutions whatever the active constraint is, so we keep strict inequalities in (9).

The main objective of this article is to show that despite each DAE, considered separately, possesses simple dynamics, the switching DAE in (9) *has, on the contrary, surprisingly complex dynamics.* As exposed in the introduction, we search to obtain absolutely continuous (AC) solutions on the switching surface $x_1 = 0$ by considering a set-valued relaxation of the constraint. It follows that the switching constraints can be embedded into a set-valued constraint obtained by convexification, or by filling-in the gap. Indeed, in a similar manner to the regularisation of solution for switching ODE [3], we construct the nonsmooth DAE system

$$\begin{cases} \dot{x}_1(t) = 1 + \mathrm{B}_1 z(t) \\ \dot{x}_2(t) = \mathrm{B}_2 z(t) \end{cases} \tag{10a}$$

$$\begin{cases} 0 = \lambda(t)(1 + x_1(t)) - x_2(t) \\ \lambda(t) \in \operatorname{sign}(x_1(t))\,, \end{cases} \tag{10b}$$

where the relaxed constraint is $0 \in \lambda(x_1+1) - x_2$ with $\lambda = \text{sign}(x_1)$ and $\text{sign}(\cdot)$ is the set-valued operator:

$$\text{sign}(x) = \begin{cases} \{-1\} & \text{if } x < 0 \\ [-1,1] & \text{if } x = 0 \\ \{1\} & \text{if } x > 0\,. \end{cases} \tag{11}$$

The set-valued algebraic constraint in (10b) equals the ones of (9) when $x_1 < 0$ (respectively $x_1 > 0$), and is a convex relaxation of both in $x_1 = 0$: that is $(x_1, x_2) \in \overline{\text{co}}(\begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} \{0\} \\ [-1,1] \end{pmatrix}$ (see Figure 1). Let us remark that (10b) is a particular case of (4) and will be further studied in Section 4.2 and 4.3.

If $x_1(t) \leq 0$ and (9b) is satisfied then we say that the nonsmooth DAE (10) is in mode 1. In addition, if $x_1(t) < 0$ and (9b) is satisfied then we say that (10) is *strictly* in mode 1. Similarly, If $x_1(t) \geq 0$ and the constraint (9c) is satisfied then we say that the nonsmooth DAE (10) is in mode 2, and strictly in mode 2 if $x_1(t) > 0$. Finally, if $x_1(t) = 0$ and $x_2(t) \in [-1,1]$ then we say that the nonsmooth DAE (10) is in mode 3, and strictly in mode 3 if $x_2(t) \in (-1,1)$.

In Section 2.1, we first study existence of AC solutions $\mathbf{x}(t)$ to (10). Then, in Section 2.2 we compare the solutions of (10) to the existing concept of solutions applied to (9). In particular, we discuss the solutions proposed by [17] and the conditions for the existence of a sliding motion on the switching surface $x_1 = 0$. Finally, in Section 2.3 we study the extension to solutions of bounded variations, and the definition of a jump law when an AC solution cannot be continued.

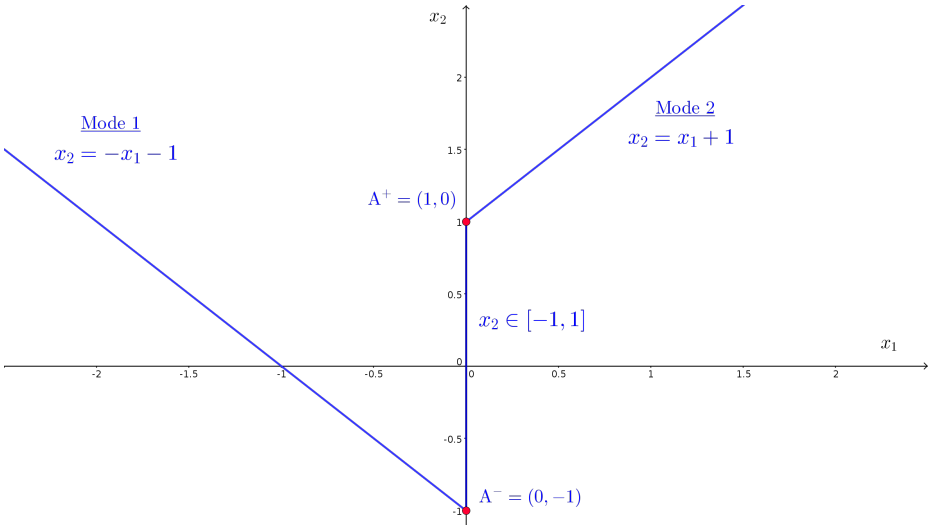


Fig. 1: Phase-space representation of the constraint of (9) and (10)

2.1 Analysis of Absolutely Continuous(AC) solutions

Let us study the conditions on the differential part (9a) of the DAE, and in particular B_1, B_2 for the existence of a sliding mode along the switching surface $x_1 = 0$; in other words, the existence of an AC solution $x_1(t)$, $x_2(t)$ for some arbitrary time interval and initial conditions to the following problem:

Problem 1. *Let $I \subset \mathbb{R}$ a compact interval such that $t_0 = \inf I$. The problem is to find an absolutely continuous(AC) function $\mathbf{x}(\cdot) = (x_1(\cdot), x_2(\cdot))$ on I, and Lebesgue integrable ($\mathcal{L}^1(I)$) functions $z(\cdot), \lambda(\cdot)$ such that for the nonsmooth DAE* (10) *the differential part* (10a) *is satisfied almost everywhere, and the constraint* (10b) *is satisfied everywhere, given the initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$.*

If it exists $\varepsilon > 0$ such that the solution exists on an interval $I = [t_0, t_0 + \varepsilon)$, the solution is called a local solution. If a local solution cannot be extended, it is called a maximal solution. If the solution exists on $I = [0, T]$ for any $T > t_0$, then the solution is called a global solution.

Let us define the concept of feasible initial condition for Problem 1.

Definition 3 (Feasible initial condition). *The initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$ in Problem 1 is said to be feasible if it exists $\lambda(t_0) \in \mathbb{R}$ such that*

$$\begin{cases} \lambda(t_0) \in \operatorname{sign}(x_1(t_0)) \\ -x_2(t_0) + \lambda(t_0) x_1(t_0) + \lambda(t_0) = 0. \end{cases} \tag{12}$$

Clearly, a feasible initial condition is a necessary condition to obtain a solution to Problem 1. The constraint (10b) in Problem 1 can be rewritten equivalently as the following set-valued equation:

$$\begin{aligned} 0 &\in -x_2(t) + |x_1(t)| + \operatorname{sign}(x_1(t)) \\ \Leftrightarrow\ x_1(t) &\in \mathcal{N}_{[-1,1]}(-x_2(t) + |x_1(t)|)\ , \end{aligned} \tag{12'}$$

where the equivalence is obtained by using the inversion of subdifferentials of convex lower-semicontinuous functions [22]. Thus, the original switching DAE in (9), which is embedded in (3), is recast in a new formalism:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathrm{A}\mathbf{x}(t) + \mathrm{B}\mathbf{z}(t) + \mathbf{b} \\ \quad 0 \in \mathcal{F}(\mathbf{x})\,, \end{cases} \tag{13}$$

with $\mathcal{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}$.

Proposition 1 (Local solutions to Problem 1). *Let us state necessary and sufficient conditions on $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^T$ such that Problem 1 has, at least, one solution on $[t_0, t_0 + \epsilon)$, for some $0 < \epsilon$, and $\mathbf{x}(t_0) = \mathbf{x}_0$ is a feasible initial condition.*

- *If* (10) *is strictly in mode 1 at t_0, then there exists a local solution to Problem 1 if and only if*

$$\mathrm{B}_1 + \mathrm{B}_2 \neq 0 \tag{14}$$

– *If* (10) *is strictly in mode 2 at* t_0*, then there exists a local solution to Problem 1 if and only if*

$$\mathrm{B}_1 - \mathrm{B}_2 \neq 0 \tag{15}$$

– *If* (10) *is strictly in mode 3 at* t_0*, then there exists a local solution to Problem 1 if and only if*

$$\mathrm{B}_1 \neq 0 \tag{16}$$

– *If* $\mathbf{x}(t_0) = (0,-1)^{\mathrm{T}} = A^-$*, then there exists a local solution to Problem 1 if and only if*

$$\frac{\mathrm{B}_2}{\mathrm{B}_1} \leq 0 \tag{17}$$

– *If* $\mathbf{x}(t_0) = (0,1)^{\mathrm{T}} = A^+$*, then there exists a local solution to Problem 1 for all* $\mathrm{B} \in \mathbb{R}^2$*.*

Proof. Let us first consider the conditions of existence of solutions when strictly in each mode: mode 1 ($x_1 < 0$), mode 2 ($x_1 > 0$), mode 3 ($x_1 = 0$, $x_2 \in (-1,1)$). Then, we consider the conditions at the switching points $A^- = (0,-1)^{\mathrm{T}}$ and $A^+ = (0,1)^{\mathrm{T}}$.

1. Assume $x_1(t_0) < 0$ for some $t_0 \geq 0$. Let us consider a local solution staying in mode 1: $x_1(t) \leq 0$, $\lambda(t) = -1$ for all $t \in [t_0, t_0+\varepsilon] \triangleq I_\varepsilon$, with $\varepsilon > 0$. For $t \in I_\varepsilon$, an AC solution of (10) in mode 1 must satisfy almost everywhere:

$$\begin{aligned} -\dot{x}_2(t) - \dot{x}_1(t) = 0 &\Leftrightarrow (\mathrm{B}_2 + \mathrm{B}_1) z(t) + 1 = 0 \\ &\Leftrightarrow z(t) = \frac{-1}{\mathrm{B}_2 + \mathrm{B}_1}\,. \end{aligned} \tag{18}$$

   Therefore, $z(\cdot)$ is Lebesgue integrable if and only if $(\mathrm{B}_2 + \mathrm{B}_1) \neq 0$, and then Problem 1 has a solution. On the contrary, we say that mode 1 is not feasible if $(\mathrm{B}_2 + \mathrm{B}_1) = 0$. Let us notice (see (9) and (10)) that $(\mathrm{B}_2 + \mathrm{B}_1) \neq 0$ corresponds to non-singular $\mathrm{C}_1\mathrm{B}$ with $\mathrm{C}_1 = (-1,-1)$. We deduce that there exists an local AC solution in mode 1 if and only if $(\mathrm{B}_2 + \mathrm{B}_1) \neq 0$, and we have

$$\dot{x}_1(t) = \frac{\mathrm{B}_2}{\mathrm{B}_2 + \mathrm{B}_1}\,. \tag{19}$$

   Let us now examine the upper bound t_1 such that a solution in mode 1 continues to exist. From (19), we obtain:
   – If $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 \leq 0$ then $\dot{x}_1(t) \leq 0$, for all $t \in [t_0, +\infty) \triangleq I_\infty$: there is a global solution in mode 1.
   – If $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 > 0$ then $\dot{x}_1(t) > 0$, for $t > t_0$. Then, for $t_1 = -x_1(t_0)\frac{\mathrm{B}_1+\mathrm{B}_2}{\mathrm{B}_2} + t_0$ we have $\mathbf{x}(t_1) = (0,-1)$ and the trajectory must leave mode 1 in a right-neighbourhood of t by continuation in another mode, if this is possible.

   To conclude this case, a sufficient and necessary condition for existence of a local solution of Problem 1 in mode 1 on $[t_0, T]$ for $T = t_1$ is:

$$\mathrm{B}_1 + \mathrm{B}_2 \neq 0\,. \tag{20}$$

2. Similarly, assume $x_1(t_0) > 0$. Let us consider a solution staying in mode 2: $x_1(t) \geq 0$, $\lambda(t) = 1$ for all $t \in I_\varepsilon$. Hence, for $t \in I_\varepsilon$, a solution of (10) in mode 2 must satisfy almost everywhere:

$$\begin{aligned} -\dot{x}_2(t) + \dot{x}_1(t) = 0 \Leftrightarrow & -\mathrm{B}_2 z(t) + \mathrm{B}_1 z(t) + 1 = 0 \\ \Leftrightarrow z(t) = & \frac{-1}{\mathrm{B}_1 - \mathrm{B}_2}\,, \end{aligned} \tag{21}$$

and this mode is feasible if and only if $(\mathrm{B}_1 - \mathrm{B}_2) \neq 0$. Similarly to the previous case, this corresponds to non-singular $\mathrm{C}_2\mathrm{B}$ with $\mathrm{C}_2 = (1, -1)$. We deduce that there exists a local AC solution in mode 2 if and only if $(\mathrm{B}_1 - \mathrm{B}_2) \neq 0$, and we have

$$\dot{x}_1(t) = \frac{-\mathrm{B}_2}{\mathrm{B}_1 - \mathrm{B}_2}\,. \tag{22}$$

Let us now examine the upper bound t_1 such that a solution in mode 2 continues to exist. It follows that :

- If $(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 \leq 0$ then $\dot{x}_1(t) \geq 0$, for all $t \in^{\triangle} I_\infty$: there is a global solution in mode 2.
- If $(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 > 0$ then $\dot{x}_1(t) < 0$, for $t > t_0$. Furthermore, for $t_1 = x_1(t_0)\frac{\mathrm{B}_1 - \mathrm{B}_2}{\mathrm{B}_2} + t_0$ we have $\mathbf{x}(t_1) = (0, 1)$ and the solution cannot be continued further in mode 2 and must leave mode 2 in a right-neighbourhood of t by continuation in another mode, if this is possible.

To conclude this case, a sufficient and necessary condition for existence of a local solution of Problem 1 in mode 2 on $[t_0, T]$ for $T = t_1$ is:

$$\mathrm{B}_1 - \mathrm{B}_2 \neq 0\,. \tag{23}$$

3. Assume there exists $t_0 \geq 0$ such that $x_1(t_0) = 0$ and $x_2(t_0) \in (-1, 1)$. Let us consider a solution staying in mode 3, $x_1(t) = 0$ and $x_2(t) \in [-1, 1]$, for all $t \in I_\varepsilon$. For $t \in I_\varepsilon$, a solution in mode 3 must satisfies almost everywhere:

$$\begin{cases} \dot{x}_1(t) = 0 \Leftrightarrow \mathrm{B}_1 z(t) = -1 \\ \dot{x}_2(t) = \dfrac{-\mathrm{B}_2}{\mathrm{B}_1}\,. \end{cases} \tag{24}$$

In addition, the constraint in (12) becomes $\lambda(t) = x_2(t) \in [-1, 1]$. In a similar way to modes 1 and 2, there exists a local solution in mode 3 if and only if:

$$\mathrm{B}_1 \neq 0 \tag{25}$$

This corresponds to non-singular $\mathrm{C}_3\mathrm{B}$ with $\mathrm{C}_3 = (1, 0)$. In addition, a solution exists in mode 3 only if $-1 \leq x_2(t) \leq 1$: if it reaches $x_2(t) = 1$ or -1 then it leaves mode 3 in a right-neighbourhood of t by continuation in another mode, if this is possible. From (24), the solutions staying in mode 3 are either:

- Constant if $\mathrm{B}_2 = 0$, and there is a global solution in mode 3.

- Going 'downward' if $\mathrm{B}_2/\mathrm{B}_1 > 0$ such that $\dot{x}_2 < 0$. Then, there exists $t_1 \geq t_0$ given by $t_1 = \frac{\mathrm{B}_1}{\mathrm{B}_2}(1+x_2(t_0))+t_0$ such that $x_2(t_1) = -1$ and the solution cannot stay in mode 3 in a right-neighbourhood of t_1: continuation, if any, must occur in mode 1.
- Going 'upward' if $\mathrm{B}_2/\mathrm{B}_1 < 0$ such that $\dot{x}_2 > 0$. Then, there exists $t_1 \geq t_0$ given by $t_1 = \frac{\mathrm{B}_1}{\mathrm{B}_2}(x_2(t_0) - 1) + t_0$ such that $x_2(t_1) = 1$ and the solution cannot stay in mode 3 in a right-neighbourhood of t_1: continuation, if any, must occur in mode 2.

4. Assume there exists $t_0 \geq 0$ such that $\mathbf{x}(t_0) = A^-$. The point A^- is both at the border of mode 1, and at the border of mode 3 so there may exist a local solution with continuation in either of these modes.
   Let us first consider the case of a continuation with a local solution in mode 1, which is equivalent to $x_1(t) \leq 0$ and $\lambda(t) = -1$, for all $t \in I_\varepsilon$. As $x_1(t_0) = 0$, we note from case 1 that there exists a local solution in mode 1 if and only if $\dot{x}_1(t) \leq 0$ almost every where, and this is equivalent to

$$(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 \leq 0\,,\ (\mathrm{B}_1 + \mathrm{B}_2) \neq 0\,. \tag{26}$$

   Let us now consider the case of a continuation with a local solution in mode 3, which is equivalent to $x_1(t) = 0$ and $x_2(t) = [-1, 1]$, for all $t \in I_\varepsilon$. As $x_2(t_0) = -1$, we note from the 'upward' case 3 that there exists a local solution in mode 3 if and only if $\dot{x}_2(t) \geq 0$ almost everywhere, and this is equivalent to

$$\frac{\mathrm{B}_2}{\mathrm{B}_1} \leq 0\,,\ \mathrm{B}_1 \neq 0. \tag{27}$$

   We can conclude that there is existence of local solution in $\mathbf{x}(t_0) = A^-$ if and only if (26) or (27) are satisfied. Then, let us notice that any $\mathrm{B} \in \mathbb{R}^2$ satisfying (26) also satisfies (27), and the union of these two conditions yields (27).
5. Assume there exists $t_0 \geq 0$ such that $\mathbf{x}(t_0) = A^+$. The point A^+ is both at the border of mode 2, and at the border of mode 3 so there may exist a local solution with continuation in either of these modes.
   Let us first consider the case of a local solution in mode 2, which is equivalent to $x_1(t) \geq 0$ and $\lambda(t) = 1$, for all $t \in I_\varepsilon$. As $x_1(t_0) = 0$, we note from case 2 that there exists a local solution in mode 2 if and only if $\dot{x}_2(t) \geq 0$ almost every where, and this is equivalent to

$$(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 \leq 0\,,\ (\mathrm{B}_1 - \mathrm{B}_2) \neq 0\,. \tag{28}$$

   Let us now consider the case of a local solution in mode 3, which is equivalent to $x_1(t) = 0$ and $x_2(t) = [-1, 1]$, for all $t \in I_\varepsilon$. As $x_2(t_0) = 1$, we note from the 'downward' case 3 that there exists a local solution in mode 3 if and only if $\dot{x}_2(t) \leq 0$ almost every where, and this is equivalent to

$$\frac{\mathrm{B}_2}{\mathrm{B}_1} \geq 0\ \ \mathrm{B}_1 \neq 0. \tag{29}$$

We can conclude that there is existence of a local solution in $\mathbf{x}(t_0) = A^+$ if and only if (28) or (29) are satisfied. The union of these two conditions is satisfied for all $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}} \in \mathbb{R}^2$. The proof is complete.

□

Let us remark that the conditions (14) and (15) in Proposition 1 are always satisfied under the assumption that (9) respects Definition 2 and is of index-2 in mode 1 and 2. In the remainder of this paper, we consider this is true, and we assume that (14) and (15) are always satisfied.

Let us denote $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^2 \text{ sol. of (12')}\}$ and $\mathcal{T}_{\mathcal{C}}(\mathbf{x})$ the contingent cone of $\mathcal{C}$ at the point $\mathbf{x}$, see Definition 14. For our particular example, the contingent cone is given in the following result.

Lemma 1. *The computation of the contingent cone to $\mathcal{C}$ at the point $\mathbf{x}_0$, i.e., $\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)$, can be separated in 5 cases as follows:*

- *If $\mathbf{x}_0$ is such that $0 = x_{2,0} + x_{1,0} + 1$ and $x_{1,0} < 0$, that is $\mathbf{x}_0$ strictly in mode 1, then:*
$$\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^2 \mid 0 = x_2 + x_1\} \tag{30}$$
- *If $\mathbf{x}_0$ is such that $0 = -x_{2,0} + x_{1,0} + 1$ and $x_{1,0} > 0$, that is $\mathbf{x}_0$ strictly in mode 2, then:*
$$\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^2 \mid 0 = -x_2 + x_1\} \tag{31}$$
- *If $\mathbf{x}_0$ is such that $0 = x_{1,0}$ and $x_{2,0} \in (-1, 1)$, that is $\mathbf{x}_0$ strictly in mode 3, then:*
$$\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^2 \mid 0 = -x_1\} \tag{32}$$
- *If $\mathbf{x}_0 = A^-$, then $\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)$, is constituted of two half lines:*
$$\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 = 0, x_2 \geq 0\} \cup \{\mathbf{x} \in \mathbb{R}^2 \mid x_2 = -x_1, x_1 \leq 0\}\,, \tag{33}$$
- *If $\mathbf{x}_0 = A^+$, then $\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)$, is constituted of two half lines:*
$$\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 = 0, x_2 \leq 0\} \cup \{\mathbf{x} \in \mathbb{R}^2 \mid x_2 = x_1, x_1 \geq 0\} \tag{34}$$

Proposition 2. *Let $\mathbf{x}(t_0) = \mathbf{x}_0$ be a feasible initial condition. A necessary and sufficient condition for a local solution to Problem 1 is:*

$$\begin{cases} \mathbf{x}_0 \in \mathcal{C}\,. \\ \exists\, z(t_0^+) \in \mathbb{R} \text{ such that } (1,0)^{\mathrm{T}} + \mathrm{B} z(t_0^+) \in \mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)\,. \end{cases} \tag{35}$$

It follows that a necessary and sufficient condition for local solutions to Problem 1 is

$$(\mathbf{x}(t_0), \dot{\mathbf{x}}(t_0^+)) \in \mathcal{C} \times \mathcal{T}_{\mathcal{C}}(\mathbf{x}(t_0))\,. \tag{36}$$

Proof. Let us consider the computation of $\mathcal{T}_{\mathcal{C}}(\mathbf{x})$ in Lemma 1. One notices that outside of the points A^- and A^+, the contingent cone is the tangent space to the manifold associated with the constraint in (9b), (9c) or $x_1 = 0$. Then, considering the classical conditions for solutions of differential equations on manifold [20, Theorem 6.2] we can assert that for an initial condition $\mathbf{x}_0 \in \mathcal{C} \setminus \{A^-, A^+\}$, a necessary and sufficient condition for the existence of a local solution to Problem 1 is:

$$\begin{cases} \mathbf{x}_0 \in \mathcal{C}\,. \\ \dot{\mathbf{x}}(t_0^+) \in \mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)\,, \end{cases} \tag{37}$$

that is solution of (35).

Let us assume that $\mathbf{x}_0 = A^-$ and $\dot{\mathbf{x}}(t_0^+) \in \mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)$, with $\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)$ given in (33), it follows that:

$$\begin{cases} \dot{x}_1(t_0^+) = 0\,,\ \dot{x}_2(t_0^+) \geq 0 & \text{(38a)} \\ \text{ or} \\ \dot{x}_2(t_0^+) = -\dot{x}_1(t_0^+)\,,\ \dot{x}_1(t_0^+) \leq 0\,. & \text{(38b)} \end{cases}$$

If $\dot{\mathbf{x}}(t_0^+)$ satisfies (38a), then from (24), it is equivalent to $\mathrm{B}_2/\mathrm{B}_1 \leq 0$ and from (27) it is equivalent to the existence of a local solution in A^- that can be continued in mode 3. Similarly, if $\dot{\mathbf{x}}(t_0^+)$ satisfies (38b), then from (19) it is equivalent to $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 \leq 0$, $(\mathrm{B}_1 + \mathrm{B}_2) \neq 0$, which is equivalent, by (26), to a local solution in A^- that can be continued in mode 1. One can proceed in a similar manner with $\mathbf{x}_0 = A^+$ and prove that $\dot{\mathbf{x}}(t_0^+) \in \mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)$, with $\mathcal{T}_{\mathcal{C}}(\mathbf{x}_0)$ given in (34), is a necessary and sufficient condition for the existence of local solutions to Problem 1 with $\mathbf{x}(t_0) = A^+$. This concludes the proof. □

Remark 1. Let us notice that the condition (36) on the initialisation $\mathbf{x}(t_0)$, and $\dot{\mathbf{x}}(t_0^+)$ the right limit of the derivatives, is close to the one stated, for example, in [20] for the existence of solutions to smooth DAEs where the state variables must be initialised on a manifold $\mathcal{M}$, and the derivative[4] must be in the tangent space of $\mathcal{M}$ in $\mathbf{x}_0$. In the context of smooth DAE, such initialisation is called consistent initialisation.

Now, let us state some additional results on the existence of global solutions to Problem 1 for any feasible initial condition. Such solutions are especially interesting in some modelling purposes as they characterise the possibility to initialize the system in any mode while keeping global solutions.

Proposition 3. *Problem 1 has, at least, one solution on* $[t_0, T)$*, for some* $T \leq +\infty$*, and for any feasible initial condition* $\mathbf{x}_0$*, if and only if* B_1*,* B_2 *satisfy:*

$$\begin{cases} \mathrm{B}_1 \neq 0 \\ \dfrac{\mathrm{B}_2}{\mathrm{B}_1} \leq 0 \\ (\mathrm{B}_1 + \mathrm{B}_2) \neq 0. \end{cases} \tag{39}$$

[4] When considering index 2 smooth DAEs with constraint of the form $\mathbf{g}(\mathbf{x}) = 0$ similar to our framework.

Furthermore, the algebraic variables $z(t)$ is function of bounded variation and $\lambda(t)$ is AC.

Proof. This proposition corresponds to conditions on $(\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$ such that there is existence of local solutions in any point of the constraint set $\mathcal{C}$. Then, the resulting condition is simply provided by the intersection of all the conditions for local solutions given in Proposition 1. We can also prove that $\lambda(t)$ is AC by referring to the case 3 of Proposition 1's proof: if a transition occurs at $t_\star$ from mode 2 to mode 3 then $x_2(t_\star) = 1$ and $\lambda(t_\star) = 1$. Similarly, if a transition occurs at $t_\star$ from mode 1 to mode 3 then $x_2(t_\star) = -1$ and $\lambda(t_\star) = -1$, ensuring that $\lambda(t)$ is AC. Finally, as seen in the study of the local solutions, $z(\cdot)$ is constant in each mode, and is defined to be right-continuous at mode switch. As it can be notice that the number of events is finite, we can conclude that $z(\cdot)$ is of bounded variations. □

Remark 2. The above study does not prove the uniqueness of the global solutions. Indeed, if $\mathbf{x}(t_0) = (0, -1)$ there also exist two global AC solutions, one continuing in mode 1 and one continuing in mode 3, for vectors B such that $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 < 0$ and $\mathrm{B}_2/\mathrm{B}_1 \leq 0$. For example, let us consider $\mathrm{B} = (-1, 0.5)^{\mathrm{T}}$ and $\mathbf{x}(t_0) = (0, -1)^{\mathrm{T}}$: then, there exists two global AC solutions, one continuing in mode 1, with $\dot{\mathbf{x}}(t_0^+) = (-1, 1)^{\mathrm{T}}$, and another continuing in mode 3 with $\dot{\mathbf{x}}(t_0^+) = (0, 0.5)^{\mathrm{T}}$.

Proposition 4 (Uniqueness of a solution to Proposition 3). *Problem 1 has a unique AC global solution $(\mathbf{x}(\cdot), \lambda(\cdot))$, and a unique global right-continuous solution $z(\cdot)$, for any $t > t_0$, and for any feasible initial condition $\mathbf{x}_0$, if and only if* B_1, B_2 *satisfy:*

$$\begin{cases} \mathrm{B}_1 \neq 0 \\ (\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 \geq 0 \\ (\mathrm{B}_1 + \mathrm{B}_2) \neq 0. \end{cases} \tag{40}$$

Proof. Let us first notice that there is a unique solution if $\mathrm{B}_2 = 0, \mathrm{B}_1 \neq 0$, resulting in $\dot{\mathbf{x}}(t) = 0$, and $\lambda(t)$, $z(t)$ unique and constant. Apart from the constant case, there is a unique solution everywhere outside A^- and A^+. The condition $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 > 0$ is equivalent to $\dot{x}_1(t) \geq 0$ in every mode, and there is a unique solution for $\mathbf{x}(t)$ in A^- switching from mode 1 to mode 3, and a unique solution in A^+ switching from mode 3 to mode 2. Then, $\lambda(t)$ is unique as it is AC, unique in mode 1 and 2, and equal to $x_2(t)$ in mode 3. Finally, let us remark that $z(t)$ is of bounded variations and has a unique right-continuous solution at the switching time t_1 determined at t_1^+ by the switch from mode 1 to mode 3 described in the case 4 of the proof of Proposition 1, or by the switch from mode 3 to mode 2 described in the case 5 of the proof of Proposition 1. □

Remark 3. From Remark 2 and Proposition 4 we can distinguish three categories of solutions in Proposition 3:

- If $\mathrm{B}_2 = 0$ one obtains the constant solutions, unique w.r.t. a given feasible initial condition.
- If $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 > 0$ one obtains "sliding-crossing" solutions that will cross the switching constraint if $x_1(t_0) < 0$ by sliding on the switching surface $x_1 = 0$, and unique w.r.t. a given initial condition.
- If $(\mathrm{B}_1+\mathrm{B}_2)\mathrm{B}_2 < 0$ one obtains "sliding-repulsive" solutions that are repulsive w.r.t to the switching surface if $x_1(t_0) < 0$ or $x_1(t_0) > 0$, but that will slide on the switching surface if $x_1(t_0) = 0$. In particular there exist two solutions for the initial condition $\mathbf{x}(t_0) = (0, -1)^{\mathrm{T}}$.

In Figure 2, we summarise the conditions of Proposition 3 for existence of a global AC solution for any feasible initial condition. In addition, we illustrate the three categories of solutions in Figures 2a and the associated conditions on the parameters $(\mathrm{B}_1, \mathrm{B}_2)$ are in Figure 2b.

In Proposition 3, conditions on B are given such that it exists global AC solutions to Problem 1, for all feasible initial conditions. For the other possible choices of B, there may exist local solutions to Problem 1, but only for a subset of the feasible initial conditions.

Proposition 5. *Let us state, the sets of feasible initial conditions* $\mathbf{x}(t_0) = \mathbf{x}_0$ *such that there exists a local solution to Problem 1 when conditions* (39) *are not satisfied. In particular, we can divide this study in two groups of values for* B*:* $\{\mathrm{B} \mid \mathrm{B}_1 = 0\}$*, and* $\{\mathrm{B} \mid \mathrm{B}_2/\mathrm{B}_1 > 0, \mathrm{B}_1 \neq \mathrm{B}_2, \mathrm{B}_1 \neq 0\}$.

- *Assume* $\mathrm{B}_1 = 0$*:* $\mathbf{x}(t_0) = \mathbf{x}_0$ *is a feasible initial condition such that there exists a local solution to Problem 1 if* $\mathbf{x}_0 \in S_1$ *with:*

$$S_1 = \left\{(x_1, x_2) \in \mathbb{R}^*_- \times \{-x_1 - 1\}\right\} \cup \{(x_1, x_2) \in \mathbb{R}_+ \times \{x_1 + 1\}\} \tag{41}$$

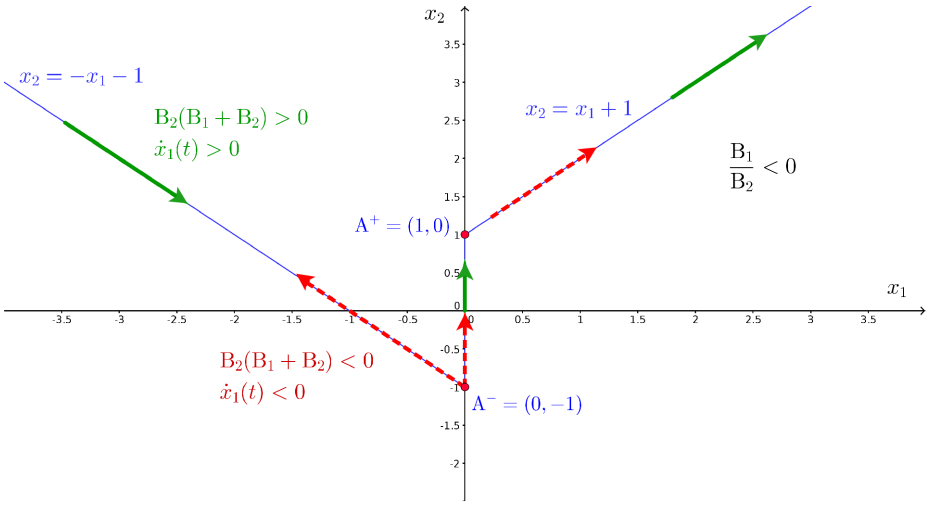
- *Assume* $\mathrm{B}_2/\mathrm{B}_1 > 0$*,* $\mathrm{B}_1 \neq \mathrm{B}_2$*, and* $\mathrm{B}_1 \neq 0$*. Then,* $\mathbf{x}(t_0) = \mathbf{x}_0$ *is a feasible initial condition such that there exists a local solution to Problem 1 if* $\mathbf{x}_0 \in S_2$ *with:*

$$S_2 = \{(x_1, x_2) \in \mathbb{R}^2 \backslash \{(0, -1)\} \mid x_2 \in |x_1| + \mathrm{sign}(x_1)\}, \tag{42}$$

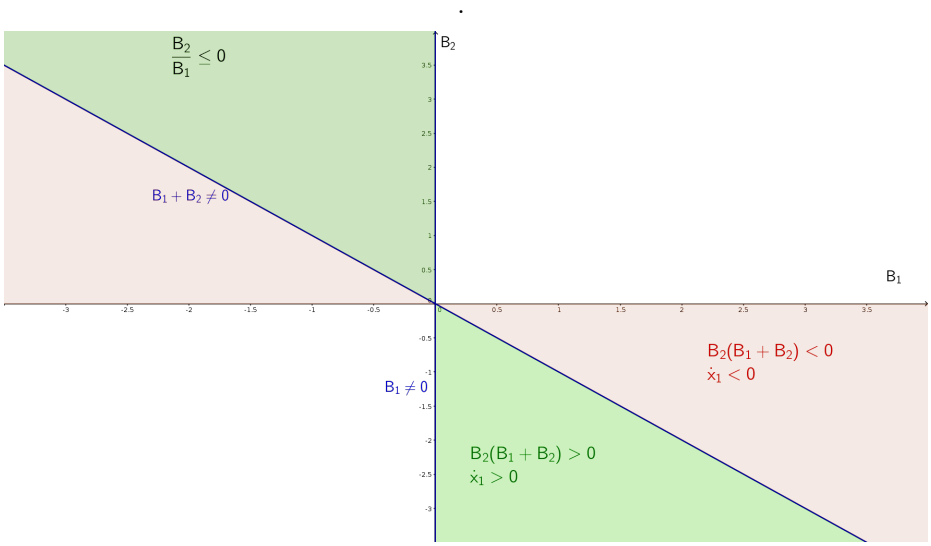
Proof. Let us notice that the first case corresponds to a singularity of the DAE in mode 3. It follows that initial conditions outside, or at the border, of this mode are necessary for the existence of local solutions to Problem 1. Then, noticing that $\dot{x}_1(t) \geq 0$ whatever the choice of B_2, we remove the initial condition $\mathbf{x}(t_0) = (0, -1)$ which leads to the definition of S_1. In the second case $\mathrm{B}_2/\mathrm{B}_1 > 0$, $\mathrm{B}_1 \neq \mathrm{B}_2$, and $\mathrm{B}_1 \neq 0$, we deduce from the proof of Proposition 1 that any point on the constraint set, with the exception of $A^- = (0, -1)$, is a valid feasible initial condition for existence of local solution to Problem 1. From this, we can define the set S_2, and the proof is complete. □

2.2 Comparison with existing concepts of solutions.

Let us first notice that the concept of solutions to discontinuous DAEs proposed in [15], and which we recalled in Section 1.2, yields no sliding



(a) In red (dashed line) the "sliding-repulsive" solutions, and in green (full line) the "sliding-crossing"solutions as defined in Remark 3



(b) As defined in Remark 3, the red cones ($\mathrm{B}_2(\mathrm{B}_1 + \mathrm{B}_2) < 0$) are the sets of parameterization $(\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$ leading to "sliding-repulsive" solutions. For example, $\mathrm{B} = (-1, 0.5)$ is such parameterization. The green cones ($\mathrm{B}_2(\mathrm{B}_1 + \mathrm{B}_2) > 0$) are the sets of parameterization $(\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$ leading to "sliding-crossing" solutions. For example, $\mathrm{B} = (-0.5, 1)$ is such parameterization.

Fig. 2: The conditions for existence of AC solutions given in Proposition 3 are given by the union of the the red and green cones.

solution to our problem as there is no continuous $z(\cdot)$ solution to (6) for our particular system (9) in case of switching. Now, let us compare the sliding mode solutions from mode 3 with the Filippov solutions [11] of the discontinuous ODE obtained by index reduction of the DAE in each mode: (9b) in mode 1 ($x_1 < 0$) and (9c) in mode 2 ($x_1 > 0$). The construction of such reduced solutions for discontinuous DAEs is close to the one defined in [17]. After the index reduction, the "left-hand" vector field $\mathbf{f}_1(\cdot)$ (for $x_1 < 0$) and "right-hand" vector field $\mathbf{f}_2(\cdot)$ (for $x_1 > 0$) are given by:

$$\mathbf{f}_1(t, \mathbf{x}(t)) = \begin{pmatrix} \frac{\mathrm{B}_2}{\mathrm{B}_1+\mathrm{B}_2} \\ \frac{-\mathrm{B}_2}{\mathrm{B}_1+\mathrm{B}_2} \end{pmatrix} \qquad (43) \qquad \mathbf{f}_2(t, \mathbf{x}(t)) = \begin{pmatrix} \frac{-\mathrm{B}_2}{\mathrm{B}_1-\mathrm{B}_2} \\ \frac{-\mathrm{B}_2}{\mathrm{B}_1-\mathrm{B}_2} \end{pmatrix} \qquad (44)$$

Let us observe from the analysis in cases 1 and 2 in the proof of Proposition 1 that $\dot{z}(t) = 0$ in mode 2 and mode 1. The variable $z(t)$ changes its value only after switching.

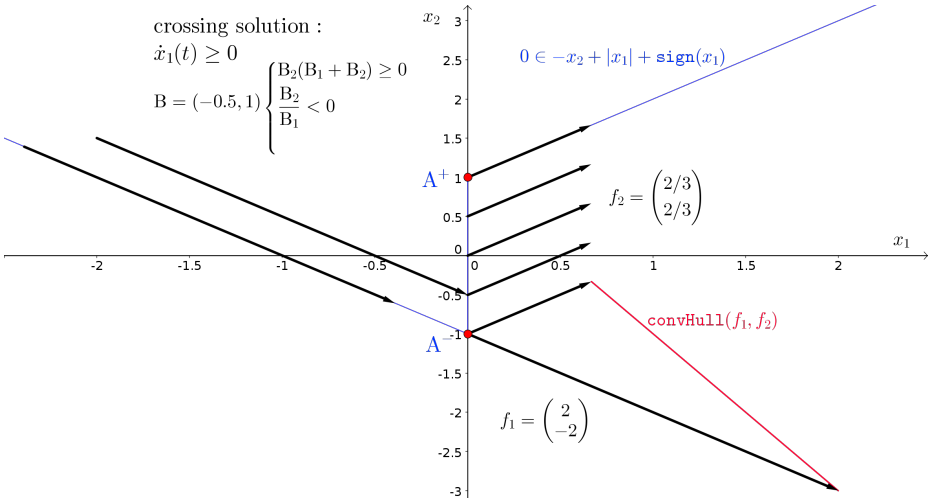
In the "sliding-repulsive" solutions case (see Remark 3), the switching ODE resulting from the index reduction yields the same sliding mode (using Filippov concept of solutions) than the one we obtain in the case 3 of Section 2.1. Indeed, in every point of the switching surface $x_1 = 0$, there exists a sliding solution associated with the vector field $f_0(\mathbf{x}) = \overline{\mathrm{co}}(\{\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x})\}) \cap \{\mathbf{x} \in \mathbb{R}^2 | x_1 = 0\}$. It follows that the solutions obtained by [17] correspond to the same solutions we obtain by our relaxation of the switching constraint with a generalised equation. A particular example with B=(-1,0.5) is shown in Figure 3b.

In the particular case of "sliding-crossing" solutions (see Remark 3), the index reduced system does not lead to any sliding motion as $\overline{\mathrm{co}}(\{\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x})\})$ does not intersect the switching surface. The solutions do not stay on the surface $x_1 = 0$, and due to the index reduction, the constraint in $x_1 > 0$ is not satisfied anymore if $x_1(t_0) < 0$. In particular, we need an explicit transition function (re-initialization rule) for the continuation of the solution as in [17] following the guidelines for sliding mode detection. The sliding solution, which we obtain using our relaxation, is not retrieved by the concept of solutions from [17]. A particular example with $\mathrm{B} = (-0.5, 1)$ is shown in Figure 3a.

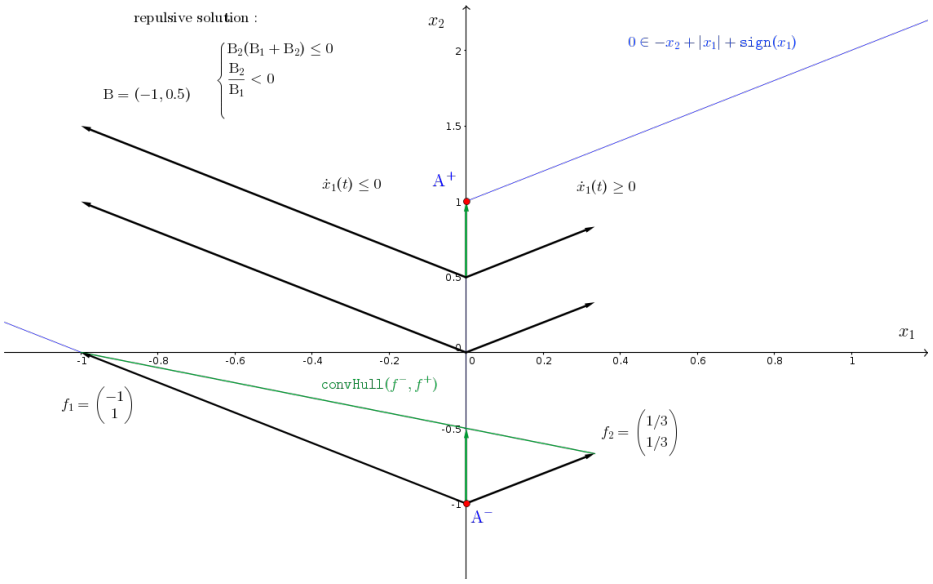
Let us note that the approach of convexifying the "left-hand" and "right-hand" reduced DAEs has already been shown problematic for some cases in [16].

2.3 Analysis of solutions with jumps

Let us study the existence of discontinuous solutions when no continuation with an AC solution exists after some time t_j. For example, this is the case if the trajectory reaches the point A^- in Figure 2a with $\mathrm{B}_2/\mathrm{B}_1 > 0$. Solutions with discontinuities make sense in a context where quite different time-scales exist in the same system. In particular, this may be found in mechanical systems with impacts [8], or in circuits with ideal diodes or set-valued electronic components [1].



(a) Example of solution with $\mathrm{B} = (-0.5, 1)$ which corresponds to solutions crossing the switching surface. The vector field of the reduced system is in black, and its convex hull is given in red.



(b) Example of solution with $\mathrm{B} = (-1, 0.5)$ which corresponds to repulsive solutions along the switching surface. The vector field of the reduced system is in black, and its convex hull is given in green.

Fig. 3: Both figures are examples of "Filippov convexification" applied to the "reduced" DAE. The top figure represents the case of crossing-solutions and the bottom figure represents the case of repulsive-solutions (see Remark 3).

2.3.1 Analysis of jump dynamics

The analysis of the dynamics with state jumps requires specific mathematical tools to describe the continuous time dynamics. First, the state $\mathbf{x}$ will be assumed to be of bounded variations (BV), see Definition 9. This functional setting is often well justified from a modelling point of view. The time derivative of a function of bounded variations is not defined in the usual sense at the discontinuity point. However, BV functions have a countable set of jump points. To give a correct meaning to the derivative, we resort to differential measures [2,18], that is a specification of the derivative in the distributional sense of a BV function. Let us denote $\mathrm{d}\mathbf{x}$ the differential measure associated with the BV function $\mathbf{x}(t)$. Using the Lebesgue decomposition of measures on the real line, a differential measure can be decomposed as

$$\mathrm{d}\mathbf{x} = \dot{\mathbf{x}}(t)\mathrm{d}t + \sum_i (\mathbf{x}(t_i^+) - \mathbf{x}(t_i^-))\delta_{t_i} + \mathrm{d}\mathbf{x}_c \tag{45}$$

where $\mathrm{d}t$ is the usual Lebesgue measure, $\dot{\mathbf{x}}(t)$ is the usual time derivative of $\mathbf{x}$ existing $\mathrm{d}t$-almost-everywhere, δ_{t_i} the Dirac measure supported at $t = t_i$, and $\mathrm{d}\mathbf{x}_c$ is the measure such that $\dot{\mathbf{x}}(t)\mathrm{d}t + \mathrm{d}\mathbf{x}_c$ is the absolutely continuous part of the decomposition of $\mathrm{d}\mathbf{x}$.

We first introduce the problem in terms of measure differential inclusions (MDI) (46) associated with (10) (see [18] for theoretical aspects of MDIs).

Problem 2. *Let $I = [t_0, T]$ be a time interval such that $T > t_0$. The problem is to find a BV function $\mathbf{x}(t) = (x_1(t), x_2(t))$ on I, and a differential measure $d\Lambda_z$ such that*

$$\begin{cases} dx_1 = dt + B_1\, d\Lambda_z \\ dx_2 = B_2\, d\Lambda_z \end{cases} \tag{46a}$$

$$0 \in -x_2(t) + |x_1(t)| + \mathrm{sign}(x_1(t))\,, \tag{46b}$$

is satisfied, in the sense of measures, given the initial condition $\mathbf{x}(t_0^-) = \mathbf{x}_0$. In addition, we wish to construct solutions $\mathbf{x}(t)$ that jump at some time t_j if and only if there is no continuous solution in $\mathbf{x}(t_j^-)$: i.e., there is no solution to Problem 1 in $\mathbf{x}(t_j^-)$. From Proposition 2 this is can be expressed by the additional condition:

$$\nexists\, z(t_j^+) \in \mathbb{R}\,, \quad \text{s.t. } (1,0)^{\mathrm{T}} + \mathrm{B}z(t_j^+) \in \mathcal{T}_{\mathcal{C}}(\mathbf{x}(t_j^-))\,. \tag{47}$$

The continuous and the discrete parts of the measure $\mathrm{d}\Lambda_z$ are identical to those of $\mathrm{d}x$. We can also write a Lebesgue decomposition

$$\mathrm{d}\Lambda_z = z(t)\mathrm{d}t + \sum_i \sigma_{z,i}\delta_{t_i} + \mathrm{d}z_c\,, \tag{48}$$

with $\sigma_{z,i}$ the jump amplitude at time t_i. For the sake of readability, we simply denote $\sigma_{z,i} = \sigma_z$. From the decomposition, we get a smooth dynamics that is equivalent to (10) $\mathrm{d}t$-almost everywhere, and an algebraic problem that partly defines the jump at impact time t_j, that is defined by the following problem:

Problem 3 (Jump algebraic problem)**.** *Let us consider the dynamics from* (46)*. Let us assume that the left limit* $\mathbf{x}(t_j^-)$ *is given and satisfies the constraints:*

$$0 \in \ \lambda(t_j^-) + |x_1(t_j^-)| - x_2(t_j^-), \quad \lambda(t_j^-) \in \ \text{sign}(x_1(t_j^-)). \tag{49}$$

The jump algebraic problem is to solve the generalized equation with unknown $\mathbf{x}(t_j^+)$*,* $\lambda(t_j^+)$*, and* σ_z*:*

$$\begin{cases} x_1(t_j^+) - x_1(t_j^-) = B_1\sigma_z \\ x_2(t_j^+) - x_2(t_j^-) = B_2\sigma_z \\ 0 \in \lambda(t_j^+) + |x_1(t_j^+)| - x_2(t_j^+) \\ \lambda(t_j^+) \in \ \text{sign}(x_1(t_j^+)). \end{cases} \tag{50}$$

It is noteworthy that writing a set-valued constraint with right-limits is a quite natural thing to do if we assume that the solutions are right-continuous at jumps[5], and it also allows us to study continuation after the jumps (see Section 2.3.3).

Multiplying the first and second lines of (50) by B_2 and B_1, respectively, one can eliminate σ_z in (50) which can be rewritten as:

$$\begin{cases} \mathrm{B}_2\left(x_1(t_j^+) - x_1(t_j^-)\right) = \mathrm{B}_1\left(x_2(t_j^+) - x_2(t_j^-)\right) \\ 0 \in \lambda(t_j^+) + |x_1(t_j^+)| - x_2(t_j^+)\,, \\ \lambda(t_j^+) \in \ \text{sign}(x_1(t_j^+))\,. \end{cases} \tag{51}$$

Note that for now we do not enforce conditions for existence of a continuous solution at t_j^+ immediately after the jump: we only define jump solutions respecting both formulation with the measures, and the index-2 constraint at t_j^+.

2.3.2 Solvability of the jump algebraic problem

Let us give a first result on the solvability of Problem 3.

Proposition 6 (Solvability of Problem 3). *Let* t_j *be a any instant and let* $\mathbf{x}(t_j^-)$ *be a given value that satisfies the condition* (49) *at* t_j^-*. The solutions* $(\mathbf{x}(t_j^+), \lambda(t_j^+), \sigma_z)$ *of Problem 3 can be characterised as follows:*

1. *If* $\mathrm{B}_1 \neq 0$*:*
   (a) *if* $\mathrm{B}_2/\mathrm{B}_1 < -1$ *there is a unique solution,*
   (b) *if* $\mathrm{B}_2/\mathrm{B}_1 \geq -1$ *there are either one or several solutions depending on* $\mathbf{x}(t_j^-)$*.*
2. *If* $\mathrm{B}_1 = 0$
   (a) *if* $x_1(t_j^-) = 0$*, there are infinitely many solutions,*
   (b) *if* $x_1(t_j^-) \neq 0$*, there is only one solution.*

[5] A property which is satisfied in jumping systems with solutions of bounded variations [18].

See Appendix A.2 for the proof.

Remark 4. Let us remark that if there is a unique solution then this solution satisfies $\mathbf{x}(t_j^+) = \mathbf{x}(t_j^-)$ and $\sigma_z = 0$ (cf. the proof of Proposition 6 in Appendix A.2). However, such result hides various behaviours in terms of solution to Problem 2. For example, if $\mathrm{B}_2/\mathrm{B}_1 \in (-1, 1)$ and $\mathbf{x}(t_j^-) = (0, -1)^{\mathrm{T}}$ then there is a unique solution solution to the algebraic jump Problem 3 and $\mathbf{x}(t_j^+) = \mathbf{x}(t_j^-)$ (cf. case (b) of the proof of Proposition 6 in Appendix A.2). This can be further separated into two sub-cases. If $\mathrm{B}_2/\mathrm{B}_1 \in (-1, 0]$, then from Proposition 1 there exists an AC solution in this point $\mathbf{x}(t_j) = (0, -1)^{\mathrm{T}}$, and the solution $\sigma_z = 0$ to the jump algebraic problem (50) is not a hinder to obtain a solution to the Problem 2. However, if $\mathrm{B}_2/\mathrm{B}_1 \in (0, 1)$, then there no AC solution in $\mathbf{x} = (0, -1)$ to Problem 1 as it is shown in Proposition 5: in this case the unique solution $\sigma_z = 0$ to the jump dynamic implies there is no solution to the Problem 2, even though in Problem 2 we consider an extension to solutions $\mathbf{x}(t)$ of bounded variations.

In the next Section 2.3.3, we search to provide an additional condition on the jump dynamics (50) to ensure the existence of an AC solution after a jump. In addition, it will restrain Problem 3 solutions, in particular when $\sigma_z = 0$, to the ones allowing the existence of a solution to Problem 2.

2.3.3 Analysis of consistent jumps

Let us now study what are the possible jumps such that an AC solution exists in the right-neighbourhood of t_j, after the jump has occurred. They are called *consistent jumps*. Let us define them more formally.

Definition 4 (Consistent jumps). *A consistent jump at time t_j^+ is defined as a solution $(\mathbf{x}(t_j^+), \lambda(t_j^+), \sigma_z)$ of Problem 3 such that a local solution of Problem 1 exists for the initial condition $\mathbf{x}(t_j^+)$.*

In Proposition 5, we expose the set of feasible initial conditions where there there exists a solution to Problem 1. Then, we give the set of consistent jumps as defined in Definition 4 as the set of reachable point in Proposition 5 given the jump dynamics from Problem 3.

Proposition 7. *Let us consider $\mathbf{x}(t_j^-)$ that satisfies (49). The set of points $\mathbf{x}(t_j^-)$ such that there is no continuation to the right with an AC solution, is equal to $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$ with:*

$$\mathcal{P}_1 = \left\{ \mathbf{x} \in \mathbb{R}^2 \mid x_1 = 0,\, x_2 \in [-1, 1) \right\}, \quad \textit{if } \mathrm{B}_1 = 0\,, \tag{52}$$

and

$$\mathcal{P}_2 = \left\{ \mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} = (0, -1)^{\mathrm{T}} = A^- \right\}, \quad \textit{if } \frac{\mathrm{B}_2}{\mathrm{B}_1} > 0,\, \mathrm{B}_1 \neq 0\,. \tag{53}$$

Then, it follows that:

a). If $\mathbf{x}(t_j^-) \in \mathcal{P}_1$ and $\mathrm{B}_1 = 0$, then $\mathbf{x}(t_j^+) = A^+$ is the only consistent jump.

b). If $\mathbf{x}(t_j^-) \in \mathcal{P}_2$, $\mathrm{B}_2/\mathrm{B}_1 > 0$, $\mathrm{B}_1 \neq 0$, *and* $\mathrm{B}_2 - \mathrm{B}_1 \leq 0$, *then there is no consistent jump.*

c). If $\mathbf{x}(t_j^-) \in \mathcal{P}_2$, $\mathrm{B}_2/\mathrm{B}_1 > 0$, $\mathrm{B}_1 \neq 0$, *and* $\mathrm{B}_2 - \mathrm{B}_1 > 0$, *then* $\mathbf{x}(t_j^+) = \left(\frac{2\mathrm{B}_1}{(\mathrm{B}_2-\mathrm{B}_1)}, \frac{(\mathrm{B}_1+\mathrm{B}_2)}{(\mathrm{B}_2-\mathrm{B}_1)}\right)^{\mathrm{T}}$ *is the only consistent jump.*

Proof. Let us first notice that the sets $\mathcal{P}_1$ and $\mathcal{P}_2$ can be easily obtained as the complementary sets in $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^2 \text{ sol. of (12')}\}$ to, respectively, S_1 and S_2 in Proposition 5.

a). Assume $\mathrm{B}_1 = 0$ and consider $\mathbf{x}(t_j^-) \in \mathcal{P}_1$. From Proposition 5, there is a consistent jump if $\mathbf{x}(t_j^+) \in S_1$ in (41). From (51), if $x_1(t_j^-) = 0$ and $\mathrm{B}_1 = 0$, it follows that $x_1(t_j^+) = 0$ and the only only solution in S_1 is $\mathbf{x}(t_j^+) = A^+$.

b) and c). Assume $\mathrm{B}_2/\mathrm{B}_1 > 0$, $\mathrm{B}_1 \neq 0$ and $\mathbf{x}(t_j^-) \in \mathcal{P}_2$. Under this assumption, Equation (51) is:

$$0 \in \ \mathrm{sign}\left(x_1(t_j^+)\right) + |x_1(t_j^+)| - \frac{\mathrm{B}_2}{\mathrm{B}_1}\left(x_1(t_j^+)\right) + 1\,. \tag{54}$$

Then, the solutions of (54), which are the possible jumps from $\mathbf{x}(t_j^-) \in \mathcal{P}_2$, are simply the intersections between the constraint $c(x_1) = |x_1| + \mathrm{sign}(x_1)$ (in blue in Figure 4) and the real line $l(x_1) = (\mathrm{B}_2/\mathrm{B}_1)x_1 - 1$ (in black in Figure 4).

b). If $(\mathrm{B}_2 - \mathrm{B}_1) \leq 0$, then there is a unique jump solution $(x_1(t_j^+), x_2(t_j^+)) = (0, -1)^{\mathrm{T}} = A^-$ which is not in S_2 in (42): there is no consistent jump.

c). If $(\mathrm{B}_2 - \mathrm{B}_1) > 0$, then there are two jump solutions: one in $(x_1(t_j^+), x_2(t_j^+)) = A^-$ that is not consistent, and a consistent one in S_2, with $(x_1(t_j^+), x_2(t_j^+))^{\mathrm{T}} = \left(\frac{2\mathrm{B}_1}{(\mathrm{B}_2-\mathrm{B}_1)}, \frac{(\mathrm{B}_1+\mathrm{B}_2)}{(\mathrm{B}_2-\mathrm{B}_1)}\right)^{\mathrm{T}}$. The proof is now complete. □

The next result is a consequence of Proposition 7.

Corollary 1. *There is no local solution to Problem 2 if and only if:*

$$\mathbf{x}(t_0^-) = A^-,\ \mathrm{B}_2/\mathrm{B}_1 > 0, \mathrm{B}_1 \neq 0, (\mathrm{B}_2 - \mathrm{B}_1) \leq 0. \tag{55}$$

Now, let us provide necessary and sufficient conditions for the existence of a consistent jump using the contingent cone condition from Proposition 2.

Proposition 8. *A necessary and sufficient condition for the solution* $(\mathbf{x}(t_j^+), \lambda(t_j^+), \sigma_z)$ *of Problem 3 to be a consistent jump is: there exists* $\gamma \in \mathbb{R}$, *such that*

$$\begin{cases} \mathbf{x}(t_j^+) - \mathbf{x}(t_j^-) = \mathrm{B}\sigma_z \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathrm{B}\gamma \in \mathcal{T}_{\mathcal{C}}(\mathbf{x}(t_j^+)) \\ 0 \in \lambda(t_j^+) + |x_1(t_j^+)| - x_2(t_j^+) \\ \lambda(t_j^+) \in \ \mathrm{sign}(x_1(t_j^+))\,, \end{cases} \tag{56}$$

with $\mathcal{T}_{\mathcal{C}}(\mathbf{x}(t_j^+))$ *as defined in Lemma 1.*

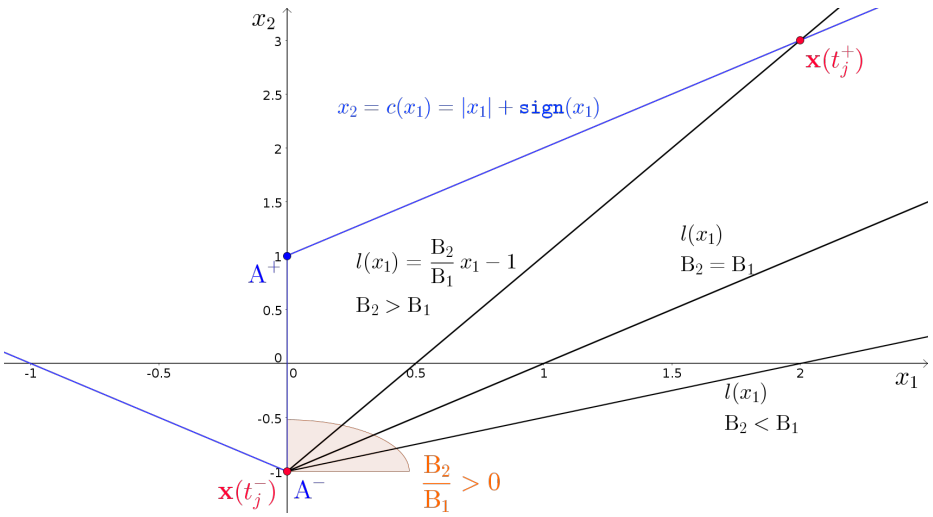


Fig. 4: Solutions of the GE (54) associated with jumps from $\mathbf{x}(t_j^-) = (0,-1)^{\mathrm{T}}$ for various choices of $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$.

Proof. Let us observe from Proposition 2, that $(\mathbf{x}(t_j^+), (1,0)^{\mathrm{T}} + \mathrm{B}\gamma) \in \mathcal{C} \times \mathcal{T}_{\mathcal{C}}(\mathbf{x}(t_j^+))$ is a necessary and sufficient condition for the existence of a local solution to Problem 1 in $\mathbf{x}(t_j^+)$, which concludes the proof. □

3 Analysis of an event-capturing backward Euler scheme

Backward (implicit) Euler schemes have proved to be efficient schemes for the simulation of nonsmooth dynamical systems [1,2]. Therefore, let us now consider the backward Euler discretization of the system (10) as follows:

$$\begin{cases} x_{1,k+1} - x_{1,k} = h(1 + \mathrm{B}_1 z_{k+1}) \\ x_{2,k+1} - x_{2,k} = h\mathrm{B}_2 z_{k+1} \\ \qquad\qquad 0 \in \ \mathrm{sign}(x_{1,k+1}) + |x_{1,k+1}| - x_{2,k+1}\,, \end{cases} \tag{57}$$

with $h > 0$ the time step. For the sake of simplicity, we choose a fixed time step, $h = \frac{T}{N}$, where $[0,T]$ is the interval of integration and N the number of fixed steps.

3.1 Well-posedness of the backward Euler discretization

One sees that (57) has a structure quite close to (46), which puts the backward scheme in a favourable perspective for the computation of solutions with jumps. This is the aim of the further analysis.

Using the same method as in the previous section, we can eliminate z_{k+1}, and we obtain the GE:

$$\begin{cases} \mathrm{B}_2(x_{1,k+1} - x_{1,k}) = h\mathrm{B}_2 + \mathrm{B}_1(x_{2,k+1} - x_{2,k}) \\ 0 \in \ \mathrm{sign}(x_{1,k+1}) + |x_{1,k+1}| - x_{2,k+1} \, . \end{cases} \tag{58}$$

Proposition 9 (Well-posedness of the Backward Euler one-step problem (58)). *Let us characterise the existence and uniqueness of solutions to the Backward Euler one-step problem* (58) *for various values of the parameterization*[6] $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$.

1. *If* $\mathrm{B}_1 \neq 0$*:*
   (a) *if* $\mathrm{B}_2/\mathrm{B}_1 < -1$*, there is a unique solution to* (58) *whatever the size of the time step* $h > 0$.
   (b) *if* $\mathrm{B}_2/\mathrm{B}_1 \in (-1, 0]$*, there are either one, or several solutions depending on* $\mathbf{x}_k$ *and* h.
   (c) *if* $\mathrm{B}_2/\mathrm{B}_1 \in (0, 1)$*, there are either none, one, or several solutions depending on* $\mathbf{x}_k$ *and* h.
   (d) *if* $\mathrm{B}_2/\mathrm{B}_1 > 1$*, there are either one, or several solutions depending on* $\mathbf{x}_k$ *and* h.
2. *If* $\mathrm{B}_1 = 0$*:*
   (a) *if* $x_{1,k} = -h$*, there are infinitely many solutions.*
   (b) *if* $x_{1,k} \neq -h$*, there is only one solution*

The complete proof can be found in Appendix A.3. From Proposition 9 and its proof, we deduce a relation in-between the lack of solution of the Euler one-step problem (58) for any $h > 0$, and the non-existence of local solution to Problem 2 (see Corollary 1). This result is given in Corollary 2.

Corollary 2. *If there is no solutions to the backward Euler one-step problem* (58) *for all* $h > 0$*, then* $\mathrm{B}_2/\mathrm{B}_1 \in (0, 1)$ *and* $\mathbf{x}_k = A^-$*, and there is no solution to Problem 2 in* $\mathbf{x}_k$.

3.2 Minimal implicit Euler discretization

As we have seen in Section 3.1, the classical implicit Euler discretization may output multiple solutions, for $h > 0$ as small as wanted. One needs to improve the implicit Euler discretization to select the discrete solution close the continuous-time solution. To this aim, we propose a minimisation over the results of the backward Euler scheme.

Definition 5. *Let us consider a non-smooth DAE system* (10)*. The minimal norm backward Euler scheme is to find at each time step* $\mathbf{x}_{k+1}$*, solution of:*

$$\begin{aligned} \underline{\mathbf{x}}_{k+1} := \mathrm{argmin}_{\mathbf{x}} \quad & \|\mathbf{x} - \mathbf{x}_k\| \, , \\ \text{s.t.} \quad & x_1 - x_{1,k} = \ h(1 + B_1 z) \\ & x_2 - x_{2,k} = \ hB_2 z \\ & 0 \in \ \mathrm{sign}(x_1) + |x_1| - x_2 \, . \end{aligned} \tag{59}$$

[6] Let us recall that we assume $\mathrm{B}_1 + \mathrm{B}_2 \neq 0$ and $\mathrm{B}_1 - \mathrm{B}_2 \neq 0$.

Although the implicit Euler scheme outputs multiple solutions, as long as one of the solutions is still on the same constraint as $\mathbf{x}_k$, we know this solution is an $O(h)$ approximation. Now, let us prove this always occurs if $\mathbf{x}_k$ is a AC consistent initial condition for our particular example.

Proposition 10 (Consistency of the Minimal Backward Euler scheme for Problem 1.). *Consider the non-smooth DAE system from* (10)*. If there exists a solution* $\mathbf{x}(t)$ *that is AC on an interval* $[t_0, t_0 + \varepsilon]$*, then there exists a time step* $h > 0$ *such that the minimal norm backward Euler scheme* (59) *provides a consistent discrete solution to* (10)*. This means that given* $\mathbf{x}_k = \mathbf{x}(t_k)$ *then* $\left\|\underline{\mathbf{x}}_{k+1} - \mathbf{x}(t_k + h)\right\| \to 0$ *when* $h \to 0$*.*

The complete proof can be found in Appendix A.4.

Remark 5 (Necessary conditions of optimality). The optimisation problem solved in Definition 5 is a mathematical program with equilibrium constraints (MPEC). Under the assumption of a mixed linear complementarity problem (MLCP) representation of the generalised equation in (59), necessary conditions of optimality depend on some constraint qualifications. We refer to [27] as a reference on this question. In the simulations presented in Section 3.3, we address this problem in a naive way by simply enumerating all the solutions and selecting the one of minimal norm, since the aim is to show the experimental convergence of such method for this problem.

3.3 Implementation and numerical results

In this section, we expose some simulation results of the minimal implicit Euler scheme (59) on the example studied in the previous sections. In particular, we show through experiments on the example associated with (10) that if there exists at least one continuous solution, then (59) converges in $O(h)$ to one of these solutions. Furthermore, if the discretization (57) yields a unique solution for any step size, then it converges in $O(h)$ to the unique solutions of (10).

Implementation has been performed using Siconos 4.2.0 [4], a platform for numerical simulation of non-smooth dynamical systems. The code of these simulations can be found in the github repository associated with the Siconos examples[7]. In this section, performance results are not discussed as the optimisation problem in (59) is currently solved by enumeration of all[8] the solutions of the generalised equation associated with the classical implicit Euler scheme (57). A way to implement this problem is to use a Linear Complementarity Problem (LCP) formulation (see Definition 19) to represent the non-smooth component of the constraint (12). Then, the constraint (12) can be expressed as an LCP with an additional equality constraint, that is, a Mixed Linear Complementarity

[7] `https://github.com/siconos/siconos-tutorials/tree/master/.sandbox/code_IFAC`
[8] In the considered cases limited to three solutions.

Problem (MLCP) as follows:

$$\begin{cases} 0 = -x_2 + |x_1| + \alpha \\ 0 \leq |x_1| + x_1 \perp |x_1| - x_1 \geq 0 \\ 0 \leq x_1^- \perp \alpha \geq -1 \\ \alpha \leq 1 \perp x_1^+ \geq 0\,, \end{cases} \tag{60}$$

with $\alpha \in \text{sign}(x_1)$. This yields a Mixed Linear Complementarity System (MLCS)[9]:

$$\begin{cases} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{z}(t) \end{pmatrix} = \begin{pmatrix} 1 + \mathrm{B}_1 z(t) \\ \mathrm{B}_2 z(t) \\ x_1 - x_2 + 1 + \lambda_1(t) - \lambda_2(t) \end{pmatrix} \\ 0 \leq 2x_1(t) + \lambda_1(t) \perp \lambda_1(t) \geq 0 \\ 0 \leq \lambda_3(t) + x_1(t) \perp \lambda_2(t) \geq 0 \\ 0 \leq 2 - \lambda_2(t) \perp \lambda_3(t) \geq 0 \end{cases} \tag{61}$$

with $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3) = \left(|x_1| - x_1, 1 - \alpha, x_1^-\right)$.

Some numerical results can be found in Figure 5 and Figure 6. In these figures, we consider the particular case of sliding-crossing solutions (here $\mathrm{B} = (-0.5, 1)^{\mathrm{T}}$) where uniqueness of AC solutions and discrete solutions is guaranteed. We notice that the resulting solutions in $\mathbf{x}(t)$ are Lipschitz continuous, and run through all the modes (the initial condition is taken with $x_1(t_0) < 0$). In Figure 7, the error term $\|\mathbf{x}(T) - \underline{\mathbf{x}}_N\|$ is depicted as a function of the step size h. The term $\underline{\mathbf{x}}_N$ is the numerical approximation of $\mathbf{x}(T)$ by the minimal implicit Euler numerical scheme (59) when the interval $[t_0, T]$ is subdivided in N steps of size h. In the case of sliding-crossing solutions, we choose $\mathrm{B} = (-0.5, 1)^{\mathrm{T}}$, $\mathbf{x}_0 = \mathbf{x}(t_0) = (-5, 4)$, and $T = 10$. In the case of sliding-repulsive solutions, we choose $\mathrm{B} = (-1, 0.5)^{\mathrm{T}}$, $\mathbf{x}_0 = \mathbf{x}(t_0) = (0, 0)$, and $T = 10$. Time step sizes are taken equally spaced in log-scale. We observe the first order convergence of the implicit Euler scheme when there is uniqueness of the numerical solutions. In addition, we also observe a first order convergence of the minimal implicit Euler scheme when there is non-uniqueness of the discrete solution (for any time step) as it is shown in Fig 7 on the curve associated with the sliding-repulsive case. Finally, we also test a variant of the studied example where the dynamic $\dot{\mathbf{x}}(t) = \mathrm{B}z(t) + b$ is replaced by $\dot{\mathbf{x}}(t) = \mathrm{A}\mathbf{x}(t) + \mathrm{B}z(t) + b$, with

$$\mathrm{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} -1 \\ 0.5 \end{pmatrix}. \tag{62}$$

Some results are exposed in Figures 8, 9, and convergence on some experiments is also in $O(h)$ as it can be seen in Figure 7.

[9] Please note that this LCP formulation is not unique as it depends of the naming convention for the λ_i variables.

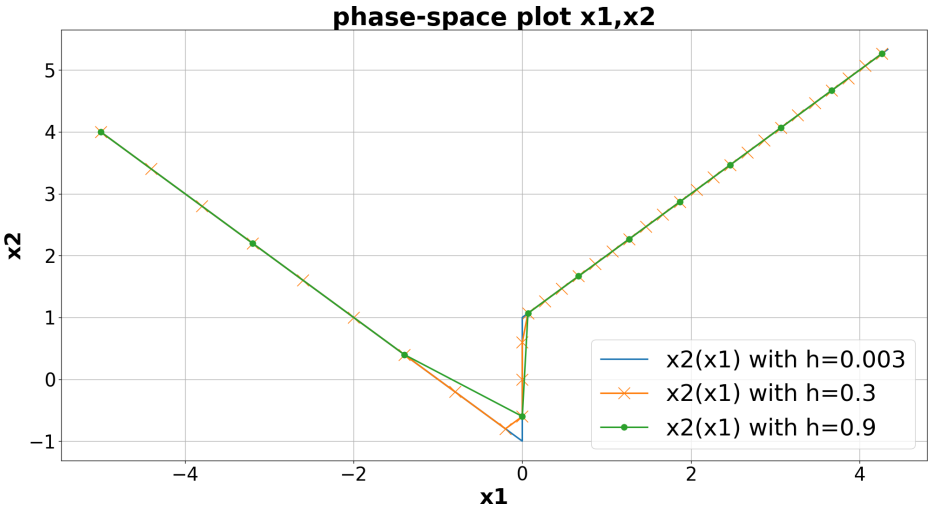


Fig. 5: Phase space plot in (x_1, x_2) of the numerical solutions for the sliding crossing case, $\mathrm{B} = (-0.5, 1)^{\mathrm{T}}$, and $h = 0.9$, $h = 0.3$, or $h = 0.003$. Initial condition is $\mathbf{x}_0 = (-5, 4)^{\mathrm{T}}$.

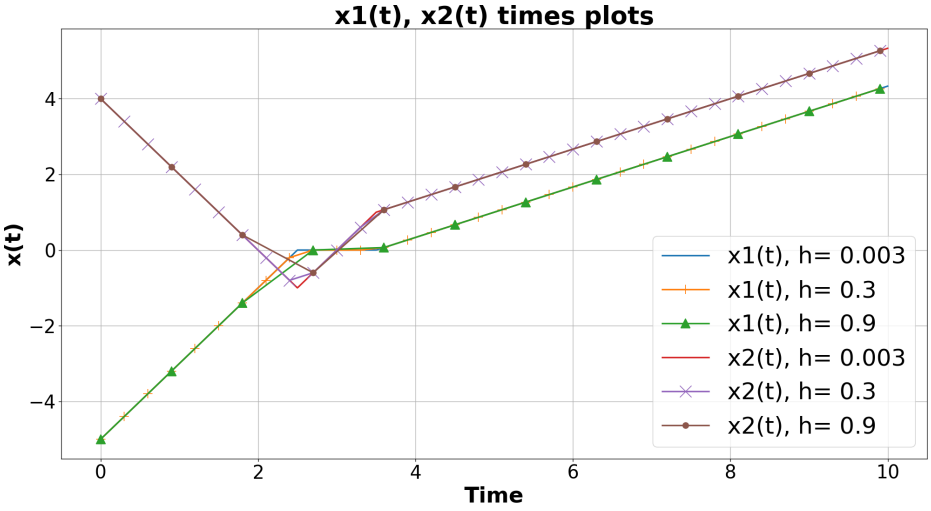


Fig. 6: Time plots of the sliding-crossing solutions $x_1(t)$ and $x_2(t)$ for $\mathrm{B} = (-0.5, 1)^{\mathrm{T}}$ and $h = 0.9$, $h = 0.3$, or $h = 0.003$. Initial condition is $\mathbf{x}_0 = (-5, 4)$.

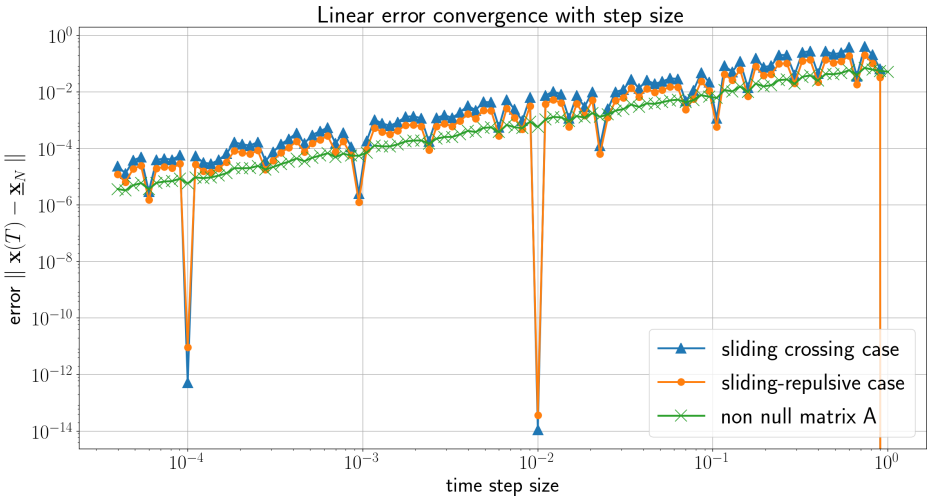


Fig. 7: Error $\|\mathbf{x}(t) - \underline{\mathbf{x}}_N\|$ with respect to time step h. We consider the two kind of AC solutions: the sliding-crossing solutions and the sliding-repulsive solutions.

4 Relaxation of switching linear equality constraints

In this section, we analyse three relaxation strategies applied to switching equality constraints in order to generate a new constraint whose associated solution set is path-connected (see Definition 18 in Appendix A.1). In Section 4.1, the possibilities of a simple continuous bond of both sides of the switching constraint is formalised by the switching constraint connected (SCC) relaxation; in Section 4.2, we study the so-called multivalued step operator (MSO) relaxation (which is (4) with set-valued step functions); in Section 4.3, a closed convex hull (CCH) relaxation (which can be seen as an extension of Filippov's regularisation of switching ODEs to switching constraints) is analyzed. In particular, in every section we discuss a few examples stemming from the piecewise-linear DAE (see Definition 1), and the generalisation of the example in Section 2, that is, a piecewise-linear DAE with switching equality constraints, defined by switching hyperplanes.

Referring to Definition 2, we restrict ourselves to piecewise-linear DAEs with a single switching. A particular example of such piecewise DAEs was studied in Section 2. In this section, we mainly discuss properties of various relaxations of

$$\begin{cases} 0 = g_1(\mathbf{x}) = \mathrm{C}_1\mathbf{x} + \mathrm{p}_1\,, & \text{if } h(\mathbf{x}) = \mathrm{H}\mathbf{x} + \mathrm{q} < 0 \qquad (63a) \\ 0 = g_2(\mathbf{x}) = \mathrm{C}_2\mathbf{x} + \mathrm{p}_2\,, & \text{if } h(\mathbf{x}) > 0\,, \qquad (63b) \end{cases}$$

where $\mathbf{x} \in \mathbb{R}^{n_1}$, $\mathrm{C}_{i=1,2}^{\mathrm{T}}$ and $\mathrm{H}^{\mathrm{T}} \in \mathbb{R}^{n_1}$, $\mathrm{p}_{i=1,2}$ and $\mathrm{q} \in \mathbb{R}$. Let us define N_{g_1}, N_{g_2}, and N_h, the solution sets associated with the constraints (63a), (63b), and the switching surface $h(\mathbf{x}) = 0$:

$$\mathrm{N}_{g_1} = \{\mathbf{x} \in \mathbb{R}^{n_1} \mid g_1(\mathbf{x}) = \mathrm{C}_1\mathbf{x} + \mathrm{p}_1 = 0\}\,, \qquad (64)$$

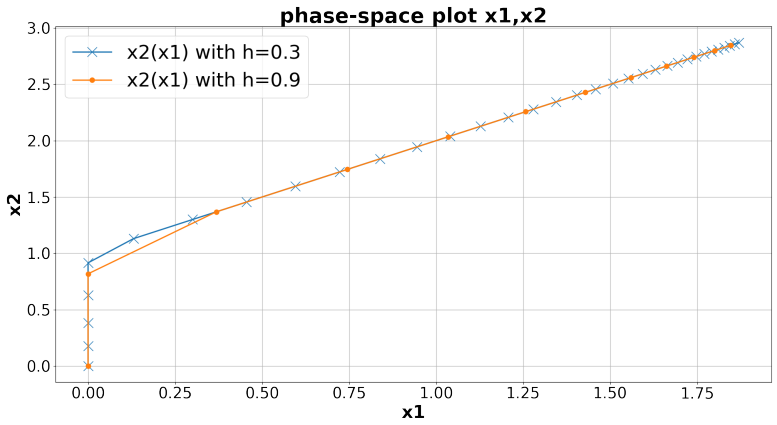


Fig. 8: Phase space plot in (x_1, x_2) of the numerical solutions for $\mathrm{B} = (-1, 0.5)$, A given in (62), and $h = 0.9$ or $h = 0.3$. Initial condition is $\mathbf{x}_0 = (0, 0)$.

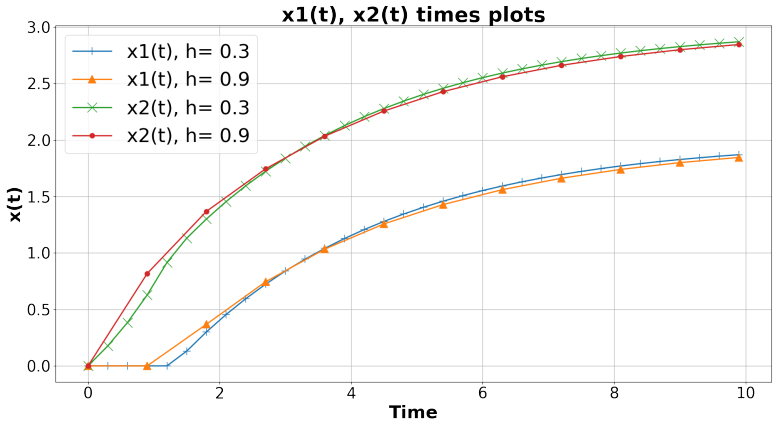


Fig. 9: Time plots of the solutions $x_1(t)$ and $x_2(t)$ for $\mathrm{B} = (-1, 0.5)$, A given in (62), and $h = 0.9$ or $h = 0.3$. Initial condition is $\mathbf{x}_0 = (0, 0)$.

$$\mathrm{N}_{g_2} = \{\mathbf{x} \in \mathbb{R}^{n_1} \mid g_2(\mathbf{x}) = \mathrm{C}_2\mathbf{x} + \mathrm{p}_2 = 0\}\,, \tag{65}$$

$$\mathrm{N}_h = \{\mathbf{x} \in \mathbb{R}^{n_1} \mid h(\mathbf{x}) = \mathrm{H}\mathbf{x} + \mathrm{q} = 0\}\,. \tag{66}$$

Let us remark that N_{g_1}, N_{g_2}, and N_h are the zero spaces associated with the functions $g_1(\cdot)$, $g_2(\cdot)$, and $h(\cdot)$, respectively.

We wish to study relaxation strategies in order to transform the switching surface into a sliding mode, and to enable the existence of paths in-between the "left-hand" constraint (63a) and the "right-hand" constraint (63b). To this aim, let us assume that both (64) and (65) intersect, not necessary at the same points, the switching surface (66):

$$\mathrm{N}_{g_1} \cap \mathrm{N}_h \neq \emptyset \text{ and } \mathrm{N}_{g_2} \cap \mathrm{N}_h \neq \emptyset\,. \tag{67}$$

Finally, let us define the solution set of the switching constraint (63), $\mathcal{G}$ by

$$\mathcal{G} = \{\mathbf{x} \in \mathbb{R}^{n_1} \text{ that solves (63)}\}. \tag{68}$$

4.1 Switching Constraint Connected (SCC) relaxation

The objective of the first relaxation is to define the extended constraint of the form :

$$\begin{cases} 0 = \mathrm{C}_1\mathbf{x} + \mathrm{p}_1\,, & \text{if } \mathrm{H}\mathbf{x} + \mathrm{q} \leq 0 \\ 0 = \mathrm{C}_2\mathbf{x} + \mathrm{p}_2\,, & \text{if } \mathrm{H}\mathbf{x} + \mathrm{q} \geq 0\,, \end{cases} \tag{69}$$

We do not enter in the details of the state-dependent switching DAE as in Definition 1, and only consider its switching constraint. However, let us remark that the path connectivity of the constraint solution set is a necessary condition (assuming (67)) for the existence of continuous solutions, in the sense of Proposition 3, to state-dependent switching DAE with constraints such as in (63).

Proposition 11 (Switching Constraint Connected (SCC) relaxation). *Let $\mathcal{S}$ be the set of solutions to the linear system:*

$$\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^{n_1} \mid 0 = \mathrm{C}_1\mathbf{x} + \mathrm{p}_1; 0 = \mathrm{C}_2\mathbf{x} + \mathrm{p}_2; 0 = \mathrm{H}\mathbf{x} + \mathrm{q}\}\,. \tag{70}$$

If $\mathcal{S}$ is non-empty, we define the Switching Constraint Connected (SCC) relaxation by the union of the switching constraint solution set $\mathcal{G}$ with $\mathcal{S}$, such that

$$\mathcal{H}_e = \mathcal{G} \cup \mathcal{S} = \{\mathbf{x} \in \mathbb{R}^{n_1} \textit{ that solves } (69)\} \tag{71}$$

Then, $\mathcal{H}_e$ is path-connected.

Proof. Let $\mathbf{x}_s$ be any point in $\mathcal{S}$. Let α be any point satisfying (63a), meaning $0 = g_1(\alpha)$ and $h(\alpha) < 0$. Similarly, let β be any point satisfying (63b), meaning $0 = g_2(\beta)$ and $h(\beta) > 0$. Then, there is a path from α to $\mathbf{x}_s$ since N_{g_1} is a hyperplane and $\mathcal{S}$ is not empty. The same is true from $\mathbf{x}_s$ to β. Finally, by transitivity we can build a path from α to β, which proves that the solution set of (69) is path-connected. □

We will now provide some illustrating examples of switching constraints.

Example 1.a. First, let us consider the trivial switching constraint:

$$\begin{cases} 0 = x_2\,, & \text{if } x_1 < 0 \\ 0 = x_2\,, & \text{if } x_1 > 0\,. \end{cases} \quad \Longleftrightarrow \quad \mathbf{x} \in \mathcal{G}\,. \tag{72}$$

In this example, $\mathrm{C}_1 = (0,1)$, $\mathrm{C}_2 = (0,1)$, $\mathrm{H} = (1,0)$, and $(\mathrm{p}_1, \mathrm{p}_2, \mathrm{q}) = (0,0,0)$. The set $\mathcal{G}$ can be obviously extended in $x_1 = 0$ by $x_2 = 0$ as given by the solution of (70) applied to this example to obtain

$$\begin{aligned} \mathcal{H}_e = & \left\{ (x_1, x_2)^{\mathrm{T}} \in \mathbb{R}^2,\ \begin{cases} 0 = x_2\,, & \text{if } x_1 \leq 0 \\ 0 = x_2\,, & \text{if } x_1 \geq 0 \end{cases} \right\} \\ & = \mathbf{x}_s = (0,0)^{\mathrm{T}}\,. \end{aligned} \tag{73}$$

This is depicted in Figure 10. Although this example seems trivial, as both constraints are the same on either side of the discontinuity, some unexpected results arise when applying the formula from (4), as we will see Section 4.2. ■

Example 2.a. Let us now consider the switching constraint associated with the example (9) from Section 2:

$$\begin{cases} 0 = -1 - x_1 - x_2\,, & \text{if } x_1 < 0 \\ 0 = 1 + x_1 - x_2\,, & \text{if } x_1 > 0\,. \end{cases} \tag{74}$$

In this example, $\mathrm{C}_1 = (-1,-1)$, $\mathrm{C}_2 = (1,-1)$, $\mathrm{H} = (1\ 0)$, and $(\mathrm{p}_1, \mathrm{p}_2, \mathrm{q}) = (-1,1,0)$. The problem (70) has no solution, and the set $\mathcal{G}$ cannot be extended by a SCC relaxation in the sense of Proposition 11 to a path connected set $\mathcal{H}_e$. This is depicted in Figure 11. ■

Example 3.a. Finally, let us consider the next switching constraint for $\mathbf{x} \in \mathbb{R}^3$:

$$\begin{cases} 0 = 1 + x_1 + x_2\,, & \text{if } x_3 < 0 \\ 0 = -1 - x_1 + x_2\,, & \text{if } x_3 > 0\,. \end{cases} \tag{75}$$

In this example, $\mathrm{C}_1 = (1,1,0)$, $\mathrm{C}_2 = (-1,1,0)$, $\mathrm{H} = (0,0,1)$, and $(\mathrm{p}_1, \mathrm{p}_2, \mathrm{q}) = (1,-1,0)$. For this example, the problem (70) has a unique solution $\mathbf{x}_s = (-1,0,0)$ and the graph of the constraint can be extended to give

$$\mathcal{H}_e = \left\{ (x_1, x_2, x_3)^{\mathrm{T}} \in \mathbb{R}^3\,,\ \begin{cases} 0 = 1 + x_1 + x_2\,, & \text{if } x_3 \leq 0 \\ 0 = -1 - x_1 + x_2\,, & \text{if } x_3 \geq 0 \end{cases} \right\}. \tag{76}$$

This is depicted in Figure 12. ■

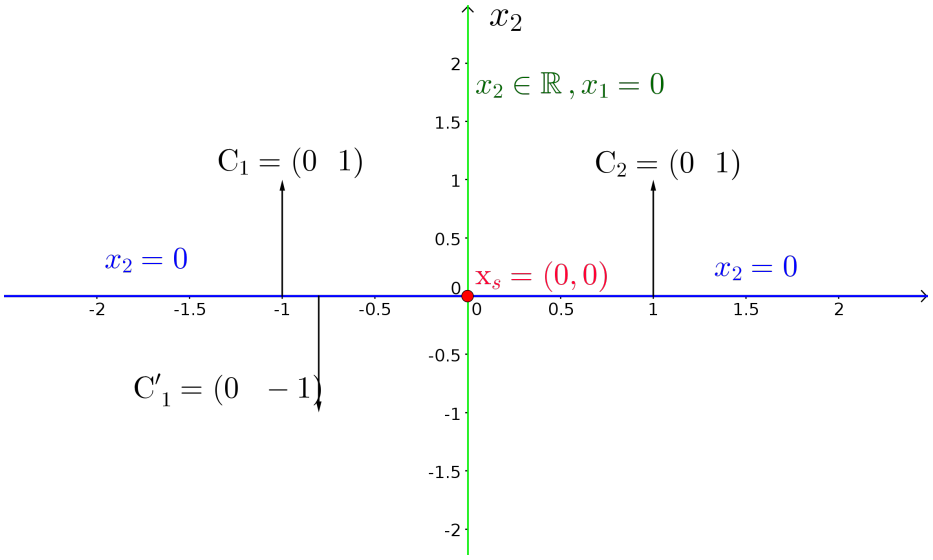


Fig. 10: This figure depicts the various relaxations applied to Example 1.a. In red, the point $\mathbf{x}_s$ is obtain by the continuous extension, the CCH relaxation as well as the MSO relaxation with the choice of normal C_1. In green, we notice the MSO relaxation with the choice of normal C'_1.

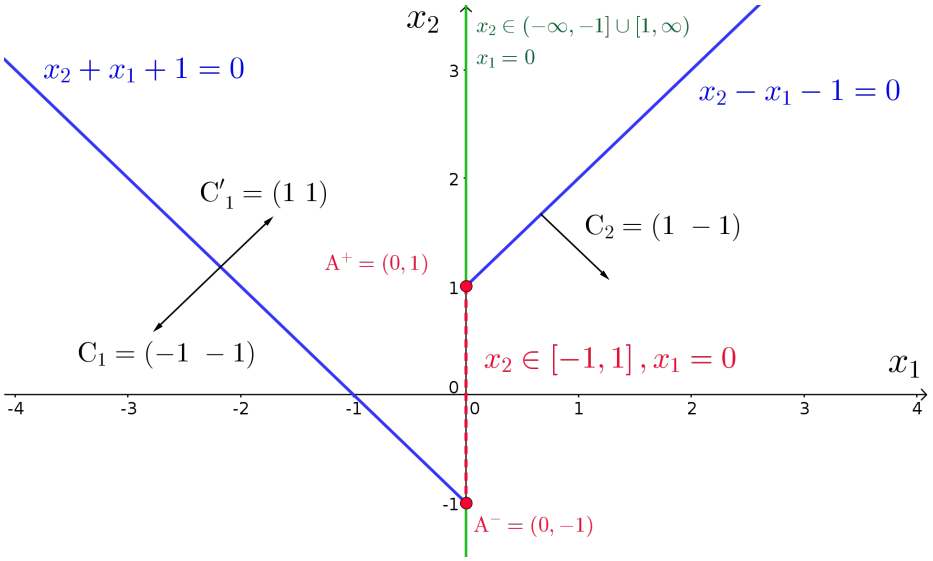


Fig. 11: This figure depicts the various relaxations applied to Example 2.a. In red, the CCH relaxation as well as the MSO relaxation with the choice of normal C_1 is represented. In green, we notice the MSO relaxation with the choice of normal C'_1.

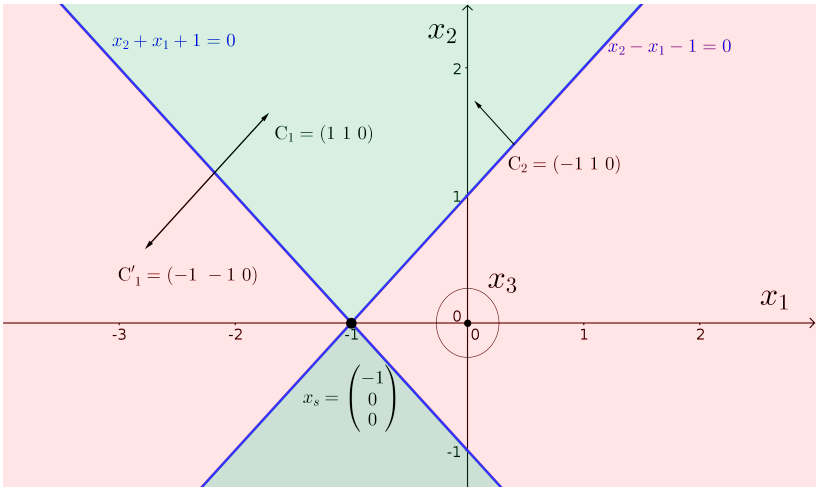


Fig. 12: This figure depicts the various relaxations applied to Example 3.a. Only the section on the switching surface $x_3 = 0$ is represented. It follows that the symbol $\odot$ indicates that the x_3 axe is oriented toward the reader. The SCC relaxation is given by the point $\mathbf{x}_s = (-1, 0, 0)^T$ in black. The red cones are the the MSO relaxation with the choice of normal vector C_1 (see Example 2.b). The green cone are the MSO relaxation with the choice of normal vector C'_1. The CCH relaxation is the whole surface $x_3 = 0$ and is not represented.

4.2 Multivalued Step Operator (MSO) relaxation

Let us now define a new kind of relaxation in the context of this section, the relaxation of switching constraints (4) where $\mathrm{s}(\cdot)$ is the set-valued step operator.

Definition 6 (Multivalued Step Operator (MSO) relaxation). *The Multivalued Step Operator (MSO) relaxation of a switching constraint* (63) *is given by the roots of the GE:*

$$0 = (1-\lambda)g_1(\mathbf{x}) + \lambda g_2(\mathbf{x})\,, \tag{77}$$

where $\lambda \in \mathrm{s}(h(\mathbf{x}))$ *and* $\mathrm{s}(\cdot)$ *is the multivalued step function:*

$$\mathrm{s}(x) = \begin{cases} \{0\} & \text{if } x < 0 \\ [0,1] & \text{if } x = 0 \\ \{1\} & \text{if } x > 0 \end{cases} \tag{78}$$

The MSO relaxation defines the set $\mathcal{H}_s$ *associated with the constraints by*

$$\mathcal{H}_s = \{\mathbf{x} \in \mathbb{R}^n_1 \mid 0 = (1-\lambda)g_1(\mathbf{x}) + \lambda g_2(\mathbf{x}),\, \lambda \in \mathrm{s}(h(\mathbf{x}))\}\,. \tag{79}$$

In Definition 6, we relax the constraint on the switching surface by the set $\mathcal{H}_s$ defined as the intersection of points satisfying (77) and the switching surface included in the zero space N_h defined by (66).

The MSO relaxation is the one intuitively built to combine two functions, and a similar approach has already been taken in the context of switching ODEs to combine two vector fields along a switching surface. In [24], the authors apply a similar relaxation to exogenous switching DAEs by using smooth step, or sign functions approximation (respectively the sigmoid or the hyperbolic tangent) instead of a multivalued operator. However, the results of such relaxation in term of dimension and path-connectivity of $\mathcal{H}_s$ are hard to predict. In particular, in Examples 1.b, 2.b and 3.b, we show that for each problem two possible relaxations can be defined by simply considering equivalent formulations[10] of the switching constraints. Furthermore, we note that in the Example 2.b, associated with the constraint of Section 2's example, an equivalent switching constraint can be considered whose relaxation in not path-connected.

Remark 6. We can also remark that solutions $\mathbf{x}_s$ of (70) are included in the ones of (77) as it can be trivially observed that if $0 = g_1(\mathbf{x})$ and $0 = g_2(\mathbf{x})$ then $0 = (1-\lambda)g_1(\mathbf{x}) + \lambda g_2(\mathbf{x})$ for all λ.

Example 1.b. Let us consider the switching constraint from Example 1.a. The GE built by the MSO relaxation is:

$$0 = (1-\lambda)x_2 + \lambda x_2\,, \lambda \in \mathrm{s}(x_1)\,. \tag{80}$$

[10] With respect to the solution set of the constraint.

This equation can trivially be reduced to $x_2 = 0$ and we retrieve the real line, as with the continuous extension from Example 1.a. In $x_1 = 0$, the set of solutions is $\{\lambda \in [0,1], x_2 = 0\}$ which is of dimension 0 in the plane (x_1, x_2). Let us now consider the equivalent switching constraint:

$$\begin{cases} 0 = -x_2\,, & \text{if } x_1 < 0 \quad (81a) \\ 0 = \ x_2 \ \ , & \text{if } x_1 > 0\,. \quad (81b) \end{cases}$$

In (81a), the equality constraint $x_2 = 0$, has been replaced by the equivalent $-x_2 = 0$. This corresponds to the change of normal vector from C_1 to $\mathrm{C}_1' = (0, -1)$ as shown in Figure 10. The GE that corresponds to the relaxation of (81) yields:

$$\mathcal{H}_s = \{\mathbf{x} \in \mathbb{R}^2, 0 \in x_2(2\lambda - 1)\,, \lambda \in \mathrm{s}(x_1)\}\,. \tag{82}$$

The set of solutions in $x_1 = 0$ is $\{\lambda = 1/2, x_2 \in \mathbb{R}\} \cup \{\lambda \in [0, 1/2], x_2 = 0\}$, which is larger than the one of (80), and is of dimension 1 in the plane (x_1, x_2). This is exposed in green in Figure 10. Although the solutions set of (82) is still path-connected, the fact that $x_2 \in \mathbb{R}$ is solution in $x_1 = 0$ may allow multiple solutions to a DAE in this point. ■

Example 2.b. Let us now consider the switching constraint from Example 2.a, which is the same as the one studied in Section 2. As seen in (10) and Figure 1, the graph of the relaxation is path connected. In fact, for this example and Example 1.a, on the switching surface, the graph of the relaxation is the closure of the convex hull of the extended left and right constraints. However, we will see in Examples 3.b and 3.c that this may not always be the case. In Figure 11, we represent the switching constraint (74) in blue with the relaxation in $x_1 = 0$ in red. Again, in a similar way to the previous example, we now consider an equivalent switching constraint where one of the equality constraints is re-written with the opposite normal $\mathrm{C}_1' = (1\ 1)$ instead of $\mathrm{C}_1 = (-1\ -1)$, that is

$$\begin{cases} 0 = x_1 + x_2 + 1\,, & \text{if } x_1 < 0 \\ 0 = x_1 - x_2 + 1\,, & \text{if } x_1 > 0\,. \end{cases} \tag{83}$$

The MSO relaxation associated with (83) is:

$$\mathcal{H}_s = \{\mathbf{x} \in \mathbb{R}^2, 0 \in x_1 + x_2 + 1 - 2\lambda x_2\,, \lambda \in \mathrm{s}(x_1)\}\,. \tag{84}$$

Let us now study the behaviour in $x_1 = 0$. Let us assume that $x_2 \geq 0$. Then, as $\lambda \in [0, 1]$ this yields the inclusion $0 \in [-x_2+1, x_2+1]$ and this implies that $x_2 \geq 1$. Similarly, let us assume that $x_2 \leq 0$, then it follows that: $0 \in [x_2 + 1, -x_2 + 1]$ and $x_2 \leq -1$ is necessary. As a conclusion, the set $\mathcal{H}_s$ is not path-connected, it is represented in green in Figure 11. This prevents the existence of AC solutions switching from one mode to the other one. ■

Example 3.b. Finally, let us consider the MSO relaxation on the switching constraint of Example 3.a. This example is a bit more complex since $\dim(\mathbf{x}) = 3$,

with a switching surface orthogonal to each mode's hyperplane. The resulting GE yields:

$$\mathcal{H}_s = \{\mathbf{x} \in \mathbb{R}^3, 0 \in 1 + x_1 + x_2 - 2\lambda(x_1 + 1)\,, \lambda \in \mathrm{s}(x_3)\}\,. \tag{85}$$

If $x_1 + 1 \geq 0$ and $x_3 = 0$, we obtain the enclosure:

$$0 \in [x_2 - x_1 - 1, 1 + x_1 + x_2]\,. \tag{86}$$

This implies that solutions *in the switching surface* for $(x_1 + 1) \geq 0$ are such that $x_2 \leq x_1 + 1$ and $x_2 \geq -(x_1 + 1)$. This is the right red cone in Figure 12. Similarly, if $x_1 + 1 \leq 0$, we then obtain:

$$0 \in [1 + x_1 + x_2, x_2 - x_1 - 1]\,. \tag{87}$$

The solutions are such that $x_2 \geq x_1 + 1$ and $x_2 \leq -(x_1 + 1)$. This is the left red cone in Figure 12. The graph of the relaxation on the switching surface $x_3 = 0$ is composed of two cones. The set $\mathcal{H}_s$ is still path connected, and is larger compared to the SCC relaxation, $\mathcal{H}_e$ from Proposition 11. As in the Examples 1.b and 2.b we can consider the equivalent switching constraint:

$$\begin{cases} 0 = -\,x_2 - x_1 - 1\,, & \text{if } x_3 < 0 \\ 0 = x_2 - x_1 - 1\,, & \text{if } x_3 > 0\,. \end{cases} \tag{88}$$

The relaxation yields:

$$\mathcal{H}_s = \{\mathbf{x} \in \mathbb{R}^3, 0 \in -x_2 - x_1 - 1 + 2\lambda x_2\,, \quad \lambda \in \mathrm{s}(x_3)\}\,. \tag{89}$$

Similarly, the set of solutions is composed of two cones. If $x_2 \geq 0$, then $x_2 \geq -x_1 - 1$ and $x_2 \geq x_1 + 1$ the top green cone in Figure 12. If $x_2 \leq 0$, then $x_2 \leq -x_1 - 1$ and $x_2 \leq x_1 + 1$ that is the bottom green cone. In this example, it is even harder to select a "correct representation" between equivalent switching constraints as both relaxations are path connected and very similar in shape. Moreover, none of the relaxation defines a convex set on the switching surface unlike what we observed on the simpler planar examples. ■

Remark 7. Let us observe that the MSO relaxation proposed in (77) appears to be the limit of the smooth singularly perturbed relaxation:

$$0 = (1 - \lambda(\alpha, \mathbf{x}))g_1(\mathbf{x}) + \lambda(\alpha, \mathbf{x})g_2(\mathbf{x})\,, \tag{90}$$

where $\lambda(\alpha, \mathbf{x})$ is a smooth approximation of the multivalued step operator $\mathrm{s}(\cdot)$ that converges to $\mathrm{s}(\cdot)$ as $\alpha \to \infty$. In Figures 13 and 14, we plot the singularly perturbed versions of the Example 2.b and the Figure 11.

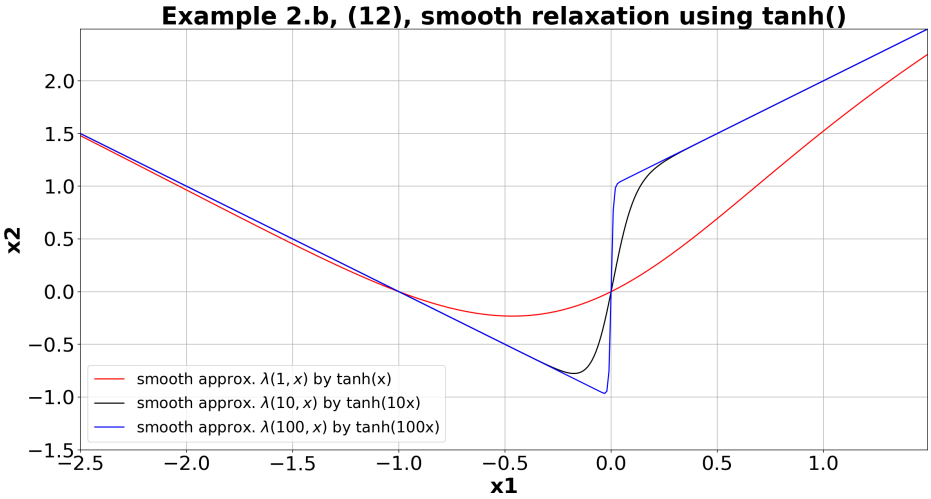


Fig. 13: Example 2.b smooth version of (12).

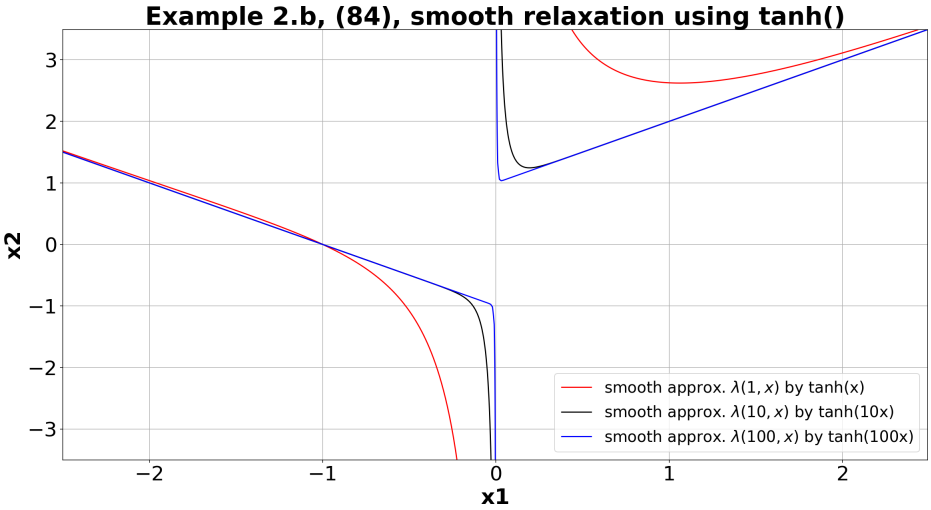


Fig. 14: Example 2.b smooth version of (84).

4.3 Closed Convex Hull (CCH) relaxation

In Section 4.2, we have seen that although the relaxation (77) seems to be the limit of the singularly perturbed relaxation (90) and has been used in previous works for switching ODEs, it may be hazardous to use it without further arguments. Indeed, two equivalent switching constraints may be relaxed differently. In the example in Section 2, we noticed that the relaxation (77) yields the same graph as a convex relaxation of the "left-hand" and "right-hand" constraints on the switching surface (in a similar manner to a Filippov relaxation in switching ODEs). In this section, we study and discuss this convex relaxation.

Definition 7 (Closed Convex Hull (CCH) relaxation). *Let us consider $\mathcal{S}^-$ the set of solutions associated with the intersection of the continuous extension of the left constraint and the switching hyperplane:*

$$\mathcal{S}^- = \{0 = \ \mathrm{C}_1\mathbf{x} + \mathrm{p}_1 = g_1(\mathbf{x}) \ \textit{and}\ 0 = \ \mathrm{H}\mathbf{x} + \mathrm{q} = h(\mathbf{x})\}\,. \tag{91}$$

Similarly, let us consider $\mathcal{S}^+$ the set of solutions associated with the intersection of the continuous extension of the right constraint and the switching hyperplane:

$$\mathcal{S}^+ = \{0 = \ \mathrm{C}_2\mathbf{x} + \mathrm{p}_2 = g_2(\mathbf{x}) \ \textit{and}\ 0 = \ \mathrm{H}\mathbf{x} + \mathrm{q}\}\ . \tag{92}$$

We define the Closed Convex Hull (CCH) relaxation of the switching constraint as :

$$\begin{cases} \{\mathbf{x} \in \mathbb{R}^{n_1} \mid 0 = \ g_1(\mathbf{x})\}\,, & \textit{if } h(\mathbf{x}) < 0 \\ \qquad \mathbf{x} \in \overline{\mathrm{co}}(\mathcal{S}^-, \mathcal{S}^+)\,, & \textit{if } h(\mathbf{x}) = 0 \\ \{\mathbf{x} \in \mathbb{R}^{n_1} \mid 0 = \ g_2(\mathbf{x})\}\,, & \textit{if } h(\mathbf{x}) > 0\,. \end{cases} \tag{93}$$

which defines the set $\mathcal{H}_c$ as

$$\mathcal{H}_c = \{\mathbf{x} \in \mathbb{R}^{n_1} \ \textit{that solves}\ (93)\}\,. \tag{94}$$

Proposition 12. *The set $\mathcal{H}_c$ from Definition 7 is path-connected.*

Proof. This is can be easily verified since the convex set $\overline{\mathrm{co}}(\mathcal{S}^-, \mathcal{S}^+)$ is path-connected and one can use a transitivity argument of the path-connectivity similar to the one in the proof of Proposition 11 to obtain the path-connectivity of $\mathcal{H}_c$. □

Again, this relaxation can be applied to the various examples from Section 4.1 and Section 4.2.

Example 1.c. Let us first consider the switching constraint (72) from Example 1.a and its equivalent alternative (81) in Example 1.b. The sets $\mathcal{S}^-$ and $\mathcal{S}^+$ are the same point $\mathbf{x}_s = (0, 0)$, and the convex hull is the singleton $\{\mathbf{x}_s\}$. The CCH relaxation has the benefit to be independent of the choice of the constraint normal vectors C_1 and C_2. ■

Remark 8. While in Section 4.2 the relaxation is the zero space of a convex combination of two affine applications, which does not have a clear structure, in the CCH relaxation, we construct a convex combination of the zero spaces of two affine applications. This definition is independent of the choice of the affine application as long as their zero spaces are the same. In particular in Example 1.c, $f : \mathbf{x} \mapsto x_2$ and $f' : \mathbf{x} \mapsto -x_2$ have the same zero space, and changing the representation of the "left-hand" constraint does not influence the relaxation.

Example 2.c. For the example considered in Section 2, the switching surface given by the CCH relaxation is the set $\{x_1 = 0, x_2 \in [-1, 1]\}$, whatever the choice of the normal vectors, unlike in Example 2.b. ■

Example 3.c. Let us emphasize, in this example, that the CCH relaxation is not necessarily included in one of the possible solutions of the MSO relaxation from Definition 6. Indeed, when applying the CCH relaxation of Definition 7 to the switching constraints from Examples 3.a and 3.b, we obtain consistently the whole hyperplane $\{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 = 0\}$. In this particular case, the solution set of the MSO relaxed constraint are included in the CCH relaxation.■

As seen in these examples, the CCH relaxation provides a consistent way to fill-in the graph of the constraint. However, unlike the automatic representation by a generalised equation in Definition 6, we need additional work to rewrite the solution set as the solution of an closed-form generalised equation.

4.4 Discussion

In the Section 4.1, 4.2, and 4.3 we have presented three different relaxation methods that may be used to obtain a path-connected constraint. Let us now discuss the effect of each relaxation on the actual solutions of state-dependent switching DAE. In the foregoing examples, only the examples 3.a, 3.b, and 3.c, yield different solution sets for each relaxation. Consequently, in order to put more emphasis on the differences between these relaxations, let us consider the switching constraint (75).

Example 3.d. Let us consider the switching constraint (75) from Example 3.a in the context of a state-dependent switching DAE. For the purpose of comparing the three possible relaxations, let us define the state-dependent switching DAE:

$$\begin{cases} \dot{x}_1 = 1 - 2z & \\ \dot{x}_2 = z & \text{(95a)} \\ \dot{x}_3 = 1 - 0.5z & \\ 0 = 1 + x_1 + x_2\,, \qquad \text{if } x_3 < 0 & \text{(95b)} \\ 0 = -1 - x_1 + x_2\,, \qquad \text{if } x_3 > 0\,. & \text{(95c)} \end{cases}$$

Unlike Section 2, we will not detail all the possible solutions for each relaxation, but we will only highlight a particular case where each successive relaxation

allows a greater set of solutions. Let us assume that $x_3(0) < 0$, it follows from (95a) and (95b) that $z(0) = 1$. On the "left-hand" constraint (95b), the solutions $\mathbf{x}(t)$ are given by:

$$\begin{cases} x_1(t) = -t + x_1(0) \\ x_2(t) = t + x_2(0) \\ x_3(t) = 0.5t + x_3(0)\,. \end{cases} \tag{96}$$

Since we are interested in the differences for each solution at the switching on the sliding surface, we study the set of initial conditions in $\mathbf{x}(\cdot)$ such that there exists a sliding motion on the surface $x_3 = 0$. Let us notice that $\dot{x}_3(t) > 0$ in (96): if there is no sliding solution, then there is no AC solution $\mathbf{x}(t)$ after $t' > 0$ with $x_3(t') = 0$. Let us first recall from Example 3.a that there exists on the switching surface a set $\mathcal{S} = \{(-1, 0, 0)\}$ that is solution of (70). From Proposition 11, we can build an SCC relaxation $\mathcal{H}_s$, and an associated relaxed DAE.

$$\begin{cases} \dot{x}_1 = 1 - 2z \\ \dot{x}_2 = z \\ \dot{x}_3 = 1 - 0.5z \end{cases} \tag{97}$$
$$\begin{cases} 0 = 1 + x_1 + x_2\,, & \text{if } x_3 \leq 0 \\ 0 = -1 - x_1 + x_2\,, & \text{if } x_3 \geq 0\,. \end{cases}$$

It can be immediately noticed that there exists an AC solution $\mathbf{x}(\cdot)$ reaching the switching surface at $t' > t_0$ only if $(x_1(t'), x_2(t')) = (-1, 0)$ when $x_3(t') = 0$. From (96) we notice this is possible if and only if the initial conditions are on the half-line:

$$\mathcal{I}_a = \left\{ \mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ -0.5 \end{pmatrix} u + \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}\ ,\ u \geq 0 \right\}. \tag{98}$$

For any other initial conditions, under the assumption $x_3(0) < 0$, there is no solution for $t' > 0$ such that $x_3(t') = 0$.

Let us now consider the MSO relaxation (77) applied to this problem. As seen in Example 3.b, the choice of the constraint normal vector in the switching constraint representation is critical, as the graph of the relaxation may change on the switching surface (see Figure 12). There exists a sliding solution on the switching surface $x_3(t) = 0$ if and only if $z(t) = 2$, and it follows that $(\dot{x}_1(t), \dot{x}_2(t)) = (-3, 2)$. It appears that apart from the initial conditions inside $\mathcal{I}_a$ in (98), we can extend the set of initial conditions to obtain a sliding motion. However, this is only in the case for the switching constraint (88) where $C_1' = (-1, -1, 0)$. Using the relaxed constraint (89) we obtain the switching DAE:

$$\begin{cases} \dot{x}_1 = 1 - 2z \\ \dot{x}_2 = z \\ \dot{x}_3 = 1 - 0.5z \\ \quad 0 \in -x_2 - x_1 - 1 + 2\lambda x_2\,, \quad \lambda \in \mathrm{s}(x_3)\,. \end{cases} \tag{99}$$

Then, the set of initial conditions such that there exists a sliding motion can be extended to $\mathcal{I}_b$ in (100) and solutions can slide temporarily in the cone $x_2 \leq 0$, $x_2 \leq -x_1 - 1$, and $x_2 \leq x_1 + 1$ (bottom green cone in Figure 12).

$$\mathcal{I}_b = \left\{ \mathbf{x} \in \mathbb{R}^3 \mid x_3 < 0 \,, x_2 \leq 2x_3 \,, x_2 = -x_1 - 1 \right\} . \tag{100}$$

It is interesting to note that for the dynamics given in (99) the solutions slide a finite amount of time on the surface $x_3(t) = 0$, and inevitably reach the "right-hand" constraint (95c) and continue on it.

Finally, let us consider the CCH relaxation in (93), i.e.,

$$\begin{cases} \dot{x}_1 = 1 - 2z \\ \dot{x}_2 = z \\ \dot{x}_3 = 1 - 0.5z \\ \quad \mathbf{x} \in \mathcal{H}'_c \,, \end{cases} \tag{101}$$

$$\mathcal{H}'_c = \{ \mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x} = (1-\lambda)\,(A_1\mathbf{u} + \mathbf{b}_1) + \lambda\,(A_2\mathbf{u}' + \mathbf{b}_2) \;,\; \lambda \in \mathrm{s}(x_3) \} \,, \tag{102}$$

where $\mathbf{u}', \mathbf{u} \in \mathbb{R}^2$, $A_1 = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{b}_1 = \mathbf{b}_2 = (-1, 0, 0)^{\mathrm{T}}$. As stated in Example 3.c, the solution set of this relaxation on the switching surface is the whole plane $x_3 = 0$. Then, the whole "left-hand" constraint (95b) defines the set initial values $\mathcal{I}_c$ in (103) such that there exists sliding motion. However, some solutions may never leave the sliding surface as they may not intersect the right constraint.

$$\mathcal{I}_c = \left\{ \mathbf{x} \in \mathbb{R}^3 \mid x_3 < 0 \,, x_2 = -x_1 - 1 \right\} . \tag{103}$$

Overall, for this example the CCH relaxation (93) is the less constraining, but may define too large a set if the goal is to reach the right constraint as this may be intended for modelling purpose. ■

To summarise, this example demonstrates that each relaxation has pros and cons. The simplest SCC relaxation from Proposition 11 is independent of the representation of the constraint, but also has the smallest set of valid initial conditions for the existence of continuous solutions. Furthermore, the SCC relaxation may not always have solutions (see Example 2.a). The MSO relaxation (Definition 6) is dependent on the choice of the constraint representation. However, assuming the correct representation, this relaxation produces a transitory sliding surface while keeping the uniqueness of the solutions. Finally, on this particular example the CCH relaxation from Definition 7 ensures the existence of solutions, and is independent from the constraint representation. However, the uniqueness of the solutions is lost and some solutions may never leave the sliding surface.

5 Summary and Conclusion

5.1 Summary

Throughout this article, we propose a concept of (AC) solutions for state-dependent switching DAEs. We define solutions by considering a path-connected relaxation on the switching surface of the "left-hand" and "right-hand" algebraic constraints. In particular, three different strategies are introduced to obtain such relaxation. The CCH relaxation from Section 4.3 can be understood as a Filippov relaxation of the constraints on the switching surface. The MSO relaxation from Section 4.2 can be understood as a Filippov in the narrow-sense (or Aizerman-Pyatnitskiy) relaxation of the constraints on the switching surface. And, the SSC relaxation from Section 4.1 can be related to a Caratheodory extension of the constraints on the switching surface.

In the Section 2, we first consider a simple example of state-dependent switching DAE where the discontinuous constraint has been relaxed using the CCH relaxation, or the MSO relaxation. Using this example we show, in Section 2.1, that we can now construct AC solutions crossing the switching surface, and obtain a sliding motion. In addition, we also show that in this context the contingent cone can be used to obtain necessary and sufficient condition for the existence of local AC solutions. Furthermore, in this example, the contingent cone acts similarly to the tangent space in the concept of solutions for smooth DAEs. Using our concept of solution based on the constraints relaxation, we obtain well-posed solutions named "sliding-crossing" solutions.

In Section 2.3, we extend the study to solutions of bounded variations in order to discuss the problem of jumping solutions, or re-initialisation, when there is no possible continuation with an AC solution in some point of the relaxed constraint. In order to analyse this re-initialisation problem, we first study in Section 2.3.1, the jump dynamic boiling down from the measure representation of the state-dependent switching DAE. Then, we study in Section 2.3.3 the consistent jumps that define valid re-initialisation points, and we introduce an associated jump law using the contingent cone to the solution set of the relaxed constraint.

In the Section 3.1, we study the solutions of the implicit Euler numerical scheme for the simulation of our newly defined solutions. This event-capturing scheme has proven to be an efficient method for the simulation of non-smooth dynamical system such as Linear complementary system (LCS), but is also widely use for the simulation of index-2 linear DAE. We show that the implicit Euler scheme cannot be used as reliable method for the simulation of such system. In Section 3.2, we provided a refinement of the Euler scheme to address the observed problems. Further properties of this numerical scheme are provided through additional simulations in Section 3.3.

In the Section 4, we expose three concepts of solutions associated with three strategies to relax the constraint on the switching surface. We show in Section 4.2, that the intuitive MSO relaxation, that is the convex combination of the "left-hand" and "right-hand" constraints, may not necessarily result in a path-connected relaxation. On the other hand in Section 4.3, the CCH relaxation,

that is a "Filippov-like" relaxation of the constraints, always provides a path-connected relaxation. Finally, we discuss the variety of different solutions as well as the pros and cons of each concept of solutions in Section 4.4.

5.2 Conclusion

Overall, in this paper we explored, through various examples, new concepts of solutions for state-dependent switching DAEs of index-2. We show that these new methods may help to provide additional well-posed solutions to such system, which can be studied using classical mathematical tools, and simulated using some improvement of the well-known implicit Euler scheme. However, we also show that providing general results on such solutions will be a difficult problem as these new constraints cannot always be expressed in a reliable mathematical framework, where well-posedness can be more easily studied, such as MLCS or differential inclusion. In order to further extend the results, various paths are possible. The first one is to study when a relaxed constraint is equivalent to an MLCP, or an inclusion into a maximal monotone operator such as in [9]. For example, one can investigate whether or not the constraint can be rewritten in absolute normal form (ANF) [23]. Then, an MLCP conversion exists and can be studied. Another research path is to look further into the role played by the contingent cone, which seems similar to the tangent space for classical DAEs.

Acknowledgement

This work was supported by the FUI ModeliScale DOS0066450/00 French national grant (2018-2021) and the Inria IPL ModeliScale large scale initiative (2017-2021, `https://team.inria.fr/modeliscale/`).

References

1. Acary, V., Bonnefon, O., and Brogliato, B. *Nonsmooth Modeling and Simulation for Switched Circuits*, vol. 69 of *Lecture Notes in Electrical Engineering*. Springer Science & Business Media, 2010.
2. Acary, V., and Brogliato, B. *Numerical Methods for Nonsmooth Dynamical Systems: Applications in Mechanics and Electronics*, vol. 35 of *Lecture Notes in Applied and Computational Mechanics*. Springer Science & Business Media, 2008.
3. Acary, V., De Jong, H., and Brogliato, B. Numerical simulation of piecewise-linear models of gene regulatory networks using complementarity systems. *Physica D: Nonlinear Phenomena 269* (2014), 103–119.
4. Acary, V., and Pérignon, F. Siconos: A software platform for modeling, simulation, analysis and control of nonsmooth dynamical systems. *HAL INRIA* (2007).
5. Aizerman, M., and Pyatnitskiy, E. Fundamentals of the theory of discontinuous systems. i. *Avtom. Telemekh 7 & 8* (1974), 33–48 & 39–62.
6. Barton, P. I., Khan, K. A., Stechlinski, P., and Watson, H. A. Computationally relevant generalized derivatives: theory, evaluation and applications. *Optimization Methods and Software 33*, 4-6 (2018), 1030–1072.

7. Benveniste, A., Caillaud, B., Elmqvist, H., Ghorbal, K., Otter, M., and Pouzet, M. Structural analysis of multi-mode DAE systems. In *Proceedings of the 20th International Conference on Hybrid Systems: Computation and Control, Pittsburgh PA, U.S.A., April 18–20, 2017* (2017), ACM, pp. 253–263.
8. Brogliato, B. *Nonsmooth Mechanics, 3^{rd} Edition*. Springer Communications and Control Engineering, 2016.
9. Camlibel, K., Iannelli, L., Tanwani, A., and Trenn, S. Differential-algebraic inclusions with maximal monotone operators. In *2016 IEEE 55th Conference on Decision and Control (CDC)* (2016), IEEE, pp. 610–615.
10. Facchinei, F., and Pang, J.-S. *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer Science & Business Media, 2003.
11. Filippov, A. F. Differential Equations with Discontinuous Right-hand Side. *Matematicheskii sbornik 93*, 1 (1960), 99–128.
12. Hamann, P., and Mehrmann, V. Numerical solution of hybrid systems of differential-algebraic equations. *Computer Methods in Applied Mechanics and Engineering 197*, 6-8 (2008), 693–705.
13. Henningsson, E., Olsson, H., and Vanfretti, L. DAE solvers for large-scale hybrid models. In *Proceedings of the 13th International Modelica Conference, Regensburg, Germany, March 4–6, 2019* (2019), no. 157, Linköping University Electronic Press.
14. Khan, K. A. Branch-locking AD techniques for nonsmooth composite functions and nonsmooth implicit functions. *Optimization Methods and Software 33*, 4-6 (2018), 1127–1155.
15. Matrosov, I. V. Existence of solutions of the algebro-differential equations. *Automation and Remote Control 67*, 9 (2006), 1408–1415.
16. Matrosov, I. V. On right-hand uniqueness of solutions to nondegenerated algebro-differential equations with discontinuances. *Automation and Remote Control 68*, 1 (2007), 9–17.
17. Mehrmann, V., and Wunderlich, L. Hybrid systems of differential-algebraic equations–analysis and numerical solution. *Journal of Process Control 19*, 8 (2009), 1218–1228.
18. Monteiro, M. D. P. *Differential Inclusions in Nonsmooth Mechanical Problems: Shocks and Dry Friction*, vol. 9. Birkhäuser, 2013.
19. Pang, J.-S., and Stewart, D. E. Differential variational inequalities. *Mathematical Programming 113*, 2 (2008), 345–424.
20. Rabier, P. J., and Rheinboldt, W. C. A geometric treatment of implicit differential-algebraic equations. *Journal of Differential Equations 109*, 1 (1994), 110–146.
21. Rocca, A., Acary, V., and Brogliato, B. Index-2 hybrid DAE: a case study with well-posedness and numerical analysis. In *IFAC World Congress 2020* (Berlin, Germany, July 2020).
22. Rockafellar, R. T., and Wets, R. J.-B. *Variational Analysis*, vol. 317. Springer Science & Business Media, 2009.
23. Scholtes, S. *Introduction to Piecewise Differentiable Equations*. Springer Science & Business Media, 2012.
24. Souza, D. F. d. S., Vieira, R. C., and Biscaia Jr, E. C. Strategies for numerical integration of discontinuous DAE models. In *Computer Aided Chemical Engineering*, vol. 20. Elsevier, 2005, pp. 151–156.
25. Stechlinski, P., Patrascu, M., and Barton, P. I. Nonsmooth differential-algebraic equations in chemical engineering. *Computers & Chemical Engineering 114* (2018), 52–68.

26. Trenn, S. Switched differential algebraic equations. In *Dynamics and Control of Switched Electronic Systems*. Springer, 2012, pp. 189–216.
27. Ye, J. J. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *Journal of Mathematical Analysis and Applications 307*, 1 (2005), 350–369.

A Appendix

A.1 Notations and Definitions

Vectors of real variables $\mathbf{x} = (x_1, \dots, x_i, \dots, x_n)^{\mathrm{T}}$ in $\mathbb{R}^n$ are noted in **bold**. In the context of algebraic differential systems, the variables $\mathbf{x}$ and $\mathbf{y}$ will denote differential variables, while $\mathbf{z}$ will denote algebraic variables. In the context of non-smooth expressions, we will in general denote as $\boldsymbol{\lambda}$ the Lagrange multipliers. Finally, for a given function $\mathbf{f} : \mathbb{R} \to \mathbb{R}^n$ we denote $\mathbf{f}(t^+)$ the right limit of $\mathbf{f}(\cdot)$ at t:

$$\mathbf{f}(t^+) = \lim_{\substack{\tau \to t \\ \tau > t}} \mathbf{f}(\tau)\,.$$

Definition 8 (Absolute Continuity). *Given a compact interval $I = [t_1, t_2] \subseteq \mathbb{R}$, a function $f : I \to \mathbb{R}$ is absolutely continuous (also noted* AC*) on I if and only if the following property holds: f has a derivative $\dot{f}$ a.e. (almost everywhere), the derivative is Lebesgue integrable, and*

$$f(t) = f(t_1) + \int_{t_1}^{t} \dot{f}(\tau) d\tau\,,$$

for all t in$[t_1, t_2]$.

Definition 9 (Bounded Variation). *Let us consider an interval $I \neq \emptyset$, $I \subset \mathbb{R}$. Let $\mathcal{P}$ denotes the set of finite partitions of I, each partition P_N associated with nodes $t_0 < t_1 < \cdots < t_N$. The variation of a function $x : I \to X$ on a partition P_N of I is defined by:*

$$\mathrm{var}(x, I) = \sup_{P_N \in \mathcal{P}} \sum_{i=1}^{N} \|x(t_i) - x(t_{i-1})\|\,, \tag{104}$$

The function $x(\cdot)$ is of bounded variations (BV) on I if and only if $\mathrm{var}(u, I) < +\infty$

We now introduce important tools of convex analysis we will use in the study of variational equations:

Definition 10 (Normal cone). *Let $K \subseteq \mathbb{R}^n$ be a closed non-empty convex set. The normal cone to K at $\mathbf{x} \in K$ is the convex set:*

$$\mathcal{N}_K(\mathbf{x}) = \{d \in \mathbb{R}^n | \langle d, \mathbf{y} - \mathbf{x} \rangle\,,\ \forall \mathbf{y} \in K\}$$

Definition 11 (Subdifferential). *A vector $\gamma \in \mathbb{R}^n$ is said to be a subgradient of a proper lower semi-continuous convex function $f(\cdot)$ at a point $\mathbf{x}$ if it satisfies:*

$$f(\mathbf{y}) - f(\mathbf{x}) \geq \gamma^{\mathrm{T}}(\mathbf{y} - \mathbf{x}),$$

for all $\mathbf{y} \in \mathbb{R}^n$. The set of all subgradients of $f(\cdot)$ at $\mathbf{x}$ is the subdifferential of $f(\cdot)$ at $\mathbf{x}$ and is noted $\partial f(\mathbf{x})$.

Definition 12 (Conjugate Function). *Let us consider $\mathbf{f} : \mathbb{R}^n \to (-\infty, +\infty]$. The conjugate of $\mathbf{f}$ is $\mathbf{f}^* : \mathbb{R}^n \to [-\infty, +\infty]$ such that:*

$$\mathbf{f}^*(\mathbf{y}) = \sup_{\mathbf{x} \in \mathbb{R}^n} \left(\langle \mathbf{x}, \mathbf{y} \rangle - \mathbf{f}(\mathbf{x}) \right) \quad , \forall \mathbf{y} \in \mathbb{R}^n . \tag{105}$$

Definition 13 (Indicator function). *The indicator function of a set $K \subseteq \mathbb{R}^n$, $\psi_K : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, is defined by :*

$$\psi_K(\mathbf{x}) = \begin{cases} 0 , & \text{if } \mathbf{x} \in K \\ +\infty , & \text{if } \mathbf{x} \notin K \end{cases}$$

If K is a closed, non-empty convex set, then $\psi_K(\mathbf{x})$ is convex proper lower semi-continuous and $\partial \psi_K = \mathcal{N}_K(\mathbf{x})$.

Definition 14 (Contingent Cone). *The contingent cone or Bouligand tangent cone to a set $K \subset X$ is given by:*

$$\begin{aligned} \mathcal{T}_K(\mathbf{x}_0) = \{ \mathbf{x} \in X \mid \exists \{t_n\} \in \mathbb{R}, \{\mathbf{x}_n\} \subset X, \text{ with } t_n \downarrow 0 , \ (t_n > 0), \\ \text{and } \mathbf{x}_n \to \mathbf{x}, \text{ s.t. } \mathbf{x}_0 + t_n \mathbf{x}_n \in K, \forall n \in \mathbb{N} \} \end{aligned} \tag{106}$$

Definition 15 (Consistency[11] of a numerical method). *Given a numerical method of the form:*

$$\mathbf{y}_{k+1} = \mathbf{y}_k + h_k \mathbf{f}(t_k, \mathbf{y}_k, h_k) \text{ for all } k \geq 0, \ t_0 = 0, \ t_{k+1} = t_k + h_k \tag{107}$$

This numerical method is said to be consistent for a DAE with $\mathbf{y}(0) = \mathbf{y}_0$, if for any solution of this DAE the consistency error

$$\sum_{k=0}^{N-1} \| \mathbf{y}(t_{k+1}) - \mathbf{y}(t_k) - h_k \mathbf{f}(t_k . \mathbf{y}(t_k), h_k) \| \tag{108}$$

tends to 0 when $h = \max\limits_{0 \leq k \leq N} h_k$ tends to 0.

Definition 16 (Connected Sets). *A topological space $\mathcal{C}$ is said to be* disconnected *if it is the union of two disjoint non-empty open sets. Otherwise, $\mathcal{C}$ is said to be* connected.

[11] Please note the difference with the notion of consistent initial conditions.

Definition 17 (Path). *Let us assume that $\mathcal{C}$ is a non-empty subset of $\mathbb{R}^n$ and that x and $y \in \mathcal{C}$. Then a* continuous *function $f : [0,1] \to \mathcal{C}$ where $f(0) = x$ and $f(1) = y$ is called a path in $\mathcal{C}$ from x to y.*

Remark 9. For two points $x, y \in \mathcal{C}$, let us denote $x \sim y$ the relation: there exists a path between x and y. Then, the relation $\sim$ is an equivalence relation on $\mathcal{C}$:

- $x \sim x$
- If $x \sim y$ then $y \sim x$
- If $x \sim z$ and $z \sim y$ then $x \sim y$.

Definition 18 (Path-Connected Sets). *A subset $\mathcal{C}$ of $\mathbb{R}^n$ is said to be path-connected if and only if, for all $x, y \in \mathcal{C}$, there is a path in $\mathcal{C}$ from x to y: there exists a* continuous *function $f : [0,1] \to \mathcal{C}$ where $f(0) = x$ and $f(1) = y$.*

Definition 19 (Linear Complementarity Problem). *A Linear Complementarity Problem (LCP) is a set of equations:*

$$\begin{aligned} w &= M\lambda + q \\ 0 \leq w \perp \lambda &\geq 0\,, \end{aligned} \tag{109}$$

where the complementarity equation $0 \leq w \perp \lambda \geq 0$ means $\langle w, \lambda \rangle = 0$, for all $w, \lambda \geq 0$. Let us remark that an LCP (109) *has a unique solution if and only if M is a P-matrix (i.e., all its principal minor are positive).*

A.2 Proof of Proposition 6

Proof.
• Assume $\underline{\mathrm{B}_1 \neq 0}$, then (51) can be further reduced into:

$$0 \in \ \mathrm{sign}\left(x_1(t_j^+)\right) + |x_1(t_j^+)| - \frac{\mathrm{B}_2}{\mathrm{B}_1}\left(x_1(t_j^+) - x_1(t_j^-)\right) - x_2(t_j^-)\,, \tag{110}$$

which is a generalised equation (GE) of the form:

$$0 \in f(x) + \mathcal{F}(x)\,, \tag{111}$$

where $\mathcal{F} : \mathbb{R} \rightrightarrows \mathbb{R}$ is the maximal monotone operator $\mathrm{sign}(\cdot)$, $f : \mathbb{R} \to \mathbb{R}$ is a continuous function, and $x = x_1(t_j^+)$. In particular, here we have:

$$f(x) = \mathrm{a}x + \mathrm{b}|x| + \mathrm{c}\,, \tag{112}$$

with $a = \frac{-\mathrm{B}_2}{\mathrm{B}_1}$, $b = 1$, and $c = \frac{\mathrm{B}_2}{\mathrm{B}_1} x_1(t_j^-) - x_2(t_j^-)$. We can first notice that by assumption, $\mathbf{x}(t_j^-)$ is solution of (110), and it follows that in this particular context of jump dynamics there is always existence of solutions. However, we will give a more in-depth study of the solutions of (112) in a general context, as it will prove to be useful for the study of numerical solutions in Appendix A.3. Let us now study the conditions for existence and/or uniqueness of solutions to

such GE. For the sake of briefness, proofs will be given in a succinct manner using Figure 15 as a support.

Assume $a-b \neq 0$ and $a+b \neq 0$[12]. Then, depending on the signs of $a-b$ and $a+b$, we define an associated continuous piecewise-linear function $h(y)$.

(a) Let us first consider $a-b>0$ and $a+b>0$, which is equivalent to $\mathrm{B}_2/\mathrm{B}_1 < -1$, and define $h(y) = f^{-1}(\mathbf{x})$ with $h(y) = (y-c)/(a-b)$ if $y \leq c$ and $h(y) = (y-c)/(a+b)$ if $y \geq c$. As we will see, an important point is that under this assumption, $h(\cdot)$ is a continuous function with domain $\mathbb{R}$.
We note that $\mathrm{sign}(x) = \partial|x|$, and the conjugate of $g(x) = |x|$ is $g^*(x) = \psi_{[-1,1]}(y)$ (see Definition 12 and 13). It follows [22] that the inverse of $\mathrm{sign}(x)$ is $\mathcal{N}_{[-1,1]}(y) = \partial\psi_{[-1,1]}(y)$ (see Definition 10). Consequently, the solutions of the GE (111) can be obtained by solving the inverse problem:

$$0 \in h(y) + \mathcal{N}_{[-1,1]}(y)\,, \tag{113}$$

with $\mathcal{N}_{[-1,1]}(y)$ the normal cone to $[-1,1]$ at y. This is the canonical form of a generalised equation as analysed in [10]. Finally, $h(\cdot)$ is continuous on $\mathbb{R}$ and as $[-1,1]$ is a compact set in $\mathbb{R}$, it follows from [10, Corollary 2.2.5] that there is always existence of solutions to the GE (113) (and equivalently (110)).
Additionally, [10, Theorem 2.3.3] provides sufficient conditions for uniqueness: if $h(\cdot)$ is strictly monotone on $[-1,1]$ then equation (113) has at most one solution. Strict monotonicity of $h(\cdot)$ holds if and only if $a+b>0$ and $a-b>0$, as it can be seen using the green lines in Figure 15. Furthermore, we already know that the trivial solution $\mathbf{x}(t_j^+) = \mathbf{x}(t_j^-)$ always exists: if there is a unique solution then this solution is $\mathbf{x}(t_j^+) = \mathbf{x}(t_j^-)$, and $\sigma_z = 0$.
We have proved that $a-b>0$ and $a+b>0$ are sufficient conditions for existence and uniqueness of solutions to (112). In addition, these sufficient conditions do not depend on c and consequently do not depends on $\mathbf{x}(t_j^-)$. However, they are not necessary conditions.

(b) Indeed, let us now consider the case where $a-b<0$ and $a+b>0$, which is equivalent to $\mathrm{B}_2/\mathrm{B}_1 \in (-1,1)$. Then, we can try to build another "piecewise linear function" $h(y)$ by inverting the equation $y = ax+b|x|+c$ for all $x \in \mathbb{R}$. If $x \leq 0$, $y = (a-b)x - c$ that is $x = (y-c)/(a-b)$ if $y \geq c$. Similarly, if $x \geq 0$ then $y = (a+b)x + c$ and $x = \frac{y-c}{a+b}$ if $y \geq c$. It follows that $h(\cdot)$ is a multi-valued operator defined on $[c, +\infty)$: it corresponds to the red line $(a-b) < 0$ and the green line $(a+b) > 0$ in Figure 15. Consequently, if $c > 1$ there are no solutions as the domains of $h(\cdot)$ and $\mathcal{N}_{[-1,1]}(\cdot)$ do not intersect. If $c \leq 1$, there is either a unique solution for $c = 1$, or multiple solutions if $c < 1$ as $(y-c)/(a-b)$ intersects the normal cone both in $y = -1$ and $y = c$. As $\mathbf{x}(t_j^-)$ satisfies (49), and consequently $\sigma_z = 0$ is always a solution, we know by construction that $c \leq 1$. However, it also possible

[12] The case $a-b=0$ (respectively $a+b=0$) corresponds to parameterization in mode 1 (respectively mode 2) where the DAE is not regular and $\mathrm{C}_1\mathrm{B}$ (respectively $\mathrm{C}_2\mathrm{B}$) is singular.

to prove it by computing c for all the values of $x_1(t_j^-)$. From the definition of c, from (49), and $\mathrm{B}_2/\mathrm{B}_1 \in (-1,1)$, we obtain $c \in (0, 2x_1(t_j^-)+1)$, and $c \leq 1$ if $x_1(t_j^-) < 0$. Similarly, if $x_1(t_j^-) = 0$, we compute $c \in [-1,1]$. If $x_1(t_j^-) > 0$ then $c \in (-2x_1(t_j^-)-1, 0)$. It follows that the case where there is the uniqueness of solution for $c = 1$ corresponds to $\mathbf{x}(t_j^-) = (0,-1)^{\mathrm{T}}$. Finally, a similar reasoning can be done for the case with $a-b>0$ and $a+b<0$. However, we notice this last case is not possible if $b = 1$ as in the particular example (110).

(c) In the case where $a-b<0$ and $a+b<0$ (which is equivalent to $\mathrm{B}_2/\mathrm{B}_1 > 1$), the associated piecewise-linear function $h(y)$ is continuous over $\mathbb{R}$ (this corresponds to the red lines in Figure 15), and again using [10, Corollary 2.2.5] we can prove there is always existence of solutions, independently of c. Under these assumptions on $a-b$ and $a+b$, it can be proved using the respective graphs of the functions $h(\cdot)$ and $-\mathcal{N}_{[-1,1]}(\cdot)$ that there is uniqueness if and only if $c > 1$ (intersection of $(y-c)/(a+b)$ with the normal cone in $y=-1$) or $c<1$ (intersection of $(y-c)/(a-b)$ with the normal cone in $y=1$). Again, these conditions for uniqueness of solutions translate into $x_1(t_j^-) < -2/(1+\mathrm{B}_2/\mathrm{B}_1) < 0$ or $x_1(t_j^-) > 2/(\mathrm{B}_2/\mathrm{B}_1 - 1) > 0$.

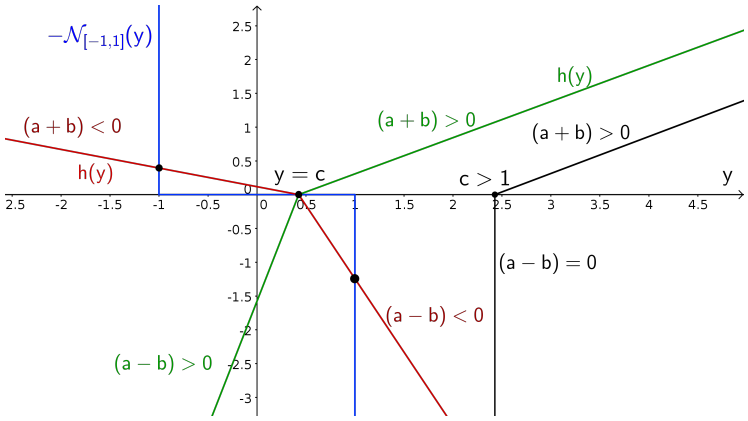


Fig. 15: Solutions of the the GE (113) for various signs of $(a+b)$ and $(a-b)$. In green an example of case (a) from the proof of Proposition 6 is depicted. In red, an example of case (b). Case (c) is constituted of a mix between the top and the bottom of the red and green graphs.

• Assume $\underline{\mathrm{B}_1 = 0}$, then from (50) it follows that:

$$\begin{cases} x_1(t_j^+) = x_1(t_j^-) \\ x_2(t_j^+) = x_2(t_j^-) + \mathrm{B}_2\sigma_z \\ x_2(t_j^+) \in |x_1(t_j^-)| + \mathrm{sign}(x_1(t_j^-))\,. \end{cases} \tag{114}$$

We note that if $x_1(t_j^-) \neq 0$ there is a unique solution, since $|x_1(t_j^-)|+\text{sign}(x_1(t_j^-))$ is uniquely defined on $\mathbb{R}\backslash\{0\}$. If $x_1(t_j^-) = 0$, then there are multiple solutions (infinitely many) with $x_2(t_j^+) \in [-1, 1]$. The proof is complete. □

A.3 Proof of Proposition 9

Proof.
• If $\underline{\mathrm{B}_1 \neq 0}$, then we obtain the GE:

$$0 \in -\frac{\mathrm{B}_2}{\mathrm{B}_1} x_{1,k+1} + |x_{1,k+1}| + \left(-x_{2,k} + \frac{\mathrm{B}_2}{\mathrm{B}_1}(x_{1,k} + h) \right) + \text{sign}(x_{1,k+1}) \,. \quad (115)$$

Sufficient conditions for uniqueness are the same as in Appendix A.2, identifying:

$$f(x_{1,k+1}) = -\frac{\mathrm{B}_2}{\mathrm{B}_1} x_{1,k+1} + |x_{1,k+1}| + \left(-x_{2,k} + \frac{\mathrm{B}_2}{\mathrm{B}_1}(x_{1,k} + h) \right) , \quad (116)$$

with $a = -\frac{\mathrm{B}_2}{\mathrm{B}_1}$, $b = 1$, and $c = (-x_{2,k} + (\mathrm{B}_2/\mathrm{B}_1)(x_{1,k} + h))$.

(a) Let us first consider the case $a - b = -(1 + \mathrm{B}_2/\mathrm{B}_1) > 0$ and $a + b = 1 - \mathrm{B}_2/\mathrm{B}_1 > 0$, which is equivalent to the set of vectors B such that $\mathrm{B}_2/\mathrm{B}_1 < -1$. This also corresponds to case (a) in Appendix A.2. In this case, there is existence and uniqueness to the GE (115), as we can use the same arguments we used in the case (a) of Appendix A.2. Furthermore, the condition $a-b > 0$ excludes the constant solutions with $\mathrm{B}_2 = 0$ where there is not necessarily uniqueness of the solution.

(b) Let us now consider the case $(a - b) < 0$ and $(a + b) > 0$, that is $\mathrm{B}_2/\mathrm{B}_1 \in (-1, 1)$. The study of this case is itself separated in two part: $\mathrm{B}_2/\mathrm{B}_1 \in (-1, 0]$ and $\mathrm{B}_2/\mathrm{B}_1 \in (0, 1)$.
If $\mathrm{B}_2/\mathrm{B}_1 \in (-1, 0]$, there is always existence of a discrete solution, as it can be proven that $c \in [-1, 1]$ for all $\mathbf{x}_k$ satisfying (12) and any $h > 0$. Furthermore, there is uniqueness if $c = 1$, for some h big enough, respectively small enough, depending of the sign of $x_{1,k}$. However, it is also possible to have multiple solutions whatever the choice of $h \geq 0$, in which case the implicit Euler scheme is not consistent with solutions of Problem 1 (see Definition 15) and this is illustrated in Figure 16.
If $\mathrm{B}_2/\mathrm{B}_1 \in (0, 1)$, we have seen in Appendix A.2 case (b) that there is either non-existence, existence, or existence and uniqueness. There is non-existence of solution to the implicit Euler scheme when $c > 1$: this corresponds to a time step size h such that there is no maximal solution to Problem 2 in $\mathbf{x}(t_0^-) = \mathbf{x}_k$ for the interval $[t_0, t_0 + h]$. In particular, in $\mathbf{x}_k = A^-$ there is no solution for all $h > 0$ and this corresponds to the case $\mathrm{B}_2 < \mathrm{B}_1$ in Figure 4. There is existence of solutions (but not uniqueness) for $c < 1$. Finally, uniqueness occurs when $c = 1$ which happens for some values of h and $\mathbf{x}_k$ suitably chosen.

(c) If $a - b < 0$ and $a + b < 0$, which correspond to $B_2/B_1 > 1$, there is always existence of solutions to the discrete scheme, and it can be proven using arguments similar to Appendix A.2 case (c). In particular, there is a unique solution in two cases: $c > 1$ or $c < -1$. The case $c > 1$ corresponds to $h > 0$ (or $x_{1,k} > 0$) sufficiently large, and the case $c < -1$ corresponds to $h > 0$ (or $x_{1,k} < 0$) sufficiently small. As noted above, outside these cases where there are unique solutions, there is existence but uniqueness fails. In particular, as it can be seen in Proposition 5, in the case where $B_2/B_1 > 1$, there exists a local AC solution to Problem 1 everywhere except in $A^- = (0, -1)^T$.

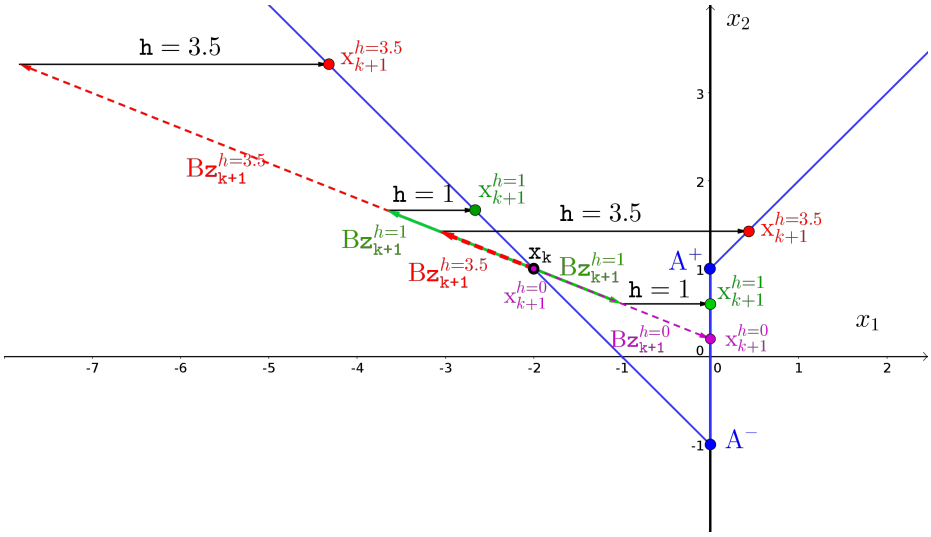


Fig. 16: Example of non consistent Euler scheme with multiple solutions whatever the choice of $h \geq 0$.

• If $B_1 = 0$, we obtain the system:

$$\begin{cases} x_{1,k+1} = h + x_{1,k} \\ hB_2 z = x_{2,k+1} - x_{2,k} \\ x_{2,k+1} \in \operatorname{sign}(h + x_{1,k}) + |h + x_{1,k}| \,. \end{cases}$$

Then, for all $x_{1,k} \neq -h$ there is a unique solution corresponding to either the continuous (Problem 1) or the discontinuous one with a jump (Problem 2). If $x_{1,k} = -h$, there is a non unique solution $x_{2,k+1} \in [-1, 1]$. The proof is complete. □

A.4 Proof of Proposition 10

Proof. In this proof, we show that the solution of the Backward Euler scheme of the linear DAE in each mode is a solution of (57) if a local AC solution exists to Problem 1. The consistency of the minimal Euler (59) follows from the fact that this foregoing solution associated with a single mode is consistent and bounds the solution of the minimal Euler Scheme.

Case 1. Let us assume that $\mathbf{x}(t_k) = \mathbf{x}_k$ *is in mode* 1*, and that is exists an AC solution in mode* 1. The existence of an AC solution in mode 1 implies $\mathrm{B}_1 + \mathrm{B}_2 \neq 0$, and if $x_{1,k} = 0$, it implies, furthermore, that $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 \leq 0$. Let us consider the following solution of the Backward Euler Scheme for the DAE in mode 1:

$$\begin{cases} x_{1,k+1} = x_{1,k} + h\dfrac{\mathrm{B}_2}{\mathrm{B}_1 + \mathrm{B}_2} \\ x_{2,k+1} = x_{2,k} - h\dfrac{\mathrm{B}_2}{\mathrm{B}_1 + \mathrm{B}_2} \\ z_{k+1} = \dfrac{-1}{\mathrm{B}_1 + \mathrm{B}_2} \end{cases} \tag{117}$$

If $x_{1,k} < 0$, then $x_{1,k+1} < 0$ for sufficiently small h, and $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ is a solution of (57) since the generalised equation reduces to $0 = -1 - x_{1,k+1} - x_{2,k+1}$. If $x_{1,k} = 0$, then $x_{1,k+1} < 0$ for sufficiently small h if $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 < 0$. This latter condition is ensured by the existence of an AC solution that continues strictly in mode 1, and again, $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ is a solution of (57). Finally, if $\mathrm{B}_2 = 0$, $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ provides us with a trivial constant solution of (57).

Case 2. Let us assume that $\mathbf{x}(t_k) = \mathbf{x}_k$ *is in mode* 2*, and that is exists an AC solution in mode* 2. The existence of an AC solution in mode 2 implies $\mathrm{B}_1 - \mathrm{B}_2 \neq 0$, and if $x_{1,k} = 0$, it implies, furthermore, that $(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 \leq 0$. Let us consider the solution of the Backward Euler Scheme applied to the DAE in mode 2:

$$\begin{cases} x_{1,k+1} = x_{1,k} - h\dfrac{\mathrm{B}_2}{\mathrm{B}_1 - \mathrm{B}_2} \\ x_{2,k+1} = x_{2,k} - h\dfrac{\mathrm{B}_2}{\mathrm{B}_1 - \mathrm{B}_2} \\ z_{k+1} = \dfrac{-1}{\mathrm{B}_1 - \mathrm{B}_2} \end{cases} \tag{118}$$

If $x_{1,k} > 0$, then $x_{1,k+1} > 0$ for sufficiently small h, and $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ is a solution of (57). If $x_{1,k} = 0$, then $x_{1,k+1} > 0$ for sufficiently small h if $(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 < 0$. This latter condition is ensured by the existence of a solution that continues strictly in mode 2. In this case, $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ is a solution of (57). Finally, if $\mathrm{B}_2 = 0$, $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ provides us with a trivial constant solution of (57).

Case 3. Let us assume that $\mathbf{x}(t_k) = \mathbf{x}_k$ *is in mode* 3*, and that is exists an AC solution in mode* 3. The existence of an AC solution in mode 3 implies $\mathrm{B}_1 \neq 0$. Furthermore, it implies that $\mathrm{B}_2/\mathrm{B}_1 \leq 0$ if $x_{2,k} = -1$ or $\mathrm{B}_2/\mathrm{B}_1 \geq 0$ if $x_{2,k} = 1$.

Let us consider the following solution of the Backward Euler Scheme for the DAE in mode 3:

$$\begin{cases} x_{1,k+1} = 0 \\ x_{2,k+1} = x_{2,k} - h\dfrac{\mathrm{B}_2}{\mathrm{B}_1} \\ z_{k+1} = \dfrac{-1}{\mathrm{B}_1} \end{cases} \tag{119}$$

If $x_{2,k} \in (-1,1)$, then we have $x_{2,k+1} \in (-1,1)$ for sufficiently small h. The generalised equation in (57) reduces to $x_{1,k+1} = 0$, and then $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ is a solution of (57). If $x_{2,k} = -1$, then we have $x_{2,k+1} > -1$ if $\mathrm{B}_2/\mathrm{B}_1 < 0$, and this is ensured by the existence of an AC solution in mode 3. Again, $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ is a solution of (57). The same reasoning applies for $x_{2,k} = 1$. Finally, if $\mathrm{B}_2 = 0$, $(x_{1,k+1}, x_{2,k+1}, z_{k+1})$ provides us with a trivial constant solution of (57).

Summary In all the cases, we found a solution of the Backward Euler Scheme, such that $\bar{z}_{k+1}$ is bounded provided that there exists a local AC solution starting from $x(t_k)$. We can then conclude that

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| = h\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} + Bz_{k+1}\| = \mathcal{O}(h) \tag{120}$$

and from

$$\|\mathbf{x}_{k+1} - \mathbf{x}(t_k + h)\| \leq \|\mathbf{x}_{k+1} - \bar{\mathbf{x}}_k\| + \|\mathbf{x}_k - \mathbf{x}(t_k + h)\| \tag{121}$$

that $\lim_{h\to 0} \|\mathbf{x}_{k+1} - \mathbf{x}(t_k)\| = 0$ if the solution is AC. To conclude the proof, we note that

$$\|\underline{\mathbf{x}}_{k+1} - \mathbf{x}(t_k + h)\| \leq \|\mathbf{x}_{k+1} - \mathbf{x}(t_k + h)\| \tag{122}$$

by the definition of $\underline{\mathbf{x}}_{k+1}$ as the minimal solution in (59). □