Index-2 hybrid DAE: a case study with well-posedness and numerical analysis

Alexandre Rocca, Vincent Acary, Bernard Brogliato

Research Report

Index-2 hybrid DAE: a case study with well-posedness and numerical analysis

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

(1) INRIA Grenoble-Alpes
655 Avenue de L'Europe 38330 Monbonnot, France

Abstract. In this work, we study differential algebraic equations with constraints defined in a piece-wise 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 are often subject to chattering at mode switch in presence of sliding modes, and can result in Zeno behaviours.
In contrast, the representation of such systems using differential inclusions and method from non-smooth dynamics are often closer to the physical theory but may be harder to interpret. Associated time-stepping 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 non-smooth methods to the simulation of piece-wise linear ODEs, we want to apply non-smooth numerical scheme to piece-wise linear DAEs. In particular, the study of a 2-D dynamical system of index-2 with a switching constraint using set-valued operators, is presented.

Keywords: Hybrid Systems, ODE, DAE, Nonsmooth Dynamical Systems, Linear Complementarity Problem

1 Introduction

The aim of this work is to study hybrid differential algebraic equations (hybrid DAE) meaning dynamical systems with some algebraic constraints switching with respect to the state variables. Such hybrid DAE systems are used in numerous field from electronics, [1] to chemical process engineering, [24]. They are especially used in model-based design through the use of language like Modelica as in [14].

Various works already studied the field of hybrid DAE. For example, DAE including complementarity constraints are a subset of the more general class system which is the differential variational inequalities (DVI). DVI are defined

and studied in [22]. In particular, they analyse the well-posedness of index one DVI and DAE of mixed index between 1 and 2 respectively in Sections 5 and 8.

Outside of the formalisation as variationnal inequalities, or complementarity problems, I.V Matrosov [19] proposes a concept of solutions, which is inspired by the one of A.F. Filippov [12], for DAE with discontinuous constraints and differential part. Then, he defines a concept of solutions by the differential inclusion into the convex hull of the set of the vector fields that are solution in $(t, \mathbf{x}, \mathbf{z})$ at the right and left limit of the discontinuity, assuming the solution of the algebraic variable $\mathbf{z}(t)$ in known. He gives sufficient conditions for existence of such solutions in [19], and sufficient condition for uniqueness in [20].

V. Merhmann et al. [13,21] provide a study of well-posedness of hybrid DAE structured as a hybrid automata. In addition, a numerical implementation of sliding modes for DAE systems is provided to avoid chattering when switching. It is interesting to note that the sliding solutions obtained in [21] are similar to the solutions from the previously introduced work of [19], assuming the solution of $\mathbf{z}(t)$ in each mode is obtained by index reduction. Furthermore, the work of [21] need explicit transition functions from one mode (DAE) to another in addition to consistent reset conditions.

S. Trenn [25] defines solutions of hybrid DAE with exogenous switching. In particular, he introduces the notion of distributional solutions which can also be used to efficiently solve inconsistent initial conditions of classical DAE as an exogenous switching at $t = 0$.

K. Camlibel et al. [10] extend results of well-posedness of differential inclusions to differential algebraic inclusions $\mathrm{P}\dot{\mathbf{x}} \in -\mathcal{F}(\mathbf{x})$ with a maximal monotone operator $\mathcal{F}(\cdot)$. Then, assuming the passivity of the Weierstrass-Kronecker form of a system (1) with $\mathcal{M}(\cdot)$ a maximal monotone operator, sufficient condition for the well-posedness of absolutely continuous (AC) solutions of (1) are given.

$$\begin{cases} \mathrm{E}\dot{\mathbf{y}}(t) = \mathrm{A}\mathbf{y}(t) + \mathrm{B}\boldsymbol{\lambda}(t) + \mathbf{b} \\ \mathbf{w}(t) = \mathrm{C}\mathbf{y}(t) + \mathrm{D}\boldsymbol{\lambda}(t) + \mathbf{q} \\ \mathbf{w}(t) \in \mathcal{M}(-\boldsymbol{\lambda}(t)). \end{cases} \tag{1}$$

It important to note this formalisation is the one which is the closest to the example studied in this paper (see (10)), with the notable difference that in our case the operator $\mathcal{M}(\cdot)$ is not maximal monotone but hypo-monotone. Additionally, its rewriting in the form of (1) with maximal monotone operator is not a passive system.

P. Stechlinski et al. [24,6] and K. Khan [15] define from the Clarke jacobian a notion of generalised differential index and an associated index reduction procedure in the context of non-smooth DAE with at least Lipschitz continuous constraints. Current implementation and theory are limited to semi-explicit index-1 non-smooth DAE.

Finally, let us cite another work for index reduction of hybrid DAE based on non standard analysis by A. 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 presence of varying index. In particular, this work

pairs well with [21], which needs the knowledge of transition and re-initialisation maps when switching from one mode to another.

Let us now define the general framework of linear hybrid DAE that we wish to study from the point of view of non-smooth dynamics, and event capturing methods. We wish to consider hybrid linear DAE defined as:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathrm{A}\mathbf{x}(t) + \mathrm{B}\mathbf{z}(t) + \mathbf{b} \\ \quad 0 = \mathbf{g}_i(\mathbf{x}(t), \mathbf{z}(t)) = \mathrm{C}_i\mathbf{x}(t) + \mathrm{D}_i\mathbf{z}(t) + \mathbf{q}_i \\ \qquad \forall (\mathbf{x}(t), \mathbf{z}(t)) \in \mathcal{X}_i. \end{cases} \tag{2}$$

The finite number of sets $\mathcal{X}_i = \{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}^{n_1+n_2=n} \mid \mathbf{h}_i(\mathbf{x}, \mathbf{z}) = \mathrm{H}_i\mathbf{x}(t) + \mathrm{F}_i\mathbf{z}(t) + \mathbf{p}_i > 0\}$ 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 $i \neq j$, $\partial\mathcal{X}_i \cap \partial\mathcal{X}_j = \emptyset$

where, $\mathbf{x}$, $\mathbf{z}$ are respectively the differential and algebraic variables. we can build using step-functions[1] in a similar fashion to [3] a generalised constraint.

$$\begin{aligned} \mathbf{g}(\mathbf{x}, \mathbf{z}) &= \sum_i \left(\prod_{j \neq i} (1 - \mathrm{s}^+(\mathbf{h}_j(\mathbf{x}, \mathbf{z}))) \right) \mathrm{s}^+(\mathbf{h}_i(\mathbf{x}, \mathbf{z})) \mathbf{g}_i(\mathbf{x}, \mathbf{z}) \\ &= 0 \end{aligned} \tag{3}$$

where $\mathrm{s}^+(\mathbf{y}) = 0$ if $\mathbf{y} < 0$ and $\mathrm{s}^+(\mathbf{y}) = 1$ if $\mathbf{y} > 0$, the behaviour in $\mathbf{y} = 0$ depending on later relaxations. In particular, in the context of piecewise ODE, the work of [3] shows that methods for non-smooth dynamics can be efficiently applied using such translation. Then, depending of the concept of solutions applied on the switching surfaces (using convexification as in [12], 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.

In this work, and its associated working example, we restrain ourselves to the simple case with only two algebraic constraints $\mathrm{C}_1\mathbf{x}+\mathrm{D}_1\mathbf{z}+\mathbf{q}_1$ and $\mathrm{C}_2\mathbf{x}+\mathrm{D}_2\mathbf{z}+\mathbf{q}_1$ and one switching condition depending on the sign of $\mathbf{h}(\mathbf{x}, \mathbf{z}) = \mathrm{H}\mathbf{x}(t) + \mathrm{F}\mathbf{z}(t)$. Then, we construct a relaxation of these two constraints along the switching surface $\mathbf{h}(\mathbf{x}, \mathbf{z}) = 0$ by "filling the graph" (see [18]).

We can construct such relaxed constraint in $\mathbf{h}(\mathbf{x}, \mathbf{z}) = 0$ 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 (3) in a similar fashion to [5] approach for discontinuous ODEs. For this working example, we consider the convexification of the constraints along the switching surface.

As we have seen, most works consider either an high index hybrid DAE framework with event-driven numerical methods and explicit transition functions, or

[1] or using sign functions

mainly index-1 DAE with non-smooth constraints aiming to rewrite the system as a differential inclusion[2] into a Liptchitz function, or a maximal monotone operator. In this paper, we are looking to make a bridge between the hybrid DAE formalism and the non-smooth DAE formalism. With this in mind, we show on a simple working example the difficulties arising with such relaxations, as well as how classical non-smooth numerical methods performs in this context. We also propose some modification to the numerical scheme to overcome the troubles observed in this context.

2 Preliminaries

Let first introduce some notations we will use through this paper. Vectors of real variables $\mathbf{x} = (x_1, \ldots, x_i, \ldots, x_n)$ 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 expression, we will in general note $\boldsymbol{\lambda}$ 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:

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

Let us now introduce important definitions for the following of the paper. First, let us define the absolute continuity:

Definition 1 (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 f' a.e. (almost everywhere), the derivative is Lebesgue integrable, and*

$$f(t) = f(t_1) + \int_{t_1}^{t} f'(\tau) d\tau\,,$$

$\forall t\ in [t_1, t_2]$.

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

Definition 2 (Normal cone). *Let $K \subseteq \mathbb{R}^n$ be a closed convex set. The normal cone to K at $\mathbf{x} \in K$ is the set:*

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

Definition 3 (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 $\partial f(\mathbf{x})$.

[2] differential algebraic inclusion in the case of [10]

Definition 4 (Indicator function). *The indicator function of a set $K \subseteq \mathbb{R}^n$, $\psi_K : \mathbb{R}^n \to \overline{\mathbb{R}}$, 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 $\psi_K(\cdot)$ is convex lower semi-continuous and $\partial\psi_K = \mathcal{N}_K(\cdot)$.

Definition 5 (Set-valued sign operator). *Let us define in this paper the* sign *function as a set-valued operator, sign $: \mathbb{R} \rightrightarrows \mathbb{R}$, such that:*

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

Finally, let us introduce important definitions concerning the differential algebraic equations considered in each of the modes.

Definition 6 (Linear Constant Coefficient Differential Algebraic Equation). *We define a linear constant coefficient DAE as:*

$$\mathrm{E}\dot{\mathbf{y}}(t) + \mathrm{F}\mathbf{y}(t) = p(t)\,, \tag{4}$$

with $\mathbf{y} \in \mathbb{R}^{n_y}$. If E *is non-singular we retrieve a linear ODE.*

In this paper we consider semi-explicit linear DAE of the form (2) in each mode: they are a particular case of linear constant coefficient DAE (also noted LCC-DAE) from Definition 6 where $E = (I_{n_1}, 0_{n_2})^{\mathrm{T}}$, $F = -\begin{pmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C}_i & \mathrm{D}_i \end{pmatrix}$, and $p = (e, q_i)^{\mathrm{T}}$.

Definition 7 (Regular Matrix Pencil). *Given an ordered pair of matrix $\{\mathrm{E}, \mathrm{F}\}$ with $\mathrm{E}, \mathrm{F} \in \mathbb{R}^{n\times n}$, the matrix pencil $\lambda\mathrm{E} + \mathrm{F}$ is said to be regular if there exists $\lambda \in \mathbb{R}$ such that $\det(\lambda\mathrm{E} + \mathrm{F}) \neq 0$.*

If the matrix pencil associated with a LCC-DAE is regular we say that the DAE is regular. Otherwise, we say the LCC-DAE is non-regular and there is either none or infinitely many solutions [16]. From now on, let consider the regular case (see proofs in [17][Chapter 1]).

Definition 8 (Canonical Weierstrass-Kronecker form). *If a matrix pencil associated with a pair of matrix $\{\mathrm{E},\mathrm{F}\}$ is regular, then there exist nonsingular matrices $\mathrm{L}, \mathrm{K} \in \mathbb{R}^{n\times n}$, and natural integers $l \leq n$, $\mu \leq l$ such that:*

$$\mathrm{LEK} = \begin{pmatrix} I & 0 \\ 0 & N \end{pmatrix}, \qquad \mathrm{LFK} = \begin{pmatrix} W & 0 \\ 0 & I \end{pmatrix} \tag{5}$$

where $N \in \mathbb{R}^{l\times l}$, $W \in \mathbb{R}^{(n-l)\times(n-l)}$, *and* N *is nilpotent of order* μ *(that is,* $N^{\mu}=0$ *and* $N^{\mu-1}\neq 0$*). Then, noting* $\mathbf{y}=\mathrm{K}(\hat{\mathbf{x}},\hat{\mathbf{z}})^{\mathrm{T}}$, *system* (4) *transforms into the Weierstrass-Kronecker form:*

$$\begin{cases} \dot{\hat{\mathbf{x}}}(t)+W\hat{\mathbf{x}}(t)=(\mathrm{L}p(t))|_{n-l}=r_1(t) \\ N\dot{\hat{\mathbf{z}}}(t)+\hat{\mathbf{z}}(t)=(\mathrm{L}p(t))|_{l}=r_2(t) \end{cases} \tag{6}$$

Proposition 1 (Solutions of LCC-DAE). *A unique solution* $\mathbf{y}(t)$ *of* (4) *is given by the solution of* (6) *through* $\mathbf{y}=\mathrm{K}(\hat{\mathbf{x}},\hat{\mathbf{z}})^{\mathrm{T}}$. *Indeed,* $\hat{\mathbf{x}}(t)$ *is simply the solution of a linear ODE, while we can observe that:*

$$\hat{\mathbf{z}}(t)=r_2(t)-N\dot{\hat{\mathbf{z}}}(t)=r_2(t)-N\frac{\mathrm{d}}{\mathrm{dt}}\left(r_2(t)-N\dot{\hat{\mathbf{z}}}(t)\right)=...$$

yielding by nilpotency of N:

$$\hat{\mathbf{z}}(t)=\sum_{i=0}^{\mu-1}(-1)^i N^i\frac{\mathrm{d}^i r_2}{\mathrm{dt}^i}(t).$$

One notes that the perturbation term $r_2(t)$ *has to be sufficiently smooth to define a solution in this way. One also notes that the nilpotency degree* μ *is equal to the number of successive differentiations to obtain an ODE system.*

Finally, the initial condition $\mathbf{y}(0)$ *of the original system is defined by the initial conditions of the successive derivatives of the perturbation term.*

In the particular case of semi-explicit linear DAE (as in (2) when considering one mode only) if the matrix D is non-singular, then one can write by differentiating the constraint once:

$$\dot{\mathbf{z}}(t)=\mathrm{D}^{-1}(\mathrm{CA}\mathbf{x}(t)+\mathrm{B}\mathbf{z}(t))\,,$$

and obtain an ODE with a unique solution with respect to the initial condition. Similarly, assuming consistent initial conditions, if D is singular (without loss of generality we can assume $\mathrm{D}=0$), one can derive twice the algebraic constraint and obtain, if CB non-singular:

$$\dot{\mathbf{z}}(t)=(\mathrm{CB})^{-1}(\mathrm{CA}^2\mathbf{x}(t)+\mathrm{CAB}\mathbf{z}(t)).$$

The number of differentiations in this process is given by the differentiation index of a DAE.

Definition 9 (Differentiation index). *The differentiation index is defined as the number of times the DAE must be derived to yield an ODE*[3] *(without eliminating* $\mathbf{z}(t)$*). The differentiation index is equal to the nilpotency degree* μ.

[3] Assuming the DAE only have first order derivatives.

Definition 10 (Consistency[4] of a numerical method). *Given a numerical method of the form:*

$$\mathbf{y}_{k+1} = \mathbf{y}_k + h_k \mathbf{f}(t_k, \mathbf{y}_k, h_k) \text{ for all } k \geq 0,\ t_0 = 0,\ t_{k+1} = t_k + h_k \tag{7}$$

This numerical method is said to be consistent for a DAE (4) *with* $\mathbf{y}(0) = \mathbf{y}_0$*, if for any solution of this DAE the consistency error*

$$\sum_{k=0}^{N-1} \|\mathbf{y}(t_{k+1}) - \mathbf{y}(t_k) - h_k \mathbf{f}(t_k.\mathbf{y}(t_k), h_k)\| \tag{8}$$

tends to 0 *when* $h = \max_{0 \leq k \leq N} h_k$ *tends to* 0.

3 Analysis of a Hybrid DAE Example

Let us consider the following switching DAE:

$$\begin{cases} \dot{x}_1(t) = 1 + \mathrm{B}_1 z(t) \\ \dot{x}_2(t) = \mathrm{B}_2 z(t) \\ \texttt{if } x_1 > 0 \texttt{ then } : \\ \qquad 0 = 1 + x_1(t) - x_2(t) \\ \texttt{if } x_1 < 0 \texttt{ then } : \\ \qquad 0 = -1 - x_1(t) - x_2(t) \end{cases} \tag{9}$$

which is a particular case of (2) with $\mathrm{A} = 0$, $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$, $\mathcal{X}_1 = \mathbb{R}^-$, $\mathcal{X}_2 = \mathbb{R}^+$, $C_1 = (1, -1)$, $C_2 = (-1, -1)$, $H_1 = (-1, 0)$, $H_2 = (1, 0)$, $\mathrm{D}_1 = 0$, and $\mathrm{D}_2 = 0$. In $x_1 = 0$ the system does not have continuous solutions whatever is the active constraint so we keep strict inequalities in (9).

As exposed in the introduction, 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 hybrid DAE system (10) below where the constraint is $0 \in -x_2 + \lambda(x_1 + 1)$ with $\lambda = \mathrm{sign}(x_1)$. This set-valued algebraic constraint 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 $x_2 \in \mathrm{convexHull}(\{x_2 = -1\}, \{x_2 = 1\}) = [-1, 1]$ (see Figure 1). This yields the non-smooth DAE system:

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

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

with the initial conditions $\mathbf{x}(t_0) = \mathbf{x}_0$.

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

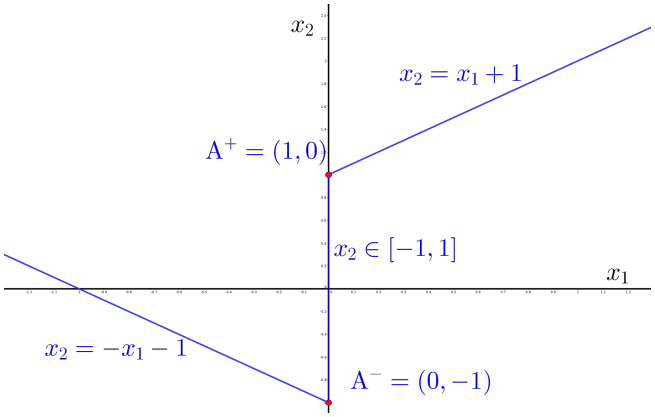


Fig. 1: Phase-space representation of the constraint of (9) and (10)

3.1 Analysis of AC solutions

Let us study the conditions on the differential part of the DAE (10), and in particular B_1, B_2 for the existence of a sliding mode along the switching surface $x_1 = 0$, and more generally the existence of AC solution $x_1(t)$, $x_2(t)$ to (10), for some arbitrary time interval and initial conditions.

Let us first observe that solutions $(x_1(t), x_2(t))$ of the non-smooth constraint of system (10) are such that:

$$\begin{aligned} &\exists \lambda(t) \in \mathrm{sign}(x_1(t)) \text{ s.t. :} \\ &\quad -x_2(t) + \lambda(t) x_1(t) + \lambda(t) = 0\,, \end{aligned} \tag{11}$$

and that the constraint of system (10) can be rewritten equivalently as the following set-valued constraint:

$$\begin{aligned} &0 \in -x_2(t) + |x_1(t)| + \mathrm{sign}(x_1(t)) \\ \Leftrightarrow\ &x_1(t) \in \mathcal{N}_{[-1,1]}(-x_2(t) + |x_1(t)|)\ , \end{aligned} \tag{11'}$$

where the equivalence is obtained by using the inversion of subdifferentials of convex lower-semicontinuous functions [23]. Thus, the original switching DAE in (9), which is embedded in (2), is recast in a new formalism:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathrm{A}\mathbf{x}(t) + \mathrm{B}\mathbf{z}(t) + \mathbf{b} \\ \quad 0 \in \mathcal{F}(\mathbf{x})\,, \end{cases} \tag{12}$$

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

Definition 11 (Global AC solutions). *We say there exists global AC solutions if for some initial conditions* $\mathbf{x}(t_0)$ *satisfying* (11)*, there exists a solution to* (10a) *almost everywhere and to* (10b) *for all* $t \in [t_0, T]$*, and any* $T > t_0$*.*

Let us state the main result of this section.

Proposition 2 (AC solutions). *The system defined in* (10) *has global AC solutions on an arbitrary interval* $[t_0, T[$, $T > t_0$ *for any consistent initial condition* $\mathbf{x}(0)$*, if and only if* B_1, B_2 *are chosen such that:*

$$\begin{cases} \mathrm{B}_1 \neq 0 \\ \dfrac{\mathrm{B}_2}{\mathrm{B}_1} \leq 0 \\ (\mathrm{B}_1 + \mathrm{B}_2) \neq 0. \end{cases} \tag{13}$$

Proof. Let us first consider the conditions of existence of solutions in each mode: mode 1 ($x_1 < 0$), mode 2 ($x_1 > 0$), mode 3 ($x_1 = 0$). Then, we consider the conditions where solutions can switch from one mode to another.

1. Assume $x_1(t') < 0$. We first consider solutions in mode 1: $x_1(t) < 0$, $\lambda(t) = -1$ for all $t \in [t', t' + \varepsilon] \triangleq I_\varepsilon$, with $\varepsilon > 0$.
   Therefore, for all $t \in I_\varepsilon$ such that $x_1(t) < 0$, solutions of (10) must also satisfy:
   $$\begin{aligned} -\dot{x}_2(t) - \dot{x}_1(t) = 0 &\Leftrightarrow (\mathrm{B}_2 + \mathrm{B}_1) z(t) + 1 = 0 \\ &\Leftrightarrow z(t) = \frac{-1}{\mathrm{B}_2 + \mathrm{B}_1}\,, \end{aligned}$$
   which has a solution if and only if $(\mathrm{B}_2 + \mathrm{B}_1) \neq 0$. We deduce that there exist AC solutions in mode 1 if and only if $(\mathrm{B}_2 + \mathrm{B}_1) \neq 0$, and we say that mode 1 is not feasible if $(\mathrm{B}_2 + \mathrm{B}_1) = 0$. We note (see (9) and (10)) that this corresponds to $\mathrm{C}_1\mathrm{B}$ non-singular with $\mathrm{C}_1 = (-1, -1)$. Furthermore,
   $$\dot{x}_1(t) = \frac{\mathrm{B}_2}{\mathrm{B}_2 + \mathrm{B}_1}\,.$$
   We deduce that either $\dot{x}_1(t) \leq 0$ and $x_1(t) < 0$ for all $t > t'$, or $\dot{x}_1(t) > 0$ and there exists $t_1 > t'$ such that $x_1(t_1) = 0$:
   - If $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 \leq 0$ then $\dot{x}_1(t) \leq 0$, for all $t \in [t', +\infty[\triangleq I_\infty$: the system stays in mode 1 forever.
   - If $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 > 0$ then $\dot{x}_1(t) > 0$, for $t > t'$. Moreover, for $t_1 = -x_1(t')\frac{\mathrm{B}_1 + \mathrm{B}_2}{\mathrm{B}_2} + t'$ we have $\mathbf{x}(t_1) = (0, -1)$ and the trajectory reaches mode 3.

   We can already provide a necessary[5] condition on $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$ for the existence of global AC solutions in mode 1:
   $$\mathrm{B}_1 + \mathrm{B}_2 \neq 0 \qquad \text{(Cond } \mathrm{A}_1\text{)}$$

[5] Note this is a sufficient and necessary condition for existence of local solution in mode 1, however this is only a necessary condition for global solutions of the hybrid DAE (10).

2. Similarly, assume $x_1(t') > 0$. Let us consider solutions in mode 2: $x_1(t) > 0$, $\lambda(t) = 1$ for all $t \in I_\varepsilon$.
   Hence, for all t such that $x_1(t) > 0$, solutions of (10) must also satisfy:

$$-\dot{x}_2(t) + \dot{x}_1(t) = 0 \Leftrightarrow -\mathrm{B}_2 z(t) + \mathrm{B}_1 z(t) + 1 = 0$$
$$\Leftrightarrow z(t) = \frac{-1}{\mathrm{B}_1 - \mathrm{B}_2}\,,$$

   and this mode is feasible if and only if $(\mathrm{B}_1 - \mathrm{B}_2) \neq 0$. Similarly to the previous case, this correspond to $\mathrm{C}_2\mathrm{B}$ non-singular with $\mathrm{C}_2 = (1, -1)$. Moreover,

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

   It follows that either $\dot{x}_1(t) \geq 0$ and $x_1(t) > 0$ for all $t > t'$, or $\dot{x}_1(t) < 0$ and there exists $t_1 > t'$ such that $x_1(t_1) = 0$:
   - If $(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 \leq 0$ then $\dot{x}_1(t) \geq 0$, for all $t \in \triangleq I_\infty$: the systems stays in mode 2 forever.
   - If $(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 > 0$ then $\dot{x}_1(t) < 0$, for $t > t'$. Furthermore, for $t_1 = x_1(t')\frac{\mathrm{B}_1 - \mathrm{B}_2}{\mathrm{B}_2} + t'$ we have $\mathbf{x}(t_1) = (0, 1)$ and the trajectory reaches the mode 3.

   Thus, another necessary condition for the existence of global AC solutions is:

$$\mathrm{B}_1 - \mathrm{B}_2 \neq 0\,. \qquad \text{(Cond A}_2\text{)}$$

   In summary, cases 1 and 2 show that solutions for initial conditions with $x_1(t') \neq 0$, exist either for all $t \in [t', +\infty[$, or attain $x_1(t_1) = 0$ (mode 3) for some finite $t_1 > t'$. Let us study what happens when the solution is in mode 3 for some $t \in [t', t' + \varepsilon[$. We recall that mode 3 is obtained by the convexification of the two constraints in (9) in $x_1 = 0$.
3. Assume there exists $t' \in \mathbb{R}$ such that $x_1(t') = 0$. For all $t \in I_\varepsilon$ such that $x_1(t) = 0$, by (10) the solution in mode 3 satisfies:

$$\begin{cases} \dot{x}_1(t) = 0 \Leftrightarrow \mathrm{B}_1 z(t) = -1 \\ \dot{x}_2(t) = \dfrac{-\mathrm{B}_2}{\mathrm{B}_1}\,. \end{cases} \tag{14}$$

   In addition, the constraint in (11) becomes:

$$x_2(t) = \lambda(t) \in [-1, 1]\,.$$

   In a similar way to modes 1 and 2, there exists a local solution in mode 3 if and only if:

$$\mathrm{B}_1 \neq 0 \qquad \text{(Cond A}_3\text{)}$$

   This corresponds to $\mathrm{C}_3\mathrm{B}$ non-singular 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 (14) the solutions staying in mode 3 are either:
- Constant if $\mathrm{B}_2 = 0$, and they stay in mode 3 for all $t \in [t', +\infty[$.
- Going 'downward' if $\mathrm{B}_2/\mathrm{B}_1 > 0$ such that $\dot{x}_2 < 0$. Then, there exists $t_1 \geq t'$ given by $t_1 = \frac{\mathrm{B}_1}{\mathrm{B}_2}(1+x_2(t'))+t'$ 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 $t_1 \geq t'$ given by $t_1 = \frac{\mathrm{B}_1}{\mathrm{B}_2}(x_2(t') - 1) + t'$ 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.

In cases 1, 2, and 3 we have studied the conditions for the existence of solutions in the modes 1, 2, 3 (respectively (Cond A_1), (Cond A_2), and (Cond A_3)). Furthermore, necessary conditions on B_1, B_2 for existence of global AC solution for solutions staying in any of this three modes are given by the intersection of these three conditions:

$$\begin{aligned} \mathrm{B}_1 + \mathrm{B}_2 &\neq 0 \\ \mathrm{B}_1 - \mathrm{B}_2 &\neq 0 \\ \mathrm{B}_1 &\neq 0 \end{aligned} \qquad \text{(Cond A)}$$

4. We now consider the conditions for existence of solutions such that there exists a continuous switching from one mode to another.

(4.1) Assume $x_1(t') < 0$: in case 1 we have seen there may exist t_1 such that the left limit $\mathbf{x}(t_1^-)$ is in mode 1 and $\mathbf{x}(t_1) = (0, -1)$ (point A^- in Figure 2) is in mode 3. Consequently, the condition for such continuous switching from mode 1 to mode 3 is $\mathrm{B}_1 \neq 0$ and $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 > 0$. This condition is already satisfied in (Cond A).

(4.2) Assume $x_1(t') > 0$: in case 2 we have seen there may exist t_1 such that the left limit $\mathbf{x}(t_1^-)$ is in mode 2, and $\mathbf{x}(t_1) = (0, 1)$ (point A^+ in Figure 2) is in mode 3. Consequently, the condition for continuous switching from mode 2 to mode 3 is $\mathrm{B}_1 \neq 0$ and $(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 > 0$. This condition is already satisfied in (Cond A).

Remark 1. In cases (4.1) and (4.2) we only discuss the continuity of the state variable $\mathbf{x}(t)$. The continuity of $\lambda(t) \in \mathrm{sign}(x_1(t))$ is given by the one of $x_2(t)$ as seen in mode 3. In addition, by construction from case 1 (respectively cases 2 and 3), we note that the algebraic variable $z(t)$ and $\dot{\mathbf{x}}(t)$ are right-continuous at t_1 since $z(t_1^-) = \frac{-1}{\mathrm{B}_2+\mathrm{B}_1}$ in mode 1 (respectively $\frac{-1}{\mathrm{B}_1-\mathrm{B}_2}$ in mode 2, and $z(t_1) = \frac{-1}{\mathrm{B}_1}$ in mode 3).

(4.3) Assume $x_1(t') = 0$ and $\mathrm{B}_2/\mathrm{B}_1 > 0$: then from case 3 there exists some time $t_1 > t'$ such that $x_2(t_1) = -1$. Furthermore, we need $\dot{x}_1(t_1^+) < 0$ (meaning that $\dot{x}_1(\cdot)$ is left continuous) to switch to mode 1: this is equivalent to the condition $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 < 0$.

However, there is no $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$ such that $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 < 0$ and $\mathrm{B}_2/\mathrm{B}_1 > 0$. When $\mathrm{B}_2/\mathrm{B}_1 > 0$, solutions cannot continue further than point A^-. Consequently, the condition (Cond A) must be further restricted to:

$$\begin{aligned} \mathrm{B}_1 + \mathrm{B}_2 &\neq 0 \\ \frac{\mathrm{B}_2}{\mathrm{B}_1} &\leq 0 \\ \mathrm{B}_1 &\neq 0\,. \end{aligned} \qquad \text{(Cond B)}$$

We note that $\frac{\mathrm{B}_2}{\mathrm{B}_1} \leq 0$ implies $\mathrm{B}_1 - \mathrm{B}_2 \neq 0$.

(4.4) Assume $x_1(t_1) = 0$ and $\mathrm{B}_2/\mathrm{B}_1 < 0$, then from case 3 there exists some time t_2 such that $x_2(t_2) = 1$. Furthermore, it is necessary that $\dot{x}_1(t_2^+) > 0$ (meaning that $\dot{x}_1(\cdot)$ is left-continuous in t_2) to switch to mode 2: this is equivalent to the condition $(\mathrm{B}_1 - \mathrm{B}_2)\mathrm{B}_2 < 0$. Finally, all $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$ satisfying $\mathrm{B}_2/\mathrm{B}_1 < 0$ also satisfy this condition and solutions can be continued in mode 2. Consequently, the conditions of (Cond B) are still satisfied and unchanged.

From this study of the conditions for each case we retrieve the condition (Cond B) given in (13) for existence of global AC solution for any initial conditions satisfying the constraint .

□

Remark 2. The above study does not prove the uniqueness of solutions. Indeed, if $\mathbf{x}(t_1) = (0, -1)$ there also exists an AC solution switching to mode 1 with $x_1(t_1^+) < 0$ and $\dot{x}_1(t_1^+) < 0$ as the set of vectors B such that $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 < 0$ and $\mathrm{B}_2/\mathrm{B}_1 \leq 0$ is not empty.

In Figure 2, we summarise the conditions of Proposition 2 for existence of global AC solutions for any initial conditions satisfying (11). The set of B_1, B_2 that satisfy these conditions can be separated in two subsets: the subset yielding "sliding-crossing" solutions which are unique with respect to the initial conditions, and the subset leading to "sliding-repulsive" solutions, which are not unique in $\mathbf{x}(t_0) = (0, -1)^{\mathrm{T}}$. These solutions are illustrated in Figure 2a, and the conditions on the parameters $(\mathrm{B}_1, \mathrm{B}_2)$ in Figure 2b.

3.2 Analysis of the index-reduced system's Filippov solutions.

Let us compare the sliding mode obtained in the proof of Proposition 2 with the Filippov solutions [12] in $x_1 = 0$ of the switching ODE obtained by index reduction of the DAE in $x_1 < 0$ and $x_1 > 0$. The construction of such solutions is defined, in [19] and [21]. The left vector field $\mathbf{f}^-(\cdot)$ (for $x_1 < 0$) and right vector field $\mathbf{f}^+(\cdot)$ (for $x_1 > 0$) are given by:

$$\begin{aligned} \mathbf{f}^-(t, \mathbf{x}(t)) &= \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\mathrm{B}_2}{\mathrm{B}_1 + \mathrm{B}_2} \\ \frac{-\mathrm{B}_2}{\mathrm{B}_1 + \mathrm{B}_2} \end{pmatrix} \end{aligned} \qquad (15)$$

$$\begin{aligned} \mathbf{f}^+(t, \mathbf{x}(t)) &= \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} \\ &= \begin{pmatrix} \frac{-\mathrm{B}_2}{\mathrm{B}_1 - \mathrm{B}_2} \\ \frac{-\mathrm{B}_2}{\mathrm{B}_1 - \mathrm{B}_2} \end{pmatrix} \end{aligned} \qquad (16)$$

Let us observe from the analysis in cases 1 and 2 that $\dot{z} = 0$ on both sides of the switching surface: only its re-initialisation changes after switching.

In the "sliding-repulsive" solutions case, 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 3.1. Indeed, in every point of the switching surface $x_1 = 0$, there exists a sliding solution associated with the vector field $f_0(\mathbf{x}) = \text{convexHull}(f^-(\mathbf{x}), f^+(\mathbf{x})) \cap \{\mathbf{x} \in \mathbb{R}^2 | x_1 = 0\}$. It follows that the solutions obtained by either [19] or [21] correspond to the same solutions we obtain by our relaxation of the switching constraint by a generalised equation. A particular example with B=(-1,0.5) is shown in Figure 3b.

In the particular case of "sliding-crossing" solutions, the index reduced system does not lead to any sliding motions as $\text{convexHull}(f^-(\mathbf{x}), f^+(\mathbf{x}))$ 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 the sense of [19] there is no absolutely continuous solutions. In contrast, following the guidelines for sliding mode detection given by [21] we need an explicit transition function to continue. The sliding solution we obtain thanks to our relaxation is not retrieved by these concept of solutions. A particular example with $\mathrm{B} = (-0.5, 1)$ is shown in Figure 3a.

We note that the approach of convexifying the left and right reduced DAEs has already been shown wrong for some cases in [20] with the index-1 non-smooth example:

$$\begin{aligned} \dot{x}(t) &= -\operatorname{sign}(x(t)) \\ y(t) &= |x(t)| \end{aligned}$$

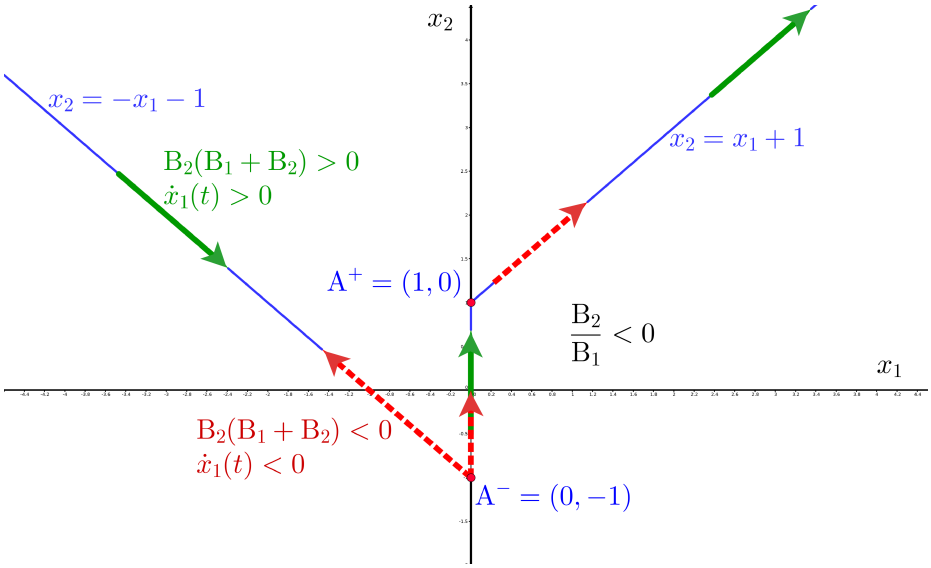
Although the equation $\dot{x}(t) = -\operatorname{sign}(x(t))$ has Filippov solutions, the adding of an output leads to some incoherence. Indeed, in $x(t) = 0$ the convexification of the left and right vector fields solutions would yield $(\dot{x}, \dot{y}) \in \{[-1, 1] \times \{-1\}\}$ where the solution $y(t \geq 1) = 0$ is not represented. However, an approach in the sense of A.P. [5] would gives $(\dot{x}, \dot{y}) \in \{[-1, 1] \times [-1, 0]\}$ which contains the correct solution $y(t \geq 1) = 0$.

3.3 Analysis of solutions with bounded discontinuities.

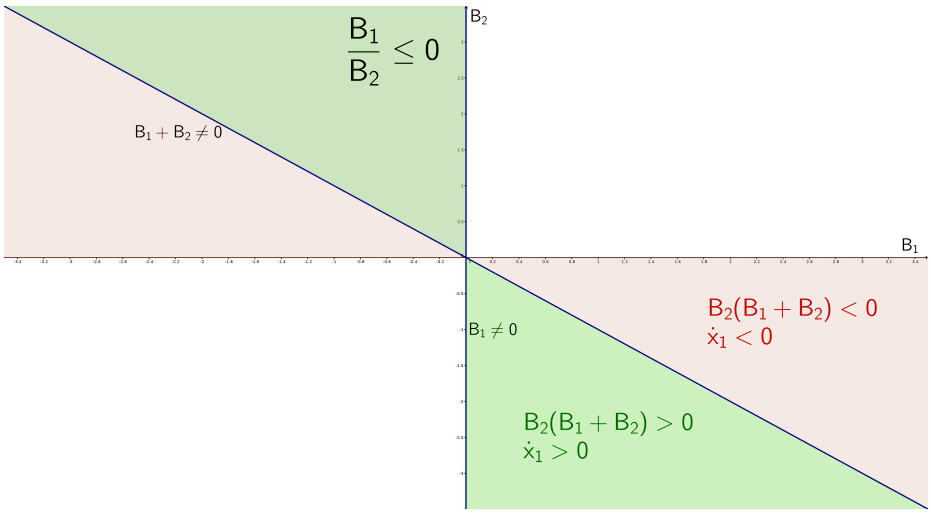
Let us study the existence of discontinuous solutions when no continuation with an AC solution exists after some time t_j. For example, this is the case if the trajectory reaches the point A^- in Figure 2 with $\mathrm{B}_2/\mathrm{B}_1 > 0$. Solutions with discontinuities make sense in a context where quite different time-scales exist in the same system. In particular, this may be found in mechanical systems with impacts [9], or in circuit with ideal diodes or set-valued electronic components [1].

3.3.1 Analysis of jump dynamics

Assume there is a state jump at some time t_j. We first introduce the measure differential inclusions (MDI) (17) associated with (10) (see [18]). Let us note $\mathrm{d}\mathbf{x}$

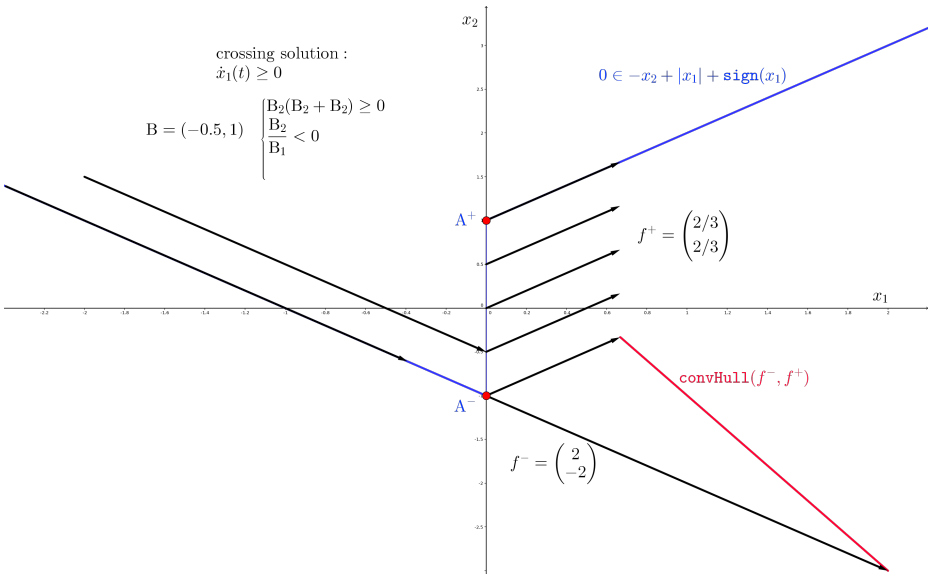


(a) In red (dashed line) the "sliding-repulsive" solutions, and in green (full line) the "sliding-crossing"solutions.

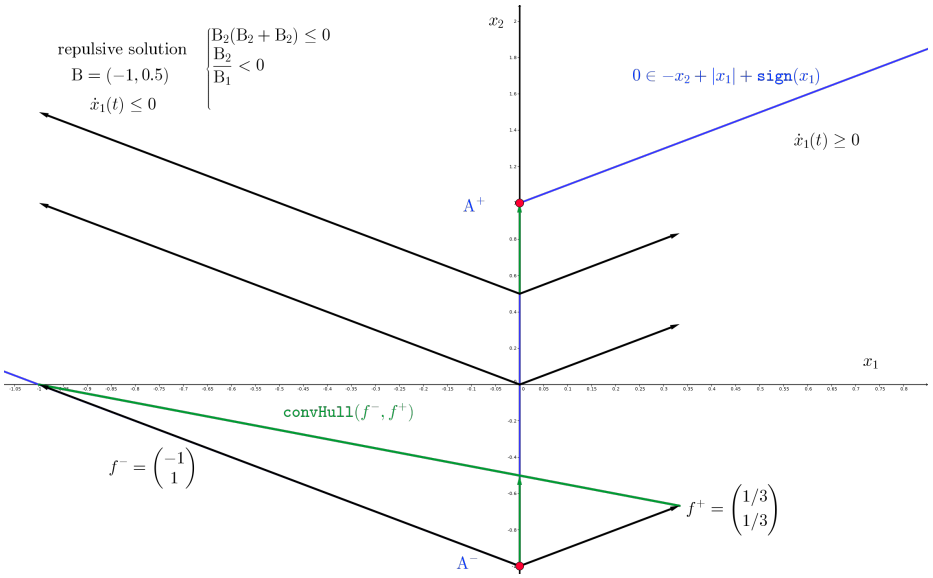


(b) The red cones $(\mathrm{B}_2(\mathrm{B}_1+\mathrm{B}_2) < 0)$ are the sets of parameterization $(\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$ leading to "sliding-repulsive" solutions. For example, $\mathrm{B} = (-1, 0.5)$ is such parameterization. The green cones $(\mathrm{B}_2(\mathrm{B}_1 + \mathrm{B}_2) > 0)$ 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 2 are given in the red and green cones.



(a) Example of solution with $\mathrm{B} = (-0.5, 1)$ which correspond 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 correspond 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 convexication' applied to the 'reduced' DAE. The top figure represents the case of crossing-solutions and the bottom figure represents the case of repulsive-solutions.

the differential measure of $\mathbf{x}(t)$. Let us notice that both sides of the dynamics are considered as Schwartz distributions. Hence z is to be considered as a measure. Using (10) we obtain the equalities of measures :

$$\begin{cases} \mathrm{d}x_1 = \mathrm{d}t + \mathrm{B}_1 \mathrm{d}\Lambda_z \\ \mathrm{d}x_2 = \mathrm{B}_2 \mathrm{d}\Lambda_z \\ \quad 0 \in \ -x_2(t) + |x_1(t)| + \mathrm{sign}(x_1(t)) \, , \end{cases} \tag{17}$$

with $\mathrm{d}\Lambda_z(t) = z(t)\mathrm{d}t + \sigma_z \delta_{t_j}$ the differential measure of z on an interval including t_j, with $\mathrm{d}t$ the Lebesgue measure associated with time, δ_{t_j} the Dirac distribution at $t = t_j$, and σ_z the amplitude of the jump. Then, at t_j the system (17) becomes the algebraic problem:

$$\begin{cases} x_1(t_j^+) - x_1(t_j^-) = \mathrm{B}_1 \sigma_z \\ x_2(t_j^+) - x_2(t_j^-) = \mathrm{B}_2 \sigma_z \\ \qquad\qquad 0 \in \lambda(t_j^+) + |x_1(t_j^+)| - x_2(t_j^+) \\ \qquad\quad \lambda(t_j^+) \in \ \mathrm{sign}(x_1(t_j^+)). \end{cases} \tag{18}$$

We assume $\mathbf{x}(t)$ is an AC solution for all $t < t_j$, that is $\mathbf{x}(t_j^-)$ is satisfying the constraint (11), and there exists $\lambda(t_j^-) \in \mathrm{sign}(x_1(t_j^-))$ such that $0 = \lambda(t_j^-) + |x_1(t_j^-)| - x_2(t_j^-)$. 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[6] and it also allows to study continuation after the jumps, see Section 3.3.3. Multiplying the first and second lines of (18) by B_2 and B_1, respectively, one can eliminate σ_z in (18) which can be rewritten as:

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

Note that for now we do not enforce conditions for existence of a continuous solutions 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^+. Let us analyse the jump dynamics in (18).

3.3.2 Well-posedness of the jump dynamics

Proposition 3 (Jump Dynamics Well-posedness). *Let consider the case* $\mathrm{B}_1 \neq 0$*: if* $\mathrm{B}_2/\mathrm{B}_1 < -1$ *there is a unique solution to* (18)*, otherwise if* $\mathrm{B}_2/\mathrm{B}_1 \geq -1$ *there are either one or several solutions depending on* $\mathbf{x}(t_j^-)$*. Let us now consider* $\mathrm{B}_1 \neq 0$*: if* $x_2(t_j^-) \in [-1, 1]$ *there are infinitely many solutions, otherwise there is only one solution.*

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

Proof.
• Assume $\underline{\mathrm{B}_1 \neq 0}$, then (18) can be further reduced into:

$$0 \in \ \operatorname{sign}\left(x_1(t_j^+)\right) + |x_1(t_j^+)| - \frac{\mathrm{B}_2}{\mathrm{B}_1}\left(x_1(t_j^+) - x_1(t_j^-)\right) - x_2(t_j^-)\,, \tag{19}$$

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

$$0 \in f(x) + \mathcal{F}(x)\,, \tag{20}$$

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

$$f(x) = \mathrm{a}x + \mathrm{b}|x| + \mathrm{c}\,, \tag{21}$$

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 this GE, 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 (21) in a general context, as it will prove to be useful for the study of numerical solutions in Section 3.4. Let now study the conditions for existence and/or uniqueness of solutions to such GE. Proofs will be given in a succinct manner using Figure 4 as a support.

Assume $a - b \neq 0$ and $a + b \neq 0$[7]. Then, depending on the sign of $a - b$ and $a + b$ we define an associated continuous piecewise linear function $h(y)$.

(a) Let 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 $h(y) = \frac{y-c}{a-b}$ if $y \leq c$ and $h(y) = \frac{y-c}{a+b}$ 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 $\operatorname{sign}(x) = \partial|x|$, and the conjugate of $g(x) = |x|$ is $g^*(x) = \psi_{[-1,1]}(y)$. It follows from [23] that the inverse of $\operatorname{sign}(x)$ is $\mathcal{N}_{[-1,1]}(y) = \partial\psi_{[-1,1]}(y)$. Consequently, the GE (20) is equivalent to:

$$0 \in h(y) + \mathcal{N}_{[-1,1]}(y)\,, \tag{22}$$

with $\mathcal{N}_{[-1,1]}(y)$ the normal cone to $[-1, 1]$ in y. This is the canonical form of a generalised equation as analysed in [11]. Finally, $h(\cdot)$ is continuous on $\mathbb{R}$ and as $[-1, 1]$ is a compact set in $\mathbb{R}$, it follows from [11, Corollary 2.2.5] that there is always existence of solutions to the GE (22) (and equivalently (19)).
Additionally, [11, Theorem 2.3.3] provides sufficient conditions for uniqueness: if $h(\cdot)$ is strictly monotone on $[-1, 1]$ then equation (22) 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 on Figure 4. Furthermore, we already know that

[7] We remark that the case $a - b = 0$ (respectively $a + b = 0$) corresponds to the vector B being parallel to the right (respectively left) constraint. This means that DAEs in modes 1 and/or 2 are non-regular with $\mathrm{C}_1\mathrm{B}$ and $\mathrm{C}_2\mathrm{B}$ singular.

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 (21). In addition, these sufficient conditions do not depend on c. 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 $x = \frac{y-c}{a-b}$ if and only if $y \geq c$. Similarly, If $x \geq 0$, $x = \frac{y-c}{a+b}$ if and only if $y \geq c$. It follows that $h(\cdot)$ is a multi-valued operator defined on $[c, +\infty[$. Consequently, if $c > 1$ there are no solutions as the domains of $h(\cdot)$ and $\mathcal{N}_{[-1,1]}(\cdot)$ does 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 $y = c$. 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 for our particular example (19).

(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 $\mathbb{R}$, and again using [11, 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 $\frac{y-c}{a+b}$ with the normal cone in $y = -1$) or $c < 1$ (intersection of $\frac{y-c}{a-b}$ with the normal cone in $y = 1$).

(d) Finally, let consider the particular cases where $a - b = 0$ or $a + b = 0$, which correspond to $\mathrm{B}_2/\mathrm{B}_1 = 1$ or -1. Under these conditions one part of $h(y)$ is multi-valued (its graph has a vertical branch). For example, if $(a - b) = 0$ and $(a + b) > 0$, then we can write $h(y) = f^{-1}(\mathbf{x})$ defined on the domain $[c, +\infty[$ with $h(y) \in \mathbb{R}^-$ if $y = c$, and $h(y) = \frac{y-c}{a+b}$ if $y \geq c$. We note that if $c > 1$ there is no solution to (22). In a similar way, if $(a + b) < 0$ then $h(y) \in \mathbb{R}^-$ if $y = c$, and $h(y) = \frac{y-c}{a+b}$ if $y \leq c$: its domain is $]-\infty, c]$ and there is not solution to (22) if $c < -1$. A similar reasoning can be done for the cases with $a + b = 0$.

We recall that the cases $a-b = 0$, $a+b = 0$, $a-b < 0$ and $a+b > 0$ with $c > 1$ or $c < -1$, which are cases without solutions, cannot occur. Indeed, in these cases the assumption $0 \in -x_2(t_j^-) + |x_1(t_j^-)| + \mathrm{sign}(x_1(t_j^-))$ enforces $c \in [-1, 1]$.

• Assume $\underline{\mathrm{B}_1 = 0}$, then from (18) it follows:

$$\begin{cases} x_1(t_j^+) = x_1(t_j^-) \\ x_2(t_j^+) = x_2(t_j^-) + \mathrm{B}_2\sigma_z \\ x_2(t_j^+) \in |x_1(t_j^-)| + \mathrm{sign}(x_1(t_j^-))\,. \end{cases} \tag{23}$$

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

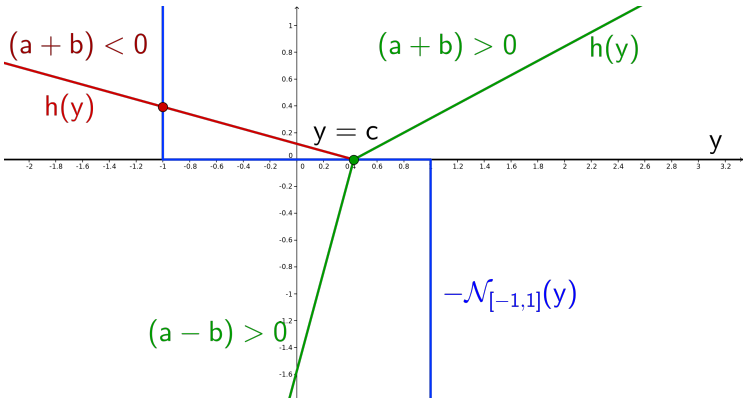


Fig. 4: solutions of the the GE (22) for various signs of $(a+b)$ and $(a-b)$.

3.3.3 Analysis of AC consistent jumps

Let 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*.

From the analysis of global AC solutions in Section 3.1, we know this study can be separated in three cases. Loss of existence of continuous solutions can happen: anywhere in mode 1 if $\mathrm{B}_1+\mathrm{B}_2=0$, anywhere in mode 2 if $\mathrm{B}_1-\mathrm{B}_2=0$, anywhere in mode 3 if $\mathrm{B}_1=0$, and only at point $A^-=(0,-1)$ if $\mathrm{B}_2/\mathrm{B}_1>0$, $\mathrm{B}_1\neq\mathrm{B}_2$, $\mathrm{B}_1\neq 0$. Without loss of generality we can also assume that B_1 and B_2 are not simultaneously null since this would implies there is no locally continuous solution.

Definition 12 (AC consistent initial conditions). *We define AC consistent initial conditions, as the set of points* $\mathbf{x}(t_0)$ *such that* $\mathbf{x}(t_0)$ *satisfies the constraint* $x_2(t_0)\in|x_1(t_0)|+\mathrm{sign}(x_1(t_0))$*, and there exists an AC solution on an interval* $I_\varepsilon=[t_0,t_0+\varepsilon]$*,* $\varepsilon>0$*, with initial condition* $\mathbf{x}(t_0)$*.*

• Assume $\mathrm{B}_1+\mathrm{B}_2=0$. From the study of global AC solutions, the AC consistent initial conditions if $\mathrm{B}_1+\mathrm{B}_2=0$, are all points in mode 2 and 3, that is:

$$S_1=\{(x_1,x_2)\in\{0\}\times[-1,1]\}\cup\{(x_1,x_2)\in\mathbb{R}_{+,*}\times\{x_1+1\}\}\ .$$

We now look for solutions $\mathbf{x}(t_j^+)$ of the jump dynamic problem (19) (we take $\mathrm{B}_1\neq 0$ by assumption) such that $\mathbf{x}(t_j^+)\in S_1$, and $\mathbf{x}(t_j^-)\in\mathbb{R}_{-,*}\times\{-x_1(t_j^-)-1\}$ (that is, initial condition is in mode 1). Under these conditions, Equation (19) can be reduced to:

$$(-x_1(t_j^+)-1)\in\mathrm{sign}(x_1(t_j^+))+|x_1(t_j^+)|\,.$$

The set of solutions for this equation is $(x_1(t_j^+), x_2(t_j^+)) \in \mathbb{R}_- \times \{-x_1(t_j^+)-1\}$, which intersects S_1 only in $(0,-1)$: there exists a unique consistent jump to A^-.
• Assume $\underline{\mathrm{B}_1 - \mathrm{B}_2 = 0}$. Then, from the study of global AC solutions, the AC consistent initial conditions are all points in mode 1 and 3 with the exception of point $A^- = (0,-1)$, that is:

$$S_2 = \{(x_1,x_2) \in \{0\}\times]-1,1]\} \cup \{(x_1,x_2) \in \mathbb{R}_{-,*} \times \{-x_1 - 1\}\} \ .$$

Let us look for solutions $\mathbf{x}(t_j^+)$ of the jump dynamic problem (19) such that $\mathbf{x}(t_j^+) \in S_2$, and $\mathbf{x}(t_j^-) \in \mathbb{R}_{+,*} \times \{x_1(t_j^-)+1\}$ (that is initial condition is in mode 2). Under these conditions, (19) can be reduced to:

$$(x_1(t_j^+)+1) \in \mathrm{sign}(x_1(t_j^+)) + |x_1(t_j^+)| \, .$$

The set of solutions to this equation is $(x_1(t_j^+), x_2(t_j^+)) \in \mathbb{R}_+ \times \{x_1(t_j^+) + 1\} \cup \{(-1,0)\}$, which intersects S_2 in $(0,1)$ and $(-1,0)$: there exist multiple consistent jumps. The case where $\mathbf{x}(t_j^-) = (0,-1)$ will be treated in the case $\mathrm{B}_2/\mathrm{B}_1 > 0$.
• Assume $\underline{\mathrm{B}_1 = 0}$, then the AC consistent initial conditions are all points in mode 1 and 2 that is:

$$S_3 = \{(x_1,x_2) \in \mathbb{R}_{-,*} \times \{-x_1 - 1\}\} \cup \{(x_1,x_2) \in \mathbb{R}_+ \times \{x_1 + 1\}\}$$

From (18) the jumps dynamics is given by (23). Then, the jump must happen if $x_1(t_j^-) = 0$. It follows that $x_1(t_j^+) = 0$ and there are solutions with an AC consistent jump only if $x_2(t_j^+) = 1$.
• Assume $\underline{\mathrm{B}_2/\mathrm{B}_1 > 0}$, $\mathrm{B}_1 \neq \mathrm{B}_2$, $\mathrm{B}_1 \neq 0$. Then the set of consistent initial conditions is:

$$S_4 = \{(x_1,x_2) \in \mathbb{R}^2 \backslash \{(0,-1)\} \mid x_2 \in |x_1| + \mathrm{sign}(x_1)\} \, ,$$

that is the whole constraint with the exception of point $A^- = (0,-1)$. It follows that the jump can occur only in A^-. Under this assumption, Equation (19) is:

$$0 \in \ \mathrm{sign}\left(x_1(t_j^+)\right) + |x_1(t_j^+)| - \frac{\mathrm{B}_2}{\mathrm{B}_1}\left(x_1(t_j^+)\right) + 1 \, . \tag{24}$$

This is equivalent to look for the intersections between the constraint $c(x_1) = |x_1| + \mathrm{sign}(x_1)$ and the real line $l(x_1) = (\mathrm{B}_2/\mathrm{B}_1)x_1 - 1$ (see Figure 5). There are up to two intersections: one in $(x_1,x_2) = (0,-1)$ which is not in S_4 and an additional one in mode 2: $(x_1,x_2) = \left(\frac{2\mathrm{B}_1}{(\mathrm{B}_2-\mathrm{B}_1)}, \frac{(\mathrm{B}_1+\mathrm{B}_2)}{(\mathrm{B}_2-\mathrm{B}_1)}\right)$ if and only if $(\mathrm{B}_2 - \mathrm{B}_1) > 0$

3.4 Analysis of a time-stepping Backward Euler scheme

Backward (implicit) Euler schemes have proved to be efficient schemes for the simulation of nonsmooth dynamical systems [1,2]. Therefore, let now consider

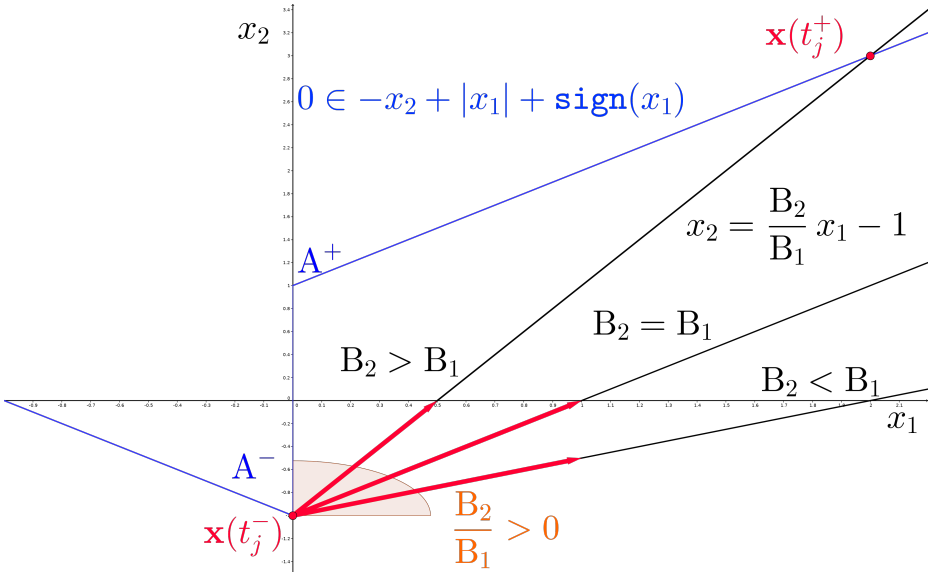


Fig. 5: Solutions of the GE (24) associated with jumps from $\mathbf{x}(t_j^-) = (0, -1)$ for various choices of $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$.

the backward Euler discretization of system (10):

$$\begin{cases} x_{1,k+1} - x_{1,k} = h(1 + \mathrm{B}_1 z_{k+1}) \\ x_{2,k+1} - x_{2,k} = h\mathrm{B}_2 z_{k+1} \\ \qquad\qquad 0 \in \ \mathrm{sign}(x_{1,k+1}) + |x_{1,k+1}| - x_{2,k+1}\,, \end{cases} \tag{25}$$

with $h > 0$ a fixed time step.

3.4.1 Well-posedness of the backward Euler discretization

One sees that (25) has a structure quite close to (17), which puts the backward scheme as favourable perspective for the computations of solutions with jumps. This is the object of the next analysis. Using the same method as in the previous section, we can eliminate z_{k+1} and obtain the GE:

$$\begin{cases} \mathrm{B}_2(x_{1,k+1} - x_{1,k}) = h\mathrm{B}_2 + \mathrm{B}_1(x_{2,k+1} - x_{2,k}) \\ \qquad\qquad 0 \in \ \mathrm{sign}(x_{1,k+1}) + |x_{1,k+1}| - x_{2,k+1}\,. \end{cases} \tag{26}$$

Proposition 4 (Backward Euler Well-posedness). *Let consider the case* $\mathrm{B}_1 \neq 0$*: if* $\mathrm{B}_2/\mathrm{B}_1 < -1$ *there is a unique solution to* (26) *whatever the size of the time step* h*. Otherwise, if* $\mathrm{B}_2/\mathrm{B}_1 \geq -1$ *there are either none, one or several solutions depending on* $\mathbf{x}_k$ *and* h*. Let us now consider* $\mathrm{B}_1 \neq 0$*: if* $x_{1,k} = -h$ *there are infinitely many solutions, otherwise there is only one solution.*

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

$$0 \in -\frac{\mathrm{B}_2}{\mathrm{B}_1} x_{1,k+1} + |x_{1,k+1}| + \left(-x_{2,k} + \frac{\mathrm{B}_2}{\mathrm{B}_1}(x_{1,k} + h) \right) + \mathrm{sign}(x_{1,k+1}) \qquad (27)$$

Sufficient conditions for uniqueness are the same as in the previous section, identifying:

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

with $a = -\frac{\mathrm{B}_2}{\mathrm{B}_1}$, $b = 1$, and $c = \left(-x_{2,k} + \frac{\mathrm{B}_2}{\mathrm{B}_1}(x_{1,k} + h) \right)$: it is sufficient that $h(y) = f^{-1}(y)$ be strictly monotone.
Consequently, *sufficient conditions* for uniqueness are: $a - b = -(1 + \frac{\mathrm{B}_2}{\mathrm{B}_1}) > 0$, and $a + b = 1 - \frac{\mathrm{B}_2}{\mathrm{B}_1} > 0$, which is case (a) in Section 3.3.2. We note that this corresponds to a subset of the valid $\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2)^{\mathrm{T}}$ for Proposition 2: $\mathrm{B}_1 \neq 0$, $\mathrm{B}_2\mathrm{B}_1 < 0$ and $(\mathrm{B}_1 + \mathrm{B}_2)\mathrm{B}_2 > 0$, the subset of "sliding-crossing" solutions. In this subset of B there is also uniqueness of the continuous solution. Furthermore, the condition $a - b > 0$ excludes the constant solutions with $\mathrm{B}_2 = 0$ where there is not necessarily uniqueness of the solution. An important remark when $a - b > 0$ and $a + b > 0$ is that well-posedness of the discrete solution is independent of the time step h as there are no condition on c.

Thanks to the study of solutions for (21) in Section 3.3.2, we remark that if $a - b < 0$ and $a + b < 0$ (case (c)), there is always existence of a discrete solution, but uniqueness is not guaranteed. This corresponds to the set of B such that $\frac{\mathrm{B}_2}{\mathrm{B}_1} > 1$, that is the condition such that there is a consistent jump in point A^- (see Figure 5).

In addition, there will be a unique solution to the discrete scheme in two cases: $c > 1$ or $c < -1$. The case $c > 1$ correspond to h (or $x_{1,k}$) sufficiently big, for example when the exact solution of (10) must jump at $t_j \in [t_0, t_0 + h]$ assuming an initial condition $\mathbf{x}(t_0) = \mathbf{x}_k$. For example, this is the case if $B = (1, 2)$, $\mathbf{x}_k = (-0.5, -0.5)$ with $h > 0.75$, $t_0 = 0$ and $t_j = 0.75$. The case $c < -1$ corresponds to h (or $x_{1,k}$) sufficiently small: for example for $B = (1, 2)$, $\mathbf{x} = (-1.5, 0.5)$, and $h < 1.25$.

Outside of these cases where there are unique solutions, there is existence of solution but not uniqueness. In particular as it can be seen in Section 3.3.3, in the case where $\frac{\mathrm{B}_2}{\mathrm{B}_1} > 1$, there exists a continuous solution everywhere except in $A = (0, -1)$. It follows in this case, there always exist a solution to the backward Euler discretization when there exists a consistent jump (See Proposition 5). However, the implicit Euler scheme provides both the solutions with, and without jumps, at each time step. For this reason, we give in Conjecture 1 an additional criterion to the backward Euler scheme, to select the discrete solution "associated" with a continuous solution if at least one exists, and otherwise choose a jumping solution.

This is critical in the case where the continuous, or jump, solution cannot exist, after a given instant t_f. Indeed, the Euler scheme may provide non-unique solutions with one of them jumping consistently to an unreachable part of the constraint. This can be the case if $a - b < 0$ and $a + b > 0$. For example if $\mathbf{x}_0 = (-5, 4)$ and $\mathrm{B} = (2, 1)$, then there is no solutions (AC or with jumps) after $t_f = 15$. However, if for example $h = 1$, then one of the solution of the Euler discretization after the first step is $\mathbf{x}_1 = (10, 11)$ after which an AC solution exists forever. Again, the refinement proposed in Conjecture 1 below should enable for $h > 0$ sufficiently small to eliminate these wrong results.

More generally, the conditions $(a - b) < 0$ and $(a + b) > 0$ can be separated into two subsets: the condition $\frac{\mathrm{B}_2}{\mathrm{B}_1} \leq 0$, which implies is $-1 < \frac{\mathrm{B}_2}{\mathrm{B}_1} \leq 0$, and the condition $\frac{\mathrm{B}_2}{\mathrm{B}_1} > 0$, which implies $0 < \frac{\mathrm{B}_2}{\mathrm{B}_1} < 1$.

If $\frac{\mathrm{B}_2}{\mathrm{B}_1} \leq 0$ there is always existence of a discrete solution, as it can be proven that $c \in [-1, 1]$ for all $\mathbf{x}_k$ satisfying (11) and any $h > 0$. Furthermore there is uniqueness if $c = -1$, or $c = 1$, for some h big enough, respectively small enough, depending of the sign of $x_{1,k}$. However, it also is possible to have multiple solutions whatever the choice of $h \geq 0$, in which case the implicit Euler scheme is not consistent: there exists some numerical solutions that do not satisfy Definition 10. For example, if $\mathbf{x}_k = (-2, 1)$ and $\mathrm{B}_2/\mathrm{B}_1 < 0$ then for all $0 \leq h \leq 2$ there exists two solutions for $\mathbf{x}_{k+1}$: one with $\mathbf{x}_{k+1} < 0$ and one with $\mathbf{x}_{k+1} = 0$. Note that if $h > 2$ there are still two solutions: one with $\mathbf{x}_{k+1} < 0$ and one with $\mathbf{x}_{k+1} > 0$.

If $\frac{\mathrm{B}_2}{\mathrm{B}_1} > 0$, as we have seen in Section 3.3.2 there is either non-existence, just existence, or uniqueness. There is non-existence of solution to the implicit Euler scheme when $c > 1$: this corresponds to a time step size h leading to a time t_f where there is no possible continuous solution or solution with jumps (cases $\mathrm{B}_2 = \mathrm{B}_1$ and $\mathrm{B}_2 < \mathrm{B}_1$ in Figure 5). There is existence of solutions (but not uniqueness) for $c < 1$: this correspond to the precedent example where the discrete solutions can "jump" to a point unreachable by a continuous solution or a solution with finite discontinuities (with respect to an initial condition). Finally, uniqueness occurs when $c = 1$ which occurs for some values of h and $\mathbf{x}_k$ well chosen: for example, if $\mathrm{B} = (0.5, 0.25)$, $\mathbf{x}_k = (-0.75, -0.25)$, and $h = 1.25$ then $\mathbf{x}_{k+1} = (0, -1)$ is the only solution (after which there is no solution).

• If $\underline{\mathrm{B}_1 = 0}$, we obtain the system:

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

Then, for all $x_{1,k} \neq -h$ there is a unique solution corresponding to either the continuous solution or the jump. If $x_{1,k} = -h$, there is a non unique solution in $[-1, 1]$. □

We conclude that in this particular example the set of vectors B where there is always uniqueness of solutions for the discrete scheme is a subset of the vectors B

for which there are global continuous solutions. Additionally, (with the exception of $\mathrm{B}_2 = 0$) this corresponds to the subset where uniqueness of the continuous solution holds.

3.4.2 Relation between backward Euler and consistent initialisation

As we have seen in Section 3.4.1 on our working example, most of the solutions obtained by implicit Euler discretization are solutions approximating a consistent jump[8] followed by an AC solution. Let us show this property in a more general context of index-2 nonsmooth DAEs.

Definition 13 (Piecewise Linear Non-smooth DAE). *We define an index-2 piecewise linear non-smooth DAE as a system of the form:*

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathrm{A}\mathbf{x}(t) + \mathrm{B}\mathbf{z}(t) + b \\ 0 &\in \mathcal{F}(\mathbf{x}(t)) \quad \forall t \geq t_0\,, \end{aligned} \tag{28}$$

where the solutions of the generalized equation $0 \in \mathcal{F}(\mathbf{x})$ can be represented as a finite union of connected hyper-planes of the form $\mathrm{C}_i\mathbf{x} + e_i = 0$.

This means that for all $\mathbf{x}$ such that $0 \in \mathcal{F}(\mathbf{x})$, $\mathbf{x}$ is also solution of at least one linear constraint $\mathrm{C}_i\mathbf{x} + e_i = 0$. In addition, we assume that for all C_i, $C_i\mathrm{B}$ is non singular[9]*.*

Proposition 5 (Jump consistency representation by the implicit Euler method). *Let us assume there exists a consistent jump in $\mathbf{x}_0^- = \mathbf{x}(t_0^-)$ to system (28): this means there exists $\mathbf{x}_0^+ = \mathbf{x}(t_0^+)$ a solution of the jump dynamic (18) from $\mathbf{x}_0^-$ at instant t_0 that is a consistent initial condition for (28). It follows there exists an AC solution $\mathbf{x}(t)$ on $]t_0 t_0 + \varepsilon]$, $\varepsilon > 0$ such that for all $t \in]t_0 t_0 + \varepsilon]$ the solution $\mathbf{x}(t)$ satisfies a local linear constraint $\mathrm{C}_i\mathbf{x}(t) + e_i = 0$ constitutive of $\mathcal{F}(\cdot)$. The implicit Euler scheme applied to (28) read as:*

$$\begin{aligned} \mathbf{x}_{k+1} - \mathbf{x}_k &= h\mathrm{A}\mathbf{x}_{k+1} + h\mathrm{B}\mathbf{z}_{k+1} + hb \\ 0 &\in \mathcal{F}(\mathbf{x}_{k+1}) \end{aligned} \tag{29}$$

Then, there exists a solution to (29), which is a O(h) approximation of this AC solution $\mathbf{x}(t)$ for $h < \varepsilon$ and $t \in]t_0, t_0 + \varepsilon]$ with initial condition x_0^+. In addition the solution $\mathbf{x}_{k+1}$ of the Euler discretization with initial condition x_0^-, and time step $h > 0$ is exactly the solution of the Euler discretization with initial condition x_0^+, and time step $h > 0$: the jump dynamic is exactly represented by the implicit Euler scheme.

Proof. By assumption of the consistent solution, for all $t \in]t_0, t_0 + h]$, $x(t)$ is solution of:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathrm{A}\mathbf{x}(t) + \mathrm{B}\mathbf{z}(t) + b \\ 0 &= \mathrm{C}_i\mathbf{x}(t) + e_i\,, \end{aligned}$$

[8] Let us recall that a consistent jump can be of amplitude null and then the solution is A.C. on the considered interval

[9] Let us recall *this is not* a sufficient condition to say that $\mathbf{y}$ is a consistent initial condition of (28). However, it is a necessary condition.

and from [8, Theorem 3.1.1][10] since we are considering locally an index-2 linear DAE with constant coefficients, there exists an $O(h)$ approximation of $\mathbf{x}(t_0+h)$ by the solution of:

$$\begin{aligned}\mathbf{x}_1 - \mathbf{x}_0^+ &= h\mathrm{A}\mathbf{x}_1 + h\mathrm{B}\gamma_1 + hb\\ 0 = \mathrm{C}_i\mathbf{x}_1 + e_i &\Rightarrow 0 \in \ \mathcal{F}(\mathbf{x}_1)\,.\end{aligned}$$

In addition, $\mathbf{x}_0^+$ is a solution of the jump dynamics:

$$\begin{aligned}\mathbf{x}_0^+ &= \mathbf{x}_0^- + \mathrm{B}\sigma\\ 0 &\in \mathcal{F}(\mathbf{x}_0^+)\end{aligned}$$

with its associated jump amplitude σ. It follows that:

$$\begin{aligned}\mathbf{x}_1 - \mathbf{x}_0^- &= h\mathrm{A}\mathbf{x}_1 + h\mathrm{B}(\gamma_1 + \frac{\sigma}{h}) + hb\\ 0 &\in \mathcal{F}(\mathbf{x}_1)\,,\end{aligned}$$

which proves that the proposed $\mathbf{x}_1$ is a solution of system (29), with $z_1 = \gamma_1 + \frac{\sigma}{h}$. □

It is important to note that we do not prove that all solutions of Euler are approximating a consistent jump. An easy counter-example can be found in Section 3.4.1 for the case $\mathrm{B}_1 = 0$, $x_{1,k} = -h$ where there is a dense set of solutions $x_{2,k+1} \in [-1, 1[$ which are not consistent initialisation.

3.4.3 Minimal implicit Euler discretization

As we have seen in Section 3.4.1, the classical implicit Euler discretization may output multiple solutions, for $h > 0$ as small as wanted (see the case with $(a-b) < 0$ and $(a+b) > 0$). One needs to refine the results of 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 in order to keep the solutions the closest in Euclidian norm.

Conjecture 1 (Minimal Backward Euler). Considering a non-smooth DAE system close to (28):

$$\begin{cases}\dot{\mathbf{x}} = \mathrm{A}\mathbf{x} + \mathrm{B}\mathbf{z} + \mathbf{b} \quad \text{a.e. on } [t_0, +\infty)\\ 0 \in \mathcal{F}(\mathbf{x}, \mathbf{z}) \quad \forall t \in [t_0, +\infty)\\ \mathbf{x}_0 = \mathbf{x}(t_0)\,,\end{cases} \tag{30}$$

with $\mathbf{x}$ the differential variables, $\mathbf{z}$ algebraic variables, and where the solutions of the generalized equation[11] $0 \in \mathcal{F}(\cdot)$ can be represented as a finite union of

[10] Note that here we have a $O(h)$ approximation after one step as the constraint is only state-dependent (see proof of [8, Theorem 3.1.1]).

[11] Here express as the solution of a Mixed Linear Complementarity Problem.

strongly connected hyper-planes of the form $\mathrm{C}_i\mathbf{x} + \mathrm{D}_i\mathbf{z} + e_i = 0$ such that (30) is of differentiation index less or equal to two. If there exists a unique solution $Y(t) = (\mathbf{x}(t), \mathbf{z}(t), \boldsymbol{\lambda}(t))$ such that $\mathbf{x}(t)$ 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:

$$\begin{aligned} p^*_{k+1} := \min_{\mathbf{x}_{k+1},\mathbf{z}_{k+1},\boldsymbol{\lambda}_{k+1}} \quad & \|\mathbf{x}_{k+1} - \mathbf{x}_k\|, \\ \text{s.t} \quad & \mathbf{x}_{k+1} - \mathbf{x}_k = h\mathrm{A}\mathbf{x}_{k+1} + h\mathrm{B}\mathbf{z}_{k+1} + h\mathbf{b} \\ & 0 \in \mathcal{F}(\mathbf{x}_{k+1}, \mathbf{z}_{k+1}), \end{aligned} \tag{31}$$

provides a consistent discrete solution to (30). This means that given $\mathrm{Y}_k = \mathrm{Y}(t_k)$ and $\mathrm{Y}^*_{k+1} = \mathrm{argmin}(p^*_{k+1})$ then $\left\|\mathrm{Y}^*_{k+1} - \mathrm{Y}(t_k + h)\right\| \to 0$ when $h \to 0$, which can be simplified to $\left\|\mathrm{Y}^*_{k+1} - \mathrm{Y}_k\right\| = O(h)$).

Let us observe that Euler gives an $O(h)$ approximation of an index-2 linear DAE with constant coefficients solution, as seen in [8, Theorem 3.1.1]. Let assume that $\mathbf{x}_k = \mathbf{x}(t_0)$ satisfies the linear constraint $\mathrm{C}_i\mathbf{x} + \mathrm{D}_i\mathbf{z} + e_i = 0$, and that $\mathbf{x}(t)$ is solution of the corresponding index-2 linear constant coefficient DAE. Then, if it exists a continuous solutions in the neighbourhood of t_0, then there is $\varepsilon > 0$ sufficiently small such that $\mathbf{x}(t_0 + \varepsilon)$ satisfies $\mathrm{C}_i\mathbf{x} + \mathrm{D}_i\mathbf{z} + e_i = 0$. As a result, if one of the solutions of the implicit Euler, for $h < \varepsilon$, is such that $\mathrm{C}_i\mathbf{x}_{k+1} + \mathrm{D}_i\mathbf{z}_{k+1} + e_i = 0$, then this solution is a O(h) approximation of $\mathbf{x}(t_0+\varepsilon)$.

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 consistent initial condition for our particular example.

Proposition 6 (Minimal Backward Euler). *Considering the non-smooth DAE system from* (10)*. If there exists a solution* $Y(t) = (\mathbf{x}(t), \mathbf{z}(t), \boldsymbol{\lambda}(t))$ *such that* $\mathbf{x}(t)$ *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:*

$$\begin{aligned} p^*_{k+1} := \min_{\mathbf{x}_{k+1},\mathbf{z}_{k+1},\boldsymbol{\lambda}_{k+1}} \quad & \|\mathbf{x}_{k+1} - \mathbf{x}_k\|, \\ \text{s.t} \quad & x_{1,k+1} - x_{1,k} = h(1 + B_1 z_{k+1}) \\ & x_{2,k+1} - x_{2,k} = hB_2 z_{k+1} \\ & 0 \in \ \mathrm{sign}(x_{1,k+1}) + |x_{1,k+1}| - x_{2,k+1}, \end{aligned} \tag{32}$$

provides a consistent discrete solution to (30) *as defined in Conjecture 1.*

Proof. Let first observe that if:

(A) If $x_{1,k} < 0$ and $x_{1,k+1} < 0$, then $\|x_{1,k+1} - x_{1,k}\| = |\frac{\mathrm{B}_2}{\mathrm{B}_1+\mathrm{B}_2}|h = O(h)$ and $\|x_{2,k+1} - x_{2,k}\| = |\frac{-\mathrm{B}_2}{\mathrm{B}_1+\mathrm{B}_2}|h = O(h)$.
(B) If $x_{1,k} = 0$ and $x_{1,k+1} = 0$, then $\|x_{1,k+1} - x_{1,k}\| = 0 = O(h)$ and $\|x_{2,k+1} - x_{2,k}\| = |\frac{-\mathrm{B}_2}{\mathrm{B}_1}|h = O(h)$.
(C) If $x_{1,k} > 0$ and $x_{1,k+1} > 0$, then $\|x_{1,k+1} - x_{1,k}\| = |\frac{-\mathrm{B}_2}{\mathrm{B}_1-\mathrm{B}_2}|h = O(h)$ and $\|x_{2,k+1} - x_{2,k}\| = |\frac{-\mathrm{B}_2}{\mathrm{B}_1-\mathrm{B}_2}|h = O(h)$.

We now have to show that in all the case for $(a-b)$, $(a+b)$ studied in Section 3.4.1: if there exists a local continuous solution, then there exists h sufficiently small such that we can find a solution corresponding to either (A), (B) or (C). Then, as this solution is consistent the minimal Euler scheme from Proposition 6 will select it for $h > 0$ sufficiently small.

(1) Let us assume $(a+b) > 0$, $(a-b) > 0$. This implies that $\mathrm{B}_2/\mathrm{B}_1 < 0$.
  (1.a) Let consider $x_{1,k} < 0$. We search h such that $x_{1,k+1} < 0$. From the resolution of the general equation this implies $c > 1$ and it follows:
$$0 \leq h < -x_{1,k}(1 + \frac{\mathrm{B}_1}{\mathrm{B}_2})$$
  (1.b) Let consider $x_{1,k} = 0$. We search h such that $x_{1,k+1} = 0$. From the resolution of the general equation this implies $-1 \leq c \leq 1$ and it follows that if $x_{2,k} \in [-1, 1[$ then:
$$0 \leq h < \frac{\mathrm{B}_1}{\mathrm{B}_2}(x_{2,k} - 1)$$
  If $x_{2,k} = 1$ then $h = 0$. This case will be treated in (1.c).
  (1.c) Let consider $x_{1,k} \geq 0$ and $x_{2,k} = x_{1,k} + 1$ to contain the precedent untreated case. We search h such that $x_{1,k+1} > 0$. From the resolution of the general equation this implies $c < -1$ and this holds for any $h \geq 0$
(2) Let us assume $(a+b) > 0$, $(a-b) > 0$. This implies that $\mathrm{B}_2/\mathrm{B}_1 > 1 (> 0)$.
  (2.a) Let consider $x_{1,k} < 0$. We search h such that $x_{1,k+1} < 0$. From the resolution of the general equation this implies $c < 1$ and it follows:
$$0 \leq h < -x_{1,k}(1 + \frac{\mathrm{B}_1}{\mathrm{B}_2})$$
  Note this does not holds in $x_{1,k} = 0$ and $x_{2,k} = -1$.
  (2.b) Let consider $x_{1,k} = 0$. We search h such that $x_{1,k+1} = 0$. From the resolution of the general equation this implies $-1 \leq c \leq 1$ and it follows that if $x_{2,k} \in]-1, 1]$ then:
$$0 \leq h < \frac{\mathrm{B}_1}{\mathrm{B}_2}(x_{2,k} + 1)$$
  If $x_{2,k} = -1$ then $h = 0$. Again there is no O(h) solution in $\mathbf{x}_k = (0, 1)$, and there is no AC solution in this point.
  (2.c) Let consider $x_{1,k} \geq 0$. We search h such that $x_{1,k+1} > 0$. From the resolution of the general equation this implies $c \geq -1$ and this holds for any $h \geq 0$. Note the point $\mathbf{x}_k = (0, 1)$ is handled both in this case and in (2.b). Indeed, there are multiple AC solutions for this initial condition.
(3) Let us assume $(a+b) > 0$, $(a-b) < 0$ and $0 < \mathrm{B}_2/\mathrm{B}_1 < 1$.
  (3.a) Let consider $x_{1,k} < 0$. We search h such that $x_{1,k+1} < 0$. From the resolution of the general equation this implies $c < 1$ and it follows:
$$0 \leq h < -x_{1,k}(1 + \frac{\mathrm{B}_1}{\mathrm{B}_2})$$
  Note this does not holds in $x_{1,k} = 0$ and $x_{2,k} = -1$.

(3.b) Let consider $x_{1,k} = 0$. We search h such that $x_{1,k+1} = 0$. From the resolution of the general equation this implies $-1 \leq c \leq 1$ and it follows that if $x_{2,k} \in]-1, 1]$ then:

$$0 \leq h < \frac{\mathrm{B}_1}{\mathrm{B}_2}(x_{2,k} + 1)$$

If $x_{2,k} = -1$ then $h = 0$. There is no O(h) solution in $\mathbf{x}_k = (0, -1)$, and there is no AC solution in this point.

(3.c) Let consider $x_{1,k} > 0$. We search h such that $x_{1,k+1} > 0$. From the resolution of the general equation this implies $c < -1$ and it follows:

$$0 \leq h < x_{1,k}(\frac{\mathrm{B}_1}{\mathrm{B}_2} - 1)$$

Note that the point $\mathbf{x}_k = (0, 1)$ is handled in (3.b).

(4) Let us assume $(a + b) > 0$, $(a - b) < 0$ and $-1 < \mathrm{B}_2/\mathrm{B}_1 < 0$.

(4.a) Let consider $x_{1,k} \leq 0$. We search h such that $x_{1,k+1} < 0$. From the resolution of the general equation this implies $c < 1$ and it follows this holds for any $h \geq 0$.

(4.b) Let consider $x_{1,k} = 0$. We search h such that $x_{1,k+1} = 0$. From the resolution of the general equation this implies $-1 \leq c \leq 1$ and it follows that if $x_{2,k} \in [-1, 1[$ then:

$$0 \leq h < \frac{\mathrm{B}_1}{\mathrm{B}_2}(x_{2,k} - 1)$$

If $x_{2,k} = 1$ then $h = 0$. This case will be treated in (4.c). In addition, in the point $\mathbf{x}_k = (0, -1)$ there are 2 possible $O(h)$ solutions one given by (4.a) and one by (4.b): this corresponds to a multiplicity of AC solutions in this particular context.

(4.c) Let consider $x_{1,k} \geq 0$ and $x_{2,k} = x_{1,k} + 1$ to contain the precedent untreated case. We search h such that $x_{1,k+1} > 0$. From the resolution of the general equation this implies $c < -1$ and this holds for any $h \geq 0$.

In each case, we can find h sufficiently small such that if there exists an AC solutions then there exists a O(h) solutions to the implicit Euler scheme. □

Remark 3 (Necessary conditions of optimality). The optimisation problem solved in Conjecture 1 is a mathematical program with equilibrium constraints (MPEC). Under the assumption of a MLCP representation of the generalised equation,necessary conditions of optimality depends on some constraint qualifications. We refers to [26] as a reference on this question. In the simulations presented in Section 3.5, we address this problem in a naive way by simply enumerating all the solutions and selecting the one of minimal norm as the aim is to show the experimental convergence of such method for this problem.

Remark 4. Let us also observe that when the generalised equation can be expressed using a MLCP, we simply check there exists a solution in the mode of

$\mathbf{x}_k$ before performing the minimal Euler resolution. This can be done by considering only the current sets active and inactive constraints at step k, instead of the whole MLCP. This would correspond to the cases (A), (B), (C) of the proof above where consistence (and convergence) is guaranteed.

3.5 Implementation and numerical results

In this section we expose some actual simulation results of the minimal implicit Euler scheme (31) on the example studied in the previous sections. In particular, we show through experiments on example (10) that if there exists at least one continuous solution, then (31) converges in $O(h)$ to one of these solutions. Furthermore, if discretization (25) yields a unique solution for any step size then it converges in $O(h)$ to this unique solutions of (10).

Implementation has been performed using the Siconos 4.2.0 [4], a platform for numerical simulation of non-smooth dynamical systems. The code of these simulation can be found in the github repository associated to the Siconos examples[12]. In this section, performance results are not discussed as the optimisation problem in (31) is currently solved by enumeration of all[13] the solutions of the generalised equation associated to the classical implicit Euler scheme (25).

To simulate the problem stated in (10) using Siconos one have multiple ways to represent the non-smooth components $|x_1|$ and $\text{sign}(x_1)$ in constraint (11). First the constraint can be represented using only the Relay approach:

$$0 \in \ -x_2 + |x_1| + \text{sign}(x_1) \Leftrightarrow \begin{cases} 0 = \ -x_2 + (1+x_1)\lambda \\ x_1 \in \mathcal{N}_{[-1,1]}(\lambda) \end{cases} \tag{33}$$

this yields the following representation in Siconos:

$$\begin{cases} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{z}(t) \end{pmatrix} = \begin{pmatrix} 1 + \mathrm{B}_1 z(t) \\ \mathrm{B}_2 z(t) \\ -x_2(t) + \mathtt{r}(\mathbf{x}(t), z(t), \lambda(t)) \end{pmatrix} \\ \mathtt{r}(\mathbf{x}(t), z(t), \lambda(t)) = (1 + x_1(t))\lambda(t) \\ \mathtt{w}(\mathbf{x}(t), z(t), \lambda(t)) = \ -x_1(t) \\ -\mathtt{w}(\mathbf{x}(t), z(t), \lambda(t)) \in \mathcal{N}_{[-1,1]}(\lambda(t)) \end{cases} \tag{34}$$

One can note that this representation of the problem contains a term $(1 + x_1(t))\lambda(t)$ non-linear in $\mathbf{x}$, λ. This implies that the solver will use a Newton scheme to linearize it. Sadly, numerical simulation using this representation of the non-smooth law gives erroneous results (See Figure 6). This may be due to the lack of convergence of the current implementation of the Newton Jacobi solver for first order problems. The second way to implement this problem is using a linear complementarity problem (LCP) formulation to represent the non-smooth component of constraint (11). Then constraint (11) can be expressed as

[12] `https://github.com/siconos/siconos-tutorials/tree/master/.sandbox/code_IFAC`
[13] In the considered cases limited to maximum 3 solutions

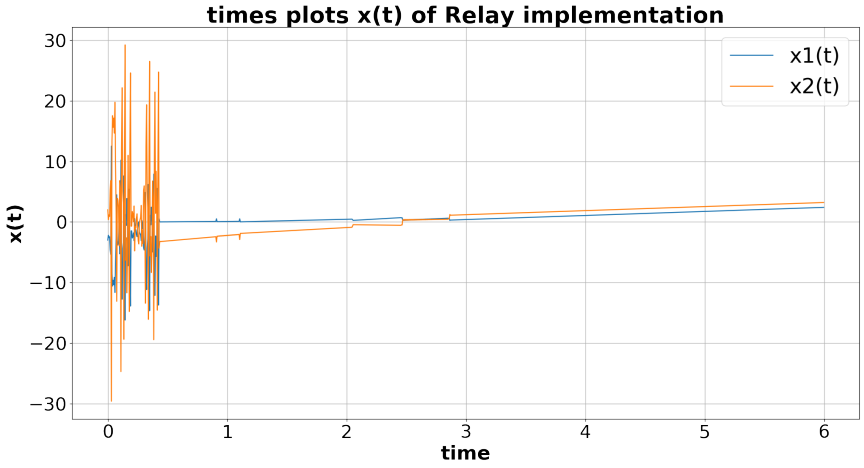


Fig. 6: Numerical simulation of the representation with relays as in (34) for $\mathbf{x}(0) = (-3, 2)$ and $\mathrm{B} = (-0.49, 1)^{\mathrm{T}}$. We note that $\mathrm{B} = (-0.5, 1)$ leads to a non-invertibility error in the Euler solving process. The numerical simulation is unstable and the Newton scheme fails to converge even with numerous steps.

a LCP with an additional equality constraint (MLCP).

$$\begin{cases} 0 = -x_2 + |x_1| + \alpha \\ 0 \leq |x_1| + x_1 \perp |x_1| - x_1 \geq 0 \\ 0 \leq x_1^- \perp \alpha \geq -1 \\ \alpha \leq 1 \perp x_1^+ \geq 0 \end{cases} \tag{35}$$

with $\alpha \in \mathrm{sign}(x_1)$. This yield a linear complementarity system (LCS)[14]:

$$\begin{cases} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{z}(t) \end{pmatrix} = \begin{pmatrix} 1 + \mathrm{B}_1 z(t) \\ \mathrm{B}_2 z(t) \\ x_1 - x_2 + 1 + \mathtt{r}(\boldsymbol{\lambda}(t)) \end{pmatrix} \\ \mathtt{r}(\boldsymbol{\lambda}(t)) = \lambda_1(t) - \lambda_2(t) \\ 0 \leq 2x_1(t) + \lambda_1(t) \perp \lambda_1(t) \geq 0 \\ 0 \leq \lambda_3(t) + x_1(t) \perp \lambda_2(t) \geq 0 \\ 0 \leq 2 - \lambda_2(t) \perp \lambda_3(t) \geq 0 \end{cases} \tag{36}$$

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 Fig. 7 and Fig. 8. In these figures, we consider the particular case of

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

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 Fig. 9, we show the error term $\|Y(T) - Y_k\|$ in function of the step size h. The term Y_k is the numerical approximation of $Y(T)$ by the minimal implicit Euler numerical scheme when the interval $[t_0, T]$ is subdivided in k 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 steps are taken equally spaced in log-scale. We observe the linear convergence rate of the implicit Euler scheme when there is uniqueness of the numerical solutions. In addition, we also observe a linear convergence rate 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 9 on the curve associated with the sliding-repulsive case. Finally, we also test a variant of the studied example

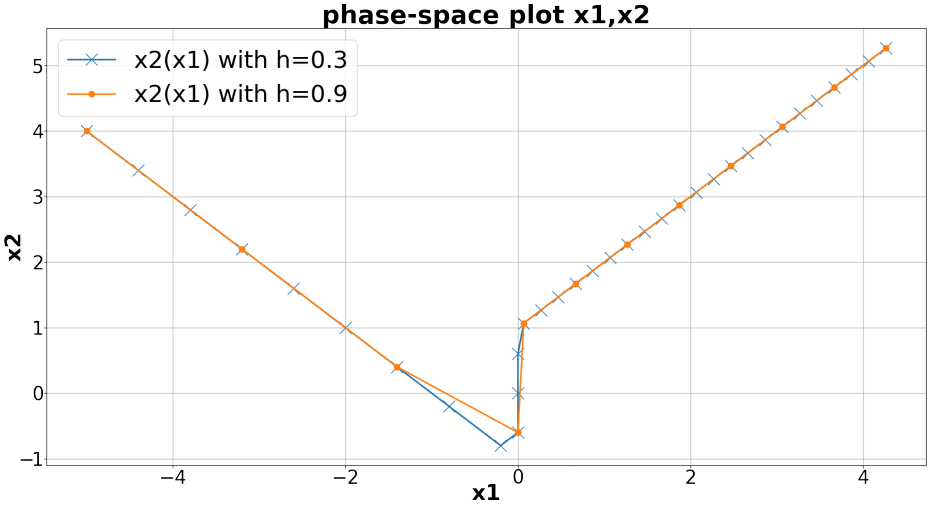


Fig. 7: Phase space plot in (x_1, x_2) of the numerical solutions for $\mathrm{B} = (-0.5, 1)$ and $h = 0.9$ or $h = 0.3$. Initial condition is $\mathbf{x}_0 = (-5, 4)$

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

$$\mathrm{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} -1 \\ 0.5 \end{pmatrix}. \tag{37}$$

Some Results are exposed in Figures 10, 11, and convergence on some experiments is also in $O(h)$ as it can be seen on Figure 9.

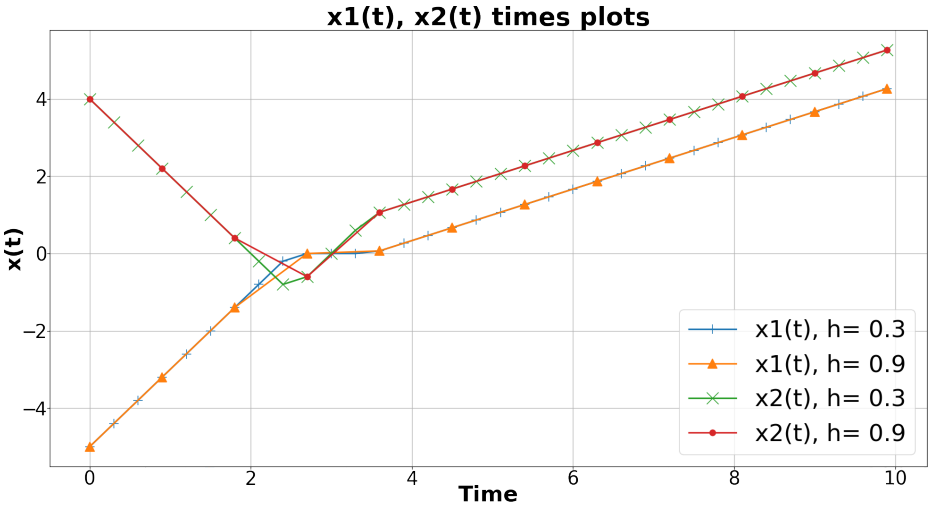


Fig. 8: Time plots of the solutions $x_1(t)$ and $x_2(t)$ for B $= (-0.5, 1)$ and $h = 0.9$ or $h = 0.3$. Initial condition is $\mathbf{x}_0 = (-5, 4)$

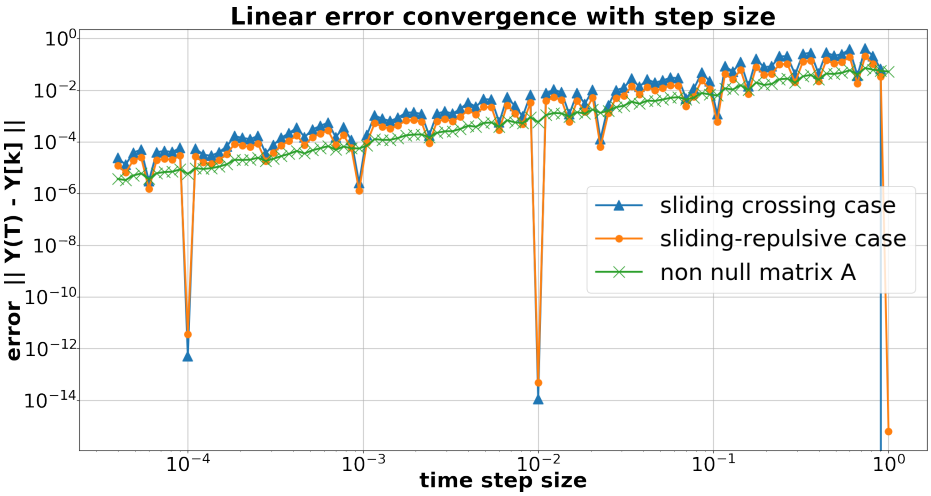


Fig. 9: Error $\|Y(t) - Y_k\|$ with respect to time step h. We consider the two kind of AC solutions: the sliding-crossing solutions and the sliding-repulsive solutions.

4 Conclusion and Future Works

In a first time, we analysed the AC solution of a 2D example of hybrid DAE. We show that in the context of piecewise linear constraints, we can observe multiplicity of AC solutions. Furthermore, it is not enough to study each mode independently to conclude on the well-posedness of AC solutions. In a second

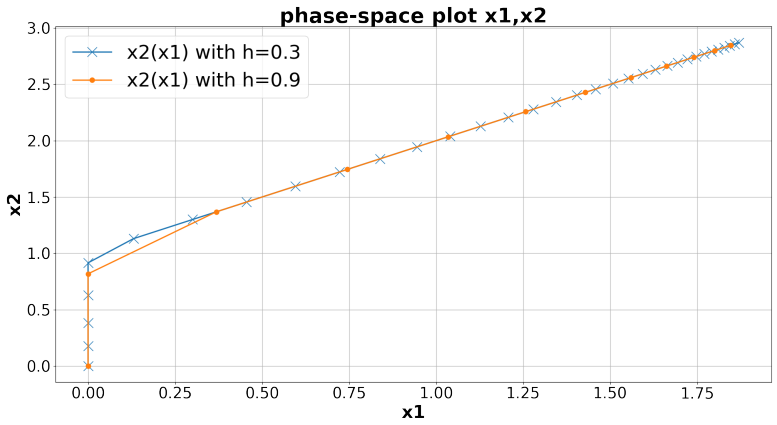


Fig. 10: Phase space plot in (x_1, x_2) of the numerical solutions for $\mathrm{B} = (-1, 0.5)$, A given in (37), and $h = 0.9$ or $h = 0.3$. Initial condition is $\mathbf{x}_0 = (0, 0)$.

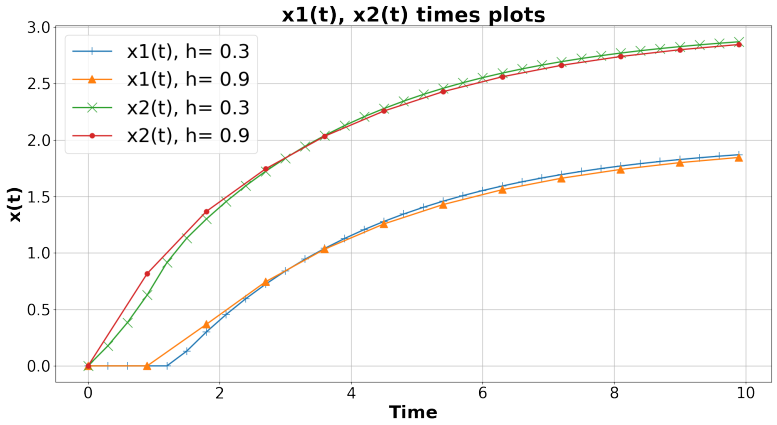


Fig. 11: Time plots of the solutions $x_1(t)$ and $x_2(t)$ for $\mathrm{B} = (-1, 0.5)$, A given in (37), and $h = 0.9$ or $h = 0.3$. Initial condition is $\mathbf{x}_0 = (0, 0)$.

time, we studied the generalised equation resulting from the example jump dynamic. We conclude on the conditions for well-posedness of such equation, and build a framework for the study of numerical solutions. Indeed, in the two last sections, we show that solutions of implicit Euler scheme, which is classically used as an event-capturing scheme for non-smooth dynamical systems, are solutions of an equation with the same structure as the jump dynamics. It follows that such numerical scheme can have either none, a unique, several, or infinitely many solutions depending of small variations on the considered problem. In particular, consistency of the numerical scheme is not preserved in some cases. However, on this example it can be proven that a "correct" discrete solution can always be retrieved from the numerous solutions of the implicit Euler scheme. Consequently, we propose a minimal implicit Euler numerical scheme to select the correct solutions assuming a time step sufficiently small. In future work, we will extend these results and observations to more general dynamics and switching constraint. Another, interesting research direction would be to make the link with [10].

Acknowledgement

We gratefully acknowledge the support of Inria Project Lab (IPL) as well as the 'Fonds unique interministériel" (FUI) through the Modeliscale projects.

References

1. Acary, V., Bonnefon, O., and Brogliato, B. *Nonsmooth modeling and simulation for switched circuits*, vol. 69. Springer Science & Business Media, 2010.
2. Acary, V., and Brogliato, B. *Numerical methods for nonsmooth dynamical systems: applications in mechanics and electronics*. 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* (2017), ACM, pp. 253–263.
8. Brenan, K. E., Campbell, S. L., and Petzold, L. R. *Numerical solution of initial-value problems in differential-algebraic equations*, vol. 14. Siam, 1996.
9. Brogliato, B., and Brogliato, B. *Nonsmooth mechanics*. Springer, 1999.

10. 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.
11. Facchinei, F., and Pang, J.-S. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.
12. Filippov, A. F. Differential equations with discontinuous right-hand side. *Matematicheskii sbornik 93*, 1 (1960), 99–128.
13. 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.
14. 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.
15. 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.
16. Kleinert, J., and Simeon, B. Differential-algebraic equations and beyond: From smooth to nonsmooth constrained dynamical systems. *arXiv preprint arXiv:1811.07658* (2018).
17. Lamour, R., März, R., and Tischendorf, C. *Differential-algebraic equations: a projector based analysis*. Springer Science & Business Media, 2013.
18. Marques, M. *Differential inclusions in nonsmooth mechanical problems: Shocks and dry friction*, vol. 9. Birkhäuser, 2013.
19. Matrosov, I. V. Existence of solutions of the algebro-differential equations. *Automation and Remote Control 67*, 9 (2006), 1408–1415.
20. 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.
21. Mehrmann, V., and Wunderlich, L. Hybrid systems of differential-algebraic equations–analysis and numerical solution. *Journal of Process Control 19*, 8 (2009), 1218–1228.
22. Pang, J.-S., and Stewart, D. E. Differential variational inequalities. *Mathematical Programming 113*, 2 (2008), 345–424.
23. Rockafellar, R. T., and Wets, R. J.-B. *Variational analysis*, vol. 317. Springer Science & Business Media, 2009.
24. Stechlinski, P., Patrascu, M., and Barton, P. I. Nonsmooth differential-algebraic equations in chemical engineering. *Computers & Chemical Engineering 114* (2018), 52–68.
25. Trenn, S. Switched differential algebraic equations. In *Dynamics and Control of Switched Electronic Systems*. Springer, 2012, pp. 189–216.
26. 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.