On preserving dissipativity properties of linear complementarity dynamical systems with the θ-method

Scott Greenhalgh, Vincent Acary, Bernard Brogliato

On preserving dissipativity properties of linear complementarity dynamical systems with the θ-method

Scott Greenhalgh · Vincent Acary · Bernard Brogliato

Abstract In this work we study the following problem: given a numerical method (an extended θ-method named the (θ, γ)-method), find the class of dissipative linear complementarity systems such that their discrete-time counterpart is still dissipative, with the same storage (energy) function, supply rate (reciprocal variables), and dissipation function. Systems with continuous solutions, and with state jumps are studied. The notion of numerical dissipation is given a rigorous meaning.

1 Introduction

Due to the importance of the notion of dissipativity in automatic control applications [9] (dissipativity being the positive real property of transfer functions in the case

---

S. Greenhalgh was supported by French National Research Agency (ANR) through COSINUS program (project SALADYN ANR-08-COSI-014).

S. Greenhalgh
INRIA Grenoble Rhône-Alpes, Bipop project-team, Inovallée de Montbonnot, 655 avenue de l'Europe, Saint Ismier Cedex 38334, France
e-mail: scott.greenhalg@yale.edu

S. Greenhalgh
Department of Mathematics and Statistics, University of Guelph, Guelph, France

V. Acary (✉) · B. Brogliato
INRIA, Bipop team-project, Inovallée de Montbonnot, 655 avenue de l'Europe, Saint Ismier Cedex 38334, France
e-mail: vincent.acary@inria.fr

B. Brogliato
e-mail: bernard.brogliato@inria.fr

of linear, time-invariant systems), the preservation of dissipativity properties (or of the positive realness) after time-discretization has been studied for a long time, see e.g. [15,17,19,20,26,28,29,35]. In the above works the question usually answered is: given a positive real system, perform a time-discretization (Euler, or zero order hold) and examine whether the obtained discrete-time system is still positive real, possibly with different storage function, supply rate and dissipation function. For instance four different types of discretizations are studied in [28]. Whether or not the continuous-time and the discrete-time systems possess the same storage function or the same dissipation function, is tackled in [15,19,20,36]. Recently the interest has focused on dissipativity of nonsmooth dynamical systems like dynamical complementarity systems [10–14,21,22,24,27], hybrid systems [6], and multivalued Lur'e systems [7,8].

In this paper we deal with linear complementarity dynamical systems, possibly with state jumps. We deal with preservation of passivity (in Willems' sense [9]) after discretization by an extended θ-method called the (θ, γ)-algorithm. In view of the state of the art on time discretization of such nonsmooth systems, higher order methods are not yet available and only first order methods (implicit or explicit Euler, trapezoidal rule [3]) have been shown to converge. Extensions towards higher-order methods is an open issue, not tackled in this paper (see [1,39] for some preliminary work in the field of nonsmooth mechanical systems). This means that Runge–Kutta, multi-step methods are outside the scope of this study. The problem that is tackled in this paper is as follows:

Given a discretization method, find the class of linear complementarity dissipative systems such that their discretized counterpart is still dissipative with the same storage function set, supply rate and dissipation function.

In addition the method, when applied to complementarity dynamical system, should guarantee that the so-called one-step-nonsmooth-problem to be solved at each time step, possesses a unique solution, and, in case the solution jumps, that the energetic properties of the jump rule are preserved. Usually, all this yields quite stringent conditions and narrow classes of continuous-time systems, and may be seen as counterpart of the problem tackled in [15,19,20] which is: find a discretization method such that any dissipative system is transformed into a dissipative discrete-time system. Finally we do not want to stick to the conservative (or lossless) case, since it is desirable to deal with systems that possess a non-zero dissipation function and to seek conditions under which the dissipation function is also preserved.

The case of linear complementarity systems (LCS) without state jumps is first dealt with, and then we focus on state jumps. In this paper we are not interested in convergence results as the time-step goes to zero, but on the algorithm properties when $h > 0$. The paper is organized as follows. In Sect. 2 the continuous-time and the discrete-time systems are presented, the definitions of dissipativity are recalled, and a definition of numerical dissipation is given. Section 3 is dedicated to the study of the conditions such that dissipativity is preserved after the discretization. In Sect. 4 we examine whether the numerical method consistently approximates state jumps. Conclusions are given in Sect. 5 and some technical details are provided in the appendix. Academic and physical examples (electrical circuits with multivalued nonsmooth components) are used throughout the paper to illustrate the theoretical developments.

All the numerical results have been obtained with the SICONOS platform[1] of INRIA, see [2,3,5].

Notation The right and left limits of a function f at t are denoted as $f(t^+)$ and $f(t^-)$ respectively. The normal cone to a convex non-empty set $K \subseteq \mathbb{R}^n$ at $x \in K$ is $N_K(x) = \{v \in \mathbb{R}^n \mid \langle v, z - x\rangle \leqslant 0 \text{ for all } z \in K\}$. The projection of $x \in \mathbb{R}^n$ on K in the metric defined by a symmetric positive definite matrix M is denoted as $\mathrm{proj}_M[K; x]$. Given a matrix $M \in \mathbb{R}^{n\times n}$ and a vector $q \in \mathbb{R}^n$, a Linear Complementarity Problem (LCP) with unknown $\lambda \in \mathbb{R}^m$ is a problem of the form $\lambda \geqslant 0$, $M\lambda + q \geqslant 0$, $\lambda^T(M\lambda + q) = 0$, written compactly as $0 \leqslant \lambda \perp M\lambda + q \geqslant 0$. This is denoted LCP$(q, M)$, and the set of solutions is SOL(q, M). Let $K \subseteq \mathbb{R}^m$ be a convex non-empty closed cone, its dual cone is the set $K^* = \{v \in \mathbb{R}^m \mid v^T z \geqslant 0 \text{ for all } z \in K\}$. A linear cone CP (LCCP) is a problem of the form $K \ni \lambda \perp M\lambda + q \in K^*$. Ker$(A)$ is the kernel of the matrix A. A positive semi definite (PSD) matrix M, possibly non-symmetric, is such that for all $x \in \mathbb{R}^n$ one has $x^T M x \geqslant 0$. It is positive definite if $x^T M x > 0$ for all $x \neq 0$. The matrix I is the identity matrix with appropriate dimension. For any matrix $A \in \mathbb{R}^{n\times n}$, let us recall that $I + \eta A$ is full-rank for a sufficiently small $\eta \in \mathbb{R}$ and that $(I + \eta A)(I - \mu A) = (I - \mu A)(I + \eta A)$ for any η and $\mu \in \mathbb{R}$.

2 The dynamical system and its discretization

In this section and in Sect. 3 we deal with the case without state jumps.

2.1 Continuous-time systems: the dynamics passivity conditions

We consider the following LCS:

$$\begin{cases} \dot{x}(t) = Ax(t) + B\lambda(t) + Eu(t) \\ w(t) = Cx(t) + D\lambda(t) + Fv(t) \\ 0 \leqslant \lambda(t) \perp w(t) \geqslant 0 \\ x(0^-) = x_0, \end{cases} \tag{1}$$

with $x(t) \in \mathbb{R}^n$, $\lambda(t) \in \mathbb{R}^m$, $w(t) \in \mathbb{R}^m$. In general one may have $n \geqslant m$ or $m \geqslant n$, depending on the application. The vector $x_0 \in \mathbb{R}^n$ denotes the given initial conditions. The well-posedness (existence and uniqueness of solutions) of such systems has been studied. Depending on the data solutions may be continuous, of class C^1, discontinuous functions, measures, or distributions, see *e.g.* [4,8,10,12,14,24,40]. A general assumption throughout the paper is that $v(\cdot)$ and $u(\cdot)$ are bounded functions of time. In the first part we suppose that the solutions are absolutely continuous consequently the first equality in (1) is satisfied almost everywhere.

[1] http://siconos.gforge.inria.fr

Definition 1 (*Passivity properties and energy storage function. Continuous-time case*) The quadruple (A, B, C, D) is said to be *passive* if there exist matrices $L \in \mathbb{R}^{n\times m}$ and $W \in \mathbb{R}^{m\times m}$ and a symmetric positive semi-definite matrix $P \in \mathbb{R}^{n\times n}$, such that:

$$\begin{cases} A^T P + PA = -LL^T & (2) \\ B^T P - C = -W^T L^T & (3) \\ -D - D^T = -W^T W. & (4) \end{cases}$$

In this case, let $V(x) = \frac{1}{2}x^T Px$ denote the corresponding *energy storage function*. The *dissipation equality*

$$V(x(T)) - V(x(0)) = -\frac{1}{2}\int_0^T (x^T(t), \lambda^T(t))Q\begin{pmatrix} x(t) \\ \lambda(t)\end{pmatrix} dt, \quad \forall\ T \geqslant 0 \tag{5}$$

in terms of the positive semi-definite matrix

$$Q \triangleq \begin{pmatrix} LL^T & W^T L^T \\ LW & W^T W \end{pmatrix}, \tag{6}$$

then implies that

$$V(x(T)) - V(x(0)) \leqslant 0. \tag{7}$$

The system is said to be *strictly passive* when Q is positive definite, and *lossless* when $Q = 0$. The system is said to be *state lossless* when $L = 0$ and *input lossless* when $W = 0$. The system is *dissipative*, *state dissipative*, and *input dissipative* when $Q \neq 0$, $L \neq 0$, or $W \neq 0$, respectively. In particular, we have

$$V(x(T)) - V(x(0)) \leqslant S(\lambda(t), w(t)), \tag{8}$$

where the supply rate $S(\lambda, w) \triangleq \lambda^T w$, since the LCS implies that $S(\lambda(t), w(t)) = 0$ for all $t \geqslant 0$.

The *infinitesimal* dissipation inequality writes as

$$\dot{V}(x(t)) = -\frac{1}{2}(x^T(t), \lambda^T(t))Q\begin{pmatrix} x(t) \\ \lambda(t)\end{pmatrix}, \tag{9}$$

which is equivalent to (5) as long as $x(\cdot)$ is differentiable or absolutely continuous (hence with a derivative almost everywhere).

Definition 2 (*Cumulative dissipation function. Continuous–time case*) We define the (continuous) cumulative dissipation function as

$$\mathcal{D}(t) \triangleq \int_0^t \frac{1}{2}(x^T(s), \lambda^T(s)) Q \begin{pmatrix} x(s) \\ \lambda(s) \end{pmatrix} ds. \tag{10}$$

As the next example shows, allowing for $P \geqslant 0$ in (2) is important because if the pair (A, C) is not observable the set of Eqs. (2)–(4) may possess positive semi definite solutions.

Example 1 Consider (A, B, C, D) defined as:

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \ C = (0 \cdots 0 \ \ 1), \ D = 0. \tag{11}$$

with $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times 1}, C \in \mathbb{R}^{1\times n}$. Suppose we want to characterize the lossless property $L = 0$. Enforcing $A^T P + PA = 0$ and $PB = C^T$, implies $p_{i,j} = \begin{cases} 1 & \text{for } (i, j) = (n, n) \\ 0 & \text{o.w.} \end{cases}$, and hence $P = \begin{pmatrix} 0_{n-1,n-1} & 0_{n-1} \\ 0_{n-1}^T & 1 \end{pmatrix}$ where the notation $0_{n,n}$ and 0_n is used to represent zero matrices and zero vectors of dimensions $n \times n$ and n respectively. Consequently (2)–(4) for the lossless case does not possess any positive definite solution, but only PSD solutions. The system is not observable from the "output" signal w since $CA = 0$. Recall that observability implies $P > 0$ in (2)–(4), because the kernel of P satisfying this set of relations is a subset of the unobservability space of (C, A) [24, Lemma 2].

Remark 1 In the following we shall sometimes write that the system in (1) is passive, by which it is abusively meant that the quadruple (A, B, C, D) is passive. Both coincide in case $u(\cdot) = v(\cdot) = 0$.

2.2 The time-discretization

In this section we present the discretization of (1) with $u(\cdot) = v(\cdot) = 0$. Considering non-zero exogenous functions $u(\cdot)$ and $v(\cdot)$ does not create much difficulty for the discretization, however we are interested in preserving the passivity properties of the quadruple (A, B, C, D) and the exogenous terms play no role.

2.2.1 The (θ, γ)-method

Let us propose the following (θ, γ)-method for (1):

$$\begin{cases} \dfrac{x_{k+1} - x_k}{h} = A x_{k+\theta} + B\lambda_{k+\gamma} \\ 0 \leqslant \lambda_{k+\gamma} \perp w_{k+\gamma} = C x_{k+\gamma} + D\lambda_{k+\gamma} \geqslant 0 \\ x_0 = x_0, \end{cases} \tag{12}$$

where θ and $\gamma \in [0, 1]$. The subscript notation $k+\theta$ means $x_{k+\theta} = \theta x_{k+1} + (1-\theta)x_k$ and similarly $\lambda_{k+\gamma} = \gamma\lambda_{k+1} + (1-\gamma)\lambda_k$. By specifying (θ, γ) we completely characterize the form of the discretization: $(\theta, \gamma) = (1, 1)$ is a fully implicit scheme, $(\theta, \gamma) = (0, 1)$ is a semi-implicit scheme, and $(\theta, \gamma) = (0, 0)$ is a fully explicit scheme. Assuming that the inverse $(I - h\theta A)^{-1}$ is well defined (a sufficient condition is $h < \frac{1}{\theta ||A||}$ where $|| \cdot ||$ is a norm for which $||I|| = 1$ [30, Theorem 1, Chapter 11], but in many cases $I - h\theta A$ may be full rank for $h > 0$ not necessarily small), we define

$$\begin{cases} \widetilde{A}(h, \theta) = (I - h\theta A)^{-1}(I + h(1-\theta)A) \\ \widetilde{B}(h, \theta) = h(I - h\theta A)^{-1} B \\ \widetilde{C}(h, \theta, \gamma) = \gamma C \widetilde{A}(h, \theta) + (1-\gamma)C \\ \widetilde{D}(h, \theta, \gamma) = \gamma C \widetilde{B}(h, \theta) + D. \end{cases} \tag{13}$$

Where there is no ambiguity, we further note that $\widetilde{A} = \widetilde{A}(h, \theta)$, $\widetilde{B} = \widetilde{B}(h, \theta)$, $\widetilde{C} = \widetilde{C}(h, \theta, \gamma)$ and $\widetilde{D} = \widetilde{D}(h, \theta, \gamma)$ to simplify the notation.

The (θ, γ)-discretization of the LCS is compactly written as:

$$\begin{cases} x_{k+1} = \widetilde{A} x_k + \widetilde{B}\lambda_{k+\gamma} \\ w_{k+\gamma} = \widetilde{C} x_k + \widetilde{D}\lambda_{k+\gamma} \\ 0 \leqslant \lambda_{k+\gamma} \perp w_{k+\gamma} \geqslant 0 \\ x_0 = x_0. \end{cases} \tag{14}$$

One can infer directly from (13) that it is necessary that $\gamma > 0$ when $D = 0$. Indeed a discrete-time system that is passive has a non-zero feedthrough matrix [9].

2.2.2 The passivity conditions in the discrete-time case

Definition 3 (*Passivity properties and energy storage function. Discrete-time case*) The quadruple $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ is said to be *passive* if there exists matrices $\widetilde{L} \in \mathbb{R}^{n\times m}$ and $\widetilde{W} \in \mathbb{R}^{m\times m}$ and a symmetric positive semi-definite matrix $R \in \mathbb{R}^{n\times n}$, such that:

$$\begin{cases} \widetilde{A}^T R \widetilde{A} - R = -\widetilde{L}\widetilde{L}^T & (15) \\ \widetilde{B}^T R \widetilde{A} - \widetilde{C} = -\widetilde{W}^T \widetilde{L}^T & (16) \\ \widetilde{B}^T R \widetilde{B} - \widetilde{D} - \widetilde{D}^T = -\widetilde{W}^T \widetilde{W}. & (17) \end{cases}$$

In this case, let $V(x_k) = \frac{1}{2} x_k^T R x_k$ denote the corresponding *energy storage function*. The *dissipation equality*

$$V(x_{k+1}) - V(x_k) = -\frac{1}{2}(x_k^T, \lambda_{k+\gamma}^T)\widetilde{Q}\begin{pmatrix} x_k \\ \lambda_{k+\gamma} \end{pmatrix}, \tag{18}$$

or equivalently

$$V(x_{k+1}) - V(x_0) = -\frac{1}{2}\sum_{i=0}^{k}(x_i^T, \lambda_{i+\gamma}^T)\widetilde{Q}\begin{pmatrix} x_i \\ \lambda_{i+\gamma} \end{pmatrix}, \tag{19}$$

in terms of the positive semi-definite matrix $\widetilde{Q} = \begin{pmatrix} \widetilde{L}\widetilde{L}^T & \widetilde{W}^T\widetilde{L}^T \\ \widetilde{L}\widetilde{W} & \widetilde{W}^T\widetilde{W} \end{pmatrix}$, and for all $k \geqslant 0$, then implies that

$$V(x_{k+1}) - V(x_k) \leqslant 0 \ \ (\Rightarrow V(x_{k+1}) - V(x_0) \leqslant 0), \tag{20}$$

for all $k \geqslant 0$. The system is said to be *strictly passive* when $\widetilde{Q}$ is positive definite, and *lossless* when $\widetilde{Q} = 0$. The system is said to be *state lossless* when $\widetilde{L} = 0$ and *input lossless* when $\widetilde{W} = 0$. The system is *dissipative*, *state dissipative*, and *input dissipative* when $\widetilde{Q} \neq 0$, $\widetilde{L} \neq 0$, or $\widetilde{W} \neq 0$, respectively. In particular, we have

$$V(x_{k+1}) - V(x_0) \leqslant \sum_{i=0}^{k} S(\lambda_{i+\gamma}, w_{i+\gamma}), \tag{21}$$

and

$$V(x_{k+1}) - V(x_k) \leqslant S(\lambda_{k+\gamma}, w_{k+\gamma}), \tag{22}$$

where the supply rate $S(\lambda_{k+\gamma}, w_{k+\gamma}) \stackrel{\Delta}{=} \lambda_{k+\gamma}^T w_{k+\gamma}$, since the LCS implies that $S(\lambda_{k+\gamma}, w_{k+\gamma}) = 0$ for all $k \geqslant 0$. Note that the dependence on h, θ and γ of $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$ and $\widetilde{D}$ implies that $\widetilde{Q}$

Definition 4 (*Cumulative dissipation function. Discrete-time case*) We define the (discrete) cumulative dissipation function as:

$$\mathcal{D}_k \stackrel{\Delta}{=} \sum_{i=0}^{k} \frac{h}{2}(x_i^T, \lambda_{i+\gamma}^T)\widetilde{Q}\begin{pmatrix} x_i \\ \lambda_{i+\gamma} \end{pmatrix}. \tag{23}$$

2.2.3 Two options for the same objective

The occurrence of an integral symbol in the dissipation inequality (5) and the sum symbol in the dissipation inequality in discrete-time (19) does not make possible to have both the same storage function and the same dissipation function. One objective can be that both systems (continuous-time and discrete-time) possess the same storage function and the other is that both systems possess the same dissipation function. Two options are possible:

1. A first option is to consider to exactly have the same storage function and a first order approximation of the cumulative dissipation function, that is
$$P = R \quad \text{and} \quad \widetilde{Q} = hQ. \tag{24}$$
In this first case, recalling that $\lambda_{k+\gamma}^T w_{k+\gamma} = 0$, the dissipation inequality (18) reads as
$$\frac{1}{2}x_{k+1}Px_{k+1} - \frac{1}{2}x_k Px_k = -\frac{h}{2}(x_k^T \ \lambda_{k+1}^T)Q\begin{pmatrix} x_k \\ \lambda_{k+1} \end{pmatrix}, \tag{25}$$
Comparing (5), (18) and (25), we note that the left-hand sides of the dissipation inequalities are conserved but the right-hand sides differ by a first order approximation of the integral.
2. A second option is to consider to exactly have the dissipation function and a first order approximation of the time derivative of the storage function, that is
$$hR = P \quad \text{and} \quad \widetilde{Q} = Q. \tag{26}$$
In this second case, the dissipation inequality (18) reads as
$$\frac{\frac{1}{2}x_{k+1}Px_{k+1} - \frac{1}{2}x_k Px_k}{h} = -\frac{1}{2}(x_k^T \ \lambda_{k+1}^T)Q\begin{pmatrix} x_k \\ \lambda_{k+1} \end{pmatrix}. \tag{27}$$
One sees that the second option has the form of the approximation of the continuous-time storage function derivative, with the instantaneous dissipation [hence it approximates the infinitesimal dissipation inequality (9)], whereas the first option rather approximates the integral form (5) of the passivity equality.

Both options are equivalent and we choose arbitrarily the second one in the sequel, *i.e.* $hR = P$ and $\widetilde{Q} = Q$. Choosing $P = hR$ or $R = P$ does not change much the analysis. Note that due to the fact that $\lambda_{k+\gamma}^T w_{k+\gamma} = 0$ and since the reciprocal variables belong to a cone, the supply rate may be scaled by any positive constant without changing the system's dissipativity properties.

2.2.4 The equivalence lemma

In the next lemma, we state equivalent formulations of the passivity conditions (2)–(4) if $hR = P$.

Lemma 1 *Let $hR = P$.*

(i) *Suppose that* (2) *is satisfied. Then, the equality (15) is equivalent to*
$$-LL^T + h^2(1-2\theta)A^T RA = -(I - h\theta A)^T \widetilde{L}\widetilde{L}^T(I - h\theta A) \tag{28}$$

(ii) *Suppose that* (2) *and* (3) *are satisfied. Then, the equality* (15) *is equivalent to*

$$
\begin{aligned}
&hB^T R\big(h(1-\theta-\gamma)A - h^2\theta(\gamma-\theta)A^2\big) - \theta B^T (I - h\theta A)^{-T} LL^T (I + h(1-\theta)A) \\
&\quad = W^T L^T (I + h(\gamma-\theta)A)(I + h\theta A) - \widetilde{W}^T \widetilde{L}^T (I - h\theta A)(I + h\theta A).
\end{aligned} \tag{29}
$$

(iii) *Suppose that* (2)–(3) *are satisfied. Then, the equality* (15) *is equivalent to:*

$$
\begin{aligned}
&h^2 B^T (I - h\theta A)^{-T}((1-2\gamma)R - \theta\gamma LL^T)(I - h\theta A)^{-1} B \\
&\quad = W^T W - \widetilde{W}^T \widetilde{W} + h\gamma W^T L^T (I - h\theta A)^{-1} B + h\gamma B^T (I - h\theta A)^{-T} LW.
\end{aligned} \tag{30}
$$

The proof is given in Appendix B.

2.2.5 The numerical dissipation

Let us focus on the dissipation functions, *i.e.* the quadratic forms with PSD matrices Q and $\widetilde{Q}$.

Definition 5 (*Numerical dissipation*) The numerical algorithm is said to produce:

- Numerical over-dissipation (NOD) if $Q < \widetilde{Q}$.
- Numerical under-dissipation (NUD) if $Q > \widetilde{Q}$.
- Numerical equal-dissipation (NED) if $Q = \widetilde{Q}$.
- Numerical indefinite-dissipation (NID) if $Q - \widetilde{Q}$ is not a definite matrix.

Usually one says that a scheme does not dissipate energy when it is of the NED type: for instance if the continuous-time system is lossless, the discrete-time system is lossless as well. We may refine Definition 5 by treating separately the state dissipation (governed by LL^T and $\widetilde{L}\widetilde{L}^T$) and the input dissipation (governed by $W^T W$ and $\widetilde{W}^T \widetilde{W}$). We may then define the notion of numerical under state dissipation (NUSD), numerical indefinite state dissipation (NISD), numerical under input dissipation (NUID), numerical equal input dissipation (NEID), numerical over input dissipation (NOID), *etc*. The study in [36] aims at characterizing such properties for the zero order hold discretization method.

2.2.6 The one-step nonsmooth problem

At each step advancing the algorithm (14) (equivalently 12) boils down to solving the following linear complementarity problem (LCP):

$$
0 \leqslant \lambda_{k+\gamma} \perp w_{k+1} = \widetilde{C}x_k + \widetilde{D}\lambda_{k+\gamma} \geqslant 0. \tag{31}
$$

From a classical result [16] this LCP has a unique solution for any $\widetilde{C}x_k$, if and only if $\widetilde{D}$ is a P-matrix (*i.e.* all its principal minors are positive). Passivity of the continuous-time system is known to be a crucial property for the discrete-time LCP well-posedness (see [11, Lemma 24] when $\theta = \gamma = 1$, see also [24,37] for $\theta = 0$, $\gamma = 1$). It is obvious from (13) that when D is a P-matrix, so is $\widetilde{D}$ for small enough h or γ.

3 Preservation of passivity properties after discretization

In this section we present conditions for the preservation of the passivity properties after discretization with the (θ, γ)-method. In other words, we assume that both the continuous and the discrete-time systems are passive, and we seek conditions such that they possess the same energy storage function (more precisely the same first order approximation of its time derivative, *i.e* $hR = P$), same dissipation functions ($\widetilde{Q} = Q$), with same supply rates. The problem that is tackled here is thus more stringent than just preserving the passivity after discretization without further constraints, see Remark 2. We first analyze the consequences of time-discretization, on the energy storage function and on the state dissipation preservation.

Let us rewrite (28) by developing its right-hand side

$$-LL^T + h^2(1-2\theta)A^T RA = -\widetilde{L}\widetilde{L}^T + h\theta(\widetilde{L}\widetilde{L}^T A + A^T \widetilde{L}\widetilde{L}^T) - h^2\theta^2 A^T \widetilde{L}\widetilde{L}^T A. \tag{32}$$

Unfortunately, since $\widetilde{L}$ depends on h, it is not possible to directly equating the coefficient of the same power of h in order to obtain necessary and sufficient conditions for the the preservation of both the energy storage function (*i.e.* $hR = P$) and the state dissipation function (*i.e.* $LL^T = \widetilde{L}\widetilde{L}^T$) after discretization. The following result aims at bridging this gap.

Proposition 1 *Let $hR = P$. Suppose further that (2) and (15) are satisfied. Then we have for all $h > 0$*

$$LL^T = \widetilde{L}\widetilde{L}^T \iff \begin{cases} \theta A^T LL^T = 0 \\ (2\theta - 1)A^T RA = 0. \end{cases} \tag{33}$$

Proof Note that if (2) and (15) are satisfied, then from Lemma 1 we know that (28) (equivalently (32)) holds. Let $LL^T = \widetilde{L}\widetilde{L}^T$ hold for all $h > 0$, (28) implies

$$h[(1-2\theta)A^T RA + \theta^2 A^T LL^T A] = \theta[LL^T A + A^T LL^T]. \tag{34}$$

For (34) to hold for any $h > 0$ one has to nullify the coefficient of the polynomial in h. Then we get

$$LL^T = \widetilde{L}\widetilde{L}^T \implies \begin{cases} \theta LL^T A = -(\theta LL^T A)^T \\ (2\theta - 1)A^T RA = \theta^2 A^T LL^T A. \end{cases} \tag{35}$$

Let us split the proof with the values of θ, *i.e.* (a) $\theta \in (0, 1]$, $\theta \neq 1/2$, (b) $\theta = 1/2$ and (c) $\theta = 0$.

Case (a) Let $LL^T = \widetilde{L}\widetilde{L}^T$ hold for all $h > 0$ and (35) for all $\theta \in (0, 1]$. Then $\theta \neq 1/2$ yields

$$LL^T = \widetilde{L}\widetilde{L}^T \implies \begin{cases} LL^T A = -(LL^T A)^T \\ A^T RA = 0 \\ A^T LL^T A = 0. \end{cases} \tag{36}$$

Since $A^T LL^T A = 0$ is equivalent to $A^T L = L^T A = 0$, the condition $\theta A^T LL^T = 0$ is satisfied. The implication in (33) is then proven.

Case (b) $\theta = 1/2$. Let $LL^T = \widetilde{L}\widetilde{L}^T$ hold for all $h > 0$. Then Eq. (35) for $\theta = 1/2$ implies $A^T LL^T A = 0$. This implies that $A^T L = 0$, and therefore $A^T LL^T = A^T \widetilde{L}\widetilde{L}^T = 0$. The implication in (33) is then satisfied.

Case (c) $\theta = 0$. Let $LL^T = \widetilde{L}\widetilde{L}^T$ hold for all $h > 0$, (35) for $\theta = 0$ implies $A^T RA = 0$. The implication in (33) is then satisfied.

Conversely, let us suppose that $\theta A^T LL^T = 0$ and $(2\theta - 1)A^T RA = 0$. Equation (28) (equivalently (32)) implies

$$-LL^T = -(I - h\theta A)^T \widetilde{L}\widetilde{L}^T (I - h\theta A). \tag{37}$$

Note that if $\theta A^T LL^T = 0$, we have $LL^T = (I - h\theta A)^T LL^T (I - h\theta A)$. From (37), we conclude that $LL^T = \widetilde{L}\widetilde{L}^T$. □

Remark 2 (Passivity preservation) As alluded to above, the problem of finding a discretization that is still passive but without taking care of whether it has the same energy storage, dissipation and supply rates as the continuous-time system, is a different problem than the one dealt with in this paper. Suppose that $I - h\theta A$ has full rank n. Let $\widetilde{R} = (I - h\theta A)^{-T} R (I - h\theta A)^{-1} \in \mathbb{R}^{n\times n}$. Then $R \geqslant 0 \Leftrightarrow \widetilde{R} \geqslant 0$. Using (13) and after some lengthy but easy manipulations we may equivalently rewrite (15)–(17) as follows:

$$\begin{cases} h(A^T\widetilde{R} + \widetilde{R}A) + h^2(1 - 2\theta)A^T\widetilde{R}A = -\widetilde{L}\widetilde{L}^T & (38) \\ hB^T\widetilde{R}(I + h(1-\theta)A) - \widetilde{C} = -\widetilde{W}^T\widetilde{L}^T & (39) \\ h^2B^T\widetilde{R}B - \widetilde{D}^T - \widetilde{D} = -\widetilde{W}^T\widetilde{W} & (40) \end{cases}$$

We know from the passivity of (A, B, C, D) that the Lyapunov equation $(A^T(h\widetilde{R}) + (h\widetilde{R})A) = -\widetilde{L}\widetilde{L}^T$ has a unique solution $(h\widetilde{R})$ for given $\widetilde{L}\widetilde{L}^T$. Thus provided that $(1 - 2\theta)A^T\widetilde{R}A = 0$ (which is satisfied if $\theta = \frac{1}{2}$) the equality (38), that is equivalent to (15), has a solution $\widetilde{R}$ such that $R = (I - h\theta A)^T\widetilde{R}(I - h\theta A)$ which defines the energy storage function of the discretized system. The state dissipation is given by $\widetilde{L}\widetilde{L}^T$. Now taking $\widetilde{W} = 0$ one may rewrite (39) as $hB^T\widetilde{R} - \widetilde{C}(I + h(1-\theta)A)^{-1} = 0$, which means that the second equality for passivity is satisfied with a new output matrix $\frac{1}{h}\widetilde{C}(I + h(1-\theta)A)^{-1}$. Then (40) boils down to $B^T\widetilde{R}B = \widetilde{D}^T + \widetilde{D}$. Clearly one can always find $\widetilde{D}$ such that this equality holds, however it may not be equal to the matrix $\widetilde{D}$ in (13), so we denote it $\bar{D}$. Changing $\widetilde{D}$ into $\bar{D}$ once again modifies the "output" $w_{k+\gamma}$

in (14). Therefore the discrete-time system does not possess the output $w_{k+\gamma} = \widetilde{C}x_k + \widetilde{D}\lambda_{k+\gamma}$ in (14), but a new output equal to $\bar{w}_{k+\gamma} \triangleq \frac{1}{h}\widetilde{C}(I+h(1-\theta)A)^{-1}x_{k+1} + \bar{D}\lambda_{k+\gamma}$. This corresponds to changing the supply rate of the system. Therefore the discrete-time system is dissipative with storage function $\frac{h}{2}x_k^T R x_k$, dissipation matrices $\widetilde{L}$ and $\widetilde{W} = 0$, supply rate $\bar{w}_{k+\gamma}^T \lambda_{k+\gamma}$.

3.1 State losslessness preservation ($L = 0$)

It is noteworthy that usually what is referred to as a conservative system in the literature corresponds to having $L = 0$ solely (the state energy is constant along trajectories). The losslessness applies here only to the state, *i.e.* the LL^T term. From (2) to (4) with $L = 0$ one obtains:

$$A^T P + PA = 0,\ PB = C^T,\ D + D^T = W^T W,\ P = P^T \geqslant 0. \tag{41}$$

Proposition 2 *Let $hR = P$. Suppose that both (A, B, C, D) and $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ are passive with $L = 0$. Then we have*

$$\begin{cases} LL^T = \widetilde{L}\widetilde{L}^T \\ \widetilde{W}^T\widetilde{L}^T = W^T L^T \end{cases} \iff \begin{cases} (2\theta - 1)A^T RA = 0 \\ (1-\theta-\gamma)B^T RA = 0 \\ \theta(\gamma-\theta)B^T RA^2 = 0. \end{cases} \tag{42}$$

Let us suppose that $LL^T = \widetilde{L}\widetilde{L}^T$ and $\widetilde{W}^T\widetilde{L}^T = W^T L^T$, then we have

$$W^T W = \widetilde{W}^T \widetilde{W} \iff (1-2\gamma)B^T RB = 0. \tag{43}$$

Proof From Lemma 1(ii), Eq. (16) with $L = 0$ is equivalent to

$$hB^T R(h(1-\theta-\gamma)A - h^2\theta(\gamma-\theta)A^2) = -\widetilde{W}^T\widetilde{L}^T(I - h\theta A)(I + h\theta A). \tag{44}$$

Then, it follows from (44) and $L = 0$:

$$LL^T = \widetilde{L}\widetilde{L}^T \implies \widetilde{L}\widetilde{L}^T = 0 \implies \widetilde{L}^T = 0 \implies \begin{cases} (1-\theta-\gamma)B^T RA = 0 \\ \theta(\gamma-\theta)B^T RA^2 = 0. \end{cases} \tag{45}$$

Conversely, using (44), one deduces

$$\begin{cases} (1-\theta-\gamma)B^T RA = 0 \\ \theta(\gamma-\theta)B^T RA^2 = 0 \end{cases} \implies -\widetilde{W}^T\widetilde{L}^T(I - h\theta A)(I + h\theta A) = 0 \implies 0 = \widetilde{W}^T\widetilde{L}^T = W^T L^T. \tag{46}$$

From Proposition 1, (46) and (45) , we get the equivalence in (42).

From Lemma 1(iii) , Eq. (16) with $L = 0$ is equivalent to

$$h^2(1-2\gamma)B^T(I - h\theta A)^{-T} R(I - h\theta A)^{-1}B = W^T W - \widetilde{W}^T\widetilde{W}. \tag{47}$$

Since $A^T R + RA = 0$, we have $(I - h\theta A)^T R(I - h\theta A) = R + h^2\theta^2 A^T RA$, so that $(I - h\theta A)^{-T} R(I - h\theta A)^{-1} = R - h^2\theta^2(I - h\theta A)^{-T} A^T RA(I - h\theta A)^{-1}$. Then, (47) is equivalent to

$$h^2(1-2\gamma)B^T \left[R - h^2\theta^2(I - h\theta A)^{-T} A^T RA(I - h\theta A)^{-1}\right] B = W^T W - \widetilde{W}^T \widetilde{W}. \tag{48}$$

Let us rewrite the left-hand side of (48):

$$\begin{aligned} &B^T[R - h^2\theta^2(I - h\theta A)^{-T} A^T RA(I - h\theta A)^{-1}]B \\ &\quad = B^T[R - h^2\theta^2 A^T(I - h\theta A)^{-T} R(I - h\theta A)^{-1} A]B \\ &\quad = B^T[R - h^2\theta^2 A^T R(I + h\theta A)^{-1}(I - h\theta A)^{-1} A]B \\ &\quad = B^T[R + h^2\theta^2 RA(I + h\theta A)^{-1}(I - h\theta A)^{-1} A]B. \end{aligned} \tag{49}$$

To obtain the first equality we used the fact that $(I - h\theta A)^T A^T = (A(I - h\theta A))^T = (A - h\theta A^2)^T = ((I - h\theta A)A)^T = A^T(I - h\theta A)^T$. Then multiplying both sides by $(I - h\theta A)^{-T}$ one obtains that $A^T(I - h\theta A)^{-T} = (I - h\theta A)^{-T} A^T$. In a similar way $A(I - h\theta A)^{-1} = (I - h\theta A)^{-1} A$. Consequently $(I - h\theta A)^{-T} A^T RA(I - h\theta A)^{-1} = A^T(I - h\theta A)^{-T} R(I - h\theta A)^{-1} A$. Now from the fact that $A^T R + RA = 0$ one has $R(I + h\theta A) = (I - h\theta A)^T R$. Multiplying on the left by $(I + h\theta A)^{-1}$ we equivalently obtain that $R = (I - h\theta A)^T R(I + h\theta A)^{-1}$. Multiplying on the right by $(I - h\theta A)^{-T}$ we equivalently get $(I - h\theta A)^{-T} R = R(I + h\theta A)^{-1}$, which allows us to pass from the first to the second equality.

Inserting (49) into (48) yields

$$h^2(1 - 2\gamma)B^T[R + h^2\theta^2 RA(I + h\theta A)^{-1}(I - h\theta A)^{-1} A]B = W^T W - \widetilde{W}^T \widetilde{W}. \tag{50}$$

If $\theta \neq 1/2$, we have from (42) that $A^T RA = 0$ and then (47) and (48) are equivalent to

$$h^2(1 - 2\gamma)B^T RB = W^T W - \widetilde{W}^T \widetilde{W}, \tag{51}$$

and then we get the equivalence (43).

For the case $\theta = 1/2$, we use (50). If $\gamma \neq 1/2$ and $\theta = 1/2$, we have from (42) that $B^T RA = 0$. Then, (47) and (50) are equivalent to (51) and we get the equivalence (43). The last case $\gamma = 1/2$ is trivial. □

The necessary and sufficient conditions in (42) and (43) impose not only that the energy function is preserved, but also the dissipation function since $\widetilde{Q} = Q = 0$ (the method is NED). Note that neither observability nor controllability nor asymptotic stability conditions are required.

In the following result, we precise under the choice $L = 0$, the necessary and sufficient conditions required on the parameters θ, γ and the system (A, B, C, D) such that the energy function and the dissipation are preserved under discretization.

Corollary 1 *Suppose that both* (A, B, C, D) *and* $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ *are passive. For the following choices of* $\theta \in [0, 1]$ *and* $\gamma \in [0, 1]$*, the conditions listed are necessary and sufficient conditions for the preservation the energy storage function* (i.e. $P = hR$) *and lossless passivity upon discretization* (i.e. $W^T W = \widetilde{W}^T \widetilde{W}$ *and* $L = \widetilde{L} = 0$)*:*

(i) *For* $\theta \neq 1/2, \theta \neq 0, \theta \neq (1-\gamma), \gamma \neq 1/2:$
$A^T RA = 0 \; B^T RA = 0 \; B^T RB = 0.$

(ii) *For* $\theta \neq 1/2, \theta \neq 0, \theta = (1-\gamma)$*:* $A^T RA = 0 \; B^T RA^2 = 0 \; B^T RB = 0.$

(iii) *For* $\theta \neq 1/2, \gamma = 1/2:$ $A^T RA = 0 \; B^T RA = 0.$

(iv) *For* $\theta = 0, \gamma \neq 1/2, \gamma \neq 1:$ $A^T RA = 0 \; B^T RA = 0 \; B^T RB = 0.$

(v) *For* $\theta = 0, \gamma = 1:$ $A^T RA = 0 \; B^T RB = 0.$

(vi) *For* $\theta = 1/2, \gamma \neq 1/2:$ $B^T RA = 0 \; B^T RB = 0.$

(vii) *For* $\theta = \gamma = \frac{1}{2}$ *(midpoint method): the conditions are satisfied for any* (A, B, C, D).

Proof The proof is given by an inspection of all possible cases in the conditions of Proposition 2. □

Note that the above conditions for preserving the whole set of passivity properties are quite stringent. Indeed from (3) one has $B^T P = C \Rightarrow B^T PB = CB = (CB)^T$, so in case $B \neq 0$ the condition $B^T RB = 0$ implies that $P = hR$ is low rank and $CB = 0$. Thus the pair (A, C) cannot be observable, since observability implies that the solutions P of (2)–(4) are positive definite [14]. Moreover since R is symmetric $RB = hPB = C^T = 0$. In particular if $m = 1$ these conditions together with (41) imply $P = D = 0$ so $C = 0$ and the system cannot be passive since it has a relative degree larger than 1 [9]. We infer that (i), (ii) and (iv) apply only to a narrow class of non-observable multi-input multi-output systems. One may thus conclude from Corollary 1 that in general it is impossible to preserve the storage function together with the input dissipation function. If one relaxes $W^T W = \widetilde{W}^T \widetilde{W}$ then the condition $B^T RB = 0$ is no longer needed (see 43 in Proposition 2). The cases (i), (ii) and (iv) apply to more general classes of passive systems. In this case the obtained discrete-time system has an input dissipation matrix $\widetilde{W}$ that is obtained from (30) and the results in the proof of the proposition: $\widetilde{W}^T \widetilde{W} = (2\gamma - 1)h^2 B^T RB + W^T W = (2\gamma - 1)hB^T PB = (2\gamma - 1)h(CB)^T + W^T W = (2\gamma - 1)hCB + W^T W$. The condition $B^T RA = 0$ that comes from (29) yields similar conclusions if A has full rank. Since one deals in Corollary 1 with marginally stable systems that may possess poles on the imaginary axis, it is possible that A has low rank.

The result in Corollary 1(vii) shows that Proposition 3.3 in [19,20] can be extended in the sense that the midpoint method preserves also the dissipation function in the lossless case, with relaxed assumptions.

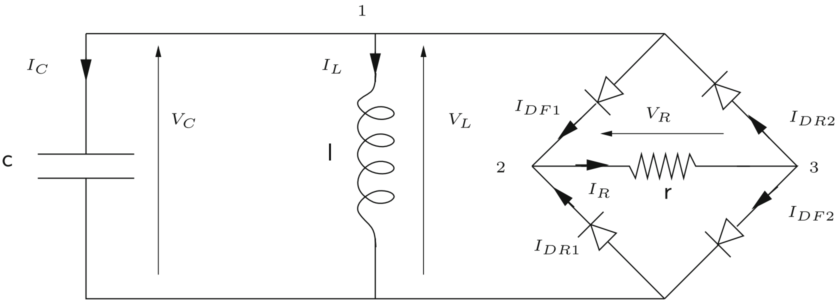

Fig. 1 LC oscillator with a load resistor

Example 2 Consider a continuous-time dissipative system with (A, B, C, D) and P defined as:

$$A = \begin{pmatrix} 0 & a_{1,2} & -a_{1,2} & 0 \\ -a_{1,2} & 0 & a_{1,2} & 0 \\ a_{1,2} & -a_{1,2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ (b_1 + b_2 + b_3)\frac{-p_2 \pm \sqrt{p_2^2 - p_3 p_1}}{p_3} \end{pmatrix}, \tag{52}$$
$$C = B^T P, \; D + D^T = 0,$$

$$P = \begin{pmatrix} p_1 & p_1 & p_1 & p_2 \\ p_1 & p_1 & p_1 & p_2 \\ p_1 & p_1 & p_1 & p_2 \\ p_2 & p_2 & p_2 & p_3 \end{pmatrix}. \tag{53}$$

Then for $R = \frac{1}{h} P$ with the added conditions $p_2^2 - p_3 p_1 \geqslant 0$ and $p_1 \neq 0$ (which are required conditions for the entries of B and C to be real), passivity is preserved under any (θ, γ)-discretization.

Example 3 Let us consider the configuration of the four-diode bridge illustrated in Fig. 1. The resistor inside the bridge is supplied by an LC oscillator. The dynamical equations are stated choosing:

$$x = \begin{pmatrix} V_L \\ I_L \end{pmatrix}, \quad w = \begin{pmatrix} I_{DR1} \\ I_{DF2} \\ V_2 - V_1 \\ V_1 - V_3 \end{pmatrix}, \quad \lambda = \begin{pmatrix} V_2 \\ -V_3 \\ I_{DF1} \\ I_{DR2} \end{pmatrix}, \tag{54}$$

and with

$$
\begin{aligned}
A &= \begin{pmatrix} 0 & -1/c \\ 1/l & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & -1/c & 1/c \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
C &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 1/r & 1/r & -1 & 0 \\ 1/r & 1/r & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.
\end{aligned}
\tag{55}
$$

The storage function matrix is:

$$
P = \begin{pmatrix} c & 0 \\ 0 & l \end{pmatrix}. \tag{56}
$$

Also:

$$
W = \begin{pmatrix} \frac{1}{\sqrt{r}} & \frac{1}{\sqrt{r}} & 0 & 0 \\ \frac{1}{\sqrt{r}} & \frac{1}{\sqrt{r}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{57}
$$

and $L = 0$. The initial conditions and parameters of the system are taken as, $x_1(0) = 10$, $x_2(0) = 0$, $l = 1 \times 10^{-2}$, $c = 1 \times 10^{-6}$, and $r = 1 \times 10^3$. Using the $(\frac{1}{2}, \frac{1}{2})$-method, one can verify that $\widetilde{L} = 0$ and $\widetilde{W} = W$. Hence we have that the system is NED for this particular discretization (since $Q - \widetilde{Q} = 0$). Figure 2 shows the NED property of the scheme. We observe the dissipative behavior as in the continuous time-case. For this configuration, the matrix D has full rank, so the solution $x(t)$ is a function of class C^1 [8,14]. One sees on Fig. 2 that the sum of the storage and the cumulative dissipation functions, is constant as expected from (5).

3.2 General preservation conditions ($L \neq 0$, $W \neq 0$)

As in the previous section, let us start with an equivalence result.

Proposition 3 *Let $hR = P$. Suppose that both (A, B, C, D) and $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ are passive. Then we have*

$$
\begin{cases} \widetilde{L}\widetilde{L}^T = LL^T \\ \widetilde{W}^T\widetilde{L}^T = W^T L^T \end{cases} \iff \begin{cases} \theta A^T LL^T = 0; \quad \theta B^T LL^T = 0 \\ (2\theta - 1)A^T RA = 0 \\ (1 - \theta - \gamma)B^T RA = 0 \\ \theta(\gamma - \theta)B^T RA^2 = 0 \\ B^T LL^T A = 0 \\ \gamma W^T L^T A = 0. \end{cases} \tag{58}
$$

Let us further suppose that $LL^T = \widetilde{L}\widetilde{L}^T$ and $\widetilde{W}^T\widetilde{L}^T = W^T L^T$. Then, we have

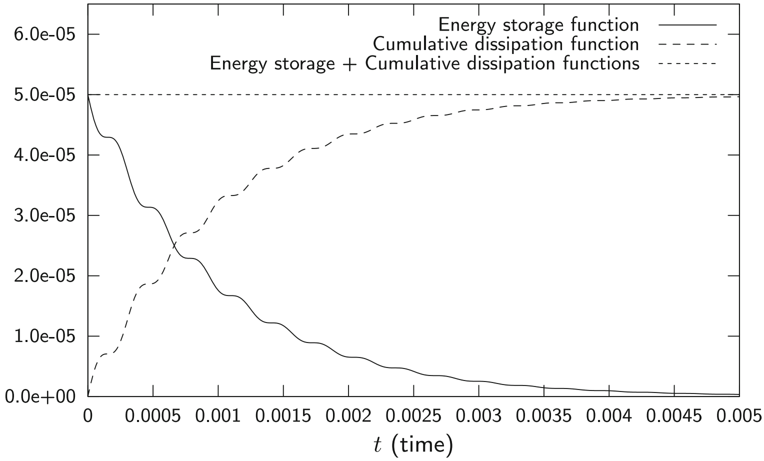

Fig. 2 Example 2. LC oscillator: $(\frac{1}{2}, \frac{1}{2})$-method approximation of storage function (*solid line*), $(\frac{1}{2}, \frac{1}{2})$-method approximation of cumulative dissipation function (*dashed line*), $(\frac{1}{2}, \frac{1}{2})$-method approximation of storage function + cumulative dissipation function (*dotted line*). Time step $h = 1 \times 10^{-6}$

$$W^T W = \widetilde{W}^T \widetilde{W} \iff \begin{cases} (1-2\gamma)B^T R B = 0 \\ \gamma W^T L^T B = \gamma B^T L W. \end{cases} \tag{59}$$

Proof From Proposition 1 and Lemma 1(ii), we have

$$\begin{cases} \widetilde{L}\widetilde{L}^T = LL^T \\ \widetilde{W}^T \widetilde{L}^T = W^T L^T \end{cases} \implies \begin{cases} \theta A^T L L^T = 0; \quad (2\theta - 1)A^T R A = 0 \\ hB^T R\big(h(1-\theta-\gamma)A - h^2\theta(\gamma-\theta)A^2\big) \\ \quad -\theta B^T (I - h\theta A)^{-T} L L^T (I + h(1-\theta)A) \\ \qquad = h\gamma\theta W^T L^T A(I + h\theta A) \end{cases} \tag{60}$$

Note that $\theta A^T L L^T = 0$ implies $(I - h\theta A)^T L L^T (I - h\theta A) = L L^T$, and therefore

$$(I - h\theta A)^{-T} L L^T = L L^T (I - h\theta A). \tag{61}$$

The last equation in the right-hand side of (60) can be rewritten as:

$$\begin{aligned} &hB^T R[h(1-\theta-\gamma)A - h^2\theta(\gamma-\theta)A^2] - \theta B^T L L^T (I - h\theta A)(I + h(1-\theta)A) \\ &\quad = h\gamma\theta W^T L^T A(I + h\theta A). \end{aligned} \tag{62}$$

Expanding the terms and grouping with the orders of h, we get

$$\begin{aligned}&-\theta B^T LL^T + h[(1-2\theta)B^T LL^T A - \gamma W^T L^T A]\\ &\quad +h^2[(1-\theta-\gamma)B^T RA - \theta(1-\theta)B^T LL^T A^2 - \gamma\theta W^T L^T A^2]\\ &\quad -h^3\theta(\gamma-\theta)B^T RA^2 = 0.\end{aligned} \tag{63}$$

The implication (60) can be then simplified to:

$$\begin{cases}\widetilde{L}\widetilde{L}^T = LL^T\\ \widetilde{W}^T\widetilde{L}^T = W^T L^T\end{cases} \implies \begin{cases}\theta A^T LL^T = 0; & \theta B^T LL^T = 0\\ (2\theta-1)A^T RA = 0; & (1-\theta-\gamma)B^T RA = 0\\ \theta(\gamma-\theta)B^T RA^2 = 0; & B^T LL^T A = 0\\ \gamma W^T L^T A = 0. & \end{cases} \tag{64}$$

Conversely, let us assume that the right-hand side of (58) holds. From Proposition 1, it follows that $\widetilde{L}\widetilde{L}^T = LL^T$. From Lemma 1(ii) and (61), we have

$$\begin{aligned}&[W^T L^T(I+h(\gamma-\theta)A) - \widetilde{W}^T\widetilde{L}^T(I-h\theta A)](I+h\theta A)\\ &\quad = -\theta B^T LL^T(I-h\theta A)(I+h(1-\theta)A) = 0.\end{aligned} \tag{65}$$

Simplifying and using $\theta B^T LL^T = 0$, it follows that

$$(W^T L^T - \widetilde{W}^T\widetilde{L}^T)(I-h\theta A) + h\gamma W^T L^T A = 0. \tag{66}$$

Since $\gamma W^T L^T A = 0$, we get $\widetilde{W}^T\widetilde{L}^T = W^T L^T$. The proof of the equivalence (58) is then completed.

Let us switch to the proof of (59) under the assumption that $LL^T = \widetilde{L}\widetilde{L}^T$ and $\widetilde{W}^T\widetilde{L}^T = W^T L^T$. In particular, we have (58) at hand. From Proposition 1 and Lemma 1(iii), we have

$$\widetilde{W}^T\widetilde{W}^T = W^T W^T \iff \begin{cases} h^2 B^T(I-h\theta A)^{-T}((1-2\gamma)R - \theta\gamma LL^T)(I-h\theta A)^{-1}B\\ \quad = h\gamma W^T L^T(I-h\theta A)^{-1}B + h\gamma B^T(I-h\theta A)^{-T}LW.\end{cases} \tag{67}$$

Under the assumption that $\theta A^T LL^T = 0$, we recall that we have $(I-h\theta A)^{-T}LL^T(I-h\theta A)^{-1} = LL^T$. Since $\theta B^T LL^T = 0$, the right-hand side in (67) can be written as

$$\begin{aligned}&(1-2\gamma)h^2 B^T(I-h\theta A)^{-T}R(I-h\theta A)^{-1}B\\ &\quad = h\gamma W^T L^T(I-h\theta A)^{-1}B + h\gamma B^T(I-h\theta A)^{-T}LW.\end{aligned} \tag{68}$$

Let us focus for a while on the left-hand side of (68). Let us perform the same algebraic manipulation as in (50) in the proof of Proposition 2. Firstly, using $A^T R + RA = -LL^T$, we have $(I-h\theta A)^T R = R(I+h\theta A) + h\theta LL^T$, and then using (61) ($\theta A^T LL^T = 0$ holds), we have

$$R(I+h\theta A)^{-1} = (I-h\theta A)^{-T}R + h\theta LL^T(I-h\theta A)(I+h\theta A)^{-1}. \tag{69}$$

Since $\theta B^T LL^T = 0$, it follows that

$$B^T R(I + h\theta A)^{-1} = B^T (I - h\theta A)^{-T} R. \tag{70}$$

Let us remark that $B^T R(I - h\theta A)(I + h\theta A) = B^T R - h^2\theta^2 B^T RA^2$. We obtain

$$B^T (I - h\theta A)^{-T} R(I - h\theta A)^{-1} = B^T R + h^2\theta^2 B^T RA^2(I + h\theta A)^{-1}(I - h\theta A)^{-1}. \tag{71}$$

Therefore the right-hand side of (68), is equivalently rewritten as

$$(1 - 2\gamma)h^2 B^T [R + h^2\theta^2 B^T RA^2(I + h\theta A)^{-1}(I - h\theta A)^{-1}]B. \tag{72}$$

For the left-hand side of (68), the equality $W^T L^T (I + h\theta A)(I - h\theta A) = W^T L^T - h^2\theta^2 W^T L^T A^2$ implies that

$$\gamma W^T L^T (I - h\theta A)^{-1} = \gamma W^T L^T (I + h\theta A) + h^2\gamma\theta^2 W^T L^T A^2(I - h\theta A)^{-1}. \tag{73}$$

Since $\gamma W^T L^T A = 0$, Eq. (73) implies

$$\gamma W^T L^T (I - h\theta A)^{-1} B = \gamma W^T L^T B, \quad \text{and } \gamma B^T (I - h\theta A)^{-T} LW = \gamma B^T LW. \tag{74}$$

Finally, the equivalence in (67) is equivalent to

$$\widetilde{W}^T \widetilde{W}^T = W^T W^T \iff \begin{cases} (1 - 2\gamma)B^T RB - h^2(1 - 2\gamma)\theta^2 B^T RA^2(I + h\theta A)^{-1}(I - h\theta A)^{-1}B \\ \quad = h\gamma(W^T L^T B + B^T LW). \end{cases} \tag{75}$$

We know from (58) that $\theta(\theta - \gamma)B^T RA^2 = 0$. Two cases can be discussed. If $\theta \neq \gamma$, then $B^T RA^2 = 0$ and then we get (59). If $\theta = \gamma$m, we get from (58) that $B^T RA = 0$, and (59) holds. □

If we let $L = 0$ we recover from Proposition 3 the results of Proposition 2. Let us now consider the following corollary when $W = 0$.

Corollary 2 *Suppose that both* (A, B, C, D) *and* $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ *with* $W = 0$ *are passive. For the following choices of* θ *and* γ *the conditions listed are sufficient and necessary for the preservation of the energy storage function* (i.e. $R = \frac{1}{h}P$) *and the state dissipation upon discretization* (i.e. $W^T L^T = \widetilde{W}^T \widetilde{L}^T = 0$, $W^T W = \widetilde{W}^T \widetilde{W} = 0$ *and* $LL^T = \widetilde{L}\widetilde{L}^T$)*:*

(i) *For* $\theta \neq 1/2, \theta \neq 0, \theta \neq 1 - \gamma, \gamma \neq 1/2$*:*
$A^T LL^T = 0, B^T LL^T = 0, A^T RA = 0, B^T RA = 0, B^T RB = 0.$
(ii) *For* $\theta \neq 1/2, \theta \neq 0, \gamma = 1/2$*:* $A^T LL^T = 0, B^T LL^T = 0, A^T RA = 0, B^T RA = 0.$

(iii) *For* $\theta \neq 1/2, \theta \neq 0, \theta = 1 - \gamma$*:* $A^T LL^T = 0, B^T LL^T = 0, A^T RA = 0, B^T RA^2 = 0, B^T RB = 0$.
(iv) *For* $\theta = 0, \gamma \neq 1, \gamma \neq 1/2$*:* $A^T RA = 0, B^T RA = 0, B^T LL^T A = 0, B^T RB = 0$.
(v) *For* $\theta = 0, \gamma = 1$*:* $A^T RA = 0, B^T LL^T A = 0, B^T RB = 0$,.
(vi) *For* $\theta = 0, \gamma = 1/2$*:* $A^T RA = 0, B^T RA = 0, B^T LL^T A = 0$,.
(vii) *For* $\theta = 1/2 \neq \gamma$*:* $A^T LL^T = 0, B^T LL^T = 0, B^T RA = 0, B^T RB = 0$.
(viii) *For* $\theta = 1/2 = \gamma$*:* $A^T LL^T = 0, B^T LL^T = 0$.

Proof The proof is given by an inspection of all possible cases in the conditions of Proposition 3 with $W = 0$. □

The conditions in Proposition 3 and Corollary 2 are quite stringent conditions on the structure of the system (A, B, C, D). The following corollary give some insights on the consequences of some of conditions of Proposition 3.

Corollary 3 *Suppose that both* (A, B, C, D) *and* $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ *are passive. Let us suppose further that the conditions of Proposition 3 are satisfied. Then, we have*

(a) $\gamma \neq 0 \implies CB = (CB)^T$
(b) $\gamma \neq 1/2 \implies CB = W^T L^T B$
(c) $\theta \neq 1/2$ *and* $\theta \neq 0 \implies PA^2 = 0$. *If* (A, B, C) *is observable, then* $A^2 = 0$.
(d) $\theta + \gamma \neq 1$ *and* $\gamma \neq 0 \implies CA = 0$.

Proof (a) If $\gamma \neq 0$, the conditions in (59) imply that $W^T L^T B$ is symmetric. From (3), one has $B^T PB - CB = -W^T L^T B = 0$ so that $CB \in \mathbb{R}^{m \times m}$ is symmetric. (b) If $\gamma \neq 1/2$, then $B^T PB = B^T hRB = 0$ so $CB = W^T L^T B$. (c) If $\theta \neq 1/2$ and $\theta \neq 0$, then $A^T RA = A^T LL^T = B^T LL^T = 0$. Using (2), we have

$$A^T PA + PA^2 = hA^T RA + PA^2 = PA^2 = -LL^T A = 0, \tag{76}$$

so $PA^2 = 0$. If (A, B, C) is observable, $P > 0$ and then $A^2 = 0$. (d) If $\theta + \gamma \neq 1$, we get $B^T RA = 0$. Then we have

$$B^T PA - CA = B^T hRA - CA = -W^T L^T A, \tag{77}$$

so $CA = -W^T L^T A$. If in addition $\gamma \neq 0$, $W^T L^T A =$ and then $CA = 0$ □

4 Systems with state jumps

Complementarity systems as in (1) may undergo state jumps (for instance initially if $D = 0$ and $Cx(0^-) < 0$). They may be seen as a switching system that switches between DAEs, where the number of constraints of the DAEs may vary: complementarity systems may live on lower-dimensional subspaces. The switches are ruled by complementarity conditions. The state jumps are necessary to re-initialize the system so that the right limit of the state is an admissible initial data for the new mode (the new DAE). The first point to fix here is a modelling issue. Depending on the application the state re-initialization may take different forms. In [2, section 1.1.5] it is shown on a circuit example that the θ-method with $\theta = \gamma = 1$ is able to approximate state jumps for inconsistent states. There are mainly two sources of state discontinuities: the first one is associated with inconsistent initial values and the second one is due to the external excitation term $Fv(t)$ in (1) which may move the state outside the feasible region defined by the complementarity condition, see [10,12].

4.1 The state jump law

If some state jumps are expected, the state $x(t)$ is usually assumed to be a right continuous function of local bounded variations (RCLBV) [31,32,34], or of special bounded variations (RCLSBV) [4]. The variable λ has to be replaced by a measure that contains Dirac distributions. In the same vein, the time–derivative of the state $x(t)$ cannot be considered in the usual sense but as a differential measure dx associated with a RCLBV function $x(t)$ [31]. In the following we shall assume that the solution of (1) is RCLSBV. Consequently the dynamics in (1) is written in terms of a measure differential equation as:

$$dx = Ax(t)dt + Eu(t)dt + Bd\Lambda, \tag{78}$$

where $d\Lambda$ is a measure associated in the following way with $\lambda(t)$. The absolutely continuous function $\lambda(t)$ is the density of $d\Lambda$ with respect to the Lebesgue measure dt, *i.e.*:

$$\frac{d\Lambda}{dt}(t) = \lambda(t). \tag{79}$$

Since we assume that solutions are RCLSBV, a decomposition of the measure can be written as [3,31,34]:

$$d\Lambda = \lambda(t)dt + \sum_i \sigma_i \delta_{t_i}, \tag{80}$$

where δ_{t_i} is the Dirac measure at time of discontinuities t_i and σ_i the amplitude. Using this decomposition, the differential measure in (78) can be written as a smooth dynamics:

$$\dot{x}(t) = Ax(t) + Eu(t) + B\lambda(t), \quad dt - \text{almost everywhere}, \tag{81}$$

and a jump dynamics at t_i:

$$x(t_i^+) - x(t_i^-) = B\sigma_i. \tag{82}$$

The jump dynamics (82) is not sufficient to determine uniquely the state $x(t_i^+)$ after a discontinuity. A jump rule needs to be stated which has to be consistent with the complementarity conditions. In the sequel, the following energy-based jump rule in Definition 7 will be used. This jump rule follows from [10,14,21], and is inspired by Moreau's generalized impact law for Lagrangian systems (see [33] and [3, §2.7]). In order to design a coherent state jump law, we need to ensure :

1. that the post-jump state $x(t^+)$ will be coherent with the constraints in the system (1), that is

$$\begin{cases} w(t^+) = Cx(t^+) + D\lambda(t^+) + Fv(t^+) \\ 0 \leqslant \lambda(t^+) \perp w(t^+) \geqslant 0 \end{cases} \tag{83}$$

   has a solution,
2. and that the energy storage function decreases at the jump time.

The second point will be checked below in Lemma 3. For the first point, we need to define the set of admissible post-jump states $x(t^+)$ in the following way.

Definition 6 (*Admissible post-jump states*) Let us define the set of admissible post-jump state as

$$\mathbf{K} = \{z \in \mathbb{R}^n \mid Cz + Fv(t^+) \in \mathbf{Q}^*\}, \tag{84}$$

with $\mathbf{Q}^*$ the dual cone of

$$\mathbf{Q} = \{z \in \mathbb{R}^m \mid z \geqslant 0,\, Dz \geqslant 0,\, z^T Dz = 0\}. \tag{85}$$

Note that $\mathbf{Q}$ is a closed convex cone, which is sometimes called the kernel of $LCP(0, D)$.

Why this definition ? If the qualification constraint

$$Fv(t) \in \mathbf{Q}^* + \mathrm{Im}(C) \tag{86}$$

is satisfied, then the set $\mathbf{K} \neq \emptyset$. More precisely, these conditions are equivalent. If $\mathbf{K} \neq \emptyset$, the LCP (83) has a non-empty solution set provided that $D \geqslant 0$.

Definition 7 (*State Jump Law*) Let us consider the dynamics in (1), and suppose that (A, B, C, D) is passive with storage function $V(x) = \frac{1}{2}x^T Px$, $P = P^T > 0$. For any $x(t^-)$, the state after the discontinuities, *i.e.* $x(t^+)$, is given by the solution of the generalized equation :

$$P(x(t^+) - x(t^-)) \in -N_{\mathbf{K}}(x(t^+)). \tag{87}$$

The following lemma gives equivalent formulations of the state jump law (87).

Lemma 2 *Under the conditions of Definition 7, the following holds:*

$$P(x(t^+) - x(t^-)) \in -N_{\mathbf{K}}(x(t^+)). \tag{88}$$

$$\Updownarrow$$

$$P(x(t^+) - x(t^-))(x(t^+) - y) \geqslant 0, \quad \textit{for all } y \in \mathbf{K}. \tag{89}$$

$$\Updownarrow$$

$$x(t^+) = \mathrm{argmin}_{x\in\mathbf{K}} \tfrac{1}{2}(x - x(t^-))^T P(x - x(t^-)), \tag{90}$$

$$\Updownarrow$$

$$\mathbf{K} \ni x(t^+) \perp P(x(t^+) - x(t^-)) \in \mathbf{K}^* \tag{91}$$

$$\Updownarrow$$

$$x(t^+) = \mathrm{proj}_P[\mathbf{K}; x(t^-)] \tag{92}$$

$$\Updownarrow$$

$$\begin{cases} P(x(t^+) - x(t^-)) = C^T\sigma \\ w = Cx(t^+) + Fv(t^+) \\ \mathbf{Q}^* \ni w \perp \sigma \in \mathbf{Q}. \end{cases} \tag{93}$$

$$\Updownarrow$$

$$\begin{cases} (x(t^+) - x(t^-)) = B\sigma \\ w = Cx(t^+) + Fv(t^+) \\ \mathbf{Q}^* \ni w \perp \sigma \in \mathbf{Q}. \end{cases} \tag{94}$$

$$\Updownarrow$$

$$\begin{cases} w = Cx(t^-) + Fv(t^+) + CB\sigma \\ \mathbf{Q}^* \ni w \perp \sigma \in \mathbf{Q}. \end{cases} \tag{95}$$

Proof The equivalence between (88) and (89) follows from the definition of a normal cone to a convex set [25, Definition 5.2.3]. The equivalences between (89), (90) and (91) can be shown using the material in [18, Chapter 1]. The equivalence between (91) and (88) is direct from convex analysis: for any convex non empty closed cone $\mathbf{K} \subset \mathbb{R}^n$ and any two vectors x and y in $\mathbb{R}^n$, $\mathbf{K} \ni x \perp y \in \mathbf{K}^* \quad\Leftrightarrow\quad y \in -N_{\mathbf{K}}(x)$. Notice that (92) is just a rewriting of (90) with the projection in the metric defined by P. The equivalence between (93) and (88) can be shown as follows: the complementarity conditions in (93) are equivalent to $\sigma \in -N_{\mathbf{Q}^*}(w)$ that is equivalent (since $P > 0$) to $P(x(t^+) - x(t^-)) \in -C^T N_{\mathbf{Q}^*}(Cx(t^+) + Fv(t^+)) = -N_{\mathbf{K}}(x(t^+))$, where the last equality follows from the chain rule of convex analysis [25, Theorem 4.2.1] and the definitions of $\mathbf{K}$ and $\mathbf{Q}$. The equivalence between (93) and (94) is true since $\sigma \in \mathbf{Q}$ implies that $\sigma^T(D + D^T)\sigma = 0$ and then $C^T\sigma = PB\sigma$ (see *e.g.* [24, Lemma 2.b]). Finally (95) is just a rewriting of (94). □

One of the equivalent expression (92) is given as a projection. If the qualification constraint (86) holds, or equivalently $\mathbf{K}(t) \neq \emptyset$, the projection is unique. Furthermore, the post-jump state $x(t^+)$ is consistent with the complementarity system's dynamics on the right of t.

Lemma 3 *The state jump law in (87) guarantees that* $V(x(t^+)) - V(x(t^-)) \leqslant 0$ *provided that* $0 \in \mathbf{K}$.

Proof Direct from (92) because $0 \in \mathbf{K}$ assures that the projection makes the norm of $x(t^+)$, in the metric defined by P, smaller than that of $x(t^-)$. □

Definition 7 is also related to the principle of maximum dissipation in the following sense. Another equivalent expression of the state jump law is given in (90) by a quadratic problem. The post-jump state $x(t^+)$ that satisfies (90) minimizes the energy storage function over the admissible jumps, *i.e.* $\mathbf{K}-x(t^-)$. In other terms, it maximizes the dissipation over the admissible jumps $x(t^+) - x(t^-)$.

4.2 Jump rule consistency

If $v(\cdot) = 0$ in (1), then the jump occurs initially and dissipates energy since the condition $0 \in \mathbf{K}$ is always satisfied [12]. The condition $0 \in \mathbf{K}$ may also be satisfied for $v(t) \neq 0$. The time-discretization of (79) has to take into account the nature of the solution to avoid point-wise evaluations of measures at atoms. A direct application of the scheme (40) is not consistent with possible jumps in the state. Let us consider that the scheme (40) is used with $x_k = x(t^-)$ and we expect to have a jump at time t such that $x(t^+) - x(t^-) = \sigma \neq 0$. If the scheme is consistent, we expect to have $\lim_{h\to 0} x_{k+1} = x(t^+)$. For $B \neq 0$, the scheme implies that $\lim_{h\to 0} \lambda_{k+\gamma} = \infty$. This reveals a point–wise evaluation of a measure. Only the measures of the time–intervals $(t_k, t_{k+1}]$ are considered such that:

$$dx((t_k, t_{k+1}]) = \int_{t_k}^{t_{k+1}} Ax(t) + Eu(t)\, dt + Bd\Lambda((t_k, t_{k+1}]). \tag{96}$$

By definition of a differential measure, we have:

$$dx((t_k, t_{k+1}]) = x(t_{k+1}^+) - x(t_k^+). \tag{97}$$

The measure of the time–interval by $d\Lambda$ is kept as an unknown variable denoted by:

$$\sigma_{k+1} \approx d\Lambda((t_k, t_{k+1}]). \tag{98}$$

Finally, the remaining Lebesgue integral in (96) is approximated by the θ-method:

$$\int_{t_k}^{t_{k+1}} Ax(t) + Eu(t)\, dt \approx h(Ax_{k+\theta} + Eu_{k+\theta}), \tag{99}$$

yielding the following integration formula for (78):

$$x_{k+1} - x_k = h(Ax_{k+\theta} + Eu_{k+\theta}) + B\sigma_{k+1}. \tag{100}$$

In the following sections, we try to answer the question: Is the scheme based on the integration rule (100) able to consistently approximate the jump rule of Definition 7? To be more precise, let us give a definition of a consistent approximation of the jump rule by a numerical scheme.

Definition 8 (*Consistent approximation of the jump law*) Let us assume that $\mathbf{K} \neq \emptyset$. A numerical scheme with a time–step h which generates the sequence $\{x_k\}$ is said to consistently approximate the jump rule of Definition 7 if for $x_k = x(t^-)$ and $v_{k+1} = v(t^+)$ we have

$$\lim_{h \to 0} \|x_{k+1} - x(t^+)\| = 0, \tag{101}$$

where $x(t^+)$ is the unique solution of (87) for given $x(t^-)$ and $v(t^+)$.

We will also assume that the RCLSBV solution exists and that the following schemes based on (100) generate a bounded sequence $\{x_k\}$ for a sufficiently small h. Especially, for one time–step, given the values of x_k and σ_k, we assume that

$$\lim_{h \to 0} x_{k+1} < +\infty \text{ and } \lim_{h \to 0} \sigma_{k+1} < +\infty. \tag{102}$$

In the following, only the most general and interesting case $D \geqslant 0$ is investigated. The simplest cases $D = 0$ and $D \geqslant 0$ with a special block structure are treated in details in [23]. The case $D > 0$ implies that the solutions are continuous of class C^1, and is of no interest in this section.

Considering the most general case $D \geqslant 0$, we propose to use the following $(\theta, 1)$-scheme:

$$\begin{cases} x_{k+1} - x_k = h\,(Ax_{k+\theta} + Eu_{k+\theta}) + B\sigma_{k+1} \\ w_{k+1} = Cx_{k+1} + Fv_{k+1} + \dfrac{D}{h}\sigma_{k+1} \\ 0 \leqslant w_{k+1} \perp \sigma_{k+1} \geqslant 0. \end{cases} \tag{103}$$

Let us denote by $\mathbf{R}$ is the set of vectors q such that LCP(q, D) has a solution, that is:

$$\mathbf{R} = \{q \in \mathbb{R}^m \mid \mathrm{SOL}(q, D) \neq \emptyset\}. \tag{104}$$

The set R is often called the range of the LCP(q, D). From Lemma 4 in Appendix A one has:

$$\mathbf{Q} = \{z \in \mathbb{R}^m \mid z \geqslant 0,\, D^T z \leqslant 0\} = \mathbf{R}^*, \tag{105}$$

and

$$\mathbf{R} = \mathbf{Q}^* = \mathbb{R}^m_+ - D\mathbb{R}^m_+. \tag{106}$$

Proposition 4 *Let us denote the limit* $\lim_{h \to 0} \sigma_{k+1} = \sigma_\infty$ *with* $\sigma_\infty < +\infty$. *Then* σ_∞ *solves the following LCCP:*

$$\mathbf{Q} \ni \sigma_\infty \perp Fv_{k+1} + Cx_k + CB\sigma_\infty \in \mathbf{Q}^*, \tag{107}$$

for $x_k = x(t^-)$ *and* $v_{k+1} = v(t^+)$, *which is equivalent to the jump law in Definition 7. Furthermore, the scheme* (103) *consistently approximates the jump rule in Definition 8.*

Proof From (103) let us consider the associated LCP one–step problem:

$$\begin{cases} hw_{k+1} = h(Cx_{k+1} + Fv_{k+1}) + D\sigma_{k+1} \\ 0 \leqslant hw_{k+1} \perp \sigma_{k+1} \geqslant 0. \end{cases} \tag{108}$$

If we assume that $\lim_{h\to 0} \sigma_{k+1} = \sigma_\infty < +\infty$ and $\lim_{h\to 0} x_{k+1} < +\infty$, we have that:

$$\lim_{h\to 0} hw_{k+1} = D\sigma_\infty, \tag{109}$$

and σ_∞ satisfies:

$$0 \leqslant D\sigma_\infty \perp \sigma_\infty \geqslant 0. \tag{110}$$

This implies that $\sigma_\infty \in \mathbf{Q}$. From (106), we note that

$$w_{k+1} - \frac{D}{h}\sigma_{k+1} \in \mathbf{R} = \mathbf{Q}^*. \tag{111}$$

From (103) one has:

$$w_{k+1} - \frac{D}{h}\sigma_{k+1} = Fv_{k+1} + C(I - h\theta A)^{-1}\left[(I + h(1-\theta)A)x_k + hEu_{k+\theta}\right] \\ +C(I - h\theta A)^{-1}B\sigma_{k+1}. \tag{112}$$

Let us denote the limit of w_{k+1} as:

$$w_\infty \triangleq \lim_{h\to 0} w_{k+1} - \frac{D}{h}\sigma_{k+1}. \tag{113}$$

It follows that:

$$w_\infty = Fv_{k+1} + Cx_k + CB\sigma_\infty \in \mathbf{Q}^*, \tag{114}$$

since $\mathbf{Q}^*$ is a closed set and we assume $\lim_{h\to 0} \sigma_{k+1} = \sigma_\infty$. It remains to prove that $w_\infty \perp \sigma_\infty$. Since $\lim_{h\to 0} \sigma_{k+1} = \sigma_\infty < +\infty$, hence $w_\infty < +\infty$, we can write:

$$\begin{aligned} \sigma_\infty^T w_\infty &= \lim_{h\to 0} \sigma_{k+1}^T\left(w_{k+1} - \frac{D}{h}\sigma_{k+1}\right) \\ &= \lim_{h\to 0} -\frac{1}{h}\sigma_{k+1}^T D\sigma_{k+1} \qquad \text{due to (103)} \\ &\leqslant 0, \end{aligned} \tag{115}$$

due to the positive semi–definiteness of D. Since $\sigma_\infty^T \in \mathbf{Q}$ and $w_\infty \in \mathbf{Q}^*$, we also have $\sigma_\infty^T w_\infty \geqslant 0$ and therefore we conclude that $\sigma_\infty^T w_\infty = 0$. To summarize, w_∞ and

σ_∞ solve the following LCCP:

$$\begin{cases} w_\infty = Fv_{k+1} + Cx_k + CB\sigma_\infty \\ \mathbf{Q}^* \ni w_\infty \perp \sigma_\infty \in \mathbf{Q}, \end{cases} \tag{116}$$

or equivalently the jump law (93). From (94) and (103), we get:

$$\|x_{k+1} - x(t^+)\| = \|h\,(Ax_{k+\theta} + Eu_{k+\theta}) + B(\sigma_{k+1} - \sigma)\| \tag{117}$$

since $x_k = x(t^-)$ and therefore, we conclude that

$$\lim_{h\to 0} \|x_{k+1} - x(t^+)\| = \lim_{h\to 0} \|B(\sigma_{k+1} - \sigma)\| \tag{118}$$

for a bounded sequence of $x_{k+\theta}$ and $u_{k+\theta}$. If $\mathbf{K} \neq \emptyset x(t^+)$ is uniquely defined by the jump law (93) and therefore $B\sigma$ is also uniquely defined. We can therefore conclude that $\lim_{h\to 0} \|B(\sigma_{k+1} - \sigma)\| = 0$ which ends the proof. □

Remark 3 In the introduction, one recalls that we are not interested in convergence results as the time–step vanishes, but on the algorithm properties when $h > 0$ is finite. Albeit the result of Proposition 4 is original as far as we know, it is not sufficient since it deals with the limit case $h \to 0$. The respect of the state jump law for a finite step size h holds if and only if

$$\mathbf{Q} \ni \sigma_{k+1} \perp Fv_{k+1} + Cx_k + CB\sigma_{k+1} \in \mathbf{Q}^* \tag{119}$$

and we have in practice

$$\begin{aligned} 0 \leqslant\ & Fv_{k+1} + C(I - h\theta A)^{-1}[(I + h(1-\theta)A)x_k + hEu_{k+\theta}] \\ & + [D/h + C(I - h\theta A)^{-1}B]\sigma_{k+1} \perp \sigma_{k+1} \geqslant 0. \end{aligned} \tag{120}$$

Sufficient conditions for (120) to imply (119) are that for any $h > 0$, the following equalities hold

$$\begin{cases} [D/h + C(I - h\theta A)^{-1}B] = CB \\ C(I - h\theta A)^{-1}[(I + h(1-\theta)A)x_k + hEu_{k+\theta}] = Cx_k \end{cases} \tag{121}$$

which implies

$$D = 0, \quad \mathbf{Q} = \mathbf{Q}^* = \mathbb{R}^m_+, \quad A = 0, \quad E = 0 \tag{122}$$

Such conditions are stringent.

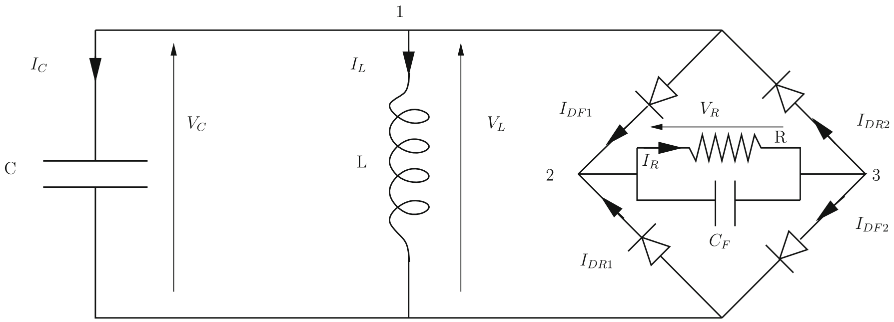

Fig. 3 LC oscillator with a load resistor filtered by a capacitor

Example 4 (Diode bridge cap filter) Let us consider the circuit in Fig. 3. Its dynamics is given by the following data:

$$A = \begin{pmatrix} 0 & -\frac{1}{c} & 0 \\ \frac{1}{l} & 0 & 0 \\ 0 & 0 & -\frac{1}{rc_f} \end{pmatrix}, \; B = \begin{pmatrix} 0 & 0 & -\frac{1}{c} & \frac{1}{c} \\ 0 & 0 & 0 & 0 \\ \frac{1}{c_f} & 0 & \frac{1}{c_f} & 0 \end{pmatrix}, \; C = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \; D = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \quad (123)$$

with

$$L = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\frac{2}{r}} \end{pmatrix} \text{ and } W = 0. \quad (124)$$

The parameter values and initial conditions of the system are taken as, $x_1(0) = 10.0, x_2(0) = 0, r = 10^3, c = 10^{-6}, c_f = 300.10^{-9}$. The discrete system (for $\theta = \gamma = \frac{1}{2}$) has:

$$\widetilde{L} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2rc_f}{2rc_f+h}\sqrt{\frac{2}{r}} \end{pmatrix} \text{ and } \widetilde{W} = \frac{\sqrt{2r}h}{(2rc_f+h)} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (125)$$

The non-zero eigenvalues of $Q - \widetilde{Q}$ are:

$$\frac{-2hr^2 + 4rc_f + h \pm \sqrt{(1+2r^2)(2h^2r^2 + 16r^2c_f^2 + 8rc_fh + h^2)h}}{r(4r^2c_f^2 + 4rc_fh + h^2)}. \quad (126)$$

Let us start with some computation for this specific example:

$$\mathbf{Q} = \{z \in \mathbb{R}^m \mid z \geqslant 0, Dz \geqslant 0, z^T Dz = 0\}. \quad (127)$$

Since $D + D^T = 0$ and $z^T D z = \frac{1}{2} z^T (D + D^T) z$, the condition $z^T D z = 0$ holds for any $z \in \mathbb{R}^m$. The computation of $\mathbf{Q}$ yields

$$\mathbf{Q} = \{z \in \mathbb{R}^4 \mid z \geqslant 0, z_2 = 0, z_1 + z_3 - z_4 \geqslant 0\}, \tag{128}$$

and the cone $\mathbf{K}$ given by

$$\mathbf{K} = \{x \in \mathbb{R}^3 \mid Cx \in \mathbf{Q}\}, \tag{129}$$

is given in our example by

$$\mathbf{K} = \{x \in \mathbb{R}^3 \mid x_1 \geqslant 0, x_3 \geqslant 0, 2(x_3 - x_1) \geqslant 0\}, \tag{130}$$

that is

$$\mathbf{K} = \{x \in \mathbb{R}^3 \mid Cx \geqslant 0\}. \tag{131}$$

We can check that

$$P = \begin{pmatrix} c & 0 & 0 \\ 0 & L & 0 \\ 0 & 0 & c_F \end{pmatrix} \tag{132}$$

solves the LMI in (2). The jump law is given by:

$$-P(x(t^+) - x(t^-)) \in N_{\mathbf{K}}(x(t^+)). \tag{133}$$

Since $\mathbf{K} = \{x \in \mathbb{R}^3 \mid Cx \geqslant 0\}$, we get:

$$\begin{cases} P(x(t^+) - x(t^-)) = C^T \sigma \\ w = Cx(t^+) \\ 0 \leqslant w \perp \sigma \geqslant 0. \end{cases} \tag{134}$$

Since $PB = C^T$ and $P > 0$, we get:

$$\begin{cases} (x(t^+) - x(t^-)) = B\sigma \\ w = Cx(t^+) \\ 0 \leqslant w \perp \sigma \geqslant 0, \end{cases} \tag{135}$$

and we conclude that the jump law amounts to solving

$$\begin{cases} w = Cx(t^-) + CB\sigma \\ 0 \leqslant w \perp \sigma \geqslant 0, \end{cases} \tag{136}$$

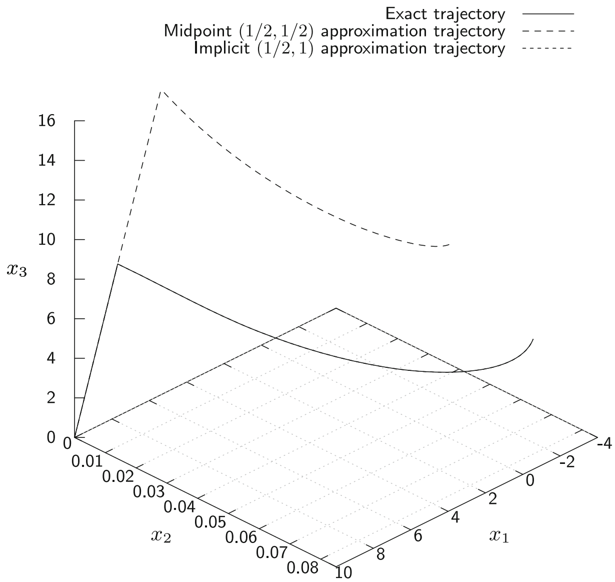

Fig. 4 Example 4: Capfilter phase portrait: exact solution (*solid line*), $(\frac{1}{2}, \frac{1}{2})$-method approximation (*dashed line*), $(\frac{1}{2}, 1)$-method approximation (*dotted line*). Initial state (10,0,0), time step $h = 1.0 \times 10^{-6}$

which is exactly what is solved by the time-stepping scheme at the first step for $h \to 0$. If the (θ, γ)-scheme is used, we solve:

$$\begin{cases} w_{k+\gamma} = M\lambda_{k+\gamma} + q \\ \mathbf{K}^* \in w_{k+\gamma} \perp \lambda_{k+\gamma} \in \mathbf{K}, \end{cases} \tag{137}$$

with

$$\begin{aligned} M = D + h\gamma C(I_n - h\theta A)^{-1}B, q = a_{k+\gamma} \\ + \gamma C(I_n - h\theta A)^{-1}\left[(I_n + h(1-\theta)A)x_k + hu_{k+\theta}\right] + C(1-\gamma)x_k, \end{aligned} \tag{138}$$

that is for $h \to 0$ and $\sigma_{k+\gamma} = h\lambda_{k+\gamma}$

$$\begin{cases} w_{k+\gamma} = \gamma CB\sigma_{k+\gamma} + a_{k+\gamma} + Cx_k \\ \mathbf{K}^* \in w_{k+\gamma} \perp \sigma_{k+\gamma} \in \mathbf{K}. \end{cases} \tag{139}$$

We can see that the matrix of the LCP is multiplied by γ. Since D is not full rank and $Cx(0^-) \leqslant 0$ the system initially undergoes a state jump. One can see in Fig. 4 that the $(\frac{1}{2}, \frac{1}{2})$-method fails to estimate the jump properly, whereas the $(\frac{1}{2}, 1)$-method jumps to the correct state. Unsurprisingly, due to the incorrect jump approximation, the storage and dissipation functions (Fig. 5) fail to be approximated by the $(\frac{1}{2}, \frac{1}{2})$-method.

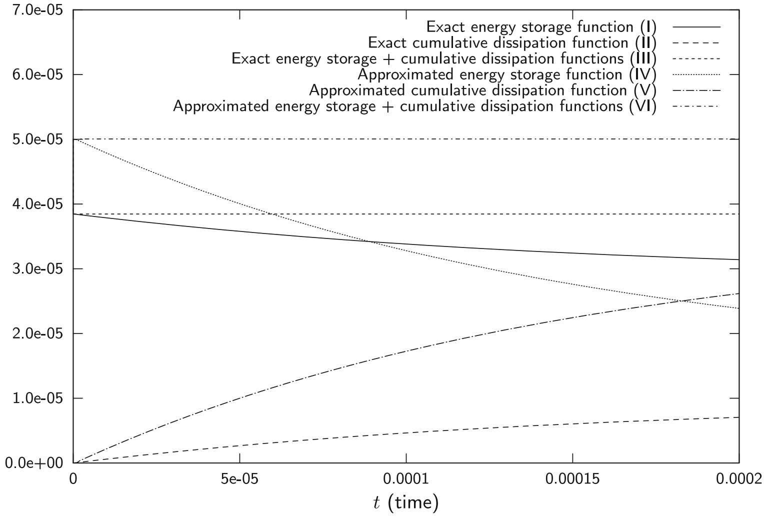

Fig. 5 Example 4. Capfilter : Exact storage function (*solid line I*), exact cumulative dissipation function (*long-dashed line II*), exact storage function + cumulative dissipation function (*short–dashed line III*), $(\frac{1}{2}, \frac{1}{2})$-method approximation of storage function (*dotted line IV*), $(\frac{1}{2}, \frac{1}{2})$-method approximation of cumulative dissipation function (*long-dashed-dotted line V*), $(\frac{1}{2}, \frac{1}{2})$-method approximation of storage function + cumulative dissipation function (*long-dashed-dotted line VI*). Time step $h = 1.0 \times 10^{-6}$

It follows from the above analysis that the (θ, γ)-method is consistently approximating the state jumps only if $\gamma = 1$, *i.e.* for schemes that are fully implicit in the multiplier.

5 Conclusions and perspectives

The main results of this paper and some perspectives may be summarized now:

- It offers a systematic framework to study the dissipativity properties preservation after the discretization with the (θ, γ)-method;
- It provides a rigorous definition of the numerical dissipation;
- It explains why state lossless continuous-time systems are more easily transformed into state lossless discrete-time systems, than state dissipative systems;
- It examines the consistency of state jumps approximations, and shows that only fully implicit (in the multiplier) methods yield consistency;
- It presents several examples of circuits containing multivalued nonsmooth components (ideal diodes, Zener diodes) to illustrate the developments;
- The framework may be extended to other numerical schemes like the zero order hold method that is used for feedback control purposes; it may also be used to study if other methods like multi-step methods (applied on the state only) may improve the dissipativity preservation.
- From a practical point of view it may be recommended to choose $\theta = \gamma = \frac{1}{2}$ for systems with no state jumps, and $\theta = \frac{1}{2}, \gamma = 1$ for systems with state jumps.

Appendix A: Some results on LCPs

Lemma 4 *Let us assume that* $D \in \mathbb{R}^{m\times m}$ *is a semi-definite positive matrix. Let us define* $\mathbf{Q} = \mathrm{SOL}(D, 0) = \{z \mid z \geqslant 0, Dz \geqslant 0, z^T Dz = 0\}$ *and* $\mathbf{R} = \{q \mid \mathrm{SOL}(D, q) \neq \emptyset\}$. *Then, we have*

a) $z^T Dz = 0 \Longleftrightarrow (D + D^T)z = 0$
b) $\mathbf{Q} = \{z \in \mathbb{R}^m \mid D^T z \leqslant 0, z \geqslant 0\}$
c) $\mathbf{R} = \mathbf{Q}^* = \mathbb{R}^m_+ - D\mathbb{R}^m_+$
d) $\mathbf{R}^* = \mathbf{Q}$

Proof a) Let us consider the following convex quadratic programming problem

$$\min \frac{1}{2} z^T Dz \tag{140}$$

which is equivalent to:

$$\min z^T (D + D^T)z. \tag{141}$$

Since $z^T (D + D^T)z \geqslant 0$ and the bound is reached for $z = 0$, the solution of (44) is also a solution of $z^T Dz = 0$. Since the problem is convex, the KKT conditions are:

$$(D + D^T)z = 0, \tag{142}$$

and they are equivalent to (141). Finally, we conclude that:

$$z^T Dz = 0 \Longleftrightarrow \min \frac{1}{2} z^T Dz = 0 \Longleftrightarrow (D + D^T)z = 0. \tag{143}$$

b) If $v \in \{z \mid D^T z \leqslant 0, z \geqslant 0\}$, we have:

$$v^T D^T v \leqslant 0, \tag{144}$$

which implies:

$$v^T D^T v = 0, \tag{145}$$

since $D^T \geqslant 0$. Hence, $Dv = -D^T v \geqslant 0$ and $v \in \mathbf{Q}$. Conversely, if $v \in \mathbf{Q}$, we have $Dv = -D^T v \geqslant 0$.

c) If $q \in \mathbf{Q}^*$, then LCP(D, q) is solvable [16, Theorem 3.8.6]. Hence, $\mathbf{Q}^* \subset \mathbf{R}$. If $q \in \mathbf{R}$, $\exists x, w \in \mathbb{R}^m$ such that

$$\begin{cases} w = Dx + q \\ 0 \leqslant x \perp w \geqslant 0. \end{cases} \tag{146}$$

Since $\mathbf{Q} = \{z \in \mathbb{R}^m \mid D^T z \leqslant 0, z \geqslant 0\}$, the dual cone $\mathbf{Q}^*$ can be expressed as [38, p 122]

$$\mathbf{Q}^* = \{v \in \mathbb{R}^m \mid v = [\, I; \;\; -D\,]\alpha, \alpha \in \mathbb{R}_+^{2m}\}, \tag{147}$$

that is

$$\mathbf{Q}^* = \mathbb{R}_+^m - D\mathbb{R}_+^m. \tag{148}$$

From (146), we get

$$q = w - Dx, w \geqslant 0, x \geqslant 0. \tag{149}$$

Hence, $q \in \mathbf{Q}^*$ if we choose $\alpha^T = [w^T x^T]$.

d) Since $D \geqslant 0$, the set of solution of LCP(0, D) is a closed convex cone; therefore $[\mathbf{Q}^*]^* = \mathbf{Q} = \mathbf{R}^*$ [38, Theorem 14.1].

□

Appendix B: Proof of Lemma 1

(i) We will first prove that

$$h(A^T R + RA) + h^2(1 - 2\theta)A^T RA = -(I - h\theta A)^T \widetilde{L}\widetilde{L}^T (I - h\theta A). \tag{150}$$

In the derivation of Eq. (150) we make use of the fact that $(I - \mu A)(I + \eta A) = (I + \eta A)(I - \mu A)$ for any reals μ and η. Recalling Eq. (15) we have,

$$\widetilde{A}^T R \widetilde{A} - R = -\widetilde{L}\widetilde{L}^T. \tag{151}$$

Using the definition that $\widetilde{A} = (I - h\theta A)^{-1}(I + h(1 - \theta)A)$ yields:

$$(I + h(1 - \theta)A)^T (I - h\theta A)^{-T} R(I - h\theta A)^{-1}(I + h(1 - \theta)A) - R = -\widetilde{L}\widetilde{L}^T. \tag{152}$$

Multiplying (on the left) by $(I - h\theta A)^T$ and (on the right) by $(I - h\theta A)$ yields:

$$\begin{aligned}(I - h\theta A)^T ((I + h(1 - \theta)A)^T (I - h\theta A)^{-T} R(I - h\theta A)^{-1} \\ (I + h(1 - \theta)A) - R)(I - h\theta A) = -(I - h\theta A)^T \widetilde{L}\widetilde{L}^T (I - h\theta A),\end{aligned} \tag{153}$$

and thus:

$$\begin{aligned}&(I - h\theta A)^T (I + h(1 - \theta)A)^T (I - h\theta A)^{-T} R(I - h\theta A)^{-1} \\ &(I + h(1 - \theta)A)(I - h\theta A) - (I - h\theta A)^T R(I - h\theta A) = -(I - h\theta A)^T \\ &\widetilde{L}\widetilde{L}^T (I - h\theta A).\end{aligned} \tag{154}$$

Using the commutativity of $(I+h(1-\theta)A)(I-h\theta A)$ (and likewise the commutativity of its transpose), we obtain:

$$(I+h(1-\theta)A)^T(I-h\theta A)^T(I-h\theta A)^{-T}R(I-h\theta A)^{-1}(I-h\theta A)(I+h(1-\theta)A) \tag{155}$$
$$-(I-h\theta A)^T R(I-h\theta A) = -(I-h\theta A)^T\widetilde{L}\widetilde{L}^T(I-h\theta A).$$

Simplifying:

$$(I+h(1-\theta)A)^T R(I+h(1-\theta)A) - (I-h\theta A)^T R(I-h\theta A) \tag{156}$$
$$= -(I-h\theta A)^T\widetilde{L}\widetilde{L}^T(I-h\theta A).$$

Expanding yields:

$$(R+h(1-\theta)(A^TR+RA)+h^2(1-\theta)^2A^TRA)-(R-h\theta(A^TR+RA)+h^2\theta^2A^TRA) \tag{157}$$
$$= -(I-h\theta A)^T\widetilde{L}\widetilde{L}^T(I-h\theta A).$$

Collecting terms by powers of h yields (150). Inserting (2) into (150) (which is equivalent to (15)) and then using $P = hR$, one finds (28). Conversely, suppose that (28) holds. By the passivity assumption one has that (150) holds also, so by equalling both (28) and (150) one finds that $-LL^T = h(A^TR+RA) = A^TP+PA$, so that indeed $R = \frac{1}{h}P$ from the uniqueness property of solutions to Lyapunov equations [Theorem A.7] [9].

(ii) Recalling Eq. (16) we have,

$$\widetilde{B}^T R\widetilde{A} - \widetilde{C} = -\widetilde{W}^T\widetilde{L}^T. \tag{158}$$

Using the definitions of $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ from (13) yields:

$$hB^T(I-h\theta A)^{-T}R(I-h\theta A)^{-1}(I+h(1-\theta)A) - \gamma C(I-h\theta A)^{-1}(I+h(1-\theta)A) - (1-\gamma)C \tag{159}$$
$$= -\widetilde{W}^T\widetilde{L}^T.$$

Using $h(A^TR+RA) = -LL^T$, we have that $(I-h\theta A)^T R = R(I+h\theta A)+\theta LL^T$ and thus $(I-h\theta A)^{-T}R = R(I+h\theta A)^{-1} - \theta(I-h\theta A)^{-T}LL^T(I+h\theta A)^{-1}$, so that we obtain:

$$B^T[hR(I+h\theta A)^{-1}-\theta(I-h\theta A)^{-T}LL^T(I+h\theta A)^{-1}](I-h\theta A)^{-1}(I+h(1-\theta)A) \tag{160}$$
$$-\gamma C(I-h\theta A)^{-1}(I+h(1-\theta)A)-(1-\gamma)C = -\widetilde{W}^T\widetilde{L}^T.$$

For the next step we multiply (on the right) by $(I-h\theta A)$ and $(I+h\theta A)$ and use the commutativity feature of matrices of the form $(I+\mu A)(I-\eta A)$ in order to cancel out matrices $(I+h\theta A)^{-1}$ and $(I-h\theta A)^{-1}$, and thus obtain:

$$\begin{aligned} B^T[hR-\theta(I-h\theta A)^{-T}LL^T](I+h(1-\theta)A)-[\gamma C(I+h(1-\theta)A) \\ +(1-\gamma)C(I-h\theta A)](I+h\theta A)=-\widetilde{W}^T\widetilde{L}^T(I-h\theta A)(I+h\theta A). \end{aligned} \tag{161}$$

Simplifying the part of the equality involving C yields:

$$\begin{aligned} B^T(hR-\theta(I-h\theta A)^{-T}LL^T)(I+h(1-\theta)A)-C(I+h(\gamma-\theta)A)(I+h\theta A) \\ =-\widetilde{W}^T\widetilde{L}^T(I-h\theta A)(I+h\theta A). \end{aligned} \tag{162}$$

Imposing the condition (3), that is $C=hB^TR+W^TL^T$, we obtain:

$$\begin{aligned} B^T(hR-\theta(I-h\theta A)^{-T}LL^T)(I+h(1-\theta)A)-(hB^TR+W^TL^T) \\ (I+h(\gamma-\theta)A)(I+h\theta A)=-\widetilde{W}^T\widetilde{L}^T(I-h\theta A)(I+h\theta A). \end{aligned} \tag{163}$$

Moving the part of the equality involving W^TL^T on the left-hand side to the right-hand side yields:

$$\begin{aligned} B^T\left(hR-\theta(I-h\theta A)^{-T}LL^T\right)(I+h(1-\theta)A)-hB^TR(I+h(\gamma-\theta)A)(I+h\theta A) \\ =W^TL^T(I+h(\gamma-\theta)A)(I+h\theta A)-\widetilde{W}^T\widetilde{L}^T(I-h\theta A)(I+h\theta A). \end{aligned} \tag{164}$$

Collecting terms involving hB^TR yields:

$$\begin{aligned} hB^TR((I+h(1-\theta)A)-(I+h(\gamma-\theta)A)(I+h\theta A))-\theta B^T(I-h\theta A)^{-T}LL^T \\ (I+h(1-\theta)A)=W^TL^T(I+h(\gamma-\theta)A)(I+h\theta A)-\widetilde{W}^T\widetilde{L}^T(I-h\theta A)(I+h\theta A). \end{aligned} \tag{165}$$

(iii) For (17) we once again use $(\widetilde{A},\widetilde{B},\widetilde{C},\widetilde{D})$ defined by (13) to get:

$$\begin{aligned} &h^2B^T(I-h\theta A)^{-T}R(I-h\theta A)^{-1}B-h\gamma C(I-h\theta A)^{-1}B \\ &-h\gamma B^T(I-h\theta A)^{-T}C^T-D^T-D=-\widetilde{W}^T\widetilde{W}. \end{aligned} \tag{166}$$

Using the continuous conditions $D^T+D=W^TW$ and $C=hB^TR+W^TL^T$ we obtain:

$$\begin{aligned} &h^2B^T(I-h\theta A)^{-T}R(I-h\theta A)^{-1}B-h\gamma(hB^TR+W^TL^T)(I-h\theta A)^{-1}B \\ &-h\gamma B^T(I-h\theta A)^{-T}(hB^TR+W^TL^T)^T-W^TW=-\widetilde{W}^T\widetilde{W}. \end{aligned} \tag{167}$$

Rearranging so that terms involving W and L are on the right-hand side yields:

$$h^2 B^T (I-h\theta A)^{-T} R(I-h\theta A)^{-1} B - h^2\gamma B^T R(I-h\theta A)^{-1} B - h^2\gamma B^T (I-h\theta A)^{-T} \tag{168}$$
$$RB = W^T W - \widetilde{W}^T \widetilde{W} + h\gamma W^T L^T (I - h\theta A)^{-1} B + h\gamma B^T (I - h\theta A)^{-T} LW.$$

Factoring the left-hand side by $B^T(I - h\theta A)^{-T}$ (from the left) and $(I - h\theta A)^{-1} B$ (from the right) yields:

$$\begin{aligned} &h^2 B^T (I - h\theta A)^{-T} (R - \gamma R(I - h\theta A) - \gamma (I - h\theta A)^T R)(I - h\theta A)^{-1} B \\ &= W^T W - \widetilde{W}^T \widetilde{W} + h\gamma W^T L^T (I - h\theta A)^{-1} B + h\gamma B^T (I - h\theta A)^{-T} LW. \end{aligned} \tag{169}$$

Collecting the left-hand side by powers of h:

$$\begin{aligned} &h^2 B^T (I - h\theta A)^{-T} ((1 - 2\gamma)R + h\theta\gamma(A^T R + RA))(I - h\theta A)^{-1} B \\ &= W^T W - \widetilde{W}^T \widetilde{W} + h\gamma W^T L^T (I - h\theta A)^{-1} B + h\gamma B^T (I - h\theta A)^{-T} LW. \end{aligned} \tag{170}$$

Finally using $h(A^T R + RA) = -LL^T$ yields (30).

References

1. Acary, V.: Higher order event capturing time-stepping schemes for nonsmooth multibody systems with unilateral constraints and impacts. Appl. Numer. Math. **62**, 1259–1275 (2012)
2. Acary, V., Bonnefon, O., Brogliato, B.: Nonsmooth modeling and simulation for switched circuits. In: Lecture Notes in Electrical Engineering, Vol. 69. Springer, Dordrecht, xxiii (2011)
3. Acary, V., Brogliato, B.: Numerical methods for nonsmooth dynamical systems. Applications in mechanics and electronics. In: Lecture Notes in Applied and Computational Mechanics, Vol. 35. Springer, Berlin, xxi (2008)
4. Acary, V., Brogliato, B., Goeleven, D.: Higher order Moreau's sweeping process: mathematical formulation and numerical simulation. Math. Programm. Ser. A **113**(1), 133–217 (2008)
5. Acary, V., Pérignon, F.: An introduction to Siconos. Technical Report TR-0340, INRIA. http://hal.inria.fr/inria-00162911/en/, (2007)
6. Bemporad, A., Bianchini, G., Brogi, F.: Passivity analysis and passification of discrete-time hybrid systems. IEEE Trans. Autom. Control **53**(4), 1004–1009 (2008)
7. Brogliato, B.: The absolute stability problem and the lagrange-dirichlet theorem with monotone multivalued mappings. Syst. Control Lett. **51**(5), 343–353 (2004)
8. Brogliato, B., Goeleven, D.: Well-posedness, stability and invariance results for a class of multivalued Lur'e dynamical systems. Nonlinear Anal Theory Methods Appl Ser A Theory Methods **74**(1), 195–212 (2011)
9. Brogliato, B., Lozano, R., Maschke, B., Egeland, O.: Dissipative Systems Analysis and Control, 2nd edn. Springer, Berlin (2007)
10. Brogliato, B., Thibault, L.: Well-posedness results for non-autonomous complementarity systems. J. Convex Anal. **17**(3–4), 961–990 (2010)
11. Camlibel, M.K., Heemels, W.P.M.H., Schumacher, J.M.: Consistency of a time-stepping method for a class of piecewise-linear networks. IEEE Trans. Circuits Syst. **I**(49), 349–357 (2002)
12. Camlibel, M.K., Heemels, W.P.M.H., Schumacher, J.M.: On linear passive complementarity systems. Eur. J. Control **8**, 220–237 (2002)
13. Camlibel, M.K., Heemels, W.P.M.H., van der Schaft, A.J., Schumacher, J.M.: Switched networks and complementarity. IEEE Trans. Circuits Syst. I Fundam. Theory Appl. **1036**(8), 46–50 (2003)

14. Camlibel, M.K., Iannelli, L., Vasca, F.: Passivity and complementarity. Technical Report 352, GRACE Internal Report. http://www.grace.ing.unisannio.it (2006)
15. Costa Catello, R., Fossas, E.: On preserving passivity in sampled-data linear systems. Eur. J. Control **13**(6), 583–590 (2007)
16. Cottle, R.W., Pang, J., Stone, R.E.: Linear Complement. Probl. Academic Press, Inc., Boston (1992)
17. de la Sen, M.: Preserving positive realness through discretization. Positivity **6**(1), 31–45 (2002)
18. Facchinei, F., Pang, J.S.: Finite-dimensional variational inequalities and complementarity problems. In: Springer Series in Operations Research, Vol. I, II. Springer, New York (2003)
19. Faurre, P.: Réalisations Markoviennes de Processus Stationnaires. PhD thesis, University Paris VI (1972)
20. Faurre, P., Clerget, M., Germain, F.: Opérateurs rationnels positifs. Application à l'hyperstabilité et aux processus aléatoires. Méthodes Mathématiques de l'Informatique, Vol. 8. Paris: Dunod, Bordas. IX, F 190.00 (1979)
21. Frasca, R., Camlibel, M.K., Goknar, I.C., Iannelli, L., Vasca, F.: Linear passive networks with ideal switches: consistent initial conditions and state discontinuities. IEEE Trans. Circuits Syst. I Regular Pap. **57**(12), 3138–3151 (2010)
22. Frasca, R., Camlibel, M.K., Goknar, I.C., Vasca, F.: State jump rules in linear passive networks with ideal switches. In: IEEE International Symposium on Circuits and Systems (ISCAS), Seattle (2008)
23. Greenhalgh, S., Acary, V., Brogliato, B.: Preservation of the dissipativity properties of a class of nonsmooth dynamical systems with the (θ, γ)-algorithm. Research Report RR-7632, INRIA, May 2011
24. Han, L., Tiwari, A., Camlibel, K., Pang, J.S.: Convergence of time-stepping schemes for passive and extended linear complementarity systems. SIAM J. Numer. Anal. **47**(5), 3768–3796 (2009)
25. Hiriart-Urruty, J.B., Lemaréchal, C.: Fundamentals of Convex Analysis. Springer, Berlin (2001)
26. Hoagg, J.B., Lacy, S.L., Erwin, R.S., Bernstein, D.S.: First-order-hold sampling of positive real systems and subspace identification of positive real models. In: Proceedings of the 2004. American Control Conference, Vol. 1, pp. 861–866, 30 2004–july 2 2004 (2004)
27. Iannelli, L., Vasca, F., Angelone, G.: Computation of steady-state oscillations in power converters through complementarity. IEEE Trans. Circuits Syst. I Regul. Papers **58**(6), 1421–1432 (2011)
28. Jiang, J.: Preservations of positive realness under discretizations. J. Franklin Inst. **330**(4), 721–734 (1993)
29. Laila, D.S., Nesic, D., Teel, A.R.: Open and closed loop dissipation inequalities under sampling and controller emulation. Eur. J. Control **8**(2), 109–125 (2002)
30. Lancaster, P., Tismenetsky, M.: Theory Of Matrices, 2nd edn. Academic Press, New York (1997)
31. Monteiro Marques, M.D.P.: Differential inclusions in nonsmooth mechanical problems. Shocks and dry friction. In: Progress in Nonlinear Differential Equations and their Applications, Vol. 9. Birkhauser, Basel (1993)
32. Moreau, J.J.: Evolution problem associated with a moving convex set in a Hilbert space. J. Differ. Equ. **26**, 347–374 (1977)
33. Moreau, J.J.: Liaisons unilatérales sans frottement et chocs inélastiques. Comptes Rendus de l'Académie des Sciences, **296**(série II), 1473–1476 (1983)
34. Moreau, J.J.: Bounded variation in time. In: Moreau, J.J., Panagiotopoulos, P.D., Strang, G. (eds.) Topics in Nonsmooth Mechanics, pp. 1–74. Bikhäuser, Basel (1988)
35. Nesic, D., Laila, D.S., Teel, A.R.: On preservation of dissipation inequalities under sampling. In: Proceedings of the 39th IEEE Conference on Decision and Control, 2000, Vol. 3, pp. 2472–2477 (2000)
36. Oishi, Y.: Passivity degradation under the discretization with the zero-order hold and the ideal sampler. In: 49th IEEE Conference on Decision and Control (CDC), pp. 7613–7617 (2010)
37. Pang, J.-S.: Three modeling paradigms in mathematical programming. Math. programm. Ser. B **125**, 297–323 (2010)
38. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
39. Schindler, T., Acary, V.: Timestepping schemes for nonsmooth dynamics based on discontinuous Galerkin methods: definition and outlook. Math. Comput. Simul. (In Press). doi:10.1016/j.matcom.2012.04.012 (2012)
40. Shen, J., Pang, J.-S.: Semicopositive linear complementarity systems. Int. J. Robust Nonlinear Control **17**(15), 1386–1667 (2007)