Sensitivity and Optimal Control Theory for Linear Complementarity Systems

Peter Stechlinski, Hesham Abdelfattah, and Sameh A. Eisa

Abstract This article focuses on sensitivity and control theory for linear complementarity systems (LCSs), a type of dynamical system that experiences hybrid continuous/discrete behavior and is therefore nonsmooth. In particular, a sensitivity theory is given that characterizes generalized derivative information of solutions of LCSs with respect to parametric perturbations. With this theory in hand, a computationally-relevant open-loop optimal control theory is provided using a direct method (i.e., the control is parametrically discretized and generalized gradients of the objective function are described). The approach here is based on lexicographic directional differentiation theory, a relatively new tool in nonsmooth analysis, being applied to nonlinear complementarity systems (NCSs). The optimal control theory is illustrated with an example. As a byproduct of the sensitivity theory, well-posedness results for a new class of hybrid dynamical system, called the lexicographic linear complementarity system (LexLCS), are also established.

Keywords Complementarity systems · Hybrid systems · Generalized derivatives · Sensitivity analysis · Optimal control

---

P. Stechlinski (✉)
University of Maine, Orono, ME, USA
e-mail: peter.stechlinski@maine.edu

H. Abdelfattah · S. A. Eisa
University of Cincinnati, Cincinnati, OH, USA
e-mail: abdelfhm@mail.uc.edu

S. A. Eisa
e-mail: eisash@ucmail.uc.edu

D. M. Kilgour et al. (eds.), *Addressing Modern Challenges in the Mathematical, Statistical, and Computational Sciences*, Springer Proceedings in Mathematics & Statistics 494, https://doi.org/10.1007/978-3-031-84869-8_44

1 Introduction

Nonlinear complementarity systems (NCSs) are systems of ODEs that are coupled with complementarity constraints that cause the system to experience "switches" between modes whenever the active constraint changes. Because of this, NCSs display hybrid-like behavior (i.e., a mixture of continuous and discrete events). The NCS modeling framework finds use in a variety of problems, such as electrical systems, mechanical systems, and economics [18]. NCSs were formalized and their basic properties were studied in [16, 17] (see also [18]).

Linear complementarity systems (LCSs) [9] are an important subclass of NCSs with linear RHS functions. Authors have investigated their well-posedness [9], non-Zenoness [19, 20], passivity [3], stability [4], observer-based stabilization control [8], and controllability and stabilizability [2]. In the paper by Vieira et al. [27], the authors developed quadratic optimal control theory and methods for LCSs by deriving necessary optimality conditions (and provided sufficient conditions in some special cases). Vieira et al. [26] established theory and methods for minimal time control problems for NCSs and then specialized their results to LCSs, because of difficult-to-verify conditions, challenges in practical implementations, etc. in the general nonlinear setting (see their discussion in [26, Sect. 3.1]).

In the papers [21, 22], the author uses nonsmooth differential-algebraic equation (DAE) theory to analyze NCSs, including their sensitivity to parameters and optimal control, under an appropriate regularity condition (strong regularity of [14]). In this paper, we are concerned with establishing optimal control theory and methods for LCSs, using the above-mentioned theory of NCSs. In particular, we use a direct method (sequential approach), wherein the control is parametrically discretized by a finite truncation of basis functions so as to use sensitivity theory to determine how to update parameters to decrease the objective function value. This is in contrast to the indirect approaches outlined above (whose methodology falls under the umbrella of calculus of variations/Pontryagin's maximum principle). See [10] for more details. In doing so, we establish some existence/uniqueness results for a new class of hybrid dynamical system that arises from the sensitivity system associated with an LCS. In summary, we make the following contributions: (i) a sensitivity theory for parametric LCSs is given (Sect. 2); (ii) a computationally-relevant optimal control method for LCSs is provided (Sect. 3); (iii) simulations are shown to illustrate the optimal control results (Sect. 4); and (iv) well-posedness theory for a new class of hybrid dynamical system, called the lexicographic linear complementarity system, is established (Sect. 5). Future directions are given in Sect. 6.

2 Parametric LCS Sensitivity Theory

In this part, we give sensitivity theory for the following parametric LCS:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(\mathbf{p})\mathbf{x} + \mathbf{B}(\mathbf{p})\mathbf{v}, \quad \mathbf{x}(0) = \mathbf{f}_0(\mathbf{p}), \tag{1a}$$

$$\mathbf{0} \le \mathbf{v} \perp \mathbf{C}(\mathbf{p})\mathbf{x} + \mathbf{D}(\mathbf{p})\mathbf{v} \ge \mathbf{0}, \tag{1b}$$

with independent variable t, differential and algebraic state variables $\mathbf{x}(t) \in \mathbb{R}^{n_x}$ and $\mathbf{v}(t) \in \mathbb{R}^{n_v}$, respectively, matrix-valued C^1 functions $\mathbf{A} : D_p \to \mathbb{R}^{n_x \times n_x}$, $\mathbf{B} : D_p \to \mathbb{R}^{n_x \times n_v}$, $\mathbf{C} : D_p \to \mathbb{R}^{n_v \times n_x}$, $\mathbf{D} : D_p \to \mathbb{R}^{n_v \times n_v}$, initial value C^1 function $\mathbf{f}_0 : D_p \to \mathbb{R}^{n_x}$, and problem parameters $\mathbf{p} \in D_p \subseteq \mathbb{R}^{n_p}$. The complementarity conditions are understood in the following sense: given two vectors $\mathbf{v}$ and $\mathbf{y}$, $\mathbf{0} \le \mathbf{v} \perp \mathbf{y} \ge \mathbf{0}$ is equivalent to $\mathbf{v} \ge \mathbf{0}$ and $\mathbf{y} \ge \mathbf{0}$ (component-wise) and $\mathbf{v}^{\mathrm{T}}\mathbf{y} = \mathbf{0}$. The complementarity conditions implicitly define an "active set" of the constraints, with the dynamical system experiencing "switches" between different modes whenever the active set changes: Given a solution $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{v}})$ on $[0, t_f]$ through initial data $(\mathbf{p}_0, \mathbf{x}_0, \mathbf{v}_0)$, we define the inactive, weakly active, and strongly active index sets associated with (1b), respectively, as follows:

$$\alpha(t) = \{i : \widetilde{v}_i > 0 = \widetilde{y}_i\},\ \beta(t) = \{i : \widetilde{v}_i = 0 = \widetilde{y}_i\},\ \gamma(t) = \{i : \widetilde{v}_i = 0 < \widetilde{y}_i\}, \tag{2}$$

where $\widetilde{\mathbf{y}}(t) = \mathbf{C}(\mathbf{p}_0)\widetilde{\mathbf{x}}(t) + \mathbf{D}(\mathbf{p}_0)\widetilde{\mathbf{v}}(t)$. The LCS system satisfies the P-property if $\mathbf{D}$ is a P-matrix (all its principal minors are positive). In this case, the embedded linear complementarity system in (1b) is uniquely solvable given any $\mathbf{Cx}$ [7], and it follows that the solution $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{v}})$ of the LCS is strongly regular in the sense of [14] (i.e., the matrix-theoretic condition on [14, Page 243]) and unique.

However, because of the complementarity conditions in (1b), the solutions $\widetilde{\mathbf{x}}$ and $\widetilde{\mathbf{v}}$ of the LCS system may behave nonsmoothly with respect to parametric perturbations. Thus, the classical (forward) parametric sensitivity functions $\mathbf{S}_{\mathbf{x}}(t) = \frac{\partial \widetilde{\mathbf{x}}}{\partial \mathbf{p}}(t, \mathbf{p}_0)$ and $\mathbf{S}_{\mathbf{v}}(t) = \frac{\partial \widetilde{\mathbf{v}}}{\partial \mathbf{p}}(t, \mathbf{p}_0)$ may not exist. Consequently, we characterize the solution parametric sensitivities using generalized first-order derivative information instead, using the concept of Clarke [5]; given a locally Lipschitz function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$, its Clarke generalized derivative [5] at $\mathbf{z} = \mathbf{z}_0$ is

$$\partial \mathbf{F}(\mathbf{z}_0) = \mathrm{conv}\{\mathbf{H} \in \mathbb{R}^{m \times n} : \exists\, \mathbf{z}_n \to \mathbf{z}_0 \text{ s.t. } \mathbf{z}_j \in \mathrm{dom}\mathbf{F} \setminus \Omega_{\mathbf{F}}\ \forall j \text{ and } \mathbf{JF}(\mathbf{z}_n) \to \mathbf{H}\},$$

where $\mathbf{JF}(\mathbf{z}_j) \in \mathbb{R}^{m \times n}$ is the Jacobian matrix of $\mathbf{F}$ evaluated at $\mathbf{z}_j$ and $\Omega_{\mathbf{F}}$ is the zero measure subset on which $\mathbf{F}$ is not differentiable. For example, for $F(z) = |z|$, $\partial F(0) = \mathrm{conv}\{-1, 1\} = [-1, 1]$. Such first-order generalized derivative information is computationally relevant in dedicated nonsmooth numerical methods (such as nonsmooth equation solving and nonsmooth optimization—see [1]). However, it is difficult to calculate generalized derivative elements in an accurate and easy-to-implement way for complex problems/models, such as dynamic LCS systems (see [1] for a detailed explanation of the difficulties).

To overcome these challenges, a new tool in nonsmooth analysis, called the lexicographic directional derivative [11] can be used (again, see [1] for details) to furnish lexicographic derivatives [12] which are just as computationally relevant as Clarke generalized derivatives in the above-mentioned nonsmooth numerical methods. Using this approach, we provide theory below that describes how to obtain "lexicographic sensitivity functions"—in particular, an auxiliary sensitivity system is obtained whose unique solution yields (lexicographic) sensitivity functions $\mathbf{S}^{\mathrm{L}}_{\mathbf{x}}(t)$ and $\mathbf{S}^{\mathrm{L}}_{\mathbf{v}}(t)$, that are analogous to $\mathbf{S}_{\mathbf{x}}(t) = \frac{\partial \widetilde{\mathbf{x}}}{\partial \mathbf{p}}(t, \mathbf{p}_0)$ and $\mathbf{S}_{\mathbf{v}} = \frac{\partial \widetilde{\mathbf{v}}}{\partial \mathbf{p}}(t, \mathbf{p}_0)$, and are indistinguishable from Clarke generalized derivative information of the solutions with respect to parameters, i.e., $\partial[\widetilde{\mathbf{x}}(t,\cdot)](\mathbf{p}_0)$ and $\partial[\widetilde{\mathbf{v}}(t,\cdot)](\mathbf{p}_0)$. Said auxiliary sensitivity system is a lexicographic linear complementarity system with (matrix-valued) lexicographic complementarity conditions [21, Sect. III.B.]: given $\mathbf{V}, \mathbf{H} \in \mathbb{R}^{m\times n}$,

$$\mathbf{0} \preceq \mathbf{V} \perp \mathbf{H} \succeq \mathbf{0} \iff (\mathbf{V} \succeq \mathbf{0}) \wedge (\mathbf{H} \succeq \mathbf{0}) \wedge (\mathrm{row}_i(\mathbf{V}) = \mathbf{0} \vee \mathrm{row}_i(\mathbf{H}) = \mathbf{0}\ \forall i) \tag{3}$$

where $\mathbf{V} \succeq \mathbf{0}$ holds if, for each i, $\mathrm{fsign}(\mathrm{row}_i(\mathbf{V})) = 1$ or $\mathrm{row}_i(\mathbf{V}) = \mathbf{0}$, i.e., the first nonzero element in each row is positive or the row is the zero vector. For example,

$$\mathbf{0} \preceq \begin{bmatrix} 1 & -1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \perp \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \succeq \mathbf{0} \ \text{(✓)} \qquad \mathbf{0} \npreceq \begin{bmatrix} -1 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \perp \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \nsucceq \mathbf{0} \ \text{(X)}$$

Before presenting the main result, we introduce the folloting notation: Let $\mathbf{I}_n$ denote the $n \times n$ identity matrix. Given two matrices $\mathbf{M} \in \mathbb{R}^{m\times n}$ and $\mathbf{N} \in \mathbb{R}^{p\times q}$, let $\mathbf{M} \otimes \mathbf{N} \in \mathbb{R}^{mp\times nq}$ denote the Kronecker product of the matrices. Let $\mathbf{M}_J \in \mathbb{R}^{|J|\times n}$ denote the submatrix of $\mathbf{M} \in \mathbb{R}^{m\times n}$ constructed by its rows as indexed by the properly ordered subset $J \subseteq \{1, \ldots, m\}$. Given a matrix-valued C^1 function $\mathbf{W} : \mathbb{R}^q \to \mathbb{R}^{n\times m}$, let $\frac{\partial \mathbf{W}}{\partial \mathbf{z}}$ denote the $mn \times q$ partial derivative matrix obtained by first vectorizing $\mathbf{W}$ and then taking its partial derivative (see [13, Eq. (2.5)]). We now give the main result of this section.

Theorem 1 *Let* $\mathbf{p}_0 \in D_p$. *Let* $\mathbf{D}(\mathbf{p}_0)$ *be a P-matrix. Let* $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{v}})$ *be a solution of* (1) *on* $[0, t_f]$ *through initial data* $(\mathbf{p}_0, \mathbf{x}_0, \mathbf{v}_0)$. *Then the (lexicographic) sensitivity functions* $(\mathbf{S}^{\mathrm{L}}_{\mathbf{x}}, \mathbf{S}^{\mathrm{L}}_{\mathbf{v}})$ *uniquely solve the following (lexicographic) sensitivity system:*

$$\frac{d\mathbf{X}}{dt} = (\mathbf{I}_{n_x} \otimes \widetilde{\mathbf{x}}^{\mathrm{T}})\frac{\partial \mathbf{A}}{\partial \mathbf{p}} + \mathbf{A}\mathbf{X} + (\mathbf{I}_{n_v} \otimes \widetilde{\mathbf{v}}^{\mathrm{T}})\frac{\partial \mathbf{B}}{\partial \mathbf{p}} + \mathbf{B}\mathbf{V}, \quad \mathbf{X}(0) = \mathbf{J}\mathbf{f}_0(\mathbf{p}_0), \tag{4a}$$

$$\mathbf{0} = (\mathbf{I}_{|\alpha|} \otimes \widetilde{\mathbf{x}}^{\mathrm{T}})\frac{\partial \mathbf{C}_\alpha}{\partial \mathbf{p}} + \mathbf{C}_\alpha\mathbf{X} + (\mathbf{I}_{|\alpha|} \otimes \widetilde{\mathbf{v}}^{\mathrm{T}})\frac{\partial \mathbf{D}_\alpha}{\partial \mathbf{p}} + \mathbf{D}_\alpha\mathbf{V}, \tag{4b}$$

$$\mathbf{0} \preceq \mathbf{V}_\beta \perp (\mathbf{I}_{|\beta|} \otimes \widetilde{\mathbf{x}}^{\mathrm{T}})\frac{\partial \mathbf{C}_\beta}{\partial \mathbf{p}} + \mathbf{C}_\beta\mathbf{X} + (\mathbf{I}_{|\beta|} \otimes \widetilde{\mathbf{v}}^{\mathrm{T}})\frac{\partial \mathbf{D}_\beta}{\partial \mathbf{p}} + \mathbf{D}_\beta\mathbf{V} \succeq \mathbf{0}, \tag{4c}$$

$$\mathbf{0} = \mathbf{V}_\gamma, \tag{4d}$$

on $[0, t_f]$ *through initial data* $(\mathbf{X}_0, \mathbf{V}_0)$*, where* $\mathbf{X}_0 = \mathbf{J}\mathbf{f}_0(\mathbf{p}_0) \in \mathbb{R}^{n_x \times n_p}$ *and* $\mathbf{V}_0 \in \mathbb{R}^{n_v \times n_p}$ *uniquely solves* (4b)–(4d) *at* $t = 0$*.*

Proof Letting $\mathbf{f}(\mathbf{p}, \mathbf{x}, \mathbf{v}) \mapsto \mathbf{A}(\mathbf{p})\mathbf{x} + \mathbf{B}(\mathbf{p})\mathbf{v}$ and $\mathbf{h}(\mathbf{p}, \mathbf{x}, \mathbf{v}) \mapsto \mathbf{C}(\mathbf{p})\mathbf{x} + \mathbf{D}(\mathbf{p})\mathbf{v}$, Eq. (1) can be cast in the form

$$\tfrac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{p}, \mathbf{x}, \mathbf{v}), \quad \mathbf{x}(0) = \mathbf{f}_0(\mathbf{p}), \tag{5a}$$

$$\mathbf{0} \leq \mathbf{v} \perp \mathbf{h}(\mathbf{p}, \mathbf{x}, \mathbf{v}) \geq \mathbf{0}. \tag{5b}$$

Moreover, (5b) is equivalent to the nonsmooth algebraic equation $\mathbf{0} = \mathbf{g}(\mathbf{p}, \mathbf{x}, \mathbf{v})$ where $\mathbf{g} : (\mathbf{p}, \mathbf{x}, \mathbf{v}) \mapsto \mathbf{min}(\mathbf{v}, \mathbf{h}(\mathbf{p}, \mathbf{x}, \mathbf{v}))$. Following the arguments in [21, Sect. 5] (based on [15, Example 17]), $\mathbf{D}(\mathbf{p}_0)$ being a P-matrix implies that the mapping $\mathbf{g}$ is completely coherently oriented (in the sense of [15]) with respect to $\mathbf{v}$ at $(\mathbf{p}_0, \mathbf{x}, \mathbf{v}) \in D_p \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_v}$. It follows that $\mathbf{g}$ is completely coherently oriented with respect to $\mathbf{v}$ at $(\mathbf{p}_0, \tilde{\mathbf{x}}(t, \mathbf{p}_0), \tilde{\mathbf{v}}(t, \mathbf{p}_0))$ for each $t \in [0, t_f]$. Hence, from [21], $(\tilde{\mathbf{x}}, \tilde{\mathbf{v}})$ is a strongly regular solution (in the sense of [14]) on $[0, t_f]$. Thus, the sensitivity theory in [22, Sect. III.B] is applicable: the sensitivity system of (5) is given by [22, Eq. (18)], which, using the definitions of $\mathbf{f}$ and $\mathbf{h}$ above, setting $\mathbf{P} = \mathbf{I}_{n_p}$, and using the product rule $\frac{\partial[\mathbf{MN}]}{\partial \mathbf{z}} = (\mathbf{M} \otimes \mathbf{I}_r)\frac{\partial \mathbf{N}}{\partial \mathbf{z}} + (\mathbf{I}_n \otimes \mathbf{N}^{\mathrm{T}})\frac{\partial \mathbf{M}}{\partial \mathbf{z}}$ for matrix-valued C^1 functions $\mathbf{M}$ and $\mathbf{N}$ [13], simplifies to (4). □

Remark 1 If $\beta(t) = \emptyset$ for all $t \in [0, t_f]$, then $\mathbf{S}^{\mathrm{L}}_{\mathbf{x}}(t) = \mathbf{S}_{\mathbf{x}}(t) = \frac{\partial \tilde{\mathbf{x}}}{\partial \mathbf{p}}(t, \mathbf{p}_0)$ and $\mathbf{S}^{\mathrm{L}}_{\mathbf{v}}(t) = \mathbf{S}_{\mathbf{v}}(t) = \frac{\partial \tilde{\mathbf{v}}}{\partial \mathbf{p}}(t, \mathbf{p}_0)$ in Theorem 1.

3 A Direct Method for Solving LCS Optimal Control Problems

In this part, we apply the sensitivity theory from the previous section to solve the following quadratic optimal control LCS problem:

$$\min_{\mathbf{p}} \quad \phi(\mathbf{p}) = \int_0^{t_f} (\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} + \mathbf{v}^{\mathrm{T}}\mathbf{R}\mathbf{v} + \mathbf{u}^{\mathrm{T}}\mathbf{N}\mathbf{u})dt, \tag{6a}$$

$$\text{s.t.} \quad \frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v} + \mathbf{E}\mathbf{u}, \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{6b}$$

$$\mathbf{0} \leq \mathbf{v} \perp \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{v} + \mathbf{F}\mathbf{u} \geq \mathbf{0}, \tag{6c}$$

with state variables $(\mathbf{x}, \mathbf{v})$ and (parametrically discretized) controls $\mathbf{u}$, constant initial conditions $\mathbf{x}_0$, positive semidefinite $\mathbf{Q}$, $\mathbf{R}$ and positive definite $\mathbf{N}$, and objective function ϕ. The controls have been parametrically discretized as follows:

$$\mathbf{u}(t, \mathbf{p}) = \mathbf{u}(t, \mathbf{p}_{(1)}, \ldots, \mathbf{p}_{(n_s)}) = \psi_{(1)}(t)\mathbf{p}_{(1)} + \cdots + \psi_{(n_s)}(t)\mathbf{p}_{(n_s)}, \tag{7}$$

where $\{\psi_{(1)}, \ldots, \psi_{(n_s)}\}$ is a finite truncation of basis functions (e.g., Lagrange polynomials, Legendre polynomials, etc.), and $\mathbf{p} = (\mathbf{p}_{(1)}, \ldots, \mathbf{p}_{(n_s)}) \in D_p^{n_s}$ is a vector of control parameters. The approach here is a sequential direct method: the optimal control problem has been transformed into a finite-dimensional nonlinear optimization problem thanks to the parametric discretization of the controls (hence the "direct" part of this approach). Said optimization can be solved by obtaining gradient information associated with $\phi(\mathbf{p})$ in order to update the control parameters, in the spirit of gradient descent, for example (hence the "sequential" part of this approach). The objective function ϕ may change nonsmoothly under parametric perturbations, because of the dependence of $\mathbf{x}$ and $\mathbf{v}$ on $\mathbf{p}$, and hence we require the (lexicographic) sensitivity functions from the previous section. The main result is given as follows.

Theorem 2 *Let $\mathbf{p}_0 \in D_p^{n_s}$. Let $\mathbf{D}$ be a P-matrix. Let $\mathbf{u}$ be parameterized according to (7) with $\{\psi_{(i)}\}$ a finite truncation of C^1 basis functions. Let $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{v}})$ be a solution of (6b) and (6c) on $[0, t_f]$ through initial data $(\mathbf{p}_0, \mathbf{x}_0, \mathbf{v}_0)$. Then a generalized Clarke derivative element of the objective function ϕ in (6a) is given by*

$$\boldsymbol{\mu} = 2\widetilde{\mathbf{x}}^{\mathrm{T}}(t_f)\mathbf{Q}\mathbf{S}_{\mathbf{x}}^{\mathrm{L}}(t_f) + 2\widetilde{\mathbf{v}}^{\mathrm{T}}(t_f)\mathbf{R}\mathbf{S}_{\mathbf{v}}^{\mathrm{L}}(t_f) + 2\mathbf{u}^{\mathrm{T}}(t_f, \mathbf{p}_0)\mathbf{N}\mathbf{U}(t_f) \in \partial\phi(\mathbf{p}_0), \tag{8}$$

where $\mathbf{U}(t) = [\psi_{(1)}(t) \quad \cdots \quad \psi_{(n_s)}(t)] \otimes \mathbf{I}_{n_u}$ and where $(\mathbf{S}_{\mathbf{x}}^{\mathrm{L}}, \mathbf{S}_{\mathbf{v}}^{\mathrm{L}})$ uniquely solve the following sensitivity system on $[0, t_f]$ through initial data $(\mathbf{X}_0, \mathbf{V}_0)$:

$$\frac{d\mathbf{X}}{dt} = \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{V} + \mathbf{E}\mathbf{U}, \quad \mathbf{X}(0) = \mathbf{0}, \tag{9a}$$

$$\mathbf{0} = \mathbf{C}_\alpha\mathbf{X} + \mathbf{D}_\alpha\mathbf{V} + \mathbf{F}_\alpha\mathbf{U}, \quad \mathbf{0} \preceq \mathbf{V}_\beta \perp \mathbf{C}_\beta\mathbf{X} + \mathbf{D}_\beta\mathbf{V} + \mathbf{F}_\beta\mathbf{U} \succeq \mathbf{0}, \quad \mathbf{0} = \mathbf{V}_\gamma, \tag{9b}$$

where $\mathbf{X}_0 = \mathbf{0}$ and $\mathbf{V}_0 \in \mathbb{R}^{n_v \times n_p}$ uniquely solves (9b) at $t = 0$.

Proof The proof follows from [22, Theorem II.1] with the following considerations: (i) the sensitivity system in Equation (6) in [22, Theorem II.1] is replaced by (9) above; (ii) we add an auxiliary variable $\mathbf{x}_{\text{aux}}$ via $\frac{d\mathbf{x}_{\text{aux}}}{dt} = \mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} + \mathbf{v}^{\mathrm{T}}\mathbf{R}\mathbf{v} + \mathbf{u}^{\mathrm{T}}\mathbf{N}\mathbf{u}$, $\mathbf{x}_{aux}(0) = \mathbf{0}$; and (iii) we set $\mathbf{Q} = \mathbf{I}_{n_p n_s}$ in [22, Theorem II.1] (which is not to be confused with the $\mathbf{Q}$ appearing in (6a) above), with $\mathbf{p}$ and $\mathbf{P}$ in [22, Theorem II.1] absent here. For Item (i), we define the mappings $\mathbf{f}(t, \mathbf{p}, \mathbf{x}, \mathbf{v}) \mapsto \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v} + \mathbf{E}\mathbf{u}(t, \mathbf{p})$ and $\mathbf{h}(\mathbf{p}, \mathbf{x}, \mathbf{v}) \mapsto \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{v} + \mathbf{F}\mathbf{u}(t, \mathbf{p})$ (in this case, $\mathbf{g}$ is absent) and proceed in a similar fashion to the proof of Theorem 1 to derive the sensitivity system (9) . Item (ii) is to transform the Lagrange form of ϕ here into the Mayer form of ϕ as used in [22]. Item (iii) is to furnish lexicographic derivatives, as is done in Theorem 1 above. With this in place, a generalized gradient of ϕ can be obtained via [22, Eq. (5)] as $\boldsymbol{\eta} = \mathbf{S}_{\mathbf{x}_{\text{aux}}}^{\mathrm{L}}(t_f)$, which simplifies to (8). □

4 Simulations

In this part, we illustrate the theory with an example. First, we provide a special case of the optimal control method in Sect. 3 for piecewise constant controls in the differential equations (i.e., $\mathbf{F} = \mathbf{0}$ in (6c)): suppose that (7) is replaced by:

$$\mathbf{u}(t, \mathbf{p}) = \mathbf{p}_{(i)} \quad \text{if } t \in (\tau_{i-1}, \tau_i], \quad i = 1, 2, \ldots, n_s, \tag{10}$$

where the time horizon $[0, t_f] = \bigcup[\tau_{i-1}, \tau_i]$ is split up into n_s subintervals, i.e., $\tau_0 = 0$ and $\tau_{n_s} = t_f$, and where $\mathbf{p}_{(i)} \in D_p$ as above. Then, following the same arguments as in the proof of Theorem 2, combined with the piecewise constant control theory from [23, Example 3.4], we get the following if $\mathbf{D}$ is a P-matrix: a generalized Clarke derivative element of the objective function ϕ in (6a) is given by

$$\boldsymbol{\mu} = 2\widetilde{\mathbf{x}}^{\mathrm{T}}(t_f)\mathbf{Q}\mathbf{S}_{\mathbf{x}}^{\mathrm{L}}(t_f) + 2\widetilde{\mathbf{v}}^{\mathrm{T}}(t_f)\mathbf{R}\mathbf{S}_{\mathbf{v}}^{\mathrm{L}}(t_f) + 2\mathbf{p}_{(n_s)}^{\mathrm{T}}\mathbf{N}\mathbf{P}(t_f) \in \partial\phi(\mathbf{p}_0),$$

where $(\mathbf{S}_{\mathbf{x}}^{\mathrm{L}}, \mathbf{S}_{\mathbf{v}}^{\mathrm{L}})$ uniquely solve the following sensitivity system on $[0, t_f]$ through initial data $(\mathbf{X}_0, \mathbf{V}_0)$:

$$\frac{d\mathbf{X}}{dt} = \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{V} + \mathbf{E}\mathbf{P}, \quad \mathbf{X}(0) = \mathbf{0}, \tag{11a}$$

$$\mathbf{0} = \mathbf{C}_\alpha\mathbf{X} + \mathbf{D}_\alpha\mathbf{V}, \quad \mathbf{0} \preceq \mathbf{V}_\beta \perp \mathbf{C}_\beta\mathbf{X} + \mathbf{D}_\beta\mathbf{V} \succeq \mathbf{0}, \quad \mathbf{0} = \mathbf{V}_\gamma, \tag{11b}$$

where $\mathbf{P}(t) = \mathbf{e}_{(i)}^{\mathrm{T}} \otimes \mathbf{I}_{n_p}$ for $t \in (\tau_{i-1}, \tau_i]$, $\mathbf{e}_{(i)}$ denotes the ith unit coordinate vector in $\mathbb{R}^{n_p}$ and $\mathbf{X}_0 = \mathbf{0}$ and $\mathbf{V}_0 \in \mathbb{R}^{n_v \times n_p}$ uniquely solves (11b) at $t = 0$.

Example 1 Consider the quadratic optimal control problem in (6) with $n_x = 2, n_v = 1$, $n_p = n_u = 1$, LCS matrices $\mathbf{A} = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$, $\mathbf{B} = \left[\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right]$, $\mathbf{C} = [-1 \quad 1]$, $\mathbf{D} = 1$, $\mathbf{E} = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, $\mathbf{F} = \mathbf{0}$, and objective function matrices $\mathbf{Q} = \mathbf{I}$, $\mathbf{R} = \mathbf{0}$, $\mathbf{N} = \mathbf{I}$. Suppose that $u(t)$ is parametrically discretized according to (10) where the time horizon $[0, t_f] = [0, 1]$ is split into $n_s = 50$ subintervals for the piecewise constant controls. The optimal control can then be constructed numerically by a sequential approach, using the theory outlined above; the optimal control and corresponding optimal trajectories are given in Fig. 1, using initial control parameters $\mathbf{p}_0 = (p_{(1)}, \ldots, p_{(n_s)}) = (-0.5, \ldots, -0.5)$, i.e., $u(t, \mathbf{p}_0) = -0.5$ for each $t \in (\tau_{i-1}, \tau_i]$, and initial conditions $\mathbf{x}(0) = \mathbf{x}_0 = (1, 3)$, $v(0) = v_0 = 0$, $\mathbf{X}(0) = [0 \quad 0]$, $\mathbf{V}(0) = 0$. Observe that $v(t) = 0$ for $t \in [0, t^*]$, and then $v(t) > 0$ for $t > t^* \approx 0.35$, i.e., $\gamma([0, t^*)) = \{1\}$, $\beta(t^*) = \{1\}$, $\alpha((t^*, t_f]) = \{1\}$, as, at $t = t^*$, the constraint sets switch from inactive through weakly active to active, and the system shifts accordingly.

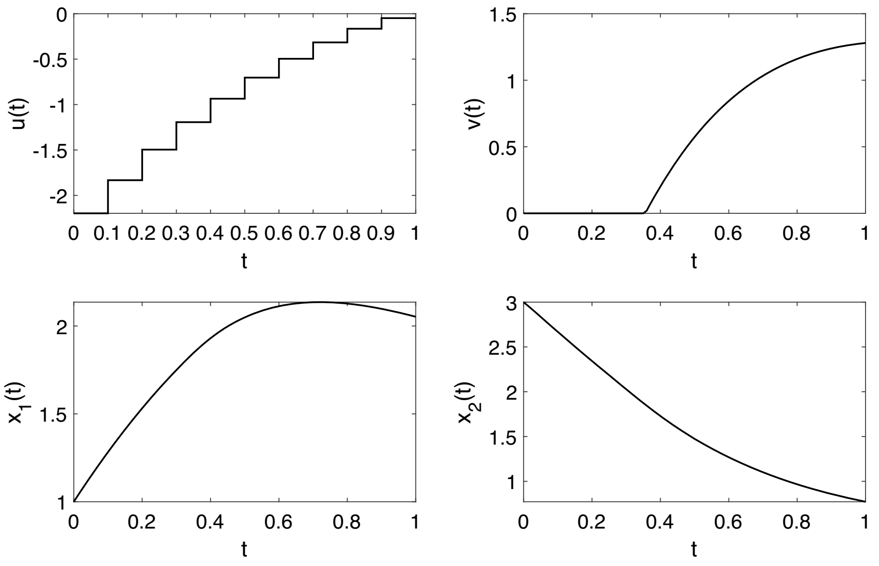

Fig. 1 Simulations of optimal control and state trajectories in Example 1, using the direct method

5 Theory of Lexicographic Linear Complementarity Systems

The NCSs and LCSs studied above are dynamical systems coupled with complementarity problems. Given $\mathbf{q} \in \mathbb{R}^{n_q}$ and $\mathbf{F} : \mathbb{R}^{n_q} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_v}$, consider the nonlinear complementarity problem (NCP) [7] of finding $\mathbf{v} \in \mathbb{R}^{n_v}$ such that $\mathbf{0} \leq \mathbf{v} \perp \mathbf{F}(\mathbf{q}, \mathbf{v}) \geq \mathbf{0}$. Linear complementarity problems (LCPs) [6] are NCPs with affine participating functions, often arising from linearizing NCPs, and can be formulated as follows: given $\mathbf{q} \in \mathbb{R}^{n_v}$ and $\mathbf{D} \in \mathbb{R}^{n_v \times n_v}$, find $\mathbf{v} \in \mathbb{R}^{n_v}$ such that

$$\mathbf{0} \leq \mathbf{v} \perp \mathbf{q} + \mathbf{D}\mathbf{v} \geq \mathbf{0}. \tag{12}$$

This LCP is uniquely solved for any $\mathbf{q} \in \mathbb{R}^{n_v}$ if and only if $\mathbf{D}$ is a P-matrix [6]. As mentioned earlier, a matrix $\mathbf{D}$ is a P-matrix if all its principal minors are positive, a condition which implies nonsingularity of $\mathbf{D}$ and which is implied if $\mathbf{D}$ is positive definite [6].

Mirroring the connection between LCSs and LCPs (i.e., LCSs being ODEs with LCPs embedded), we note that the lexicographic sensitivity systems associated with LCSs contain lexicographic complementarity conditions, and hence are dynamical systems with "lexicographic complementarity problems" embedded. Motivated by this, we formally introduce lexicographic complementarity problems and systems. The lexicographic linear complementarity problem (LexLCP) is stated as follows: given $k \in \mathbb{N}$, $\mathbf{Q} \in \mathbb{R}^{n_v \times k}$ and $\mathbf{D} \in \mathbb{R}^{n_v \times n_v}$, find $\mathbf{V} \in \mathbb{R}^{n_v \times k}$ such that

$$\mathbf{0} \preceq \mathbf{V} \perp \mathbf{Q} + \mathbf{D}\mathbf{V} \succeq \mathbf{0}. \tag{13}$$

Solution properties of the LexLCP are captured in the next result.

Proposition 1 *If* $\mathbf{D}$ *is a P-matrix, then the LexLCP in* (13) *admits a unique solution for any* $\mathbf{Q} \in \mathbb{R}^{n_v \times k}$.

Proof Let $\mathbf{D}$ be a P-matrix. Choose $\mathbf{q}_0 = \mathbf{0} \in \mathbb{R}^{n_v}$. Then the LCP (12) has the unique solution $\mathbf{v}_0 = \mathbf{0} \in \mathbb{R}^{n_v}$. Thus, the piecewise affine (and thus PC^1) function $\mathbf{g} : (\mathbf{q}, \mathbf{v}) \mapsto \mathbf{min}(\mathbf{v}, \mathbf{q} + \mathbf{D}\mathbf{v})$ satisfies $\mathbf{g}(\mathbf{q}_0, \mathbf{v}_0) = \mathbf{0}$ and is completely coherently oriented (in the sense of [15]) with respect to $\mathbf{v}$ by the arguments made in the proof of Theorem 1 above. Hence, [25, Theorem 3.4] is applicable; there exists a PC^1 implicit function $\mathbf{r} : N \to \mathbb{R}^{n_v}$, for some neighborhood N of $\mathbf{q}_0$, such that for any $\mathbf{q} \in N$, $(\mathbf{q}, \mathbf{r}(\mathbf{q}))$ is the unique solution of $\mathbf{g}(\mathbf{q}, \mathbf{r}(\mathbf{q})) = \mathbf{min}(\mathbf{q}, \mathbf{r}(\mathbf{q})) = \mathbf{0}$. Moreover, the LD-derivative $\mathbf{r}'(\mathbf{q}_0; \mathbf{Q}) = \mathbf{V}_0$ is the unique solution of $\mathbf{g}'(\mathbf{q}_0, \mathbf{v}_0; (\mathbf{Q}, \mathbf{V}_0)) = \mathbf{0}$. By [25, Lemma 4.2], $\mathbf{g}'(\mathbf{q}_0, \mathbf{v}_0; (\mathbf{Q}, \mathbf{V}_0)) = \mathbf{slmmin}([\mathbf{v}_0\ \mathbf{V}_0], [\mathbf{q}_0 + \mathbf{D}\mathbf{v}_0\ \mathbf{Q} + \mathbf{D}\mathbf{V}_0])$ where **slmmin** is the shifted-lexicographic-matrix minimum function [25]. Then, $\mathbf{g}'(\mathbf{q}_0, \mathbf{v}_0; (\mathbf{Q}, \mathbf{V}_0)) = \mathbf{0}$ is equivalent to $\mathbf{lmmin}(\mathbf{V}_0, \mathbf{Q} + \mathbf{D}\mathbf{V}_0) = \mathbf{0}$, where **lmmin** is the lexicographic-matrix minimum [25], because $\mathbf{v}_0 = \mathbf{0} = \mathbf{q}_0 + \mathbf{D}\mathbf{v}_0$. The result then holds by [22, Proposition A.1]. □

We can establish existence and uniqueness of solutions of lexicographic linear complementarity systems (LexLCSs), introduced as follows:

$$\frac{d\mathbf{X}}{dt}(t) = \mathbf{A}\mathbf{X}(t) + \mathbf{B}\mathbf{V}(t), \qquad \mathbf{X}(0) = \mathbf{X}_0, \tag{14a}$$

$$\mathbf{0} \preceq \mathbf{V}(t) \perp \mathbf{C}\mathbf{X}(t) + \mathbf{D}\mathbf{V}(t) \succeq \mathbf{0}, \tag{14b}$$

for some $k \in \mathbb{N}$ and $\mathbf{X}_0 \in \mathbb{R}^{n_x \times k}$, and matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ with dimensions outlined earlier. A mapping $\widetilde{\mathbf{Z}} = (\widetilde{\mathbf{X}}, \widetilde{\mathbf{V}})$ is a solution of (14) on $T \subseteq \mathbb{R}$ if (i) $\widetilde{\mathbf{X}}$ is absolutely continuous on T and $\widetilde{\mathbf{V}}$ is Lebesgue integrable on T; and (ii) $(\widetilde{\mathbf{X}}, \widetilde{\mathbf{V}})$ satisfies (14a)–(14b) for almost every $t \in T$ and satisfies the ICs in (14a) at $t = 0$.

Theorem 3 *If* $\mathbf{D}$ *is a P-matrix, then* (14) *admits a unique solution* $(\widetilde{\mathbf{X}}, \widetilde{\mathbf{V}}) : \mathbb{R} \to \mathbb{R}^{(n_x+n_v)\times k}$ *on* $T = (-\infty, +\infty)$ *through initial data* $(\mathbf{X}_0, \mathbf{V}_0)$*, where* $\mathbf{V}_0$ *is the unique solution of* $\mathbf{0} \preceq \mathbf{V}_0 \perp \mathbf{C}\mathbf{X}_0 + \mathbf{D}\mathbf{V}_0 \succeq \mathbf{0}$.

Proof Consider (1) with $\mathbf{f}_0(\mathbf{p}) = \mathbf{p}$ and $\mathbf{p}_0 = \mathbf{0}$. Then, since $\mathbf{D}$ is a P-matrix, $\mathbf{v}_0 = \mathbf{0}$ is the unique solution of the embedded LCP $\mathbf{0} \leq \mathbf{v}_0 \perp \mathbf{C}\mathbf{p}_0 + \mathbf{D}\mathbf{v}_0 \geq \mathbf{0}$ and $\widetilde{\mathbf{z}}(t) = (\widetilde{\mathbf{x}}(t), \widetilde{\mathbf{v}}(t)) = (\mathbf{0}, \mathbf{0})$ is the corresponding unique solution of (1) on $(-\infty, +\infty)$ through $(\mathbf{p}_0, \mathbf{x}_0, \mathbf{v}_0) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$. In this case, $\alpha(t) = \gamma(t) = \emptyset$ and $\beta(t) = \{1, \ldots, n_v\}$ for all t. Choose any $[a, b] \subset \mathbb{R}$ with $a < 0 < b$, then according to [21, Theorem 4.5], an additional conclusion in Theorem 1 above is that $\widetilde{\mathbf{z}}(t) = \widetilde{\mathbf{z}}(t, \mathbf{p})$ is lexicographically smooth (in the sense of [12]) with respect to $\mathbf{p}$ near $\mathbf{p}_0$, for each fixed $t \in T_1 = [a, b]$, and, given the directions matrix $\mathbf{P} = \mathbf{X}_0 \in \mathbb{R}^{n_x \times k}$, Eq. (4) admits unique solution $\widetilde{\mathbf{Z}} = (\widetilde{\mathbf{X}}, \widetilde{\mathbf{V}})$ on $T_1 = [a, b]$, where the mappings $\widetilde{\mathbf{X}}$ and

$\widetilde{\mathbf{V}}$ are absolutely continuous and Lebesgue integrable, respectively. Moreover, with $\frac{\partial \mathbf{A}}{\partial \mathbf{p}} = \mathbf{0}$, $\frac{\partial \mathbf{C}}{\partial \mathbf{p}} = \mathbf{0}$, etc., and $\alpha(t) = \gamma(t) = \emptyset$ for all t, Eq. (4) simplifies to Eq. (14), with (14a) satisfied for almost every $t \in T$ and (14b) satisfied for every $t \in T$.

It is straightforward to construct a continuation of $\widetilde{\mathbf{Z}}$ to $T_2 = [-2a, 2b]$ by repeating the arguments above. Hence, the augmented graphs of continuations, $\Gamma_{\text{ext}} := \{\{(t, \widehat{\mathbf{Z}}(t)) : t \in \widehat{T}\} : \widehat{\mathbf{Z}}$ is a continuation of $\widetilde{\mathbf{Z}}$ on $\widehat{T}\}$ is nonempty. Then solutions can be continued to $\mathbf{Z}_{\max}$ on a maximal interval of existence $T_{\max} = (t_L, t_U)$, where $t_L \in \mathbb{R} \cup \{-\infty\}$ and $t_U \in \mathbb{R} \cup \{+\infty\}$, by arguments involving Zorn's Lemma (see the proof of [24, Theorem 6.4]). In fact, it must hold that $T_{\max} = (-\infty, +\infty)$ here, since if not, i.e., $T_{\max} = (t_L, t_U)$ with t_L or $t_U \in \mathbb{R}$, there exists a continuation to $T_k = [-ka, kb]$ where $k \in \mathbb{N}$ is chosen so that $-ka < t_L$ or $kb > t_U$, a contradiction to $\mathbf{Z}_{\max}$ being the maximal continuation. □

6 Future Work

One future direction for this work would be comparing the methods here with those put forth by Vieira et al. in [26, 27] (which can accommodate "high-index" LCS systems, since $\mathbf{D}$ need not be a P-matrix in their work—e.g. see [27, Example 2]).

Acknowledgements This material is based upon work supported by the National Science Foundation under Award No. 2318772 and 2318773.

References

1. Barton, P.I., Khan, K.A., Stechlinski, P., Watson, H.A.J.: Computationally relevant generalized derivatives: theory, evaluation and applications. Optim. Methods Softw. **33**, 1030–1072 (2018)
2. Camlibel, M.K.: On the controllability and stabilizability of linear complementarity systems. IFAC-PapersOnLine **52**(17), 1–6 (2019)
3. Camlibel, M.K., Heemels, W.P.M.H., Schumacher, J.M.: On linear passive complementarity systems. Eur. J. Control **8**, 220–237 (2002)
4. Camlibel, M.K., Pang, J.S., Shen, J.: Lyapunov stability of complementarity and extended systems. SIAM J. Optim. **17**(4), 1056–1101 (2007)
5. Clarke, F.H.: Optimization and Nonsmooth Analysis. SIAM, Philadelphia (1990)
6. Cottle, R.W., Pang, J.S., Stone, R.E.: The Linear Complementarity Problem. SIAM, Philadelphia (2009)
7. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I. Springer, New York (2003)
8. Heemels, W.M.H., Camlibel, M.K., Schumacher, J.M., Brogliato, B.: Observer-based control of linear complementarity systems. Int. J. Robust Nonlinear Control **21**(10), 1193–1218 (2011)
9. Heemels, W.P.M.H., Schumacher, J.M., Weiland, S.: Linear complementarity systems. SIAM J. Appl. Math. **60**(4), 1234–1269 (2000)
10. Kelly, M.: An introduction to trajectory optimization: how to do your own direct collocation. SIAM Rev. **59**, 849–904 (2017)

11. Khan, K.A., Barton, P.I.: A vector forward mode of automatic differentiation for generalized derivative evaluation. Optim. Methods Softw. **30**(6), 1185–1212 (2015)
12. Nesterov, Y.: Lexicographic differentiation of nonsmooth functions. Math. Program. **104**, 669–700 (2005)
13. Özyurt, D.B., Barton, P.I.: Cheap second order directional derivatives of stiff ODE embedded functionals. SIAM J. Sci. Comput. **26**(5), 1725–1743 (2005)
14. Pang, J.S., Shen, J.: Strongly regular differential variational systems. IEEE Trans. Autom. Control **52**(2), 242–255 (2007)
15. Ralph, D., Scholtes, S.: Sensitivity analysis of composite piecewise smooth equations. Math. Program. **76**, 593–612 (1997)
16. van der Schaft, A.J., Schumacher, J.M.: The complementary-slackness class of hybrid systems. Math. Control Signals Syst. **9**(3), 266–301 (1996)
17. van der Schaft, A.J., Schumacher, J.M.: Complementarity modeling of hybrid systems. IEEE Trans. Autom. Control **43**(4), 483–490 (1998)
18. Schumacher, J.M.: Complementarity systems in optimization. Math. Program. **101**, 263–295 (2004)
19. Shen, J.: Robust non-zenoness of piecewise affine systems with applications to linear complementarity systems. SIAM J. Optim. **24**(4), 2023–2056 (2014)
20. Shen, J., Pang, J.S.: Linear complementarity systems: Zeno states. SIAM J. Control Optim. **44**(3), 1040–1066 (2005)
21. Stechlinski, P.: Theory of index-one nonlinear complementarity systems. J. Differ. Equ. **285**, 99–127 (2021)
22. Stechlinski, P.: Dynamic optimization of complementarity systems. IEEE Trans. Autom. Control **68**(2), 1122–1129 (2022)
23. Stechlinski, P., Barton, P.I.: Generalized derivatives of optimal control problems with nonsmooth differential-algebraic equations embedded. In: 55th IEEE Conference on Decision and Control, pp. 592–597 (2016)
24. Stechlinski, P., Barton, P.I.: Dependence of solutions of nonsmooth differential-algebraic equations on parameters. J. Differ. Equ. **262**(3), 2254–2285 (2017)
25. Stechlinski, P., Khan, K.A., Barton, P.I.: Generalized sensitivity analysis of nonlinear programs. SIAM J. Optim. **28**(1), 272–301 (2018)
26. Vieira, A., Brogliato, B., Prieur, C.: Optimality conditions for the minimal time problem for complementarity systems. IFAC-PapersOnLine **52**(16), 239–244 (2019)
27. Vieira, A., Brogliato, B., Prieur, C.: Quadratic optimal control of linear complementarity systems: first order necessary conditions and numerical analysis. IEEE Trans. Autom. Control **65**(6), 2743–2750 (2020)