OPTIMIZATION OF A PERTURBED SWEEPING PROCESS BY CONSTRAINED DISCONTINUOUS CONTROLS

GIOVANNI COLOMBO¹ BORIS S. MORDUKHOVICH² DAO NGUYEN³

ABSTRACT. This paper deals with optimal control problems described by a controlled version of Moreau's sweeping process governed by convex polyhedra, where measurable control actions enter additive perturbations. This class of problems, which addresses unbounded discontinuous differential inclusions with intrinsic state constraints, is truly challenging and underinvestigated in control theory while being highly important for various applications. To attack such problems with constrained measurable controls, we develop a refined method of discrete approximations with establishing its well-posedness and strong convergence. This approach, married to advanced tools of first-order and second-order variational analysis and generalized differentiations, allows us to derive adequate collections of necessary optimality conditions for local minimizers, first in discrete-time problems and then in the original continuous-time controlled sweeping process by passing to the limit. The new results include an appropriate maximum condition and significantly extend the previous ones obtained under essentially more restrictive assumptions. We compare them with other versions of the maximum principle for controlled sweeping processes that have been recently established for global minimizers in problems with smooth sweeping sets by using different techniques. The obtained necessary optimality conditions are illustrated by several examples.

Key words. Optimal control, sweeping process, variational analysis, discrete approximations, generalized differentiation, necessary optimality conditions.

AMS Subject Classifications. 49M25, 49J53, 90C30.

1 Introduction and Problem Formulation

This paper addresses the following optimal control problem labeled as (P):

Minimize the Mayer-type cost functional

$$J[x,u] := \varphi(x(T)) \tag{1.1}$$

over the corresponding (described below) pairs $(x(\cdot), u(\cdot))$ satisfying

$$\begin{cases} \dot{x}(t) \in -N(x(t); C) + g(x(t), u(t)) \text{ a.e. } t \in [0, T], \ x(0) = x_0 \in C \subset \mathbb{R}^n, \\ u(t) \in U \subset \mathbb{R}^d \text{ a.e. } t \in [0, T], \end{cases}$$
(1.2)

where the set C is a convex *polyhedron* given by

$$C := \bigcap_{j=1}^{s} C^{j} \text{ with } C^{j} := \left\{ x \in \mathbb{R}^{n} \middle| \langle x_{*}^{j}, x \rangle \le c_{j} \right\},$$
(1.3)

and where N(x; C) stands for the normal cone of convex analysis defined by

$$N(x;C) := \left\{ v \in \mathbb{R}^n \mid \langle v, y - x \rangle \le 0, \ y \in C \right\} \text{ if } x \in C \text{ and } N(x;C) := \emptyset \text{ if } x \notin C.$$
(1.4)

Observe that due to the second part of definition (1.4) mandatory yields the presence of the hidden *pointwise state constraints* on the trajectories of (1.2):

$$x(t) \in C$$
, i.e. $\langle x_*^j, x(t) \rangle \le c_j$ for all $t \in [0, T]$ and $j = 1, \dots, s$. (1.5)

¹Dipartimento di Matematica "Tullio Levi-Civita", Università di Padova, via Trieste 63, 35121 Padova, Italy (colombo@math.unipd.it) and G.N.A.M.P.A. of INdAM.

²Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA (boris@math.wayne.edu). Research of this author was partly supported by the USA National Science Foundation under grants DMS-1512846 and DMS-1808978, and by the USA Air Force Office of Scientific Research grant #15RT0462.

³Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA (dao.nguyen2@wayne.edu). Research of this author was partly supported by the USA National Science Foundation under grant DMS-1808978 and by the USA Air Force Office of Scientific Research grant #15RT0462.

Considering the differential inclusion in (1.2) without the additive perturbation term g(x, u), we arrive at the framework of the *sweeping process* introduced by Jean-Jacques Moreau who was motivated by applications to problems of elastoplasticity; see [23]. It has been well recognized that the (uncontrolled) Moreau's sweeping process has a *unique* absolutely continuous (or even Lipschitz continuous) solution for convex and mildly nonconvex sets C; see, e.g., [13] and the references therein. Thus there is no room for optimization of the sweeping process unless some additional functions or parameters of choice are inserted into its description. It is very different from control theory for Lipschitzian differential inclusions

$$\dot{x}(t) \in F(x(t))$$
 a.e. $t \in [0, T], x(0) = x_0 \in \mathbb{R}^n$, (1.6)

which have multiple solutions. The latter type of dynamics extends the classical ODE control setting with F(x) := f(x, U) in (1.6), where the choice of measurable controls $u(t) \in U \subset \mathbb{R}^d$ a.e. $t \in [0, T]$ creates the possibility to find an optimal one with respect to a prescribed performance. The main issue here is that the normal cone mapping $N(\cdot; C)$ in the sweeping process is *highly non-Lipschitzian* (even discontinuous) while being *maximal monotone*. On the other hand, the well-developed optimal control theory for differential inclusions (1.6) strongly depends on Lipschitzian behavior of $F(\cdot)$; see, e.g., [21, 27] with the references therein as well as more recent publications.

Introducing controls into the perturbation term of (1.2) allows us to have multiple solutions $x(\cdot)$ of this system by the choice of feasible control functions $u(\cdot)$ and thus to minimize the cost functional (1.1) over feasible control-trajectory pairs. Problems of this type were considered in the literature from the viewpoint of the existence of optimal solutions and relaxation; see [1, 9, 15, 26] among other publications.

More recently, necessary optimality conditions for local minimizers were derived in [6, 7] by the method of discrete approximations for problems of type (P) with smooth (in fact $W^{2,\infty}$) control functions without any constraints. Later on these results were further extended in [8] to nonconvex (and hence nonpolyhedral) problems with prox-regular sets C in the same control setting. Note that both C and g in (1.2) may be time-dependent; we discuss the autonomous case just for simplicity. The discrete approximation approach implemented in [6]–[8] was based on the scheme from [11] developed for the unperturbed sweeping process with controls in the moving set. The later was in turn a sweeping control version of the original discrete approximations method to derive necessary optimality conditions for Lipschitzian differential inclusions (1.6) suggested and implemented in [19]; see also [21].

Quite recently, other approximation procedures were developed to derive necessary optimality conditions for global minimizers of (P) in the class of measurable controls while under rather strong assumptions. The first paper [3] assumes, among other requirements, that the boundary of the sweeping set Cin (1.2) is \mathcal{C}^3 -smooth, the control set U is compact and convex, and its image g(x, U) under g is convex as well. The \mathcal{C}^3 -smoothness assumption on C was relaxed in [14], by employing a smooth approximation procedure not relying on the distance function as in [3], for the case of $C := \{x \in \mathbb{R}^n \mid \psi(x) \leq 0\}$ with ψ being a \mathcal{C}^2 - smooth convex function. The necessary optimality conditions obtained in both papers [3, 14] can be treated as somewhat different counterparts of the celebrated Pontryagin Maximum Principle (PMP) for state-constrained controlled differential equations $\dot{x} = f(x, u)$.

Note that necessary optimality conditions in some other classes of optimal control problems governed by various controlled versions of the sweeping process were developed in [2, 5, 6, 7, 8, 10, 11, 18].

The main goal of this paper is to derive necessary optimality conditions for local minimizers (in the senses specified below) of the formulated problem (P), with the constraint set U in (1.2) given by an *arbitrary compact* and with the (nonsmooth) *polyhedral* set C from (1.3), by significantly reducing regularity assumptions on the reference control. Although problem (1.2) is stated in the class of *measurable* feasible control actions, we assume that the local *optimal* control under consideration is of *bounded variation*, hence allowing to be discontinuous.

Our approach is based on developing the *method of discrete approximations*, which is certainly of its own interest and has never been implemented before in control theory for sweeping processes with discontinuous controls. The novel results in this direction establish a *strong* approximation of *every feasible* control-state pair for (P) in the sense of the L^2 -norm convergence of discretized controls and the $W^{1,2}$ -norm convergence of the corresponding piecewise linear trajectories. Furthermore, we justify such a strong convergence of *optimal* solutions for discrete problems to the given local minimizer of (P).

Dealing further with intrinsically nonsmooth and nonconvex discrete-time approximation problems, we derive for them necessary optimality conditions of the discrete Euler-Lagrange type by using appropriate unconvexified tools of first-order and second-order variational analysis and generalized differentiation. Employing these tools and passing to the limit from discrete approximations lead us to new nondegenerate necessary optimality conditions for local optimal solutions of the sweeping control problem (P). The obtained results significantly extend those recently established in [7] for unconstrained $W^{2,\infty}$ optimal controls in (P), contain a maximum condition, while being essentially different from the necessary optimality conditions derived in [3, 14] for problems of type (P) with smooth sets C in addition to other assumptions. We present nontrivial examples that illustrate the efficiency of the new results. Further applications to some practical models are considered in our subsequent paper [12].

The rest of the paper is organized as follows. In Section 2 we formulate the standing assumptions, discuss the types of local minimizers under consideration, and present some preliminary results.

Section 3 is devoted to the construction of discrete approximations of the controlled constrained sweeping dynamics (1.2) that allows us to deal with measurable controls (in fact of bounded variation) and to strongly approximate any feasible solutions of (P) as mentioned above. This result plays a major role in the justification of the developed version the method of discrete approximations for problem (P).

In Section 4 we construct a sequence of discrete approximation of a given "intermediate" local minimizer for (P) that occupies an intermediate position between weak and strong minimizers in variational and control problems. The major result of this section justifies the strong $W^{1,2} \times L^2$ approximation of the given local minimum pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ by extended optimal solutions to the discretized problems. It makes a bridge between the continuous-time sweeping control problem (P) and its discrete-time counterparts.

It occurs that the discrete-time approximating problems are unavoidably nonsmooth and nonconvex, even when the initial data are differentiable. It is due to the presence of increasingly many geometric constraints generated by the normal cone graph. To deal with them, we need adequate tools of variational analysis involving not only first-order but also second-order generalized differentiation. The latter is because of the normal cone description of the sweeping process. In Section 5 we present the corresponding definitions of the first-order and second- order generalized differential constructions taken from [20] together with the results of their computations entirely in terms of the given data of (1.2).

Section 6 provides the derivation of necessary optimality conditions for discrete-time problems by reducing them to problems of nondifferentiable programming with many geometric constraints, using necessary optimality conditions for them obtained via variational/ extremal principles, and then expressing the latter in terms of the given data of (P) by employing calculus rules of generalized differentiation.

Section 7 is the culmination. We pass to the limit from the necessary optimality conditions for discretetime problems by using stability of discrete approximations, robustness of then generalized differential constructions, and establishing an appropriate convergence of adjoint functions, which is the most difficult part. In this way we arrive at new necessary conditions for local minimizers of (P) expressed in terms of the given data of the original problem. The usefulness of the nondegenerated optimality conditions obtained is illustrated in Section 8 by nontrivial examples.

Throughout the paper we use standard notations of variational analysis and optimal control; see, e.g., [20, 21]. Recall that $\mathbb{N} := \{1, 2, \ldots\}$.

2 Standing Assumptions and Basic Notions

Dealing with the polyhedron C from (1.3) and having $\bar{x} \in C$, consider the set of active constraint indices

$$I(\bar{x}) := \{ j \in \{1, \dots, s\} \mid \langle x_*^j, \bar{x} \rangle = c_j \}.$$
 (2.1)

Recall that the linear independence constraint qualification (LICQ) holds at \bar{x} if

$$\left[\sum_{j\in I(\bar{x})}\alpha_j x_*^j = 0, \ \alpha_j \in \mathbb{R}\right] \Longrightarrow \left[\alpha_j = 0 \text{ for all } j \in I(\bar{x})\right].$$
(2.2)

Our *standing assumptions* in this paper are as follows:

(H1) The control region $U \neq \emptyset$ is a compact set in \mathbb{R}^d (in fact it may be an arbitrary metric compact). (H2) The perturbation mapping $g \colon \mathbb{R}^n \times U \to \mathbb{R}^n$ is continuous in (x, u) while being also Lipschitz continuous with respect to x uniformly on U whenever x belongs to a bounded subset of \mathbb{R}^n and satisfies there the sublinear growth condition

$$||g(x,u)|| \le \beta (1+||x||) \text{ for all } u \in U$$

with some positive constant β .

(H3) The LICQ condition (2.2) holds along the reference trajectory $\bar{x}(t)$ of (1.2) for all $t \in [0, T]$.

It follows from [15, Theorem 1] that for each measurable control $u(\cdot)$ there is a unique solution $x(\cdot) \in W^{1,2}([0,T],\mathbb{R}^n)$ to the Cauchy problem in (1.2). Thus by a *feasible process* for (P) we understand a pair $(x(\cdot), u(\cdot))$ such that $u(\cdot)$ is measurable, $x(\cdot) \in W^{1,2}([0,T],\mathbb{R}^n)$, and all the constraints in (1.2) are satisfied. The above discussion tells us that the set of feasible pairs for (P) is nonempty.

Furthermore, it follows from [15, Theorem 2] that under the assumptions above the sweeping control problem (P) admits an *optimal solution* provided that the image set

$$g(x,U) := \left\{ x \in \mathbb{R}^n \mid x = g(x,u) \text{ for some } u \in U \right\}$$

is convex. Since in this paper we are interested in deriving necessary optimality conditions for a given local minimizer of (P), we do not impose the aforementioned convexity assumption.

Let us now specify what we mean by a local minimizer of (P).

Definition 2.1 We say that a feasible pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ for (P) is a $W^{1,2} \times L^2$ -LOCAL MINIMIZER in this problem if $\bar{x}(\cdot) \in W^{1,2}([0,T]; \mathbb{R}^n)$ and there exists $\epsilon > 0$ such that $J[\bar{x}, \bar{u}] \leq J[x, u]$ for all feasible pairs $(x(\cdot), u(\cdot))$ satisfying the condition

$$\int_0^T \left(\|\dot{x}(t) - \dot{\bar{x}}(t)\|^2 + \|u(t) - \bar{u}(t)\|^2 \right) dt < \epsilon.$$

For the case of differential inclusions of type (1.6) with no explicit controls, this notion corresponds to *intermediate local minimizers* of rank two introduced in [19] and then studied there and in other publications; see, e.g., [21, 27] and the references therein. Quite recently, such minimizers have been revisited in [18] for controlled sweeping processes different from (1.2); namely, for those where continuous control actions enter the moving set C(t) = C(u(t)). It is easy to see that *strong* $\mathcal{C} \times L^2$ - local minimizers of (P) with $\bar{x}(\cdot) \in W^{1,2}([0,T]; \mathbb{R}^n)$ fall into the category of Definition 2.1, but not vice versa.

In the general setting of $W^{1,2} \times L^2$ -local minimizers we need to use a certain relaxation procedure in the line of Bogolyubov and Young that has been well understood in the calculus of variations and optimal control; see, e.g., [16, 15, 21, 26, 27] for more recent publications in the case of differential inclusions. Taking into account the convexity and closedness of the normal cone N(x; C) and the compactness of the set g(x, U), the *relaxed* version (R) of problem (P) consists of minimizing the cost functional (1.1) on absolutely continuous trajectories of the convexified differential inclusion

$$\dot{x}(t) \in -N(x(t); C) + \cos g(x(t), U)$$
 a.e. $t \in [0, T], x(0) = x_0 \in C \subset \mathbb{R}^n$, (2.3)

where 'co' signifies the convex hull of the set. Then we come up with the following notion.

Definition 2.2 Let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be a feasible pair for (P). We say that it is a RELAXED $W^{1,2} \times L^2$ -LOCAL MINIMIZER for (P) if $\bar{x}(\cdot) \in W^{1,2}([0,T]; \mathbb{R}^n)$ and there is $\epsilon > 0$ such that

$$\varphi(\bar{x}(T)) \leq \varphi(x(T)) \quad whenever \quad \int_0^T \left(\left\| \dot{x}(t) - \dot{\bar{x}}(t) \right\|^2 + \left\| u(t) - \bar{u}(t) \right\|^2 \right) dt < \epsilon,$$

where $u(\cdot)$ is a measurable control with $u(t) \in \operatorname{co} U(t)$ a.e. on [0,T], and where $x(\cdot)$ is a trajectory of the convexified inclusion (2.3) that can be strongly approximated in $W^{1,2}([0,T];\mathbb{R}^n)$ by feasible trajectories to (P) generated by piecewise constant controls $u_m(\cdot)$ on [0,T] with

$$\int_0^T \|u_m(t) - u(t)\|^2 dt \to 0 \quad as \quad m \to \infty.$$

Since step functions are dense in the space $L^2([0,T]; \mathbb{R}^d)$, we obviously have that there is no difference between $W^{1,2} \times L^2$ -local minimizers for (P) and their relaxed counterparts provided that the sets g(x, U)and U are *convex*, which is not assumed in what follows. Moreover, it is possible to deduce from the proofs of [15, Theorem 2] and [26, Theorem 4.2] that any strong local minimizer for (P) is automatically a relaxed one under the assumptions made, but we are not going to pursue this issue here.

Consider further a set-valued mapping $F: C \times U \rightrightarrows \mathbb{R}^n$ defined by

$$F(x,u) := N(x;C) - g(x,u) \text{ for all } x \in C, \ u \in U$$

$$(2.4)$$

and deduce from the Motzkin's theorem of the alternative the representation

$$F(x,u) := \left\{ \sum_{j \in I(x)} \lambda^j x^j_* \mid \lambda^j \ge 0 \right\} - g(x,u), \quad x \in C, \ u \in U,$$

$$(2.5)$$

3 Discrete Approximations of Feasible Solutions

In this section we start developing the *method of discrete approximations* to study the sweeping control problem (P) under our standing assumptions. For simplicity, consider the standard Euler explicit scheme for the replacement of the time derivative in (1.2) by

$$\dot{x}(t) \approx \frac{x(t+h) - x(t)}{h}$$
 as $h \downarrow 0$

which we formalize as follows. For any $m \in \mathbb{N}$ denote by

$$\Delta_m := \left\{ 0 = t_m^0 < t_m^1 < \ldots < t_m^{2^m} = T \right\} \text{ with } h_m := t_m^{i+1} - t_m^i$$

the discrete mesh on [0, T] and define the sequence of discrete-time systems

$$x_m^{i+1} \in x_m^i - h_m F(x_m^i, u_m^i), \ i = 0, \dots, 2^m - 1,$$
(3.1)

where we have $u_m^i \in U$ due to the definition of F in (2.4). Let $I_m^i := [t_m^{i-1}, t_m^i)$.

The next result provides a constructive approximation of any feasible process for (P) by feasible solutions to (3.1) that are appropriately extended to the continuous-time interval [0, T]. This result plays a major role in the entire subsequent procedure to derive necessary optimality conditions for (P) while certainly being of its independent interest. Recall that a *representative* of a given measurable function on [0, T] is a function that agrees with the given one for a.e. $t \in [0, T]$.

Theorem 3.1 Let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be a feasible pair for problem (P) such that $\bar{x}(\cdot) \in W^{1,2}([0,T];\mathbb{R}^n)$ and that $\bar{u}(\cdot)$ is of bounded variation (BV) while admitting a right continuous representative on [0,T], which we keep denoting by $\bar{u}(\cdot)$. In addition to (H1)–(H3), suppose that the mapping g(x, u) is locally Lipschitzian

in both variables around $(\bar{x}(t), \bar{u}(t))$ for all $t \in [0, T]$. Then for each $i = 1, \ldots, 2^m$ there exist sequences of unit vectors $z_m^{ji} \in \mathbb{R}^n$, real numbers c_m^{ji} , and state-control pairs $(x_m(t), u_m(t)), 0 \le t \le T$, such that

$$z_m^{ji} \to x_*^j \text{ and } c_m^{ji} \to c_j \text{ as } m \to \infty$$
 (3.2)

and the following properties are fulfilled:

(a) The sequence of control mappings $u_m: [0,T] \to U$, which are constant on each interval I_m^i , converges to $\bar{u}(\cdot)$ strongly in $L^2([0,T]; \mathbb{R}^d)$ and pointwise on [0,T].

(b) The sequence of continuous state mappings $x_m: [0,T] \to \mathbb{R}^n$, which are affine on each interval I_m^i , converges strongly in $W^{1,2}([0,T];\mathbb{R}^n)$ to $\bar{x}(\cdot)$ while satisfying the inclusions

$$x_m(t_m^i) = \bar{x}(t_m^i) \in C_m^i \text{ for each } i = 1, \dots, 2^m \text{ with } x_m(0) = x_0,$$
(3.3)

where the perturbed polyhedra C_m^i are given by

$$C_m^i := \bigcap_{j=1}^s \left\{ x \in \mathbb{R}^n \mid \langle z_m^{ji}, x \rangle \le c_m^{ji} \right\} \text{ for } i = 1, \dots, 2^m \text{ with } C_m^0 := C.$$
(3.4)

(c) For all $t \in (t_m^{i-1}, t_m^i)$ and $i = 1, ..., 2^m$ we have the differential inclusions

$$\dot{x}_m(t) \in -N(x_m(t_m^i); C_m^i) + g(x_m(t_m^i), u_m(t)).$$
(3.5)

As a part of the proof of Theorem 3.1, we establish the following lemma, which is of its own interest. Lemma 3.2 Given a feasible solution $(\bar{x}(\cdot), \bar{u}(\cdot))$ to (P) under the assumptions of Theorem 3.1, we have:

(i) $\bar{x}(\cdot)$ is Lipschitz continuous on [0,T] and right differentiable for every $t \in [0,T]$, and its right derivative denoted by $\dot{\bar{x}}(\cdot)$ is also right continuous on [0,T].

(ii) The sweeping differential inclusion

$$\dot{\bar{x}}(t) \in -N(\bar{x}(t); C) + g(\bar{x}(t), \bar{u}(t))$$

with $\dot{x}(t)$ taken from (i) and the right continuous representative of $\bar{u}(t)$, is satisfied for each $t \in [0,T]$.

Proof. Considering the differential inclusion

$$\dot{x}(t) \in -N(x(t); C) + g(x(t), \bar{u}(t)), \ x(0) = x_0 \in C,$$

we deduce from, e.g., [4, Propositions 3.8 and 3.12] that there exists one and only one Lipschitz continuous solutions on [0, T], which therefore agrees with the given trajectory $\bar{x}(\cdot)$. Furthermore, $\bar{x}(\cdot)$ is a unique solution of the differential inclusion

$$\dot{x}(t) \in -N(x(t); C) + g(\bar{x}(t), \bar{u}(t))$$
 a.e. $t \in [0, T], x(0) = x_0 \in C,$

where the perturbation term depends only on t. Then the assumptions imposed on $\bar{u}(\cdot)$ and g ensure that the mapping $t \mapsto g(\bar{x}(t), \bar{u}(t))$ is BV on [0, T]. The result of [4, Proposition 3.3] tells us that $\bar{x}(\cdot)$ is right differentiable at each $t \in [0, T)$ and satisfies the equalities

$$\dot{\bar{x}}(t) = g(\bar{x}(t), \bar{u}(t)) - \operatorname{proj}_{N(\bar{x}(t);C)} \left(g(\bar{x}(t), \bar{u}(t)) \right) = \operatorname{proj}_{T(\bar{x}(t);C)} \left(g(\bar{x}(t), \bar{u}(t)) \right) \text{ a.e. } t \in [0,T]$$
(3.6)

written via the (unique) projection onto the convex set $N(\bar{x}(t); C)$, where the second one can be easily verified. Our goal is to show that $\dot{\bar{x}}(t)$ is right continuous on [0, T] while satisfying (3.6) for each $t \in [0, T]$.

Observe preliminary that, thanks to LICQ, the polyhedron C has nonempty interior and so the normal cone N(x; C) is pointed at each $x \in \partial C$. Denote $v(t) := g(\bar{x}(t), \bar{u}(t))$, fix $0 \le \bar{t} < T$, and let $t_k \to \bar{t}^+$ as $k \to \infty$. We need to verify that $v(t_k) \to v(\bar{t})$, which is equivalent by (3.6) to

$$\operatorname{proj}_{N(\bar{x}(\bar{t});C)}(v(\bar{t})) = \lim_{k \to \infty} \operatorname{proj}_{N(\bar{x}(t_k);C)}(v(t_k)).$$
(3.7)

Note that there is nothing to prove if $\bar{x}(\bar{t}) \in \text{int } C$, since $N(\bar{x}(t_k); C) = \{0\}$ for all k sufficiently large. To proceed further, assume that $\bar{x}(\bar{t}) \in \text{bd } C$ and observe easily that $N(\bar{x}(t_k); C) \subset N(\bar{x}(\bar{t}); C)$ for all large k. Consider now the following three possible cases:

(1) If $v(\bar{t}) \in T(\bar{x}(\bar{t}); C)$, then $\langle v(\bar{t}), x_*^j \rangle \leq 0$ for all $j \in I(\bar{x}(\bar{t}))$, and so $\lim_{k\to\infty} \langle v(t_k), x_*^j \rangle \leq 0$ for all $j \in I(\bar{x}(\bar{t}))$. Since $N(\bar{x}(t_k); C) \subset N(\bar{x}(\bar{t})); C)$, we get (3.7).

(2) If $v(\bar{t}) \in N(\bar{x}(\bar{t}); C)$, then arguing similarly to (1) and using the second equality in (3.6) show that $\operatorname{proj}_{T(\bar{x}(t_k);C)}(v(t_k)) \to 0 = \operatorname{proj}_{T(\bar{x}(\bar{t});C)}(v(\bar{t}))$ as $k \to \infty$ and hence verifies (3.7) directly.

(3) Let now $v(\bar{t}) \notin T(\bar{x}(\bar{t}); C) \cup N(\bar{x}(\bar{t}); C)$. Then LICQ ensures the unique representation

$$\operatorname{proj}_{N(\bar{x}(\bar{t});C)}(v(\bar{t})) = \sum_{j \in I(\bar{x}(\bar{t}))} \lambda_j(\bar{t}) x_*^j,$$
(3.8)

where $\lambda_j(\bar{t}) \ge 0$ for all $j \in I(\bar{x}(\bar{t}))$. If in this case $\lambda_j(\bar{t}) = 0$ for some $j \in I(\bar{x}(\bar{t}))$, then

$$\left\langle \operatorname{proj}_{T(\bar{x}(\bar{t});C)}(v(\bar{t})), x_*^j \right\rangle < 0$$

due to the LICQ assumption. This implies that $\langle \bar{x}(t_k), x_*^j \rangle \leq \langle \bar{x}(\bar{t}), x_*^j \rangle$ for all k sufficiently large. Consequently, the corresponding vector x_*^j appears being multiplied by zero in the representation

$$\operatorname{proj}_{N(\bar{x}(t_k);C)}(v(t_k)) = \sum_{j \in I(\bar{x}(t_k))} \lambda_j(t_k) x_*^j.$$
(3.9)

Recalling that $I(\bar{x}(t_k)) \subset I(\bar{x}(\bar{t}))$ for all large k, it turns out that the set of active indices in (3.9) is the same as in (3.8). Finally, it follows from [24, Theorems 2.1 and 4.1] under the imposed LICQ that the coefficients $\lambda_j(\cdot)$ are continuous with respect to $v(\cdot)$. This verifies (3.7) in case (3).

Now we are ready to proceed with the proof of the major Theorem 3.1.

Proof of Theorem 3.1. Fix $m \in \mathbb{N}$ and for all $t \in [t_m^i, t_m^{i+1})$ and $i = 0, \ldots, 2^m - 1$ define

$$u_m(t) := \bar{u}(t_m^{i+1}), \quad x_m(t) := \bar{x}(t_m^i) + (t - t_m^i) \frac{\bar{x}(t_m^{i+1}) - \bar{x}(t_m^i)}{h_m}.$$

Then denote $\omega_m(t) := \dot{\bar{x}}_m(t)$ for which we have the representation

$$\omega_m(t) = \omega_m^i := \frac{\bar{x}(t_m^{i+1}) - \bar{x}(t_m^i)}{h_m} \text{ whenever } t \in [t_m^i, t_m^{i+1}), \ i = 0, \dots, 2^m - 1$$

It follows from the right continuity of \bar{u} that $u_m(t) \to \bar{u}(t)$ as $m \to \infty$ for all $t \in [0, T)$. Hence we get that $u_m(\cdot) \to \bar{u}(\cdot)$ strongly in $L^2(0, T)$ by the dominated convergence theorem, which verifies (a). To prove (b) and (c), let \bar{t} be a nodal point of the *m*-th mesh that by construction remains a nodal point for all m'-mesh with $m' \ge m$. Denote by $i_m(\bar{t})$ the index i such that $\bar{t} = i \frac{T}{2^m}$ and observe by Lemma 3.2 that

$$\lim_{m \to \infty} \omega_m^{i_m(\bar{t})} = \dot{\bar{x}}(\bar{t}) \text{ and } \lim_{m \to \infty} \|\omega_m - \dot{\bar{x}}\|_{L^2(0,T)} = 0.$$
(3.10)

Indeed, the first equality in (3.10) is a consequence of the construction of $\omega_m(t)$ and the right continuity of the derivative $\dot{\bar{x}}(t)$ on [0, T] by Lemma 3.2(i). This in turn yields the second equality therein by basic real analysis and thus justifies the strong $W^{1,2}$ convergence of $x_m(\cdot)$ to $\bar{x}(\cdot)$.

To proceed further, for any $t \in [0,T)$ denote $\tau_m(t) := \min\{t_m^i \mid t \leq t_m^i\}$ and get by assertion (a) that

$$\dot{x}_m(t) - g(x_m(\tau_m(t)), u_m(t)) = \omega_m(\tau_m(t)) - g(x_m(\tau_m(t)), u_m(t)) \to \dot{\bar{x}}(t) - g(\bar{x}(t), \bar{u}(t)) \in -N(\bar{x}(t); C)$$

as $m \to \infty$. Thus there are unit vectors $z_m^{ji} \in \mathbb{R}^n$ and constants $c_m^{ji} \in \mathbb{R}$ as $j = 1, \ldots, k$ and $i = 1, \ldots, 2^m$ satisfying (3.2) and the relationships in (3.3), (3.5) with the sets C_m^i defined in (3.4).

4 Discrete Approximations of Local Optimal Solutions

As seen above, Theorem 3.1 provides a constructive discrete approximation of any feasible solution to problem (P) by feasible solutions to discrete-time problems, with no connections to optimization. The main goal here is to study a given local optimal solution to (P) by using discrete approximations as a vehicle to derive further necessary optimality conditions for it. To proceed in this direction, we construct a sequence of discrete-time optimization problems such that their optimal solutions always exist and strongly converge in the sense below to the given local minimizer of the original sweeping control problem.

Our main attention in this section is paid to relaxed $W^{1,2} \times L^2$ -local minimizers $(\bar{x}(\cdot), \bar{u}(\cdot))$ for (P) introduced in Definition 2.2 while recalling that the relaxation is not needed if either the set g(x, U) is convex, or $(\bar{x}(\cdot), \bar{u}(\cdot))$ is a strong local minimizer for (P); see the discussions in Section 2.

Given a relaxed $W^{1,2} \times L^2$ -local minimizer $(\bar{x}(\cdot), \bar{u}(\cdot))$, we construct the following family of discretetime problems $(P_m), m \in \mathbb{N}$, where F is defined in (2.4), and where z_m^{ji}, c_m^{ji} are taken from Theorem 3.1:

minimize
$$J_m[x_m, u_m] := \varphi(x_m(T)) + \frac{1}{2} \sum_{i=0}^{2^m - 1} \int_{t_m^i}^{t_m^{i+1}} \left(\left\| \frac{x_m^{i+1} - x_m^i}{h_m} - \dot{x}(t) \right\|^2 + \left\| u_m^i - \bar{u}^i \right\|^2 \right) dt$$
 (4.1)

over discrete trajectories $(x_m, u_m) = (x_m^0, x_m^1, \dots, x_m^{2^m}, u_m^0, u_m^1, \dots, u_m^{2^m-1})$ subject to the constraints

$$x_m^{i+1} \in x_m^i - h_m F(x_m^i, u_m^i) \text{ for } i = 0, \dots, 2^m - 1,$$
(4.2)

$$\langle z_m^{ji}, x_m^i \rangle \le c_m^{ji}$$
 for all $j = 1, \dots, s$ and $i = 1, \dots, 2^m$ with $x_m^0 := x_0 \in C, \ u_m^0 := \bar{u}(0),$

$$\sum_{i=0}^{m-1} \int_{t_m^i}^{t_m^{i+1}} \left(\left\| \frac{x_m^{i+1} - x_m^i}{h_m} - \dot{x}(t) \right\|^2 + \left\| u_m^i - \bar{u}^i \right\|^2 \right) dt \le \frac{\epsilon}{2},\tag{4.3}$$

$$u_m^i \in U$$
 for $i = 0, \dots, 2^m - 1.$ (4.4)

To implement the method of discrete approximation, we have to make sure that each problem (P_m) admits an optimal solution. By taking into account Theorem 3.1, we deduce it from the classical Weierstrass existence theorem in finite dimensions due to the construction of (P_m) and the assumptions made.

Proposition 4.1 In addition to the assumptions of Theorem 3.1, suppose that the cost function φ is lower semicontinuous (l.s.c.) on bounded sets. Then each problem (P_m) admits an optimal solution provided that $m \in \mathbb{N}$ is sufficiently large.

Proof. It follows from Theorem 3.1 that the set of feasible solutions (x_m, u_m) to (P_m) is nonempty for any large m. It follows from the constraint structures in (P_m) and the assumptions imposed on U and g that the feasible sets are closed. Furthermore, it easy to deduce from the localization in (4.3) that the feasible sets are bounded as well. Thus the lower semicontinuity assumption on the cost function φ ensures the existence of optimal solutions to (P_m) by the Weierstrass theorem. \Box

Now we are ready to derive the main result of this section that establishes the strong $W^{1,2}$ convergence of any sequence $(\bar{x}_m(\cdot), \bar{u}_m(\cdot))$ of optimal solutions to (P_m) , which are extended to the entire interval [0, T], to the given local minimizer $(\bar{x}(\cdot), \bar{u}(\cdot))$ for the original problem (P).

Theorem 4.2 Let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be a relaxed $W^{1,2} \times W^{1,2}$ -local minimizer for the sweeping control problem (P), and let φ be continuous around $\bar{x}(T)$ in addition to the assumptions of Theorem 3.1. Consider any sequence of optimal solutions $(\bar{x}_m(\cdot), \bar{u}_m(\cdot))$ to problems (P_m) and extend them to [0, T] piecewise linearly for $\bar{x}_m(\cdot)$ and piecewise constantly for $\bar{u}_m(\cdot)$ without relabeling. Then we have the convergence

$$(\bar{x}_m(\cdot), \bar{u}_m(\cdot)) \to (\bar{x}(\cdot), \bar{u}(\cdot))$$
 as $m \to \infty$

in the strong topology of $W^{1,2}([0,T];\mathbb{R}^n) \times L^2([0,T];\mathbb{R}^d)$.

Proof. It is sufficient to show that

$$\lim_{m \to \infty} \int_0^T \left(\left\| \dot{\bar{x}}_m(t) - \dot{\bar{x}}(t) \right\|^2 + \left\| \bar{u}_m(t) - \bar{u}(t) \right\|^2 \right) dt = 0.$$
(4.5)

Arguing by contradiction, suppose that there exists a subsequence of the integral values γ_m in (4.5) that converges, without relabeling, to some number $\gamma > 0$. Due to (4.3), the sequence of extended optimal solutions $\{(\dot{x}_m(\cdot), \bar{u}_m(\cdot))\}$ to (P_m) is bounded in the reflexive space $L^2([0, T]; \mathbb{R}^n) \times L^2([0, T]; \mathbb{R}^d)$, and thus it contains a weakly convergence subsequence in this product space, again without relabeling. Denote by $(\tilde{v}(\cdot), \tilde{u}(\cdot))$ the limit of the latter subsequence and then let

$$\widetilde{x}(t) := x_0 + \int_0^T \widetilde{v}(\tau) d\tau \text{ for all } t \in [0, T].$$

Since $\dot{\widetilde{x}}(t) = \widetilde{v}(t)$ for a.e. $t \in [0, T]$, we have that

$$(\bar{x}_m(\cdot), \bar{u}_m(\cdot)) \to (\tilde{x}(\cdot), \tilde{u}(\cdot))$$
 as $m \to \infty$

in the topology of $W^{1,2}([0,T];\mathbb{R}^n) \times L^2([0,T];\mathbb{R}^d)$. Invoking the Mazur weak closure theorem tells us that there is a sequence of convex combinations of $(\bar{x}_m(\cdot), \bar{u}_m(\cdot))$, which converges to $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ strongly in $W^{1,2}([0,T];\mathbb{R}^n) \times L^2([0,T];\mathbb{R}^d)$, and thus $(\dot{\bar{x}}_m(t), \bar{u}_m(t)) \to (\dot{\bar{x}}(t), \tilde{u}(t))$ for a.e. $t \in [0,T]$ along a subsequence. Furthermore, we can clearly replace above the piecewise linear extensions of the discrete trajectories $\bar{x}_m(\cdot)$ to the interval [0,T] by the trajectories of (1.2) generated by the controls $\bar{u}_m(\cdot)$ piecewise constantly extended to [0,T]. The obtained pointwise convergence of convex combinations allows us to conclude that $\tilde{u}(t) \in \text{co } U$ for a.e. $t \in [0,T]$ and that $\tilde{x}(\cdot)$ satisfies the convexified differential inclusion (2.3). Passing now to the limit as $m \to \infty$ in the cost functional and constraints (4.1)–(4.4) of problem (P_m) with taking into account the assumed local continuity of φ and the constructions above, we conclude that the pair $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ belongs to the prescribed $W^{1,2} \times L^2$ -neighborhood of the given local minimizer $(\bar{x}(\cdot), \bar{u}(\cdot))$ and satisfies the inequality

$$J[\tilde{x}, \tilde{u}] + \gamma/2 \le J[\bar{x}, \bar{u}] \Longrightarrow J[\tilde{x}, \tilde{u}] < J[\bar{x}, \bar{u}]$$

$$\tag{4.6}$$

due the aforementioned strong convergence of $(\bar{x}_m(\cdot), \bar{u}_m(\cdot))$ to $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ and the structure of (4.1). Appealing to Definition 2.2 tells us that (4.6) contradicts the very fact that $(\bar{x}(\cdot), \bar{u}(\cdot))$ is a relaxed $W^{1,2} \times L^2$ -local minimizer of (P). Thus we get (4.5) and complete the proof of the theorem. \Box

Recalling the discussion after Definition 2.2 leads us to the following consequence of Theorem 4.2, which provides the strong approximation of local minimizers for (P) without an explicit relaxation.

Corollary 4.3 In addition to the assumptions of Theorem 4.2, suppose that the sets g(x, U) and U are convex. Then the convergence result of Theorem 4.2 holds true.

5 Tools of Variational Analysis

The results of Section 4 make a bridge between the given local minimizer $(\bar{x}(\cdot), \bar{u}(\cdot))$ of the original problem (P) and (global) optimal solutions for the sequence of discrete approximations (P_m) that exist by Proposition 4.1 and strongly converge to $(\bar{x}(\cdot), \bar{u}(\cdot))$ by Theorem 4.2. This supports our approach to derive necessary optimality conditions for $(\bar{x}(\cdot), \bar{u}(\cdot))$ by establishing firstly necessary conditions for optimal solutions to the discrete-time problems (P_m) and then passing to the limit in them as $m \to \infty$.

Looking at the structures of each problem (P_m) and the equivalent problem of finite-dimensional mathematical programming defined in Section 6, we observe that they are always *nonsmooth* and *nonconvex*, even when the initial data of (P) possess these properties. This is due to the *graphical* set constraints associated with the discrete-time inclusions (4.2) that are generated by the normal cone mapping in (2.4).

To proceed with deriving necessary optimality conditions for (P_m) and then for (P) by passing to the limit, we have to employ appropriate generalized differential constructions of variational analysis. These

constructions should be *robust*, enjoy comprehensive *calculus rules*, and such that the corresponding normal cone is *not too large* while being applied to–specifically–graphical sets. It does hold, in particular, for the Clarke normal cone \overline{N} , which is always a linear subspace of a maximum dimension for sets that are graphically homeomorphic to graphs of Lipschitzian functions; see [20, 25] for more details and references. For example, we have $\overline{N}((0,0); \text{gph} |x|) = \mathbb{R}^2$ for the graph of the simplest convex function on \mathbb{R} .

All the required properties are satisfied for the generalized differential constructions initiated by the second author. Elements of the first-order theory and various applications can be found by now in many books; see, e.g., [20]–[22], [25], [27]. We refer the reader to [21, 22] and the bibliographies therein for second-order constructions used in what follows.

To briefly overview the needed notions, recall first the (Painlevé-Kuratowski) *outer limit* of a set-valued mapping/multifunction $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ at \bar{x} with $F(\bar{x}) \neq \emptyset$ given by

$$\lim_{x \to \bar{x}} \sup F(x) := \left\{ y \in \mathbb{R}^m \mid \exists \text{ sequences } x_k \to \bar{x}, \ y_k \to y \text{ such that } y_k \in F(x_k), \ k \in \mathbb{N} \right\}$$
(5.1)

Given now a set $\Omega \subset \mathbb{R}^n$ locally closed around $\bar{x} \in \Omega$, we define by using (5.1) the (basic, limiting, Mordukhovich) normal cone to Ω at \bar{x} by

$$N(\bar{x};\Omega) = N_{\Omega}(\bar{x}) := \limsup_{x \to \bar{x}} \{\operatorname{cone}[x - \Pi(x;\Omega)]\}.$$
(5.2)

where $\Pi(x;\Omega) := \{ u \in \Omega \mid ||x - u|| = \operatorname{dist}(x;\Omega) \}$ is the Euclidean projection of x onto Ω , and where 'cone' stands for the (nonconvex) conic hull of the set. When Ω is convex, (5.2) reduces to the normal cone of convex analysis, but it is often nonconvex otherwise.

Given further a set-valued mapping $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with its domain and graph

dom
$$F := \left\{ x \in \mathbb{R}^n \mid F(x) \neq \emptyset \right\}$$
 and $\operatorname{gph} F := \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in F(x) \right\}$

locally closed around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, the *coderivative* of F at (\bar{x}, \bar{y}) is generated by (5.2) as

$$D^*F(\bar{x},\bar{y})(u) := \left\{ v \in \mathbb{R}^n \mid (v,-u) \in N\left((\bar{x},\bar{y}); \operatorname{gph} F\right) \right\}, \quad u \in \mathbb{R}^m.$$
(5.3)

When $F: \mathbb{R}^n \to \mathbb{R}^m$ is single-valued and continuously differentiable (\mathcal{C}^1 -smooth) around \bar{x} , we have

$$D^*F(\bar{x})(u) = \left\{ \nabla F(\bar{x})^*u \right\} \text{ for all } u \in \mathbb{R}^m$$

via the adjoint/transposed Jacobian matrix $\nabla F(\bar{x})^*$, where $\bar{y} = F(\bar{x})$ is omitted.

Let $\phi \colon \mathbb{R}^n \to \overline{\mathbb{R}} := (-\infty, \infty]$ be an extended-real-valued l.s.c. function $\phi \colon \mathbb{R}^n \to \overline{\mathbb{R}} := (-\infty, \infty]$ with

dom
$$\phi := \left\{ x \in \mathbb{R}^n \mid \varphi(x) < \infty \right\}$$
 and $\operatorname{epi} \phi := \left\{ (x, \alpha) \in \mathbb{R}^{n+1} \mid \alpha \ge \phi(x) \right\}$

standing for its domain and epigraph. The (first-order) subdifferential of ϕ at $\bar{x} \in \text{dom } \phi$ is defined geometrically via the normal cone (5.2) by

$$\partial \phi(\bar{x}) := \left\{ v \in \mathbb{R}^m \mid (v, -1) \in N((\bar{x}, \phi(\bar{x})); \operatorname{epi} \phi) \right\}$$
(5.4)

while admitting equivalent analytic representations; see, e.g., [20, 25]. Note that $N(\bar{x};\Omega) = \partial \delta(\bar{x};\Omega)$ for any $\bar{x} \in \Omega$, where $\delta(x;\Omega)$ denotes the indicator function of Ω equal to 0 for $x \in \Omega$ and ∞ otherwise. Then given a subgradient $\bar{v} \in \partial \phi(\bar{x})$ and following [20, 22], we define the *second-order subdifferential* (or *generalized Hessian*) of ϕ at \bar{x} relative to \bar{v} by

$$\partial^2 \phi(\bar{x}, \bar{v})(u) := (D^* \partial \phi)(\bar{x}, \bar{v})(u), \quad u \in \mathbb{R}^n,$$

via the coderivative (5.3) of the first-order subdifferential mapping $x \mapsto \partial \phi(x)$ from (5.4). If the function ϕ is \mathcal{C}^2 -smooth around \bar{x} , then we have the representation

$$\partial^2 \phi(\bar{x}, \bar{v})(u) = \left\{ \nabla^2 \phi(\bar{x})u \right\} \text{ for all } u \in \mathbb{R}^n,$$

where $\nabla^2 \phi(\bar{x})$ stands for the classical (symmetric) Hessian of ϕ at \bar{x} with $\bar{v} = \nabla \phi(\bar{x})$. If $\phi(x) := \delta(x;\Omega)$, then $\partial^2 \varphi(\bar{x},\bar{v})(u) = (D^*N_{\Omega})(\bar{x},\bar{v})(u)$ for any $\bar{v} \in N(\bar{x};\Omega)$ and $u \in \mathbb{R}^n$. The latter second-order construction is evaluated below in the case of the polyhedral set $\Omega = C$ from (1.3). To proceed, define the index sets corresponding the generating vectors x_i^* in (1.3) by

$$I_0(w) := \left\{ j \in I(x) \mid \langle x_*^j, w \rangle = c_j \right\} \text{ and } I_>(w) := \left\{ j \in I(x) \mid \langle x_*^j, w \rangle > c_j \right\}, \ w \in \mathbb{R}^n.$$
(5.5)

where I(x) is taken from (2.1) with $\bar{x} := x \in C$. The next theorem provides an effective upper estimate of the coderivative of F from (2.4) with ensuring the equality under an additional assumption on x_*^j .

Theorem 5.1 Given F in (2.4) with C from (1.3), denote G(x) := N(x; C) and suppose in addition to standing assumptions that g is C^1 -smooth around the reference points. Then for any $(x, u) \in C \times U$ and $\omega + g(x, u) \in G(x)$ we have the coderivative upper estimate

$$D^*F(x, u, \omega)(w) \subset \Big\{ z = \Big(-\nabla_x g(x, u)^* w + \sum_{j \in I_0(w) \cup I_>(w)} \gamma^j x^j_*, -\nabla_u g(x, u)^* w \Big) \Big\},$$
(5.6)

where $w \in \text{dom} D^*G(x, \omega + g(x, u))$, where $I_0(w)$ and $I_>(w)$ are taken from (5.5), and where $\gamma^j \in \mathbb{R}$ for $j \in I_0(w)$, while $\gamma^j \ge 0$ for $j \in I_>(w)$. Furthermore, (5.6) holds as an equality and the domain $\text{dom} D^*G(x, \omega + g(x, u))$ can be computed by

$$\operatorname{dom} D^*G(x,\omega+g(x,u)) = \left\{ w \middle| \exists \lambda^j \ge 0 \text{ with } \omega + g(x,u) = \sum_{j \in I(x)} \lambda^j x^j_*, \ \lambda^j > 0 \Longrightarrow \langle x^j_*, w \rangle = c_j \right\}$$
(5.7)

provided that the generating vectors $\{x_*^j \mid j \in I(x)\}$ of the polyhedron C are linearly independent.

Proof. Picking any $w \in \text{dom } D^*G(x, \omega + g(x, u))$ and $z \in D^*F(x, u, y)(w)$ and then denoting $\widetilde{G}(x, u) := G(x)$ and $\widetilde{f}(x, u) := -g(x, u)$, we deduce from [21, Theorem 1.62] that

$$z \in \nabla \widetilde{f}(x, u)^* w + D^* \widetilde{G}(x, u, \omega + g(x, u))(w).$$

Observe then the obvious composition representation

$$\widetilde{G}(x,u) = G \circ \widetilde{g}(x,u) \text{ with } \widetilde{g}(x,u) := x,$$

where the latter mapping has the surjective derivative. It follows from [21, Theorem 1.66] that

$$z \in \nabla \widetilde{f}(x, u)^* w + \nabla \widetilde{g}(x, u)^* D^* G((x, \omega + g(x, u))(w).$$
(5.8)

Employing now in (5.8) the coderivative estimate for the normal cone mapping G obtained in [17, Theorem 4.5] with the exact coderivative calculation given in [17, Theorem 4.6] under the linear independence of the generating vectors x_*^j and also taking into account the structure of the mapping \tilde{f} in (5.8), we arrive at (5.6) and the equality therein under the aforementioned assumption.

6 Necessary Optimality Conditions for Discrete-Time Problems

Here we derive necessary optimality conditions for solutions to each problem (P_m) , $m \in \mathbb{N}$, formulated in (4.1)–(4.4). It will be done by reducing each (P_m) to a nondynamic problem of nondifferentiable programming with functional and many geometric constraints, then employing necessary optimality conditions for the latter problem obtained in terms of generalized differential constructions of Section 5, and finally expressing the obtained conditions in terms of the given data of (P_m) by using calculus rules of generalized differentiation. In this way we arrive at the following necessary conditions, which will be further specified below by applying the second-order calculations presented in Section 5. **Theorem 6.1** Let $(\bar{x}_m, \bar{u}_m) = (\bar{x}_m^0, \dots, \bar{x}_m^{2^m}, \bar{u}_m^0, \dots, \bar{u}_m^{2^m-1})$ be an optimal solution to problem (P_m) . Assume that gph F is closed and the function φ is Lipschitz continuous around the point $\bar{x}_m(T)$. Then there are elements $\lambda_m \geq 0$, $\psi_m = (\psi_m^0, \dots, \psi_m^{2^m-1})$ with $\psi_m^i \in N(\bar{u}_m^i; U)$ as $i = 0, \dots, 2^m - 1$, $\xi_m = (\xi_m^1, \dots, \xi_m^s) \in \mathbb{R}^s_+$, and $p_m^i \in \mathbb{R}^n$ as $i = 0, \dots, 2^m$ satisfying the conditions

$$\lambda_m + \|\xi_m\| + \sum_{i=0}^{2^m - 1} \|p_m^i\| + \|\psi_m\| \neq 0,$$
(6.1)

$$\xi_m^j \left(\langle z_m^{j2^m}, x_m^{2^m} \rangle - c_m^{j2^m} \right) = 0, \quad j = 1, \dots, s,$$
(6.2)

$$-p_m^{2^m} = \lambda_m \vartheta_m^{2^m} + \sum_{j=1}^s \xi_m^j z_m^{j2^m} \in \lambda_m \partial \varphi(\bar{x}_m^{2^m}) + \sum_{j=1}^s \xi_m^j z_m^{j2^m},$$
(6.3)

$$\left(\frac{p_m^{i+1} - p_m^i}{h_m}, -\frac{1}{h_m}\lambda_m\theta_m^{iu}, \frac{1}{h_m}\lambda_m\theta_m^{iy} - p_m^{i+1}\right) \\ \in \left(0, \frac{1}{h_m}\psi_m^i, 0\right) + N\left(\left(\bar{x}_m^i, \bar{u}_m^i, -\frac{\bar{x}_m^{i+1} - \bar{x}_m^i}{h_m}\right); \operatorname{gph} F\right)$$
(6.4)

for $i = 0, ..., 2^m - 1$, where we use the notation

$$\theta_m^i = \left(\theta_m^{iy}, \theta_m^{iu}\right) := \left(\int_{t_m^i}^{t_m^{i+1}} \left\|\frac{\bar{x}_m^{i+1} - \bar{x}_m^i}{h_m} - \dot{x}(t)\right\| dt, \int_{t_m^i}^{t_m^{i+1}} \|\bar{u}_m^i - \bar{u}(t)\| dt\right).$$
(6.5)

Proof. Denote $z := (x_m^0, \ldots, x_m^{2^m}, u_m^0, \ldots, u_m^{2^m-1}, y_m^0, \ldots, y_m^{2^m-1}) \in \mathbb{R}^{(2 \cdot 2^m + 1)n + 2^m \cdot d}$, where the starting point x_m^0 is fixed. Taking $\epsilon > 0$ from (P_m) , consider the following problem of mathematical programming (MP) with respect to the variable z:

minimize
$$\phi_0(z) := \varphi(x(T)) + \frac{1}{2} \sum_{i=0}^{2^m - 1} \int_{t_m^i}^{t_m^{i+1}} \left\| \left(y_m^i - \dot{x}(t), u_m^i - \bar{u}(t) \right) \right\|^2 dt$$

subject to finitely many equality, inequality, and geometric constraints given by

$$\begin{split} \phi(z) &:= \sum_{i=0}^{2^m - 1} \int_{t_m^i}^{t_m^{i+1}} \left\| \left(y_m^i, u_m^i \right) - \left(\dot{\bar{x}}(t), \bar{u}(t) \right) \right\|^2 dt - \frac{\epsilon}{2} \le 0, \\ g_i(z) &:= x_m^{i+1} - x_m^i - h_m y_m^i = 0, \quad i = 0, \dots, 2^m - 1, \\ h_j(z) &:= \langle z_m^{j2^m}, x_m^{2^m} \rangle - c_m^{j2^m} \le 0, \quad j = 1, \dots, s, \\ z &\in \Xi_i := \left\{ (x_m^0, \dots, y_m^{2^m - 1}) \in \mathbb{R}^{(2 \cdot 2^m + 1)n + 2^m \cdot d} \ \right| - y_m^i \in F(x_m^i, u_m^i) \right\}, \quad i = 0, \dots, 2^m - 1, \\ z &\in \Xi_{2^m} := \left\{ (x_m^0, \dots, y_m^{2^m - 1}) \in \mathbb{R}^{(2 \cdot 2^m + 1)n + 2^m \cdot d} \ \right| \ u_m^i \in U \right\}, \quad i = 0, \dots, 2^m - 1, \end{split}$$

Necessary optimality conditions for problem (MP) in terms of the generalized differential tools reviewed above can deduced from [22, Proposition 6.4 and Theorem 6.5]. We specify them for the optimal solution

$$\bar{z} := \left(\bar{x}_m^0, \dots, \bar{x}_m^{2^m}, \bar{u}_m^0, \dots, \bar{u}_m^{2^m-1}, \bar{y}_m^0, \dots, \bar{y}_m^{2^m-1}\right)$$

to (MP). It follows from Theorem 4.2 that the inequality constraint in (MP) defined by ϕ is inactive for large m, and so the corresponding multiplier does not appear in the optimality conditions. Thus we can find $\lambda_m \geq 0$, $\xi_m = (\xi_m^1, \ldots, \xi_m^s) \in \mathbb{R}^s_+$, $p_m^i \in \mathbb{R}^n$ as $i = 1, \ldots, 2^m$, and

$$z_i^* = \left(x_{0i}^*, \dots, x_{2^m i}^*, u_{0i}^*, \dots, u_{(2^m - 1)i}^*, y_{0i}^*, y_{1i}^*, \dots, y_{(2^m - 1)i}^*\right), \quad i = 0, \dots, 2^m,$$

which are not zero simultaneously while satisfying the conditions

$$z_i^* \in \begin{cases} N(\bar{z}; \Xi_i) + N(\bar{z}; \Omega_i) & \text{if } i \in \{0, \dots, 2^m - 1\}, \\ N(\bar{z}; \Xi_i); & \text{if } i = 2^m, \end{cases}$$
(6.6)

$$-z_{0}^{*} - \dots - z_{2^{m}}^{*} \in \lambda_{m} \partial \phi_{0}(\bar{z}) + \sum_{j=1}^{s} \xi_{m}^{j} \nabla h_{j}(\bar{z}) + \sum_{i=0}^{2^{m}-1} \nabla g_{i}(\bar{z})^{*} p_{m}^{i+1},$$

$$\xi_{m}^{j} h_{j}(\bar{z}) = 0, \quad j = 1, \dots, s.$$
(6.7)

Note that the first line in (6.6) comes by applying the normal cone intersection formula from [20, Corollary 3.5] to $\bar{z} \in \Omega_i \cap \Xi_i$ for $i = 0, \ldots, 2^m - 1$. It follows from the structure of the sets Ω_i and Ξ_i that the inclusions in (6.6) can be equivalently written as

$$\left(x_{ii}^{*}, u_{ii}^{*} - \psi_{m}^{i}, -y_{ii}^{*}\right) \in N\left(\left(\bar{x}_{m}^{i}, \bar{u}_{m}^{i}, -\frac{\bar{x}_{m}^{i+1} - \bar{x}_{m}^{i}}{h_{m}}\right); \operatorname{gph} F\right) \text{ for } i = 0, \dots, 2^{m} - 1$$
(6.8)

with every other components of z_i^* equal to zero, where $\psi_m^i \in N(\bar{u}_m^i; U)$ for all $i = 0, \ldots, 2^m - 1$. Observe furthermore that x_{0m}^* and u_{0m}^* determined by the normal cone to Ξ_m are the only nonzero components of z_m^* . This implies by using (6.6)–(6.7) that

$$-z_0^* - \ldots - z_{2^m}^* \in \lambda_m \partial \phi_0(\bar{z}) + \sum_{j=1}^s \xi_m^j \nabla h_j(\bar{z}) + \sum_{i=0}^{2^m - 1} \nabla g_i(\bar{z})^* p_m^{i+1},$$

with $\xi_m^j \left(\langle z_m^{j2^m}, x_m^{2^m} \rangle - c_m^{j2^m} \right) = 0$, $j = 1, \ldots, s$. Using the expressions for ϕ_0 , g_i , and h_j above together with the elementary subdifferential sum rule from [20, Proposition 1.107] gives the calculations

$$\begin{split} \left(\sum_{j=1}^{s} \xi_{m}^{j} \nabla h_{j}(\bar{z})\right)_{x_{m}^{2m}} &= \left(\sum_{j=1}^{s} \xi_{m}^{j} z_{m}^{j2^{m}}\right), \\ \left(\sum_{i=0}^{2^{m}-1} \nabla g_{i}(\bar{z})^{*} p_{m}^{i+1}\right)_{x_{m}^{i}} &= \begin{cases} p_{m}^{i} - p_{m}^{1}; \text{ if } i = 0, \\ p_{m}^{i+1} \text{ if } i = 1, \dots, 2^{m} - 1, \\ p_{m}^{2^{m}} \text{ if } i = 2^{m}, \end{cases} \\ \left(\sum_{i=0}^{2^{m}-1} \nabla g_{i}(\bar{z})^{*} p_{m}^{i+1}\right)_{y_{m}^{i}} &= \left(-h_{m} p_{m}^{1}, -h_{m} p_{m}^{2}, \dots, -h_{m} p_{m}^{2^{m}}\right), \\ \partial(\phi_{0})(\bar{z}) &= \partial\varphi(\bar{x}_{m}^{m}) + \frac{1}{2} \sum_{i=0}^{2^{m}-1} \nabla \rho_{i}(\bar{z}) \text{ with } \rho_{i}(\bar{z}) := \int_{t_{m}^{i}}^{t_{m}^{i+1}} \left\|\left(\frac{\bar{x}_{m}^{i+1} - \bar{x}_{m}^{i}}{h_{m}} - \dot{x}(t), \bar{u}_{m}^{i} - \bar{u}(t)\right)\right\|^{2} dt. \end{split}$$

The set $\lambda_m \partial \phi_0(\bar{z})$ is represented as the collection of

$$\lambda_m \left(0, \dots, 0, \vartheta_m^{2^m}, \theta_m^{0u}, \dots, \theta_m^{(2^m - 1)u}, \theta_m^{0y}, \dots, \theta_m^{(2^m - 1)y} \right) \text{ with } \vartheta_m^{2^m} \in \partial \varphi(\bar{x}_m^{2^m}),$$
$$(\theta_m^{iu}, \theta_m^{iy}) = \left(\int_{t_m^i}^{t_m^{i+1}} \|\bar{u}_m^i - \bar{u}(t)\| dt, \int_{t_m^i}^{t_m^{i+1}} \left\| \frac{\bar{x}_m^{i+1} - \bar{x}_m^i}{h_m} - \dot{\bar{x}}(t) \right\| dt \right), \quad i = 0, \dots, 2^m - 1.$$

Thus we obtain the following relationships

$$-x_{00}^* - x_{02^m}^* = -p_m^{1x}, (6.9)$$

$$-x_{ii}^* = p_m^i - p_m^{i+1}, \quad i = 1, \dots, 2^m - 1,$$
(6.10)

$$0 = \lambda_m \vartheta_m^{2^m} + p_m^{2^m} + \sum_{j=1}^s \xi_m^j z_m^{j2^m} \quad \text{with} \quad \vartheta_m^{2^m} \in \partial \varphi(\bar{x}_m^{2^m}), \tag{6.11}$$

$$-u_{00}^{*} = \lambda_{m} \theta_{m}^{0u} \text{ and } -u_{ii}^{*} = \lambda_{m} \theta_{m}^{iu}, \quad i = 1, \dots, 2^{m} - 1,$$
(6.12)

$$-y_{ii}^* = \lambda_m \theta_m^{iy} - h_m p_m^{i+1}, \quad i = 0, \dots, 2^m - 1,$$
(6.13)

which allow us to arrive at all the necessary optimality conditions claimed in the theorem. Indeed, observe first that (6.7) yields (6.2). Extending p_m by $p_m^0 := x_{02^m}^*$ ensures that (6.3) follows from (6.11). Then we deduce from (6.10), (6.12), and (6.13) that

$$\frac{x_{ii}^*}{h_m} = \frac{p_m^{i+1} - p_m^i}{h_m}, \quad \frac{u_{ii}^*}{h_m} = -\frac{1}{h_m} \lambda_m \theta_m^{iu}, \text{ and } \frac{y_{ii}^*}{h_m} = -\frac{1}{h_m} \lambda_m \theta_m^{iy} + p_m^{i+1}.$$

Substituting this into the left-hand side of (6.8) justifies the discrete-time adjoint inclusion (6.4).

Finally, to verify (6.1) we argue by contradiction and suppose that $\lambda_m = 0, \xi_m = 0, \psi_m = 0$, and $p_m^i = 0$ as $i = 0, \ldots, 2^m - 1$, which yield $x_{02^m}^* = p_m^0 = 0$. Then it follows from (6.11) that $p_m^{2^m} = 0$, and so $p_m^i = 0$ whenever $i = 0, \ldots, 2^m$. By (6.9) and (6.10) we get $x_{ii}^* = 0$ for all $i = 0, \ldots, 2^m - 1$. Using (6.12) tells us that $u_{ii}^* = 0$ as $i = 1, \ldots, 2^m - 1$. Since the first condition in (6.12) yields also $u_{00}^* = 0$, it follows that $u_{ii}^* = 0$ for $i = 0, \ldots, 2^m - 1$. In addition we have by (6.13) that $y_{ii}^* = 0$ for all $i = 0, \ldots, 2^m - 1$. Remembering that the components of z_i^* different from $(x_{ii}^*, u_{ii}^*, y_{ii}^*)$ are zero for $i = 0, \ldots, 2^m - 1$ ensures that $z_i^* = 0$ for $i = 0, \ldots, 2^m - 1$ and similarly $z_{2^m}^* = 0$. Therefore $z_i^* = 0$ for all $i = 0, \ldots, 2^m$, which violates the nontriviality condition for (MP) and thus completes the proof.

The next theorem applies to (6.4) the calculation result of Theorem 4.2 and provides in this way necessary optimality conditions for problem (P_m) expressed entirely via its initial data.

Theorem 6.2 Let (\bar{x}_m, \bar{u}_m) be an optimal solution to problem (P_m) formulated in (4.1)–(4.4), where the cost function φ is locally Lipschitzian around $\bar{x}_m(T)$, and where the sweeping mapping F is defined in (2.4). Using the notation and assumptions of Theorem 5.1, take $(\theta_m^{iu}, \theta_m^{iy})$ from (6.5). Then there exists dual elements (λ_m, ψ_m, p_m) as in Theorem 6.1 together with vectors $\eta_m^i \in \mathbb{R}^s$ for $i = 0, \ldots, 2^m$ and $\gamma_m^i \in \mathbb{R}^s$ for $i = 0, \ldots, 2^m - 1$ satisfying the nontriviality condition

$$\lambda_m + \left\| \eta_m^{2^m} \right\| + \sum_{i=0}^{2^m - 1} \left\| p_m^i \right\| + \left\| \psi_m \right\| \neq 0,$$
(6.14)

the primal-dual relationships given for all $i = 0, ..., 2^m - 1$ and j = 1, ..., s by

$$-\frac{\bar{x}_m^{i+1} - \bar{x}_m^i}{h_m} + g(\bar{x}_m^i, \bar{u}_m^i) = \sum_{j \in I(\bar{x}_m^i)} \eta_m^{ij} z_m^{ji},$$
(6.15)

$$\frac{p_m^{i+1} - p_m^i}{h_m} = -\nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^* \left(-\frac{1}{h_m} \lambda_m \theta_m^{iy} + p_m^{i+1} \right) \\
+ \sum_{j \in I_0(p_m^{i+1} - \frac{1}{h_m} \lambda_m \theta_m^{iy}) \cup I_>(p_m^{i+1} - \frac{1}{h_m} \lambda_m \theta_m^{iy})} \gamma_m^{ij} z_m^{ji},$$
(6.16)

$$-\frac{1}{h_m}\lambda_m\theta_m^{iu} - \frac{1}{h_m}\psi_m^i = -\nabla_u g(\bar{x}_m^i, \bar{u}_m^i)^* \left(-\frac{1}{h_m}\lambda_m\theta_m^{iy} + p_m^{i+1}\right)$$
(6.17)

with $\psi_m^i \in N(\bar{u}_m^i; U)$ as $i = 0, \dots, 2^m - 1$ taken from Theorem 6.1, the transversality condition

$$-p_m^{2^m} = \lambda_m \vartheta_m^{2^m} + \sum_{j=1}^s \eta_m^{2^m j} z_m^{j2^m} \in \lambda_m \partial \varphi(\bar{x}_m^{2^m}) + \sum_{j=1}^s \eta_m^{2^m j} z_m^{j2^m},$$
(6.18)

and such that the following implications hold for $i = 0, ..., 2^m - 1$ and j = 1, ..., s:

$$\left[\langle z_m^{ji}, \bar{x}_m^i \rangle < c_m^{ji} \right] \Longrightarrow \eta_m^{ij} = 0, \tag{6.19}$$

$$\begin{cases} \left[j \in I_0(p_m^{i+1} - \frac{1}{h_m}\lambda_m\theta_m^{iy}) \right] \Longrightarrow \gamma_m^{ij} \in \mathbb{R}, \\ \left[j \in I_>(p_m^{i+1} - \frac{1}{h_m}\lambda_m\theta_m^{iy}) \right] \Longrightarrow \gamma_m^{ij} \ge 0, \\ \left[j \notin I_0(p_m^{i+1} - \frac{1}{h_m}\lambda_m\theta_m^{iy}) \cup I_>(p_m^{i+1} - \frac{1}{h_m}\lambda_m\theta_m^{iy}) \right] \Longrightarrow \gamma_m^{ij} = 0. \end{cases}$$
(6.20)

We also have the complementary slackness condition (6.2) together with

$$\left[\langle z_m^{ji}, \bar{x}_m^i \rangle < c_m^{ji}\right] \Longrightarrow \gamma_m^{ij} = 0 \quad for \quad i = 0, \dots, 2^m - 1 \quad and \quad j = 1, \dots, s,$$

$$(6.21)$$

$$\left[\langle z_m^{j2^{m}}, \bar{x}_m^{2^{m}} \rangle < c_m^{j2^{m}} \right] \Longrightarrow \eta_m^{2^{m}j} = 0 \quad for \quad j = 1, \dots, s,$$
(6.22)

Furthermore, the linear independence of the vectors $\{z_m^{ji} \mid j \in I(\bar{x}_m^i)\}$ ensures the implication

$$\eta_m^{ij} > 0 \Longrightarrow \left[\left\langle z_m^{ji}, p_m^{i+1} - \frac{1}{h_m} \lambda_m \theta_m^{iy} \right\rangle = c_m^{ji} \right]$$
(6.23)

Assuming in addition that the matrices $\nabla_u g(\bar{x}_m^i, \bar{u}_m^i)$ are of full rank for all $i = 0, \ldots, 2^m - 1$ and $m \in \mathbb{N}$ sufficiently large, we get the enhanced nontriviality condition

$$\lambda_m + \|p_m^0\| + \|\psi_m\| \neq 0.$$
(6.24)

Proof. Using the necessary optimality conditions of Theorem 5.1, we can rewrite (6.4) as

$$\left(\frac{p_m^{i+1} - p_m^i}{h_m}, -\frac{1}{h_m}\lambda_m\theta_m^{iu} - \frac{1}{h_m}\psi_m^i\right) \in D^*F\left(\bar{x}_m^i, \bar{u}_m^i, -\frac{\bar{x}_m^{i+1} - \bar{x}_m^i}{h_m}\right)\left(-\frac{1}{h_m}\lambda_m\theta_m^{iy} + p_m^{i+1}\right)$$
(6.25)

for all $i = 0, ..., 2^m - 1$ by the coderivative definition (5.3). Taking into account that

$$-\frac{\bar{x}_m^{i+1} - \bar{x}_m^i}{h_m} + g(\bar{x}_m^i, \bar{u}_m^i) \in G(\bar{x}_m^i) \text{ for } i = 0, \dots, 2^m - 1$$
(6.26)

with G(x) = N(x;C), we find vectors $\eta_m^i \in \mathbb{R}^s_+$ as $i = 0, \ldots, 2^m - 1$ such that conditions (6.15) and (6.19) hold. Employing now the coderivative evaluation (5.6) from Theorem 5.1 with $x := \bar{x}_m^i$, $u := \bar{u}_m^i$, $\omega := -\frac{\bar{x}_m^{i+1} - \bar{x}_m^i}{h_m}$, and $w := -\frac{1}{h_m} \lambda_m \theta_m^{iy} + p_m^{i+1}$ for $i = 0, \ldots, 2^m - 1$ gives us $\gamma_m^i \in \mathbb{R}^s$ and the relationships

$$\left(\frac{p_m^{i+1}-p_m^i}{h_m},-\frac{1}{h_m}\lambda_m\theta_m^{iu}-\frac{\psi_m^{iu}}{h_m}\right)$$

$$=\left(-\nabla_x g(\bar{x}_m^i,\bar{u}_m^i)^*\left(-\frac{1}{h_m}\lambda_m\theta_m^{iy}+p_m^{i+1}\right)+\sum_{\substack{j\in I_0(p_m^{i+1}-\frac{1}{h_m}\lambda_m\theta_m^{iy})\cup I_>(p_m^{i+1}-\frac{1}{h_m}\lambda_m\theta_m^{iy})}{0}\right)$$

$$-\nabla_u g(\bar{x}_m^i,\bar{u}_m^i)^*\left(-\frac{1}{h_m}\lambda_m\theta_m^{iy}+p_m^{i+1}\right)$$

$$\psi_m^{iu}\left(\nu-\bar{u}_m^i\right)\leq 0 \text{ for all } \nu\in U \text{ and } i=0,\ldots,2^m-1.$$

This ensures the validity of all the conditions in (6.16), (6.17), (6.20), and (6.21). Denoting $\eta_m^{2^m} := \xi_m$ with ξ_m taken from Theorem 6.1, we get $\eta_m^i \in \mathbb{R}^s_+$ for all $i = 0, \ldots, 2^m$ and deduce (6.14) and (6.18) from those in (6.1) and (6.3). Implication (6.22) follows directly from (6.2) and the definition of $\eta_m^{2^m}$.

Assume finally that the generating vectors $\{z_m^{ji} \mid j \in I(\bar{x}_m^i)\}$ are linear independent. In this case we deduce from (5.7) and (6.25) that condition (6.23) is satisfied. It remains to verify the enhanced nontriviality (6.24) under the additional assumption on the full rank of the matrices $\nabla_u g(\bar{x}_m^i, \bar{u}_m^i)$. Suppose on the contrary that $\lambda_m = 0$, $p_m^0 = 0$, and $\psi_m = 0$. Then $p_m^{i+1} = 0$ as $i = 0, \ldots, 2^m - 1$ by (6.17). Then it follows from (6.16) the equality

$$\sum_{j\in I_0(p_m^{i+1}-\frac{1}{h_m}\lambda_m\theta_m^{iy})\cup I_>(p_m^{i+1}-\frac{1}{h_m}\lambda_m\theta_m^{iy})}\gamma_m^{ij}z_m^{ji}=0.$$

Invoking now (6.18) and $p_m^{2^m} = 0$ tells us that $\sum_{j=1}^s \eta_m^{2^m j} z_m^{j2^m} = 0$. This implies by definition (2.1) of the active constraint indices and the imposed linear independence of z_m^{ji} over this index set that $\eta_m^{2^m} = 0$. Thus (6.14) is violated, which verifies (6.24) and completes the proof of the theorem. \Box

7 Optimality Conditions for the Controlled Sweeping Process

In this section we derive necessary optimality conditions for the local minimizer under consideration in the original problem (P) by passing to the limit as $m \to \infty$ in the necessary optimality conditions of Theorem 6.1 for the discrete-time problems (P_m). Furnishing the limiting procedure requires the usage of Theorem 4.2 and the tools of generalized differentiation reviewed in Section 5.

Theorem 7.1 Let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be a relaxed $W^{1,2} \times L^2$ -local minimizer of problem (P) such that $\bar{u}(\cdot)$ is of bounded variation and admits a right continuous representative on [0, T]. In addition to (H1) and (H2), assume that LICQ holds along $\bar{x}(\cdot)$ on [0,T], that $g(\cdot, \cdot)$ is C^1 -smooth around $(\bar{x}(t), \bar{u}(t))$ with the full rank of the matrices $\nabla_u g(\bar{x}(t), \bar{u}(t))$ on [0,T], and that φ is locally Lipschitzian around $\bar{x}(T)$. Then there exist a multiplier $\lambda \geq 0$, a signed vector measure $\gamma = (\gamma^1, \ldots, \gamma^n) \in C^*([0,T]; \mathbb{R}^n)$ as well as adjoint arcs $p(\cdot) \in W^{1,2}([0,T]; \mathbb{R}^n)$ and $q(\cdot) \in BV([0,T]; \mathbb{R}^n)$ such that the following conditions are fulfilled:

(i) The PRIMAL-DUAL DYNAMIC RELATIONSHIPS:

$$-\dot{\bar{x}}(t) = \sum_{j=1}^{s} \eta^{j}(t) x_{*}^{j} - g(\bar{x}(t), \bar{u}(t)) \quad \text{for a.e.} \quad t \in [0, T],$$
(7.1)

where the functions $\eta^{j}(\cdot) \in L^{2}([0,T];\mathbb{R}_{+})$ are well defined at t = T while being uniquely determined by the representation in (7.1);

$$\dot{p}(t) = -\nabla_x g(\bar{x}(t), \bar{u}(t))^* q(t) \text{ for a.e. } t \in [0, T],$$
(7.2)

where the right continuous representative of $q(\cdot)$, with the same notation, satisfies

$$q(t) = p(t) - \int_{(t,T]} d\gamma(\tau)$$
(7.3)

for all $t \in [0,T]$ except at most a countable subset;

$$\psi(t) := \nabla_u g(\bar{x}(t), \bar{u}(t))^* q(t) \in N(\bar{u}(t); U) \text{ for a.e. } t \in [0, T],$$
(7.4)

which gives us the MAXIMIZATION CONDITION

$$\langle \psi(t), \bar{u}(t) \rangle = \max_{u \in U} \langle \psi(t), u \rangle \text{ for a.e. } t \in [0, T]$$
 (7.5)

provided that the set U is convex. Furthermore, for a.e. $t \in [0,T]$ including t = T and for all j = 1, ..., s we have the COMPLEMENTARITY CONDITIONS

$$\langle x_*^j, \bar{x}(t) \rangle < c_j \Longrightarrow \eta^j(t) = 0 \quad and \quad \eta^j(t) > 0 \Longrightarrow \langle x_*^j, q(t) \rangle = c_j.$$
 (7.6)

(ii) The TRANSVERSALITY CONDITIONS at the right endpoint:

$$-p(T) - \sum_{j \in I(\bar{x}(T))} \eta^j(T) x^j_* \in \lambda \partial \varphi(\bar{x}(T)) \quad and \quad \sum_{j \in I(\bar{x}(T))} \eta^j(T) x^j_* \in N(\bar{x}(T); C).$$
(7.7)

- (iii) The MEASURE NONATOMICITY CONDITION: If $t \in [0,T)$ and $\langle x_*^j, \bar{x}(t) \rangle < c_j$ for all $j = 1, \ldots, s$, then there is a neighborhood V_t of t in [0,T] such that $\gamma(V) = 0$ for all the Borel subsets V of V_t .
- (iv) NONTRIVIALITY CONDITIONS: It always holds that

$$\lambda + \|p(T)\| + \|q(0)\| > 0.$$
(7.8)

Assuming in addition that $\langle x_*^j, x_0 \rangle < c_j$ for all $j = 1, \ldots, s$, we have the ENHANCED NONTRIVIALITY

$$\lambda + \|p(T)\| > 0. \tag{7.9}$$

Proof. Given the local minimizer $(\bar{x}(\cdot), \bar{u}(\cdot))$ for (P), construct the discrete-time problems (P_m) for which optimal solutions $(\bar{x}_m(\cdot), \bar{u}_m(\cdot))$ exist by Proposition 3.1 and converge to $(\bar{x}(\cdot), \bar{u}(\cdot))$ in the sense of Theorem 4.2. We derive each of the claimed necessary conditions in (P) by passing to the limit from those in Theorem 6.1. Let us split the derivation into several steps.

Step 1: Verifying the primal equation and complementarity condition. First we prove (7.1) together with the first complementarity condition in (7.6). Based on (6.5), define the functions

$$\theta_m(t) := \frac{\theta_m^i}{h_m}$$
 for $t \in [t_m^i, t_m^{i+1})$ and $i = 0, \dots, 2^m - 1$

on [0,T] whenever $m \in \mathbb{N}$. It is easy to see that

$$\begin{split} \int_{0}^{T} \|\theta_{m}^{y}(t)\|^{2} dt &= \sum_{i=0}^{2^{m}-1} \frac{\left\|\theta_{m}^{iy}\right\|^{2}}{h_{m}} \leq \frac{1}{h_{m}} \sum_{i=0}^{2^{m}-1} \left(\int_{t_{m}^{i}}^{t_{m}^{i+1}} \left\|\dot{\bar{x}}(t) - \frac{\bar{x}_{m}^{i+1} - \bar{x}_{m}^{i}}{h_{m}}\right\| dt\right)^{2} \\ &\leq \sum_{i=0}^{2^{m}-1} \int_{t_{m}^{i}}^{t_{m}^{i+1}} \left\|\dot{\bar{x}}(t) - \frac{\bar{x}_{m}^{i+1} - \bar{x}_{m}^{i}}{h_{m}}\right\|^{2} dt = \int_{0}^{T} \left\|\dot{\bar{x}}(t) - \dot{\bar{x}}_{m}(t)\right\|^{2} dt. \end{split}$$

Using the strong convergence $(\bar{x}_m(\cdot), \bar{u}_m(\cdot)) \to (\bar{x}(\cdot), \bar{u}(\cdot))$ in Theorem 4.2 ensures that

$$\int_{0}^{T} \|\theta_{m}^{y}(t)\|^{2} dt \leq \int_{0}^{T} \|\dot{\bar{x}}(t) - \dot{\bar{x}}_{m}(t)\|^{2} dt \to 0 \text{ as } m \to \infty.$$
(7.10)

This implies that a subsequence of $\{\theta_m^y(t)\}$ converges, without relabeling, to zero a.e. on [0, T]. Likewise

$$\int_{0}^{T} \left\| \theta_{m}^{u}(t) \right\|^{2} dt = \sum_{i=0}^{2^{m}-1} \frac{\left\| \theta_{m}^{iu} \right\|^{2}}{h_{m}} \leq \frac{1}{h_{m}} \sum_{i=0}^{2^{m}-1} \left(\int_{t_{m}^{i}}^{t_{m}^{i+1}} \left\| \bar{u}_{m}^{i} - \bar{u}(t) \right\| dt \right)^{2}$$
$$\leq \sum_{i=0}^{2^{m}-1} \int_{t_{m}^{i}}^{t_{m}^{i+1}} \left\| \bar{u}_{m}^{i} - \bar{u}(t) \right\|^{2} dt = \int_{0}^{T} \left\| \bar{u}_{m}(t) - \bar{u}(t) \right\|^{2} dt,$$

which tells us, again by using Theorem 4.2, that

$$\int_{0}^{T} \left\|\theta_{m}^{u}(t)\right\|^{2} dt \leq \int_{0}^{T} \left\|\bar{u}_{m}(t) - \bar{u}(t)\right\|^{2} dt \to 0 \text{ as } m \to \infty,$$
(7.11)

and so $\theta_m^u(t) \to 0$ for a.e. $t \in [0,T]$ along a subsequence. The assumed LICQ along $\bar{x}(\cdot)$ and the robustness of this condition yields by the choice of z_m^{ji} and the convergence in Theorem 4.2 that the vectors $\{z_m^{ji} \mid j \in I(\bar{x}_m^i)\}$ are linearly independent for each $i = 1, \ldots, 2^m$ and $m \in \mathbb{N}$ sufficiently large.

Taking $\eta_m^i \in \mathbb{R}^s_+$ from Theorem 6.2, we construct the piecewise constant functions $\eta_m(\cdot)$ on [0,T] by $\eta_m(t) := \eta_m^i$ for $t \in [t_m^i, t_m^{i+1})$ with $\eta_m(T) := \eta_m^{2^m}$. It follows from (6.15) that

$$-\dot{\bar{x}}_m(t) = \sum_{j=1}^s \eta_m^j(t) z_m^{ji} - g\left(\bar{x}_m(t_m^i), \bar{u}_m(t_m^i)\right) \text{ whenever } t \in (t_m^i, t_m^{i+1}), \quad m \in \mathbb{N}.$$
(7.12)

Furthermore, we get $-\dot{x}(t) \in G(\bar{x}(t)) - g(\bar{x}(t), \bar{u}(t))$ for a.e. $t \in [0, T]$ with the mapping $G(\cdot) = N(\cdot; C)$, which is measurable by [25, Theorem 4.26]. The well-known measurable selection result (see, e.g., [25, Corollary 4.6]) allows us to find nonnegative measurable functions $\eta^{j}(\cdot)$ on [0, T] for $j = 1, \ldots, s$ such that equation (7.1) holds. Combining (7.12) and (7.1) implies that

$$\dot{\bar{x}}(t) - \dot{\bar{x}}_m(t) = \sum_{j=1}^s \left[\eta_m^j(t) z_m^{ji} - \eta^j(t) x_*^j \right] + g(\bar{x}(t), \bar{u}(t)) - g(\bar{x}_m(t_m^i), \bar{u}_m(t_m^i))$$

for $t \in (t_m^i, t_m^{i+1})$ and $i = 0, ..., 2^m - 1$. It follows from the imposed LICQ that the functions $\eta_m^j(t)$ and $\eta^j(t)$ are uniquely defined for a.e. $t \in [0, T]$ and belong to $L^2([0, T]; \mathbb{R}_+)$. The constructions above yield the estimate

$$\left\|\sum_{j=1}^{s} \left[\eta^{j}(t)x_{*}^{j} - \eta_{m}^{j}(t)z_{m}^{ji}\right]\right\|_{L^{2}} \leq \left\|\dot{\bar{x}}_{m}(t) - \dot{\bar{x}}(t)\right\|_{L^{2}} + \left\|g\left(\bar{x}(t), \bar{u}(t)\right) - g\left(\bar{x}_{m}(t), \bar{u}_{m}(t)\right)\right\|_{L^{2}}$$

whenever $t \in (t_m^i, t_m^{i+1})$. Passing to the limit therein with the usage of Theorem 4.2 gives us

$$\sum_{j\in I(\bar{x}(t))} \left[\eta^j(t) x^j_* - \eta^j_m(t) z^{ji}_m \right] \to 0 \text{ as } m \to \infty \text{ for a.e. } t \in [0,T]$$

and ensures the a.e. convergence $\eta_m(t) \to \eta(t)$ on [0,T] by the imposed LICQ. We also have that the sequence $\{\eta_m^{2^m}\}$ converges to the well-defined vector $(\eta^1(T), \ldots, \eta^{2^m}(T))$. Then the first complementarity condition in (7.6) follows from (6.19) and (6.22).

Step 2: Continuous-time extensions of approximating dual elements. In the notation of Theorem 6.1, define $q_m(t)$ by extending $p_m^i(t)$ piecewise linearly on [0,T] with $q_m(t_m^i) := p_m^i$ for $i = 0, \ldots, 2^m$. Construct further $\gamma_m(t)$ and $\psi_m(t)$ on [0,T] by

$$\gamma_m(t) := \gamma_m^i, \quad \psi_m(t) := \frac{1}{h_m} \psi_m^i \text{ for } t \in [t_m^i, t_m^{i+1}) \text{ and } i = 0, \dots, 2^m - 1$$
 (7.13)

with $\gamma_m(T) := 0$ and $\psi_m(T) := 0$. Define now the function

$$\nu_m(t) := \max\left\{ t_m^i \mid t_m^i \le t, \ 0 \le i \le 2^m - 1 \right\} \text{ for all } t \in [0, T], \quad m \in \mathbb{N},$$

and deduce respectively from (6.16) and (6.17) that

$$\dot{q}_{m}(t) = -\nabla_{x}g(\bar{x}_{m}(\nu_{m}(t)), \bar{u}_{m}(\nu_{m}(t)))^{*}(-\lambda_{m}\theta_{m}^{y}(t) + q_{m}(\nu_{m}(t) + h_{m})) + \sum_{j \in I_{0}(-\lambda_{m}\theta_{m}^{y}(t) + q_{m}(\nu_{m}(t) + h_{m})) \cup I_{>}(-\lambda_{m}\theta_{m}^{y}(t) + q_{m}(\nu_{m}(t) + h_{m}))} \gamma_{m}^{j}(t)z_{m}^{ji}, \quad \text{and}$$

$$(7.14)$$

$$-\lambda_m \theta_m^u(t) - \psi_m(t) = -\nabla_u g \left(\bar{x}_m(\nu_m(t)), \bar{u}_m(\nu_m(t)) \right)^* \left(-\lambda_m \theta_m^y(t) + q_m(\nu_m(t) + h_m) \right)$$
(7.15)

for every $t \in (t_m^i, t_m^{i+1})$ and $i = 0, \ldots, 2^m - 1$. Next we extend the adjoint arcs $p_m(\cdot)$ to [0, T] by

$$p_m(t) := q_m(t) + \int_t^T \left(\sum_{j=1}^s \gamma_m^j(\tau) z_m^{ji}\right) d\tau \text{ for every } t \in [0, T].$$
(7.16)

This shows that $p_m(T) = q_m(T)$ and that

$$\dot{p}_m(t) = \dot{q}_m(t) - \sum_{j=1}^s \gamma_m^j(t) z_m^{ji}$$
 a.e. $t \in [0, T].$ (7.17)

The latter implies due to (7.14), (7.15), and the index definitions in (5.5) that

$$\dot{p}_m(t) = -\nabla_x g \left(\bar{x}_m(\nu_m(t)), \bar{u}_m(\nu_m(t)) \right)^* \left(-\lambda_m \theta_m^y(t) + q_m(\nu_m(t) + h_m) \right)$$
(7.18)

for every $t \in (t_m^i, t_m^{i+1})$ and $i = 0, \ldots, 2^m - 1$. Define now the vector measures γ_m^{mes} on [0, T] by

$$\int_{B} d\gamma_m^{mes} := \int_{B} \sum_{j=1}^{s} \gamma_m^j(t) z_m^{ji} dt$$
(7.19)

for every Borel subset $B \subset [0, T]$ and then drop for simplicity the index "mes" in what follows if no confusion arises. Since all the expressions in the statement of Theorem 6.1 are positively homogeneous of degree one with respect to λ_m , p_m , γ_m , and ψ_m , the enhanced nontriviality condition (6.24) and the constructions above allow us to normalize them by imposing the sequential equality

$$\lambda_m + \|p_m(T)\| + \|q_m(0)\| + \int_0^T \left\| \sum_{j=1}^s \gamma_m^j(t) z_m^{ji} \right\| dt + \int_0^T \|\psi_m(t)\| dt = 1, \quad m \in \mathbb{N},$$
(7.20)

which tells us, in particular, that all the terms in (7.20) are uniformly bounded.

Step 3: Verifying the dual dynamic relationships and the maximization condition. By (7.20), suppose without loss of generality that $\lambda_m \to \lambda$ as $m \to \infty$ for some $\lambda \ge 0$. To prove the uniform boundedness of the sequence $\{p_m^0, \ldots, p_m^{2^m}\}_{m \in \mathbb{N}}$ for all $i = 0, \ldots, 2^m - 1, m \in \mathbb{N}$, observe first from (6.16) that

$$p_m^{i+1} = p_m^i - h_m \nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^* \left(-\frac{1}{h_m} \lambda_m \theta_m^{iy} + p_m^{i+1} \right) + h_m \sum_{j=1}^s \gamma_m^{ij} z_m^{ji}$$

for all $i = 0, \ldots, 2^m - 1$. This implies that

$$\begin{aligned} \|p_m^i\| &\leq \|p_m^{i+1}\| + h_m \|\nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^*\| \cdot \left\| \left(-\frac{1}{h_m} \lambda_m \theta_m^{iy} + p_m^{i+1} \right) \right\| + h_m \left\| \sum_{j=1}^s \gamma_m^{ij} z_m^{ji} \right\| \\ &= \left(1 + h_m \|\nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^*\| \right) \|p_m^{i+1}\| + h_m \lambda_m \|\theta_m^y(t_m^i)\| \cdot \|\nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^*\| + h_m \left\| \sum_{j=1}^s \gamma_m^{ij} z_m^{ji} \right\| \end{aligned}$$

whenever $i = 0, ..., 2^m - 1$. It follows from (7.10) and (7.20) that the quantities $\nabla_x g(\bar{x}_m^i, \bar{u}_m^i), \lambda_m \theta_m^{iy}$, and $\sum_{j=1}^s \gamma_m^{ij} z_m^{ji}$ are uniformly bounded for $i = 0, ..., 2^m - 1$. Thus we find a constant $M_1 > 0$ such that

$$h_m \lambda_m \left\| \theta_m^y(t_m^i) \right\| \cdot \left\| \nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^* \right\| \le M_1 h_m \left\| \theta_m^y(t_m^i) \right\| = M_1 \sqrt{h_m \int_{t_m^i}^{t_m^{i+1}} \left\| \theta_m^y(t) \right\|^2 dt}$$

for all $i = 0, \ldots, 2^m - 1$ and $m \in \mathbb{N}$. It implies that

$$\sum_{i=0}^{2^{m}-1} h_{m} \lambda_{m} \left\| \theta_{m}^{y}(t_{m}^{i}) \right\| \cdot \left\| \nabla_{x} g(\bar{x}_{m}^{i}, \bar{u}_{m}^{i})^{*} \right\| \leq M_{1} \sqrt{\int_{0}^{T} \left\| \theta_{m}^{y}(t) \right\|^{2} dt} \to 0 \text{ as } m \to \infty.$$

On the other hand, we get due to (7.20) that

$$\sum_{i=0}^{2^{m}-1} h_{m} \Big\| \sum_{j=1}^{s} \gamma_{m}^{ij} z_{m}^{ji} \Big\| = \int_{0}^{T} \Big\| \sum_{j=1}^{s} \gamma_{m}^{j}(t) z_{m}^{ji} \Big\| dt \le 1.$$
(7.21)

Considering now the numbers

$$A_m^i \coloneqq h_m \lambda_m \left\| \theta_m^y(t_m^i) \right\| \cdot \left\| \nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^* \right\| + h_m \left\| \sum_{j=1}^s \gamma_m^{ij} z_m^{ji} \right\|$$

for $i = 0, \ldots, 2^m - 1$ and using the aforementioned uniform boundedness, find a constant $M_2 > 0$ such that $\sum_{i=0}^{2^m-1} A_m^i \leq M_2$. Combining the latter with the estimates above tells us that

$$\|p_m^i\| \le \left(1 + M_1 h_m\right) \|p_m^{i+1}\| + A_m^i, \quad i = 0, \dots, 2^m - 1.$$
(7.22)

Proceeding further by induction, we get the inequalities

$$\begin{aligned} \|p_m^i\| &\leq \left(1+M_1h_m\right)^{2^m-i} \|p_m^{2^m}\| + \sum_{j=i}^{2^m-1} A_m^j (1+M_1h_m)^{j-i} \\ &\leq e^{M_1} + e^{M_1} \sum_{i=0}^{2^m-1} A_m^i \leq e^{M_1} (1+M_2) \text{ for } i=2,\ldots,2^m-1, \end{aligned}$$

which imply in turn the estimate

 $||p_m^i|| \le M_3$ for some $M_3 > 0$ and all $i = 2, \dots, 2^m - 1$.

Hence the boundedness of $\{p_m^0\}$ and $\{p_m^1\}$ follows from (7.22) and the boundedness of $\{p_m^i\}_{2 \le i \le 2^m}$, which thus justifies the boundedness of the whole bundle $\{(p_m^0, \ldots, p_m^{2^m})\}_{m \in \mathbb{N}}$.

To verify the uniform boundedness of $q_m(\cdot)$, derive from their constructions and (6.16) that

$$\sum_{i=0}^{2^{m}-1} \left\| q_{m}(t_{m}^{i+1}) - q_{m}(t_{m}^{i}) \right\| \leq h_{m} \sum_{i=0}^{2^{m}-1} \left\| \nabla_{x} g(\bar{x}_{m}^{i}, \bar{u}_{m}^{i})^{*} (-\lambda_{m} \theta_{m}^{y}(t^{i}) + p_{m}^{i+1}) \right\| + \int_{0}^{T} \left\| \sum_{j=1}^{s} \gamma_{m}^{j}(t) z_{m}^{ji} \right\| dt$$

$$(7.23)$$

and observe furthermore that

$$h_m \sum_{i=0}^{2^m-1} \|\nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^* (-\lambda_m \theta_m^y(t^i) + p_m^{i+1})\| \le T \max_{0 \le i \le 2^m-1} \{ \|\nabla_x g(\bar{x}_m^i, \bar{u}_m^i)^* (-\lambda_m \theta_m^y(t^i) + p_m^{i+1})\| \}.$$

The latter ensures the boundedness of the first term on the right-hand side of (7.23) due to the boundedness of $\{p_m^i\}_{m \in \mathbb{N}}$, while the boundedness of the second term therein follows from (7.21). Thus we get from (7.23) that the functions $q_m(\cdot)$ on [0, T] are of uniform bounded variation on [0, T] and that

$$2 \|q_m(t)\| - \|q_m(0)\| - \|q_m(T)\| \le \|q_m(t) - q_m(0)\| + \|q_m(T) - q_m(t)\| \le \operatorname{var}(q_m; [0, T])$$

for all $t \in [0, T]$. Thus the sequence $\{q_m(\cdot)\}$ is bounded on [0, T] since the boundedness of $\{q_m(0)\}$ and $\{q_m(T)\}$ follows from (7.20). Applying now Helly's selection theorem gives us a function of bounded variation $q(\cdot)$ such that $q_m(t) \to q(t)$ as $m \to \infty$ pointwise on [0, T].

We see from (7.19) and (7.20) that the measure sequence $\{\gamma_m\}$ is bounded in $C^*([0,T];\mathbb{R}^n)$. Thus the weak^{*} sequential compactness of bounded sets in this space allows us to find a measure $\gamma \in C^*([0,T];\mathbb{R}^n)$ such that $\{\gamma_m\}$ weak^{*} converges to γ in $C^*([0,T];\mathbb{R}^n)$ along a subsequence. It follows from (7.18), (7.20), and the uniform boundedness of $q_m(\cdot)$ on [0,T] that the sequence $\{p_m(\cdot)\}$ is bounded in $W^{1,2}([0,T];\mathbb{R}^n)$ and thus weakly compact in this space. By Mazur's theorem we conclude that a sequence of convex combinations of $\dot{p}_m(\cdot)$ converges to some $\dot{p}(\cdot) \in W^{1,2}([0,T];\mathbb{R}^n)$ a.e. pointwise on [0,T]. This gives us (7.2) by passing to the limit along (7.18) as $m \to \infty$ with the usage of (7.10) and (7.11). Note also that

$$\left\|\int_{t}^{T}\sum_{j=1}^{s}\gamma_{m}^{j}(\tau)z_{m}^{ji}d\tau - \int_{(t,T]}d\gamma(\tau)\right\| = \left\|\int_{t}^{T}d\gamma_{m}(\tau) - \int_{(t,T]}d\gamma(\tau)\right\| \to 0 \quad \text{as} \quad m \to \infty$$
(7.24)

for all $t \in [0, T]$ except a countable subset of [0, T] by the weak^{*} convergence of the measures γ_m to γ in $C^*([0, T]; \mathbb{R}^n)$; cf. [27, p. 325] for similar arguments. Hence we get the convergence5

$$\int_{t}^{T} \sum_{j=1}^{s} \gamma_{m}^{j}(\tau) z_{m}^{ji} d\tau \to \int_{(t,T]} d\gamma(\tau) \quad \text{on} \quad [0,T] \quad \text{as} \quad m \to \infty$$
(7.25)

and thus arrive at (7.3) by passing to the limit in (7.16). The second (dual) complementarity condition in (7.6) follows from (6.23) while arguing by contradiction with the usage of the established a.e. pointwise convergence of the functions involved therein.

To finish the proof of (i), it remains verifying the validity of the inclusion in (7.4) and the maximization condition (7.5). Using the strong convergence of the discrete optimal solutions from Theorem 4.2, the convergence of $(\theta_m^u(t), \theta_m^u(t)) \to (0, 0)$ for a.e. $t \in [0, T]$ obtained above as well as the robustness of the normal cone (5.2), we arrive at (7.4) by passing the limit in (6.17) and in the inclusions $\psi_m^i \in N(\bar{u}_m^i; U)$, $i = 0, \ldots, 2^m - 1$, of Theorem 6.2. If U is convex, the maximization condition (7.5) follows directly from (7.4) due to the structure (1.4) of the normal cone to convex sets.

Step 4: Verifying transversality inclusions. It follows from (6.18) and representation (2.5) that

$$-p_m^{2^m} - \lambda_m \vartheta_m^{2^m} = \sum_{j=1}^s \eta_m^{2^m j} z_m^{j2^m} = \sum_{j \in I(\bar{x}_m^{2^m})} \eta_m^{2^m j} z_m^{j2^m} \in N(\bar{x}_m^{2^m}; C_m^{2^m}),$$
(7.26)

where $\eta_m^{2^m j} = 0$ for $j \in \{1, \ldots, s\} \setminus I(\bar{x}_m^{2^m})$. Denoting $\zeta_m := \sum_{j \in I(\bar{x}_m^{2^m})} \eta_m^{2^m j} z_m^{j2^m}$, observe that a subsequence $\{\zeta_m\}$ converges to some $\zeta \in \mathbb{R}^n$ due to the boundedness of λ_m by (7.20) and the convergence of $\{p_m^{2^m}\}$ and $\{\bar{x}_m^{2^m}\}$ with taking into account the robustness of the subdifferential. It follows from the robustness of the normal cone in (7.26), the convergence of $\bar{x}_m^{2^m} \to \bar{x}(T)$, and the inclusion $I(\bar{x}_m^{2^m}) \subset I(\bar{x}(T))$ for all m sufficiently large that $\zeta \in N(\bar{x}(T); C)$. Thus we get from (6.18)) that

$$-p_m^{2^m} - \zeta_m \in \lambda_m \partial \varphi(\bar{x}_m^{2^m}) \text{ for all } m \in \mathbb{N}.$$

Passing now to the limit therein as $m \to \infty$ verifies both inclusions in (7.7).

Step 5: Verifying measure nonatomicity. Take $t \in [0,T]$ with $\langle x_s^j, \bar{x}(t) \rangle < c_j$ for all $j = 1, \ldots, s$ and by continuity of $\bar{x}(\cdot)$ find a neighborhood V_t of t such that $\langle x_s^j, \bar{x}(\tau) \rangle < c_j$ whenever $\tau \in V_t$ and $j = 1, \ldots, s$. Invoking Theorem 4.2 tells us that $\langle z_m^{ji}, \bar{x}_m(t_m^i) \rangle < c_m^{ji}$ if $t_m^i \in V_t$ for all $j = 1, \ldots, s$ and $m \in \mathbb{N}$ sufficiently large. Then we deduce from (6.21) that $\gamma_m(t) = 0$ on any Borel subset V of V_t . Hence

$$\|\gamma_m\|(V) = \int_V d\|\gamma_m\| = \int_V \|\gamma_m(t)\|dt = 0$$
(7.27)

by the construction of γ_m in (7.19). Passing now to limit therein and taking into account the measure convergence established above, we get $\|\gamma\|(V) = 0$, which justifies the claimed measure nonatomicity.

Step 6: Verifying nontriviality conditions. First we establish (7.8) under the general assumptions of the theorem. Arguing by contradiction, suppose that $\lambda = 0$, p(T) = 0 and q(0) = 0. Thus $\lambda_m \to 0$, $p_m(T) \to 0$, and $q_m(0) \to 0$ as $m \to \infty$. It follows from (7.13) that

$$\int_{0}^{T} \|\gamma_{m}(t)\| dt = \sum_{i=0}^{2^{m}-1} h_{m} \left\|\gamma_{m}^{i}\right\| \text{ and } \int_{0}^{T} \|\psi_{m}(t)\| dt = \sum_{i=0}^{2^{m}-1} h_{m} \frac{\|\psi_{m}^{i}\|}{h_{m}} = \sum_{i=0}^{2^{m}-1} \left\|\psi_{m}^{i}\right\|.$$
(7.28)

Let us now verify the limiting condition

$$\int_0^T \left\| \sum_{j=1}^s \gamma_m^j(t) z_m^{ji} \right\| dt \to 0 \quad \text{as} \quad m \to \infty.$$
(7.29)

Indeed, by $q_m(T) = p_m(T)$ and the assumption above we get $q_m(T) \to 0$ and thus deduce from (6.21) and (7.28) that $\int_0^T \|\gamma_m(t)\| dt \to 0$ as $m \to \infty$. Recalling that $\gamma_m^{ij} = 0$ for $i = 0, \ldots, 2^m - 1$ and $j = 1, \ldots, s$ by (6.21) and remembering the weak^{*} convergence of $\gamma_m(\cdot) \to \gamma(\cdot)$ in $C^*([0,T];\mathbb{R}^n)$ yield $\dot{p}(t) = \dot{q}(t)$ for a.e. $t \in [0,T]$ by passing to the limit in (7.17). Thus (7.2) reduces in this case to the linear ODE

$$\dot{q}(t) = -\nabla_x g(\bar{x}(t), \bar{u}(t))^* q(t)$$
 with $q(0) = 0$,

which has only the trivial solution $q(t) \equiv 0$ on [0, T]. This implies that

$$\max_{i=0,\dots,2^m} \{ \|p_m^i\| \} = \max_{t \in [0,T]} \{ \|q_m(t)\| \} \to 0 \text{ as } m \to \infty.$$
(7.30)

By the constructions above we can estimate the left-hand side of (7.29) by

$$\begin{split} &\int_{0}^{T} \left\| \sum_{j=1}^{s} \gamma_{m}^{j}(t) z_{m}^{ji} \right\| dt = \sum_{i=0}^{2^{m}-1} \left\| h_{m} \sum_{j=1}^{s} \gamma_{m}^{ij} z_{m}^{ji} \right\| \\ &\leq \sum_{i=0}^{2^{m}-1} \left\| p_{m}^{i+1} - p_{m}^{i} + h_{m} \nabla_{x} g(\bar{x}_{m}^{i}, \bar{u}_{m}^{i})^{*} \left(-\frac{1}{h_{m}} \lambda_{m} \theta_{m}^{iy} + p_{m}^{i+1} \right) \right\| \\ &\leq \sum_{i=0}^{2^{m}-1} \left\| p_{m}^{i+1} \right\| \cdot \left\| \left(1 + h_{m} \nabla_{x} g(\bar{x}_{m}^{i}, \bar{u}_{m}^{i})^{*} \right) \right\| + \sum_{i=0}^{2^{m}-1} \left\| p_{m}^{i} \right\| + \sum_{i=0}^{2^{m}-1} \left\| \nabla_{x} g(\bar{x}_{m}^{i}, \bar{u}_{m}^{i})^{*} \lambda_{m} \theta_{m}^{iy} \right\|. \end{split}$$

Then (7.30) and the uniform boundedness of $\nabla_x g(\bar{x}_m^i, \bar{u}_m^i)$ ensure that the first two terms in the last line of the obtained estimate disappear as $m \to \infty$. To deal with the third term therein, we get by the definition of θ_m^{iy} and Theorem 4.2 that

$$\sum_{i=0}^{2^{m}-1} \left\| \theta_{m}^{iy} \right\| = \sum_{i=0}^{2^{m}-1} \int_{t_{m}^{i}}^{t_{m}^{i+1}} \left\| \frac{\bar{x}_{m}^{i+1} - \bar{x}_{m}^{i}}{h_{m}} - \dot{\bar{x}}(t) \right\| dt = \int_{0}^{T} \left\| \dot{\bar{x}}_{m}(t) - \dot{\bar{x}}(t) \right\| dt \to 0 \text{ as } m \to \infty,$$
(7.31)

and therefore (7.29) is justified. To proceed further with ψ_m^i in (7.28), we get by (6.17) that

$$\sum_{i=0}^{2^{m}-1} \|\psi_{m}^{i}\| \leq \sum_{i=0}^{2^{m}-1} \|\lambda_{m}\theta_{m}^{iu}\| + \sum_{i=0}^{2^{m}-1} \|\lambda_{m}\nabla_{u}g(\bar{x}_{m}^{i},\bar{u}_{m}^{i})^{*}\theta_{m}^{iy}\| + \sum_{i=0}^{2^{m}-1} \|\lambda_{m}\nabla_{u}g(\bar{x}_{m}^{i},\bar{u}_{m}^{i})^{*}p_{m}^{i+1}\|,$$

which yields $\sum_{i=0}^{2^m-1} \|\psi_m^i\| \to 0$ due to (7.30), (7.31), and

$$\sum_{i=0}^{2^{m}-1} \|\theta_{m}^{iu}\| = \int_{0}^{T} \left\|\bar{u}_{m}(t) - \bar{u}(t)\right\| dt \to 0 \text{ as } m \to \infty$$

by Theorem 4.2. This shows that the violation of (7.8) implies the failure of (7.20), a contradiction.

To complete the proof of the theorem, it remains to verify the validity of the enhanced nontriviality condition (7.9) under the additional assumption made. Suppose on the contrary that $(\lambda, p(T)) = 0$ while $\langle x_*^j, x_0 \rangle < c_j$ for all $j = 1, \ldots, s$. It follows from the above arguments in the step, by using the complementarity conditions (6.21), that $\dot{p}(t) = 0$ for a.e. $t \in [0, T]$, which yields p(t) = p(T) = 0 on [0, T]. Then we get by (7.3) and (7.27) that

$$q(t) = \int_{(t,T]} d\gamma(\tau) = 0 \text{ for all } t \in [0,T] \setminus A,$$
(7.32)

where $A \subset [0, T]$ is a countable set. Consider the two possible cases regarding (7.32):

• $0 \notin A$, and thus q(0) = 0.

• $0 \in A$. In this case the measure nonatomicity condition and the fact that A is at most countable allow us to find $\tau > 0$, $\tau \notin A$, with $\int_{(0,\tau]} d\gamma(t) = 0$, and thus $q(0) = \int_{(\tau,T]} d\gamma(t) = 0$.

Hence we always have q(0) = 0 in (7.32) while showing in this way that the failure of (7.9) contradicts the validity of (7.8) established above.

8 Numerical Examples

In this section we consider two examples illustrating some characteristic features and strength of the necessary optimality conditions for the sweeping control problem (P) obtained in Theorem 7.1.

Prior to dealing with specific examples, let us present the following useful assertion, which is a consequence of the measure nonatomicity condition.

Proposition 8.1 Assume that $\langle x^*, \bar{x}(\tau) \rangle < c_j$ for all $\tau \in [t_1, t_2]$ with $t_1, t_2 \in [0, T)$ and some vector $x^* \in \{x_*^j \mid j = 1, ..., s\}$, and that the measure nonatomicity condition of Theorem 7.1 is satisfied with the measure γ . Then we have $\gamma([t_1, t_2]) = 0$ and $\gamma(\{\tau\}) = 0$ whenever $\tau \in [t_1, t_2]$, and so $\gamma((t_1, t_2)) = \gamma([t_1, t_2]) = \gamma((t_1, t_2)) = 0$.

Proof. Pick any $\tau \in [t_1, t_2]$ with $\langle x_1^*, \bar{x}(t) \rangle < c_j$ and find by the measure nonatomicity condition a neighborhood V_{τ} of τ in [0, T] such that $\gamma(V) = 0$ for all the Borel subsets V of V_{τ} ; in particular, $\gamma(\{\tau\}) = 0$. By $[t_1, t_2] \subset \bigcup_{\tau \in [t_1, t_2]} V_{\tau}$ and the compactness of $[t_1, t_2]$ we find $\tau_1, \ldots, \tau_l \in [t_1, t_2]$ with $[t_1, t_2] \subset \bigcup_{i=1}^l V_{\tau_i}$. Fix $i = 1, \ldots, l-1$ and take $\tilde{\tau}_i \in V_{\tau_i} \cap V_{\tau_{i+1}}$ with $[\tau_i, \tilde{\tau}_i] \subset V_{\tau_i}$ and $[\tilde{\tau}_i, \tau_{i+1}] \subset V_{\tau_{i+1}}$, where $\tau_1 := t_1$ and $\tau_l := t_2$. Then we arrive at the equalities

$$\gamma([t_1, t_2]) = \gamma\Big(\bigcup_{i=1}^{p-1} [\tau_i, \widetilde{\tau}_i) \cup [\widetilde{\tau}_i, \tau_{i+1})\Big) = \sum_{i=1}^{p-1} \Big(\gamma([\tau_i, \widetilde{\tau}_i)) + \gamma([\widetilde{\tau}_i, \tau_{i+1}))\Big) = 0,$$

which verify the claimed properties of the measure.

Our first example is two-dimensional with respect to both state and control variables.

Example 8.2 Consider the sweeping control problem of minimizing the cost functional

$$x_1(1) + x_2(1) \text{ subject to}$$

$$\begin{cases} \left(\dot{x}_1 \\ \dot{x}_2 \right) = \left(u_1 \\ u_2 \right) - N_C \left(x_1 \\ x_2 \right) \\ \text{with} \left(x_1 \\ x_2 \right) (0) = \left(\begin{matrix} 0 \\ x_2^0 \right) \end{cases}$$

where $C := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \ge 0\}$ and $(u_1, u_2) \in U := [-1, 1] \times [-1, 1]$. We rewrite the dynamics as

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix}(t) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(t) + \eta(t) \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \eta(t) \ge 0 \text{ a.e. } t \in [0,1].$$

A direct checking shows that if $x_2^0 \ge 1$ then the constraint is irrelevant and the optimal control is constant being equal to (-1, -1). If instead $0 \le x_2^0 < 1$, then the optimal couple is $\bar{u}_1(t) \equiv -1$ together with any measurable component $\bar{u}_2(t)$ such that $\bar{x}_2(1) = 0$.

The conditions of Theorem 7.1 tell us that:

(1)
$$p = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$$
 is constant on [0, 1] (by (7.2));
(2) $\begin{pmatrix} -p_1 \\ -p_2 \end{pmatrix} - \begin{pmatrix} 0 \\ -\eta(1) \end{pmatrix} = \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \lambda \ge 0$ (by (7.7));
(3) $x_2^0 > 0 \Longrightarrow \lambda + \|p\| > 0$ (by (7.9));
(4) $q(t) = p - \int_{[t,1]} d\gamma(\tau) = \psi(t) \in N_{[-1,1]^2} \begin{pmatrix} \bar{u}_1 \\ \bar{u}_2 \end{pmatrix}$ (by (7.4) and (7.3));
(5) $\lambda + \|p\| + \|q(0)\| > 0$ (by (7.8));
(6)] $\eta(t) = 0$ for a.e. $t \in [0, 1]$ with $\bar{x}_2(t) > 0$ and $\left[\eta(t) > 0 \Longrightarrow q(t) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0\right]$ a.e. $t \in [0, 1]$ (by (7.6));

(7) $d\gamma|_{\{t \mid \bar{x}_2(t) > 0\}} = 0$ (by the measure nonatomicity condition).

To apply these conditions, consider first the case where $x_2^0 > 1$ in which the constraint is automatically satisfied for all the trajectories. Since $\bar{x}_2(1) > 0$, we get $\eta(1) = 0$ from (6)). If $\lambda = 0$, then $p \equiv 0$ and the nontriviality condition (3) is violated. Thus we can suppose that $\lambda = 1$, and so $p = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Condition (7) implies that $d\gamma = 0$ on the set in question; hence $q \equiv p = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \equiv \psi$. This shows that $\psi = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Since $\psi \in N_{[-1,1]^2}\begin{pmatrix} \bar{u}_1 \\ \bar{u}_2 \end{pmatrix}$, the optimal control is $\bar{u}(t) \equiv \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. It conforms that in this case we do not loose information with respect to the classical PMP.

Consider now the case where $0 < x_2^0 \leq 1$. Assuming that $x_2(1) > 0$ yields $\eta(1) = 0$. Repeating the above arguments with the usage of (4) gives us the control $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ on [0, 1] while implying that $x_2(1) = 0$, a contradiction. Thus we get $x_2(1) = 0$, which tells us that $\bar{u}_2 \equiv -1$ in the case where $x_2^0 = 1$. Let us now deal with the first component u_1 . Again, $\eta(1) = 0$ implies that $\lambda = 1$ and that $\bar{u}_1 \equiv -1$ on [0, 1]. If $\eta(1) > 0$, then $\lambda = 0$ is forbidden by taking $p = -\eta(1)$ in (3), so $\bar{u}_1 \equiv -1$ is obtained as well. The case where $x_2^0 = 0$ requires a longer discussion, which we omit here for the sake of brevity.

This example was treated also in [3], and the given discussion allows us to compare the two sets of necessary conditions obtained in [3] and in this paper. Actually most of them, including the adjoint equation and the transversality condition, are different. Those presented here deal only with reference trajectories where the control has bounded variation, but are more detailed and-at least in this exampleare more effective for the control u_2 while being more difficult to use for u_1 . This difference can be explained by the methods that are used to obtain the necessary conditions. Actually, the argument presented here takes into account the constraint at all the steps of the procedure. On the contrary, the method used in [3] is based on penalization, and so it does not see the hard constraint in the approximation steps. This explains why it behaves well with respect to u_1 that is not influenced by the constraint, while it is almost degenerate with respect to u_2 .

The next example is also two-dimensional while addressing a more complicated polyhedral set C in comparison with the halfspace in Example 8.2.

Example 8.3 Consider problem (P) with the following initial data:

$$n = m = 2, \ T = 1, \ x_0 := \left(-\frac{1}{2}, -\frac{1}{2}\right), \ x_*^1 := (1, 0), \ x_*^2 := (0, 1), \ c_1 = c_2 = 0, \ \varphi(x) := \frac{\|x\|^2}{2}, \ g(u) = u$$

with feasible controls $u(t) = (u^1(t), u^2(t)) \in U$ a.e. $t \in [0, 1]$ taking values in the unit square $U \subset \mathbb{R}^2$ with respect to the maximum norm

$$U := \{ (u^1, u^2) \in \mathbb{R}^2 \mid \max\{u^1, u^2\} \le 1 \}.$$

Applying necessary optimality conditions of Theorem 7.1, we seek for solutions to (P) such that

$$\langle x_*^j, \bar{x}(t) \rangle < c_j = 0 \text{ for all } t \in [0, 1), \ j = 1, 2, \text{ and } \bar{x}(1) \in \mathrm{bd}(C),$$

$$(8.1)$$

and show that (8.1) holds for $\bar{x}(\cdot)$ found below. In the case of (P) under consideration these conditions say that there exist $\lambda \geq 0$ and $\eta(\cdot) = (\eta^1(\cdot), \eta^2(\cdot)) \in L^2([0,1]; \mathbb{R}^2_+)$ well defined at t = 1 such that:

(1)
$$\langle x_j^*, \bar{x}(t) \rangle \langle c_j \Longrightarrow \eta^j(t) = 0 \text{ for } j = 1, 2 \text{ and a.e. } t \in [0, 1] \text{ including } t = 1;$$

- (2) $\eta^j(t) > 0 \Longrightarrow q^j(t) = c_j \text{ for } j = 1, 2 \text{ and a.e. } t \in [0, 1];$
- (3) $-\dot{\bar{x}}(t) = \left(-\dot{\bar{x}}^{1}(t), -\dot{\bar{x}}^{2}(t)\right) = (\eta^{1}(t), \eta^{2}(t)) \left(\bar{u}^{1}(t), \bar{u}^{2}(t)\right)$ for a.e. $t \in [0, 1];$
- (4) $(\dot{p}^1(t), \dot{p}^2(t)) = (0, 0)$ for a.e. $t \in [0, 1];$
- (5) $(q^1(t), q^2(t)) \in N(\bar{u}(t); U)$ for a.e. $t \in [0, 1];$

(6) $q(t) = p(t) - \gamma([t, 1])$ for a.e. $t \in [0, 1];$

(7)
$$-p(1) = \lambda(\bar{x}^1(1), \bar{x}^2(1)) + (\eta^1(1), \eta^2(1)) \text{ with } (\eta^1(1), \eta^2(1)) \in N(\bar{x}(1); C);$$

(8) $\lambda + ||p(1)|| \neq 0.$

Employing the first condition in (8.1) together with (1) and (3), gives us $\dot{\bar{x}}(t) = \bar{u}(t)$ for a.e. $t \in [0, 1]$. It also follows from (5) and (6) that

$$q(t) = p(t) - \gamma([t, 1]) \in N(\bar{u}(t); U)$$
 for a.e. $t \in [0, 1]$,

which can be written in the maximization form (7.5). It follows from (4) that $p(\cdot)$ is constant on [0, 1], i.e., $p(t) \equiv p(1)$. This allows us to deduce that

$$q(t) = p(1) - \gamma([t, 1]) = p(1) - \gamma(\{1\})$$
 for a.e. $t \in [0, 1]$

by using the measure nonatomicity condition of Theorem 7.1 and Proposition 8.1. Considering control functions $\bar{u}(t) = (\vartheta_1, \vartheta_2)$ on [0, 1] and remembering the control constraints, we have $|\vartheta_1| \le 1$ and $|\vartheta_2| \le 1$. Thus $\bar{x}(t) = (-\frac{1}{2} + \vartheta_1 t, -\frac{1}{2} + \vartheta_2 t)$ for all $t \in [0, 1]$, and by the second condition in (8.1) provides the following two possibilities:

(1) $\bar{x}^1(1) = 0$. Then $\vartheta_1 = \frac{1}{2}$ and the cost functional reduces is $J[\bar{x}, \bar{u}] = \frac{(\vartheta_2 - \frac{1}{2})^2}{2}$. It obviously achieves its absolutely minimum value $\bar{J} = 0$ at the point $\vartheta_2 = \frac{1}{2}$.

(2) $\bar{x}^2(1) = 0$. Then $\vartheta_2 = \frac{1}{2}$ and the minimum cost is $\bar{J} = 0$ that is achieved at $\vartheta_1 = \frac{1}{2}$.

As a result, we arrive at a feasible solution giving the optimal value to the cost functionals:

$$\bar{u}(t) = \left(\frac{1}{2}, \frac{1}{2}\right)$$
 and $\bar{x}(t) = \left(-\frac{1}{2} + \frac{1}{2}t, -\frac{1}{2} + \frac{1}{2}t\right), \quad t \in [0, 1].$

satisfying all the assumptions above.

Acknowledgements. The authors are grateful to Tan Cao for many useful discussions.

References

- S. Adly, T. Haddad and L. Thibault, Convex sweeping process in the framework of measure differential inclusions and evolution variational inequalities, *Math. Program.* 148 (2014), 5–47.
- [2] L. Adam and J. V. Outrata, On optimal control of a sweeping process coupled with an ordinary differential equation, *Discrete Contin. Dyn. Syst. Ser. B* 19 (2014), 2709–2738.
- [3] C. E. Arroud and G. Colombo, A maximum principle of the controlled sweeping process, Set-Valued Var. Anal. 26 (2018), DOI 10.1007/s11228-017-0400-4.
- [4] H. Brézis, Operateurs Maximaux Monotones et Semi-Groupes de Contractions les Espaces de Hilbert, North-Holland, Amsterdam, 1973.
- [5] M. Brokate and P. Krejčí, Optimal control of ODE systems involving a rate independent variational inequality, Disc. Contin. Dyn. Syst. Ser. B 18 (2013), 331–348.
- [6] T. H. Cao and B. S. Mordukhovich, Optimal control of a perturbed sweeping process via discrete approximations, *Discrete Contin. Dyn. Syst. Ser. B* 21 (2016), pp. 3331-3358.
- [7] T. H. Cao and B. S. Mordukhovich, Optimality conditions for a controlled sweeping process with applications to the crowd motion model, *Discrete Contin. Dyn. Syst. Ser. B* 21 (2017), pp. 267-306.
- [8] T. H. Cao and B. S. Mordukhovich, Optimal control of a nonconvex perturbed sweeping process, to appear in J. Diff. Eqs., https://arxiv.org/abs/1711.02267.

- [9] C. Castaing, M. D. P. Monteiro Marques and P. Raynaud de Fitte, Some problems in optimal control governed by the sweeping process, J. Nonlinear Convex Anal. 15 (2014), 1043–1070.
- [10] G. Colombo, R. Henrion, N. D. Hoang and B. S. Mordukhovich, Optimal control of the sweeping process, Dyn. Contin. Discrete Impuls. Syst. Ser. B 19 (2012), 117–159.
- [11] G. Colombo, R. Henrion, N. D. Hoang and B. S. Mordukhovich, Optimal control of the sweeping process over polyhedral controlled sets, J. Diff. Eqs. 260 (2016), 3397–3447.
- [12] G. Colombo, B. S. Mordukhovich and D. Nguyen, Applications of controlled perturbed sweeping processes to practical modeling, in preparation.
- [13] G. Colombo and L. Thibault, Prox-regular sets and applications, in: D.Y. Gao and D. Motreanu (Eds.), *Handbook of Nonconvex Analysis*, International Press, Boston, 2010, pp. 99–182.
- [14] M. d. R. de Pinho, M. M. A. Ferreira and G. V. Smirnov, Optimal control involving sweeping processes, Set-Valued Var. Anal., to appear.
- [15] J. F. Edmond and L. Thibault, Relaxation of an optimal control problem involving a perturbed sweeping process, *Math. Program.* **104** (2005), 347–373.
- [16] T. Donchev, E. Farkhi and B. S. Mordukhovich, Discrete approximations, relaxation, and optimization of one-sided Lipschitzian differential inclusions in Hilbert spaces, J. Diff. Eqs. 243 (2007), 301–328.
- [17] R. Henrion, B. S. Mordukhovich and N. M. Nam, Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities, SIAM J. Optim. 20 (2010), 2199–2227.
- [18] N. D. Hoang and B. S. Mordukhovich, Extended Euler-Lagrange and Hamiltonian formalisms in optimal control of sweeping processes with controlled sweeping sets, to appear in *J. Optim. Theory Appl.*, arxiv:1804.10635.
- [19] B. S. Mordukhovich, Discrete approximations and refined Euler-Lagrange conditions for differential inclusions, SIAM J. Control Optim. 33 (1995), 882–915.
- [20] B. S. Mordukhovich, Variational Analysis and Generalized Differentiation, I: Basic Theory, Springer, Berlin, 2006.
- [21] B. S. Mordukhovich, Variational Analysis and Generalized Differentiation, II: Applications, Springer, Berlin, 2006.
- [22] B. S. Mordukhovich, Variational Analysis and Applications, Springer, Cham, Switzerland, 2018.
- [23] J. J. Moreau, On unilateral constraints, friction and plasticity, in: G. Capriz and G. Stampacchia (Eds.), New Variational Techniques in Mathematical Physics, Proceedings of C.I.M.E. Summer Schools, Cremonese, Rome, 1974, pp. 173–322.
- [24] S. M. Robinson, Strongly regular generalized equations, Math. Oper. Res. 5 (1980), 43–62.
- [25] R. T. Rockafellar and R. J-B. Wets, Variational Analysis, Springer, Berlin, 1998.
- [26] A. A. Tolstonogov, Control sweeping process, J. Convex Anal. 23 (2016), 1099–1123.
- [27] R. B. Vinter, Optimal Control, Birkhaüser, Boston, 2000.