ARTICLE IN PRESS
YJDEQ:10640

Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations ••• (••••) •••–•••

Journal of Differential Equations

www.elsevier.com/locate/jde

Optimization and discrete approximation of sweeping processes with controlled moving sets and perturbations

Tan H. Cao [a,1], Giovanni Colombo [b,2], Boris S. Mordukhovich [c,*,3], Dao Nguyen [c,4]

[a] *Department of Applied Mathematics and Statistics, State University of New York–Korea, Yeonsu-Gu, Incheon, Republic of Korea*

[b] *Dipartimento di Matematica "Tullio Levi-Civita", Università di Padova, via Trieste 63, 35121 Padua, Italy*

[c] *Department of Mathematics, Wayne State University, Detroit, MI 48202, USA*

Received 13 February 2020; revised 18 August 2020; accepted 12 October 2020

Abstract

This paper addresses a new class of optimal control problems for perturbed sweeping processes with measurable controls in additive perturbations of the dynamics and smooth controls in polyhedral moving sets. We develop a constructive discrete approximation procedure that allows us to strongly approximate any feasible trajectory of the controlled sweeping process by feasible discrete trajectories and also establish a $W^{1,2}$-strong convergence of optimal trajectories for discretized control problems to a given local minimizer of the original continuous-time sweeping control problem of the Bolza type. Employing advanced tools of first-order and second-order variational analysis and generalized differentiation, we derive necessary

* Corresponding author.

E-mail addresses: tan.cao@stonybrook.edu (T.H. Cao), colombo@math.unipd.it (G. Colombo), boris@math.wayne.edu (B.S. Mordukhovich), dao.nguyen2@wayne.edu (D. Nguyen).

1 Research of this author is supported by the National Research Foundation of Korea grant funded by the Korea Government (MIST) NRF-2020R1F1A1A01071015.

2 Research of this author is partially supported by the University of Padova grant SID 2018 "Controllability, stabilizability and infimum gaps for control systems," BIRD 187147, and is affiliated to Istituto Nazionale di Alta Matematica (GNAMPA).

3 Research of this author was partially supported by the USA National Science Foundation under grants DMS-1007132 and DMS-1512846, by the USA Air Force Office of Scientific Research grant #15RT0462, and by the Australian Research Council under Discovery Project DP-190100555.

4 Research of this author was partially supported by the USA National Science Foundation under grant DMS-1512846 and by the USA Air Force Office of Scientific Research grant #15RT0462.

https://doi.org/10.1016/j.jde.2020.10.017

optimality conditions for discrete optimal solutions under fairly general assumptions formulated entirely in terms of the given data. The obtained results give us efficient suboptimality ("almost optimality") conditions for the original sweeping control problem that are illustrated by a nontrivial numerical example.

MSC: 49J52; 49J53; 49K24; 49M25; 90C30

Keywords: Optimal control; Sweeping process; Discrete approximation; Variational analysis; Generalized differentiation; Necessary optimality conditions

1. Problem formulation and initial discussions

This paper is devoted to the study of *optimal control* problems for *sweeping processes* with *controlled perturbations* and *controlled moving sets*. The basic *uncontrolled* sweeping process was introduced by Moreau in the 1970s as the dissipative differential inclusion

$$\dot{x}(t) \in -N\big(x(t); C(t)\big) \text{ a.e. } t \in [0, T] \text{ with } x(0) := x_0 \in C(0) \tag{1.1}$$

describing the motion of a particle that belongs to a continuously moving set $C(t)$, where the normal cone N in (1.1) is understood in the sense of convex analysis

$$N(x; C) = N_C(x) := \big\{v \in \mathbb{R}^n \big|\ \langle v, y - x\rangle \le 0,\ y \in C\big\} \text{ if } x \in C \text{ and } N(x; C) := \emptyset \text{ if } x \notin C. \tag{1.2}$$

The sweeping inclusion (1.1) tells us that, depending on the motion of the set, the particle stays where it is in the case when it does not hit the set; otherwise, it is swept towards the interior of the set. We refer the reader to [31] and to the subsequent work in, e.g., [1,4,17,18,23–26,38] with the bibliographies therein for further developments and applications. The original motivation for Moreau came from applications to elastoplasticity, but later on the sweeping process and its modifications have been well recognized for many applications to other problems in mechanics, hysteresis, ferromagnetism, electric circuits, phase transitions, traffic equilibria, social and economic modelings, etc.; see, e.g., the references above among numerous publications.

Since the Cauchy problem in (1.1) has a *unique* solution [31] under the absolute continuity of $C(t)$, it does not make any sense to formulate optimization problems for the basic Moreau sweeping process. This is a striking difference between the discontinuous differential inclusion (1.1) and the ones $\dot{x}(t) \in F(x(t))$ described by *Lipschitzian* set-valued mappings/multifunctions $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ for which optimal control theory has been well developed; see, e.g., the books [11,28,39] for various methods and results on necessary optimality conditions.

It seems that optimal control problems for sweeping differential inclusions were first formulated and studied in the case of control actions entering *additive perturbations* [21] for which existence and relaxation results, while not optimality conditions, were obtained; see [10,36,37] for subsequent developments in this direction. To the best of our knowledge, the theory of necessary optimality conditions for sweeping processes has been started with [12], where a new class of dynamic optimization problems with *controlled moving sets* $C(t) = C(u(t))$ in (1.1) was first formulated with deriving necessary optimality conditions in the case when $C(u)$ is a half-space. Soon after that, necessary optimality conditions were obtained for another class of sweeping process without controlled in either moving sets or perturbations, but in a coupling linear ODE.

Further necessary optimality conditions and their applications for all the three types of controlled sweeping processes were developed in [2,3,6–9,13–16,19,22].

This paper concerns the following class of optimal control problems of the *generalized Bolza type* for the perturbed version of the sweeping process in (1.1). Given an extended real-valued terminal cost function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}} := (-\infty, \infty]$ and a running cost function $\ell\colon [0,T] \times \mathbb{R}^{2(n+nm+m)+d} \to \overline{\mathbb{R}}$, our basic problem (P) is defined by:

$$\text{minimize} \quad J[x,a,b,u] := \varphi\big(x(T)\big) + \int_0^T \ell\big(t, x(t), a(t), b(t), u(t), \dot{x}(t), \dot{a}(t), \dot{b}(t)\big)dt \tag{1.3}$$

over control actions $a(\cdot) = (a_1(\cdot), \ldots, a_m(\cdot)) \in W^{1,2}([0,T]; \mathbb{R}^{mn})$ and $b(\cdot) = (b_1(\cdot), \ldots, b_m(\cdot)) \in W^{1,2}([0,T]; \mathbb{R}^m)$ entering the moving set $C(t)$ and measurable controls $u(\cdot) \in L^2([0,T]; \mathbb{R}^d)$ entering additive perturbations that generate the corresponding trajectories $x(\cdot) \in W^{1,2}([0,T]; \mathbb{R}^n)$ of the sweeping differential inclusion

$$\begin{cases} \dot{x}(t) \in -N\big(x(t); C(t)\big) + g\big(x(t), u(t)\big) \ \text{ a.e. } \ t \in [0,T], \\ x(0) := x_0 \in C(0), \ u(t) \in U \ \text{ a.e. } \ t \in [0,T], \end{cases} \tag{1.4}$$

where the moving set is given in the polyhedral form as

$$C(t) := \Big\{ x \in \mathbb{R}^n \Big|\ \langle a_i(t), x \rangle \le b_i(t),\ i = 1, \ldots, m \Big\}, \tag{1.5}$$

and where the initial point $x_0 \in \mathbb{R}^n$ and the final time $T > 0$ are fixed. All such quadruples $(x(\cdot), a(\cdot), b(\cdot), u(\cdot))$ for which the running cost $\ell(\cdot)$ is integrable are *feasible solutions* to problem (P).

In addition to the above dynamical system (1.4) with the *pointwise/hard constraints* on the controls $u(\cdot)$ in perturbations, we impose the pointwise constraints on the controls $a_i(\cdot)$ in the moving set:

$$\|a_i(t)\| = 1 \ \text{ for all } \ t \in [0,T] \ \text{ and } \ i = 1, \ldots, m. \tag{1.6}$$

Furthermore, problem (P) also contains the implicit pointwise *mixed state-control constraints*

$$\langle a_i(t), x(t) \rangle \le b_i(t) \ \text{ for all } \ t \in [0,T] \ \text{ and } \ i = 1, \ldots, m, \tag{1.7}$$

which are due to construction (1.2) of the normal cone in (1.4).

Our approach to the dynamic optimization problem (P) is based on the *method of discrete approximation*, which was developed in [27,28] for optimization of nonconvex and nonautonomous Lipschitzian differential inclusions; see also the references and commentaries therein. Referring to [27,28] for more details, we specially mention here the books [32] and [35], where the former addressed convex and convex-valued inclusions under restrictive assumptions, while the latter overcame these assumptions for convex-valued, autonomous, and uniformly bounded Lipschitzian inclusions. For the case of various versions of (highly non-Lipschitzian and unbounded) sweeping differential inclusions, significant modifications of the method of discrete approximations were given in [6–8,12–14,16,22] to handle various optimal control problems for sweeping processes. There are *four major steps* in the realization of this approach to the study of continuous-time systems:

(i) Firstly, we construct a *well-posed* discrete approximation of the sweeping control system from (1.1), (1.5) in such a way that *any feasible solution* to the continuous-time sweeping inclusion can be *appropriately approximated* by feasible solutions to the discretized sweeping control systems. This step may be also considered from the *numerical viewpoint* as a finite-dimensional approximation of the discontinuous constrained differential inclusion.

(ii) The second approximation step is to construct, with the usage of (i), a sequence of discrete-time optimal control problems (P_k), $k \in I\!N := \{1, 2, \ldots\}$, for discretized sweeping inclusions such that the approximating problems admit optimal solutions whose continuous-time extensions *strongly converge* as $k \to \infty$ in the requited topology to a *chosen local minimizer* of the original sweeping control problem (P).

(iii) The next step is to derive *necessary conditions* that hold for optimal solutions of each discrete-time problem (P_k), which can be reduced to a finite-dimensional format of *mathematical programming* with increasingly many *geometric constraints* of the graphical type. To deal with such problems, we employ appropriate tools of *first-order* and *second-order variational analysis* and *generalized differentiation*. Due to (ii), the obtained results can be viewed as constructive *suboptimality* (almost optimality) conditions for (P) that practically provide, for large $k \in I\!N$, about the same amount of information as the exact optimality conditions for local minimizers of (P).

(iv) The last step is highly challenging mathematically while being of undoubted importance. It furnishes the limiting procedure to pass from the necessary conditions for the optimal solutions of the discrete-time problems (P_k) obtained in (iii) to the *exact necessary optimality conditions* for the designated *local minimizer* of the original sweeping control problem (P). This step strongly involves advanced calculus and computation results of variational analysis and generalized differentiation, especially of the *second order*.

In this paper we comprehensively resolve the issues listed in steps (i)–(iii) for the general sweeping control problem (P) formulated in (1.3)–(1.7) (which is certainly of its independent interest and own importance), while step (iv) is furnished in our forthcoming paper [5]. Note that some particular cases of problem (P) were investigated by discrete approximation techniques in the papers [6,8,14,16] mentioned above, but the general setting of our consideration is significantly more complicated and thus requires careful elaborations, which are provided in this paper and subsequently in [5].

The rest of the paper is organized as follows. In Section 2 we formulate the *standing assumptions* on the given data of (P) and present preliminary results on the well-posedness of the controlled sweeping process under consideration. Section 3 establishes the *existence of optimal solutions* to (P) and discusses its relaxation stability. In Section 4 we construct a discrete approximation of the sweeping control system in (1.1), (1.5) that allows us to *strongly approximate* any *feasible solution* to it by feasible solutions to its discrete counterparts. Section 5 develops the discrete approximation procedure at the *level of optimality* while leading us to the strong convergence of optimal solutions for the discrete-time problems to the prescribed *local minimizer* of (P). In Section 6 we first review the tools of *generalized differentiation* needed for our variational analysis and then obtain *second-order calculation formulas* that are crucial for deriving necessary optimality conditions. Such conditions are obtained in Section 7 for the constructed discrete approximation problems. Finally, we illustrate in Section 8 by a nontrivial example the efficiency of the obtained optimality conditions to solve sweeping control problems.

2. Standing assumptions and preliminaries

Throughout the entire paper we use *standard notation* of variational analysis and control theory (see, e.g., [29,34,39]), except a few special symbols, which are defined where they appear.

In this section we present some results on well-posedness of the sweeping differential inclusions in the aforementioned classes of feasible controls and formulate the standing assumptions on problem (P) that allow us to establish further the main achievements of the paper.

Denoting by $d(x;\Omega)$ the distance between a given point $x\in\mathbb{R}^n$ and a nonempty set $\Omega\subset\mathbb{R}^n$, observe first that the conventional assumption on the moving set $C(t)$ ensuring the existence of absolutely continuous solutions to the sweeping differential inclusion (1.4) is formulated as follows:

$$|d\big(x;C(t)\big)-d\big(x;C(s)\big)|\le|v(t)-v(s)|\ \text{ for all }\ t,s\in[0,T],\tag{2.1}$$

where $v\colon[0,T]\to\mathbb{R}$ is an absolutely continuous function; see [17,25] and the references therein. However, assumption (2.1) is rather restrictive and may fail for polyhedral moving sets $C(t)$ as in (1.5), even in the case of half-spaces. An improvement of (2.1) ensuring the existence of absolutely continuous solutions to (1.4) was obtained in [13] with the verification of the imposed assumption in the case of half-spaces $C(t)$ in [13] and then for general convex polyhedral sets (1.5) in [14] under the linear independence constraint qualification (LICQ) meaning that the vectors $\{a_i(t)\}$ are linearly independent for all $t\in[0,T]$ along the active constraints. Following the approach of Tolstonogov [36], we derive below an advanced result on the existence and uniqueness of $W^{1,2}$ solutions to (1.4) with the polyhedral moving sets (1.5) generated by $W^{1,\infty}$ controls $(a_i(t),b_i(t))$ and measurable controls $u(t)$ under a major assumption that is significantly weaker than LICQ. This result justifies the well-posedness of the sweeping dynamical systems under consideration, which is required for the subsequent study of the optimal control problem (P).

Now we formulate the *standing assumptions* of this paper that include those ensuring the existence of the aforementioned solutions to the sweeping system (1.4) and (1.5).

(H1) The control set U from (1.4) is closed and bounded in $\mathbb{R}^d$.

(H2) The derivatives $(\dot a_i(t),\dot b_i(t))$ are uniformly bounded for all $i=1,\ldots,m$ and a.e. $t\in[0,T]$ with the fixed initial points $a_0:=(a_1(0),\ldots,a_m(0))$ and $b_0:=(b_1(0),\ldots,b_m(0))$.

(H3) The perturbation mapping $g\colon\mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}^n$ is uniformly Lipschitz continuous with respect to both variables x and $u\in U$, i.e., there exists $L>0$ for which

$$\|g(x_1,u_1)-g(x_2,u_2)\|\le L\,(\|x_1-x_2\|+\|u_1-u_2\|)\ \text{ for all }\ (x_1,u_1)\ \text{ and }\ (x_2,u_2)\in\mathbb{R}^n\times U.\tag{2.2}$$

Furthermore, g satisfies the sublinear growth condition

$$\|g(x,u)\|\le M\,(1+\|x\|)\ \text{ for all }\ u\in U\ \text{ with some }\ M>0.$$

(H4) There exists a continuous function $\vartheta\colon [0,T]\to\mathbb{R}$ for which $\sup\limits_{t\in[0,T]}\vartheta(t)<0$ and

$$C^0(t):=\Big\{x\in\mathbb{R}^n\Big|\ \langle a_i(t),x\rangle-b_i(t)<\vartheta(t),\ i=1,\dots,m\Big\}\neq\emptyset\ \text{ for all }\ t\in[0,T]. \tag{2.3}$$

(H5) The terminal cost $\varphi\colon\mathbb{R}^n\to\overline{\mathbb{R}}$ is lower semicontinuous (l.s.c.), while the running cost/integrand $\ell(t,\cdot)\colon\mathbb{R}^{2(n+nm+m)+d}\to\bar{\mathbb{R}}$ is bounded from below and l.s.c. for a.e. $t\in[0,T]$ around the reference feasible solution to (P). Further, ℓ is a.e. continuous in t and is uniformly majorized by a summable function on $[0,T]$.

Before presenting the aforementioned well-posedness (existence and uniqueness) theorem for the sweeping process in (1.4) and (1.5), we discuss the imposed condition (2.3) in (H4). Recall that the *positive linear independence constraint qualification* (PLICQ) condition holds at $x\in C(t)$ if

$$\left[\sum_{i\in I(x,a(t),b(t))}\alpha_i a_i(t)=0,\ \alpha_i\ge 0\right]\Longrightarrow\Big[\alpha_i=0\ \text{ for all }\ i\in I\big(x,a(t),b(t)\big)\Big], \tag{2.4}$$

where the set of *active constraint indices* for (1.5) is defined by

$$I\big(x,a(t),b(t)\big):=\Big\{i\in\{1,\dots,m\}\Big|\ \langle a_i(t),x\rangle=b_i(t)\Big\},\quad t\in[0,T]. \tag{2.5}$$

The essentially more restrictive *linear independence constraint qualification* (LICQ) condition at $x\in C(t)$ used in [14] reads as (2.4) with the replacement of $\alpha_i\ge 0$ by $\alpha_i\in\mathbb{R}$ therein.

It is easy to see the Slater-type condition (2.3) reduces to PLICQ if the polyhedron (1.5) does not depend on t, which is the case considered in [16]. In the general nonautonomous case, (2.3) may be stronger than PLICQ (2.4) while being always weaker than its LICQ counterpart. Note also that in our setting, (2.4) corresponds to the *Mangasarian-Fromovitz constraint qualification*, which is classical in nonlinear programming. Furthermore, imposing PLICQ at $x\in C(t)$ is equivalent to the so-called *inverse triangle inequality* at this point defined by

$$\sum_{i\in I(x,a(t),b(t))}\lambda_i\,\|a_i(t)\|\le\gamma\left\|\sum_{i\in I(x,a(t),b(t))}\lambda_i a_i(t)\right\|\ \text{ for all }\lambda_i\ge 0 \tag{2.6}$$

with some constant $\gamma>0$ depending on t; see [38] for more discussions.

We claim now that γ can be chosen *uniformly* on $[0,T]$, i.e., that there exists a constant $\gamma>0$ such that (2.6) holds for all $t\in[0,T]$ simultaneously. Indeed, assuming the contrary gives us sequences $\{t^k\}$ from $[0,T]$ and of nonnegative numbers $\{\lambda_i^k\}$ such that

$$\sum_{i\in I(a(t^k),b(t^k))}\lambda_i^k\|a_i(t^k)\|\ge k\left\|\sum_{i\in I(a(t^k),b(t^k))}\lambda_i^k a_i(t^k)\right\|\ \text{ for all }\ k\in I\!N. \tag{2.7}$$

Since the above inequality is positively homogeneous of degree one with respect to the variables λ_i, we suppose without loss of generality that for each k the largest among numbers λ_i^k is 1. It easily follows from (2.7) and the structures of the active index sets (2.5) generated by the continuous functions $a_i(t)$ and $b_i(t)$ that there exist subsequences (with no relabeling) $t_k \to \bar{t}$ and $\lambda_i^k \to \bar{\lambda}_i \geq 0$, not all zero, such that

$$\Big\| \sum_{i \in I(a(\bar{t}), b(\bar{t}))} \bar{\lambda}_i a_i(\bar{t}) \Big\| = 0,$$

which clearly contradicts the imposed PLICQ and thus verifies the claim.

Now we are ready to present the aforementioned *well-posedness* result for the sweeping system (1.4), (1.5).

Theorem 2.1 *(well-posedness of the controlled sweeping process). Let all the assumptions in* (H1)–(H4) *be satisfied, and let* $(a(\cdot), b(\cdot)) \in W^{1,2}([0,T]; \mathbb{R}^{mn} \times \mathbb{R}^m)$ *and* $u(\cdot) \in L^2([0,T]; \mathbb{R}^d)$ *be fixed control actions in* (1.4) *and* (1.5)*. Then the sweeping differential inclusion* (1.4) *admits the unique solution* $x(\cdot) \in W^{1,2}([0,T]; \mathbb{R}^n)$ *generated by the control triple* $(a(\cdot), b(\cdot), u(\cdot))$.

Proof. Following [36], it is said that a set-valued mapping $C\colon [0,T] \rightrightarrows \mathbb{R}^n$ is *r-uniformly lower semicontinuous from the right*, if there exists a function $v_r(\cdot) \in W^{1,2}([0,T]; \mathbb{R}^n)$ such that for any $s, t \in [0,T]$ with $s \leq t$, and any $x \in \mathbb{R}^n$ with $\|x\| \leq r$ we have the inequality

$$d\big(x; C(t)\big) \leq d\big(x; C(s)\big) + |v_r(t) - v_r(s)|.$$

Let us show that assumption (H4) implies that the polyhedral mapping $C(\cdot)$ defined in (1.5) is r-uniformly lower semicontinuous from the right. To proceed, define the function $\phi\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}$ by

$$\phi(t,x) := \max_{1 \leq i \leq m} \big\{ \langle a_i(t), x \rangle - b_i(t) \big\}, \quad t \in [0,T],\ x \in \mathbb{R}^n, \tag{2.8}$$

which gives us the representation $C(t) = \{x \in \mathbb{R}^n \,|\, \phi(t,x) \leq 0\}$ of the set $C(t)$ from (1.5) for each $x \in \mathbb{R}^n$. Let us show that function (2.8) satisfies the hypothesis $H(\phi)$ formulated in [36, p. 297]. Indeed, the convexity of $x \mapsto \phi(t,x)$ and estimates in [36, (4.2)] imposed in $H(\phi)$ follow directly from the construction of ϕ. Furthermore, we deduce from (2.3) that the required condition [36, (2)] in $H(\phi)$ is also satisfied. To verify $H(\phi)$, it remains checking the validity of [36, (4.1)]. Since $b_i(\cdot) \in W^{1,2}([0,T]; \mathbb{R})$, we clearly have that $\max\limits_{1 \leq i \leq m} \max\limits_{t \in [0,T]} |b_i(t)| < \infty$. Moreover, it follows from (2.8) for all $x \in \mathbb{R}^n$ for which $\|x\| \leq r$ and all $t, s \in [0,T]$ that

$$|\phi(t,x) - \phi(s,x)| \leq \max_{1 \leq i \leq m} \|a_i(t) - a_i(s)\| \cdot \|x\| + \max_{1 \leq i \leq m} |b_i(t) - b_i(s)|.$$

The assumption (H2) implies $a_i(\cdot)$ is Lipschitz continuous on $[0,T]$ with Lipschitz constant L_i^a. As a consequence,

$$|\phi(t,x) - \phi(s,x)| \leq \max_{1 \leq i \leq m} L_i^a |t - s| r + \max_{1 \leq i \leq m} |b_i(t) - b_i(s)|.$$

Denote further

$$\xi_r(t) := \int_0^t \left(r \max_{1\le i\le m} L_i^a + \max_{1\le i\le m} |\dot{b}_i(\tau)| \right) d\tau, \quad r \ge 0.$$

Then $\xi_r(\cdot) \in W^{1,2}([0,T];\mathbb{R})$ for all $r \ge 0$ and

$$|\phi(t,x) - \phi(s,x)| \le |\xi_r(t) - \xi_r(s)| \;\text{ whenever }\; \|x\| \le r,\; t,s \in [0,T],$$

which completes the verification of all the assumptions in $H(\phi)$ of [36]. Employing now [36, Theorem 4.1] verifies that our polyhedral mapping $C(\cdot)$ is r-uniformly lower semicontinuous from the right on $[0,T]$. Finally, the existence and uniqueness result claimed in the theorem follow from [36, Lemma 3.1 and Theorem 4.1]. $\square$

3. Existence of optimal solutions and relaxation

This section addresses the existence issue for (global) *optimal solutions* to the sweeping control problem (P). Then we define an appropriate notion of *local minimizers* to (P) and discuss its *relaxed* counterpart.

Before establishing the existence of optimal solutions to (P) in the aforementioned class of feasible solutions, let us reformulate the sweeping differential inclusion (1.4) in a more convenient way. Consider the image of the control set U under the perturbation mapping $g: \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ defined by

$$g(x,U) := \left\{ v \in \mathbb{R}^m \,\middle|\, v = g(x,u) \;\text{ for some }\; u \in U \right\},\; x \in \mathbb{R}^n.$$

Then the sweeping inclusion (1.4) with the moving set (1.5) is equivalently represented as

$$-\dot{x}(t) \in N\big(x(t); C(t)\big) - g\big(x(t), U\big) \;\text{ a.e. } t \in [0,T],\; x_0 \in C(0). \tag{3.1}$$

To elaborate more rigorously upon this statement, we need to use standard facts of the theory of measurable multifunctions; see, e.g., [34, Chapter 14]. Recall that the measurability of a closed-valued multifunction $S\colon [0,T] \rightrightarrows \mathbb{R}^n$ can be described as follows (see [34, Theorem 14.10(b)]): For every $\varepsilon > 0$ there is a closed set $T_\varepsilon \subset [0,T]$ with $\text{mes}([0,T] \setminus T_\varepsilon) < \varepsilon$ such that $S\colon T_\varepsilon \rightrightarrows \mathbb{R}^n$ is of closed graph. Fix any $x(\cdot) \in W^{1,2}([0,T];\mathbb{R}^n)$ satisfying (3.1) with some $(a(\cdot), b(\cdot))$ from (1.5) and define the closed-valued mapping $S\colon [0,T] \rightrightarrows \mathbb{R}^n$ by

$$S(t) := \left\{ u \in U \,\middle|\, -\dot{x}(t) \in N\big(x(t); C(t)\big) - g\big(x(t), u\big) \right\} \;\text{ a.e. }\; t \in [0,T]. \tag{3.2}$$

Applying to $-\dot{x}(\cdot)$ the classical Luzin property of measurable functions in real analysis, for any $\varepsilon > 0$ we find a closed set $T_\varepsilon \subset [0,T]$ with $\text{mes}([0,T] \setminus T_\varepsilon) < \varepsilon$ for its Lebesgue measure such that $-\dot{x}(\cdot)$ is continuous on T_ε. Using the assumed continuity of $g(x,u)$ together with the closed-graph property of the normal cone mapping in (3.2) with $C(t)$ taken from (1.5) shows that the graph of the mapping $S\colon T_\varepsilon \rightrightarrows \mathbb{R}^n$ from (3.2) is closed. It tells us that the full mapping $S(\cdot)$ defined in (3.2) for a.e. $t \in [0,T]$ is measurable on $[0,T]$. Employing the measurable selection theorem from [34, Corollary 14.6] ensures the existence of a measurable control $u(t) \in U$ such that the pair $(x(\cdot), u(\cdot))$ together with the corresponding $(a(\cdot), b(\cdot))$ generated the moving set

$C(t)$ in (1.5) is feasible to (1.4). The converse implication from (1.4) to (3.1) is obvious, and hence we verify the claimed equivalence.

Now we are ready to obtain the existence theorem for optimal solutions to (P) under certain additional *convexity* assumptions with respect to velocities. For simplicity we suppose here that the integrand ℓ does not depend on the control variable u. If it does, we have to impose the convexity of an extended velocity set that includes the integrand component.

Theorem 3.1 *(existence of optimal solutions to controlled sweeping processes). Let (P) be the optimal control problem formulated in Section 1 with the equivalent form (3.1) of the sweeping differential inclusion over all the $W^{1,2}([0,T];\mathbb{R}^n)\times W^{1,2}([0,T];\mathbb{R}^{mn})\times W^{1,2}([0,T];\mathbb{R}^m)\times L^2([0,T];\mathbb{R}^d)$ quadruples $(x(\cdot),a(\cdot),b(\cdot),u(\cdot))$. In addition to the standing assumptions* (H1)–(H5)*, suppose that the integrand ℓ in (1.3) does not depend on the u-variable while being convex with respect to the velocity variables $(\dot x,\dot a,\dot b)$. Suppose furthermore that along a minimizing sequence of $\big(x^k(\cdot),a^k(\cdot),b^k(\cdot),u^k(\cdot)\big)$ as $k\in I\!N$ we have that $\ell(t,\cdot)$ is majorized by a summable function, that $\{(x^k(\cdot),a^k(\cdot),b^k(\cdot))\}$ is bounded in $W^{1,2}([0,T];\mathbb{R}^n\times\mathbb{R}^{mn}\times\mathbb{R}^m)$, and that the set $g(x^k(t);U)$ is convex for all $t\in[0,T]$. Then (P) admits an optimal solution in $W^{1,2}([0,T];\mathbb{R}^{n+mn+m})\times L^2([0,T];\mathbb{R}^d)$.*

Proof. Since the set of feasible solutions to problem (P) is nonempty by Theorem 2.1, we can take the minimizing sequence of quadruples $(x^k(\cdot),a^k(\cdot),b^k(\cdot),u^k(\cdot))$ in (P) from the formulation of the theorem. It follows from the boundedness of $\{x^k(\cdot),(a^k(\cdot),b^k(\cdot))\}$ in $W^{1,2}([0,T];\mathbb{R}^n\times\mathbb{R}^{mn}\times\mathbb{R}^m)$ and the weak compactness of the dual ball in $L^2([0,T];\mathbb{R}^n\times\mathbb{R}^{mn}\times\mathbb{R}^m)$ that $\dot x^k(\cdot)\to v^x(\cdot)$, $\dot a^k(\cdot)\to v^a(\cdot)$, and $\dot b^k(\cdot)\to v^b(\cdot)$ weakly in $L^2([0,T];\mathbb{R}^n)$, $L^2([0,T];\mathbb{R}^{mn})$, and $L^2([0,T];\mathbb{R}^m)$ along subsequences (without relabeling) for some functions $v^x(\cdot)$, $v^a(\cdot)$, and $v^b(\cdot)$ from the corresponding spaces. Employing Mazur's weak closure theorem, we conclude that there are sequences of convex combinations of $\dot x^k(\cdot)$, $\dot a^k(\cdot)$, and $\dot b^k(\cdot)$, which strongly converge in the corresponding spaces to $v^x(\cdot)$, $v^a(\cdot)$, and $v^b(\cdot)$, respectively. Furthermore, standard real analysis tells us that there exists a subsequence of these convex combinations (no relabeling again), which converges to $(v^x(\cdot),v^a(\cdot),v^b(\cdot)$ as $k\to\infty$ a.e. pointwise on $[0,T]$. Define now $\bar x(\cdot)\in W^{1,2}([0,T];\mathbb{R}^n)$, $\bar a(\cdot)\in W^{1,2}([0,T];\mathbb{R}^{mn})$, and $\bar b(\cdot)\in W^{1,2}([0,T];\mathbb{R}^m)$ by

$$\bar x(t):=x_0+\int_0^t v^a(s)ds,\ \bar a(t):=a_0+\int_0^t v^a(s)ds,\ \text{ and } \bar b(t):=b_0+\int_0^t v^b(s)ds,\quad t\in[0,T],$$

and observe that they satisfy the pointwise constraints in (1.6) and (1.7). Furthermore, it follows from the closedness and convexity of the normal cone (1.2) to the moving convex polyhedral set $C(t)$ in (1.5) and the assumed convexity of the compact sets $g(x^k(t),U)$ on $[0,T]$ that the right-hand site velocity set in (3.6) is convex along the selected minimizing sequence, and we have

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

for the limiting trajectory $\bar x(\cdot)$ with $\bar x(t)\in\overline{C}(t):=\{x\in\mathbb{R}^n|\ \langle\bar a_i(t),x\rangle\le\bar b_i(t),\ i=1,\dots,m\}$ on $[0,T]$. Employing now the aforementioned measurable selection allows us to find a measurable control $\bar u(\cdot)$ such that $\bar u(t)\in U$ and

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

It remains to show that the limiting quadruple $(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot))$, which is proved to be feasible for (P), is an optimal solution to this problem. This is a consequence of the inequality

$$J[\bar{x}, \bar{a}, \bar{b}, \bar{u}] \le \liminf_{k\to\infty} J[x^k, a^k, b^k, u^k] \tag{3.3}$$

for the cost functional (1.3). To verify (3.3), we use the assumptions in (H5) ensuring the application of the Lebesgue dominated convergence theorem together with the imposed convexity of integrand with respect to $(\dot{x}, \dot{a}, \dot{b})$. This allows us to apply the classical lower semicontinuity result for integral functionals with respect to the weak topology in L^2. Observe that there is no need to care about the convergence with respect to u-controls in our setting due to the independence of the integrand ℓ on the u-component. Thus the proof is complete. □

Justifying the existence of global optimal solutions to the controlled sweeping process under (P), recall that our goal is the derivation of necessary optimality conditions for suitable *local* minimizers of (P) by employing the method of discrete approximations. An appropriate concept from this viewpoint goes back to *intermediate* local minimizers introduced in [27] for Lipschitzian differential inclusions that occupies an intermediate position between the conventional notions of weak and strong minimizers in dynamic optimization while covering the latter; see the books [28,39] and the references therein for more details on this notion for Lipschitzian inclusions. In the case of our problem (P), a natural implementation of this concept reads as follows.

Definition 3.2 *(intermediate local minimizers for sweeping optimal control).* Let $(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot))$ be a feasible solution to problem (P) under the standing assumptions made. We say that $(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot))$ is an INTERMEDIATE LOCAL MINIMIZER (i.l.m.) for (P) if $(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot)) \in W^{1,2}([0, T]; \mathbb{R}^n) \times W^{1,2}([0, T]; \mathbb{R}^{mn}) \times W^{1,2}([0, T]; \mathbb{R}^m) \times L^2([0, T]; \mathbb{R}^d)$ and there exists $\varepsilon > 0$ such that

$$J[\bar{x}, \bar{a}, \bar{b}, \bar{u}] \le J[x, a, b, u]$$

for any feasible solutions $(x(\cdot), a(\cdot), b(\cdot), u(\cdot))$ to (P) satisfying

$$\big\|x(\cdot) - \bar{x}(\cdot)\big\|_{W^{1,2}} + \big\|\big(a(\cdot), b(\cdot)\big) - \big(\bar{a}(\cdot), \bar{b}(\cdot)\big)\big\|_{W^{1,2}} + \|u(\cdot) - \bar{u}(\cdot)\|_{L^2} \le \varepsilon. \tag{3.4}$$

If the term $\|x(\cdot) - \bar{x}(\cdot)\|_{W^{1,2}}$ in (3.4) is replaced by $\|x(\cdot) - \bar{x}(\cdot)\|_{\mathcal{C}}$, the norm in the space of continuous functions $\mathcal{C}([0, T]; \mathbb{R}^n)$, we speak about *strong local minimizers* for (P). It is clear that any strong local minimizer for (P) is an intermediate one, but not vice versa as can be confirmed by examples.

To implement our approach to study local minimizers of (P), we need a certain *relaxation stability* of the i.l.m. under consideration. The idea of relaxation of variational problems, related to convexification with respect to derivative variables, goes back to Bogolyubov and Young for the classical calculus of variations and to Gamkrelidze and Warga for optimal control problems governed by ordinary differential equations; see, e.g., the books [28,39] for more discussions and references, where relaxation of control problems for Lipschitzian differential inclusions were also investigated and discussed in detail. Relaxation results for non-Lipschitzian differential inclusions were more recently developed in [20,21,37].

To proceed in the case of our optimal control problem (P), consider vectors $x := (x_1, \ldots, x_n) \in \mathbb{R}^n$, $a := (a_1, \ldots, a_m) \in \mathbb{R}^{mn}$, $b := (b_1, \ldots, b_m) \in \mathbb{R}^m$, and $u := (u_1, \ldots, u_d) \in \mathbb{R}^d$, and then define the set-valued mapping $F\colon \mathbb{R}^n \times \mathbb{R}^{mn} \times \mathbb{R}^m \times \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ by

$$F(x,a,b,u) := N\big(x; C(a,b)\big) - g(x,u), \tag{3.5}$$

where $N(x; C(a,b))$ is taken from in (1.2), and where $C(a,b) := \{x \in \mathbb{R}^n \mid \langle a_i, x\rangle \le b_i,\ i = 1, \ldots, m\}$. It is not hard to see that F admits the following explicit representation:

$$F(x,a,b,u) = \Big\{ \sum_{i \in I(x,a,b)} \eta_i a_i \Big|\ \eta_i \ge 0 \Big\} - g(x,u) \tag{3.6}$$

via the active index set (2.5) at $x \in C(a,b)$. Let $\ell_F(t,a,b,u,\dot{x},\dot{a},\dot{b})$ be the restriction of the integral ℓ on the set $F(x,a,b,u)$ with $\ell_F(t,a,b,u,\dot{x},\dot{a},\dot{b}) := \infty$ if $\dot{x} \notin F(x,a,b,u)$. Denoting by $\widehat{\ell}_F$ the *convexification* of the integrand (i.e., the largest l.s.c. convex function majorized by $\ell(t,x,a,b,\cdot,\cdot,\cdot,\cdot)$) with respect to the velocity variables $(\dot{x},\dot{a},\dot{b})$ as well as to the control u on the convex hull $\operatorname{co} U$, define the *relaxed optimal control problem* (R) by:

$$\text{minimize}\quad \widehat{J}[x,a,b,u] := \varphi\big(x(T)\big) + \int_0^T \widehat{\ell}_F\big(t, x(t), a(t), b(t), u(t), \dot{x}(t), \dot{a}(t), \dot{b}(t)\big)dt \tag{3.7}$$

over quadruples $(x(\cdot), a(\cdot), b(\cdot), u(\cdot)) \in W^{1,2}([0,T]; \mathbb{R}^n) \times W^{1,2}([0,T]; \mathbb{R}^{mn}) \times W^{1,2}([0,T]; \mathbb{R}^m) \times L^2([0,T]; \mathbb{R}^d)$ satisfying (1.6) and giving a finite value of the extended running cost in (3.7). All such quadruples are said to be *feasible* to (R). It follows from (3.7) and the construction of $\widehat{\ell}_F$ with F taken from (3.5) that $u(t) \in \operatorname{co} U$ for a.e. $t \in [0,T]$, and that $x(\cdot)$ is a trajectory of the convexified differential inclusion

$$-\dot{x}(t) \in N\big(x(t); C(t)\big) - \operatorname{co} g\big(x(t), U\big) \ \text{ a.e. } t \in [0,T],\ x_0 \in C(0) \tag{3.8}$$

with $\langle a_i(t), x(t)\rangle \le b_i(t)$ for $i = 1, \ldots, m$ and all $t \in [0,T]$. The precise justification of this is similar to that given above for (3.1) being based on measurable selections.

Now we introduce a new notion of *relaxed* intermediate local minimizers for (P); cf. [27] for Lipschitzian differential inclusions and [16] for a version of problem (P) with $\ell \equiv 0$ and an uncontrolled polyhedron $C(t) \equiv C$.

Definition 3.3 *(relaxed intermediate local minimizers).* We say that $(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot))$ is a RELAXED INTERMEDIATE LOCAL MINIMIZER (r.i.l.m.) for problem (P) if it is feasible for (P) and there exists $\varepsilon > 0$ such that

$$J[\bar{x}, \bar{a}, \bar{b}, \bar{u}] = \widehat{J}[\bar{x}, \bar{a}, \bar{b}, \bar{u}] \le \widehat{J}[x,a,b,u]$$

whenever a feasible quadruple $(x(\cdot), a(\cdot), b(\cdot), u(\cdot))$ for (R) satisfies (3.4).

It follows from Definitions 3.2 and 3.3 in view of the constructions in (3.7) and (3.8) that any i.l.m. of (P) is also its r.i.l.m. provided that the sets U and $g(x(t);U)$ are convex and the integrand $\ell(t,x(t),a(t),b(t),\cdot,\cdot,\cdot,\cdot)$ is convex along feasible solutions to (P). The well-recognized beauty of relaxation procedures in variational and control problems is that they keep global or local optimal values of cost functionals under relaxation in important situations *without any convexity assumptions*. It is strongly related to deep measure-theoretical results of the Lyapunov-Aumann type ensuring the automatic convexity of integrals of arbitrary set-valued mappings over nonatomic measures. In particular, it has been realized in this way that every strong local minimizer in control problems for Lipschitzian differential inclusions with no constraint of right ends of trajectories is always a relaxed one; see, e.g., [28,39]. Similar results for controlled sweeping processes of different types were obtained in [21, Theorem 2] and [37, Theorem 4.2]. We *conjecture* that modifying the proofs of the aforementioned theorems lead us to the fact that any *strong* local minimizer of the nonconvex sweeping control problem (P) is a relaxed strong local minimizer of this problem under the imposed standing assumptions in (H1)–(H5) with the replacement of the lower semicontinuity of φ and ℓ in (H5) by their continuity.

4. Strong discrete approximation of feasible solutions

In this section we start our detailed development of the method of discrete approximations to study the sweeping optimal control problem (P) formulated in Section 1. In fact, this section does not concern the optimization part of (P) while dealing only with constructive approximations of *feasible solutions*. Our main goal here is to show that the standing assumptions imposed allow us to *strongly* approximate *any* feasible solution to (P) by feasible solutions to discrete-time problems extended to the continuous-time interval. The result established below significantly improves similar ones obtained in [6,8,14] for particular types of sweeping control problems, and so its proof is more involved in comparison with those given in [6,8,14]. Note that another discrete approximation scheme was developed in [16] for problem (P) with $\ell \equiv 0$ and an uncontrolled polyhedral convex set $C(t) \equiv C$.

To proceed, for each $k \in \mathbb{N}$ define the discrete partition of $[0,T]$ by

$$\begin{aligned} &\Delta_k := \left\{ 0 = t_0^k < t_1^k < \ldots < t_{\nu(k)-1}^k < t_{\nu(k)}^k = T \right\} \\ &\text{with } \ h_j^k := t_{j+1}^k - t_j^k \le \frac{\widetilde{\nu}}{\nu(k)} \ \text{ for } \ j = 0,\ldots,\nu(k)-1, \end{aligned} \tag{4.1}$$

where $\nu = \nu(k) \ge k$, and where $\widetilde{\nu} > 0$ is some constant.

Here is a major approximation result, which is of its own interest (also from a numerical viewpoint), while being important for the further developments of this paper and its continuation in [5]. In this and subsequent theorems we add to our assumptions the bounded variation requirement on the reference control and velocities, which is a natural and not restrictive assumption to get the strong convergence of discrete approximations. Recall that the mapping F used below is defined in (3.5).

Theorem 4.1 *(strong discrete approximation of feasible sweeping solutions). Under the standing assumptions in* (H1)–(H4)*, fix any feasible solution* $(\bar{x}(\cdot),\bar{a}(\cdot),\bar{b}(\cdot),\bar{u}(\cdot))$ *to* (P) *such that the functions* $\dot{\bar{x}}(\cdot),\dot{\bar{a}}(\cdot),\dot{\bar{b}}(\cdot)$ *and* $\bar{u}(\cdot)$ *are of bounded variation on* $[0,T]$*, i.e.,*

$$\max\left\{\mathrm{var}\big(\dot{\bar{x}}(\cdot);[0,T]\big),\mathrm{var}\big(\dot{\bar{a}}(\cdot);[0,T]\big),\mathrm{var}\big(\dot{\bar{b}}(\cdot);[0,T]\big),\mathrm{var}\big(\bar{u}(\cdot);[0,T]\big)\right\}\le K \tag{4.2}$$

for some constant $K>0$. Then there exist partitions Δ_k, $k\in I\!N$, together with sequences of piecewise linear functions $(x^k(t),a^k(t),b^k(t))$ and piecewise constant functions $u^k(\cdot)$ on $[0,T]$, as well as a sequence of positive numbers $\delta_k\downarrow 0$ such that $(x^k(0),a^k(0),b^k(0))=(x_0,a_0,b_0)$ for all $k\in I\!N$, and we have the relationships:

$$1-\delta_k\le\left\|a_i^k(t_j^k)\right\|\le 1+\delta_k\ \textit{for all}\ t_j^k\in\Delta_k,\quad i=1,\ldots,m, \tag{4.3}$$

$$x^k(t)=x^k(t_j^k)+(t-t_j^k)v_j^k,\ \ t_j^k\le t\le t_{j+1}^k\ \textit{with}\ -v_j^k\in F\big(x^k(t_j^k),a^k(t_j^k),b^k(t_j^k),u^k(t_j^k)\big)$$

for $j=0,\ldots,\nu(k)-1$ together with the convergence $\{(x^k(\cdot),a^k(\cdot),b^k(\cdot))\}\to(\bar{x}(\cdot),\bar{a}(\cdot),\bar{b}(\cdot))$ in the $W^{1,2}$-norm topology on $[0,T]$, and $\{u^k(\cdot)\}\to\bar{u}(\cdot)$ in the L^2-norm topology on $[0,T]$ as $k\to\infty$.

Proof. We split the proof into the following six major steps. To begin with the proof, observe that due to (4.2) the functions $\dot{\bar{x}}$, $\dot{\bar{a}}$, $\dot{\bar{b}}$ and $\bar{u}$ are defined everywhere.

Step 1: *Constructing $\big(u^k(\cdot),a^k(\cdot)\big)$ to approximate $(\bar{u}(\cdot),\bar{a}(\cdot))$.* Since step functions are dense in $L^2[0,T]$, there are sequences of step functions $\{u^k(\cdot)\}=\{\big(u_1^k(\cdot),\ldots,u_d^k(\cdot)\big)\}$ and $\{\alpha^k(\cdot)\}=\{\big(\alpha_1^k(\cdot),\ldots,\alpha_m^k(\cdot)\big)\}$ with

$$\mu_k:=\max\left\{\int_0^T\left\|u^k(t)-\bar{u}(t)\right\|^2dt,\int_0^T\left\|\alpha^k(t)-\dot{\bar{a}}(t)\right\|^2dt\right\}\to 0\ \text{as}\ k\to\infty. \tag{4.4}$$

Furthermore, for each $k\in I\!N$ we find a partition Δ_k of the interval $[0,T]$ from (4.1) for which the step functions $\{u^k(\cdot)\}$ and $\{\alpha^k(\cdot)\}$ are constant on $[t_j,t_{j+1})$ for $j=0,\ldots,\nu(k)-1$. This gives us the strong convergence of $\{\big(u^k(\cdot),\alpha^k(\cdot)\big)\}$ to $(\bar{u}(\cdot),\alpha(\cdot))$ in $L^2([0,T])$, where $\alpha(t):=\dot{\bar{a}}(t)$ for a.e. $t\in[0,T]$. Since the intervals (t_j,t_{j+1}) are not prescribed a priori and since $\bar{x}(\cdot)$ is a Carathéodory solution to (3.1), up to possibly increasing the number of intervals of the partition, we suppose without loss of generality that the differential inclusion (3.1) is satisfied at all endpoints of Δ_k that are contained in the open interval $(0,T)$. Next we define the functions $a^k(\cdot)$ by

$$a^k(t):=a_0+\int_0^t\alpha^k(s)ds,\quad t\in[0,T]. \tag{4.5}$$

It tells us that each $a^k(\cdot)$ is piecewise linear on $[0,T]$, since its derivative $\dot{a}^k(\cdot)=\alpha^k(\cdot)$ is piecewise constant on $[0,T]$. By (4.4) we have the strong convergence in $L^2([0,T])$ of $\{\dot{a}^k(\cdot)\}$ to $\dot{\bar{a}}(\cdot)$. Moreover, it follows from (4.5) and the classical Hölder inequality that

$$\left|a_{ip}^k(t)-\bar{a}_{ip}(t)\right|^2=\left|\int_0^t\left[\alpha_{ip}^k(s)-\bar{\alpha}_{ip}(s)\right]ds\right|^2\le\left[\int_0^T\left|\alpha_{ip}^k(s)-\bar{\alpha}_{ip}(s)\right|^2ds\right]T\le\mu_kT \tag{4.6}$$

for all $t \in [0, T]$, $i = 1, \dots, m$, and each component index $p = 1, \dots, n$. Hence the sequence of functions $a^k(\cdot)$ converges strongly to $\bar{a}(\cdot)$ in $W^{1,2}([0, T])$ and satisfies the estimates in (4.3) with

$$\delta_k := \sqrt{n \mu_k T}. \tag{4.7}$$

Step 2: *Constructing* $\big(x^k(\cdot), b^k(\cdot)\big)$ *to approximate* $\big(\bar{x}(\cdot), \bar{b}(\cdot)\big)$. While proceeding recurrently, fix any $j \in \{0, \dots, \nu(k) - 1\}$, suppose that the pairs (x_j^k, b_j^k) are known for all indexes $0, \dots, j$, and then construct the pair (x_{j+1}^k, b_{j+1}^k). Define the numbers

$$b_{ij}^k := \left\langle a_i^k(t_j), x_j^k \right\rangle + \bar{b}_i(t_j) - \langle \bar{a}_i(t_j), \bar{x}(t_j) \rangle \ \text{ for all } \ i = 1, \dots, m, \tag{4.8}$$

$$b_i^k(0) := b_{i0}, \ \text{ and } \ b_i^k(t) := b_{ij}^k + \frac{t - t_j}{h_j^k} \left(b_{i,j+1}^k - b_{ij}^k \right) \ \text{ for all } \ t \in [t_j, t_{j+1}] \ \text{ and } \ i = 1, \dots, m. \tag{4.9}$$

It gives us $b_{ij}^k - \left\langle a_i^k(t_j), x_j^k \right\rangle = \bar{b}_i(t_j) - \langle \bar{a}_i(t_j), \bar{x}(t_j) \rangle$, and hence

$$I(x_j^k, a^k(t_j), b_j^k) = I\big(\bar{x}(t_j), \bar{a}(t_j), \bar{b}(t_j)\big) \ \text{ for all } \ j = 0, \dots, \nu(k). \tag{4.10}$$

It follows from the fulfillment of $-\dot{\bar{x}}(t) \in F(\bar{x}(t), \bar{a}(t), \bar{b}(t), \bar{u}(t))$ for a.e. $t \in [0, T]$ including the mesh points of Δ_k with F given in (3.5), the measurability of the set-valued mapping $t \mapsto F(\bar{x}(t), \bar{a}(t), \bar{b}(t), \bar{u}(t))$ on $[0, T]$ due to [34, Theorem 14.26] with the representation of F in (3.6), and the measurable selection result from [34, Corollary 14.6] that there exist nonnegative measurable functions $\eta_i(\cdot)$ on $[0, T]$ as $i = 1, \dots, m$ ensuring the equality

$$-\dot{\bar{x}}(t) = \sum_{i \in I(\bar{x}(t), \bar{a}(t), \bar{b}(t))} \eta_i(t) \bar{a}_i(t) - g(\bar{x}(t), \bar{u}(t)) \ \text{ for a.e. } \ t \in [0, T].$$

Define now the vectors v_j^k for all indices $j = 0, \dots, \nu(k)$ by

$$\begin{aligned} -v_j^k &:= \sum_{i \in I(\bar{x}(t_j), \bar{a}(t_j), \bar{b}(t_j))} \eta_i(t_j) a_i^k(t_j) - g\big(x_j^k, u^k(t_j)\big) \\ &= \sum_{i \in I(x_j^k, a^k(t_j), b_j^k)} \eta_i(t_j) a_i^k(t_j) - g\big(x_j^k, u^k(t_j)\big), \end{aligned} \tag{4.11}$$

where the second equality comes from (4.10). Note that, since the BV functions $-\dot{\bar{x}}(t)$ and $\bar{u}(t)$ are defined everywhere on $[0, T]$, the usage of the pointwise PLICQ condition coming from (H4) justifies that $\eta_i(t_j) \geq 0$ in (4.11) are well-defined. It is obvious furthermore that $-v_j^k \in F(x_j^k, a^k(t_j), b^k(t_j), u^k(t_j))$ for such indices j. Using again the BV property of $\dot{\bar{x}}(t)$, we have by (4.2) that

$$\left\|\dot{\bar{x}}(t) - \dot{\bar{x}}(0)\right\| \leq \left\|\dot{\bar{x}}(t) - \dot{\bar{x}}(0)\right\| + \left\|\dot{\bar{x}}(T) - \dot{\bar{x}}(t)\right\| \leq \text{var}\big(\dot{\bar{x}}(\cdot); [0, T]\big) \leq K,$$

which in turn yields the estimate

$$\left\|\dot{\bar{x}}(t)\right\| \le \left\|\dot{\bar{x}}(0)\right\| + K := M_1^x$$

for a.e. $t \in [0, T]$ including the mesh points of Δ_k. Using the inverse triangle inequality (2.6) implies that

$$\begin{aligned}\eta_i(t) = \eta_i(t)\,\|\bar{a}_i(t)\| &\le \sum_{i\in I(\bar{x}(t),\bar{a}(t),\bar{b}(t))} \eta_i(t)\,\|\bar{a}_i(t)\| \le \gamma \left\| \sum_{i\in I(\bar{x}(t),\bar{a}(t),\bar{b}(t))} \eta_i(t)\bar{a}_i(t)\right\| \\ &\le \gamma\left\|\dot{\bar{x}}(t)\right\| + \gamma\left\|g\big(\bar{x}(t),\bar{u}(t)\big)\right\| \le \gamma M_1^x + \gamma M\left(1+\|\bar{x}(t)\|\right) \\ &\le \gamma M_1^x + \gamma M\left(1 + \max_{t\in[0,T]}\|\bar{x}(t)\|\right) =: M_2^x\end{aligned}$$

for a.e. $t \in [0, T]$ and for all $i \in I(\bar{x}(t), \bar{a}(t), \bar{b}(t))$, where $\gamma > 0$ can be chosen independently of $t \in [0, T]$ as proved in Section 2. By (2.2) it yields the estimates

$$\begin{aligned}\left\|v_j^k - \dot{\bar{x}}(t_j)\right\| &\le \sum_{i\in I(\bar{x}(t_j),\bar{a}(t_j),\bar{b}(t_j))} \eta_i(t_j)\left\|\bar{a}_i(t_j) - a_i^k(t_j)\right\| + \left\|g\big(\bar{x}(t_j),\bar{u}(t_j)\big) - g\big(x_j^k, u^k(t_j)\big)\right\| \\ &\le M_2^x \sum_{i=1}^m \left\|\bar{a}_i(t_j) - a_i^k(t_j)\right\| + L\left(\left\|\bar{x}(t_j) - x_j^k\right\| + \left\|\bar{u}(t_j) - u^k(t_j)\right\|\right).\end{aligned} \tag{4.12}$$

Letting now $x_{j+1}^k := x_j^k + h_j^k v_j^k$, we define the arcs $x^k(t)$ on $[0, T]$

$$x^k(t) := x_j^k + \frac{t - t_j}{h_j^k}\left(x_{j+1}^k - x_j^k\right) = x_j^k + (t - t_j)v_j^k, \ \text{ for } t \in [t_j, t_{j+1}] \tag{4.13}$$

and thus complete the construction of the pairs $\big(x^k(\cdot), b^k(\cdot)\big)$ in this step.

Step 3: *Estimates for trajectories.* Having in mind the subsequent proof of the strong $W^{1,2}$-convergence of discrete trajectories, we first derive the uniform estimates of the distance of the approximating discrete trajectories from the given one for (1.4). Similarly to $\bar{\alpha}(t)$ above, denote $\bar{\beta}(t) := \dot{\bar{b}}(t)$ on $[0, T]$ and then for each index $j = 0, \ldots, \nu(k) - 1$ and $i = 1, \ldots, m$ consider the functions on $[t_j, t_{j+1})$ defined by

$$\begin{gathered} f_j^x(s) := \left\|\dot{\bar{x}}(t_j) - \dot{\bar{x}}(s)\right\|, \ f_j^u(s) := \left\|\bar{u}(t_j) - \bar{u}(s)\right\|, \ f_{ij}^a(s) := \left\|\bar{\alpha}_i(t_j) - \bar{\alpha}_i(s)\right\|, \\ f_{ij}^b(s) := \left\|\bar{\beta}_i(t_j) - \bar{\beta}_i(s)\right\| \end{gathered}$$

and then select $s_j^x, s_j^u, s_{ij}^a, s_{ij}^b$ from the subintervals $[t_j, t_{j+1})$ such that

$$
\begin{cases}
\sup\limits_{s\in[t_j,t_{j+1}]} f_j^x(s) \le \left\|\dot{\bar{x}}(t_j)-\dot{\bar{x}}(s_j^x)\right\| + 2^{-k},\\
\sup\limits_{s\in[t_j,t_{j+1}]} f_j^u(s) \le \left\|\bar{u}(t_j)-\bar{u}(s_j^u)\right\| + 2^{-k},\\
\sup\limits_{s\in[t_j,t_{j+1}]} f_{ij}^a(s) \le \left\|\bar{\alpha}(t_j)-\bar{\alpha}(s_{ij}^a)\right\| + 2^{-k},\\
\sup\limits_{s\in[t_j,t_{j+1}]} f_{ij}^b(s) \le \left|\bar{\beta}(t_j)-\bar{\beta}(s_{ij}^b)\right| + 2^{-k}.
\end{cases}
\tag{4.14}
$$

With $h_k := \max\limits_{0\le j\le \nu(k)-1}\{h_j^k\}$, we get from the above the following relationships:

$$
\begin{aligned}
\left\|x_{j+1}^k-\bar{x}(t_{j+1})\right\| &= \left\|x_j^k+h_j^k v_j^k-\bar{x}(t_j)-\int_{t_j}^{t_{j+1}}\dot{\bar{x}}(s)ds\right\| \le \left\|x_j^k-\bar{x}(t_j)\right\| + \int_{t_j}^{t_{j+1}}\left\|v_j^k-\dot{\bar{x}}(s)\right\|ds\\
&\le \left\|x_j^k-\bar{x}(t_j)\right\| + \int_{t_j}^{t_{j+1}}\left\|v_j^k-\dot{\bar{x}}(t_j)\right\|ds + \int_{t_j}^{t_{j+1}}\left\|\dot{\bar{x}}(t_j)-\dot{\bar{x}}(s)\right\|ds\\
&\le \left\|x_j^k-\bar{x}(t_j)\right\| + h_j^k L\left(\left\|x_j^k-\bar{x}(t_j)\right\| + \left\|\bar{u}(t_j)-u^k(t_j)\right\|\right)\\
&+ h_j^k M_2^x\sum_{i=1}^m\left\|\bar{a}_i(t_j)-a_i^k(t_j)\right\| + \int_{t_j}^{t_{j+1}} f_j^x(s)ds\\
&\le \left(1+Lh_j^k\right)\left\|x_j^k-\bar{x}(t_j)\right\| + L\int_{t_j}^{t_{j+1}}\left\|\bar{u}(t_j)-\bar{u}(s)\right\|ds\\
&+ L\int_{t_j}^{t_{j+1}}\left\|\bar{u}(s)-u^k(s)\right\|ds + M_2^x m h_j^k\delta_k + \int_{t_j}^{t_{j+1}} f_j^x(s)ds\\
&\le (1+Lh_k)\left\|x_j^k-\bar{x}(t_j)\right\| + L\int_{t_j}^{t_{j+1}}\left\|\bar{u}(s)-u^k(s)\right\|ds + L\int_{t_j}^{t_{j+1}} f_j^u(s)ds+\\
&+ \int_{t_j}^{t_{j+1}} f_j^x(s)ds + M_2^x m h_j^k\delta_k.
\end{aligned}
\tag{4.15}
$$

Let $A := 1+Lh_k$, and for each $j=0,\ldots,\nu(k)-1$ denote $\gamma_j := \left\|x_j^k-\bar{x}(t_j)\right\|$ and

$$
\lambda_j := L\int_{t_j}^{t_{j+1}}\left\|\bar{u}(s)-u^k(s)\right\|ds + L\int_{t_j}^{t_{j+1}} f_j^u(s)ds + \int_{t_j}^{t_{j+1}} f_j^x(s)ds + M_2^x m h_j^k\delta_k.
$$

Then the final estimate in (4.15) reads as

$$\gamma_{j+1} \le A\gamma_j + \lambda_j \text{ for } j = 0, \dots, \nu(k) - 1,$$

which in turn implies the conditions

$$\gamma_j \le A^j \gamma_0 + A^{j-1}\lambda_0 + A^{j-2}\lambda_1 + \dots A^0\lambda_j = A^{j-1}\lambda_0 + A^{j-2}\lambda_1 + \dots A^0\lambda_j.$$

Since $A^j = (1 + Lh_k)^j \le (1 + Lh_k)^{\nu(k)} \le e^{L\widetilde{\nu}}$, we get $\gamma_j \le e^{L\widetilde{\nu}}\left(\lambda_0 + \lambda_1 + \dots + \lambda_j\right) \le e^{L\widetilde{\nu}} \sum_{j=0}^{\nu(k)-1} \lambda_j$.

Let us next estimate the quantity

$$\sum_{j=0}^{\nu(k)-1} \lambda_j = \sum_{j=0}^{\nu(k)-1} \left[L \int_{t_j}^{t_{j+1}} \left\| \bar{u}(s) - u^k(s) \right\| ds + L \int_{t_j}^{t_{j+1}} f_j^u(s) ds + \int_{t_j}^{t_{j+1}} f_j^x(s) ds + M_2^x m h_j^k \delta_k \right]. \tag{4.16}$$

To proceed, we deduce from (4.4) and (4.14) that

$$\begin{aligned} \sum_{j=0}^{\nu(k)-1} \int_{t_j}^{t_{j+1}} \left\| \bar{u}(s) - u^k(s) \right\| ds &\le \sqrt{T} \sqrt{\sum_{j=0}^{\nu(k)-1} \int_{t_j}^{t_{j+1}} \left\| \bar{u}(s) - u^k(s) \right\|^2 ds} \\ &= \sqrt{T} \sqrt{\int_0^T \left\| u^k(t) - \bar{u}(t) \right\|^2 dt} \le \sqrt{T\mu_k}, \end{aligned} \tag{4.17}$$

$$\begin{aligned} \sum_{j=0}^{\nu(k)-1} \int_{t_j}^{t_{j+1}} f_j^u(s) ds &\le h_k \sum_{j=0}^{\nu(k)-1} \left(\left\| \bar{u}(t_j) - \bar{u}(s_j^u) \right\| + \left\| \bar{u}(s_j^u) - \bar{u}(t_{j+1}) \right\| + 2^{-k} \right) \\ &\le h_k \operatorname{var}\big(\bar{u}; [0, T]\big) + h_k \nu(k) 2^{-k} \le h_k \mu + \widetilde{\nu} 2^{-k}. \end{aligned} \tag{4.18}$$

Using the same arguments leads us to the inequalities

$$\begin{aligned} \sum_{j=0}^{\nu(k)-1} \int_{t_j}^{t_{j+1}} f_j^x(s) ds &\le h_k \sum_{j=0}^{\nu(k)-1} \left(\left\| \dot{\bar{x}}(t_j) - \dot{\bar{x}}(s_j^x) \right\| + \left\| \dot{\bar{x}}(s_j^x) - \dot{\bar{x}}(t_{j+1}) \right\| + 2^{-k} \right) \\ &\le h_k \operatorname{var}\big(\dot{\bar{x}}; [0, T]\big) + h_k \nu(k) 2^{-k} \le h_k \mu + \widetilde{\nu} 2^{-k}. \end{aligned}$$

On the other hand, we clearly have that

$$\sum_{j=0}^{\nu(k)-1} M_2^x m h_j^k \delta_k \le M_2^x m \nu(k) h_k \delta_k \le M_2^x m \widetilde{\nu} \delta_k.$$

Combining all the above brings us to the desired estimate of the quantity (4.16) and hence of $\|x_j^k - \bar{x}(t_j)\|$:

$$\sum_{j=0}^{\nu(k)-1} \lambda_j \le L(\sqrt{T\mu_k} + h_k\mu + \widetilde{\nu}2^{-k}) + h_k\mu + \widetilde{\nu}2^{-k} + M_2^x m\widetilde{\nu}\delta_k$$

$$\le \left(h_k\mu + \widetilde{\nu}2^{-k}\right)(L+1) + L\sqrt{T\mu_k} + M_2^x m\widetilde{\nu}\delta_k,$$

$$\left\|x_j^k - \bar{x}(t_j)\right\| \le \vartheta_k := e^{L\widetilde{\nu}}\left[\left(h_k\mu + \widetilde{\nu}2^{-k}\right)(L+1) + L\sqrt{T\mu_k} + M_2^x m\widetilde{\nu}\delta_k\right] \tag{4.19}$$

for all $j = 0, \ldots, \nu(k)$. Employing this together with (4.13), (4.15), and (4.19) gives us

$$\left\|x^k(t) - \bar{x}(t)\right\| = \left\|x_j^k + h_k v_j^k - \bar{x}(t_j) - \int_{t_j}^{t} \dot{\bar{x}}(s)ds\right\| \le \left\|x_j^k - \bar{x}(t_j)\right\| + \int_{t_j}^{t}\left\|v_j^k - \dot{\bar{x}}(s)\right\| ds$$

$$\le \left\|x_j^k - \bar{x}(t_j)\right\| + \int_{t_j}^{t_{j+1}}\left\|v_j^k - \dot{\bar{x}}(s)\right\| ds \le (1 + Lh_k)\left\|x_j^k - \bar{x}(t_j)\right\| + \lambda_j$$

$$\le (1 + Lh_k)\,\vartheta_k + \lambda_j \;\text{ whenever }\; t \in (t_j, t_{j+1}] \;\text{ and }\; j = 0, \ldots, k-1,$$

which justifies by $\lambda_j \to 0$ the uniform convergence of the sequence $\left\{x^k(\cdot)\right\}$ to $\bar{x}(\cdot)$ as $k \to \infty$.

Step 4: *Verifying the strong $W^{1,2}$-convergence of $x^k(\cdot)$ to $\bar{x}(\cdot)$ on* $[0, T]$. To establish further the L^2-strong convergence of $\left\{\dot{x}^k(\cdot)\right\}$ to $\dot{\bar{x}}(\cdot)$ on $[0, T]$ as $k \to \infty$, deduce first from (4.12) that

$$h_j^k\left\|v_j^k - \dot{\bar{x}}(t_j)\right\|^2 \le h_j^k\left[M_2^x\sum_{i=1}^{m}\left\|\bar{a}_i(t_j) - a_i^k(t_j)\right\| + L\left(\left\|\bar{x}(t_j) - x_j^k\right\| + \left\|\bar{u}(t_j) - u^k(t_j)\right\|\right)\right]^2$$

$$\le 3(M_2^x)^2 h_j^k m\sum_{i=1}^{m}\left\|\bar{a}_i(t_j) - a_i^k(t_j)\right\|^2 + 3Lh_j^k\left\|\bar{x}(t_j) - x_j^k\right\|^2$$

$$+ 3Lh_j^k\left\|\bar{u}(t_j) - u^k(t_j)\right\|^2$$

$$\le 3(M_2^x)^2 m^2\delta_k^2 h_k + 3Lh_k\delta_k^2 + 3Lh_j^k\left\|\bar{u}(t_j) - u^k(t_j)\right\|^2$$

for $j = 0, \ldots, \nu(k) - 1$ and then subsequently derive the estimates

$$\int_0^T\left\|\dot{x}^k(t) - \dot{\bar{x}}(t)\right\|^2 dt = \sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|v_j^k - \dot{\bar{x}}(t)\right\|^2 dt$$

$$\le \sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left(\left\|v_j^k - \dot{\bar{x}}(t_j)\right\| + \|\dot{\bar{x}}(t_j) - \dot{\bar{x}}(t)\|\right)^2 dt$$

$$
\begin{aligned}
&\le 2\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|v_j^k-\dot{\bar{x}}(t_j)\right\|^2dt+2\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}}\left\|\dot{\bar{x}}(t_j)-\dot{\bar{x}}(t)\right\|^2dt\\
&\le 2\sum_{j=0}^{\nu(k)-1}\left[3(M_2^x)^2m^2\delta_k^2h_k+3Lh_k\delta_k^2+3Lh_j^k\left\|\bar u(t_j)-u^k(t_j)\right\|^2\right]+2\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left[f_j^x(t)\right]^2dt\\
&\le 6(M_2^x)^2m^2\delta_k^2h_k\nu(k)+6Lh_k\nu(k)\delta_k^2+6L\sum_{j=0}^{\nu(k)-1}h_j^k\left\|\bar u(t_j)-u^k(t_j)\right\|^2\\
&+2\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left[f_j^x(t)\right]^2dt\\
&\le 6(M_2^x)^2m^2\delta_k^2\widetilde{\nu}+6L\widetilde{\nu}\delta_k^2+6L\sum_{j=0}^{\nu(k)-1}h_j^k\left\|\bar u(t_j)-u^k(t_j)\right\|^2+2\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left[f_j^x(t)\right]^2dt.
\end{aligned}
$$

Since the control set U is compact, there exists a number $\overline{M}>0$ such that $\max\left\{\|\bar u(t)\|,\left\|u^k(t)\right\|\right\}\le\overline{M}$ for all $t\in[0,T]$. On the other hand, it follows from (4.17) and (4.18) that

$$
\begin{aligned}
\sum_{j=0}^{\nu(k)-1}h_j^k\left\|\bar u(t_j)-u^k(t_j)\right\|^2&\le 2\overline{M}\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|\bar u(t_j)-u^k(t_j)\right\|dt\\
&\le 2\overline{M}\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|\bar u(t_j)-\bar u(t)\right\|dt+2\overline{M}\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|\bar u(t)-u^k(t)\right\|dt\\
&\le 2\overline{M}\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}f_j^u(t)dt+2\overline{M}\sqrt{T\mu_k}\\
&\le 2\overline{M}\left(h_k\mu+\widetilde{\nu}2^{-k}+\sqrt{T\mu_k}\right).
\end{aligned}
$$

In addition we get from the constructions and notation above that

$$
\begin{aligned}
\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left[f_j^x(t)\right]^2dt&\le\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left[f_j^x(s_j)+2^{-k}\right]^2dt=\sum_{j=0}^{\nu(k)-1}h_j^k\left[f_j^x(s_j)+2^{-k}\right]^2\\
&\le 2h_k\sum_{j=0}^{\nu(k)-1}\left\{\left[f_j^x(s_j)\right]^2+4^{-k}\right\}\le 2h_k\left[\sum_{j=0}^{\nu(k)-1}f_j^x(s_j)\right]^2+2h_k\nu(k)4^{-k}
\end{aligned}
$$

$$\leq 2h_k\left[\sum_{j=0}^{\nu(k)-1}\left(\left\|\dot{\bar{x}}(t_j)-\dot{\bar{x}}(s_j^x)\right\|+\left\|\dot{\bar{x}}(s_j^x)-\dot{\bar{x}}(t_{j+1})\right\|\right)\right]^2+2\widetilde{\nu}4^{-k}$$
$$\leq 2h_k\mathrm{var}^2\left(\dot{\bar{x}}(\cdot);[0,T]\right)+2\widetilde{\nu}4^{-k}\leq 2h_k\mu^2+2\widetilde{\nu}4^{-k}.$$

This finally brings us to the estimate

$$\int_0^T\left\|\dot{x}^k(t)-\dot{\bar{x}}(t)\right\|^2dt\leq 6(M_2^x)^2m^2\delta_k^2\widetilde{\nu}+6L\widetilde{\nu}\delta_k^2+12\overline{M}\left(h_k\mu+\widetilde{\nu}2^{-k}+\sqrt{T\mu_k}\right)$$
$$+4h_k\mu^2+4\widetilde{\nu}4^{-k},$$

which justifies the L^2-strong convergence of $\left\{\dot{x}^k(\cdot)\right\}$ to $\dot{\bar{x}}(\cdot)$ in the norm topology as claimed at Step 3.

Step 5: *Uniform estimates for b-controls.* As a part of the verification of the $W^{1,2}$-convergence of $b^k(\cdot)$ to $\bar{b}(\cdot)$, we establish first the needed estimates for approximating b-controls. Picking any $t\in(t_j,t_{j+1}]$ and then using (4.8) and (4.9), we immediately observe that

$$\left|b_i^k(t)-\bar{b}_i(t)\right|=\left|b_{ij}^k+\frac{t-t_j}{h_j^k}\left(b_{i,j+1}^k-b_{ij}^k\right)-\bar{b}(t)\right|$$
$$\leq\left|\bar{b}_i(t_j)-\bar{b}_i(t)\right|+\left|\left\langle a_i^k(t_j),x_j^k\right\rangle-\left\langle\bar{a}_i(t_j),\bar{x}(t_j)\right\rangle\right|+\left|b_{i,j+1}^k-b_{ij}^k\right|.$$

Since $\bar{b}(\cdot)$ is uniformly continuous on $[0,T]$, for any $\varepsilon>0$ there exists $\delta>0$ ensuring that

$$\max\{|t-s|,h_k\}<\delta\Longrightarrow\left\|\bar{b}(t)-\bar{b}(s)\right\|\leq\varepsilon,$$

which implies that $\left|\bar{b}_i(t_j)-\bar{b}_i(t)\right|\leq\varepsilon$. Furthermore, it follows from (4.6) and (4.19) that

$$\left|\left\langle a_i^k(t_j),x_j^k\right\rangle-\left\langle\bar{a}_i(t_j),\bar{x}(t_j)\right\rangle\right|=\left|\left\langle a_i^k(t_j)-\bar{a}_i(t_j),x_j^k\right\rangle+\left\langle\bar{a}_i(t_j),x_j^k-\bar{x}(t_j)\right\rangle\right|$$
$$\leq\left\|a_i^k(t_j)-\bar{a}_i(t_j)\right\|\cdot\left\|x_j^k\right\|+\left\|\bar{a}_i(t_j)\right\|\left\|x_j^k-\bar{x}(t_j)\right\|$$
$$\leq M_1\delta_k+\vartheta_k,$$

where $M_1>0$ is chosen so that $\left\|x_j^k\right\|\leq M_1$ for all $j=0,\ldots,k-1$, δ_k was defined in (4.7), and ϑ_k was defined in (4.19). Consequently we have

$$\left|b_{i,j+1}^k-b_{ij}^k\right|=\left|\bar{b}_i(t_{j+1})-\bar{b}_i(t_j)+\left\langle a_i^k(t_{j+1}),x_{j+1}^k\right\rangle-\left\langle\bar{a}_i(t_{j+1}),\bar{x}(t_{j+1})\right\rangle\right.$$
$$\left.-\left\langle a_i^k(t_j),x_j^k\right\rangle+\left\langle\bar{a}_i(t_j),\bar{x}(t_j)\right\rangle\right|$$
$$\leq\left|\bar{b}_i(t_{j+1})-\bar{b}_i(t_j)\right|+\left|\left\langle a_i^k(t_{j+1}),x_{j+1}^k\right\rangle-\left\langle\bar{a}_i(t_{j+1}),\bar{x}(t_{j+1})\right\rangle\right|$$

$$+ \left| \left\langle a_i^k(t_j), x_j^k \right\rangle - \langle \bar{a}_i(t_j), \bar{x}(t_j) \rangle \right|$$
$$\leq \varepsilon + 2\,(M_1\delta_k + \vartheta_k)\,,$$

which justifies the fulfillment of the claimed estimate

$$\left| b_i^k(t) - \bar{b}_i(t) \right| \leq 2\varepsilon + 3\,(M_1\delta_k + \vartheta_k)$$

and thus justifies the uniform convergence of $\left\{ b^k(\cdot) \right\}$ to $\bar{b}(\cdot)$ on $[0, T]$, thanks to (4.7), (4.4), and (4.19).

Step 6: *Verifying the convergence of $b^k(\cdot)$ to $\bar{b}(\cdot)$ in $W^{1,2}([0,T]; \mathbb{R}^m)$.* It remains to prove the L^2-strong convergence of $\dot{b}^k(\cdot)$ to $\dot{\bar{b}}(\cdot)$ on $[0, T]$. For any $t \in [t_j, t_{j+1})$ we get

$$\begin{aligned}
\left| \dot{b}_i^k(t) - \dot{\bar{b}}_i(t) \right| &= \left| \frac{b_{i,j+1}^k - b_{ij}^k}{h_j^k} - \dot{\bar{b}}_i(t) \right| \\
&\leq \left| \frac{\bar{b}_i(t_{j+1}) - \bar{b}_i(t_j)}{h_j^k} - \dot{\bar{b}}_i(t) \right| \\
&+ \left| \left\langle \frac{a_i^k(t_{j+1}) - a_i^k(t_j)}{h_j^k} - \frac{\bar{a}_i(t_{j+1}) - \bar{a}_i(t_j)}{h_j^k}, x_{j+1}^k \right\rangle \right| \\
&+ \left| \left\langle \frac{\bar{a}_i(t_{j+1}) - \bar{a}_i(t_j)}{h_j^k}, x_{j+1}^k - \bar{x}(t_{j+1}) \right\rangle \right| \\
&+ \left| \left\langle a_i^k(t_j), \frac{x_{j+1}^k - x_j^k}{h_j^k} - \frac{\bar{x}(t_{j+1}) - \bar{x}(t_j)}{h_j^k} \right\rangle \right| \\
&+ \left| \left\langle a_i^k(t_j) - \bar{a}_i(t_j), \frac{\bar{x}(t_{j+1}) - \bar{x}(t_j)}{h_j^k} \right\rangle \right| \\
&\leq \left| \frac{\bar{b}_i(t_{j+1}) - \bar{b}_i(t_j)}{h_j^k} - \dot{\bar{b}}_i(t) \right| + M_1 \left\| \dot{a}_i^k(t) - \frac{\bar{a}_i(t_{j+1}) - \bar{a}_i(t_j)}{h_j^k} \right\| \\
&+ \vartheta_k \left\| \frac{\bar{a}_i(t_{j+1}) - \bar{a}_i(t_j)}{h_j^k} \right\| + (1 + \delta_k) \left\| \dot{x}^k(t) - \frac{\bar{x}(t_{j+1}) - \bar{x}(t_j)}{h_j^k} \right\| \\
&+ \delta_k \left\| \frac{\bar{x}(t_{j+1}) - \bar{x}(t_j)}{h_k} \right\|
\end{aligned}$$

due to (4.9), (4.3), (4.6), and (4.19). Since $\bar{\alpha}(\cdot)$ is a BV function, it follows that

$$\|\bar{\alpha}_i(s)\| \leq M_2 := \frac{1}{2}\left[\|\bar{\alpha}_{i0}\| + \|\bar{\alpha}_i(T)\| + \mathrm{var}\big(\bar{\alpha}_i(\cdot); [0, T]\big) \right] \text{ for all } s \in [0, T],$$

and therefore $\left\|\frac{\bar{a}_i(t_{j+1})-\bar{a}_i(t_j)}{h_k}\right\| = \frac{1}{h_k}\left\|\int_{t_j}^{t_{j+1}} \bar{\alpha}_i(s)ds\right\| \le M_2$. Arguing in the same way for the BV function $\dot{\bar{x}}(\cdot)$ shows that $\left\|\frac{\bar{x}(t_{j+1})-\bar{x}(t_j)}{h_k}\right\| \le M_3$ with some constant $M_3>0$.

Next we estimate the quantities $\left|\frac{\bar{b}_i(t_{j+1})-\bar{b}_i(t_j)}{h_j^k} - \dot{\bar{b}}_i(t)\right|$ and $\left\|\dot{x}^k(t) - \frac{\bar{x}(t_{j+1})-\bar{x}(t_j)}{h_j^k}\right\|$. Observe that

$$\begin{aligned}\left|\frac{\bar{b}_i(t_{j+1})-\bar{b}_i(t_j)}{h_j^k} - \dot{\bar{b}}_i(t)\right| &\le \frac{1}{h_j^k}\int_{t_j}^{t_{j+1}} \left|\bar{\beta}_i(s)-\bar{\beta}_i(t_j)\right|ds + \left|\bar{\beta}_i(t_j)-\bar{\beta}_i(t)\right| \\ &= \frac{1}{h_j^k}\int_{t_j}^{t_{j+1}} f_{ij}^b(s)ds + f_{ij}^b(t) \le 2f_{ij}^b(s_{ij}^b) + 2^{-k+1} \\ &\le 2\left[\left|\bar{\beta}_i(s_{ij}^b)-\bar{\beta}_i(t_j)\right| + \left|\bar{\beta}_i(t_{j+1})-\bar{\beta}_i(s_{ij}^b)\right|\right] + 2^{-k+1},\end{aligned} \tag{4.20}$$

which allows us while arguing as above to get the estimates

$$\begin{aligned}\left\|\dot{a}_i^k(t) - \frac{\bar{a}_i(t_{j+1})-\bar{a}_i(t_j)}{h_j^k}\right\| &\le \left\|\dot{a}_i^k(t)-\dot{\bar{a}}_i(t)\right\| + \left\|\dot{\bar{a}}_i(t) - \frac{\bar{a}_i(t_{j+1})-\bar{a}_i(t_j)}{h_j^k}\right\| \\ &\le \left\|\dot{a}_i^k(t)-\dot{\bar{a}}_i(t)\right\| + 2\left[\left\|\bar{\alpha}_i(s_{ij}^a)-\bar{\alpha}_i(t_j)\right\| + \left\|\bar{\alpha}_i(t_{j+1})-\bar{\alpha}_i(s_{ij}^a)\right\|\right] \\ &\quad + 2^{-k+1}, \\ \left\|\dot{x}^k(t) - \frac{\bar{x}(t_{j+1})-\bar{x}(t_j)}{h_j^k}\right\| &\le \left\|\dot{x}^k(t)-\dot{\bar{x}}(t)\right\| + \left\|\dot{\bar{x}}(t) - \frac{\bar{x}(t_{j+1})-\bar{x}(t_j)}{h_j^k}\right\| \\ &\le \left\|\dot{x}^k(t)-\dot{\bar{x}}(t)\right\| + 2\left[\left\|\dot{\bar{x}}(s_j^x)-\dot{\bar{x}}(t_j)\right\| + \left\|\dot{\bar{x}}(t_{j+1})-\dot{\bar{x}}(s_j^x)\right\|\right] \\ &\quad + 2^{-k+1}.\end{aligned} \tag{4.21}$$

It then follows by combining all the estimates in (4.20)–(4.21) that

$$\begin{aligned}\int_0^T \left|\dot{b}_i^k(t)-\dot{\bar{b}}_i(t)\right|^2 dt &= \sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}} \left|\frac{b_{i,j+1}^k - b_{ij}^k}{h_j^k} - \dot{\bar{b}}_i(t)\right|^2 dt \\ \le \sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}} &\left[1+M_1^2+(1+\delta_k)^2+1\right]\left[\left|\frac{\bar{b}_i(t_{j+1})-\bar{b}_i(t_j)}{h_j^k} - \dot{\bar{b}}_i(t)\right|^2\right. \\ + &\left\|\dot{a}_i^k(t) - \frac{\bar{a}_i(t_{j+1})-\bar{a}_i(t_j)}{h_j^k}\right\|^2 + \vartheta_k^2\left\|\frac{\bar{a}_i(t_{j+1})-\bar{a}_i(t_j)}{h_j^k}\right\|^2\end{aligned}$$

$$
\begin{aligned}
&+\left\|\dot{x}^k(t)-\frac{\bar{x}(t_{j+1})-\bar{x}(t_j)}{h_j^k}\right\|^2+\delta_k^2\left\|\frac{\bar{x}(t_{j+1})-\bar{x}(t_j)}{h_j^k}\right\|^2\Bigg]dt\\
&\le\left[2+M_1^2+(1+\delta_k)^2\right]\Bigg[\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left|\frac{\bar{b}_i(t_{j+1})-\bar{b}_i(t_j)}{h_j^k}-\dot{\bar{b}}_i(t)\right|^2dt\\
&+\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|\dot{a}_i^k(t)-\frac{\bar{a}_i(t_{j+1})-\bar{a}_i(t_j)}{h_j^k}\right\|^2dt+\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\vartheta_k^2M_2^2dt\\
&+\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|\dot{x}^k(t)-\frac{\bar{x}(t_{j+1})-\bar{x}(t_j)}{h_j^k}\right\|^2dt+\delta_k^2\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|\frac{\bar{x}(t_{j+1})-\bar{x}(t_j)}{h_j^k}\right\|^2dt\Bigg]\\
&\le\left[2+M_1^2+(1+\delta_k)^2\right]\Bigg[4h_k\sum_{j=0}^{\nu(k)-1}\left[\left|\bar{\beta}_i(s_{ij}^b)-\bar{\beta}_i(t_j)\right|+\left|\bar{\beta}_i(t_{j+1})-\bar{\beta}_i(s_{ij}^b)\right|+2^{-k}\right]^2\\
&+\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|\dot{a}_i^k(t)-\dot{\bar{a}}_i(t)\right\|^2dt\\
&+4h_k\sum_{j=0}^{\nu(k)-1}\left[\left\|\bar{\alpha}_i(s_{ij}^a)-\bar{\alpha}_i(t_j)\right\|+\left\|\bar{\alpha}_i(t_{j+1})-\bar{\alpha}_i(s_{ij}^a)\right\|+2^{-k}\right]^2+\vartheta_k^2M_2^2\widetilde{\nu}\\
&+\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}\left\|\dot{x}^k(t)-\dot{\bar{x}}(t)\right\|^2dt\\
&+4h_k\sum_{j=0}^{\nu(k)-1}\left[\left\|\dot{\bar{x}}(s_j^x)-\dot{\bar{x}}(t_j)\right\|+\left\|\dot{\bar{x}}(t_{j+1})-\dot{\bar{x}}(s_j^x)\right\|\right]^2+\delta_k^2\sum_{j=0}^{\nu(k)-1}\int_{t_j}^{t_{j+1}}M_3^2dt\Bigg].
\end{aligned}
$$

Finally, we arrive at the relationships

$$
\begin{aligned}
&\int_0^T\left|\dot{b}_i^k(t)-\dot{\bar{b}}_i(t)\right|^2dt\\
&\le\left[2+M_1^2+(1+\delta_k)^2\right]\Bigg\{8h_k\sum_{j=0}^{\nu(k)-1}\left[\left|\bar{\beta}_i(s_{ij}^b)-\bar{\beta}_i(t_j)\right|+\left|\bar{\beta}_i(t_{j+1})-\bar{\beta}_i(s_{ij}^b)\right|\right]^2\\
&+4^{-k}h_k\nu(k)+\int_0^T\left\|\dot{a}_i^k(t)-\dot{\bar{a}}_i(t)\right\|^2dt
\end{aligned}
$$

$$+ 8h_k \sum_{j=0}^{\nu(k)-1} \Big[\Big\|\bar{\alpha}_i(s_{ij}^a) - \bar{\alpha}_i(t_j)\Big\| + \Big\|\bar{\alpha}_i(t_{j+1}) - \bar{\alpha}_i(s_{ij}^a)\Big\| \Big]^2$$

$$+ 4^{-k} h_k \nu(k) + \vartheta_k^2 M_2^2 \widetilde{\nu} + \int_0^T \Big\|\dot{x}^k(t) - \dot{\bar{x}}(t)\Big\|^2 dt$$

$$+ 8h_k \sum_{j=0}^{\nu(k)-1} \Big[\Big\|\dot{\bar{x}}(s_j^x) - \dot{\bar{x}}(t_j)\Big\| + \Big\|\dot{\bar{x}}(t_{j+1}) - \dot{\bar{x}}(s_j^x)\Big\| \Big]^2 + 4^{-k} h_k \nu(k) + \delta_k^2 M_3^2 \widetilde{\nu} \Big\}$$

$$\le \big[2 + M_1^2 + (1+\delta_k)^2\big] \Big[8h_k \Big(\mathrm{var}^2(\bar{\beta}(\cdot); [0,T]) + \mathrm{var}^2(\bar{\alpha}(\cdot); [0,T]) + \mathrm{var}^2(\dot{\bar{x}}(\cdot); [0,T]) \Big)$$

$$+ \frac{3}{4^k}\widetilde{\nu} + \vartheta_k^2 M_2^2 \widetilde{\nu} + \delta_k^2 M_3^2 \widetilde{\nu} + \int_0^T \Big\|\dot{a}_i^k(t) - \dot{\bar{a}}_i(t)\Big\|^2 dt + \int_0^T \Big\|\dot{x}^k(t) - \dot{\bar{x}}(t)\Big\|^2 dt \Big]$$

$$\le \big[2 + M_1^2 + (1+\delta_k)^2\big] \left[24 h_k \mu^2 + \frac{3}{4^k}\widetilde{\nu} + \vartheta_k^2 M_2^2 \widetilde{\nu} + \delta_k^2 M_3^2 \widetilde{\nu} + \mu_k + \int_0^T \Big\|\dot{x}^k(t) - \dot{\bar{x}}(t)\Big\|^2 dt \right],$$

which ensures the convergence of the sequence $\{\dot{b}^k(\cdot)\}$ to $\dot{\bar{b}}(\cdot)$ strongly in $L^2([0,T];\mathbb{R}^m)$ as claimed in Step 4. This therefore completes the proof of the theorem. $\square$

As we see, the entire proof of the theorem is technically involved. It occurs nevertheless that the most important and challenging task is the construction of a sequence of piecewise linear functions $x^k(\cdot)$, which are feasible to the discrete differential inclusion (4.3). The main point is in approximating the continuous velocity $\dot{\bar{x}}(t_j) \in -F(\bar{x}(t_j), \bar{a}(t_j), \bar{b}(t_j), \bar{u}(t_j))$ by its discrete counterpart $v_j^k \in -F(x^k(t_j), a^k(t_j), b^k(t_j), u^k(t_j))$, where the velocity mapping F is discontinuous. Using the construction of v_j^k in (4.11) ensures that the distance between $\dot{\bar{x}}(t_j)$ and v_j^k converges to 0 as $k \to \infty$, which is the key.

5. Discrete approximation for relaxed local minimizers

The discrete approximation procedure and results developed in the previous section do not require any relaxation stability and do not concern optimal versus feasible solutions. The discrete approximation construction and the main result of this section address *relaxed local minimizers* of the sweeping optimal control problem (P).

Let $\big(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot)\big)$ be a given r.i.l.m., and let Δ_k be the discrete mesh defined in (4.1). For all $k \in I\!N$ we construct a sequence of approximating problems (P_k) as follows:

$$\text{minimize } J_k[x^k, a^k, b^k, u^k]$$

$$:= \varphi(x_{\nu(k)}^k) + \sum_{j=0}^{\nu(k)-1} h_j^k \ell\left(t_j^k, x_j^k, a_j^k, b_j^k, u_j^k, \frac{x_{j+1}^k - x_i^k}{h_j^k}, \frac{a_{j+1}^k - a_j^k}{h_j^k}, \frac{b_{j+1}^k - b_j^k}{h_j^k}\right) \tag{5.1}$$

$$+\frac{1}{2}\sum_{j=0}^{\nu(k)-1}\int_{t_j^k}^{t_{j+1}^k}\left\|\left(\frac{x_{j+1}^k-x_j^k}{h_j^k},\frac{a_{j+1}^k-a_j^k}{h_j^k},\frac{b_{j+1}^k-b_j^k}{h_j^k},u_j^k\right)-\left(\dot{\bar{x}}(t),\dot{\bar{a}}(t),\dot{\bar{b}}(t),\bar{u}(t)\right)\right\|^2 dt$$

over discrete quadruples (x^k,a^k,b^k,u^k) represented by

$$(x^k,a^k,b^k,u^k):=(x_0^k,x_1^k,\ldots,x_{\nu(k)}^k,a_0^k,a_1^k,\ldots,a_{\nu(k)}^k,b_0^k,b_1^k,\ldots,b_{\nu(k)}^k,u_0^k,u_1^k,\ldots,u_{\nu(k)-1}^k)$$

subject to the geometric and functional constraints given by

$$x_{j+1}^k\in x_j^k-h_j^kF(x_j^k,a_j^k,b_j^k,u_j^k),\ j=0,\ldots,\nu(k)-1, \tag{5.2}$$

$$\left\langle a_{i\nu(k)}^k,x_{\nu(k)}^k\right\rangle\le b_{i\nu(k)}^k,\ i=1,\ldots,m, \tag{5.3}$$

$$x_0^k=x_0\in C(0),\ a_0^k=a_0,\ b_0^k=b_0,\ u_0^k=\bar{u}(0), \tag{5.4}$$

$$\sum_{j=0}^{\nu(k)-1}\int_{t_j^k}^{t_{j+1}^k}\left\|\left(x_j^k,a_j^k,b_j^k,u_j^k\right)-\left(\bar{x}(t),\bar{a}(t),\bar{b}(t),\bar{u}(t)\right)\right\|^2 dt\le\frac{\varepsilon}{2}, \tag{5.5}$$

$$\sum_{j=0}^{\nu(k)-1}\int_{t_j^k}^{t_{j+1}^k}\left\|\left(\frac{x_{j+1}^k-x_j^k}{h_j^k},\frac{a_{j+1}^k-a_j^k}{h_j^k},\frac{b_{j+1}^k-b_j^k}{h_j^k}\right)-\left(\dot{\bar{x}}(t),\dot{\bar{a}}(t),\dot{\bar{b}}(t)\right)\right\|^2 dt\le\frac{\varepsilon}{2}, \tag{5.6}$$

$$u_j^k\in U,\ j=0,\ldots,\nu(k)-1, \tag{5.7}$$

$$1-\delta_k\le\|a_{ij}^k\|\le 1+\delta_k,\ i=1,\ldots,m,\ j=0,\ldots,\nu(k), \tag{5.8}$$

where $\varepsilon>0$ is taken from Definition 3.3 of the relaxed intermediate local minimizer $\big(\bar{x}(\cdot),\bar{a}(\cdot),\bar{b}(\cdot),\bar{u}(\cdot)\big)$, where F is defined in (3.5), and where the perturbation sequence $\delta_k\downarrow 0$ as $k\to\infty$ is constructed in the proof of Theorem 4.1 for the given quadruple $\big(\bar{x}(\cdot),\bar{a}(\cdot),\bar{b}(\cdot),\bar{u}(\cdot)\big)$.

To proceed further, first we need to make sure that for each $k\in I\!N$ sufficiently large the discrete control problem (P_k) defined in (5.1)–(5.8) admits an optimal solution. It is verified in the next proposition.

Proposition 5.1 *(existence of optimal solutions to discrete sweeping control problems). Under the assumptions in Theorem 4.1 holding along the given r.i.l.m. $\big(\bar{x}(\cdot),\bar{a}(\cdot),\bar{b}(\cdot),\bar{u}(\cdot)\big)$, each problem (P_k) for all sufficiently large $k\in I\!N$ admits an optimal solution.*

Proof. It follows from Theorem 4.1 that the set of feasible solutions of problem (P_k) is nonempty for all large k. We see in addition that this set is bounded due to the constraint structures in (P_k). Furthermore, the cost function in (P_k) is obviously lower semicontinuous for each $t_j^k\in\Delta_k$ due to (H5). To apply the classical Weierstrass existence theorem in (P_k), it remains to ensure that the feasible set in this problem is closed. But it is a direct consequence of the constraint structures in (P_k) due to the robustness (closed-graph) property of the normal cone mapping (1.2). Thus we arrive at the claimed existence result. □

Now we are ready to establish the desired theorem on the strong convergence of optimal solutions for (P_k) to the given r.i.l.m. of the original sweeping control problem (P).

Theorem 5.2 *(strong convergence of discrete optimal solutions). Let $(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot))$ be an r.i.l.m. for problem (P), and let all the assumptions of Proposition 5.1 be satisfied for this quadruple. Suppose in addition that the terminal cost φ is continuous around $\bar{x}(T)$, that the running cost ℓ is continuous at $\big(t, \bar{x}(t), \bar{a}(t), \bar{b}(t), \bar{u}(t), \dot{\bar{x}}(t), \dot{\bar{a}}(t), \dot{\bar{b}}(t)\big)$ for a.e. $t \in [0, T]$, and that $\ell\big(\cdot, x, a, b, u, \dot{x}, \dot{a}, \dot{b}\big)$ is uniformly majorized around $\big(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot)\big)$ by a summable function on $[0, T]$. Take any sequence of optimal solutions $\big(\bar{x}^k(\cdot), \bar{a}^k(\cdot), \bar{b}^k(\cdot), \bar{u}^k(\cdot)\big)$ to the discrete problems (P_k) and extend it to the entire interval $[0, T]$ piecewise linearly for $\big(\bar{x}^k(\cdot), \bar{a}^k(\cdot), \bar{b}^k(\cdot)\big)$ and piecewise constantly for $\bar{u}^k(\cdot)$. Then the extended sequence of $\big(\bar{x}^k(\cdot), \bar{a}^k(\cdot), \bar{b}^k(\cdot), \bar{u}^k(\cdot)\big)$ converges to $\big(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot)\big)$ as $k \to \infty$ in the norm topology of $W^{1,2}([0, T]; \mathbb{R}^n) \times W^{1,2}([0, T]; \mathbb{R}^{mn}) \times W^{1,2}([0, T]; \mathbb{R}^m) \times L^2([0, T]; \mathbb{R}^d)$.*

Proof. Picking any sequence $\big(\bar{x}^k(\cdot), \bar{a}^k(\cdot), \bar{b}^k(\cdot), \bar{u}^k(\cdot)\big)$ of extended optimal solutions to (P_k), we claim that

$$\lim_{k\to\infty} \int_0^T \Big\| \Big(\dot{\bar{x}}^k(t), \dot{\bar{a}}^k(t), \dot{\bar{b}}^k(t), \bar{u}^k(t)\Big) - \Big(\dot{\bar{x}}(t), \dot{\bar{a}}(t), \dot{\bar{b}}(t), \bar{u}(t)\Big) \Big\|^2 dt = 0, \tag{5.9}$$

which clearly ensures the convergence of the quadruples $\big(\bar{x}^k(\cdot), \bar{a}^k(\cdot), \bar{b}^k(\cdot), \bar{u}^k(\cdot)\big)$ to $\big(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot)\big)$ in the norm topology of $W^{1,2}([0, T]; \mathbb{R}^{n+mn+m}) \times L^2([0, T]; \mathbb{R}^d)$. To proceed, assume on the contrary that the limit in (5.9), along a subsequence (without relabeling), equals to some $\gamma > 0$. Then it follows from the weak compactness of the unit ball in $L^2([0, T]; \mathbb{R}^{n+mn+m+d})$ that there exist functions $(v^x(\cdot), v^a(\cdot), v^b(\cdot), \widetilde{u}(\cdot)) \in L^2([0, T]; \mathbb{R}^{n+mn+m+d})$ for which the quadruples $(\dot{\bar{x}}^k(\cdot), \dot{\bar{a}}^k(\cdot), \dot{\bar{b}}^k(\cdot), \bar{u}^k(\cdot))$ converge weakly to $\big(v^x(\cdot), v^a(\cdot), v^b(\cdot), \widetilde{u}(\cdot)\big)$ in the corresponding spaces. Recall that Mazur's weak closure theorem and basic real analysis yield the existence of sequences of convex combinations of these quadruples that converge to $(v^x(\cdot), v^a(\cdot), v^b(\cdot), \widetilde{u}(\cdot))$ in the L^2-norm topology with their subsequences (no relabeling) converging to $(v^x(t), v^a(t), v^b(t), \widetilde{u}(t))$ for a.e. $t \in [0, T]$. Define further the triple $(\widetilde{x}(\cdot), \widetilde{a}(\cdot), \widetilde{b}(\cdot)) \in W^{1,2}([0, T]; \mathbb{R}^{n+mn+m})$ by

$$\big(\widetilde{x}(t), \widetilde{a}(t), \widetilde{b}(t)\big) := (x_0, a_0, b_0) + \int_0^t \Big(v^x(s), v^a(s), v^b(s)\Big) ds \quad \text{for all } t \in [0, T].$$

Then $(\dot{\widetilde{x}}(t), \dot{\widetilde{a}}(t), \dot{\widetilde{b}}(t)) = (v^x(t), v^a(t), v^b(t))$ for a.e. $t \in [0, T]$, which ensures the weak convergence of $(\dot{\bar{x}}^k(\cdot), \dot{\bar{a}}^k(\cdot), \dot{\bar{b}}^k(\cdot))$ to $(\dot{\widetilde{x}}(\cdot), \dot{\widetilde{a}}(\cdot), \dot{\widetilde{b}}(\cdot))$ in $L^2([0, T]; \mathbb{R}^{n+mn+m})$. Observe that $\widetilde{u}(t) \in \operatorname{co} U$ for a.e. $t \in [0, T]$ and that the limiting triple $(\widetilde{x}(\cdot), \widetilde{a}(\cdot), \widetilde{b}(\cdot))$ satisfies the differential inclusion (3.8) with

$$C(t) = \widetilde{C}(t) := \big\{x \in \mathbb{R}^n \mid \langle \widetilde{a}_i(t), x\rangle \le \widetilde{b}_i(t),\ i = 1, \ldots, m\big\} \quad \text{for all } t \in [0, T].$$

Taking into account the convexity of the norm function and hence its lower semicontinuity in the L^2-weak topology, we get by passing to the limit in (5.5) and (5.6), respectively, that

$$\int_0^T \left\|\big(\widetilde{x}(t), \widetilde{a}(t), \widetilde{b}(t), \widetilde{u}(t)\big) - \big(\bar{x}(t), \bar{a}(t), \bar{b}(t), \bar{u}(t)\big)\right\|^2 dt$$

$$\le \liminf_{k\to\infty} \sum_{j=0}^{\nu(k)-1} \int_{t_j^k}^{t_{j+1}^k} \left\|\big(x_j^k, a_j^k, b_j^k, u_j^k\big) - \big(\bar{x}(t), \bar{a}(t), \bar{b}(t), \bar{u}(t)\big)\right\|^2 dt \le \frac{\varepsilon}{2},$$

$$\int_0^T \left\|\Big(\dot{\widetilde{x}}(t), \dot{\widetilde{a}}(t), \dot{\widetilde{b}}(t)\Big) - \Big(\dot{\bar{x}}(t), \dot{\bar{a}}(t), \dot{\bar{b}}(t)\Big)\right\|^2 dt$$

$$\le \liminf_{k\to\infty} \sum_{j=0}^{\nu(k)-1} \int_{t_j^k}^{t_{j+1}^k} \left\|\left(\frac{x_{j+1}^k - x_j^k}{h_j^k}, \frac{a_{j+1}^k - a_j^k}{h_j^k}, \frac{b_{j+1}^k - b_j^k}{h_j^k}\right) - \Big(\dot{\bar{x}}(t), \dot{\bar{a}}(t), \dot{\bar{b}}(t)\Big)\right\|^2 dt \le \frac{\varepsilon}{2}$$

This implies that the limiting quadruple $(\widetilde{x}(\cdot), \widetilde{a}(\cdot), \widetilde{b}(\cdot), \widetilde{u}(\cdot))$ belongs to the given ε-neighborhood of the r.i.l.m. $(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot))$ in the space $W^{1,2}([0,T];\mathbb{R}^{n+mn+m}) \times L^2([0,T];\mathbb{R}^d)$. It is clear furthermore that $\widetilde{a}(\cdot)$ satisfies the pointwise constraint (1.6). Applying now Theorem 4.1 to the r.i.l.m. $\big(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot)\big)$ gives us a sequence $(x^k(\cdot), a^k(\cdot), b^k(\cdot), u^k(\cdot))$ of the extended feasible solutions to (P_k) such that $x^k(\cdot), a^k(\cdot), b^k(\cdot)$ and $u^k(\cdot)$ strongly approximate $\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot)$ and $\bar{u}(\cdot)$ in $W^{1,2}([0,T];\mathbb{R}^{n+mn+m})$ and $L^2([0,T];\mathbb{R}^d)$ respectively. It then follows from the imposed convexity of $\widehat{\ell}_F$ and the optimality of $(\bar{x}^k(\cdot), \bar{a}^k(\cdot), \bar{b}^k(\cdot), \bar{u}^k(\cdot))$ to (P_k) that

$$\begin{aligned}
&\widehat{J}\big[\widetilde{x}, \widetilde{a}, \widetilde{b}, \widetilde{u}\big] + \frac{\gamma}{2} = \varphi\big(\widetilde{x}(T)\big) + \int_0^T \widehat{\ell}_F\left(t, \widetilde{x}(t), \widetilde{a}(t), \widetilde{b}(t), \widetilde{u}(t), \dot{\widetilde{x}}(t), \dot{\widetilde{a}}(t), \dot{\widetilde{b}}(t)\right) dt + \frac{\gamma}{2}\\
&\le \liminf_{k\to\infty}\left[\varphi\left(\bar{x}_{\nu(k)}^k\right) + h_k \sum_{j=0}^{\nu(k)-1} \ell\left(t_j^k, \bar{x}_j^k, \bar{a}_j^k, \bar{b}_j^k, u_j^k, \frac{\bar{x}_{j+1}^k - \bar{x}_i^k}{h_j^k}, \frac{\bar{a}_{j+1}^k - \bar{a}_j^k}{h_j^k}, \frac{\bar{b}_{j+1}^k - \bar{b}_j^k}{h_j^k}\right) + \frac{\gamma}{2}\right]\\
&= \liminf_{k\to\infty} J_k\left[\bar{x}^k, \bar{a}^k, \bar{b}^k, \bar{u}^k\right] \le \liminf_{k\to\infty} J_k\left[x^k, a^k, b^k, u^k\right],
\end{aligned} \tag{5.10}$$

which ensures, in particular, that the quadruple $(\widetilde{x}, \widetilde{a}, \widetilde{b}, \widetilde{u})$ is feasible for the relaxed problem (R). On the other hand, the strong convergence of $(x^k(\cdot), a^k(\cdot), b^k(\cdot), u^k(\cdot))$ to $\big(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot)\big)$ in $W^{1,2}([0,T];\mathbb{R}^{n+mn+n}) \times L^2([0,T];\mathbb{R}^d)$ from Theorem 4.1 and the imposed continuity assumptions on φ and ℓ imply that $J_k\left[x^k, a^k, b^k, u^k\right] \to J[\bar{x}, \bar{a}, \bar{b}, \bar{u}]$ as $k \to \infty$. Combining it with (5.10) tells us that

$$\widehat{J}\big[\widetilde{x}, \widetilde{a}, \widetilde{b}, \widetilde{u}\big] < \widehat{J}\big[\widetilde{x}, \widetilde{a}, \widetilde{b}, \widetilde{u}\big] + \frac{\gamma}{2} \le J\left[\bar{x}, \bar{a}, \bar{b}, \bar{u}\right] = \widehat{J}\left[\bar{x}, \bar{a}, \bar{b}, \bar{u}\right],$$

which clearly contradicts the fact that $(\bar{x}(\cdot), \bar{a}(\cdot), \bar{b}(\cdot), \bar{u}(\cdot))$ is an r.i.l.m. for problem (P) and hence verifies the limiting condition (5.9). This completes the proof of the theorem. □

6. Generalized differentiation and second-order calculations

Having in hands the strong approximation results of Theorem 5.2, our subsequent goal is to derive necessary optimality conditions for the discrete-time approximating problems (P_k) that provide constructive suboptimality conditions for the original sweeping control problem (P). Looking at problem (P_k) for each fixed number $k \in I\!N$, we see that it is a finite-dimensional optimization problem with various types of constraints. The most important and challenging of these constraints, that are characteristic for sweeping differential and finite-difference inclusions, are described by *graphs* of *normal cone mappings*. Such sets are *nonconvex* regardless of the convexity and/or smoothness of the given data of (P). To deal with the problems under consideration, we need to employ appropriate constructions of *generalized differentiation* in variational analysis with paying the major attention to *second-order* ones. This section briefly reviews the concepts and results of generalized differentiation used in what follows. We are mainly based on [29], while related first-order constructions can be also found in [34].

Recall that for a set-valued (in particular, single-valued) mapping $S\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ the symbol

$$\mathop{\mathrm{Lim\,sup}}_{x\to\bar{x}} S(x) := \left\{z \in \mathbb{R}^m \,\middle|\, \exists \text{ sequences } x_k \to \bar{x},\ z_k \to z \text{ such that } z_k \in S(x_k),\ k \in I\!N\right\} \tag{6.1}$$

signifies the (Kuratowski-Painlevé) *outer limit* of S at $\bar{x}$. Given a nonempty set $\Omega \subset \mathbb{R}^n$ locally closed around $\bar{x} \in \Omega$, the (Mordukhovich basic/limiting) *normal cone* to Ω at $\bar{x}$ is defined via the outer limit (6.1) by

$$N(\bar{x};\Omega) = N_\Omega(\bar{x}) := \mathop{\mathrm{Lim\,sup}}_{x\to\bar{x}} \left\{\mathrm{cone}\big[x - \Pi(x;\Omega)\big]\right\}, \tag{6.2}$$

where $\Pi(\bar{x};\Omega)$ stands for the Euclidean projection of $\bar{x}$ onto Ω and is defined by

$$\Pi(\bar{x};\Omega) := \left\{y \in \Omega \,\middle|\, \|\bar{x} - y\| = d(\bar{x};\Omega)\right\},$$

and where ‘cone’ denotes the conic hull of a set. If Ω is convex, the limiting normal cone (6.2) reduces to the normal cone of convex analysis (1.2), but in general this cone is nonconvex. Nevertheless, in vast generality the normal cone (6.2) as well as the associated subdifferential and coderivative constructions enjoy comprehensive *calculus rules* based on variational and extremal principles of variational analysis; see [28,29,34] for more details.

Given further a set-valued mapping $S\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ whose graph

$$\mathrm{gph}\, S := \left\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m \,\middle|\, y \in S(x)\right\}$$

is locally closed around $(\bar{x},\bar{y})$, the *coderivative* of S at $(\bar{x},\bar{y})$ is defined by

$$D^*S(\bar{x},\bar{y})(u) := \left\{v \in \mathbb{R}^n \,\middle|\, (v,-u) \in N\big((\bar{x},\bar{y});\mathrm{gph}\, S\big)\right\}, \quad u \in \mathbb{R}^m. \tag{6.3}$$

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

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

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

For an extended-real-valued l.s.c. function $\phi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ with the domain and epigraph defined by

$$\operatorname{dom}\phi := \big\{x \in \mathbb{R}^n \,\big|\, \phi(x) < \infty\big\} \text{ and } \operatorname{epi}\phi := \big\{(x,\alpha) \in \mathbb{R}^{n+1} \,\big|\, \alpha \geq \phi(x)\big\},$$

the *first-order subdifferential* of ϕ at $\bar{x} \in \operatorname{dom}\phi$ is generated geometrically via (6.2) as

$$\partial\phi(\bar{x}) := \big\{v \in \mathbb{R}^n \,\big|\, (v,-1) \in N\big((\bar{x},\phi(\bar{x})); \operatorname{epi}\phi\big)\big\};$$

see [28,29,34] for equivalent analytic representations. The *second-order subdifferential*, or *generalized Hessian*, of ϕ at $\bar{x}$ relative to $\bar{v} \in \partial\phi(\bar{x})$ is the mapping $\partial^2\phi(\bar{x},\bar{v})\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with the values

$$\partial^2\phi(\bar{x},\bar{v})(u) := \big(D^*\partial\phi\big)(\bar{x},\bar{v})(u), \quad u \in \mathbb{R}^n. \tag{6.4}$$

If ϕ is a $\mathcal{C}^2$-smooth around $\bar{x}$, then (6.4) with $\bar{v} = \nabla\phi(\bar{x})$ reduces to the classical (symmetric) Hessian matrix:

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

Our main interest in this paper corresponds to the case where $\phi(x) = \delta_\Omega(x)$ is the indicator function of a set that equals to 0 for $x \in \Omega$ and ∞ otherwise. In this case we have $\partial\delta_\Omega(\bar{x}) = N_\Omega(\bar{x})$ whenever $\bar{x} \in \Omega$. The following result presents evaluations of the coderivative (6.3) of the normal cone mapping

$$G\colon \mathbb{R}^n \times \mathbb{R}^{mn} \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n \text{ with } G(x,a,b) := N\big(x; C(a,b)\big) \tag{6.5}$$

associated with the moving polyhedral set (1.5). In fact, we get an efficient upper estimate of the coderivative under PLICQ (2.4) and its precise calculation under the corresponding LICQ. The proof of this result, given in [14, Lemmas 4.1 and 4.2], is based on the second-order calculus obtained in [30] and the seminal theorem by Robinson [33] on the upper Lipschitzian stability of polyhedral multifunctions. To proceed, consider the matrix $A := [a_{ij}]$ as $i = 1,\ldots,m$ and $j = 1,\ldots,n$ with the vector columns a_i, $i = 1,\ldots,n$. Recall that the symbol $^\perp$ indicates the orthogonal complement of a vector in the space in question.

Lemma 6.1 *(coderivative evaluations of the normal cone mapping). Let G be defined in (6.5) with $x \in C(a,b)$ for $(x,a,b) \in \mathbb{R}^n \times \mathbb{R}^{mn} \times \mathbb{R}^m$, and let $v \in G(x,a,b)$. Suppose that the active constraint vectors $\{a_i \,|\, i \in I(x,a,b)\}$ are positively linearly independent. Then we have the coderivative upper estimate*

$$D^*G(x,a,b,v)(w)$$
$$\subset \bigcup\left\{\left(\begin{array}{c} A^*q \\ \hline p_1w + q_1x \\ \vdots \\ p_mw + q_mx \\ \hline -q \end{array}\right) \,\middle|\, p \in N_{\mathbb{R}^m_-}(Ax-b),\ A^*p = v,\ q \in D^*N_{\mathbb{R}^m_-}(Ax-b,p)(Aw)\right\}$$

for $w \in \bigcap_{\{i \mid p_i > 0\}} a_i^{\perp} \subset \mathbb{R}^n$ *and* $D^*G(x,a,b,v)(w) = \emptyset$ *otherwise.*

If the active constraint vectors $\{a_i \mid i \in I(x,a,b)\}$ *are linearly independent, then we have the precise formula*

$$D^*G(x,a,b,v)(w) = \bigcup_{q \in D^*N_{\mathbb{R}^m_-}(Ax-b,p)(Aw)} \begin{pmatrix} A^*q \\ \hline p_1 w + q_1 x \\ \vdots \\ p_m w + q_m x \\ \hline -q \end{pmatrix} \text{ for all } w \in \bigcap_{\{i \mid p_i > 0\}} a_i^{\perp},$$

where the vector $p \in N_{\mathbb{R}^m_-(Ax-b)}$ *is uniquely determined by* $A^*p = v$. *Furthermore, the coderivative of the normal cone mapping* (6.5) *generated by the nonpositive orthant* $\mathbb{R}^m_-$ *is computed by*

$$D^*N_{\mathbb{R}^m_-}(x,v)(w) = \begin{cases} \emptyset & \text{if } \exists\, i \text{ with } v_i w_i \neq 0 \\ \{\gamma \mid \gamma_i = 0\, \forall\, i \in I_1(w),\ \gamma_i \geq 0\, \forall\, i \in I_2(w)\} & \text{otherwise} \end{cases}, \tag{6.6}$$

whenever $(x,v) \in \operatorname{gph} N_{\mathbb{R}^m_-}$ *with the index subsets in* (6.6) *defined by*

$$I_1(w) := \{i \mid x_i < 0\} \cup \{i \mid v_i = 0,\ w_i < 0\}, \quad I_2(w) := \{i \mid x_i = 0,\ v_i = 0,\ w_i > 0\}. \tag{6.7}$$

The following theorem, which is strongly used in deriving necessary optimality conditions in the next section, provides constructive evaluations of the coderivative of the sweeping control mapping F taken from (3.5) entirely in terms of the given problem data.

Theorem 6.2 *(coderivative evaluations of the sweeping control mapping). Consider the multifunction* F *from* (3.5) *with the polyhedral set* C *defined in* (1.5)*, where the perturbation mapping* $g(x,u)$ *is* $\mathcal{C}^1$*-smooth around the reference points, and where* G *is defined in* (6.5)*. Suppose that the vectors* $\{a_i \mid i \in I(x,a,b)\}$ *are positively linearly independent for any triple* $(x,a,b) \in \mathbb{R}^n \times \mathbb{R}^{mn} \times \mathbb{R}^m$*. Then for all such triples and all* $(w,u) \in \mathbb{R}^m \times U$ *with* $w + g(x,u) \in G(x,a,b)$ *we have the coderivative upper estimate*

$$D^*F(x,a,b,u,w)(y) \subset \bigcup_{\substack{p \in N_{\mathbb{R}^m_-}(Ax-b),\ A^*p = w + g(x,u) \\ q \in D^*N_{\mathbb{R}^m_-}(Ax-b,p)(Ay)}} \begin{pmatrix} A^*q - \nabla_x g(x,u)^* y \\ \hline p_1 y + q_1 x \\ \vdots \\ p_m y + q_m x \\ \hline -q \\ -\nabla_u g(x,u)^* y \end{pmatrix} \tag{6.8}$$

for any $y \in \bigcap_{\{i \mid p_i > 0\}} a_i^{\perp}$*, where the vector* $q \in \mathbb{R}^m$ *satisfies the conditions*

$$\begin{cases} q_i = 0 \text{ for all } i \text{ such that either } \langle a_i, x\rangle < b_i \text{ or } p_i = 0, \text{ or } \langle a_i, y\rangle < 0, \\ q_i \geq 0 \text{ for all } i \text{ such that } \langle a_i, x\rangle = b_i,\ p_i = 0, \text{ and } \langle a_i, y\rangle > 0. \end{cases} \tag{6.9}$$

Furthermore, the equality holds in (6.8) if the vectors $\{a_i \,|\, i \in I(x,a,b)\}$ are linearly independent in which case the vector $p \in N_{\mathbb{R}^m_-(Ax-b)}$ is uniquely determined by $A^*p = w + g(x,u)$.

Proof. Pick any $y \in \bigcap_{\{i\,|\,p_i>0\}} a_i^\perp$ and any $z^* \in D^*F(x,a,b,u,w)(y)$. It follows from the coderivative sum rules of the equality type given in [29, Theorem 3.9] that

$$z^* \in \begin{pmatrix} -\nabla_x g(x,u)^* y \\ \hline 0 \\ \vdots \\ 0 \\ \hline -\nabla_u g(x,u)^* y \end{pmatrix} y + \left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right) D^*G\big(x,a,b,w+g(x,u)\big)(y).$$

Employing further Lemma 6.1 tells us that

$$z^* \in \begin{pmatrix} -\nabla_x g(x,u)^* y \\ \hline 0 \\ \vdots \\ 0 \\ \hline -\nabla_u g(x,u)^* y \end{pmatrix} y + \left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right) \begin{pmatrix} A^*q \\ \hline p_1 y + q_1 x \\ \vdots \\ p_m y + q_m x \\ \hline -q \end{pmatrix} = \begin{pmatrix} A^*q - \nabla_x g(x,u)^* y \\ \hline p_1 y + q_1 x \\ \vdots \\ p_m y + q_m x \\ \hline -q \\ -\nabla_u g(x,u)^* y \end{pmatrix}$$

for some $p \in N_{\mathbb{R}^m_-}(Ax-b)$ with $A^*p = w + g(x,u)$ and $q \in D^*N_{\mathbb{R}^m_-}(Ax-b,\, p)(Ay)$. Finally, conditions (6.9) for the vector q follow from (6.6) and (6.7). This completes the proof of the theorem. $\square$

7. Optimality conditions via discrete approximations

This section is devoted to deriving necessary optimality conditions for each discrete-time problem (P_k) as $k \in I\!N$. As followed from Theorem 5.2, the results obtained below give us suboptimality conditions for the selected r.i.l.m. of the original sweeping optimal control problem (P) provided that the discretization index k is sufficiently large.

We establish here two results in this direction. The first theorem provides necessary optimality conditions to each problem (P_k) defined in Section 5 that are expressed in terms of the normal cone to the graph of the velocity mapping F from (3.5), i.e., via the coderivative of this mapping. The second theorem is the main result of this section. It derives verifiable necessary conditions for the given r.i.l.m. of problem (P) expressed entirely in terms of the initial data of the original sweeping control problem along the strongly converging sequence of optimal solutions to the discrete approximation problems (P_k).

Let us start with the first result, which proof is based on the reduction of (P_k) to nonsmooth finite-dimensional mathematical programming with increasingly many geometric constraints and employing calculus rules of first-order generalized differentiation. As seen below, the proof of the main result is largely based on second-order calculations. For convenience we use the notation $\mathrm{rep}_m(x) := (x,\ldots,x) \in \mathbb{R}^{mn}$.

Theorem 7.1 *(necessary conditions for discrete optimal solutions). Fix any $k \in I\!N$ and let*

$$(\bar{x}^k, \bar{a}^k, \bar{b}^k, \bar{u}^k) = (\bar{x}_0^k, \ldots, \bar{x}_{\nu(k)}^k, \bar{a}_0^k, \ldots, \bar{a}_{\nu(k)}^k, \bar{b}_0^k, \ldots, \bar{b}_{\nu(k)}^k, \bar{u}_0^k, \ldots, \bar{u}_{\nu(k)-1}^k)$$

be an optimal solution to (P_k) along which the general assumptions of Theorem 6.2 are fulfilled. Suppose in addition that the cost functions φ and ℓ are locally Lipschitzian around the corresponding components of the optimal solution. Then there exist a number $\lambda^k \geq 0$ and vectors $\alpha^{1k} = \left(\alpha_0^{1k}, \ldots, \alpha_{\nu(k)}^{1k}\right) \in \mathbb{R}_+^{(\nu(k)+1)m}$, $\psi^k = (\psi_0^k, \ldots, \psi_{\nu(k)-1}^k) \in \mathbb{R}^{\nu(k)d}$, $\alpha^{2k} = \left(\alpha_0^{2k}, \ldots, \alpha_{\nu(k)}^{2k}\right) \in \mathbb{R}_-^{(\nu(k)+1)m}$, $\xi^k = (\xi_1^k, \ldots, \xi_m^k) \in \mathbb{R}_+^m$, and $p_j^k = \left(p_j^{xk}, p_j^{ak}, p_j^{bk}\right) \in \mathbb{R}^{n+mn+m}$ as $j = 0, \ldots, \nu(k)$ satisfying the relationships:

$$\lambda^k + \left\|\xi^k\right\| + \|\alpha^{1k} + \alpha^{2k}\| + \sum_{j=0}^{\nu(k)-1} \left\|p_j^{xk}\right\| + \|p_0^{ak}\| + \|p_0^{bk}\| + \left\|\psi^k\right\| \neq 0, \tag{7.1}$$

$$\xi_i^k \left(\left\langle \bar{a}_{ik}^k, \bar{x}_k^k \right\rangle - \bar{b}_{ik}^k\right) = 0, \ \ i = 1, \ldots, m, \tag{7.2}$$

$$\alpha_{ij}^{1k} \left(\|\bar{a}_{ij}^k\| - (1 + \delta_k)\right) = 0, \ \ i = 1, \ldots, m, \ \ j = 0, \ldots, \nu(k), \tag{7.3}$$

$$\alpha_{ij}^{2k} \left(\|\bar{a}_{ij}^k\| - (1 - \delta_k)\right) = 0, \ i = 1, \ldots, m, \ \ j = 0, \ldots, \nu(k), \tag{7.4}$$

$$-p_{\nu(k)}^{xk} = \lambda^k v_{\nu(k)}^k + \sum_{i=1}^m \xi_i^k \bar{a}_{i\nu(k)}^k \in \lambda^k \partial\varphi(\bar{x}_{\nu(k)}^k) + \sum_{i=1}^m \xi_i^k \bar{a}_{i\nu(k)}^k, \tag{7.5}$$

$$p_{\nu(k)}^{ak} = -2\left[\alpha_{\nu(k)}^{1k} + \alpha_{\nu(k)}^{2k}, \bar{a}_{i\nu(k)}^k\right] - \left[\xi^k, \mathrm{rep}_m(\bar{x}_{\nu(k)}^k)\right], \tag{7.6}$$

$$p_{\nu(k)}^{bk} = \xi^k, \tag{7.7}$$

$$p_{j+1}^{ak} = \lambda^k \left(v_j^{ak} + \frac{1}{h_j^k}\theta_j^{Ak}\right), \ \ p_{j+1}^{bk} = \lambda^k \left(v_j^{bk} + \frac{1}{h_j^k}\theta_j^{Bk}\right), \ \ j = 0, \ldots, \nu(k) - 1, \tag{7.8}$$

$$\begin{aligned} &\left(\frac{p_{j+1}^{xk} - p_j^{xk}}{h_j^k} - \lambda^k w_j^{xk}, \frac{p_{j+1}^{ak} - p_j^{ak}}{h_j^k} - \lambda^k w_j^{ak}, \frac{p_{j+1}^{bk} - p_j^{bk}}{h_j^k} - \lambda^k w_j^{bk}, -\frac{1}{h_k}\lambda^k \theta_j^{uk}\right. \\ &\left. - \lambda^k w_j^{uk}, -p_{j+1}^{xk} + \lambda^k\left(v_j^{xk} + \frac{1}{h_j^k}\theta_j^{Xk}\right)\right) \in \left(0, \frac{2}{h_j^k}\left[\alpha_j^{1k} + \alpha_j^{2k}, \bar{a}_j^k\right], 0, \frac{1}{h_j^k}\psi_j^k, 0\right) \\ &+ N\left(\left(\bar{x}_j^k, \bar{a}_j^k, \bar{b}_j^k, \bar{u}_j^k, -\frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{h_j^k}\right); \mathrm{gph}\, F\right), \ \ j = 0, \ldots, \nu(k) - 1, \end{aligned} \tag{7.9}$$

$$\psi_j^k \in N\left(u_j^k; U\right), \ \ j = 0 \ldots, \nu(k) - 1, \tag{7.10}$$

where the quadruple $(\theta_j^{uk}, \theta_j^{Xk}, \theta_j^{Ak}, \theta_j^{Bk})$ is defined by

$$\left(\int_{t_j^k}^{t_{j+1}^k}\Big(\bar{u}_j^k-\bar{u}(t)\Big)\,dt,\ \int_{t_j^k}^{t_{j+1}^k}\left(\frac{\bar{x}_{j+1}^k-\bar{x}_j^k}{h_j^k}-\dot{\bar{x}}(t)\right)dt,\ \int_{t_j^k}^{t_{j+1}^k}\left(\frac{\bar{a}_{j+1}^k-\bar{a}_j^k}{h_j^k}-\dot{\bar{a}}(t)\right)dt,\right.$$
$$\left.\int_{t_j^k}^{t_{j+1}^k}\left(\frac{\bar{b}_{j+1}^k-\bar{b}_j^k}{h_j^k}-\dot{\bar{b}}(t)\right)dt\right)$$

with the running cost subgradient collections

$$\Big(w_j^{xk},w_j^{ak},w_j^{bk},w_j^{uk},v_j^{xk},v_j^{ak},v_j^{bk}\Big)\in\partial\ell\left(\bar{x}_j^k,\bar{a}_j^k,\bar{b}_j^k,\bar{u}_j^k,\frac{\bar{x}_{j+1}^k-\bar{x}_j^k}{h_j^k},\frac{\bar{a}_{j+1}^k-\bar{a}_j^k}{h_j^k},\frac{\bar{b}_{j+1}^k-\bar{b}_j^k}{h_j^k}\right).\tag{7.11}$$

Proof. Take $\varepsilon>0$ from the definition of the r.i.l.m. $(\bar{x}(\cdot),\bar{a}(\cdot),\bar{b}(\cdot),\bar{u}(\cdot))$ in problem (P_k) for any fixed $k\in I\!N$ and define the "long" vector reflecting the set of feasible solutions to each discrete-time problem (P_k) by

$$z:=\Big(x_0^k,\ldots,x_{\nu(k)}^k,a_0^k,\ldots,a_{\nu(k)}^k,b_0^k,\ldots,b_{\nu(k)}^k,u_0^k,\ldots,u_{\nu(k)-1}^k,X_0^k,\ldots,X_{\nu(k)-1}^k,A_0^k,\ldots,$$
$$A_{\nu(k)-1}^k,B_0^k,\ldots,B_{\nu(k)-1}^k\Big)$$

with the fixed starting point as in (5.4). It is clear that each problem (P_k) can be equivalently written as the nondynamic problem of mathematical programming (MP) with respect to vector z:

$$\begin{aligned}&\text{minimize } \phi_0(z):=\varphi\big(x(T)\big)+\sum_{j=0}^{\nu(k)-1}h_j^k\ell(x_j^k,a_j^k,b_j^k,u_j^k,X_j^k,A_j^k,B_j^k)\\&+\frac{1}{2}\sum_{j=0}^{\nu(k)-1}\int_{t_j^k}^{t_{j+1}^k}\Big\|\Big(X_j^k-\dot{\bar{x}}(t),A_j^k-\dot{\bar{a}}(t),B_j^k-\dot{\bar{b}}(t),u_j^k-\bar{u}(t)\Big)\Big\|^2dt\end{aligned}\tag{7.12}$$

subject to finitely many equality, inequality, and geometric constraints

$$\kappa(z):=\sum_{j=0}^{\nu(k)-1}\int_{t_j^k}^{t_{j+1}^k}\Big\|\Big(x_j^k,a_j^k,b_j^k,u_j^k\Big)-\big(\bar{x}(t),\bar{a}(t),\bar{b}(t),\bar{u}(t)\big)\Big\|^2dt\le\frac{\varepsilon}{2},\tag{7.13}$$

$$\phi(z):=\sum_{j=0}^{\nu(k)-1}\int_{t_j^k}^{t_{j+1}^k}\Big\|\Big(X_j^k,A_j^k,B_j^k,u_j^k\Big)-(\dot{\bar{x}}(t),\dot{\bar{a}}(t),\dot{\bar{b}}(t),\bar{u}(t))\Big\|^2dt-\frac{\varepsilon}{2}\le 0,\tag{7.14}$$

$$g_j^x(z) := x_{j+1}^k - x_j^k - h_j^k X_j^k = 0,\ j = 0, \ldots, \nu(k) - 1, \tag{7.15}$$

$$g_j^a(z) := a_{j+1}^k - a_j^k - h_j^k A_j^k = 0,\ j = 0, \ldots, \nu(k) - 1, \tag{7.16}$$

$$g_j^b(z) := b_{j+1}^k - b_j^k - h_j^k B_j^k = 0,\ j = 0, \ldots, \nu(k) - 1, \tag{7.17}$$

$$q_i(z) := \left\langle a_{i\nu(k)}^k, x_{\nu(k)}^k \right\rangle - b_{i\nu(k)}^k \le 0,\ i = 1, \ldots, m, \tag{7.18}$$

$$l_{ij}^1(z) := \left\| a_{ij}^k \right\|^2 - (1+\delta_k)^2 \le 0,\ i = 1, \ldots, m,\ j = 0, \ldots, \nu(k), \tag{7.19}$$

$$l_{ij}^2(z) := \left\| a_{ij}^k \right\|^2 - (1-\delta_k)^2 \ge 0,\ i = 1, \ldots, m,\ j = 0, \ldots, \nu(k), \tag{7.20}$$

$$z \in \Xi_j := \left\{ z \middle|\ -X_j^k \in F(x_j^k, a_j^k, b_j^k, u_j^k) \right\},\ j = 0, \ldots, \nu(k) - 1, \tag{7.21}$$

$$z \in \Xi_{\nu(k)} := \left\{ z \middle|\ x_0^k \text{ is fixed},\ (a_0, b_0, u_0) = \left(\bar{a}(0), \bar{b}(0), \bar{u}(0)\right) \right\}, \tag{7.22}$$

$$z \in \Omega_j = \left\{ z \middle|\ u_j^k \in U \right\},\quad j = 0, \ldots, \nu(k) - 1. \tag{7.23}$$

Next we apply the necessary conditions from [28, Proposition 6.4(ii) and Theorem 6.5(ii)] to the optimal solution

$$\bar{z} := \big(\bar{x}_0^k, \ldots, \bar{x}_{\nu(k)}^k, \bar{a}_0^k, \ldots, \bar{a}_{\nu(k)}^k, \bar{b}_0^k, \ldots, \bar{b}_{\nu(k)}^k, \bar{u}_0^k, \ldots, \bar{u}_{\nu(k)-1}^k, \bar{X}_0^k, \ldots, \bar{X}_{\nu(k)-1}^k, \bar{A}_0^k, \ldots, \bar{A}_{\nu(k)-1}^k, \bar{B}_0^k, \ldots, \bar{B}_{\nu(k)-1}^k\big)$$

of problem (MP) in (7.12)–(7.23) corresponding to the one for (P_k) given in the theorem. It follows from Theorem 5.2 that the inequality constraints in (7.13) and (7.14) are inactive for large k, and so the corresponding multipliers do not appear in the necessary optimality conditions. Thus we find dual elements $\lambda^k \ge 0$, $\xi^k = (\xi_1^k, \ldots, \xi_m^k) \in \mathbb{R}_+^m$, $\alpha^{1k} = (\alpha_0^{1k}, \ldots, \alpha_{\nu(k)}^{1k}) \in \mathbb{R}_+^{\nu(k)+1}$, $\alpha^{2k} = (\alpha_0^{2k}, \ldots, \alpha_{\nu(k)}^{2k}) \in \mathbb{R}_-^{\nu(k)+1}$, $p_j^k = \left(p_j^{xk}, p_j^{ak}, p_j^{bk}\right) \in \mathbb{R}^{n+mn+m}$ for $j = 1, \ldots, \nu(k)$, and $z_j^* = \big(x_{0j}^*, \ldots, x_{\nu(k)j}^*,\ a_{0j}^*, \ldots, a_{\nu(k)j}^*,\ b_{0j}^*, \ldots, b_{\nu(k)j}^*,\ u_{0j}^*, \ldots, u_{(\nu(k)-1)j}^*,\ X_{0j}^*, \ldots, X_{(\nu(k)-1)j}^*,\ A_{0j}^*, \ldots, A_{(\nu(k)-1)j}^*,\ B_{0j}^*, \ldots, B_{(\nu(k)-1)j}^*\big)$ for $j = 0, \ldots, \nu(k)$, which are not zero simultaneously, such that the following relationships are satisfied:

$$z_j^* \in \begin{cases} N(\bar{z}, \Xi_j) + N(\bar{z}, \Omega_j) & \text{if} \quad j \in \{0, \ldots, \nu(k) - 1\} \\ N(\bar{z}, \Xi_j) & \text{if} \quad j = \nu(k) \end{cases}, \tag{7.24}$$

$$\begin{aligned} -z_0^* - \ldots - z_{\nu(k)}^* \in{}& \lambda^k \partial\phi_0(\bar{z}) + \sum_{i=1}^m \xi_i^k \nabla q_i(\bar{z}) + \sum_{j=0}^{\nu(k)} \sum_{i=1}^m \alpha_{ij}^{1k} \nabla l_{ij}^1(\bar{z}) \\ &+ \sum_{j=0}^{\nu(k)} \sum_{i=1}^m \alpha_{ij}^{2k} \nabla l_{ij}^2(\bar{z}) + \sum_{j=0}^{\nu(k)-1} (\nabla g_j(\bar{z}))^* p_{j+1}^k, \end{aligned} \tag{7.25}$$

$$\xi_i^k q_i(\bar{z}) = 0,\ \ i = 1, \ldots, m, \tag{7.26}$$

$$\alpha_{ij}^{1k} \left(\left\| a_{ij}^k \right\| - (1+\delta_k) \right) = 0,\ i = 1, \ldots, m,\ \ j = 0, \ldots, \nu(k), \tag{7.27}$$

$$\alpha_{ij}^{2k}\left(\left\|a_{ij}^k\right\|-(1-\delta_k)\right)=0,\ i=1,\dots,m,\ j=0,\dots,\nu(k). \tag{7.28}$$

Note that the first line in (7.24) comes from applying the normal cone intersection rule from [29, Theorem 2.16] to $\bar{z}\in\Omega_j\cap\Xi_j$ for $j=0,\dots,\nu(k)-1$, where the qualification condition

$$N(\bar{z};\Xi_j)\cap\big(-N(\bar{z};\Omega_j)\big)=\{0\},\quad j=0,\dots,\nu(k)-1, \tag{7.29}$$

imposed therein is fulfilled. Indeed, for any vector $z_j^*\in N(\bar{z};\Xi_j)\cap(-N(\bar{z};\Omega_j))$ we clearly have the inclusions

$$(x_{jj}^*,a_{jj}^*,b_{jj}^*,u_{jj}^*,-X_{jj}^*)\in N\Big(\Big(\bar{x}_j^k,\bar{a}_j^k,\bar{b}_j^k,\bar{u}_j^k,-\frac{\bar{x}_j^{k+1}-\bar{x}_j^{k+1}}{h_k}\Big);\operatorname{gph}F\Big),\quad -u_{ij}^*\in N(\bar{u}_j^k;U), \tag{7.30}$$

while the other components of z_j^* are zero. It immediately follows from (7.30) that

$$x_{jj}^*=0,\ a_{jj}^*=0,\ b_{jj}^*=0,\ \text{ and }\ X_{jj}^*=0.$$

Substituting this into the first inclusion in (7.30) and using the coderivative definition (6.3) give us

$$(0,0,0,u_{jj}^*)\in D^*F\Big(\bar{x}_j^k,\bar{a}_j^k,\bar{b}_j^k,\bar{u}_j^k,-\frac{\bar{x}_j^{k+1}-\bar{x}_j^k}{h_k}\Big)(0),\quad j=0,\dots,\nu(k)-1.$$

Then we deduce directly from the coderivative estimate (6.8) for the velocity mapping F in (3.5) under the imposed PLICQ that $u_{jj}^*=0$ for all $j=0,\dots,\nu(k)-1$. It shows that $z_j^*=0$ for such indices j, and therefore the qualification condition (7.29) is verified.

To proceed further, observe from the structure of the sets Ξ_j and Ω_j in (7.21)–(7.23), respectively, that the inclusions in (7.24) are equivalent to

$$\begin{cases}(x_{jj}^*,a_{jj}^*,b_{jj}^*,u_{jj}^*-\psi_j^k,-X_{jj}^*)\in N\Big(\Big(\bar{x}_j^k,\bar{a}_j^k,\bar{b}_j^k,\bar{u}_j^k,-\frac{\bar{x}_{j+1}^k-\bar{x}_j^k}{h_j^k}\Big);\operatorname{gph}F\Big)\\ \quad\text{for }\ j=0,\dots,\nu(k)-1,\\ (x_{jj}^*,a_{jj}^*,b_{jj}^*,u_{jj}^*,-X_{jj}^*)\in N\Big(\Big(\bar{x}_j^k,\bar{a}_j^k,\bar{b}_j^k,\bar{u}_j^k,-\frac{\bar{x}_{j+1}^k-\bar{x}_j^k}{h_j^k}\Big);\operatorname{gph}F\Big)\\ \quad\text{for }\ j=\nu(k)\end{cases} \tag{7.31}$$

with ψ_j^k taken from (7.10), while the other components of z_j^* equal to zero. Similarly we get that the vectors $x_{0\nu(k)}^*, a_{0\nu(k)}^*, b_{0\nu(k)}^*$, and $u_{0\nu(k)}^*$ determined by the normal cone to $\Xi_{\nu(k)}$ might be the only nonzero components of $z_{\nu(k)}^*$. This readily yields the representation

$$\begin{aligned}-z_0^*-\ldots-z_{\nu(k)}^*=\big(&-x_{0\nu(k)}^*-x_{00}^*,-x_{11}^*,\dots,-x_{\nu(k)-1,\nu(k)-1}^*,0,-a_{0\nu(k)}^*-a_{00}^*,-a_{11}^*,\dots,\\ &-a_{\nu(k)-1\,\nu(k)-1}^*,0,-b_{0\nu(k)}^*-b_{00}^*,-b_{11}^*,\dots,-b_{\nu(k)-1\,\nu(k)-1}^*,0,\\ &-u_{0\nu(k)}^*-u_{00}^*,\dots,-u_{\nu(k)-1\,\nu(k)-1}^*,-X_{00}^*,\dots,-X_{\nu(k)-1\,\nu(k)-1}^*,0,\dots,0\big).\end{aligned}$$

Next we represent the right-hand side of the inclusion in (7.25) by

$$\lambda^k\partial\phi_0(\bar z)+\sum_{i=1}^m\xi_i^k\nabla q_i(\bar z)+\sum_{j=0}^{\nu(k)-1}\alpha_{ij}^{1k}\nabla l_{ij}^1(\bar z)+\sum_{j=0}^{\nu(k)-1}\alpha_{ij}^{2k}\nabla l_{ij}^2(\bar z)+\sum_{j=0}^{\nu(k)-1}\nabla g_j(\bar z)^*p_{j+1}^k$$

with the complementary slackness conditions

$$\xi_i^k\left(\left\langle a_{i\nu(k)}^k,x_{\nu(k)}^k\right\rangle-b_{i\nu(k)}^k\right)=0,\ i=1,\dots,m.$$

Unifying the above representations and denoting

$$\rho_j(\bar z):=\int_{t_j^k}^{t_{j+1}^k}\left\|\left(\frac{\bar x_{j+1}^k-\bar x_j^k}{h_j^k}-\dot{\bar x}(t),\frac{\bar a_{j+1}^k-\bar a_j^k}{h_j^k}-\dot{\bar a}(t),\frac{\bar b_{j+1}^k-\bar b_j^k}{h_j^k}-\dot{\bar b}(t),\bar u_j^k(t)-\bar u(t)\right)\right\|^2dt,$$

we arrive at the following relationships:

$$\left(\sum_{i=1}^m\xi_i^k\nabla q_i(\bar z)\right)_{(x_{\nu(k)},a_{\nu(k)},b_{\nu(k)},u_{\nu(k)})}=\left(\sum_{i=1}^m\xi_i^k\bar a_{ik}^k,\left[\xi^k,\mathrm{rep}_m(\bar x_{\nu(k)}^k)\right],-\xi^k,0\right),$$

$$\left(\sum_{j=0}^{\nu(k)}\sum_{i=1}^m\alpha_{ij}^{1k}\nabla l_{ij}^1(\bar z)\right)_{(a_j)}=2\left[\alpha_j^{1k},\bar a_j^k\right],\ \ j=0,\dots,\nu(k)-1,$$

$$\left(\sum_{j=0}^{\nu(k)}\sum_{i=1}^m\alpha_{ij}^{2k}\nabla l_{ij}^2(\bar z)\right)_{(a_j)}=2\left[\alpha_j^{2k},\bar a_j^k\right],\ \ j=0,\dots,\nu(k)-1,$$

$$\left(\sum_{j=0}^{\nu(k)-1}\nabla g_j(\bar z)^*p_{j+1}^k\right)_{(x_j,a_j,b_j)}=\begin{cases}-p_1^k & \text{if}\quad j=0\\ p_j^k-p_{j+1}^k & \text{if}\quad j=1,\dots,\nu(k)-1\\ p_{\nu(k)}^k & \text{if}\quad j=\nu(k)\end{cases},$$

$$\left(\sum_{j=0}^{\nu(k)-1}\nabla g_j(\bar z)^*p_{j+1}^k\right)_{(X_j,A_j,B_j)}=(-h_0^kp_1^{xk},-h_1^kp_2^{xk},\dots,-h_{\nu(k)-1}^kp_{\nu(k)}^{xk},$$
$$-h_0^kp_1^{ak},-h_1^kp_2^{ak},\dots,-h_{\nu(k)-1}^kp_{\nu(k)}^{ak},-h_0^kp_1^{bk},-h_1^kp_2^{bk},\dots,-h_{\nu(k)-1}^kp_{\nu(k)}^{bk}),$$

$$\partial\phi_0(\bar z)\subset\partial\varphi(\bar x_{\nu(k)}^k)+\sum_{j=0}^{\nu(k)-1}h_j^k\partial\ell\left(\bar x_j^k,\bar a_j^k,\bar b_j^k,\bar u_j^k,\bar X_j^k,\bar A_j^k,\bar B_j^k\right)+\frac{1}{2}\sum_{j=0}^{\nu(k)-1}\nabla\rho_j(\bar z).$$

Furthermore, the set $\lambda^k\partial\phi_0(\bar z)$ is represented as the collection of vectors

$$\lambda^k(h_0^k w_0^{xk}, h_1^k w_1^{xk}, \dots, h_{\nu(k)-1}^k k w_{\nu(k)-1}^{xk}, v_{\nu(k)}^k, h_0^k w_0^{ak}, h_1^k w_1^{ak}, \dots, h_{\nu(k)-1}^k w_{\nu(k)-1}^{ak}, 0,$$
$$h_0^k w_0^{bk}, h_1^k w_1^{bk}, \dots, h_{\nu(k)-1}^k w_{\nu(k)-1}^{bk}, 0, \theta_0^{uk} + h_0^k w_0^{uk}, \theta_1^{uk} + h_1^k w_1^{uk}, \dots,$$
$$\theta_{\nu(k)-1}^{uk} + h_{\nu(k)-1}^k w_{\nu(k)-1}^{uk}, \theta_0^{Xk} + h_0^k v_0^{xk}, \theta_1^{Xk} + h_1^k v_1^{xk}, \dots, \theta_{\nu(k)-1}^{Xk} + h_{\nu(k)-1}^k v_{\nu(k)-1}^{xk},$$
$$\theta_0^{Ak} + h_0^k v_0^{ak}, \theta_1^{Ak} + h_1^k v_1^{ak}, \dots, \theta_{\nu(k)-1}^{Ak} + h_{\nu(k)-1}^k v_{\nu(k)-1}^{ak}, \theta_0^{Bk} + h_0^k v_0^{bk},$$
$$\theta_1^{Bk} + h_1^k v_1^{bk}, \dots, \theta_{\nu(k)-1}^{Bk} + h_{\nu(k)-1}^k v_{\nu(k)-1}^{bk}),$$

where the components above are such that

$$v_{\nu(k)}^k \in \partial\varphi(\bar{x}_{\nu(k)}^k), \quad \text{and}$$

$$\left(w_j^{xk}, w_j^{ak}, w_j^{bk}, w_j^{uk}, v_j^{xk}, v_j^{ak}, v_j^{bk}\right) \in \partial\ell\left(\bar{x}_j^k, \bar{a}_j^k, \bar{b}_j^k, \bar{u}_j^k, \frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{h_j^k}, \frac{\bar{a}_{j+1}^k - \bar{a}_j^k}{h_j^k}, \frac{\bar{b}_{j+1}^k - \bar{b}_j^k}{h_j^k}\right),$$

$$\left(\theta_j^{uk}, \theta_j^{Xk}, \theta_j^{Ak}, \theta_j^{Bk}\right) :=$$
$$\left(\int_{t_j^k}^{t_{j+1}^k} \left(\bar{u}_j^k - \bar{u}(t)\right) dt, \int_{t_j^k}^{t_{j+1}^k} \left(\frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{h_j^k} - \dot{\bar{x}}(t)\right) dt, \int_{t_j^k}^{t_{j+1}^k} \left(\frac{\bar{a}_{j+1}^k - \bar{a}_j^k}{h_j^k} - \dot{\bar{a}}(t)\right) dt,\right.$$
$$\left.\int_{t_j^k}^{t_{j+1}^k} \left(\frac{\bar{b}_{j+1}^k - \bar{b}_j^k}{h_j^k} - \dot{\bar{b}}(t)\right) dt\right)$$

for $j = 0, \dots, \nu(k) - 1$. Unifying all of this gives us the conditions

$$-x_{00}^* - x_{0\nu(k)}^* = \lambda^k h_0^k w_0^{xk} - p_1^{xk}, \tag{7.32}$$
$$-x_{jj}^* = \lambda^k h_j^k w_j^{xk} + p_j^{xk} - p_{j+1}^{xk}, \ j = 1, \dots, \nu(k) - 1, \tag{7.33}$$
$$0 = \lambda^k v_{\nu(k)}^k + p_{\nu(k)}^{xk} + \sum_{i=1}^m \xi_i^k \bar{a}_{ik}^k, \ \text{where} \ v_{\nu(k)}^k \in \partial\varphi(\bar{x}_{\nu(k)}^k), \tag{7.34}$$
$$-a_{00}^* - a_{0\nu(k)}^* = \lambda^k h_0^k w_0^{ak} + 2\left[\alpha_0^{1k} + \alpha_0^{2k}, \bar{a}_0^k\right] - p_1^{ak}, \ i = 1, \dots, m, \tag{7.35}$$
$$-a_{jj}^* = \lambda^k h_j^k w_j^{ak} + 2\left[\alpha_j^{1k} + \alpha_j^{2k}, \bar{a}_j^k\right] + p_j^{ak} - p_{j+1}^{ak}, \ i = 1, \dots, m, \ j = 1, \dots, \nu(k) - 1, \tag{7.36}$$
$$0 = 2\left(\alpha_{\nu(k)}^{1k} + \alpha_{\nu(k)}^{2k}\right)\bar{a}_{i\nu(k)}^k + p_{\nu(k)}^{ak} + \left[\xi^k, \mathrm{rep}_m(\bar{x}_{\nu(k)}^k)\right], \ i = 1, \dots, m, \tag{7.37}$$
$$-b_{00}^* - b_{0\nu(k)}^* = \lambda^k h_0^k w_0^{bk} - p_1^{bk}, \tag{7.38}$$
$$-b_{jj}^* = \lambda^k h_j^k w_j^{bk} + p_j^{bk} - p_{j+1}^{bk}, \ j = 1, \dots, \nu(k) - 1, \tag{7.39}$$
$$0 = p_{\nu(k)}^{bk} - \xi^k, \tag{7.40}$$
$$-u_{00}^* = \lambda^k \theta_0^{uk} + \lambda^k h_0^k w_0^{uk}, \tag{7.41}$$

$$-u^*_{jj} = \lambda^k \theta^{uk}_j + \lambda^k h^k_j w^{uk}_j, \quad j = 1, \ldots, \nu(k) - 1, \tag{7.42}$$

$$-X^*_{jj} = \lambda^k \theta^{Xk}_j + \lambda^k h_k v^{xk}_j - h^k_j p^{xk}_{j+1}, \quad j = 0, \ldots, \nu(k) - 1, \tag{7.43}$$

$$0 = \lambda^k \theta^{Ak}_j + \lambda^k h_k v^{ak}_j - h^k_j p^{ak}_{j+1}, \quad j = 0, \ldots, \nu(k) - 1, \tag{7.44}$$

$$0 = \lambda^k \theta^{Bk}_j + \lambda^k h_k v^{bk}_j - h^k_j p^{bk}_{j+1}, \quad j = 0, \ldots, \nu(k) - 1. \tag{7.45}$$

Now we are ready to justify all the necessary optimality conditions claimed in this theorem. First observe that (7.5), (7.6), and (7.7) follow from (7.34), (7.37), and (7.40), respectively. Next let us extend each vector p^k by adding the zero component $p^k_0 := \left(x^*_{0\nu(k)}, a^*_{0\nu(k)}, b^*_{0\nu(k)}, u^*_{0\nu(k)}\right)$. It follows from the relationships in (7.33), (7.36), (7.39), (7.43), (7.44), and (7.45) that

$$\begin{aligned}
\frac{x^*_{jj}}{h^k_j} &= \frac{p^{xk}_{j+1} - p^{xk}_j}{h^k_j} - \lambda^k w^{xk}_j, \\
\frac{a^*_{jj}}{h^k_j} &= \frac{p^{ak}_{j+1} - p^{ak}_j}{h^k_j} - \lambda^k w^{ak}_j - \frac{2}{h^k_j}\left(\alpha^{1k}_j + \alpha^{2k}_j\right)\bar{a}^k_{ij}, \\
\frac{b^*_{jj}}{h^k_j} &= \frac{p^{bk}_{j+1} - p^{bk}_j}{h^k_j} - \lambda^k w^{bk}_j, \\
\frac{u^*_{jj}}{h^k_j} &= -\frac{1}{h^k_j}\lambda^k \theta^{uk}_j - \lambda^k w^{uk}_j, \\
\frac{X^*_{jj}}{h^k_j} &= -\frac{1}{h^k_j}\lambda^k \theta^{Xk}_j + p^{xk}_{j+1} - \lambda^k v^{xk}_j, \\
0 &= -\frac{1}{h^k_j}\lambda^k \theta^{Ak}_j + p^{ak}_{j+1} - \lambda^k v^{ak}_j, \\
0 &= -\frac{1}{h^k_j}\lambda^k \theta^{Bk}_j + p^{bk}_{j+1} - \lambda^k v^{bk}_j.
\end{aligned}$$

Substituting this into the left-hand side of (7.31) and taking into account the equalities in (7.26)–(7.28), (7.34), (7.37), and (7.40) justify the claims made in (7.2)–(7.9).

To verify finally the nontriviality condition (7.1), suppose on the contrary that $\lambda^k = 0$, $\xi^k = 0$, $\alpha^{1k} + \alpha^{2k} = 0$, $p^{xk}_j = 0$, $p^{ak}_j = 0$, $p^{bk}_j = 0$, $\psi^k = 0$ for all $j = 0, \ldots, \nu(k) - 1$, which yields in turn $x^*_{0k} = p^{xk}_0 = 0$, $a^*_{0k} = p^{ak}_0 = 0$, and $b^*_{0k} = p^{bk}_0 = 0$. Then it follows from (7.34), (7.37), and (7.40) that $\left(p^{xk}_{\nu(k)}, p^{ak}_{\nu(k)}, p^{bk}_{\nu(k)}\right) = 0$, and hence $\left(p^{xk}_j, p^{ak}_j, p^{bk}_j\right) = 0$, for all $j = 0, \ldots, \nu(k)$. We see also that the conditions in (7.32), (7.33), (7.35), (7.36), (7.38), (7.39), (7.41), and (7.42) imply that $\left(x^*_{jj}, a^*_{jj}, b^*_{jj}, u^*_{ij}\right) = 0$ for all $j = 0, \ldots, \nu(k) - 1$. In addition, it follows from (7.43), (7.44), and (7.45) that $X^*_{jj} = 0$, $A^*_{jj} = 0$, $B^*_{jj} = 0$ for all $j = 0, \ldots, \nu(k) - 1$. Furthermore, all the components of z^*_j different from $(x^*_{jj}, a^*_{jj}, b^*_{jj}, u^*_{jj}, X^*_{jj}, A^*_{jj}, B^*_{jj})$ are clearly zero for $j = 0, \ldots, \nu(k) - 1$, and hence $z^*_j = 0$ for $j = 0, \ldots, \nu(k) - 1$. We similarly conclude that $z^*_k = 0$, since $x^*_{0k} = p^{xk}_0 = 0$ while all the other components of this vector obviously reduce to zero. Thus

$z_j^* = 0$ for all $j = 0, \ldots, \nu(k)$, which violates the nontriviality condition for (MP) and completes the proof of the theorem. $\square$

Our next theorem provides verifiable necessary optimality conditions for solutions $(\bar{x}^k, \bar{a}^k, \bar{b}^k, \bar{u}^k)$ to problems (P_k) that strongly approximate the given r.i.l.m. $(\bar{x}, \bar{a}, \bar{b}, \bar{u})$ for the original sweeping control problem (P). The proof is based on the results of Theorem 7.1 and the second-order calculations from Theorem 6.2.

Theorem 7.2 *(optimality conditions for discretized sweeping processes via their initial data). Let $(\bar{x}^k, \bar{a}^k, \bar{b}^k, \bar{u}^k)$ be an optimal solution to problem (P_k) under the notation and assumptions of Theorem 7.1 for each fixed index $k \in I\!N$. Then there exist dual elements $(\lambda^k, \alpha^{1k}, \alpha^{2k}, \psi^k, p^k)$ as in Theorem 7.1 together with vectors $\eta_j^k \in \mathbb{R}_+^m$ as $j = 0, \ldots, \nu(k) - 1$ and $\gamma_j^k \in \mathbb{R}^m$ as $j = 0, \ldots, \nu(k) - 1$ satisfying the following conditions:*

- *The* PRIMAL ARC REPRESENTATION:

$$-\frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{h_j^k} + g(\bar{x}_j^k, \bar{u}_j^k) = \sum_{i=1}^m \eta_{ij}^k \bar{a}_{ij}^k, \quad j = 0, \ldots, \nu(k) - 1. \tag{7.46}$$

- *The* ADJOINT DYNAMIC RELATIONSHIPS:

$$\frac{p_{j+1}^{xk} - p_j^{xk}}{h_j^k} - \lambda^k w_j^{xk} \in \nabla_x g(\bar{x}_j^k, \bar{u}_j^k)^* \Big(\frac{1}{h_j^k} \lambda^k \theta_j^{Xk} + \lambda^k v_j^{xk} - p_{j+1}^{xk} \Big) + \sum_{i=1}^m \gamma_{ij}^k \bar{a}_{ij}^k, \tag{7.47}$$

$$\begin{aligned} &\frac{p_{j+1}^{ak} - p_j^{ak}}{h_j^k} - \lambda^k w_j^{ak} - \frac{2}{h_j^k} \Big[\alpha_j^{1k} + \alpha_j^{2k}, \bar{a}_j^k \Big] \\ &= \Big[\gamma_j^k, \operatorname{rep}_m(\bar{x}_j^k) \Big] + \Big[\eta_j^k, \operatorname{rep}_m \Big(-\frac{1}{h_j^k} \lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk} \Big) \Big], \end{aligned} \tag{7.48}$$

$$\frac{p_{j+1}^{bk} - p_j^{bk}}{h_j^k} - \lambda^k w_j^{bk} = -\gamma_j^k, \quad j = 0, \ldots, \nu(k) - 1, \tag{7.49}$$

where the components of the vectors γ_j^k are such that

$$\begin{cases} \gamma_{ij}^k = 0 \ \text{if} \ \Big\langle \bar{a}_{ij}^k, \bar{x}_j^k \Big\rangle < \bar{b}_{ij}^k, \ \text{or} \ \eta_{ij}^k = 0 \ \text{and} \ \Big\langle \bar{a}_{ij}^k, -\frac{1}{h_j^k} \lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk} \Big\rangle < 0, \\ \gamma_{ij}^k \geq 0 \ \text{if} \ \Big\langle \bar{a}_{ij}^k, \bar{x}_j^k \Big\rangle = \bar{b}_{ij}^k, \ \eta_{ij}^k = 0, \ \text{and} \ \Big\langle \bar{a}_{ij}^k, -\frac{1}{h_j^k} \lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk} \Big\rangle > 0, \\ \gamma_{ij}^k \in \mathbb{R} \ \text{if} \ \eta_{ij}^k > 0 \ \text{and} \ \Big\langle \bar{a}_{ij}^k, -\frac{1}{h_j^k} \lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk} \Big\rangle = 0 \end{cases}$$

for the indices $j = 0, \ldots, \nu(k) - 1$ and $i = 1, \ldots, m$.

• *The* LOCAL MAXIMUM PRINCIPLE:

$$\psi_j^k \in N(\bar{u}_j^k; U) \;\; with \;\; -\frac{1}{h_j^k}\psi_j^k - \frac{1}{h_j^k}\lambda^k\theta_j^{uk} - \lambda^k w_j^{uk} \in \nabla_u g(\bar{x}_j^k, \bar{u}_j^k)^*\Big(\frac{1}{h_j^k}\lambda^k\theta_j^{Xk} + \lambda^k v_j^{xk} - p_{j+1}^{xk}\Big) \tag{7.50}$$

for $j = 0, \ldots, \nu(k) - 1$, *where the subgradients* $(w_j^{xk}, w_j^{ak}, w_j^{bk}, w_j^{uk}, v_j^{xk}, v_j^{ak}, v_j^{bk})$ *are taken from* (7.11). *If furthermore the normal cone* $N(\bar{u}_j^k; U)$ *is tangentially generated, i.e.,*

$$N(\bar{u}_j^k; U) = T^*(\bar{u}_j^k; U) := \big\{ v \in \mathbb{R}^d \big|\; \langle v, u\rangle \le 0 \;\; for\ all\;\; u \in T(\bar{u}_j^k; U)\big\},$$

for some tangent cone $T(\bar{u}_j^k; U)$, *then the first inclusion in* (7.50) *is written as*

$$\Big\langle \psi_j^k, \bar{u}_j^k \Big\rangle = \max_{u \in T(\bar{u}_j^k; U)} \Big\langle \psi_j^k, u \Big\rangle, \qquad j = 0, \ldots, \nu(k) - 1, \tag{7.51}$$

which reduces to the GLOBAL MAXIMUM PRINCIPLE

$$\langle \psi_j^k, \bar{u}_j^k \rangle = \max_{u \in U} \Big\langle \psi_j^k, u \Big\rangle, \qquad j = 0, \ldots, \nu(k) - 1, \tag{7.52}$$

provided that the control set U *is convex.*

• *The* TRANSVERSALITY CONDITIONS *at the right endpoint:*

$$-p_{\nu(k)}^{xk} \in \lambda^k \partial\varphi(\bar{x}_{\nu(k)}^k) + \sum_{i=1}^m \eta_{i\nu(k)}^k \bar{a}_{i\nu(k)}^k, \tag{7.53}$$

$$p_{\nu(k)}^{ak} = -2\Big[\alpha_{\nu(k)}^{1k} + \alpha_{\nu(k)}^{2k}, \bar{a}_{i\nu(k)}^k\Big] - \Big[\eta_{\nu(k)}^k, \operatorname{rep}_m(\bar{x}_{\nu(k)}^k)\Big], \tag{7.54}$$

$$p_{i\nu(k)}^{bk} = \eta_{i\nu(k)}^k \ge 0, \; \Big\langle \bar{a}_{i\nu(k)}^k, \bar{x}_{\nu(k)}^k \Big\rangle < \bar{b}_{i\nu(k)}^k \Longrightarrow p_{i\nu(k)}^{bk} = 0 \;\; for \;\; i = 1, \ldots, m \tag{7.55}$$

with dual vectors $\alpha_{\nu(k)}^{1k}$ *and* $\alpha_{\nu(k)}^{2k}$ *satisfying*

$$\alpha_{i\nu(k)}^{1k} \in N_{[0, 1+\delta_k]}(\|\bar{a}_{i\nu(k)}^k\|) \;\; and \;\; \alpha_{i\nu(k)}^{2k} \in N_{[1-\delta_k, \infty]}(\|\bar{a}_{i\nu(k)}^k\|), \; i = 1, \ldots, m, \tag{7.56}$$

where the normal cone to the convex sets is explicitly expressed in form (1.2).

• *The* COMPLEMENTARITY SLACKNESS CONDITIONS:

$$\Big[\Big\langle a_{ij}^k, \bar{x}_j^k \Big\rangle < \bar{b}_{ij}^k\Big] \Longrightarrow \eta_{ij}^k = 0, \tag{7.57}$$

$$\Big[\Big\langle \bar{a}_{i\nu(k)}^k, \bar{x}_{\nu(k)}^k \Big\rangle < \bar{b}_{i\nu(k)}^k\Big] \Longrightarrow \eta_{i\nu(k)}^k = 0, \tag{7.58}$$

$$\eta_{ij}^k > 0 \Longrightarrow \left[\left\langle \bar{a}_{ij}^k, -\frac{1}{h_j^k}\lambda^k\theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk} \right\rangle = 0\right] \tag{7.59}$$

for all the indices $j = 0, \ldots, \nu(k) - 1$ *and* $i = 1, \ldots, m$.

• *The* NONTRIVIALITY CONDITIONS:

$$\lambda^k + \left\|\alpha^{1k} + \alpha^{2k}\right\| + \left\|\eta^k_{\nu(k)}\right\| + \sum_{j=0}^{\nu(k)-1} \left\|p^{xk}_j\right\| + \left\|p^{ak}_0\right\| + \left\|p^{bk}_0\right\| + \left\|\psi^k\right\| \neq 0, \tag{7.60}$$

$$\lambda^k + \left\|\alpha^{1k} + \alpha^{2k}\right\| + \left\|\gamma^k\right\| \neq 0. \tag{7.61}$$

Proof. It follows from condition (7.9) of Theorem 7.1 and the coderivative definition (6.3) that

$$\begin{aligned}&\left(\frac{p^{xk}_{j+1} - p^{xk}_j}{h^k_j} - \lambda^k w^{xk}_j, \frac{p^{ak}_{j+1} - p^{ak}_j}{h^k_j} - \lambda^k w^{ak}_j - \frac{2}{h^k_j}\left(\alpha^{1k}_j + \alpha^{2k}_j\right)\bar{a}^k_{ij}, \frac{p^{bk}_{j+1} - p^{bk}_j}{h^k_j} - \lambda^k w^{bk}_j,\right.\\ &\left. -\frac{1}{h^k_j}\lambda^k\theta^{uk}_j - \lambda^k w^{uk}_j - \frac{1}{h^k_j}\psi^k_j\right)\\ &\in D^*F\left(\bar{x}^k_j, \bar{a}^k_j, \bar{b}^k_j, \bar{u}^k_j, -\frac{\bar{x}^k_{j+1} - \bar{x}^k_j}{h^k_j}\right)\left(-\frac{1}{h^k_j}\lambda^k\theta^{Xk}_j - \lambda^k v^{xk}_j + p^{xk}_{j+1}\right)\end{aligned}$$

for all $j = 0, \ldots, \nu(k) - 1$, $i = 1, \ldots, m$. Using the inclusion

$$-\frac{\bar{x}^k_{j+1} - \bar{x}^k_j}{h^k_j} + g(\bar{x}^k_j, \bar{u}^k_j) \in G(\bar{x}^k_j, \bar{a}^k_j, \bar{b}^k_j)$$

via the normal cone mapping G from (6.5) and employing the PLICQ property of the vectors $\left\{\bar{a}^k_i \,|\, i \in I(\bar{x}^k, \bar{a}^k, \bar{b}^k)\right\}$ give us a unique vector $\eta^k_j \in \mathbb{R}^m_+$ such that for all $i = 1, \ldots, m$ we have

$$\sum_{i=1}^m \eta^k_{ij}\bar{a}^k_{ij} = -\frac{\bar{x}^k_{j+1} - \bar{x}^k_j}{h^k_j} + g(\bar{x}^k_j, \bar{u}^k_j) \text{ with } \eta^k_{ij} \in N_{\mathbb{R}_-}\left(\left\langle\bar{a}^k_{ij}, \bar{x}^k_j\right\rangle - \bar{b}^k_{ij}\right),\ j = 0, \ldots, \nu(k) - 1,$$

which verifies the implications in (7.46) and (7.57). Applying now the coderivative upper estimate (6.8) from Theorem 6.2 with $x := \bar{x}^k_j$, $a := \bar{a}^k_j$, $b := \bar{b}^k_j$, $u := \bar{u}^k_j$, $w := -\frac{\bar{x}^k_{j+1} - \bar{x}^k_j}{h^k_j}$, and $y := -\frac{1}{h_k}\lambda^k\theta^{Xk}_j - \lambda^k v^{xk}_j + p^{xk}_{j+1}$ as $j = 0, \ldots, \nu(k) - 1$ shows that $\gamma^k_j \in \mathbb{R}^m$ and that the relationships

$$\left(\begin{array}{c}\dfrac{p^{xk}_{j+1} - p^{xk}_j}{h^k_j} - \lambda^k w^{xk}_j, \dfrac{p^{ak}_{j+1} - p^{ak}_j}{h^k_j} - \lambda^k w^{ak}_j - \dfrac{2}{h^k_j}\left[\alpha^{1k}_j + \alpha^{2k}_j, \bar{a}^k_j\right], \dfrac{p^{bk}_{j+1} - p^{bk}_j}{h^k_j} - \lambda^k w^{bk}_j,\\ -\dfrac{1}{h^k_j}\lambda^k\theta^{uk}_j - \lambda^k w^{uk}_j - \dfrac{1}{h^k_j}\psi^k_j\end{array}\right)$$

$$\in \begin{pmatrix} -\nabla g_x(\bar{x}_j^k, \bar{u}_j^k)^* \left(-\frac{1}{h_j^k}\lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk}\right) + \sum_{i=1}^{m} \gamma_{ij}^k \bar{a}_{ij}^k, \\ \left[\gamma_j^k, \mathrm{rep}_m(\bar{x}_j^k)\right] + \left[\eta_j^k, \mathrm{rep}_m\left(-\frac{1}{h_j^k}\lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk}\right)\right], \\ -\gamma_j^k,\ -\nabla g_u(\bar{x}_j^k, \bar{u}_j^k)^* \left(-\frac{1}{h_j^k}\lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk}\right) \end{pmatrix}$$

are satisfied, where $\psi_j^k \in N(\bar{u}_j^k; U)$ for all $j = 0, \ldots, \nu(k) - 1$, and where the components γ_{ij}^k of the vectors $\gamma_j^k \in \mathbb{R}^m$ as $i = 1, \ldots, m$ are taken from

$$\gamma_{ij}^k \in D^* N_{\mathbb{R}_-}\left(\left\langle \bar{a}_{ij}^k, \bar{x}_j^k \right\rangle - \bar{b}_{ij}^k, \eta_{ij}^k\right)\left(\left\langle \bar{a}_{ij}^k, -\frac{1}{h_j^k}\lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk} \right\rangle\right). \tag{7.62}$$

The obtained relationships together with the direct calculation of the coderivative $D^* N_{\mathbb{R}_-}$ in (7.62) ensure the validity of all the conditions in (7.47) as well as the inclusion in (7.50). The latter together with (7.10) constitutes an appropriate version of the (linearized) local maximum principle for nonconvex discrete-time systems. It immediately gives us the local maximality condition (7.51) in the case of tangentially generated normals, which surely holds for the class of normally regular sets U; see, e.g., [28,34]. The global form of the discrete maximum principle in (7.52) is a direct consequence of (7.50) and the normal cone representation (1.2) for convex sets. Furthermore, conditions (7.53), (7.54), and (7.55) clearly follow from (7.5), (7.6), and (7.7) due to (7.2).

Defining now $\eta_{\nu(k)}^k := \xi^k$ via ξ^k from the statement of Theorem 7.1 yields $\eta_j^k \in \mathbb{R}_+^m$ for $j = 0, \ldots, \nu(k)$ and allows us to deduce the nontriviality condition (7.60) from that in (7.1) and also the transversality conditions in (7.53)–(7.55) from those in (7.5)–(7.7). Implication (7.58) is a direct consequence of (7.2) and the definition of $\eta_{\nu(k)}^k$. Observing that (7.59) follows from the fact that

$$-\frac{1}{h_j^k}\lambda^k \theta_j^{Xk} - \lambda^k v_j^{xk} + p_{j+1}^{xk} \in \bigcap_{\left\{i \mid \eta_{ij}^k > 0\right\}} (\bar{a}_{ij}^k)^\perp,$$

we get from (7.3) and (7.4) that both inclusions in (7.56) hold.

It remains to verify the nontriviality condition (7.61). Suppose on the contrary that $\lambda^k = 0$, $\alpha^{1k} + \alpha^{2k} = 0$, and $\gamma^k = 0$. We deduce from (7.8) that $p_{\nu(k)}^{ak} = 0$ and $p_{\nu(k)}^{bk} = 0$, which clearly yield $\eta_{\nu(k)}^k = p_{\nu(k)}^{bk} = 0$. Then it follows from (7.53) that $p_{\nu(k)}^{xk} = 0$, and thus $\left(p_j^{xk}, p_j^{ak}\right) = (0, 0)$ for all $j = 0, \ldots, \nu(k) - 1$ by (7.47) and (7.48). This implies that $\psi^k = 0$ by (7.50). Using finally (7.49) tells us that $p_0^{bk} = 0$. It means that (7.60) is violated, which is a contradiction that justifies the validity of (7.61) and therefore completes the proof of the theorem. □

8. Numerical illustration

In this section we present a nontrivial example illustrating the application of the obtained results to solve the sweeping optimal control problem (P). We consider this problem with the following data, where the a-components and b-components of controls are fixed, and only the u-components are used for optimization:

$$\begin{cases} n=2,\ m=1,\ T=1,\ x_0=\left(\frac{3}{2},1\right),\ a=\left(-\frac{1}{\sqrt{5}},-\frac{2}{\sqrt{5}}\right),\ b=-\frac{2}{\sqrt{5}}, \\ g(x,u):=u,\ \ \varphi(x):=x_1+x_2,\ \ \ell(t,x,a,b,u,\dot{x},\dot{a},\dot{b}):=\frac{1}{2}u_1^2+u_2^2, \\ U:=[-1,1]\times[-1,1]. \end{cases} \tag{8.1}$$

The set $C(t)$ in the sweeping inclusion (1.1) is described now by

$$C(t)=C:=\left\{(x_1,x_2)\in\mathbb{R}^2\,\middle|\, x_1+2x_2\geq 2\right\}\ \text{ for all }\ t\in[0,1].$$

In what follows we are going to show that applying the optimality conditions of Theorem 7.2 allows us to find optimal solutions to problems (P_k), *for each* $k\in\mathbb{N}$, that can be viewed as (sub)optimal solutions to the original sweeping control problem (P). By construction, the velocity of the object and the u-components of controls are piecewise constant functions on $[0,1]$.

The structure of the problem suggests that the object only changes its velocity when it hits the boundary at some time $t^*\in[0,1]$. Moreover, if $t^*<1$, the object slides on the boundary of C for the whole interval $[t^*,1]$. In this case, by construction of t^* it must be one of the mesh points t_j^k of some partition Δ_k in Theorem 4.1. Assume that $t^*=t_s^k$ for some $s\in\{0,1,\ldots,\nu(k)\}$. It is easy to see that all the assumptions of Theorem 7.2 are satisfied for (8.1), and we can employ the obtained necessary optimality conditions, where the superscript "k" is dropped, and where the quadruple $\left(\theta_j^u,\theta_j^X,\theta_j^A,\theta_j^B\right)$ is supposed to be 0 for large k due to the established convergence of optimal solutions. Then we have the existence of $\lambda\geq 0$, $\eta_j\geq 0$, $\gamma_j\in\mathbb{R}$, $\alpha_j^1,\alpha_j^2\in\mathbb{R}$, $\psi_j\in\mathbb{R}^2$, $(p_j^x,p_j^a,p_j^b)\in\mathbb{R}^5$, $(w_j^x,w_j^a,w_j^b,w_j^u)\in\mathbb{R}^7$, and $(x_j^x,v_j^a,v_j^b)\in\mathbb{R}^5$ as $j=0,\ldots,\nu(k)-1$ satisfying the following relationships:

1. $\left(w_j^x,w_j^a,w_j^b,w_j^u\right)=(0,0,0,0,0,\bar{u}_1,2\bar{u}_2)$ for $j=0,\ldots,\nu(k)-1$.
2. $\left(v_j^x,v_j^a,v_j^b\right)=(0,0,0,0,0,0)$ for $j=0,\ldots,\nu(k)-1$.
3. $\dot{\bar{x}}(t)=\begin{cases}\bar{u}(t)+\dfrac{\eta_0}{\sqrt{5}}(1,2) & \text{if}\quad t\in(0,t_s)\\ \bar{u}(t)+\dfrac{\eta_s}{\sqrt{5}}(1,2) & \text{if}\quad t\in(t_s,1)\end{cases}$, where $\bar{u}(t)=\begin{cases}\bar{u}_0 & \text{if}\quad t\in[0,t_s)\\ \bar{u}_s & \text{if}\quad t\in(t_s,1]\end{cases}$.
4. $\begin{cases}\dfrac{1}{h_j^k}\left(p_{j+1}^x-p_j^x\right)=\gamma_j(1,2),\\ \dfrac{1}{h_j^k}\left(p_{j+1}^a-p_j^a\right)-\dfrac{2}{h_j^k}\left(-\left(\alpha_j^1+\alpha_j^2\right)\frac{1}{\sqrt{5}},-\left(\alpha_j^1+\alpha_j^2\right)\frac{2}{\sqrt{5}}\right)\\ \quad=\left(\gamma_j\bar{x}_{1j},\gamma_j\bar{x}_{2j}\right)+\left(\eta_j p_{1,j+1}^x,\eta_j p_{2,j+1}^x\right),\\ \dfrac{1}{h_j^k}\left(p_{j+1}^b-p_j^b\right)=-\gamma_j\ \text{ for }\ j=0,\ldots,\nu(k)-1.\end{cases}$
5. $\dfrac{1}{h_j^k}\psi_j+\lambda\left(\bar{u}_{1j},2\bar{u}_{2j}\right)=p_{j+1}^x$ for $j=0,\ldots,\nu(k)-1$.

6. $\psi_j \in N\left(\bar{u}_j; [-1,1]\times[-1,1]\right)$ for $j = 0, \ldots, \nu(k)-1$, which is equivalent to $\psi_{1j}\bar{u}_{1j} + \psi_{2j}\bar{u}_{2j} = \max\limits_{(u_1,u_2)\in[-1,1]\times[-1,1]} \left\{\psi_{1j}u_1 + \psi_{2j}u_2\right\}$.
7. $\bar{x}_{1j} + 2\bar{x}_{2j} > 2 \Longrightarrow \gamma_j = 0$ and $\eta_j = 0$ for $j = 0, \ldots, \nu(k)-1$.
8. $\eta_j > 0 \Longrightarrow \left\langle (-1,-2), p^x_{j+1}\right\rangle = 0$ for $j = 0, \ldots, \nu(k)-1$.
9. $\bar{x}_{1\nu(k)} + 2\bar{x}_{2\nu(k)} > 2 \Longrightarrow \eta_{\nu(k)} = \eta_s = 0$.
10. $\begin{cases} -p^x_{\nu(k)} = \lambda(1,1) + \eta_{\nu(k)}(1,2), \\ p^a_{\nu(k)} = -2\left(-\left(\alpha^1_{\nu(k)} + \alpha^2_{\nu(k)}\right)\frac{1}{\sqrt{5}}, -\left(\alpha^1_{\nu(k)} + \alpha^2_{\nu(k)}\right)\frac{2}{\sqrt{5}}\right) - \left(\eta_{\nu(k)}\bar{x}_{1\nu(k)}, \eta_{\nu(k)}\bar{x}_{2\nu(k)}\right), \\ p^b_{\nu(k)} = \eta_{\nu(k)} \geq 0. \end{cases}$
11. $\alpha^1_{\nu(k)} \in N_{[0,1+\delta_k]}(1)$, $\alpha^2_{\nu(k)} \in N_{[1-\delta_k,\infty)}(1)$, which implies that $\left(\alpha^1_{\nu(k)}, \alpha^2_{\nu(k)}\right) = (0,0)$.
12. $\lambda + \left\|\alpha^1 + \alpha^2\right\| + \|\gamma\| > 0$.

It clearly follows from (3) that

$$\bar{x}(t) = \begin{cases} \left(\frac{3}{2} + t\bar{u}_{10}, 1 + t\bar{u}_{20}\right) \\ \quad \text{if } t \in [0, t_s) \\ \left(\frac{3}{2} + t_s\bar{u}_{10} + (t-t_s)\left(\bar{u}_{1s} + \eta_s/\sqrt{5}\right), 1 + t_s\bar{u}_{20} + (t-t_s)\left(\bar{u}_{2s} + 2\eta_s/\sqrt{5}\right)\right) \\ \quad \text{if } t \in [t_s, 1], \end{cases}$$

where we get $\eta_0 = 0$ due to the conditions in (7). Since $t = t_s$ is the time when the moving particle hits the boundary, i.e., $\bar{x}_1(t_s) + 2\bar{x}_2(t_s) = 2$, then

$$\eta_j = \begin{cases} \eta_0 = 0 \text{ if } j < s \\ \eta_s \geq 0 \text{ if } j \geq s. \end{cases}$$

Of course, the normal vectors are inactive before the hitting time. Consequently, we have that

$$\frac{7}{2} + t_s(\bar{u}_{10} + 2\bar{u}_{20}) = 2 \text{ if } t < t_s \text{ and}$$
$$\frac{7}{2} + t_s(\bar{u}_{10} + 2\bar{u}_{20}) + (t-t_s)\left(\bar{u}_{1s} + 2\bar{u}_{2s} + \sqrt{5}\eta_s\right) = 2 \text{ if } t \geq t_s.$$

This allows us to calculate the hitting time as

$$t_s = -\frac{3}{2(\bar{u}_{10} + 2\bar{u}_{20})}, \tag{8.2}$$

which implies in turn the condition

$$\bar{u}_{10} + 2\bar{u}_{20} \leq -\frac{3}{2} \text{ due to } 0 \leq t_s \leq 1. \tag{8.3}$$

When $\bar{x}(\cdot)$ hits the boundary of C, it would stay there while pointing in the direction shown in Fig. 1. Thus

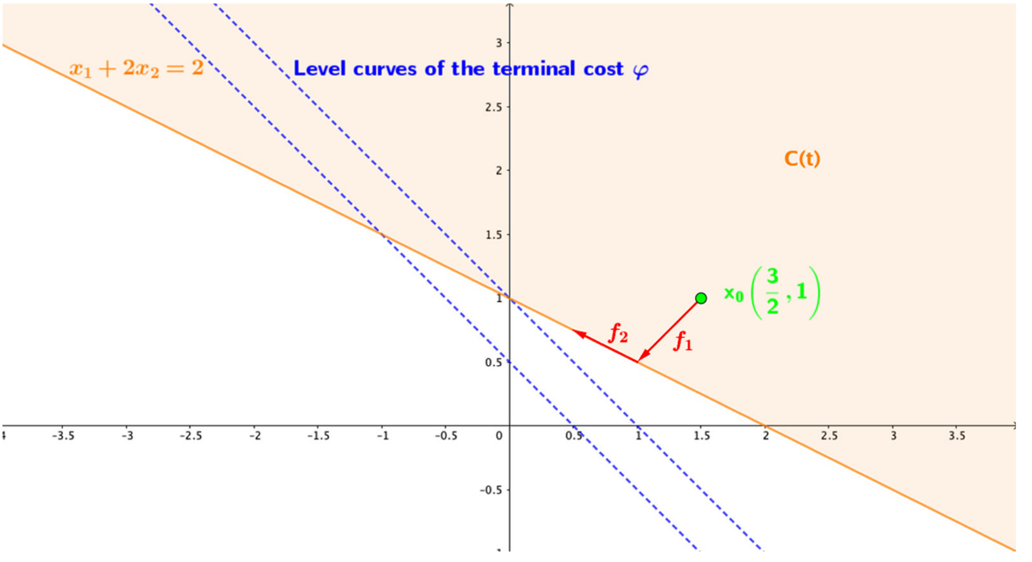

Fig. 1. Dynamics of the controlled sweeping process.

$$\bar{u}_{1s} + 2\bar{u}_{2s} + \sqrt{5}\eta_s = 0 \text{ and hence } \eta_s = -\frac{\bar{u}_{1s} + 2\bar{u}_{2s}}{\sqrt{5}}. \tag{8.4}$$

The cost functional is calculated therefore by

$$J[\bar{x}, \bar{u}] = t_s\left(\frac{\bar{u}_{10}^2}{2} + \bar{u}_{20}^2 + \bar{u}_{10} + \bar{u}_{20}\right) + (1 - t_s)\left(\frac{\bar{u}_{1s}^2}{2} + \bar{u}_{2s}^2 + \bar{u}_{1s} + \bar{u}_{2s} + \frac{3\eta_s}{\sqrt{5}}\right) + \frac{5}{2},$$

where t_s and η_s are given in (8.2) and (8.4), respectively.

When the object has not hit the boundary of the set C yet, we have $t < t_s$. In this case, the implication in (8) cannot be used and the control u should be chosen by taking into account the only requirement that the corresponding trajectory hits the boundary exactly at the time t_s. We examine the following two possibilities:

- **Case 1a:** If the object is pushed to reach the boundary as soon as possible and it slides there after that to reduce the cost functional, we expect to use more energy. Then the control $\bar{u}_0$ should be on the boundary of the control set $U = [-1, 1] \times [-1, 1]$. It follows from the structure of the problem that either $\bar{u}_{10} = -1$, or $\bar{u}_{20} = -1$. In this way we get:

  If $\bar{u}_{10} = -1$, then $-1 \le \bar{u}_{20} \le -\frac{1}{4}$.

  If $\bar{u}_{20} = -1$, then $-1 \le \bar{u}_{10} \le \frac{1}{2}$.
- **Case 1b:** If we wish to save energy, then the control selection is $\bar{u}_0 = (\bar{u}_{10}, \bar{u}_{20}) \in (-1, 1) \times (-1, 1)$ such that $\bar{u}_{10}$ and $\bar{u}_{20}$ satisfy inequality (8.3).

When the object hits the boundary, it then slights there until the end of the process, i.e., $t \ge t_s$. To proceed further, consider the following two situations:

- **Case 2a:** $\eta_s = 0$, i.e., the normal vector $\dfrac{\eta_s}{\sqrt{5}}(-1,-2)$ taken from the normal cone $N(x(t);C(t))$ is not active for $t \geq t_s$. It then follows from (8.4) that $\bar{u}_{1s} = -2\bar{u}_{2s}$. The cost functional in this case is

$$J[\bar{x},\bar{u}] = t_s\left(\frac{\bar{u}_{10}^2}{2} + \bar{u}_{20}^2 + \bar{u}_{10} + \bar{u}_{20}\right) + (1-t_s)\left(3\bar{u}_{2s}^2 - \bar{u}_{2s}\right) + \frac{5}{2},$$

which achieves the minimum value at $\bar{u}_{2s} = \frac{1}{6}$ implying in turn that $\bar{u}_{1s} = -\frac{1}{3}$ and

$$J[\bar{x},\bar{u}] = t_s\left(\frac{\bar{u}_{10}^2}{2} + \bar{u}_{20}^2 + \bar{u}_{10} + \bar{u}_{20}\right) + \frac{t_s-1}{12} + \frac{5}{2}.$$

- **Case 2b:** $\eta_s > 0$. Using (8) gives us $p_{1j}^x + 2p_{2j}^x = 0$ for all $j \geq s$. When the object hits the boundary, we do not need to use the maximum energy. That is, $\bar{u}_s$ should be selected in $(-1,1)\times(-1,1)$, and so $\psi_j = 0$ due to (6). It then follows from (4) and (5) that

$$\frac{1}{h_j^k}\psi_{1j} + \frac{2}{h_j^k}\psi_{2j} + \lambda\left(\bar{u}_{1s} + 4\bar{u}_{2s}\right) = 0, \tag{8.5}$$

which therefore gives us $\bar{u}_{1s} + 4\bar{u}_{2s} = 0$, or equivalently

$$\bar{u}_{1s} = -4\bar{u}_{2s}$$

while assuming that $\lambda > 0$; otherwise we do not have enough information to proceed. Then

$$\eta_s = \frac{2\bar{u}_{2s}}{\sqrt{5}} \geq 0.$$

In this case the cost functional is

$$J[\bar{x},\bar{u}] = t_s\left(\frac{\bar{u}_{10}^2}{2} + \bar{u}_{20}^2 + \bar{u}_{10} + \bar{u}_{20}\right) + (1-t_s)\left(9\bar{u}_{2s}^2 - \frac{9}{5}\bar{u}_{2s}\right) + \frac{5}{2},$$

which achieves the minimum value at $\bar{u}_{2s} = \frac{1}{10}$. Thus we get $\bar{u}_{1s} = -\frac{2}{5}$ and

$$J[\bar{x},\bar{u}] = t_s\left(\frac{\bar{u}_{10}^2}{2} + \bar{u}_{20}^2 + \bar{u}_{10} + \bar{u}_{20}\right) + \frac{9(t_s-1)}{100} + \frac{5}{2}.$$

It is clear that in Case 2b the cost functional has a smaller value than in Case 2a, and thus latter case can be ruled out. We then chose $\bar{u}_{10}$ and $\bar{u}_{20}$ in either Case 1a or Case 1b to minimize $J[\bar{x},\bar{u}]$. To simplify the computations, we select $\bar{u}_{10}$ and $\bar{u}_{20}$ as the mesh points of a uniform partition of $[-1,1]$ with the step size $h = \dfrac{2}{N}$ as N is sufficiently large and then compute the corresponding hitting time t_s from (8.2). Let us present the results of computations that are provided by writing a code in Python for the case of $N = 20$ and hence the step size $h = 0.1$. After running the code in Python, we get the following table:

$\bar{u}_{10}$	$\bar{u}_{20}$	t_s	$J[\bar{x}, \bar{u}]$	$\bar{u}_{10}$	$\bar{u}_{20}$	t_s	$J[\bar{x}, \bar{u}]$
−1	−1	0.5	2.205	−1	−0.9	0.5357	2.1421
−1	−0.8	0.5769	2.0812	−1	−0.7	0.625	2.0225
−1	−0.6	0.6818	1.9668	−1	−0.5	0.75	1.915
−1	−0.4	0.8333	1.8683	−1	−0.3	0.9375	1.8288
−0.9	−1	0.5172	2.2005	−0.9	−0.9	0.5556	2.135
−0.9	−0.8	0.6	2.071	−0.9	−0.7	0.6522	2.0089
−0.9	−0.6	0.7143	1.9493	−0.9	−0.5	0.7895	1.8929
−0.9	−0.4	0.8824	1.8409	−0.9	−0.3	1	1.795
−0.8	−1	0.5357	2.2011	−0.8	−0.9	0.5769	2.1331
−0.8	−0.8	0.625	2.0662	−0.8	−0.7	0.6818	2.0009
−0.8	−0.6	0.75	1.9375	−0.8	−0.5	0.8333	1.8767
−0.8	−0.4	0.9375	1.8194	−0.7	−1	0.5556	2.2072
−0.7	−0.9	0.6	2.137	−0.7	−0.8	0.6522	2.0676
−0.7	−0.7	0.7143	1.9993	−0.7	−0.6	0.7895	1.9324
−0.7	−0.5	0.8824	1.8674	−0.7	−0.4	1	1.805
−0.6	−1	0.5769	2.2196	−0.6	−0.9	0.625	2.1475
−0.6	−0.8	0.6818	2.0759	−0.6	−0.7	0.75	2.005
−0.6	−0.6	0.8333	1.935	−0.6	−0.5	0.9375	1.8662
−0.5	−1	0.6	2.239	−0.5	−0.9	0.6522	2.1654
−0.5	−0.8	0.7143	2.0921	−0.5	−0.7	0.7895	2.0192
−0.5	−0.6	0.8824	1.9468	−0.5	−0.5	1	1.875
−0.4	−1	0.625	2.2662	−0.4	−0.9	0.6818	2.1918
−0.4	−0.8	0.75	2.1175	−0.4	−0.7	0.8333	2.0433
−0.4	−0.6	0.9375	1.9694	−0.3	−1	0.6522	2.3024
−0.3	−0.9	0.7143	2.2279	−0.3	−0.8	0.7895	2.1534
−0.3	−0.7	0.8824	2.0791	−0.3	−0.6	1	2.005
−0.2	−1	0.6818	2.3486	−0.2	−0.9	0.75	2.275
−0.2	−0.8	0.8333	2.2017	−0.2	−0.7	0.9375	2.1288
−0.1	−1	0.7143	2.4064	−0.1	−0.9	0.7895	2.335
−0.1	−0.8	0.8824	2.2644	−0.1	−0.7	1	2.195
0	−1	0.75	2.4775	0	−0.9	0.8333	2.41
0	−0.8	0.9375	2.3444	0.1	−1	0.7895	2.5639
0.1	−0.9	0.8824	2.5026	0.1	−0.8	1	2.445
0.2	−1	0.8333	2.6683	0.2	−0.9	0.9375	2.6162
0.3	−1	0.8824	2.7938	0.3	−0.9	1	2.755
0.4	−1	0.9375	2.9444	0.5	−1	1	3.125

The next table collects the values of the optimal control and the corresponding costs with different choices of N.

N	$\bar{u}_{10}$	$\bar{u}_{20}$	t_s	$J[\bar{x}, \bar{u}]$
40	−0.85	−0.35	0.9677	1.804
80	−0.825	−0.35	0.9836	1.798
160	−0.825	−0.3375	1	1.7917
320	−0.825	−0.3375	1	1.7917
640	−0.8344	−0.3344	0.99679	1.7924

Keep running the code in Python with $N = 2000$, it then follows that the optimal control before the hitting time is $\bar{u}_{10} = -\dfrac{5}{6}$ and $\bar{u}_{20} = -\dfrac{1}{3}$. In this case, the object reaches the boundary at the ending time $t_s = 1$ and the minimum cost is 1.79167. Our computation of the optimal control is also supported by an educated guess that it can be done by computing the exact minimum for small k, supposing that the hitting time $t_s = 1$.

Acknowledgments

The authors are very grateful to two anonymous referees and to René Henrion for their careful reading of the paper and giving us many useful remarks and suggestions, which allowed us to improve the original presentation.

References

[1] S. Adly, T. Haddad, 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, 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, G. Colombo, A maximum principle of the controlled sweeping process, Set-Valued Var. Anal. 26 (2018) 607–629.
[4] M. Brokate, J. Sprekels, Hysteresis and Phase Transitions, Springer, Berlin, 1996.
[5] T.H. Cao, G. Colombo, B.S. Mordukhovich, D. Nguyen, Necessary optimality conditions for controlled sweeping processes, preprint, 2020.
[6] T.H. Cao, B.S. Mordukhovich, Optimal control of a perturbed sweeping process via discrete approximations, Discrete Contin. Dyn. Syst., Ser. B 21 (2016) 3331–3358.
[7] T.H. Cao, 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) 267–306.
[8] T.H. Cao, B.S. Mordukhovich, Optimal control of a nonconvex perturbed sweeping process, J. Differ. Equ. 266 (2019) 1003–1050.
[9] T.H. Cao, B.S. Mordukhovich, Applications of the controlled sweeping process to optimal control of the planar crowd motion model, Discrete Contin. Dyn. Syst., Ser. B 24 (2019) 4191–4216.
[10] C. Castaing, M.D.P. Monteiro Marques, P. Raynaud de Fitte, Some problems in optimal control governed by the sweeping process, J. Nonlinear Convex Anal. 15 (2014) 1043–1070.
[11] F.H. Clarke, Necessary conditions in dynamic optimization, Mem. Am. Math. Soc. 173 (816) (2005).
[12] G. Colombo, R. Henrion, N.D. Hoang, B.S. Mordukhovich, Optimal control of the sweeping process, Dyn. Contin. Discrete Impuls. Syst., Ser. B 19 (2012) 117–159.
[13] G. Colombo, R. Henrion, N.D. Hoang, B.S. Mordukhovich, Discrete approximations of a controlled sweeping process, Set-Valued Var. Anal. 23 (2015) 69–86.
[14] G. Colombo, R. Henrion, N.D. Hoang, B.S. Mordukhovich, Optimal control of the sweeping process over polyhedral controlled sets, J. Differ. Equ. 260 (2016) 3397–3447.
[15] G. Colombo, B.S. Mordukhovich, D. Nguyen, Optimal control of sweeping processes in robotics and traffic flow models, J. Optim. Theory Appl. 182 (2019) 439–472.
[16] G. Colombo, B.S. Mordukhovich, D. Nguyen, Optimization of a perturbed sweeping process by discontinuous controls, SIAM J. Control Optim. (2020), https://doi.org/10.1137/18M1207120.
[17] G. Colombo, L. Thibault, Prox-regular sets and applications, in: D.Y. Gao, D. Motreanu (Eds.), Handbook of Nonconvex Analysis, International Press, Boston, 2010, pp. 99–182.
[18] E. Cristiani, B. Piccoli, A. Tosin, Multistep Modeling of Pedestrian Dynamics, Springer, Cham, Switzerland, 2014.
[19] M.d.R. de Pinho, M.M.A. Ferreira, G.V. Smirnov, Optimal control involving sweeping processes, Set-Valued Var. Anal. 27 (2019) 523–548.
[20] T. Donchev, E. Farkhi, B.S. Mordukhovich, Discrete approximations, relaxation, and optimization of one-sided Lipschitzian differential inclusions in Hilbert spaces, J. Differ. Equ. 243 (2007) 301–328.
[21] J.F. Edmond, L. Thibault, Relaxation of an optimal control problem involving a perturbed sweeping process, Math. Program. 104 (2005) 347–373.
[22] N.D. Hoang, B.S. Mordukhovich, Extended Euler-Lagrange and Hamiltonian formalisms in optimal control of sweeping processes with controlled sweeping sets, J. Optim. Theory Appl. 180 (2019) 256–289.
[23] A. Jourani, E. Vilches, A differential equation approach to implicit sweeping processes, J. Differ. Equ. 266 (2019) 5168–5184.

[24] M.A. Krasnosel'skiĭ, A.V. Pokrovskiĭ, Systems with Hysteresis, Springer, New York, 1989.
[25] M. Kunze, M.D.P. Monteiro Marques, An introduction to Moreau's sweeping process, in: B. Brogliato (Ed.), Impacts in Mechanical Systems, in: Lecture Notes in Phys., vol. 551, Springer, Berlin, 2000, pp. 1–60.
[26] B. Maury, J. Venel, A discrete model for crowd motions, ESAIM: M2AN 45 (2011) 145–168.
[27] B.S. Mordukhovich, Discrete approximations and refined Euler-Lagrange conditions for differential inclusions, SIAM J. Control Optim. 33 (1995) 882–915.
[28] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications, Springer, Berlin, 2006.
[29] B.S. Mordukhovich, Variational Analysis and Applications, Springer, Cham, Switzerland, 2018.
[30] B.S. Mordukhovich, J.V. Outrata, Coderivative analysis of quasi-variational inequalities with applications to stability and optimization, SIAM J. Optim. 18 (2007) 389–412.
[31] J.J. Moreau, On unilateral constraints, friction and plasticity, in: G. Capriz, G. Stampacchia (Eds.), New Variational Techniques in Mathematical Physics, in: Proc. C.I.M.E. Summer Schools, Cremonese, Rome, 1974, pp. 173–322.
[32] B.N. Pshenichnyi, Convex Analysis and Extremal Problems, Nauka, Moscow, 1980 (in Russian).
[33] S.M. Robinson, Some continuity properties of polyhedral multifunctions, Math. Program. Stud. 19 (1981) 206–214.
[34] R.T. Rockafellar, R.J-B. Wets, Variational Analysis, Springer, Berlin, 1998.
[35] G.V. Smirnov, Introduction to the Theory of Differential Inclusions, American Mathematical Society, Providence, RI, 2002.
[36] A.A. Tolstonogov, Sweeping process with unbounded nonconvex perturbation, Nonlinear Anal. 108 (2014) 291–301.
[37] A.A. Tolstonogov, Control sweeping processes, J. Convex Anal. 23 (2016) 1099–1123.
[38] J. Venel, A numerical scheme for a class of sweeping process, Numer. Math. 118 (2011) 451–484.
[39] R.B. Vinter, Optimal Control, Birkhäuser, Boston, 2000.