Optimal control of a nonconvex perturbed sweeping process

Tan H. Cao [a,1], B.S. Mordukhovich [b,c,*,2]

[a] *Department of Applied Mathematics and Statistics, State University of New York–Korea, Yeonsu-Gu, Incheon, Republic of Korea*
[b] *Department of Mathematics, Wayne State University, Detroit, MI, USA*
[c] *RUDN University, Moscow, Russia*

Abstract

The paper concerns optimal control of discontinuous differential inclusions of the normal cone type governed by a generalized version of the Moreau sweeping process with control functions acting in both nonconvex moving sets and additive perturbations. This is a new class of optimal control problems in comparison with previously considered counterparts where the controlled sweeping sets are described by convex polyhedra. Besides a theoretical interest, a major motivation for our study of such challenging optimal control problems with intrinsic state constraints comes from the application to the crowd motion model in a practically adequate planar setting with nonconvex but prox-regular sweeping sets. Based on a constructive discrete approximation approach and advanced tools of first-order and second-order variational analysis and generalized differentiation, we establish the strong convergence of discrete optimal solutions and derive a complete set of necessary optimality conditions for discrete-time and continuous-time sweeping control systems that are expressed entirely via the problem data.
© 2018 Elsevier Inc. All rights reserved.

---

* Corresponding author.
  E-mail addresses: tan.cao@stonybrook.edu (T.H. Cao), boris@math.wayne.edu (B.S. Mordukhovich).

1. Introduction, problem formulation, and discussions

The sweeping process was introduced and investigated by Jean-Jacques Moreau in the 1970s
(see, in particular, [28–30]) in the differential inclusion form

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

where $C(t)$ is a (Lipschitz or absolutely) continuous moving *convex* set, and where the normal
cone N to it is understood in the sense of convex analysis for which Moreau was one of the
creators and major players. The original Moreau's motivation came mainly from applications to
elastoplasticity, but it has been well recognized over the years that the sweeping process is im-
portant for many other applications to various problems in mechanics, hysteresis systems, traffic
equilibria, social and economic modelings, etc.; see, e.g., [14,19–22,34,37] and the references
therein.

Due to the maximal monotonicity of the normal cone operator in convex analysis, the sweep-
ing system (1.1) is described by a dissipative discontinuous differential inclusion and can for-
mally be related to control theory for dynamical systems governed by differential inclusions of
the type $\dot{x} \in F(x)$, which has been broadly developed in variational analysis and optimal con-
trol; see, e.g., the books [11,26,38] with their extensive bibliographies. However, the results of
the latter theory, obtained under certain Lipschitzian assumptions on F, are not applicable to
the discontinuous sweeping process (1.1). Moreover, it is well known that the Cauchy problem
for (1.1) admits a unique solution, which excludes any optimization and control of the sweeping
process in form (1.1) with a given moving set $C(t)$.

The authors of [12] introduced a *control version* of the sweeping process by inserting control
actions into the moving set $C(t)$ with considering its *polyhedral* evolution

$$C(t) := \big\{x \in \mathbb{R}^n \,\big|\, \langle u_i(t), x \rangle \le b_i(t), \ i = 1, \ldots, m \big\}, \ \|u_i(t)\| = 1 \text{ for all } t \in [0, T], \tag{1.2}$$

where optimal control functions $u_i(t)$ and $b_i(t)$ ought to be selected in order to minimize some
cost functional. Formulated in this way optimization models for the controlled sweeping process
in (1.1) and (1.2) can be written as optimal control problems for unbounded discontinuous dif-
ferential inclusions with pointwise state constraints of inequality and equality types, which have
never been considered before in optimal control theory. The discrete approximation approach of
variational analysis, which significantly extends the one developed in [24,26] for Lipschitzian
differential inclusions, allowed the authors of [13] to derive an adequate set of necessary opti-
mality conditions for such polyhedral sweeping control problems with detailed illustrations of
new phenomena by nontrivial examples.

A *perturbed version* of the polyhedral controlled sweeping process was considered in [7,8] in
the form

$$-\dot{x}(t) \in N\big(x(t); C(t)\big) + f\big(x(t), a(t)\big) \text{ a.e. } t \in [0, T], \quad x(0) = x_0 \in C(0), \tag{1.3}$$

with controls $a \colon [0, T] \to \mathbb{R}^d$ acting in perturbations and controls $u \colon [0, T] \to \mathbb{R}^n$ acting in the polyhedral moving set generated by the fixed vectors x_i^* as

$$C(t) := C + u(t) \ \text{with} \ C := \big\{ x \in \mathbb{R}^n \ \big| \ \langle x_i^*, x \rangle \le 0 \ \text{for all} \ i = 1, \dots, m \big\}. \qquad (1.4)$$

The necessary optimality conditions for the controlled sweeping process governed by (1.3) and (1.4), which were derived in [7,8] by using discrete approximations and appropriate tools of generalized differentiation, were then applied in [8] to an optimal control problem for the *crowd motion model in a corridor* [21,37] admitted a sweeping polyhedral description of type (1.3), (1.4).

Note also that other types of optimization problems for some versions of the sweeping process were considered in the literature without using control parameterizations of the moving sets. Control functions appeared there either in additive perturbations [2,4,10,16,32], or in associated ordinary differential equations [1,5]. Necessary optimality conditions for optimal controls in such controlled sweeping models were derived in [4,5] by employing some other methods different from [7,8,12,13] under certain strong smoothness assumptions on the boundaries of compact uncontrolled sweeping sets. In the more recent paper [36], the author addressed relaxation issues for sweeping optimization problems with controls in additive perturbations and uncontrolled convex moving sets that were also included in optimization.

In this paper we study a perturbed sweeping process of type (1.3), where controls enter both perturbations and the moving set given now in the *nonconvex* (and hence nonpolyhedral) form as

$$C(t) := C + u(t) = \bigcap_{i=1}^{m} C_i + u(t) \ \text{with} \ C_i := \big\{ x \in \mathbb{R}^n \big| \ g_i(x) \ge 0 \big\} \ \text{for all} \ i = 1, \dots, m \quad (1.5)$$

defined by some convex and $\mathcal{C}^2$-smooth functions $g_i \colon \mathbb{R}^n \to \mathbb{R}$. Since the set $C(t)$ is nonconvex, an appropriate normal cone notion to $C(t)$ should be specified in (1.3). For definiteness we choose the *proximal* normal cone construction to describe the nonconvex sweeping process under consideration, but actually all the major normal cone notions agree in our setting due the *prox-regularity* of the set $C(t)$ under the assumptions made; see Section 2. Besides being of its own theoretical interest and importance, the controlled sweeping process version from (1.3) and (1.5) arises in applications to optimal control of the *planar crowd motion model*, which is more adequate for the practical use in comparison of the polyhedral corridor version treated in [8]. In fact, this has been our primary motivation for the developments of this paper. The results of the crowd motion applications will be presented in the separate paper [9].

This paper concerns the problem of minimizing the Bolza-type functional

$$\text{minimize} \ \ J[x, u, a] := \varphi\big(x(T)\big) + \int_0^T \ell\big(t, x(t), u(t), a(t), \dot{x}(t), \dot{u}(t), \dot{a}(t)\big) \, dt \qquad (1.6)$$

over the control functions $u(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^n)$ and $a(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^d)$ generating the corresponding trajectories $x(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^n)$ of the sweeping differential inclusion (1.3) with the controlled moving set (1.5), where the time final time $T > 0$ and the initial vector $x_0 \in \mathbb{R}^n$ are fixed. The precise assumptions on terminal extended-real-valued cost function

$\varphi \colon \mathbb{R}^n \to \overline{\mathbb{R}} := (-\infty, \infty]$ and the running cost/integrand $\ell \colon [0, T] \times \mathbb{R}^{4n+2d} \to \overline{\mathbb{R}}$ will be formulated in Section 2. In addition to the above, the *optimal control problem* (P) under consideration contains the pointwise constraints on the controls

$$r_1 \leq \|u(t)\| \leq r_2 \text{ for all } t \in [0, T] \tag{1.7}$$

with the fixed constraint bounds $0 < r_1 \leq r_2$. Note that the two inequality constraints in (1.7) collapse to the equality one when $r_1 = r_2$. The positivity requirement on r_1 in (1.7) is motivated by applications.

It is important to emphasize that due to the construction of the (proximal) normal cone in (1.3) and the moving set structure in (1.5), we implicitly have the pointwise constraints of the other type

$$g_i\big(x(t) - u(t)\big) \geq 0 \text{ for all } t \in [0, T] \text{ and } i = 1, \ldots, m, \tag{1.8}$$

which should be taken into account in the subsequent derivations.

We pursue a *threefold goal* in this paper. The *first one* is to develop the method of *discrete approximations* to study the nonpolyhedral sweeping system in (1.3), (1.5), and (1.7) as well as the optimal control problem (P) for it by constructing a well-posed sequence of discrete-time control systems such that any sweeping feasible solution can be strongly approximated (in the $W^{1,2}$-norm) by feasible ones for discrete systems and that optimal solutions to the discrete counterparts of (P) strongly converge to an optimal solution for the original sweeping control problem (P). The *second goal* is to justify the *existence* of optimal solutions to the discrete problems and to derive *optimality conditions* for them by employing advanced tools of *variational analysis* and *generalized differentiation*, which are appropriate and in fact unavoidable in this framework. The *final goal* is to use discrete approximations as a *vehicle* to establish *necessary optimality conditions* for the *given local optimal* (in an appropriate sense) solution to (P) by passing to the limit from those obtained for their discrete counterparts. The achievement of the latter goal is also heavily based on employing appropriate generalized differential techniques and, in particular, on *second-order subdifferential computations* in variational analysis.

As mentioned, the discrete approximation approach to deriving necessary optimality conditions has been implemented before for Lipschitzian differential inclusions [24,26] as well as for various versions of the polyhedral sweeping process [7,8,13], where the polyhedrality of the moving set was strongly exploited. The nonpolyhedral case of the controlled sweeping process treated below is significantly more involved in comparison with the previous developments in all the major steps of our approach.

The rest of the paper is organized as follows. In Section 2 we recall the needed definitions from variational analysis and generalized differentiation, formulate and discuss the basic assumptions on the initial data of (P), and present some preliminary results widely used below. Section 3 constructs a sequence of discrete approximations for all the constraints of problem (P) simultaneously, without touching optimality so far, and show that *any feasible* solution to (P) can be strongly approximated in the $W^{1,2}$-norm by feasible solutions to the discrete-time inclusions that are piecewise linearly extended to the continuous-time interval $[0, T]$ under fairly general assumptions. This line of the strong approximation is continued in Section 5 for *local optimal* (in the designated sense) solutions to (P), while the preceding Section 4 is devoted to justifying the existence of global optimal solutions to (P) as well as the definition of "intermediate" (between

weak and strong, the latter included) local minimizers and their relaxation that can be covered by the developed method of discrete approximations.

In Section 6 we start preparations to deriving necessary optimality conditions first for discrete-time sweeping problems and then for the original one. These preparations, which are important for their own sake, include computations of the second-order constructions of generalized differentiation that play a significant role in the subsequent results. Section 7 presents the derivation of necessary optimality conditions for discrete approximation problems by reducing them to non-dynamic models of mathematical programming with nonsmooth and nonconvex data together with the usage of generalized differentiation calculus and the second-order computations given above.

Section 8 is a culmination, which establishes a complete set of necessary optimality conditions for the original sweeping control problem (P) by passing to the limit from those obtained in Section 7 for discrete approximations together with rather involved techniques of variational analysis ensuring the appropriate convergence of adjoint trajectories and the validity of the limiting relationships. In Section 9 we present two examples, which are related to practical modeling while illustrating the scheme of applications of the obtained necessary optimality conditions to determine optimal solutions. More important practical applications of the obtained optimality conditions appear in [9].

Throughout the paper we use standard notation of variational analysis and control theory; see, e.g., [26,33,38]. Let us mention that $B(x, r)$ stands for the closed ball of the space in question centered at x with radius $r > 0$, $I\!N := \{1, 2, \ldots\}$, and $x \xrightarrow{\varphi} \bar{x}$ means that $x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$.

2. Basic definitions, assumptions, and preliminaries

First we recall some definitions from variational analysis systematically used in what follows. The framework of this paper is Euclidean and finite-dimensional. We refer the reader to the books [11,25,33] for more details on generalized differentiation and related issues of variational analysis and to the excellent survey by Colombo and Thibault [14] on prox-regularity and its applications.

Let $\Omega \subset \mathbb{R}^n$ be a nonempty set that is locally closed around $\bar{x} \in \Omega$, and let $\text{dist}(x; \Omega) := \inf_{y \in \Omega} \|x - y\|$ be the distance between $x \in \mathbb{R}^n$ and Ω. The *Euclidean projector* of x onto Ω is

$$\Pi(x; \Omega) := \big\{ w \in \Omega \big| \ \|x - w\| = \text{dist}(x; \Omega) \big\}, \quad x \in \mathbb{R}^n,$$

which is nonempty if x is sufficiently close to $\bar{x}$. The *proximal normal cone* to Ω at $\bar{x}$ is given by

$$N^P(\bar{x}; \Omega) := \big\{ v \in \mathbb{R}^n \big| \ \exists \alpha > 0 \ \text{such that} \ \bar{x} \in \Pi(\bar{x} + \alpha v; \Omega) \big\}, \quad \bar{x} \in \Omega, \qquad (2.1)$$

with $N^P(\bar{x}; \Omega) := \emptyset$ if $\bar{x} \notin \Omega$. Another geometric construction of generalized differentiation used below is the (basic/limiting/Mordukhovich) *normal cone* to Ω at $\bar{x} \in \Omega$ defined by

$$N(\bar{x}; \Omega) := \big\{ v \in \mathbb{R}^n \big| \ \exists x_k \to \bar{x}, \ w_k \in \Pi(x_k; \Omega), \ \alpha_k \geq 0 \ \text{s.t.} \ \alpha_k(x_k - w_k) \to v \ \text{as} \ k \to \infty \big\} \qquad (2.2)$$

with $N(\bar{x}; \Omega) := \emptyset$ if $\bar{x} \notin \Omega$. In contrast to the proximal normal cone (2.1), the limiting one (2.2) and the corresponding subdifferential and coderivative constructions for nonsmooth functions

and set-valued mappings, which are generated by (2.2) and are presented in Section 6, enjoy *full calculi* in general settings that are based on the *variational/extremal principles* of variational analysis; see, e.g., [25,33]. There is the following relationship between the limiting and proximal normal cone notions:

$$N(\bar{x}; \Omega) = \left\{ v \in \mathbb{R}^n \,\middle|\, \exists\, x_k \to \bar{x},\ v_k \to v\ \text{with}\ v_k \in N^P(x_k; \Omega)\ \text{for all}\ k \in I\!N \right\}.$$

If the set Ω is convex, both constructions (2.1) and (2.2) reduce to the classical normal cone of convex analysis. But the convexity of Ω is not the only case when $N^P(\bar{x}; \Omega) = N(\bar{x}; \Omega)$, and thus we can combine nice properties of both cones; in particular, the convexity of (2.1) and the rich calculus for (2.2).

It has been well realized in variational analysis that the cones (2.1) and (2.2) agree for a remarkable class of nonconvex sets introduced in variational analysis by Poliquin and Rockafellar [31] under the name of *prox-regularity*. In fact, this notion was first developed by Federer [17] in geometric measure theory under the name of "sets with positive reach." The reader can find more information in [14] with its abounded bibliographies therein. Besides many other applications, prox-regular moving sets have been used in the sweeping process theory; see, e.g., [4,14,16,35, 37]. Our main attention is paid to *uniformly* prox-regular sets, the notion that was probably first developed by Canino [6] in the study of geodesics.

Definition 2.1 *(Uniform prox-regularity).* Let Ω be a closed subset of $\mathbb{R}^n$, and let $\eta > 0$. Then Ω is η-PROX-REGULAR if for all $x \in \mathrm{bd}\,\Omega$ and $v \in N^P(x; \Omega)$ with $\|v\| = 1$ we have $B(x + \eta v, \eta) \cap \Omega = \{x\}$. Equivalently, the η-prox-regularity of Ω can be defined via the validity of the estimate

$$\langle v, y - x \rangle \leq \frac{\|v\|}{2\eta} \|y - x\|^2 \ \text{for all}\ y \in \Omega,\ x \in \mathrm{bd}\,\Omega,\ \text{and}\ v \in N^P(x; \Omega).$$

Recall that any closed convex subset in $\mathbb{R}^n$ is ∞-prox-regular and that in this case for every $\eta > 0$ the well-defined Euclidean projection operator $\Pi(x; \Omega)$ is single-valued if $\mathrm{dist}(x; \Omega) < \eta$.

Throughout the paper we impose the following assumptions on the given data of the optimal control problem (P) ensuring, in particular, that for each $t \in [0, T]$ the controlled moving set $C(t)$ in (1.5) is uniformly prox-regular, and thus the proximal and limiting normal cones agree for it. This allows us to use the normal cone notation "N" in (1.3) in the rest of the paper and employ the results available in variational analysis for either one of these cones. Here are our *standing assumptions*:

(H1) The perturbation mapping $f : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$ in (1.3) is continuous on $\mathbb{R}^n \times \mathbb{R}^d$ and locally Lipschitzian with respect to the first argument, i.e., for every $\varepsilon > 0$ there is a constant $K > 0$ such that

$$\|f(x, a) - f(y, a)\| \leq K\|x - y\| \ \text{whenever}\ (x, y) \in B(0, \varepsilon) \times B(0, \varepsilon),\ a \in \mathbb{R}^d. \qquad (2.3)$$

Furthermore, there is a constant $M > 0$ ensuring the growth condition

$$\|f(x, a)\| \leq M(1 + \|x\|) \ \text{for any}\ x \in \bigcup_{t \in [0,T]} C(t),\ a \in \mathbb{R}^d. \qquad (2.4)$$

(H2) There exist positive constants $c > 0$ and M_j, $j = 1, 2, 3$, together with open sets $V_i \supset C_i$ such that $d_H(C_i, \mathbb{R}^n \setminus V_i) > c$ and the functions $g_i(\cdot)$, $i = 1, \ldots, m$, are twice continuously differentiable (C^2-smooth) satisfying the estimates

$$M_1 \leq \|\nabla g_i(x)\| \leq M_2, \text{ and } \|\nabla^2 g_i(x)\| \leq M_3 \text{ for all } x \in V_i, \tag{2.5}$$

where d_H stands for the Hausdorff distance between sets.

(H3) There exist positive numbers β and ρ such that

$$\sum_{i \in I_\rho(x)} \lambda_i \|\nabla g_i(x)\| \leq \beta \Big\| \sum_{i \in I_\rho(x)} \lambda_i \nabla g_i(x) \Big\| \text{ for all } x \in C \text{ and } \lambda_i \geq 0, \tag{2.6}$$

where the index set for the perturbed constraints is defined by

$$I_\rho(x) := \{ i \in \{1, \ldots, m\} \mid g_i(x) \leq \rho \}. \tag{2.7}$$

(H4) The terminal cost $\varphi \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is lower semicontinuous (l.s.c.), while the running cost ℓ in (1.6) is such that $\ell_t := \ell(t, \cdot) \colon \mathbb{R}^{4n+2d} \to \overline{\mathbb{R}}$ is l.s.c. for a.e. $t \in [0, T]$, bounded from below on bounded sets, and $t \mapsto \ell(t, x(t), u(t), a(t), \dot{x}(t), \dot{u}(t), \dot{a}(t))$ is summable on $[0, T]$ for each feasible triple $(x(t), u(t), a(t))$.

Observe that the simultaneous validity of (2.5) and (2.6) imply the *positive linear independence* of the gradients $\nabla g_i(x)$ of the active inequality constraints on C, and that it reduces to the validity of the classical Mangasarian–Fromovitz constraint qualification on C in the setting under consideration.

The following proposition is due to the result by Venel [37, Proposition 2.9]; see also [3] for more discussions and further developments.

Proposition 2.2 (*Uniform prox-regularity of the moving set*). *Under the validity of* (H2) *and* (H3) *we have that for each* $t \in [0, T]$ *the set* $C(t)$ *is* η-*prox-regular with* $\eta = \dfrac{\alpha}{M_3 \beta}$.

Proof. It follows from [37, Proposition 2.9] that the set C in (1.5) is η-prox-regular with the modulus η defined in the proposition. Thus it holds for the moving set $C(t) = C + u(t)$ as a translation of C. $\square$

Proposition 2.2 allows us to verify the next proposition based on the well-posedness result by Edmond and Thibault taken from [16, Theorem 1].

Proposition 2.3 (*Existence and estimates for sweeping trajectories*). *Consider the perturbed sweeping process* (1.3) *with the fixed controls* $u(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^n)$ *and* $a(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^d)$ *under the validity of* (H1)–(H3) *with the constant* $M > 0$ *taken from* (2.4). *There is a unique solution* $x(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^n)$ *to* (1.3) *generated by the controls* $(u(\cdot), a(\cdot))$. *Furthermore, we have*

$$\|x(t)\| \leq l := \|x_0\| + e^{2MT}\left(2MT(1 + \|x_0\|) + \int_0^T \|\dot{u}(s)\| ds \right) \text{ for all } t \in [0, T],$$

$$\|\dot{x}(t)\| \leq 2(1 + l)M + \|\dot{u}(t)\| \text{ a.e. } t \in [0, T]. \tag{2.8}$$

Proof. With the fixed pair $(u(\cdot), a(\cdot))$, both existence and estimate statements of the theorem follow from [16, Theorem 1] under the validity of (H1), the uniform prox-regularity of $C(t)$, and the property

$$\left| \text{dist}(y; C(t)) - \text{dist}(y; C(\tau)) \right| \leq \|v(t) - v(\tau)\| \text{ for all } t, \tau \in [0, T] \text{ with } v(t) := \int_0^t \|\dot{u}(s)\| ds$$

and with the chosen $W^{1,2}$ control $u(\cdot)$. The latter fact was proved in [7, Proposition 1], while the uniform prox-regularity of $C(t)$ follows from Proposition 2.2 under assumptions (H2) and (H3). □

3. Discrete approximation of nonconvex sweeping process

This section deals with the constrained nonconvex sweeping process described by (1.3), (1.5), and (1.7) without considering its optimization. In what follows we construct a sequence of discrete-time counterparts of the constrained sweeping process in such a way that any feasible triple to the continuous-time process (including controls $u(t)$, $a(t)$ and the corresponding trajectory $x(t)$ satisfying rather unrestrictive conditions) can be *strongly* approximated in the $W^{1,2}$-norm by feasible solutions to the discrete-time systems that are piecewise linearly extended on the continuous-time interval $[0, T]$.

The first step of this procedure is used in all the results presented below. Unifying the control and state variables, we introduce the triple $z := (x, u, a) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$ and show that (1.3) with $C(t)$ from (1.5) can be written in the usual form of differential inclusions with respect to the new variable z. Indeed, define the set-valued mapping $F \colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ by

$$F(z) = F(x, u, a) := N(x - u; C) + f(x, a) \tag{3.1}$$

and deduce from [37, Proposition 2.8] that F admits the explicit representation

$$F(z) = \left\{ - \sum_{i \in I(x-u)} \lambda_i \nabla g_i(x - u) \,\middle|\, \lambda_i \geq 0 \right\} + f(x, a) \tag{3.2}$$

via the index set of active constraints

$$I(y) := \left\{ i \in \{1, \ldots, m\} \,\middle|\, g_i(y) = 0 \right\} \tag{3.3}$$

at the point $y := x - u \in C$. Then we can equivalently rewrite the sweeping inclusion (1.3) with the set C from (1.5) in the following equivalent form involving one variable $z \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$:

$$-\dot{z}(t) \in F(z(t)) \times \mathbb{R}^n \times \mathbb{R}^d \text{ a.e. } t \in [0, T] \text{ with } z(0) := (x_0, u(0), a(0)), \ x_0 - u(0) \in C, \tag{3.4}$$

where the last condition means that $g_i(x_0 - u(0)) \geq 0$ as $i = 1, \ldots, m$. Proposition 2.3 tells us that the Cauchy problem in (3.4) has solutions in the class of $W^{1,2}$ functions $z(t) = (x(t), u(t), a(t))$ on $[0, T]$.

For each $k \in I\!N$ we define the discrete partitions of $[0, T]$ by

$$\Delta_k := \{0 = t_0^k < t_1^k < \ldots < t_k^k\} \text{ with } h_k := t_{j+1}^k - t_j^k \downarrow 0 \text{ as } k \to \infty. \tag{3.5}$$

The next theorem justifies the desired strong discrete approximation for a large class of feasible solutions to the continuous-time inclusions (3.4) generated by the nonconvex sweeping set (1.5) with perturbing the pointwise control constraints (1.7) and keeping the implicit state constraints (1.8).

Theorem 3.1 (*Strong discrete approximation of feasible solutions to the constrained sweeping process*). *Let the assumptions in* (H1)–(H3) *be satisfied. Consider a triple* $\bar{z}(\cdot) = (\bar{x}(\cdot), \bar{u}(\cdot), \bar{a}(\cdot))$ *satisfying* (3.4), (1.5), *and* (1.7) *as well as the following properties, which all hold whenever* $\bar{z} \in W^{2,\infty}([0, T])$: *the sweeping inclusion* (3.4) *is fulfilled for* $\bar{z}(\cdot)$ *at the partition points* t_j^k *from* (3.5) *as* $j = 0, \ldots, k - 1$ *with the right-side derivative at* $t_0 = 0$ *and for some constant* $\mu > 0$ *independent of k we have*

$$\sum_{j=0}^{k-1} \left\| \frac{\bar{x}(t_{j+1}^k) - \bar{x}(t_j^k)}{h_k} - \dot{\bar{x}}(t_j^k) \right\| \le \mu, \quad \left\| \frac{\bar{u}(t_1^k) - \bar{u}(t_0^k)}{h_k} \right\| \le \mu,$$
$$\sum_{j=0}^{k-2} \left\| \frac{\bar{u}(t_{j+2}^k) - \bar{u}(t_{j+1}^k)}{h_k} - \frac{\bar{u}(t_{j+1}^k) - \bar{u}(t_j^k)}{h_k} \right\| \le \mu. \tag{3.6}$$

Then there exist a sequence of piecewise linear functions $z^k(t) := (x^k(t), u^k(t), a^k(t))$ *on* $[0, T]$ *and a sequence of positive numbers* $\varepsilon_k \le 2h_k \mu e^K \downarrow 0$ *as* $k \to \infty$ *with the constant* $K > 0$ *taken from* (2.3) *so that* $(x^k(0), u^k(0), a^k(0)) = (x_0, \bar{u}(0), \bar{a}(0))$,

$$r_1 - \varepsilon_k \le \|u^k(t_j^k)\| \le r_2 + \varepsilon_k, \quad t_j^k \in \Delta_k, \tag{3.7}$$

$$x^k(t) = x^k(t_j^k) - (t - t_j^k)v_j^k, \ x^k(0) = x_0, \ t_j^k \le t \le t_{j+1}^k \text{ with } v_j^k \in F(z^k(t_j^k)) \tag{3.8}$$

for $j = 0, \ldots, k - 1$ *and that* $z^k(\cdot) \to \bar{z}(\cdot)$ *in the* $W^{1,2}$-*norm, i.e.,*

$$z^k(\cdot) \to \bar{z}(\cdot) \quad \text{uniformly on} \quad [0, T] \quad \text{and} \quad \int_0^T \left\| \dot{z}^k(t) - \dot{\bar{z}}(t) \right\|^2 dt \to 0 \quad \text{as} \quad k \to \infty. \tag{3.9}$$

Moreover, with $\widetilde{\mu} := \max\{3\mu(1 + 4KT)e^K, 4\mu(e^K + 1)\}$ *we have the estimates*

$$\left\| \frac{u^k(t_1^k) - u^k(t_0^k)}{h_k} \right\| \le \widetilde{\mu} \text{ and } \mathrm{var}\big(\dot{u}^k; [0, T]\big) \le \widetilde{\mu}, \tag{3.10}$$

where "var" denotes the total variation on $[0, T]$ *of the derivative function in* (3.10).

Proof. It mainly follows the lines in the proof of [7, Theorem 3.1], although the problems under consideration in [7] and here are significantly different. We just present the major constructions in the new setting of this theorem. Define first a sequence of piecewise linear functions $y^k(\cdot) := (y_1^k(\cdot), y_2^k(\cdot), y_3^k(\cdot))$ on $[0, T]$ via their values at the mesh points of Δ_k by

$$\left(y_1^k(t_j^k), y_2^k(t_j^k), y_3^k(t_j^k)\right) := \left(\bar{x}(t_j^k), \bar{u}(t_j^k), \bar{a}(t_j^k)\right) \text{ for all } j = 0, \ldots, k \text{ and } k \in I\!N.$$

Define further $w^k(t) = (w_1^k(t), w_2^k(t), w_3^k(t)) := \dot{y}^k(t)$ as piecewise constant and right continuous functions on $[0, T]$ and easily deduce from these constructions that

$$y^k(\cdot) \to \bar{z}(\cdot) \text{ uniformly on } [0, T] \text{ and } w^k(\cdot) \to \dot{\bar{z}}(\cdot) \text{ strongly in } L^2([0, T]; \mathbb{R}^{2n+d}).$$

For each fixed $k \in I\!N$, denote $a^k(t) := y_3^k(t)$ on $[0, T]$ and use for simplicity the notation $t_j := t_j^k$ as $j = 1, \ldots, k$ in what follows. We construct the desired trajectories $x^k(t)$ of (3.8) by induction. Suppose that the value of $x^k(t_j)$ is known and define the vectors

$$u^k(t_j) := x^k(t_j) - y_1^k(t_j) + y_2^k(t_j) = x^k(t_j) - \bar{x}(t_j) + \bar{u}(t_j), \quad j = 0, \ldots, k,$$

for which we clearly have the relationships

$$x^k(t_j) - u^k(t_j) = \bar{x}(t_j) - \bar{u}(t_j) \text{ as } j = 0, \ldots, k \text{ and} \tag{3.11}$$

$$r_1 - \varepsilon_k \leq \|u^k(t_j)\| \leq r_2 + \varepsilon_k \text{ with } \varepsilon_k := \|x^k(t_j) - \bar{x}(t_j)\|.$$

This yields $g_i(x^k(t_j) - u^k(t_j)) = g_i(\bar{x}(t_j) - \bar{u}(t_j)) \geq 0$ for $i = 1, \ldots, m$ and thus $x^k(t_j) - u^k(t_j) \in C$, which shows that the values $F(z^k(t_j)) = F(x^k(t_j), u^k(t_j), a^k(t_j))$ of the mapping (3.1) are well defined whenever $j = 0, \ldots, k$. Set $x^k(t) := x^k(t_j) - (t - t_j)v_j^k$ with $v_j^k \in \Pi(-w_{1j}^k; F(z^k(t_j)))$ for all $t \in [t_j, t_{j+1})$ and $j = 0, \ldots, k - 1$ and then deduce from (3.2) and (3.11) that

$$F\left(x^k(t_j), u^k(t_j), a^k(t_j)\right) = F\left(\bar{x}(t_j), \bar{u}(t_j), \bar{a}(t_j)\right) + f\left(x^k(t_j), \bar{a}(t_j)\right) - f\left(\bar{x}(t_j), \bar{a}(t_j)\right).$$

Employing now the arguments similar to the proof of [7, Theorem 3.1], we readily verify that the triples $z^k(t) = (x^k(t), u^k(t), a^k(t))$ for $t \in [0, T]$ and $k \in I\!N$ constructed above satisfy all the conclusions of this theorem and thus complete the proof of this theorem. $\quad\square$

4. Existence of optimal solutions and local relaxation

This section starts the study of the entire sweeping optimal control problem (P) formulated in Section 1, not only its feasible solutions. First we establish the existence of optimal solutions to (P) with adding the convexity of the running cost ℓ of (1.6) in the velocity variables to our standing assumptions.

Theorem 4.1 (*Existence of optimal solutions to the sweeping optimal control problem*). *Let (P) be the optimal control problem defined in Section 1 considering in the equivalent form (3.4) of the sweeping differential inclusion over all the $W^{1,2}([0, T])$ triples $z(\cdot) = (x(\cdot), u(\cdot), a(\cdot))$. In addition to the standing assumptions in* (H1)–(H4), *suppose that the integrand ℓ in (1.6) is*

convex with respect to the velocity variables $(\dot{x}, \dot{u}, \dot{a})$ *and that along a minimizing sequence of* $z^k(\cdot) = (x^k(\cdot), u^k(\cdot), a^k(\cdot))$ *as* $k \in I\!N$ *we have that* $\ell(t, \cdot)$ *is majorized by a summable function,* $\{\dot{u}^k(\cdot)\}$ *is bounded in* $L^2([0, T]; \mathbb{R}^n)$ *and* $\{a^k(\cdot)\}$ *is bounded in* $W^{1,2}([0, T]; \mathbb{R}^d)$. *Then problem* (P) *admits an optimal solution in* $W^{1,2}([0, T]; \mathbb{R}^{2n+d})$.

Proof. It follows from Proposition 2.3 that the set of feasible solutions to (P) is nonempty. Take the minimizing sequence $(x^k(\cdot), u^k(\cdot), a^k(\cdot)) \in W^{1,2}([0, T]; \mathbb{R}^{2n+d})$ in (P) from the formulation of the theorem, where $x^k(\cdot)$ is uniquely generated by $(u^k(\cdot), a^k(\cdot))$ in Proposition 2.3. The imposed boundedness assumptions on $\{u^k(\cdot)\}$ and $\{a^k(\cdot)\}$ yield by standard functional analysis that the sequence $\{(\dot{u}^k(\cdot), \dot{a}^k(\cdot))\}$ is weakly compact in $L^2([0, T]; \mathbb{R}^{n+d})$, and so we have – by passing to subsequences if necessary – the weak convergence $\dot{u}^k(\cdot) \to \vartheta^u(\cdot)$ and $\dot{a}^k(\cdot) \to \vartheta^a(\cdot)$ in $L^2([0, T]; \mathbb{R}^n)$ and $L^2([0, T]; \mathbb{R}^d)$, respectively, for some functions $\vartheta^u(\cdot)$ and $\vartheta^a(\cdot)$ from the corresponding spaces. Due to pointwise constraints (1.7) and the boundedness of $\{a^k(0)\}$, suppose without loss of generality that $u^k(0) \to u_0$ and $a^k(0) \to a_0$ as $k \to \infty$ for some $u_0 \in \mathbb{R}^n$ and $a_0 \in \mathbb{R}^d$. Then $(u^k(\cdot), a^k(\cdot)) \to (\bar{u}(\cdot), \bar{a}(\cdot))$ in the norm of $W^{1,2}([0, T]; \mathbb{R}^{n+d})$ for the functions $\bar{u}(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^n)$ and $\bar{a}(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^d)$ defined by

$$\bar{u}(t) := u_0 + \int_0^t \vartheta^u(s)ds \ \text{ and } \ \bar{a}(t) := a_0 + \int_0^t \vartheta^a(s)ds, \tag{4.1}$$

which implies, in particular, that $\bar{u}(\cdot)$ satisfies the constraints in (1.7). Furthermore, it follows from estimate (2.8) in Proposition 2.3 that the sequence $\{\dot{x}^k(\cdot)\}$ is bounded in $L^2([0, T]; \mathbb{R}^n)$, and hence we get $\vartheta^x(\cdot) \in L^2([0, T]; \mathbb{R}^n)$ for which $\dot{x}^k(\cdot) \to \vartheta^x(\cdot)$ weakly in $L^2([0, T]; \mathbb{R}^n)$ along a subsequence. It yields the convergence $x^k(\cdot) \to \bar{x}(\cdot)$ in the norm topology of $W^{1,2}([0, T]; \mathbb{R}^n)$ for $\bar{x}(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^n)$ defined by

$$\bar{x}(t) := x_0 + \int_0^t \vartheta^x(s)ds, \quad t \in [0, T].$$

The next step is to verify that the limiting triple $\bar{z}(\cdot) = (\bar{x}(\cdot), \bar{u}(\cdot), \bar{a}(\cdot))$ satisfies the differential inclusion (3.4) with mapping $F(z)$ given in (3.1). Since the derivative sequences $\{\dot{z}^k(\cdot)\}$ converges to $\dot{\bar{z}}(\cdot)$ weakly in $L^2([0, T]; \mathbb{R}^{2n+d})$, the classical Mazur theorem ensures the strong convergence to $\dot{\bar{z}}(\cdot)$ in this space of some sequence of convex combinations of the functions $\dot{z}^k(t)$. Thus there is a subsequence of these convex combinations that converges to $\dot{\bar{z}}(t)$ for a.e. $t \in [0, T]$. It follows from the above that there exist a function $\nu : I\!N \to I\!N$ and sequences of real numbers $\{\alpha(k)_j\}$, $j = k, \ldots, \nu(k)$, such that

$$\alpha(k)_j \geq 0, \ \sum_{j=k}^{\nu(k)} \alpha(k)_j = 1, \ \text{ and } \ \sum_{j=k}^{\nu(k)} \alpha(k)_j \dot{z}^j(t) \to \dot{\bar{z}}(t) \ \text{ a.e. } t \in [0, T]$$

as $k \to \infty$. By taking into account that $\bar{x}(t) - \bar{u}(t) = \lim_{k \to \infty} (x^k(t) - u^k(t)) \in C$ we get

$$-\dot{\bar{x}}(t) - f(\bar{x}(t), \bar{a}(t)) = \lim_{k \to \infty} \left(-\sum_{j=k}^{\nu(k)} \alpha(k)_j \dot{x}^j(t) - \sum_{j=k}^{\nu(k)} \alpha(k)_j f(x^j(t), a^j(t)) \right)$$

$$= \lim_{k \to \infty} \Big(- \sum_{j=k}^{v(k)} \sum_{i \in I(x^j(t) - u^j(t))} \alpha(k)_j \lambda_i^j \nabla g_i \big(x^j(t) - u^j(t) \big) \Big)$$

$$= \lim_{k \to \infty} \Big(- \sum_{j=k}^{v(k)} \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \alpha(k)_j \lambda_i^j \nabla g_i \big(x^j(t) - u^j(t) \big) \Big)$$

$$= \lim_{k \to \infty} \Big(- \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(x^j(t) - u^j(t) \big) \Big),$$

where $I(\cdot)$ is taken from (3.3), and where $\lambda_i^j = 0$ if $i \in I(\bar{x}(t) - \bar{u}(t)) \setminus I(x^j(t) - u^j(t))$ due to the clear inclusion $I(x^j(t) - u^j(t)) \subset I(\bar{x}(t) - \bar{u}(t))$ for all $j = k, \ldots, v(k)$ and all large $k \in I\!N$.

Let us now show that the numerical sequence $\Big\{ \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \Big\}$ is bounded for all $i \in I(\bar{x}(t) - \bar{u}(t))$. Indeed, we deduce from (2.5) and (2.6) the validity of the estimates

$$\sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \leq \frac{1}{M_1} \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \Big(\sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \Big) \big\| \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) \big\|$$

$$\leq \frac{\beta}{M_1} \Big\| \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \Big(\sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \Big) \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) \Big\|$$

$$\leq \frac{\beta}{M_1} \Big\| \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big)$$

$$- \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(x^j(t) - u^j(t) \big) \Big\|$$

$$+ \frac{\beta}{M_1} \Big\| \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(x^j(t) - u^j(t) \big) \Big\|$$

$$\leq \frac{\beta}{M_1} \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \big\| \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) - \nabla g_i \big(x^j(t) - u^j(t) \big) \big\| + \frac{\beta \widetilde{M}}{M_1}$$

$$\leq \frac{1}{2} \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j + \frac{\beta \widetilde{M}}{M_1}$$

for all k sufficiently large, where $\widetilde{M}$ is an upper bound of $\Big\{ \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i (x^j(t) - u^j(t)) \Big\}$. This justifies the boundedness of the sequence $\Big\{ \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \Big\}$. Thus there exists $\widetilde{\alpha} > 0$

such that $\left\| \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \right\| \leq \widetilde{\alpha}$ and $\sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \to \beta_i$ as $k \to \infty$ along some subsequence, with no relabeling.

Next we verify that $\sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i (x^j(t) - u^j(t)) \to \beta_i \nabla g_i(\bar{x}(t) - \bar{u}(t))$ as $k \to \infty$. Observe that

$$
\left\| \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(x^j(t) - u^j(t) \big) - \beta_i \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) \right\|
$$

$$
\leq \left\| \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(x^j(t) - u^j(t) \big) - \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) \right\|
$$

$$
+ \left\| \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) - \beta_i \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) \right\|
$$

$$
\leq \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \left\| \nabla g_i \big(x^j(t) - u^j(t) \big) - \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) \right\| + M_2 \left\| \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j - \beta_i \right\|
$$

$$
\leq \widetilde{\alpha}\varepsilon + M_2 \varepsilon \quad \text{for all large } k \in I\!N,
$$

where M_2 is taken in (H2) while $\varepsilon > 0$ is an arbitrary small number with

$$
\max \left\{ \left\| \nabla g_i \big(x^j(t) - u^j(t) \big) - \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) \right\|, \left\| \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j - \beta_i \right\| \right\} \leq \varepsilon
$$

whenever $k \leq j \leq v(k)$ and k is large enough. Thus we have

$$
-\dot{\bar{x}}(t) - f\big(\bar{x}(t), \widetilde{a}(t) \big) = \lim_{k \to \infty} \left(- \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \sum_{j=k}^{v(k)} \alpha(k)_j \lambda_i^j \nabla g_i \big(x^j(t) - u^j(t) \big) \right)
$$

$$
= - \sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \beta_i \nabla g_i \big(\bar{x}(t) - \bar{u}(t) \big) \in N \big(\bar{x}(t) - \bar{u}(t); C \big) \quad \text{a.e. } t \in [0, T],
$$

which verifies by (3.2) that $\bar{z}(\cdot)$ satisfies the differential inclusion (3.4) and hence the constraints in (1.8).

To justify the optimality of the triple $\bar{z}(\cdot)$ in (P), it remains to show that

$$
J[\bar{x}, \bar{u}, \bar{a}] \leq \liminf_{k \to \infty} J[x^k, u^k, a^k]
$$

for the Bolza functional (1.6). But this is a clear consequence of the aforementioned Mazur theorem due the imposed convexity of the integrand ℓ in the velocity variables and the Lebesgue dominated convergence theorem for passing to the limit under the integral sign. □

Note that the convexity of the running cost ℓ in (1.6) is not among the standing assumptions of the paper and is not needed for deriving our main results on necessary optimality conditions in (P). Such conditions established below address the so-called "relaxed intermediate local minimizers" introduced in [24] and then studied in many publications. To recall this notion, we first consider the relaxed optimal control problem for (P) following the Bogoluybov–Young relaxation/convexification procedure, which has been well understood in the classical calculus of variations and optimal control; see, e.g., [26] with the references therein and [10,15,16,36] for more recent results in this direction. To this end, denote by $\widehat{\ell}_F(t, x, u, a, \dot{x}, \dot{u}, \dot{a})$ the convexification (the largest l.s.c. convex function majorized by $\ell(t, x, u, a, \cdot, \cdot, \cdot)$) of the running cost in (1.6) on the set $F(x, u, a)$ from (3.1) with respect to the velocity variables $(\dot{x}, \dot{u}, \dot{a})$ and put $\widehat{\ell} :\equiv \infty$ at points out of $F(x, u, a)$. Define the *relaxed optimal control problem* (R) by

$$\text{minimize } \widehat{J}[z] = \widehat{J}[x, u, a] := \varphi\big(x(T)\big) + \int_0^T \widehat{\ell}_F\big(t, x(t), u(t), a(t), \dot{x}(t), \dot{u}(t), \dot{a}(t)\big)dt \quad (4.2)$$

over all $z(\cdot) = (x(\cdot), u(\cdot), a(\cdot)) \in W^{1,2}([0, T])$ satisfying (1.5), (1.7), and (3.4). Besides the obvious case of integrands that are convex in velocity variables, there are broad classes of variational and control problems over continuous-time intervals where optimal values of the cost functionals in the (nonconvex) original and relaxed problems agree; it is known as "relaxation stability." This is due to some "hidden convexity" for such problems (we refer again to [26, Chapter 6] and the commentaries therein), which allows us, in particular, to verify the relaxation stability of nonconvex Bolza problems for Lipschitzian differential inclusions and also for those satisfying a certain one-sided Lipschitzian condition [15]. Unfortunately, neither of the aforementioned conditions is fulfilled for the controlled sweeping process under consideration. For a sweeping process over prox-regular moving sets with controls only in additive perturbations, the relaxation stability follows from the result by Edmond and Thibault [16, Theorem 2]; see also [10] for similar relaxation results concerning BV solutions. It seems that the closest to our setting is a recent result by Tolstonogov [36, Theorem 4.2], which establishes the relaxation stability for a sweeping process over variable convex moving sets involved in optimization together with controls in perturbations.

Implementing the general scheme of [24,26] for the case of the sweeping control problem under consideration, recall now the notion of local minimizers studied in this paper.

Definition 4.2 *(Local minimizers).* We say that $\bar{z}(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^{2n+d})$ is a RELAXED INTERMEDIATE LOCAL MINIMIZER (r.i.l.m.) of rank 2 for (P) if it is feasible to (P), $J[\bar{z}] = \widehat{J}[\bar{z}]$, and there is $\varepsilon > 0$ such that $J[\bar{z}] \leq J[z]$ for any feasible solution $z(\cdot)$ to (P) satisfying the conditions

$$\|z(t) - \bar{z}(t)\| < \varepsilon \text{ whenever } t \in [0, T] \text{ and } \int_0^T \|\dot{z}(t) - \dot{\bar{z}}(t)\|^2 dt < \varepsilon. \quad (4.3)$$

For convenience we concentrate here on relaxed intermediate local minimizers of rank 2 and skip mentioning the rank in what follows. The name "intermediate" comes from the fact that the introduced notion clearly lies (strictly) between the conventional notions of weak and strong local minimizers in the calculus of variations and optimal control; see [26] for more discussions.

Note also that in the case of $J[\bar{z}] = \widehat{J}[\bar{z}]$ (in particular, if (P) has the property of relaxation stability), there is no difference between relaxed intermediate local minimizers and merely intermediate local minimizers (without relaxation), which were also defined in [24]. Thus we can treat r.i.l.m. as a *local* version of relaxation stability. It can be distilled from the proofs of [16, Theorem 2] and [36, Theorem 4.2] that no relaxation is needed provided that $\bar{z}(\cdot)$ is a *strong* local minimizer of (P) and, in addition to our standing assumptions, either controls are presented only in perturbations, or the set C in (1.5) is convex.

5. Discrete approximation of intermediate local minimizers

In this section we continue with developing the method of discrete approximation, while now paying our main attention not to constructing such approximations of *any feasible* solution to the constrained system (3.4), (1.5), and (1.7) as in Section 3 but to the given *local optimal solution* $\bar{z}(\cdot) = (\bar{x}(\cdot), \bar{u}(\cdot), \bar{a}(\cdot))$ (in the *r.i.l.m.* sense) of this system with respect to the cost functional (1.6), which we can write as

$$\text{minimize } J[z] = \varphi\big(z(T)\big) + \int_0^T \ell\big(t, z(t), \dot{z}(t)\big) dt \tag{5.1}$$

with $\varphi(z) := \varphi(x)$. It means that our goal here is to construct discrete approximations of the entire problem (P) including its cost functional and the given r.i.l.m. for which we aim subsequently to derive necessary optimality conditions by employing the discrete approximation method.

For any fixed $k \in I\!N$ we define the *discrete sweeping control problem* (P_k) as follows: minimize

$$J_k[z^k] := \varphi(x_k^k) + h_k \sum_{j=0}^{k-1} \ell\left(t_j^k, x_j^k, u_j^k, a_j^k, \frac{x_{j+1}^k - x_j^k}{h_k}, \frac{u_{j+1}^k - u_j^k}{h_k}, \frac{a_{j+1}^k - a_j^k}{h_k}\right)$$

$$+ \sum_{j=0}^{k-1} \int_{t_j^k}^{t_{j+1}^k} \left(\left\|\frac{x_{j+1}^k - x_j^k}{h_k} - \dot{\bar{x}}(t)\right\|^2 + \left\|\frac{u_{j+1}^k - u_j^k}{h_k} - \dot{\bar{u}}(t)\right\|^2 + \left\|\frac{a_{j+1}^k - a_j^k}{h_k} - \dot{\bar{a}}(t)\right\|^2\right)$$

$$+ \text{dist}^2\left(\left\|\frac{u_1^k - u_0^k}{h_k}\right\|; (-\infty, \widetilde{\mu}]\right) + \text{dist}^2\left(\sum_{j=0}^{k-2}\left\|\frac{u_{j+2}^k - 2u_{j+1}^k + u_j^k}{h_k}\right\|; (-\infty, \widetilde{\mu}]\right) \tag{5.2}$$

over elements $z^k := (x_0^k, x_1^k, \ldots, x_k^k, u_0^k, u_1^k, \ldots, u_{k-1}^k, a_0^k, a_1^k, \ldots, a_{k-1}^k)$ satisfying the constraints

$$x_{j+1}^k \in x_j^k - h_k F(x_j^k, u_j^k, a_j^k) \text{ for } j = 0, \ldots, k-1 \text{ with } (x_0^k, u_0^k, a_0^k) = \big(x_0, \bar{u}(0), \bar{a}(0)\big), \tag{5.3}$$

$$g_i(x_k^k - u_k^k) \geq 0 \text{ for } i = 1, \ldots, m, \tag{5.4}$$

$$r_1 - \varepsilon_k \leq \|u_j^k\| \leq r_2 + \varepsilon_k \text{ for } j = 0, \ldots, k, \tag{5.5}$$

$$\|(x_j^k, u_j^k, a_j^k) - (\bar{x}(t_j^k), \bar{u}(t_j^k), \bar{a}(t_j^k))\| \leq \varepsilon/2 \text{ for } j = 0, \ldots, k-1, \tag{5.6}$$

$$\sum_{j=0}^{k-1} \int_{t_j^k}^{t_{j+1}^k} \left(\left\| \frac{x_{j+1}^k - x_j^k}{h_k} - \dot{x}(t) \right\|^2 + \left\| \frac{u_{j+1}^k - u_j^k}{h_k} - \dot{u}(t) \right\|^2 + \left\| \frac{a_{j+1}^k - a_j^k}{h_k} - \dot{a}(t) \right\|^2 \right) dt \leq \frac{\varepsilon}{2},$$

(5.7)

$$\left\| \frac{u_1^k - u_0^k}{t_1^k - t_0^k} \right\| \leq \widetilde{\mu} + 1 \text{ and } \sum_{j=0}^{k-2} \left\| \frac{u_{j+2}^k - 2u_{j+1}^k + u_j^k}{h_k} \right\| \leq \widetilde{\mu} + 1,$$

(5.8)

where ε is taken from (4.3) while ε_k and $\widetilde{\mu}$ are taken from Theorem 3.1 applied to the given r.i.l.m. $\bar{z}(\cdot)$.

To study $\bar{z}(\cdot)$ via the method of discrete approximations, we have to verify first that all the problems (P_k) for each $k \in I\!N$ sufficiently large admit optimal solutions.

Proposition 5.1 *(Existence of optimal solutions to discrete sweeping control problems). Under the standing assumptions* (H1)–(H4) *holding along the r.i.l.m.* $\bar{z}(\cdot)$ *each problem* (P_k) *for large* $k \in I\!N$ *admits an optimal solution* $\bar{z}^k(\cdot)$.

Proof. It follows directly from Theorem 3.1 and the construction of (P_k) along the given r.i.l.m. $\bar{z}(\cdot)$ that the set of feasible solutions to (P_k) is nonempty for all large $k \in I\!N$. Furthermore, the imposed constraints (5.5)-(5.7) ensure that this set is bounded for each k. Thus the existence of optimal solutions to finite-dimensional problems (P_k) is ensures by the classical the Weierstrass existence theorem provided that the feasible solution set to each problem (P_k) is closed. To check it, take a sequence $z^\nu(\cdot) = z^\nu := (x_0^\nu, \ldots, x_k^\nu, u_0^\nu, \ldots, u_{k-1}^\nu, a_0^\nu, \ldots, a_{k-1}^\nu)$ of feasible solutions for (P_k) converging to some $z(\cdot) = z := (x_0, \ldots, x_k, u_0, \ldots, u_{k-1}, a_0, \ldots, a_{k-1})$ as $\nu \to \infty$. We need to show that z is feasible to (P_k). Observing that $g_i(x_j - u_j) = \lim_{\nu \to \infty} g_i(x_j^\nu - u_j^\nu) \geq 0$ for all $i = 1, \ldots, m$ and $j = 0, \ldots, k-1$ gives us that $x_j - u_j \in C$ for all $j = 0, \ldots, k-1$. It is not hard to see that $I(x_j^\nu - u_j^\nu) \subset I(x_j - u_j)$ for $\nu \in I\!N$ sufficiently large. Taking (3.1) and (3.4) into account, gives us for all such indices j that

$$\frac{x_{j+1}^\nu - x_j^\nu}{-h_k} - f(x_j^\nu, a_j^\nu) \in N_C(x_j^\nu - u_j^\nu),$$

which implies therefore by passing to the limit as $\nu \to \infty$ that

$$\frac{x_{j+1}^\nu - x_j^\nu}{-h_k} - f(x_j^\nu, a_j^\nu) \to \frac{x_{j+1} - x_j}{-h_k} - f(x_j, a_j) \text{ and } x_j^\nu - u_j^\nu \to x_j - u_j.$$

This allows us to arrive at the inclusions

$$\frac{x_{j+1} - x_j}{-h_k} - f(x_j, a_j) \in \limsup_{x-u \to x_j - u_j} N_C(x - u) = N_C(x_j - u_j),$$

ensuring that $x_{j+1} - x_j \in F(x_j, u_j, a_j)$ for all $j = 0, \ldots, k-1$ and thus completing the proof. $\square$

The next key result makes a bridge between optimal solutions to the control problems (P) and (P_k) by showing on one hand that the optimal solutions $\bar{z}^k(\cdot)$ to (P_k) are *approximately optimal/suboptimal* solutions to (P) and that necessary optimality conditions for $\bar{z}^k$ can be treated as "almost optimality conditions" for $\bar{z}(\cdot)$. On the other hand, the necessary optimality conditions for the discrete solutions $\bar{z}^k(\cdot)$ obtained below will serve as the basis to derive the exact necessary optimality conditions for $\bar{z}(\cdot)$ by passing to the limit in their relationships as $k \to \infty$.

Theorem 5.2 (*Strong $W^{1,2}$ convergence of discrete optimal solutions*). *Suppose that all the standing assumptions* (H1)–(H4) *and those of Theorem 3.1 hold along the given r.i.l.m. $\bar{z}(\cdot)$ of problem (P). Assume in addition that the terminal cost φ is continuous at $\bar{x}(T)$, that the running cost ℓ in (5.1) is continuous at $(t, \bar{z}(t), \dot{\bar{z}}(t))$ for a.e. $t \in [0, T]$, and that $\ell(\cdot, z, \dot{z})$ is uniformly majorized around $\bar{z}(\cdot)$ by a summable function on $[0, T]$. Then any sequence of optimal solutions $\bar{z}^k(\cdot) = (\bar{x}^k(\cdot), \bar{u}^k(\cdot), \bar{a}^k(\cdot))$ of (P_k), piecewise linearly extended to the whole interval $[0, T]$, converges to $\bar{z}(\cdot)$ in the strong topology of $W^{1,2}([0, T]; \mathbb{R}^{2n+d})$ with the validity of the estimates*

$$\left\| \frac{\bar{u}_1^k - \bar{u}_0^k}{h_k} \right\| \le \widetilde{\mu} \quad \text{and} \quad \limsup_{k \to \infty} \sum_{j=0}^{k-2} \left\| \frac{\bar{u}_{j+2}^k - 2\bar{u}_{j+1}^k + \bar{u}_j^k}{h_k} \right\| \le \widetilde{\mu}, \tag{5.9}$$

where the number $\widetilde{\mu} > 0$ is taken from Theorem 3.1.

Proof. Take any sequence $\{\bar{z}^k(\cdot)\}$ of optimal solutions to (P_k), the existence of which is ensured by Proposition 5.1, and then extend each $\bar{z}^k(\cdot)$ piecewise linearly to the continuous-time interval $[0, T]$. All the statements of the theorem follow from the following limiting equality:

$$\lim_{k \to \infty} \int_0^T \left(\left\| \dot{\bar{x}}(t) - \dot{\bar{x}}^k(t) \right\|^2 + \left\| \dot{\bar{u}}(t) - \dot{\bar{u}}^k(t) \right\|^2 + \left\| \dot{\bar{a}}(t) - \dot{\bar{a}}^k(t) \right\|^2 \right) dt$$

$$+ \left\| \frac{\bar{x}_1^k - \bar{x}_0^k}{h_k} - \dot{\bar{x}}(0) \right\|^2 + \text{dist}^2\left(\left\| \frac{\bar{u}_1^k - \bar{u}_0^k}{h_k} \right\|; (-\infty, \widetilde{\mu}] \right) \tag{5.10}$$

$$+ \text{dist}^2\left(\sum_{j=0}^{k-2} \left\| \frac{\bar{u}_{j+2}^k - 2\bar{u}_{j+1}^k + \bar{u}_j^k}{h_k} \right\|; (-\infty, \widetilde{\mu}] \right) = 0.$$

Arguing by contradiction, suppose that (5.10) fails and thus find a subsequence of $k \in IN$ along which the limit in (5.10) equals to some $c > 0$. The weak compactness of the unit ball in $L^2([0, T]; \mathbb{R}^{2n+d})$ yields the existence of a triple $(v(\cdot), w(\cdot), q(\cdot)) \in L^2([0, T]; \mathbb{R}^{2n+d})$ such that

$$\left(\dot{\bar{x}}^k(\cdot), \dot{\bar{u}}^k(\cdot), \dot{\bar{a}}^k(\cdot) \right) \to \left(v(\cdot), w(\cdot), q(\cdot) \right) \text{ weakly in } L^2([0, T]; \mathbb{R}^{2n+d})$$

along a subsequence of optimal velocities $\{\dot{\bar{z}}^k(\cdot)\}$. It is clear that $\dot{\widetilde{z}}(t) = (v(t), w(t), q(t))$ a.e. on $[0, T]$ for the absolutely continuous triple defined by

$$\widetilde{z}(t) := \left(x_0, \bar{u}(0), \bar{a}(0) \right) + \int_0^t (v(s), w(s), q(s)) ds \text{ for all } t \in [0, T].$$

The latter implies that $\dot{\tilde{z}}^k(\cdot) \to \dot{\tilde{z}}(\cdot) = (\dot{\tilde{x}}(\cdot), \dot{\tilde{u}}(\cdot), \dot{\tilde{a}}(\cdot))$ weakly in $L^2([0, T]; \mathbb{R}^{2n+d})$, and hence it yields $\tilde{z}(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^{2n+d})$. Similarly to the proof of Theorem 4.1 we verify that $\tilde{z}(\cdot)$ satisfies inclusion (3.4) with the constraints in (1.7) and (1.8). The rest of the proof of this theorem follows the lines in the proof of [7, Theorem 5.1], which show that $\tilde{z}(\cdot)$ is feasible to the relaxed problem (R), belongs to the selected $W^{1,2}$ neighborhood of $\bar{z}(\cdot)$, and gives a smaller value to (4.2) than $\bar{z}(\cdot)$ with $\tilde{J}[\tilde{z}] = J[\bar{z}]$. Thus the assumed failure of (5.10) contradicts the choice of $\bar{z}(\cdot)$ as a r.i.l.m. in (P), and we are done. $\quad\square$

6. Second-order subdifferential computations

Optimization problems (P) and (P_k) are intrinsically *nonsmooth* and *nonconvex*, even for smooth and/or convex terminal and running costs. The unavoidable source of nonsmoothness and nonconvexity comes from the sweeping differential inclusion (3.4) and its discrete approximations (5.3), which constitute nonconvex geometric constraints of the *graphical* type. Furthermore, the first-order normal cone (subdifferential, variational structure) of the sweeping inclusions (3.4) and (5.3) calls for appropriate *second-order* subdifferential constructions to derive and analyze optimality conditions for their solutions. In this section we recall the corresponding generalized differential constructions and present the results of their computations in terms of the initial problem data that play a significant role in what follows.

Given a set-valued mapping $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, we always assume that its graph

$$\text{gph } F := \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \,\big|\, y \in F(x)\big\}$$

is locally closed around the reference point $(\bar{x}, \bar{y}) \in \text{gph } F$ and define its *coderivative* of F at this point via the (limiting) normal cone (2.2) by

$$D^* F(\bar{x}, \bar{y})(u) := \big\{v \in \mathbb{R}^n \mid (v, -u) \in N\big((\bar{x}, \bar{y}); \text{gph } F\big)\big\} \text{ for all } u \in \mathbb{R}^m. \qquad (6.1)$$

When $F: \mathbb{R}^n \to \mathbb{R}^m$ is single-valued (then $\bar{y} = F(\bar{x})$ is omitted in the coderivative notation) and continuously differentiable ($\mathcal{C}^1$-smooth) around $\bar{x}$, we have the representation

$$D^* F(\bar{x})(u) = \big\{\nabla F(\bar{x})^* u\big\}, \quad u \in \mathbb{R}^m,$$

via the adjoint/transposed Jacobian matrix $\nabla F(\bar{x})^*$. The corresponding (first-order) *subdifferential* of an l.s.c. function $\phi: \mathbb{R}^n \to \overline{\mathbb{R}}$ at $\bar{x} \in \text{dom } \phi := \{x \in \mathbb{R}^n \mid \phi(x) < \infty\}$ can be defined geometrically

$$\partial \phi(\bar{x}) := \big\{v \in \mathbb{R}^n \,\big|\, (v, -1) \in N\big((\bar{x}, \phi(\bar{x})); \text{epi } \phi\big)\big\} \qquad (6.2)$$

via the normal cone (2.2) of its epigraph $\text{epi } \phi := \{(x, \alpha) \in \mathbb{R}^{n+1} \mid \alpha \geq \phi(x)\}$ while admitting, together with the coderivative (6.1) equivalent analytical representations and – the crucial issue – satisfy comprehensive *calculus rules* despite the nonconvexity of their values; see [25,33] with the references therein.

Now we turn to the *second-order subdifferential*/generalized Hessian of $\phi: \mathbb{R}^n \to \overline{\mathbb{R}}$ at $\bar{x} \in \text{dom } \phi$ relative to $\bar{v} \in \partial \phi(\bar{x})$, which plays an underlying role in this paper and is defined, following the "dual derivative-of-derivative" scheme [23], as the coderivative of the first-order subdifferential of ϕ by

$$\partial^2 \phi(\bar{x}, \bar{v})(u) := (D^* \partial \phi)(\bar{x}, \bar{v})(u), \quad u \in \mathbb{R}^n. \tag{6.3}$$

If ϕ is C^2-smooth around $\bar{x}$, then (6.3) reduces to the (symmetric) Hessian matrix $\partial^2 \phi(\bar{x})(u) = \{\nabla^2 \phi(\bar{x}) u\}$ for all $u \in \mathbb{R}^n$, while in general it is a positively homogeneous set-valued mapping of u satisfying well-developed *second-order subdifferential calculus*; see [26,27].

To present next the second-order computations needed in what follows, we recall the definition of calmness, which is a weak "one-point" stability property that has been well-understood by now in variational analysis and optimization; see [33], [18] and the references therein. A set-valued mapping $F \colon \mathbb{R}^s \rightrightarrows \mathbb{R}^q$ is *calm* at $(\bar{\vartheta}, \bar{q}) \in \text{gph}\, F$ if there are positive numbers μ and η such that

$$F(\vartheta) \cap (\bar{q} + \eta I\!B) \subset F(\bar{\vartheta}) + \mu \|\vartheta - \bar{\vartheta}\| I\!B \text{ whenever } \vartheta \in \bar{\vartheta} + \eta I\!B,$$

where $I\!B$ stands for the unit closed ball of the space in question. Due to the normal cone description of the sweeping process, in this paper we employ the second-order subdifferential (6.3) just for the set indicator function $\delta_\Omega(x)$ of $\Omega \subset \mathbb{R}^n$ that is equal to 0 if $x \in \Omega$ and to ∞ if $x \notin \Omega$. The second-order subdifferential of the indicator function clearly reduces to $D^* N_\Omega$, where we use the notation $N_\Omega(x) := N(x; \Omega)$ for convenience. In this case the following upper estimates and exact computations of $D^* N_\Omega$ can be deduced from [18] and the previous developments mentioned therein. We use the standard notation

$$\mathbb{R}^m_- := \big\{ (y_1, \ldots, y_m) \in \mathbb{R}^m \,\big|\, y_i \leq 0 \text{ for all } i = 1, \ldots, m \big\}.$$

Proposition 6.1 (*Coderivative of the normal cone mapping to inequality constraints*). *Consider the set* $\Omega := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0\}$ *defined by the* C^2*-smooth functions* $g = (g_1, \ldots, g_m) \colon \mathbb{R}^n \to \mathbb{R}^m$ *around* $\bar{x} \in \Omega$ *so that the vectors* $\nabla g_1(\bar{x}), \ldots \nabla g_m(\bar{x})$ *are positively linearly independent, which amounts to saying that the Mangasarian–Fromovitz constraint qualification (MFCQ) is satisfied at* $\bar{x}$*. Given a normal* $\bar{v} \in N_\Omega(\bar{x})$*, suppose in addition that the multifunction* $M \colon \mathbb{R}^{2m} \rightrightarrows \mathbb{R}^{n+m}$ *defined by*

$$M(\vartheta) := \big\{ (x, \lambda) \,\big|\, \big(-g(x), \lambda\big) + \vartheta \in \text{gph}\, N_{\mathbb{R}^m_-} \big\} \tag{6.4}$$

is calm at $(0, \bar{x}, \bar{\lambda})$ *for all* $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m) \geq 0$ *satisfying the equation* $-\nabla g(\bar{x})^* \bar{\lambda} = \bar{v}$*. Then we have the second-order upper estimate*

$$D^* N_\Omega(\bar{x}, \bar{v})(u)$$

$$\subset \bigcup_{\bar{\lambda} \geq 0, -\nabla g(\bar{x})\bar{\lambda} = \bar{v}} \left\{ \left(-\sum_{i=1}^m \bar{\lambda}_i \nabla^2 g_i(\bar{x}) \right) u - \nabla g(\bar{x})^* D^* N_{\mathbb{R}^m_-}\big(-g(\bar{x}), \bar{\lambda}\big)\big(-\nabla g(\bar{x})u\big) \right\}.$$

Strengthening the calmness assumption by the full rank of the Jacobian $\nabla g(\bar{x})$ *gives us the precise formula*

$$D^* N_\Omega(\bar{x}, \bar{v})(u) = \left(-\sum_{i=1}^m \bar{\lambda}_i \nabla^2 g_i(\bar{x}) \right) u - \nabla g(\bar{x})^* D^* N_{\mathbb{R}^m_-}\big(-g(\bar{x}), \bar{\lambda}\big)\big(-\nabla g(\bar{x})u\big),$$

where $\bar{\lambda} \geq 0$ is a unique solution to the equation $-\nabla g(\bar{x})^*\bar{\lambda} = \bar{v}$. Furthermore, the coderivative of the normal cone mapping generated by the nonpositive orthant $\mathbb{R}^m_-$ above is computed by

$$D^* N_{\mathbb{R}^m_-}(x, v)(y) = \begin{cases} \emptyset & \text{if } \exists i \text{ with } v_i y_i \neq 0, \\ \{\gamma \mid \gamma_i = 0 \, \forall i \in I_1(y), \; \gamma_i \geq 0 \, \forall i \in I_2(y)\} & \text{otherwise} \end{cases} \quad (6.5)$$

whenever $(x, v) \in \text{gph } N_{\mathbb{R}^m_-}$ with the index subsets in (6.5) defined by

$$I_1(y) := \{i \mid x_i < 0\} \cup \{i \mid v_i = 0, \; y_i < 0\}, \quad I_2(y) := \{i \mid x_i = 0, \; v_i = 0, \; y_i > 0\}. \quad (6.6)$$

Proof. Compare [18, Theorem 3.3], its proof, and the further references therein. □

The next result presents the crucial second-order computations of the coderivative of the sweeping process under consideration entirely in terms of its given data.

Theorem 6.2 (*Second-order computations for the sweeping process*). *Consider the set-valued mapping F associated with the sweeping process* (1.3) *by* (3.1), *where the nonconvex set C is taken from* (1.5), *and where the perturbation mapping that f is C^1-smooth. Given $x, u \in \mathbb{R}^n$ with $x - u \in C$ as well as $w \in N_C(x - u)$ and $a \in \mathbb{R}^d$, suppose that the vectors $\nabla g_1(x - u), \ldots, \nabla g_m(x - u)$ are positively linearly independent and that the multifunction M from* (6.4) *is calm at $(0, x - u, \lambda)$ for all $\lambda = (\lambda_1, \ldots, \lambda_m) \geq 0$ satisfying the equation $-\nabla g(x - u)^*\lambda = w - f(x, a)$. Then we have the upper estimate*

$$D^* F(x, u, a, w)(y)$$
$$\subset \bigcup_{\lambda \geq 0, -\nabla g(x-u)\lambda = w - f(x,a)} \left\{ \left(\nabla_x f(x, a)^* y - \left(\sum_{i=1}^m \lambda \nabla^2 g_i(x - u) \right) y - \nabla g(x - u)^*\gamma, \right.\right.$$
$$\left.\left. \left(\sum_{i=1}^m \lambda \nabla^2 g_i(x - u) \right) y + \nabla g(x - u)^*\gamma, \nabla_a f(x, a)^* y \right) \right\}$$

for all $y \in \text{dom } D^* N_C(x - u, w - f(x, a))$, \quad (6.7)

where the coderivative domain is satisfied the inclusion

$$\text{dom } D^* N_C(x - u, w - f(x, a)) \subset \{y \mid \exists \lambda \geq 0 \text{ such that } -\nabla g(x - u)\lambda = w - f(x, a),$$
$$\lambda_i \langle \nabla g_i(x - u), y \rangle = 0 \text{ for } i = 1, \ldots, m\}, \quad (6.8)$$

and where we have in (6.7) *that $\gamma_i = 0$ if either $g_i(x - u) > 0$ or $\lambda_i = 0$ and $\langle \nabla g_i(x - u), y \rangle > 0$, and that $\gamma_i \geq 0$ if $g_i(x - u) = 0$, $\lambda_i = 0$, and $\langle \nabla g_i(x - u), y \rangle < 0$.*

Furthermore, replacing the calmness of (6.4) *by the stronger assumptions on the full rank of the Jacobian matrix $\nabla g(x - u)$ (which is actually the classical LICQ – linear independence constraint qualification) ensures that the equalities hold in both inclusions* (6.7) *and* (6.7) *with the collection of nonnegative multipliers $\lambda = (\lambda_1, \ldots, \lambda_m) \geq 0$ uniquely determined by the equation $-\nabla g(x - u)^*\lambda = w - f(x, a)$.*

Proof. Consider the mappings $G(x, u, a) := N_C(x - u)$ and $\widetilde{f}(x, u, a) := f(x, a)$. It follows from the coderivative sum rule in [25, Theorem 1.62] that

$$z^* \in \nabla \widetilde{f}(x, u, a)^* y + D^* G\big(x, u, a, w - f(x, a)\big)(y)$$

for any $y \in \operatorname{dom} D^* N_C(x - u, w - f(x, a))$ and $z^* \in D^* F(x, u, a, w)(y)$. Observe further that

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

where the Jacobian of latter mapping is obviously of full rank. It follows from the coderivative chain rule of [25, Theorem 1.66] applied to the above composition that

$$z^* \in \nabla \widetilde{f}(x, u, a)^* y + \nabla \widetilde{g}(x, u, a)^* D^* N_C\big(x - u, w - f(x, a)\big)(y). \tag{6.9}$$

Substituting now into (6.9) the corresponding results of Proposition 6.1 and taking into account the structure of the mapping $\widetilde{f}$ in (6.9) give us all the statements claimed in the theorem. □

Note that for the linear case of $g_i(x) = -\langle x_i^*, x \rangle$, corresponding to the polyhedral sweeping process, the upper estimate in (6.7) reduces to our previous computations in [7, Theorem 6.1].

7. Necessary optimality conditions for discrete-time problems

This section is devoted to deriving necessary optimality conditions for local optimal solution to the discrete-time control problems (P_k), for each fixed $k \in I\!N$. First we establish, under minimal assumptions, necessary optimality conditions in the extended Euler–Lagrange form for a general class of problems (P_k) with an arbitrary discrete velocity map F via its coderivative by reducing such problems to nonsmooth mathematical programming with many geometric constraints of the graphical type. Then we exploit the special normal cone structure (3.1) of F to obtain optimality conditions for the discrete sweeping control problems (P_k) defined in Section 5 expressed entirely via the given data by using the second-order computations of Section 6. Due to the approximation results of Section 5, the optimality conditions for (P_k) obtained in this way can be treated as *suboptimality* conditions for the given r.i.l.m. of the original sweeping control problem (P), while our main goal in what follows is to derive necessary optimality conditions for such local minimizers of (P) by passing to the limit from those in discrete approximations.

Theorem 7.1 (*Necessary conditions of the Euler–Lagrange type for discrete-time optimal control*). *Let $\bar{z}^k = (x_0, \bar{x}_1^k, \ldots, \bar{x}_k^k, \bar{u}_0^k, \ldots, \bar{u}_k^k, \bar{a}_0^k, \ldots, \bar{a}_k^k)$ be a local optimal solution to problem (P_k) for whenever $k \in I\!N$, where F is an arbitrary closed-graph mapping, and where φ and $\ell_t := \ell(t, \cdot, \cdot)$ are locally Lipschitzian around the optimal points for any $t \in \Delta_k$. Then there exist dual elements $\alpha^k = (\alpha_1^k, \ldots, \alpha_m^k) \in \mathbb{R}_+^m$, $\xi^{1k} = (\xi_0^{1k}, \ldots, \xi_k^{1k}) \in \mathbb{R}_+^{k+1}$, $\xi^{2k} = (\xi_0^{2k}, \ldots, \xi_k^{2k}) \in \mathbb{R}_-^{k+1}$, and $p_j^k = (p_j^{xk}, p_j^{uk}, p_j^{ak}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$ for $j = 0, \ldots, k$ satisfying the conditions*

$$\lambda^k + \|\alpha^k\| + \|\beta^k\| + \|\xi^{1k} + \xi^{2k}\| + \sum_{j=0}^{k-1} \|p_j^{xk}\| + \|p_0^{uk}\| + \|p_0^{ak}\| \neq 0, \tag{7.1}$$

$$\alpha_i^k g_i(\bar{x}_k^k - \bar{u}_k^k) = 0, \quad i = 1, \ldots, m, \tag{7.2}$$

$$\xi_j^{1k}(\|u_j^k\| - r_2 - \varepsilon_k) = 0 \; for \; j = 0, \ldots, k, \tag{7.3}$$

$$\xi_j^{2k}(\|u_j^k\| - r_1 + \varepsilon_k) = 0 \; for \; j = 0, \ldots, k, \tag{7.4}$$

$$\begin{cases} -p_k^{xk} \in \lambda^k \partial\varphi(\bar{x}_k^k) - \sum_{i=1}^m \alpha_i^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k), \\ \\ p_k^{uk} = -\sum_{i=1}^m \alpha_i^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) - 2(\xi_k^{1k} + \xi_k^{2k})\bar{u}_k^k, \; p_k^{ak} = 0, \end{cases} \tag{7.5}$$

$$p_{j+1}^{uk} = \lambda^k(v_j^{uk} + h_k^{-1}\theta_j^{uk}), \; p_{j+1}^{ak} = \lambda^k(v_j^{ak} + h_k^{-1}\theta_j^{ak}), \; j = 0, \ldots, k-1, \tag{7.6}$$

$$\left(\frac{p_{j+1}^{xk} - p_j^{xk}}{h_k} - \lambda^k w_j^{xk}, \frac{p_{j+1}^{uk} - p_j^{uk}}{h_k} - \lambda^k w_j^{uk}, \frac{p_{j+1}^{ak} - p_j^{ak}}{h_k} - \lambda^k w_j^{ak}, \right.$$

$$\left. p_{j+1}^{xk} - \lambda^k(v_j^{xk} + h_k^{-1}\theta_j^{xk}) \right)$$

$$\in \left(0, \frac{2}{h_k}(\xi_j^{1k} + \xi_j^{2k})\bar{u}_j^k, 0, 0 \right) + N\left(\left(\bar{x}_j^k, \bar{u}_j^k, \bar{a}_j^k, \frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{-h_k} \right); \mathrm{gph}\, F \right) \tag{7.7}$$

for $j = 0, \ldots, k-1$ with the triples

$$(\theta_j^{xk}, \theta_j^{uk}, \theta_j^{ak}) := 2 \int_{t_j}^{t_{j+1}} \left(\frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{h_k} - \dot{\bar{x}}(t), \frac{\bar{u}_{j+1}^k - \bar{u}_j^k}{h_k} - \dot{\bar{u}}(t), \frac{\bar{a}_{j+1}^k - \bar{a}_j^k}{h_k} - \dot{\bar{a}}(t) \right) dt \tag{7.8}$$

and the subgradient collections

$$(w_j^{xk}, w_j^{uk}, w_j^{uk}, v_j^{xk}, v_j^{uk}, v_j^{ak}) \in \partial\ell_t\left(\bar{z}_j^k, \frac{\bar{z}_{j+1}^k - \bar{z}_j^k}{h_k} \right), \; j = 0, \ldots, k-1, \tag{7.9}$$

where the sequence $\{\varepsilon_k\} \downarrow 0$ as $k \to \infty$ is taken from Theorem 3.1.

Proof. For simplicity we drop the upper index "k" in the notation below and consider the "long" vector y reflecting the collection of feasible solutions to each discrete-time problem (P_k):

$$y = (x_0, \ldots, x_k, u_0, \ldots, u_k, a_0, \ldots, a_k, X_0, \ldots, X_{k-1}, U_0, \ldots, U_{k-1}, A_0, \ldots, A_{k-1}).$$

We now reduce (P_k) to the following equivalent nondynamic problem of mathematical programming (MP) with respect to the variable vector y, where the starting point x_0 is fixed:

$$\text{minimize } \varphi_0[y] := \varphi(x_k) + h_k \sum_{j=0}^{k-1} \ell(x_j, u_j, a_j, X_j, U_j, A_j)$$

$$+ \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \left\| (X_j, U_j, A_j) - \dot{\bar{z}}(t) \right\|^2 dt$$

$$+ \left\| \frac{x_1^k - x_0^k}{h_k} - \dot{\bar{x}}(0) \right\|^2 + \text{dist}^2 \left(\left\| \frac{u_1^k - u_0^k}{h_k} \right\| ; (-\infty, \widetilde{\mu}] \right)$$

$$+ \text{dist}^2 \left(\sum_{j=0}^{k-2} \left\| U_{j+1} - U_j \right\| ; (-\infty, \widetilde{\mu}] \right)$$

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

$$b_j^x(y) := x_{j+1} - x_j - h_k X_j = 0 \text{ for } j = 0, \dots, k-1,$$

$$b_j^u(y) := u_{j+1} - u_j - h_k U_j = 0 \text{ for } j = 0, \dots, k-1,$$

$$b_j^a(y) := a_{j+1} - a_j - h_k A_j = 0 \text{ for } j = 0, \dots, k-1,$$

$$c_i(y) := -g_i(x_k - u_k) \le 0 \text{ for } i = 1, \dots, m,$$

$$d_j^1(y) := \|u_j\|^2 - (r_2 + \varepsilon_k)^2 \le 0 \text{ for } j = 0, \dots, k,$$

$$d_j^2(y) := \|u_j\|^2 - (r_1 - \varepsilon_k)^2 \ge 0 \text{ for } j = 0, \dots, k,$$

$$\phi_j(y) := \left\| (x_j, u_j, a_j) - \bar{z}(t_j) \right\| - \varepsilon/2 \le 0 \text{ for } j = 0, \dots, k,$$

$$\phi_{k+1}(y) := \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \left(\left\| (X_j, U_j, A_j) - \dot{\bar{z}}(t) \right\|^2 \right) dt - \frac{\varepsilon}{2} \le 0,$$

$$\phi_{k+2}(y) := \sum_{j=0}^{k-2} \left\| U_{j+1} - U_j \right\| \le \widetilde{\mu} + 1, \quad \phi_{k+3}(y) := \|u_1 - u_0\| \le (\widetilde{\mu} + 1)(t_1^k - t_0^k),$$

$$y \in \Xi_j := \left\{ y \mid -X_j \in F(x_j, u_j, a_j) \right\} \text{ for } j = 0, \dots, k-1,$$

$$y \in \Xi_k := \left\{ y \mid x_0 \text{ is fixed, } (u_0, a_0) = (\bar{u}(0), \bar{a}(0)) \right\}.$$

Let us apply the necessary optimality conditions from [26, Theorem 5.24] to any local optimal solution $\bar{y}$ of the finite-dimensional problem (MP) written above with taking into account that by Theorem 5.2 all the inequality constraints in (MP) relating to functions ϕ_j as $j = 0, \dots, k+2$ are *inactive* for large $k \in \mathbb{N}$, and hence the corresponding multipliers do not appear in the optimality conditions. In this way we find dual elements $\lambda \ge 0$, $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}_+^m$, $\xi^{1k} = (\xi_0^{1k}, \dots, \xi_k^{1k}) \in \mathbb{R}_+^{k+1}$, $\xi^{2k} = (\xi_0^{2k}, \dots, \xi_k^{2k}) \in \mathbb{R}_-^{k+1}$, $p_j = (p_j^x, p_j^u, p_j^a) \in \mathbb{R}^{2n+d}$ as $j = 0, \dots, k$, and

$$y_j^* = (x_{0j}^*, \dots, x_{kj}^*, u_{0j}^*, \dots, u_{kj}^*, a_{0j}^*, \dots, a_{kj}^*, X_{0j}^*, \dots, X_{(k-1)j}^*, U_{0j}^*, \dots, U_{(k-1)j}^*,$$
$$A_{0j}^*, \dots, A_{(k-1)j}^*),$$

$j = 0, \ldots, k$, which are not zero simultaneously, satisfy the conditions in (7.5) and the inclusions

$$y_j^* \in N(\bar{y}; \Xi_j) \text{ for } j = 0, \ldots, k, \tag{7.10}$$

$$-y_0^* - \ldots - y_k^* \in \lambda \partial \varphi_0(\bar{y}) + \sum_{i=1}^{m} \alpha_i \nabla c_i(\bar{y}) + \sum_{j=0}^{k} \xi_j^1 \nabla d_j^1(\bar{y}) + \sum_{j=0}^{k} \xi_j^2 \nabla d_j^2(\bar{y}) + \sum_{j=0}^{k-1} \nabla b_j(\bar{y})^* p_{j+1}, \tag{7.11}$$

$$\alpha_i c_i(\bar{y}) = 0 \text{ for } i = 1, \ldots, m. \tag{7.12}$$

It is easy to see that the validity of (7.3) follows directly from the structure of the sets Ξ_j, and that the inclusions in (7.10) can be equivalently rewritten as

$$\left(x_{jj}^*, u_{jj}^*, a_{jj}^*, -X_{jj}^* \right) \in N \left(\left(\bar{x}_j^k, \bar{u}_j^k, \bar{a}_j^k, \frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{-h_k} \right); \text{gph } F \right), \; j = 0, \ldots, k, \tag{7.12}$$

while every other components of y_j^* equals to zero. We conclude similarly that the only nonzero component of y_k^* might be $(x_{0k}^*, u_{0k}^*, a_{0k}^*)$. This gives us the equality

$$-y_0^* - y_1^* - \ldots - y_k^* = \left(-x_{0k}^* - x_{00}^*, -x_{11}^*, \ldots, -x_{k-1,k-1}^*, 0, -u_{0k}^* - u_{00}^*, \ldots, -u_{k-1,k-1}^*, 0, \right.$$
$$\left. - a_{0k}^* - a_{00}^*, -a_{11}^*, \ldots, -a_{k-1,k-1}^*, 0, -X_{00}^*, \ldots, -X_{k-1,k-1}^*, 0, \ldots, 0 \right). \tag{7.13}$$

Next we calculate the sums on the right-hand side of (7.11). It follows from the constructions above that

$$\left(\sum_{i=1}^{m} \alpha_i \nabla c_i(\bar{y}) \right)_{(x_k, u_k, a_k)} = \left(-\sum_{i=1}^{m} \alpha_i \nabla g_i(x_k - u_k), \sum_{i=1}^{m} \alpha_i \nabla g_i(x_k - u_k), 0 \right),$$

$$\left(\sum_{j=0}^{k} \xi_j^1 \nabla d_j^1(\bar{y}) + \sum_{j=0}^{k} \xi_j^2 \nabla d_j^2(\bar{y}) \right)_{u_j} = (2\xi_j^1 + 2\xi_j^2) \bar{u}_j \text{ for } j = 0, \ldots, k,$$

$$\left(\sum_{j=0}^{k-1} (\nabla b_j(\bar{y}))^* p_{j+1} \right)_{(x_j, u_j, a_j)} = \begin{cases} -p_1 & \text{if } j = 0, \\ p_j - p_{j+1} & \text{if } j = 1, \ldots, k-1, \\ p_k & \text{if } j = k, \end{cases}$$

$$\left(\sum_{j=0}^{k-1} (\nabla b_j(\bar{y}))^* p_{j+1} \right)_{(X, U, A)} = -h_k p = \left(-h_k p_1^x, \ldots, -h_k p_k^x, -h_k p_1^u, \ldots, \right.$$
$$\left. -h_k p_k^u, -h_k p_1^a, \ldots, -h_k p_k^a \right),$$

where the subscripts on the left-hand side refer to the corresponding components of the vectors therein. Applying the subdifferential sum rule from [25, Theorem 2.13] yields the inclusion

$$\partial\varphi_0(\bar{y}) \subset \partial\varphi(\bar{x}_k) + h_k \sum_{j=0}^{k-1} \partial\ell_t(\bar{x}_j, \bar{u}_j, \bar{a}_j, \bar{X}_j, \bar{U}_j, \bar{A}_j) + \sum_{j=0}^{k-1} \nabla\rho_j(\bar{y}) + \partial\sigma(\bar{y})$$

with the real-valued functions $\rho_j(\cdot)$ and $\sigma(\cdot)$ given by

$$\rho_j(y) := \int_{t_j}^{t_{j+1}} \left\| (X_j, U_j, A_j) - \dot{\bar{z}}(t) \right\|^2 dt,$$

$$\sigma(y) := \operatorname{dist}^2\left(\left\| \frac{u_1^k - u_0^k}{h_k} \right\|; (-\infty, \widetilde{\mu}] \right) + \operatorname{dist}^2\left(\sum_{j=0}^{k-2} \|U_{j+1} - U_j\|; (-\infty, \widetilde{\mu}] \right).$$

Using now the differentiability of $\psi(x) := \operatorname{dist}^2(x; (-\infty, \widetilde{\mu}])$ with the gradient $\nabla\psi(x) = 0$ whenever $x \le \widetilde{\mu}$ and combining it with second condition in (5.9) tells us that $\partial\sigma(\bar{y}) = \{0\}$. Furthermore, we get

$$\nabla\rho_j(\bar{y}) = \nabla_{X_j, U_j, A_j}\rho(\bar{y}) = (\theta_j^x, \theta_j^u, \theta_j^a)$$

for nonzero components. Putting this together shows that $\lambda\partial\varphi_0(\bar{y})$ in (7.11) is represented as

$$\lambda(h_k w_0^x, h_k w_1^k, \ldots, h_k w_{k-1}^x, \vartheta^k, h_k w_0^u, \ldots, h_k w_{k-1}^u, 0, h_k w_0^a, \ldots, h_k w_{k-1}^a, 0, \theta_0^x + h_k v_0^x, \ldots,$$
$$\theta_{k-1}^x + h_k v_{k-1}^x, \theta_0^u + h_k v_0^u, \ldots, \theta_{k-1}^u + h_k v_{k-1}^u, \theta_0^a + h_k v_0^a, \ldots, \theta_{k-1}^a + h_k v_{k-1}^a)$$

with $\vartheta^k \in \partial\varphi(\bar{x}_k)$ and with the components of $(w^x, w^u, w^a, v^x, v^u, v^a)$ satisfying (7.9). Involving (7.13), we derive from (7.10) the following relationships:

$$-x_{0k}^* - x_{00}^* = \lambda h_k w_0^x - p_1^x,$$

$$-x_{jj}^* = \lambda h_k w_j^x + p_j^x - p_{j+1}^x \quad \text{for } j = 1, \ldots, k-1,$$

$$0 = \lambda\vartheta^k - \sum_{i=1}^{m} \alpha_i \nabla_x \widetilde{g}_i(x_k, u_k) + p_k^x,$$

$$-u_{0k}^* - u_{00}^* = \lambda h_k w_0^u + (2\xi_0^1 + 2\xi_0^2)\bar{u}_0 - p_1^u,$$

$$-u_{jj}^* = \lambda h_k w_j^u + 2(\xi_j^1 + \xi_j^2)\bar{u}_j + p_j^u - p_{j+1}^u \quad \text{for } j = 1, \ldots, k-1,$$

$$0 = \sum_{i=1}^{m} \alpha_i \nabla_u \widetilde{g}_i(x_k, u_k) + p_k^u + (2\xi_k^1 + 2\xi_k^2)\bar{u}_k,$$

$$-a_{0k}^* - a_{00}^* = \lambda h_k w_0^a - p_1^a,$$

$$-a_{jj}^* = \lambda h_k w_j^a + p_j^a - p_{j+1}^a \quad \text{for } j = 1, \ldots, k-1,$$

$$0 = p_k^a,$$

$$-X_{jj}^* = \lambda(h_k v_j^x + \theta_j^x) - h_k p_{j+1}^x \quad \text{for } j = 0, \ldots, k-1,$$

$$0 = \lambda(h_k v_j^u + \theta_j^u) - h_k p_{j+1}^u \ \text{ for } \ j = 0, \ldots, k-1,$$

$$0 = \lambda(h_k v_j^a + \theta_j^a) - h_k p_{j+1}^a \ \text{ for } \ j = 0, \ldots, k-1.$$

To complete the proof, we proceed similarly to the last part in the proof of [7, Theorem 7.1].   □

Employing the second-order calculations conducted in Section 6 allows us to derive from Theorem 7.1 necessary optimality conditions for each discrete control problem (P_k) expressed entirely via its given data. For simplicity of the formulation we assume that the perturbation mapping f is smooth with respect to the state variable x. The reader can proceed with the case of Lipschitz continuous mapping f by using well-developed calculus rules for our basic first-order generalized differential constructions [25].

Theorem 7.2 (*Optimality conditions for discretized sweeping control problems via their original data*). *Given an optimal control $\bar{z}^k = (\bar{x}^k, \bar{u}^k, \bar{a}^k)$ to discrete-time problem (P_k) with any fixed $k \in I\!N$ and with the sweeping velocity mapping F defined in* (3.1), *suppose that the functions g_i in* (1.5) *are of class $\mathcal{C}^2$ and the perturbation mapping $f(\cdot, a)$ is of class $\mathcal{C}^1$ around the optimal points. Then there are dual elements $(\lambda^k, \beta^k, \xi^{1k}, \xi^{2k}, p^k)$ together with vectors $\eta_j^k \in \mathbb{R}_+^m$ as $j = 0, \ldots, k$ and $\gamma_j^k \in \mathbb{R}^m$ as $j = 0, \ldots, k-1$ satisfying* (7.3) *and such that the following conditions hold:*

NONTRIVIALITY CONDITION

$$\lambda^k + \|\eta_k^k\| + \|\xi^{1k} + \xi^{2k}\| + \sum_{j=0}^{k-1} \|p_j^{xk}\| + \|p_0^{uk}\| + \|p_0^{ak}\| \neq 0; \tag{7.14}$$

PRIMAL-DUAL DYNAMIC RELATIONSHIP *for all $j = 0, \ldots, k-1$:*

$$\frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{h_k} + f(\bar{x}_j^k, \bar{a}_j^k) = \sum_{i \in I(\bar{x}_j^k - \bar{u}_j^k)} \eta_{ji}^k \nabla g_i(\bar{x}_j^k - \bar{u}_j^k), \tag{7.15}$$

$$\frac{p_{j+1}^{xk} - p_j^{xk}}{h_k} - \lambda^k w_j^{xk} = \nabla_x f(\bar{x}_j^k, \bar{a}_j^k)^*(\lambda^k(v_j^{xk} + h_k^{-1}\theta_j^{xk}) - p_{j+1}^{xk})$$

$$- \sum_{i=1}^m \eta_{ji}^k \nabla^2 g_i(\bar{x}_j^k - \bar{u}_j^k)$$

$$\times (\lambda^k(v_j^{xk} + h_k^{-1}\theta_j^{xk}) - p_{j+1}^{xk})$$

$$- \sum_{i=1}^m \gamma_{ji}^k \nabla g_i(\bar{x}_j^k - \bar{u}_j^k), \tag{7.16}$$

$$\frac{p_{j+1}^{uk} - p_j^{uk}}{h_k} - \lambda^k w_j^{uk} - \frac{2}{h_k}(\xi_j^{1k} + \xi_j^{2k})\bar{u}_j^k = \sum_{i=1}^m \eta_{ji}^k \nabla^2 g_i(\bar{x}_j^k - \bar{u}_j^k)(\lambda^k(v_j^{xk} + h_k^{-1}\theta_j^{xk}) - p_{j+1}^{xk})$$

$$+ \sum_{i=1}^m \gamma_{ji}^k \nabla g_i(\bar{x}_j^k - \bar{u}_j^k), \tag{7.17}$$

$$\frac{p_{j+1}^{ak} - p_j^{ak}}{h_k} - \lambda^k w_j^{ak} = \nabla_a f(\bar{x}_j^k, \bar{a}_j^k)^* (\lambda^k (v_j^{xk} + h_k^{-1}\theta_j^{xk}) - p_{j+1}^{xk}), \quad (7.18)$$

where $(w_j^{xk}, w_j^{uk}, w_j^{ak}, v_j^{xk}, v_j^{uk}, v_j^{ak})$ are taken from (7.9) while the active constraint index set $I(\cdot)$ and the triples $\theta_j^{xk}, \theta_j^{uk}, \theta_j^{ak}$ are defined in (3.3) and (7.8), respectively;

TRANSVERSALITY CONDITIONS

$$\begin{cases} -p_k^{xk} \in \lambda^k \partial \varphi(\bar{x}_k^k) - \sum_{i=1}^{m} \eta_{ki}^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k), \\ p_k^{uk} = -\sum_{i=1}^{m} \eta_{ki}^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) - 2(\xi_k^{1k} + \xi_k^{2k})\bar{u}_k^k, \quad p_k^{ak} = 0 \end{cases} \quad (7.19)$$

with dual vectors ξ_k^{1k} and ξ_k^{2k} satisfying the inclusions

$$\xi_k^{1k} \in N_{[0, r_2+\varepsilon_k]}(\|\bar{u}_k^k\|), \quad \xi_k^{2k} \in N_{[r_1-\varepsilon_k, \infty)}(\|\bar{u}_k^k\|); \quad (7.20)$$

COMPLEMENTARITY SLACKNESS CONDITIONS

$$[g_i(\bar{x}_j^k - \bar{u}_j^k) > 0] \Longrightarrow \eta_{ji}^k = 0, \quad (7.21)$$

$$\begin{cases} [i \in I_1(-p_{j+1}^{xk} + \lambda^k(h_k^{-1}\theta_j^{xk} + v_j^{xk}))], \text{ i.e., } [g_i(\bar{x}_j^k - \bar{u}_j^k) > 0 \text{ or} \\ \eta_{ji}^k = 0, \langle \nabla g_i(\bar{x}_j^k - \bar{u}_j^k), -p_{j+1}^{xk} + \lambda^k(h_k^{-1}\theta_j^{xk} + v_j^{xk})\rangle > 0] \Longrightarrow [\gamma_{ji}^k = 0], \end{cases} \quad (7.22)$$

$$\begin{cases} [i \in I_2(-p_{j+1}^{xk} + \lambda^k(h_k^{-1}\theta_j^{xk} + v_j^{xk}))], \text{ i.e., } [g_i(\bar{x}_j^k - \bar{u}_j^k) = 0, \eta_{ji}^k = 0, \text{ and} \\ \langle \nabla g_i(\bar{x}_j^k - \bar{u}_j^k), -p_{j+1}^{xk} + \lambda^k(h_k^{-1}\theta_j^{xk} + v_j^{xk})\rangle < 0] \Longrightarrow [\gamma_{ji}^k \geq 0] \end{cases} \quad (7.23)$$

for $j = 0, \ldots, k-1$ and $i = 1, \ldots, m$, where the index subsets $I_1(\cdot)$ and $I_2(\cdot)$ are taken from (6.6), in addition to (7.3) together with the implications

$$[g_i(\bar{x}_j^k - \bar{u}_j^k) > 0] \Longrightarrow \gamma_{ji}^k = 0 \text{ for } j = 0, \ldots, k-1 \text{ and } i = 1, \ldots, m, \quad (7.24)$$

$$[g_i(\bar{x}_k^k - \bar{u}_k^k) > 0] \Longrightarrow \eta_{ki}^k = 0 \text{ for } i = 1, \ldots, m, \text{ and} \quad (7.25)$$

$$\eta_{ji}^k > 0 \Longrightarrow [\langle \nabla g_i(\bar{x}_j^k - \bar{u}_j^k), -p_{j+1}^{xk} + \lambda^k(h_k^{-1}\theta_j^{xk} + v_j^{xk})\rangle = 0]. \quad (7.26)$$

Furthermore, assuming the surjectivity of the Jacobian matrix $\{\nabla g(\bar{x}_j^k - \bar{u}_j^k)\}$ ensures the validity of the ENHANCED NONTRIVIALITY CONDITION

$$\lambda^k + \|\xi^{1k} + \xi^{2k}\| + \|p_0^{uk}\| \neq 0. \quad (7.27)$$

Proof. It follows from (7.7) and the coderivative definition (6.1) that

$$\left(\frac{p_{j+1}^{xk} - p_j^{xk}}{h_k} - \lambda^k w_j^{xk}, \frac{p_{j+1}^{uk} - p_j^{uk}}{h_k} - \lambda^k w_j^{uk} - \frac{2}{h_k}(\xi_j^{1k} + \xi_j^{2k})\bar{u}_j^k, \frac{p_{j+1}^{ak} - p_j^{ak}}{h_k} - \lambda^k w_j^{ak} \right)$$

$$\in D^* F\left(\bar{x}_j^k, \bar{u}_j^k, \bar{a}_j^k, \frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{-h_k} \right)(\lambda^k(h_k^{-1}\theta_j^{xk} + v_j^{xk}) - p_{j+1}^{xk}), \quad j = 0, \ldots, k-1. \quad (7.28)$$

By $\dfrac{\bar{x}_{j+1}^k - \bar{x}_j^k}{-h_k} - f(\bar{x}_j^k, \bar{a}_j^k) \in N(\bar{x}_j^k - \bar{u}_j^k; C)$ for $j = 0, \ldots, k-1$ and the representation of F in (3.2), we find vectors $\eta_j^k \in \mathbb{R}_+^m$, $j = 0, \ldots, k-1$, such that the conditions in (7.15) and (7.21) are satisfied. Employing the second-order upper estimate from Theorem 6.2 with $x := \bar{x}_j^k$, $u := \bar{u}_j^k$, $a := \bar{a}_j^k$, $w := \dfrac{\bar{x}_{j+1}^k - \bar{x}_j^k}{-h_k}$, and $y := \lambda^k (h_k^{-1} \theta_j^{xk} + v_j^{xk}) - p_{j+1}^{xk}$ and combining it with (6.8) give us $\gamma_j^k \in \mathbb{R}^m$ for which

$$
\left(\frac{p_{j+1}^{xk} - p_j^{xk}}{h_k} - \lambda^k w_j^{xk}, \frac{p_{j+1}^{uk} - p_j^{uk}}{h_k} - \lambda^k w_j^{uk} - \frac{2}{h_k}(\xi_j^{1k} + \xi_k^{2k})\bar{u}_j^k, \frac{p_{j+1}^{ak} - p_j^{ak}}{h_k} - \lambda^k w_j^{ak} \right)
$$

$$
= \left(\nabla_x f(\bar{x}_j^k, \bar{a}_j^k)^* (\lambda^k (v_j^{xk} + h_k^{-1} \theta_j^{xk}) - p_{j+1}^{xk}) \right.
$$

$$
- \sum_{i=1}^m \eta_{ji}^k \nabla^2 g_i (\bar{x}_j^k - \bar{u}_j^k)(\lambda^k (v_j^{xk} + h_k^{-1} \theta_j^{xk}) - p_{j+1}^{xk})
$$

$$
- \sum_{i=1}^m \gamma_{ji}^k \nabla g_i (\bar{x}_j^k - \bar{u}_j^k), \sum_{i=1}^m \eta_{ji}^k \nabla^2 g_i (\bar{x}_j^k - \bar{u}_j^k)(\lambda^k (v_j^{xk} + h_k^{-1} \theta_j^{xk}) - p_{j+1}^{xk})
$$

$$
\left. + \sum_{i=1}^m \gamma_{ji}^k \nabla g_i (\bar{x}_j^k - \bar{u}_j^k), \nabla_a f(\bar{x}_j^k, \bar{a}_j^k)^* (\lambda^k (v_j^{xk} + h_k^{-1} \theta_j^{xk}) - p_{j+1}^{xk}) \right), \quad j = 0, \ldots, k-1.
$$

This yields the validity of all the conditions in (7.16), (7.17), (7.18), (7.22), and (7.23). Put $\eta_k^k := \alpha_k$ with α_k taken from the statement of Theorem 7.1 and observe that $\eta_j^k \in \mathbb{R}_+^m$ for $j = 0, \ldots, k$. In this way we deduce the nontriviality condition (7.14) from (7.1) and the transversality conditions (7.19) from (7.5). Furthermore, (7.25) follows immediately from (7.2) and the definition of η_k^k, while (7.28) yields

$$
\lambda^k (h_k^{-1} \theta_j^{xk} + v_j^{xk}) - p_{j+1}^{xk} \in \mathrm{dom}\, D^* N_C \left(\bar{x}_j^k - \bar{u}_j^k, \frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{-h_k} - f(\bar{x}_j^k, \bar{a}_j^k) \right).
$$

This implies by (6.8) that (7.26) holds. Inclusions (7.20) follow from (7.3) and (7.4).

To complete the proof of the theorem, it remains to justify the enhanced nontriviality condition (7.27) under the surjectivity of the Jacobians $\{ \nabla g(\bar{x}_j^k - \bar{u}_j^k) \}$. Arguing by contradiction, suppose that (7.27) fails, i.e., $\lambda^k = 0$, $\xi^{1k} + \xi^{2k} = 0$, and $p_0^{uk} = 0$. Then it follows from (7.6) that $p_j^{uk} = 0$ for $j = 0, \ldots, k$ and $p_j^{ak} = 0$ for $j = 1, \ldots, k$. Employing the second condition in (7.19) with $p_k^{uk} = 0$ tells us that $\sum_{i=1}^m \eta_{ki}^k \nabla g_i (\bar{x}_k^k - \bar{u}_k^k) = 0$, and so $p_k^{xk} = 0$ by the first condition therein. We also get from (7.17) that

$$
\sum_{i=1}^m \eta_{ji}^k \nabla^2 g_i (\bar{x}_j^k - \bar{u}_j^k)(\lambda^k (v_j^{xk} + h_k^{-1} \theta_j^{xk}) - p_{j+1}^{xk}) + \sum_{i=1}^m \gamma_{ji}^k \nabla g_i (\bar{x}_j^k - \bar{u}_j^k) = 0,
$$

$$
j = 0, \ldots, k-1.
$$

Using this together with (7.16) and $p_k^{xk} = 0$ shows that $p_j^{xk} = 0$ for all $j = 0, \ldots, k - 1$. Finally, it follows from (7.18) that $p_0^{ak} = 0$, which contradicts the validity of (7.14) and thus verifies (7.27). □

8. Necessary conditions for sweeping optimal solutions

This section is the culmination of the paper. Given an arbitrary *relaxed intermediate local minimizer* $\bar{z}(\cdot)$ of the sweeping optimal control problem (P) and using the discrete approximation method, we derive verifiable necessary optimality conditions for $\bar{z}(\cdot)$, which expressed entirely via the problem data, by combining the strong convergence of discrete optimal solutions to $\bar{z}(\cdot)$ established in Section 5 and the necessary optimality conditions in discrete approximations taken from Section 7 with a rather involved technique to justify a proper convergence of the adjoint trajectories from Theorem 7.2. The latter technique developed here is heavily based on the underlying properties of our basic generalized differential constructions and the second-order calculations of Section 6. As discussed in Section 4, *no relaxation* is needed if $\bar{z}(\cdot)$ is a *strong* local minimizer of (P) and either controls are located only in perturbations, or the set C in (1.5) is convex in addition to the standing assumptions formulated in Section 2.

We now add to the standing assumptions the one on time dependence of the basic subdifferential $\partial\ell$ of the running cost in (1.6) taken below with respect to all but t variables. It is well known that the subdifferential mapping (6.2) is *robust* (which reduces to the graph-closedness for continuous functions) with respect to the variables of subdifferentiation. We suppose that this robustness property keeps holding when the time parameter is involved into the limiting procedure. Precisely it amounts to saying that

$$\partial\ell\big(t, \bar{z}(t), \dot{\bar{z}}(t)\big) = \underset{(\tau,u,v) \overset{\ell}{\to} (t,z(t),\dot{z}(t))}{\operatorname{Lim\,sup}} \partial\ell(\tau, u, v) \text{ a.e. } t \in [0, T]$$

around the given local optimal solution to (P), where "Lim sup" stands for the Painlevé–Kuratowski outer/upper limit [33]. This assumption is not restrictive and is satisfied, in particular, for smooth functions with time-continuous derivatives as well as in broad nonsmooth settings; see [24,26].

Theorem 8.1 (*Optimality conditions for the nonconvex sweeping process*). *Given an r.i.l.m.* $\bar{z}(\cdot) = (\bar{x}(\cdot), \bar{u}(\cdot), \bar{a}(\cdot))$ *for problem* (P), *suppose in addition to the standing assumptions and those in Theorem 3.1 that ℓ is continuous in t a.e. on $[0, T]$ and is represented as*

$$\ell(t, z, \dot{z}) = \ell_1(t, z, \dot{x}) + \ell_2(t, \dot{u}) + \ell_3(t, \dot{a}) \tag{8.1}$$

where the local Lipschitz constants of $\ell_1(t, \cdot, \cdot)$ and $\ell_3(t, \cdot)$ are essentially bounded on $[0, T]$ and continuous at a.e. $t \in [0, T]$ including $t = 0$, and where ℓ_2 is differentiable in $\dot{u}$ on $\mathbb{R}^n$ with the estimates

$$\|\nabla_{\dot{u}}\ell_2(t, \dot{u}, \dot{a})\| \le L\|\dot{u}\| \text{ and } \|\nabla_{\dot{u}}\ell_2(t, \dot{u}_1) - \nabla_{\dot{u}}\ell_2(s, \dot{u}_2)\| \le L|t - s| + L\|\dot{u}_1 - \dot{u}_2\| \tag{8.2}$$

holding for all $t, s \in [0, T]$, $\dot{a} \in \mathbb{R}^d$, and $\dot{u}, \dot{u}_1, \dot{u}_2 \in \mathbb{R}^n$ with some uniform constant $L > 0$. Then there are $\lambda \ge 0$, $p(\cdot) = (p^x(\cdot), p^u(\cdot), p^a(\cdot)) \in W^{1,2}([0, T]; \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d)$, $w(\cdot) =$

$(w^x(\cdot), w^u(\cdot), w^a(\cdot)) \in L^2([0, T]; \mathbb{R}^{2n+d})$, *and* $v(\cdot) = (v^x(\cdot), v^u(\cdot), v^a(\cdot)) \in L^2([0, T]; \mathbb{R}^{2n+d})$ *satisfying*

$$\left(w(t), v(t)\right) \in \text{co } \partial\ell\left(t, \bar{z}(t), \dot{\bar{z}}(t)\right) \ a.e. \ t \in [0, T] \tag{8.3}$$

as well as measures $\gamma = (\gamma_1, \ldots, \gamma_n) \in C([0, T]; \mathbb{R}^n)^*$, $\xi^1 \in C([0, T]; \mathbb{R}_+)^*$, *and* $\xi^2 \in C([0, T]; \mathbb{R}_-)^*$ *on* $[0, T]$ *such that the following conditions hold:*

- PRIMAL-DUAL DYNAMIC RELATIONSHIPS:

$$\dot{\bar{x}}(t) + f\left(\bar{x}(t), \bar{a}(t)\right) = \sum_{i=1}^m \eta_i(t)\nabla g_i\left(\bar{x}(t) - \bar{u}(t)\right) \ for \ a.e. \ t \in [0, T] \tag{8.4}$$

with $\eta_i(\cdot) \in L^2([0, T]; \mathbb{R}^+)$ *a.e. uniquely determined by representation* (8.4) *and well defined at* $t = T$;

$$\dot{p}(t) = \lambda w(t) + \left(\nabla_x f(\bar{x}(t), \bar{a}(t))^*(\lambda v^x(t) - q^x(t)), 0, \nabla_a f(\bar{x}(t), \bar{a}(t))^*(\lambda v^x(t) - q^x(t))\right), \tag{8.5}$$

$$q^u(t) = \lambda\nabla_{\dot{u}}\ell\left(t, \dot{\bar{u}}(t)\right), \ q^a(t) \in \lambda\partial_{\dot{a}}\ell_3\left(t, \dot{\bar{a}}(t)\right) \ a.e. \ t \in [0, T], \tag{8.6}$$

where $q = (q^x, q^u, q^a) : [0, T] \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$ *is a vector function of bounded variation, and its left-continuous representative is given for all* $t \in [0, T]$, *except at most a countable subset, by*

$$q(t) := p(t) - \int\limits_{[t, T]} \left(-d\gamma(s), 2\bar{u}(s)d(\xi^1(s) + \xi^2(s)) + d\gamma(s), 0\right). \tag{8.7}$$

Furthermore, for a.e. $t \in [0, T]$ *including* $t = T$ *and for all* $i = 1, \ldots, m$ *we have*

$$g_i\left(\bar{x}(t) - \bar{u}(t)\right) > 0 \Longrightarrow \eta_i(t) = 0, \ \eta_i(t) > 0 \Longrightarrow \left\langle\nabla g_i\left(\bar{x}(t) - \bar{u}(t), \lambda v^x(t) - q^x(t)\right)\right\rangle = 0. \tag{8.8}$$

- TRANSVERSALITY CONDITIONS

$$\begin{cases} -p^x(T) + \displaystyle\sum_{i \in I(\bar{x}(T) - \bar{u}(T))} \eta_i(T)\nabla g_i\left(\bar{x}(T) - \bar{u}(T)\right) \in \lambda\partial\varphi\left(\bar{x}(T)\right), \quad p^a(T) = 0, \\ p^u(T) - \displaystyle\sum_{i \in I(\bar{x}(T) - \bar{u}(T))} \eta_i(T)\nabla g_i(\bar{x}(T) - \bar{u}(T)) \\ \quad \in -2\bar{u}(T)\left(N_{[0, r_2]}(\|\bar{u}(T)\|) + N_{[r_1, \infty)}(\|\bar{u}(T)\|)\right) \end{cases} \tag{8.9}$$

with the validity of the inclusion

$$-\sum_{i \in I(\bar{x}(T) - \bar{u}(T))} \eta_i(T)\nabla g_i\left(\bar{x}(T) - \bar{u}(T)\right) \in N_C\left(\bar{x}(T) - \bar{u}(T)\right). \tag{8.10}$$

- MEASURE NONATOMICITY CONDITIONS:

   (a) Take $t \in [0, T]$ *with* $g_i(\bar{x}(t) - \bar{u}(t)) > 0$ *whenever* $i = 1, \ldots, m$. *Then there is a neighborhood* V_t *of* t *in* $[0, T]$ *such that* $\gamma(V) = 0$ *for all the Borel subsets* V *of* V_t.

(b) *Take* $t \in [0, T]$ *with* $r_1 < \|\bar{u}(t)\| < r_2$. *Then there is a neighborhood* W_t *of* t *in* $[0, T]$ *such that* $\xi^1(W) = 0$ *and* $\xi^2(W) = 0$ *for all the Borel subsets* W *of* W_t.

• NONTRIVIALITY CONDITIONS: *We always have:*

$$\lambda + \|q^u(0)\| + \|p(T)\| + \|\xi^1\|_{TV} + \|\xi^2\|_{TV} > 0. \tag{8.11}$$

Furthermore, the following implications hold while ensuring the ENHANCED NONTRIVIALITY*:*

$$\left[g_i\big(x_0 - \bar{u}(0)\big) > 0, \ i = 1, \dots, m\right] \Longrightarrow \left[\lambda + \|p(T)\| + \|\xi^1\|_{TV} + \|\xi^2\|_{TV} > 0\right], \tag{8.12}$$

$$\left[g_i\big(\bar{x}(T) - \bar{u}(T)\big) > 0, \ r_1 < \|\bar{u}(T)\| < r_2, \ i = 1, \dots, m\right]$$
$$\Longrightarrow \left[\lambda + \|q^u(0)\| + \|\xi^1\|_{TV} + \|\xi^2\|_{TV} > 0\right], \tag{8.13}$$

where $\|\xi\|_{TV}$ *stands for the measure* ξ *total variation on* $[0, T]$.

Proof. We split the proof into the following major steps in accordance to the statement of the theorem.

Step 1: *Subgradients of the running cost.* To verify the subdifferential inclusion (8.3), take the subgradient sequence $\{w_j^k, v_j^k\}$ from Theorem 7.2 and consider the piecewise constant extensions $w^k, v^k \colon [0, T] \to \mathbb{R}^{2n+d}$. It follows from (7.9) therein. The imposed assumptions and the structure of ℓ in (8.1) with estimates (8.2) ensure that the subgradient sets $\partial \ell(t, \cdot)$ are uniformly L^2-bounded near $\bar{z}(\cdot)$, and hence the sequence $\{(w^k(\cdot), v^k(\cdot))\}$ is weakly compact in $L^2([0, T]; \mathbb{R}^{2(2n+d)}) := L^2[0, T]$. Without relabeling we get the weak convergence $(w^k(\cdot), v^k(\cdot)) \to (w(\cdot), v(\cdot)) \in L^2[0, T]$ and thus, by Mazur's theorem, the strong L^2-convergence to $(w(\cdot), v(\cdot))$ of a sequence of convex combinations of $(w^k(\cdot), v^k(\cdot))$; that is, the a.e. convergence on $[0, T]$ to the above limiting pair of some subsequence of the latter. This readily verifies (8.3) by taking into account the assumed a.e. continuity of ℓ in t and the robustness of its subdifferential.

Step 2: *Verification of the primal dynamic limiting relationships.* We claim two of them: the differential equation (8.4) and the first implication in (8.8). We proceed by passing to the limit in (7.15), (7.26) and first construct the piecewise constant functions on $[0, T]$ by

$$\theta^k(t) := \frac{\theta_j^k}{h_k} \text{ as } t \in [t_j^k, t_{j+1}^k), \ j = 0, \dots, k-1, \quad k \in I\!N,$$

where θ_j^k are taken from (7.8). It follows from Theorem 7.2 that

$$\int_0^T \|\theta^{xk}(t)\|^2 dt = \sum_{j=0}^{k-1} \frac{\|\theta_j^{xk}\|^2}{h_k} \leq \frac{4}{h_k} \sum_{j=0}^{k-1} \left(\int_{t_j^k}^{t_{j+1}^k} \left\| \dot{\bar{x}}(t) - \frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{h_k} \right\| dt \right)^2$$

$$\leq 4 \sum_{j=0}^{k-1} \left(\int_{t_j^k}^{t_{j+1}^k} \left\| \dot{\bar{x}}(t) - \frac{\bar{x}_{j+1}^k - \bar{x}_j^k}{h_k} \right\|^2 \right) dt = 4 \int_0^T \left\| \dot{\bar{x}}(t) - \dot{x}^k(t) \right\|^2 dt \to 0 \tag{8.14}$$

as $k \to \infty$ with the same conclusion for $\theta^{uk}(\cdot)$ and $\theta^{ak}(\cdot)$. Thus some subsequences of these functions converge to zero a.e. on $[0, T]$. Invoking further the vectors $\eta_j^k \in \mathbb{R}_+^m$ from Theorem 7.2, define the piecewise constant functions $\eta^k(\cdot)$ on $[0, T]$ by $\eta^k(t) := \eta_j^k$ as $t \in [t_j^k, t_{j+1}^k)$ with $\eta^k(T) := \eta_k^k$ and deduce from (7.15) for each $k \in I\!N$ we have the relationships

$$\dot{\bar{x}}^k(t) + f\big(\bar{x}^k(t), \bar{a}^k(t)\big) = \sum_{i=1}^m \eta_i^k(t) \nabla g_i\big(\bar{x}^k(t) - \bar{u}^k(t)\big) \text{ if } t \in (t_j^k, t_{j+1}^k). \tag{8.15}$$

The feasibility of $\bar{z}(\cdot)$ in (P) tells us that $-\dot{\bar{x}}(t) \in N_C(\bar{x}(t) - \bar{u}(t)) + f(\bar{x}(t), \bar{a}(t))$ for a.e. $t \in [0, T]$, where the closed-valued normal cone mapping $N_C(\cdot)$ is measurable by [33, Theorem 14.26]. The classical measurable selection result (see, e.g., [33, Corollary 14.6]) gives us nonnegative measurable functions $\eta_i(\cdot)$ on $[0, T]$ as $i = 1, \ldots, m$ for which the differential equation (8.4) and the first implication in (8.8) are satisfied. Then invoking (8.15) and (8.4) yields the equalities

$$\dot{\bar{x}}(t) - \dot{\bar{x}}^k(t) + f\big(\bar{x}(t), \bar{a}(t)\big) - f\big(\bar{x}^k(t), \bar{a}^k(t)\big)$$

$$= \sum_{i=1}^m \Big[\eta_i(t) \nabla g_i\big(\bar{x}(t) - \bar{u}(t)\big) - \eta_i^k(t) \nabla g_i\big(\bar{x}^k(t) - \bar{u}^k(t)\big) \Big]$$

whenever $t \in (t_j^k, t_{j+1}^k)$ and $j = 0, \ldots, k-1$, which imply the estimate

$$\Big\| \sum_{i=1}^m \Big[\eta_i(t) \nabla g_i\big(\bar{x}(t) - \bar{u}(t)\big) - \eta_i^k(t) \nabla g_i\big(\bar{x}^k(t) - \bar{u}^k(t)\big) \Big] \Big\|$$

$$\leq \big\| \dot{\bar{x}}(t) - \dot{\bar{x}}^k(t) \big\| + \big\| f\big(\bar{x}(t), \bar{a}(t)\big) - f\big(\bar{x}^k(t), \bar{a}^k(t)\big) \big\|$$

on (t_j^k, t_{j+1}^k). Passing to the limit as $k \to \infty$ in this estimate with replacing t_j^k by

$$\nu^k(t) := \max\big\{ t_j^k \,\big|\, t_j^k \leq t, \, 0 \leq j \leq k \big\}, \quad t \in [0, T], \tag{8.16}$$

and taking into account that $I(\bar{x}^k(\cdot) - \bar{u}^k(\cdot)) \subset I(\bar{x}(\cdot) - \bar{u}(\cdot))$ for $k \in I\!N$ sufficiently large and that the sequence $\{\bar{z}^k(\cdot)\}$ converges to $\bar{x}(\cdot)$ strongly in $W^{1,2}([0, T])$, we get

$$\sum_{i \in I(\bar{x}(t) - \bar{u}(t))} \Big[\eta_i(t) \nabla g_i\big(\bar{x}(t) - \bar{u}(t)\big) - \eta_i^k(t) \nabla g_i\big(\bar{x}^k(t) - \bar{u}^k(t)\big) \Big] \to 0 \text{ a.e. } t \in [0, T].$$

Observe also that $\eta^k(\cdot) \to \eta(\cdot)$ a.e. on $[0, T]$ by (2.5) and (2.6). Postponing till Step 5 the verification of the claim that the sequence $\{\eta_k^k\}$ converges to the well-defined vector $(\eta_1(T), \ldots, \eta_m(T))$, we check now that $\eta(\cdot) \in L^2([0, T]; \mathbb{R}_+^m)$. In fact, it follows from the estimates

$$\eta_i(t) \leq \frac{1}{M_1}\eta_i(t)\big\|\nabla g_i\big(\bar{x}(t) - \bar{u}(t)\big)\big\| \leq \frac{1}{M_1}\sum_{i \in I(\bar{x}(t)-\bar{u}(t))}\eta_i(t)\big\|\nabla g_i\big(\bar{x}(t) - \bar{u}(t)\big)\big\|$$

$$\leq \frac{\beta}{M_1}\Big\|\sum_{i \in I(\bar{x}(t)-\bar{u}(t))}\eta_i(t)\nabla g_i\big(\bar{x}(t) - \bar{u}(t)\big)\Big\|$$

$$\leq \frac{\beta}{M_1}\|\dot{\bar{x}}(t)\| + \frac{\beta}{M_1}\big\|f\big(\bar{x}(t), \bar{a}(t)\big)\big\|$$

valid for a.e. $t \in [0, T]$ and all $i = 1, \ldots, m$, which are consequences of (8.4), (2.5), and (2.6). The uniqueness of $\eta(t)$ a.e. on $[0, T]$ follows from the positive linear independence of the gradients $\nabla g_i(x)$ on C, which is a consequence of the standing assumptions in Section 2.

Step 3: *Constructions of approximating dual elements on* $[0, T]$. The next step is to extend the discrete dual elements from Theorem 7.2 on the continuous-time interval $[0, T]$ in the way appropriate for the subsequent limiting procedure. Define $q^k(t) = (q^{xk}(t), q^{uk}(t), q^{ak}(t))$ on $[0, T]$ as the piecewise linear extensions of $q^k(t_j^k) := p_j^k$ when $j = 0, \ldots, k$. Then construct $\gamma^k(t)$ on $[0, T]$ by

$$\gamma^k(t) := \gamma_j^k \ \text{ for } \ t \in [t_j^k, t_{j+1}^k), \ j = 0, \ldots, k-1, \ \text{ with } \ \gamma^k(t_k^k) := 0$$

and define further $\xi^{1k}(t) := \dfrac{\xi_j^{1k}}{h_k}$, $\xi^{2k}(t) := \dfrac{\xi_j^{2k}}{h_k}$ for $t \in [t_j^k, t_{j+1}^k)$ and $j = 0, \ldots, k-1$ with $\xi^{1k}(t_k^k) := \xi_k^{2k}$ and $\xi^{1k}(t_k^k) := \xi_k^{2k}$. It follows from the relationships in (7.16)–(7.18) with $\nu^k(t)$ given in (8.16) that

$$\dot{q}^{xk}(t) - \lambda^k w^{xk}(t)$$
$$= \nabla_x f\big(\bar{x}^k(\nu^k(t)), \bar{a}^k(\nu^k(t))\big)^*\big(\lambda^k(\nu^{xk}(t) + \theta^{xk}(t)) - q^{xk}\big(\nu^k(t) + h_k\big)\big)$$
$$- \sum_{i=1}^m \eta_i^k(t)\nabla^2 g_i\big(\bar{x}^k(\nu^k(t)) - \bar{u}^k(\nu^k(t))\big)\big(\lambda^k(\nu^{xk}(t) + \theta^{xk}(t)) - q^{xk}\big(\nu^k(t) + h_k\big)\big)$$
$$- \sum_{i=1}^m \gamma_i^k(t)\nabla g_i\big(\bar{x}^k(\nu^k(t)) - \bar{u}^k(\nu^k(t))\big),$$

$$\dot{q}^{uk}(t) - \lambda^k w^{uk}(t)$$
$$= \sum_{i=1}^m \eta_i^k(t)\nabla^2 g_i\big(\bar{x}^k(\nu^k(t)) - \bar{u}^k(\nu^k(t))\big)\big(\lambda^k(\nu^{xk}(t) + \theta^{xk}(t)) - q^{xk}\big(\nu^k(t) + h_k\big)\big)$$
$$+ \sum_{i=1}^m \gamma_i^k(t)\nabla g_i\big(\bar{x}^k(\nu^k(t)) - \bar{u}^k(\nu^k(t))\big) + 2\big(\xi^{1k}(t) + \xi^{2k}(t)\big)\bar{u}^k(\nu^k(t)),$$

$$\dot{q}^{ak}(t) - \lambda^k w^{ak}(t) = \nabla_a f\big(\bar{x}^k(\nu^k(t)), \bar{a}^k(\nu^k(t))\big)^*\big(\lambda^k(\nu^{xk}(t) + \theta^{xk}(t)) - q^{xk}\big(\nu^k(t) + h_k\big)\big)$$

for $t \in (t_j^k, t_{j+1}^k)$ and $j = 0, \ldots, k-1$. The next triple is $p^k(t) = (p^{xk}(t), p^{uk}(t), p^{ak}(t))$ defined by

$$p^k(t) := q^k(t) + \int_{[t,T]} \left(-\sum_{i=1}^{m} \eta_i^k(t) \nabla^2 g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big)\big(\lambda^k(v^{xk}(t) + \theta^{xk}(t)\big) \right.$$

$$- q^{xk}\big(v^k(t) + h_k)\big) - \sum_{i=1}^{m} \gamma_i^k(s)\nabla g_i\big(\bar{x}^k(v^k(s)) - \bar{u}^k(v^k(s))\big),$$

$$\sum_{i=1}^{m} \eta_i^k(t)\nabla^2 g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big)\big(\lambda^k(v^{xk}(t) + \theta^{xk}(t)) - q^{xk}(v^k(t) + h_k)\big)$$

$$\left. + \sum_{i=1}^{m} \gamma_i^k(t)\nabla g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big) + 2\big(\xi^{1k}(t) + \xi^{2k}(t)\big)\bar{u}^k\big(v^k(t)\big), 0 \right) ds$$

for all $t \in [0, T]$. We clearly have $p^k(T) = q^k(T)$ together with the differential condition

$$\dot{p}^k(t) = \dot{q}^k(t) - \left(-\sum_{i=1}^{m} \eta_i^k(t)\nabla^2 g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big)\big(\lambda^k(v^{xk}(t) + \theta^{xk}(t)) \right.$$

$$- q^{xk}\big(v^k(t) + h_k)\big) - \sum_{i=1}^{m} \gamma_i^k(t)\nabla g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big),$$

$$\sum_{i=1}^{m} \eta_i^k(t)\nabla^2 g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big)\big(\lambda^k(v^{xk}(t) + \theta^{xk}(t)) - q^{xk}\big(v^k(t) + h_k)\big)$$

$$\left. + \sum_{i=1}^{m} \gamma_i^k(t)\nabla g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big) + 2\big(\xi^{1k}(t) + \xi^{2k}(t)\big)\bar{u}^k\big(v^k(t)\big), 0 \right)$$

valid for a.e. $t \in [0, T]$. It follows from the relationships above that

$$\dot{p}^{xk}(t) - \lambda^k w^{xk}(t) = \nabla_x f\big(\bar{x}^k(v^k(t)), \bar{a}^k(v^k(t))\big)^* \big(\lambda^k(v^{xk}(t) + \theta^{xk}(t)) - q^{xk}\big(v^k(t) + h_k)\big), \tag{8.17}$$

$$\dot{p}^{uk}(t) - \lambda^k w^{uk}(t) = 0, \tag{8.18}$$

$$\dot{p}^{ak}(t) - \lambda^k w^{ak}(t) = \nabla_a f\big(\bar{x}^k(v^k(t)), \bar{a}^k(v^k(t))\big)^* \big(\lambda^k(v^{xk}(t) + \theta^{xk}(t)) - q^{xk}\big(v^k(t) + h_k)\big), \tag{8.19}$$

for every $t \in (t_j^k, t_{j+1}^k)$, $j = 0, \ldots, k-1$. Now we get the measures γ^k, ξ^{1k}, and ξ^{2k} on $[0, T]$ given by

$$\begin{cases} \displaystyle\int_A d\gamma^k := \int_A \left(\sum_{i=1}^{m} \eta_i^k(t)\nabla^2 g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big)\big(\lambda^k(v^{xk}(t) + \theta^{xk}(t)) \right. \\ \left. - q^{xk}\big(v^k(t) + h_k)\big)\right)dt + \int_A \left(\sum_{i=1}^{m} \gamma_i^k(t)\nabla g_i\big(\bar{x}^k(v^k(t)) - \bar{u}^k(v^k(t))\big)\right)dt, \qquad (8.20) \\ \displaystyle\int_A d\xi^{1k}(t)dt := \int_A \xi^{1k}(t)dt, \ \int_A d\xi^{2k}(t)dt := \int_A \xi^{2k}(t)dt \end{cases}$$

for any Borel subset $A \subset [0, T]$. Finally in the step, we employ the standard normalization procedure to equivalently rewrite the nontriviality condition (7.27) in the form

$$\lambda^k + \left\| p^k(T) \right\| + \left\| q^{uk}(0) \right\| + \int_0^T \left| \xi^{1k}(t) \right| dt + \int_0^T \left| \xi^{2k}(t) \right| dt + |\xi_k^{1k}| + |\xi_k^{2k}| = 1, \ \ k \in IN. \ (8.21)$$

Step 4: *Verifying the dual dynamic conditions.* By (8.21) we get $\lambda^k \to \lambda \geq 0$ along a subsequence. To verity next that $\{q^{uk}(\cdot)\}$ is of uniformly bounded variations on $[0, T]$, observe by (7.6) that

$$\sum_{j=0}^{k-1} \| q^{uk}(t_{j+1}) - q^{uk}(t_j) \|$$

$$= \sum_{j=0}^{k-1} \| p_{j+1}^{uk} - p_j^{uk} \| \leq \sum_{j=1}^{k-1} \| p_{j+1}^{uk} - p_j^{uk} \| + \| p_1^{uk} \| + \| p_0^{uk} \|$$

$$\leq \lambda^k \sum_{j=1}^{k-1} \left\| \frac{\theta_j^{uk} - \theta_{j-1}^{uk}}{h_k} \right\| + \lambda^k \frac{\| \theta_0^{uk} \|}{h_k} + \lambda^k \sum_{j=1}^{k-1} \| v_j^{uk} - v_{j-1}^{uk} \| + \lambda^k \| v_0^{uk} \| + \| p_0^{uk} \|$$

$$\leq 2\lambda^k \sum_{j=1}^{k-1} \left\| \frac{\bar{u}_{j+1}^k - 2\bar{u}_j^k + \bar{u}_{j-1}^k}{h_k} \right\| + 2\lambda^k \sum_{j=1}^{k-1} \left\| \frac{\bar{u}^k(t_{j+1}^k) - 2\bar{u}^k(t_j^k) + \bar{u}^k(t_{j-1}^k)}{h_k} \right\|$$

$$+ \lambda^k \frac{\| \theta_0^{uk} \|}{h_k} + \lambda^k \sum_{j=1}^{k-1} \| v_j^{uk} - v_{j-1}^{uk} \| + \lambda^k \| v_0^{uk} \| + \| p_0^{uk} \|$$

$$\leq 4\widetilde{\mu}\lambda^k + \lambda^k \frac{\| \theta_0^{uk} \|}{h_k} + \lambda^k \sum_{j=1}^{k-1} \| v_j^{uk} - v_{j-1}^{uk} \| + \lambda^k \| v_0^{uk} \| + \| p_0^{uk} \|$$

with $\widetilde{\mu}$ taken from Theorem 3.1. The differentiability of ℓ_2 in (8.1) with respect to $\dot{u}$ yields

$$v_j^{uk} = \nabla_{\dot{u}} \ell_2 \left(t_j, \frac{\bar{u}_{j+1}^k - \bar{u}_j^k}{h_k}, \frac{\bar{a}_{j+1}^k - \bar{a}_j^k}{h_k} \right) \text{ for } j = 0, \ldots, k-1.$$

Then the third estimate above ensures that

$$\lambda^k \sum_{j=1}^{k-1} \| v_j^{uk} - v_{j-1}^{uk} \| \leq \lambda^k \sum_{j-1}^{k-1} L \left(t_{j+1}^k - t_j^k + \left\| \frac{\bar{u}_{j+1}^k - 2\bar{u}_j^k + \bar{u}_{j-1}^k}{h_k} \right\| \right) \leq \lambda^k L(T + \widetilde{\mu}),$$

which in turn yields by the construction of θ^{uk} in (7.8) the relationships

$$\frac{\theta_0^{uk}}{h_k} = \frac{2(\bar{u}_1^k - \bar{u}_0^k)}{h_k} - \frac{2(\bar{u}(h_k) - \bar{u}(0))}{h_k} \text{ and } \lambda^k \frac{\| \theta_0^{uk} \|}{h_k} \leq 2\lambda^k(\mu + \widetilde{\mu})$$

due to the first estimate in (5.9). Furthermore

$$\|v_0^{uk}\| = \left\| \nabla_{\ddot{u}} \ell_2 \left(0, \frac{\bar{u}_1^k - \bar{u}_0^k}{h_k}, \frac{\bar{a}_1^k - \bar{a}_0^k}{h_k} \right) \right\| \leq L \|\frac{\bar{u}_1^k - \bar{u}_0^k}{h_k}\| \leq L\widetilde{\mu}$$

due to the first estimate in (8.2). This tells us that

$$
\begin{aligned}
\sum_{j=0}^{k-1} \|q^{uk}(t_{j+1}) - q^{uk}(t_j)\| &= \sum_{j=0}^{k-1} \|p_{j+1}^{uk} - p_j^{uk}\| \\
&\leq 4\widetilde{\mu}\lambda^k + 2\lambda^k(\mu + \widetilde{\mu}) + \lambda^k L(T + \widetilde{\mu}) + \lambda^k L\widetilde{\mu} + \|p_0^{uk}\| \\
&= \lambda^k(2\mu + 6\widetilde{\mu} + LT + L\widetilde{\mu}) + \|p_0^{uk}\|
\end{aligned}
\tag{8.22}
$$

and verifies therefore the uniform bounded variations of the sequence $\{q^{uk}(\cdot)\}$ on $[0, T]$.

Our next goal is to prove the boundedness of $\{(p_0^{xk}, \ldots, p_k^{xk})\}$. It follows from (7.16) and (7.17) that

$$
\begin{aligned}
p_{j+1}^{xk} - p_j^{xk} &= h_k \lambda^k (w_j^{xk} + w_j^{uk}) + 2(\xi_j^{1k} + \xi_j^{2k})\bar{u}_j^k \\
&\quad + h_k \nabla_x f(\bar{x}_j^k, \bar{a}_j^k)^*(\lambda^k(v_j^{xk} + h_k^{-1}\theta_j^{xk}) - p_{j+1}^{xk}) - (p_{j+1}^{uk} - p_j^{uk}), \quad
\end{aligned}
\tag{8.23}
$$

and readily implies the estimates

$$
\begin{aligned}
\|p_j^{xk}\| &\leq (1 + h_k \|\nabla_x f(\bar{x}_j^k, \bar{a}_j^k)\|)\|p_{j+1}^{xk}\| + h_k \lambda^k (\|w_j^{xk}\| + \|w_j^{uk}\|) + 2r_2(\xi_j^{1k} + \xi_j^{2k}) \\
&\quad + \|\nabla_x f(\bar{x}_j^k, \bar{a}_j^k)\| \cdot \|h_k v_j^{xk}\| + h_k \|\nabla_x f(\bar{x}_j^k, \bar{a}_j^k)\|\lambda^k \|\theta^{xk}(t_j)\| + \|p_{j+1}^{uk} - p_j^{uk}\|
\end{aligned}
\tag{8.24}
$$

valid for all $j = 0, \ldots, k - 1$. Denoting further

$$
\begin{aligned}
A_j^k &:= h_k \lambda^k \left(\|w_j^{xk}\| + \|w_j^{uk}\| \right) + 2r_2(\xi_j^{1k} + \xi_j^{2k}) + \|\nabla_x f(\bar{x}_j^k, \bar{a}_j^k)\| \cdot \|h_k v_j^{xk}\| \\
&\quad + h_k \|\nabla_x f(\bar{x}_j^k, \bar{a}_j^k)\|\lambda^k \|\theta^{xk}(t_j)\| + \|p_{j+1}^{uk} - p_j^{uk}\|
\end{aligned}
$$

and selecting a $\widetilde{M}_1$ such that $\|\nabla_x f(\bar{x}_j^k, \bar{a}_j^k)\| \leq \widetilde{M}_1$ for all $j = 0, \ldots, k - 1$ and $k \in I\!N$, we have by (8.21)

$$
\begin{aligned}
\sum_{j=0}^{k-1} h_k \left\| \nabla_x f(\bar{x}_j^k, \bar{a}_j^k) \right\| \lambda^k \left\| \theta^{xk}(t_j) \right\| &\leq \lambda^k \widetilde{M}_1 \sum_{j=0}^{k-1} \sqrt{h_k \int_{t_j}^{t_{j+1}} \left\| \theta^{xk}(t) \right\|^2 dt} \\
&\leq \lambda^k \widetilde{M}_1 \sqrt{\int_0^T \left\| \theta^{xk}(t) \right\|^2 dt} \downarrow 0 \text{ as } k \to \infty.
\end{aligned}
$$

The structure (8.1) of the running cost in (1.6) and the imposed assumptions on its Lipschitz constant $L(t)$ yield by (8.3) the relationships

$$\sum_{j=0}^{k-1} \left\| h_k w_j^{xk} \right\| = \sum_{j=0}^{k-1} h_k \left\| w^{xk}(t_j) \right\| \le \sum_{j=0}^{k-1} L(t_j) h_k \le 2 \int_0^T L(t) dt := \widetilde{L} < \infty,$$

which ensure furthermore that $\sum_{j=0}^{k-1} \left\| h_k w_j^{uk} \right\| \le \widetilde{L}$ and $\sum_{j=0}^{k-1} \left\| h_k v_j^{xk} \right\| \le \widetilde{L}$. We also have

$$\sum_{i=0}^{k-1} |\xi_j^{1k} + \xi_j^{2k}| = \int_0^T |\xi^{1k}(t) + \xi^{2k}(t)| dt \le 1.$$

As follows from the above arguments, the boundedness of $\{\|\nabla_x f(\bar{x}_j^k, \bar{a}_j^k)\|\}$ with the usage of (8.22) that

$$\sum_{j=0}^{k-1} A_j^k \le \widetilde{M}_2 \tag{8.25}$$

for some constant $\widetilde{M}_2 > 0$. Combining it with (8.24) gives us the estimates

$$\|p_j^{xk}\| \le (1 + \widetilde{M}_1 h_k) \|p_{j+1}^{xk}\| + A_j^k \quad \text{for all} \quad j = 0, \ldots, k-1.$$

Proceeding now by induction shows that

$$\|p_j^{xk}\| \le (1 + \widetilde{M}_1 h_k)^{k-j} \|p_k^{xk}\| + \sum_{i=j}^{k-1} A_i^k (1 + \widetilde{M}_1 h_k)^{i-j}$$

$$\le e^{\widetilde{M}_1} + e^{\widetilde{M}_1} \sum_{i=0}^{k-1} A_i^k \le e^{\widetilde{M}_1} (1 + \widetilde{M}_2)$$

for $j = 0, \ldots, k-1$, which justifies the boundedness of $\{(p_0^{xk}, \ldots, p_k^{xk})\}$.

Next we show that the functional sequences $\{q^{xk}(\cdot)\}$ and $\{q^{ak}(\cdot)\}$ are of uniform bounded variations on $[0, T]$. It follows from (8.23) that

$$\sum_{j=0}^{k-1} \left\| q^{xk}(t_{j+1}) - q^{xk}(t_j) \right\| = \sum_{j=0}^{k-1} \left\| p_{j+1}^{xk} - p_j^{xk} \right\| \le \sum_{j=0}^{k-1} A_j^k + \sum_{j=0}^{k-1} h_k \left\| \nabla_x f(\bar{x}_j^k, \bar{a}_j^k) \right\| \left\| p_{j+1}^{xk} \right\|,$$

which verifies the claimed property for $\{q^{xk}(\cdot)\}$ due to (8.25) and the boundedness of $\{(p_0^{xk}, \ldots, p_k^{xk})\}$. Furthermore, (7.18) leads us to the estimate

$$\sum_{j=0}^{k-1} \left\| q^{ak}(t_{j+1}) - q^{ak}(t_j) \right\|$$

$$\leq h_k \sum_{j=0}^{k-1} \left\| \lambda^k w_j^{ak} \right\| + h_k \sum_{j=0}^{k-1} \left\| \nabla_a f(\bar{x}_j^k, \bar{a}_j^k) \right\| \left(\lambda^k \left(\left\| v_j^{xk} \right\| + \left\| \theta^{xk}(t_j^k) \right\| \right) + \left\| p_{j+1}^{xk} \right\| \right),$$

which ensures the uniform bounded variations of $\{q^{ak}(\cdot)\}$ and hence of the whole sequence of triples $\{q^k(\cdot)\}$. We clearly get the validity of

$$2 \left\| q^k(t) \right\| - \left\| q^k(0) \right\| - \left\| q^k(T) \right\| \leq \left\| q^k(t) - q^k(0) \right\| + \left\| q^k(T) - q^k(t) \right\| \leq \mathrm{var}\left(q^k; [0, T] \right)$$

whenever $t \in [0, T]$. Hence $\{q^k(\cdot)\}$ is bounded on $[0, T]$ due to the boundedness of $\{q^k(0)\}$ and $\{q^k(T)\}$. The classical Helly theorem provides a function of bounded variation $q(\cdot)$ such that $q^k(t) \to q(t)$ for all $t \in [0, T]$. It follows from (8.21) that the sequences $\{\xi^{1k}\}$ and $\{\xi^{2k}\}$ are bounded in $C([0, T]; \mathbb{R}_+)^*$ and $C([0, T]; \mathbb{R}_-)^*$, respectively. It allows us to get the boundedness of $\{\gamma^k\}$ in $C([0, T]; \mathbb{R}^n)^*$ from (7.17) and the uniform bounded variations of $\{q^{uk}(\cdot)\}$ on $[0, T]$. We derive from the weak* sequential compactness of balls in these spaces that there are $\gamma \in C([0, T]; \mathbb{R}^n)^*$, $\xi^1 \in C([0, T]; \mathbb{R}_+)^*$, and $\xi^2 \in C([0, T]; \mathbb{R}_-)^*$ such that the triples $(\gamma^k, \xi^{1k}, \xi^{2k})$ weak* converge to (γ, ξ^1, ξ^2) along some subsequence.

Using (8.17)–(8.19) and (8.21) together with the uniform boundedness of $q^k(\cdot)$, $w^k(\cdot)$, and $v^k(\cdot)$ on $[0, T]$ ensures the boundedness of the sequence $\{p^k(\cdot)\}$ in $W^{1,2}([0, T]; \mathbb{R}^{3n})$ and hence its weak compact in this space. By Mazur's theorem we find a function $p(\cdot) \in W^{1,2}([0, T]; \mathbb{R}^{3n})$ such that a sequence of convex combinations of $\dot{p}^k(t)$ converges to $\dot{p}(t)$ for a.e. $t \in [0, T]$. Then the passage to the limit in (8.17)–(8.19) justifies the claimed representation of $\dot{p}(\cdot)$ in (8.5).

Our next aim is to derive the optimality conditions of the theorem that involve the dual arc $q(\cdot)$ of bounded variation on $[0, T]$. Observe that the condition $\eta_i(t) > 0$ for some $t \in [0, T]$ and $i \in \{1, \ldots, m\}$ yields $\eta_i^k(t) > 0$ for large k by the a.e. convergence $\eta_i^k(\cdot) \to \eta_i(\cdot)$ on $[0, T]$. This implies by (7.26) that

$$\left\langle \nabla g_i\left(x^k(t) - u^k(t)\right), -q^{xk}\left(v(t) + h_k\right) + \lambda^k\left(\theta^{xk}(t) + v^{xk}(t)\right) \right\rangle = 0$$

for such k and t, and hence we get by passing to the limit that

$$\left\langle \nabla g_i\left(\bar{x}(t) - \bar{u}(t)\right), \lambda v^x(t) - q^x(t) \right\rangle = 0,$$

which verifies the second implication in (8.8). By the construction of $q^k(\cdot)$ we get

$$q^{uk}\left(v(t) + h_k\right) = \lambda^k\left(v^{uk}(t) + \theta^{uk}(t)\right) \text{ and } q^{ak}\left(v(t) + h_k\right) = \lambda^k\left(v^{ak}(t) + \theta^{ak}(t)\right) \qquad (8.26)$$

whenever $t \in (t_j^k, t_{j+1}^k)$ and $j = 0, \ldots, k - 1$. Involving (8.3) and the assumptions on ℓ_2, ℓ_3 in (8.1) gives us both conditions in (8.6) by passing to the limit in (8.26). Proceeding similarly to [38, p. 325] yields

$$\left\| \int\limits_{[t,T]} \left(\sum_{i=1}^{m} \eta_i^k(s)\nabla^2 g_i\left(\bar{x}^k(v^k(s)) - \bar{u}^k(v^k(s))\right)\left(\lambda^k(v^{xk}(s) + \theta^{xk}(s)) - q^{xk}\left(v^k(s) + h_k\right)\right) \right. $$

$$\left. + \sum_{i=1}^{m} \gamma_i^k(s)\nabla g_i\left(\bar{x}^k(v^k(s)) - \bar{u}^k(v^k(s))\right) \right) ds - \int\limits_{[t,T]} d\gamma(s) \right\|$$

$$= \left\| \int\limits_{[t,T]} d\gamma^k(s) - \int\limits_{[t,T]} d\gamma(s) \right\| \to 0 \text{ as } k \to \infty$$

for all $t \in [0, T]$ except a countable subset of $[0, T]$. It tells us by using (8.20) that

$$\int\limits_{[t,T]} \left(\sum_{i=1}^{m} \eta_i^k(s)\nabla^2 g_i\left(\bar{x}^k(v^k(s)) - \bar{u}^k(v^k(s))\right)\left(\lambda^k(v^{xk}(s) + \theta^{xk}(s)) - q^{xk}\left(v^k(s) + h_k\right)\right) \right.$$

$$\left. + \sum_{i=1}^{m} \gamma_i^k(s)\nabla g_i\left(\bar{x}^k(v^k(s)) - \bar{u}^k(v^k(s))\right) \right) ds \to \int\limits_{[t,T]} d\gamma(s) \text{ as } k \to \infty. \tag{8.27}$$

To derive (8.7) by passing to the limit in the differential condition for $p^k(t)$, consider the estimate

$$\left\| \int\limits_{[t,T]} \left(\xi^{1k}(s) + \xi^{2k}(s)\right)\bar{u}^k\left(v^k(s)\right)ds - \int\limits_{[t,T]} \bar{u}(s)d(\xi^1(s) + \xi^2(s)) \right\|$$

$$\leq \left\| \int\limits_{[t,T]} \left(\xi^{1k}(s) + \xi^{2k}(s)\right)\bar{u}^k\left(v^k(s)\right)ds - \int\limits_{[t,T]} \left(\xi^{1k}(s) + \xi^{2k}(s)\right)\bar{u}(s)ds \right\|$$

$$+ \left\| \int\limits_{[t,T]} \left(\xi^{1k}(s) + \xi^{2k}(s)\right)\bar{u}(s)ds - \int\limits_{[t,T]} \bar{u}(s)d\left(\xi^1(s) + \xi^2(s)\right) \right\| \tag{8.28}$$

$$= \left\| \int\limits_{[t,T]} \left(\xi^{1k}(s) + \xi^{2k}(s)\right)\left[\bar{u}^k\left(v^k(s)\right) - \bar{u}(s)\right]ds \right\|$$

$$+ \left\| \int\limits_{[t,T]} \left(\xi^{1k}(s) + \xi^{2k}(s)\right)\bar{u}(s)ds - \int\limits_{[t,T]} \bar{u}(s)d\left(\xi^1(s) + \xi^2(s)\right) \right\|.$$

Note that the first summand after the equality sign in (8.28) vanishes as $k \to \infty$ due to the uniform convergence $\bar{u}^k(\cdot) \to \bar{u}(\cdot)$ on $[0, T]$ and the uniform boundedness of $\int\limits_0^T |\xi^{1k}(t)|dt +$

$\int_0^T |\xi^{2k}(t)|dt$ by (8.21). The second summand there also converges to zero for all $t \in [0, T]$ except some countable subset by the weak* convergence of $\xi^{1k} \to \xi^1$ in $C([0, T]; \mathbb{R}_+)^*$ and $\xi^{2k} \to \xi^2$ in $C([0, T]; \mathbb{R}_-)^*$. Thus we get

$$\int_{[t,T]} (\xi^{1k}(s) + \xi^{2k}(s))\bar{u}^k(\tau^k(s))ds \to \int_{[t,T]} \bar{u}(s)d(\xi^1(s) + \xi^2(s)) \text{ as } k \to \infty$$

and hence obtain (8.7) by passing to the limit in the above differential condition for $p^k(t)$.

Step 5: *Verifying transversality.* Observe first that the sequence $\{\eta_{ki}^k\}$ admits a convergent subsequence. To show it, we deduce from the second discrete transversality condition in (7.19) that

$$\sum_{i \in I(\bar{x}(T) - \bar{u}(T))} \eta_{ki}^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) = \sum_{i \in I(\bar{x}_k^k - \bar{u}_k^k)} \eta_{ki}^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) = -p_k^{uk} - 2(\xi_k^{1k} + \xi_k^{2k})\bar{u}_k^k$$

with $\eta_{ki}^k = 0$ for $i \in \{1, \ldots, m\}\setminus I(\bar{x}_k^k - \bar{u}_k^k)$. This justifies by (8.21) the boundedness of the sequence $\left\{ \sum_{i \in I(\bar{x}(T) - \bar{u}(T))} \eta_{ki}^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) \right\}$. It follows from the standing assumptions in (2.5) and (2.6) that

$$\eta_{ki}^k \leq \frac{1}{M_1} \sum_{j \in I(\bar{x}(T) - \bar{u}(T))} \eta_{kj}^k \left\| \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) \right\| \leq \frac{\beta}{M_1} \left\| \sum_{i \in I(\bar{x}(T) - \bar{u}(T))} \eta_{ki}^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) \right\|,$$

which ensures the boundedness of $\{\eta_{ki}^k\}$ for $i = 1, \ldots, m$, and thus we get $\eta_{ki}^k \to \tilde{\eta}_i$ along a subsequence of $k \to \infty$. Denote $\eta(T) := (\tilde{\eta}_1, \ldots, \tilde{\eta}_n)$ and observe that it is well defined due to the positive linear independence of vectors $\nabla g_i(\bar{x}(T) - \bar{u}(T))$, which follows from the standing assumptions. Then

$$\vartheta_k := \sum_{i \in I(\bar{x}(T) - \bar{u}(T))} \eta_{ki}^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) \to \vartheta := \sum_{i \in I(\bar{x}(T) - \bar{u}(T))} \eta_i(T) \nabla g_i(\bar{x}(T) - \bar{u}(T))$$

as $k \to \infty$, where the defined vector ϑ satisfies the inclusion in (8.10). Further, it follows from (7.20) and the second equation in (7.19) that

$$p_k^{uk} + \vartheta_k = -2(\xi_k^{1k} + \xi_k^{2k}) \in -2\bar{u}_k^k \left(N_{[0, r_2 + \varepsilon_k]}(\|\bar{u}_k^k\|) + N_{[r_1 - \varepsilon_k, \infty)}(\|\bar{u}_k^k\|) \right). \tag{8.29}$$

Passing now to the limit in (8.29) and (7.19) with taking into account the subdifferential robustness as well as the convergence of $\{\xi_k^{1k}\}$ and $\{\xi_k^{2k}\}$ (7.20), we justify the transversality conditions (8.9).

Step 6: *Verifying measure nonatomicity.* We provide the verification of the nonatomicity condition (a) while observing that the case of (b) is similar. To proceed, pick $t \in [0, T]$ with $g_i(\bar{x}(t) - \bar{u}(t)) > 0$ whenever $i = 1, \ldots, m$ and employing the continuity of g_i and $(\bar{x}(\cdot), \bar{u}(\cdot))$

get a neighborhood V_t of t such that $g_i(\bar{x}(s) - \bar{u}(s)) > 0$ for all $s \in V_t$ and $i = 1, \ldots, m$. The obtained convergence of the discrete optimal solutions tells us that $g_i(\bar{x}^k(t_j^k) - \bar{u}^k(t_j^k)) > 0$ if $t_j^k \in V_t$ for $i = 1, \ldots, m$ and all large k. It follows from (7.21) and (7.24) that $\eta_{ji}^k = 0$ and $\gamma_{ji}^k = 0$ for $i = 1, \ldots, m$. Thus

$$\|\gamma^k\|_{TV} = \int_V d\left\|\gamma^k\right\| = \int_V \left\|\gamma^k(t)\right\| dt = 0$$

by (8.20). Passing to the limit as $k \to \infty$ with taking into account the measure convergence established in Step 3, we get $\|\gamma\|(V) = 0$, which verifies the claimed measure nonatomicity.

Step 7: *Verifying nontriviality.* To start with the verification of the general nontriviality condition (8.11), suppose on the contrary that $\lambda = 0$, $q^u(0) = 0$, $p(T) = 0$, $\|\xi^1\|_{TV} = 0$, and $\|\xi^2\|_{TV} = 0$. Then

$$\lambda^k \to 0, \ q^{uk}(0) \to 0, \ p^k(T) \to 0, \ \int_{[0,T]} |\xi^{1k}(t)| dt \to 0, \ \int_{[0,T]} |\xi^{2k}(t)| dt \to 0$$

as $k \to \infty$. Let us check that in this case we get $\xi_k^{1k} + \xi_k^{2k} \to 0$. Indeed, observe that the convergence $p^k(T) \to 0$, $\lambda^k \to 0$ implies by the first condition in (7.19) that $p_k^{xk} \to 0$, $p_k^{uk} \to 0$, and $\sum_{i=1}^m \eta_{ki}^k \nabla g_i(\bar{x}_k^k - \bar{u}_k^k) \to 0$. Then the second condition in (7.19) yields $(\xi_k^{1k} + \xi_k^{2k})\bar{u}_k^k \to 0$, and $\xi_k^{1k} + \xi_k^{2k} \to 0$ due to $\bar{u}_k^k \neq 0$ for all large k. This clearly contradicts (8.21), and so we are done with (8.11).

It remains to verify the enhanced nontriviality conditions of the theorem under the additional assumptions made. To proceed with (8.12), suppose that $g_i(x_0 - \bar{u}(0)) > 0$ for $i = 1, \ldots, m$ and, arguing by contradiction, that $\lambda = 0$, $p(T) = 0$, $\|\xi^1\|_{TV} = 0$, and $\|\xi^2\|_{TV} = 0$. Then $g_i(\bar{x}_0^k - \bar{u}_0^k) > 0$ for large k. It follows from (7.21) and (7.22) that $\eta_{ji}^k = 0$ and $\gamma_{ji}^k = 0$ for $i = 1, \ldots, m$. Unifying it with (7.17) and the construction of $q^{uk}(\cdot)$ in Step 3 shows that

$$q^{uk}(0) = p_0^{uk} = p_1^{uk} + 2(\xi_0^{1k} + \xi_0^{2k})\bar{u}_0^k + h_k \lambda^k w_0^{uk}$$

whenever $k \in I\!N$ is sufficiently large. This yields the estimates

$$\|q^{uk}(0)\| \leq \lambda^k \|v_0^{uk}\| + \lambda^k \frac{\|\theta_0^{uk}\|}{h_k} + \lambda^k h_k \left\|w_0^{uk}\right\| + 2r_2 \left|\xi_0^{1k} + \xi_0^{2k}\right|$$

$$\leq \lambda^k L\tilde{\mu} + 2\lambda^k(\mu + \tilde{\mu}) + \lambda^k \tilde{\mu} + \int_{[0,T]} |\xi^{1k}(t)| dt + \int_{[0,T]} |\xi^{2k}(t)| dt,$$

which imply in turn that $q^u(0) = \lim_{k \to \infty} q^{uk}(0) = 0$ while contradicting the nontriviality condition (8.11). The verification of the other enhanced nontriviality condition (8.13) is similar. $\square$

-250   0
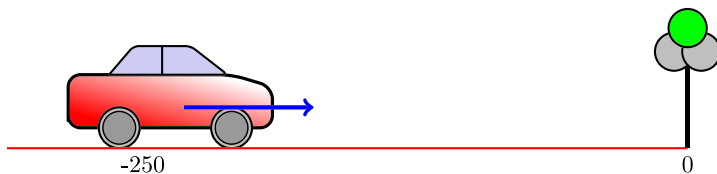
Fig. 1. Direction of optimal control. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

9. Numerical examples

In this section we present two examples, which are related to real-life models while illustrating some special features and applications of the obtained necessary optimality conditions for the controlled sweeping process under consideration. In both examples, the obtained optimality conditions allow us to determine optimal solutions and explicitly calculate their parameters.

The first example addresses a one-dimensional sweeping control model of type (P).

Example 9.1 *(Optimal control of car motion)*. Consider a car moving towards the traffic light with the initial speed $s = 9$ m/s (about 20 mi/h). When the car is 250 meters away from the traffic light, the light turns into green color and it lasts for 30 seconds. We need to control the motion of the car in such a way that after 20 seconds it must be as close to the traffic light as possible and the energy used to adjust the speed must be minimized as well; see Fig. 1.

To solve this problem numerically, let us specify the initial data in problem (P) as follows:

$$\begin{cases} n = m = d = 1, \ T = 20, \ x_0 := -250, \ g_1(x) := -x, \ f(x,a) := sa = 9a, \\ r_1 := 10^{-3}, \ r_2 := 50, \ \varphi := \dfrac{x^2}{2}, \ \text{and } \ell(t,x,u,a,\dot{x},\dot{u},\dot{a}) := \tfrac{1}{2}a^2 \end{cases} \tag{9.1}$$

with $C = \mathbb{R}_-$ in this case. For definiteness, suppose that the traffic light is located at the origin. It follows from (9.1) that the set of a-controls can be assumed to be uniformly bounded. Thus (P) admits an optimal solution $(\bar{x}(\cdot), \bar{u}(\cdot), \bar{a}(\cdot)) \in W^{1,2}([0,20]; \mathbb{R}^3)$ by Theorem 4.1. It is also easy to check that all the assumptions of Theorem 8.1 are satisfied. Supposing now that $\bar{x}(t) \in \text{int}(C + \bar{u}(t))$ for any $t \in [0, 20)$ and that $\bar{x}(20) - \bar{u}(20) \in \text{bd}\,C$, we see that these assumptions are realized for the optimal solution found via the necessary optimality conditions of Theorem 8.1. Since we expect the car to be as close to the traffic light as possible after 20 seconds with $\bar{x}(20) = \bar{u}(20)$, the value $\bar{u}(20)$ should be small. The choice of r_1 and r_2 in (9.1) ensures that the validity of the constraints $r_1 \le |\bar{u}(t)| \le r_2$ for all $t \in [0, 20]$.

Applying the necessary optimality conditions of Theorem 8.1 gives us the following relationships with a number $\lambda \ge 0$ and a function $\eta(\cdot) \in L^2([0,20]; \mathbb{R}_+)$ well defined at $t = 20$:

1. $w(t) = (0, 0, \bar{a}(t)), \ v(t) = (0, 0, 0)$ a.e. $t \in [0, 20]$;
2. $\bar{x}(t) < \bar{u}(t) \Longrightarrow \eta(t) = 0$ a.e. $t \in [0, 20]$;
3. $\eta(t) > 0 \Longrightarrow q^x(t) = 0$ a.e. $t \in [0, 20]$ including $t = 20$;
4. $\dot{\bar{x}}(t) + 9\bar{a}(t) = -\eta(t)$ a.e. $t \in [0, 20]$;
5. $(\dot{p}^x(t), \dot{p}^u(t), \dot{p}^a(t)) = (0, 0, \lambda\bar{a}(t) - 9q^x(t))$ a.e. $t \in [0, 20]$;
6. $q^u(t) = 0, \ q^a(t) = 0$ a.e. $t \in [0, 20]$;

7. $(q^x(t), q^u(t), q^a(t)) = (p^x(t), p^u(t), p^a(t)) - \int\limits_{[t,20]} (-d\gamma, 2\bar{u}(s)d(\xi^1(s) + \xi^2(s)), 0)$ a.e. $t \in$ [0, 20];

8. $-p^x(20) - \eta(20) = \lambda\bar{x}(20);$

9. $p^u(20) + \eta(20) \in -2\bar{u}(20)\left(N_{[0,1]}(|\bar{u}(20)|) + N_{[10^{-3},\infty)}(|\bar{u}(20)|)\right);$

10. $-\eta(20) \in N_C(\bar{x}(20) - \bar{u}(20));$

11. $\lambda + |q^u(0)| + |p(20)| + \|\xi^1\|_{TV} + \|\xi^2\|_{TV} > 0.$

We get from (5)–(7) that $p^x(\cdot)$ is a constant function on [0, 20] and that

$$\lambda\bar{a}(t) = 9q^x(t) = 9p^x(20) + 9\int\limits_{[t,20]} d\gamma.$$

Proceeding similarly to [8, Example 1] gives us the relationship

$$\lambda\bar{a}(t) = 9p^x(20) + 9\gamma(\{20\}) \quad \text{a.e.} \quad [0, 20]. \tag{9.2}$$

If $\lambda > 0$, we can deduce from (9.2) that $\bar{a}(\cdot)$ must be a constant function, $\bar{a}(\cdot) \equiv \vartheta$, on [0, 20] due to its continuity on this interval. In the case $\lambda = 0$, we may assume that we control the speed of the car at the initial time as $\dot{\bar{x}}(0) = -10\vartheta$ and maintain this speed till the end of the process, i.e., $\dot{\bar{x}}(t) = -10\vartheta$ for a.e. $t \in [0, 20]$. This results in $\bar{a}(\cdot) \equiv \vartheta$ on [0, 20]. Then (2) and (4) yield

$$\bar{x}(t) = x_0 + \int\limits_0^t \dot{\bar{x}}(s)ds = -250 - 9\vartheta t \quad \text{for all } t \in [0, 20].$$

Consequently, the cost functional in our problem (P) is computed as

$$J[\bar{x}, \bar{u}, \bar{a}] = \frac{(-250 - 180\vartheta)^2}{2} + \frac{20\vartheta^2}{2}$$

and clearly achieves its absolute minimum at $\vartheta = -\frac{45000}{32420} \approx -1.388$. Thus in this case we arrive by the necessary optimality conditions of Theorem 8.1 at the (local) optimal solution written as

$$\bar{x}(t) = -250 + 12.492t, \quad \bar{a}(t) = -1.388 \quad \text{on } [0, 20]$$

with $\bar{u}(\cdot)$ being an absolutely continuous function on [0, 20] such that $\bar{u}(20) = \bar{x}(20) = -0.16$. This tells us that at the moment when the car is 250 meters away from the traffic light its speed should be switched to $1.388 \times 9 = 12.492$ m/s, and we should maintain this speed till the end of the process. After 20 seconds, the car is just 0.16 meter away from the traffic light. As seen, the value 12.492 m/s is very close to $12.5 = \frac{250}{20}$. To save more energy needed to get such speed, it makes sense to adjust the running cost as follows: $\ell(t, x, u, a, \dot{x}, \dot{u}, \dot{a}) := \frac{100}{2}a^2$. In this case the cost functional is represented by

$$J[\bar{x}, \bar{u}, \bar{a}] = \frac{(-250 - 180\vartheta)^2}{2} + \frac{2000\vartheta^2}{2}$$

and achieves its absolute minimum at $\vartheta = -\frac{45000}{34400} \approx -1.308$ by using the optimal solution

$$\bar{x}(t) = -250 + 11.772t, \quad \bar{a}(t) = -1.308,$$

where the optimal control $\bar{u}(\cdot)$ is absolutely continuous on $[0, 20]$ and such that $\bar{u}(20) = \bar{x}(20) = -14.56$. Hence we should switch the speed of the car to 11.772 m/s when it is 250 meters away from the traffic light and maintain this new constant speed till the end of the process. In this way the car is about 14.56 meters away from the traffic light after 20 seconds.

The next example concerns a rather particular case of the two-dimensional crowd motion model that is formalized as a nonconvex controlled sweeping process of type (P). A more general controlled crowd motion model on the plane is the subject of our adjacent paper [9]. Note that the given example demonstrates the usefulness for calculating optimal solutions of the necessary optimality conditions from Theorem 8.1 with the general nontriviality condition (8.11).

Example 9.2 (*Case for optimal control of the planar crowd motion model*). We refer the reader to [21,37] for describing an uncontrolled microscopic version of the crowd motion model as a sweeping process. Here we introduce (following the previous corridor version in [8]) controls entering both the moving set and perturbations. In what follows we restrict ourselves to the case of two participants identified with rigid disks of the same radius $R = 3$. The center of the i-th disk is denoted by $x_i \in \mathbb{R}^2$. To fulfill the nonoverlapping condition crucial in the crowd motion model, the vector of positions $x = (x_1, x_2) \in \mathbb{R}^4$ has to belong to the nonconvex set of feasible configurations (see Fig. 2) defined by

$$C := \left\{ x = (x_1, x_2) \in \mathbb{R}^4 \,\middle|\, g(x) := \|x_1 - x_2\| - 2R \geq 0 \right\}.$$

Suppose that the initial positions of the two participants are

$$x_1(0) := \left(-48 - \frac{6}{\sqrt{2}}, 48 + \frac{6}{\sqrt{2}} \right), \quad x_2(0) := (-48, 48),$$

and that the exit is located at the origin $(0, 0)$.

Assume also that the participants exhibit the same behavior and aim to reach the exit by the shortest path. To regulate the participant speeds under the nonoverlapping condition, we use control functions in the moving set

$$u_1(t) = u_2(t) \text{ for all } t \in [0, 6] \text{ with the bounds } r_1 = 1 \text{ and } r_2 = 10 \tag{9.3}$$

as well as in the perturbations $a(\cdot) = (a_1(\cdot), a_2(\cdot)) : [0, 20] \to \mathbb{R}^2$ entering via the velocity function

$$f(x, a) := \left(\frac{s_1 a_1}{\|x_1\|} x_1, \frac{s_2 a_2}{\|x_2\|} x_2 \right) = (s_1 a_1 \cos\theta_1, s_1 a_1 \sin\theta_1, s_2 a_2 \cos\theta_2, s_2 a_2 \sin\theta_2),$$

where $s_1 = 6$, $s_2 = 3$, and $\theta_1 = \theta_2 = 135°$. Then the controlled dynamics is:

$$\begin{cases} -\dot{x}(t) \in N_{C(t)}\big(x(t)\big) + f(x(t), a(t)) & \text{a.e. } t \in [0, 6], \\ C(t) := C + u(t), \ u_1(t) = u_2(t), \ r_1 \leq \|u(t)\| \leq r_2 & \text{on } [0, 6], \ x(0) = x_0 \in C(0). \end{cases}$$
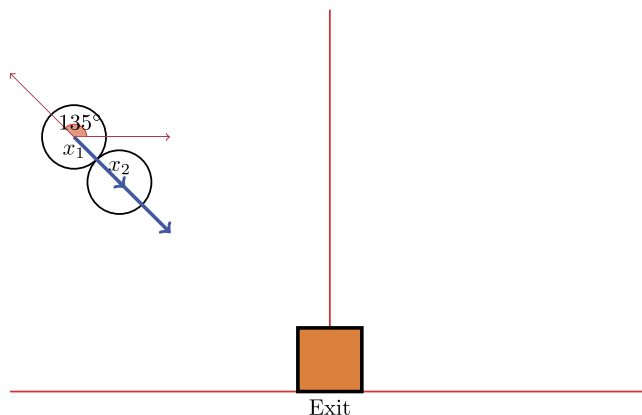
Fig. 2. Planar crowd motion model.

We consider the cost functional given of the form

$$J[x, u, a] := \frac{1}{2} \left(\|x(6)\|^2 + \int_0^6 \|a(t)\|^2 dt \right)$$

the meaning of which is to minimize the distance of the two participants to the exit together with the energy of feasible controls $a(\cdot)$ after six seconds. Applying now the necessary conditions of Theorem 8.1 yields the following, where $\lambda \geq 0$ and $\eta_{12} \in L^2([0, 6]; \mathbb{R}_+)$ are well defined at $t = 6$:

1. $w(t) = (0, 0, \bar{a}(t))$, $v(t) = (0, 0, 0)$ a.e. $t \in [0, 6]$;
2. $\dot{\bar{x}}(t) + f(\bar{x}(t), \bar{a}(t)) = \eta_{12}(t) \nabla g(\bar{x}(t) - \bar{u}(t)) = \left(-\eta_{12}(t) \frac{\bar{x}_2(t) - \bar{x}_1(t)}{\|\bar{x}_2(t) - \bar{x}_1(t)\|}, \eta_{12}(t) \frac{\bar{x}_2(t) - \bar{x}_1(t)}{\|\bar{x}_2(t) - \bar{x}_1(t)\|} \right)$;
3. $\|\bar{x}_2(t) - \bar{x}_1(t)\| > 2R \Longrightarrow \eta_{12}(t) = 0$ a.e. $t \in [0, 6]$;
4. $\eta_{12}(t) > 0 \Longrightarrow \langle q_2^x(t) - q_1^x(t), \bar{x}_2(t) - \bar{x}_1(t) \rangle = 0$ a.e. $t \in [0, 6]$;
5. $\dot{p}(t) = \left(0, 0, \lambda \bar{a}_1(t) - 6 \left(-\frac{\sqrt{2}}{2} q_{11}^x(t) + \frac{\sqrt{2}}{2} q_{12}^x(t) \right), \lambda \bar{a}_2(t) - 3 \left(-\frac{\sqrt{2}}{2} q_{21}^x(t) + \frac{\sqrt{2}}{2} q_{22}^x(t) \right) \right)$
   a.e. $t \in [0, 6]$;
6. $q^x(t) = p^x(t) + \gamma([t, 6])$ a.e. $t \in [0, 6]$;
7. $q^u(t) = p^u(t) - \int_{[t,6]} 2\bar{u}(s) d \left(\xi^1(s) + \xi^2(s) \right) + d\gamma(s) = 0$ a.e. $t \in [0, 6]$;
8. $q^a(t) = p^a(t) = 0$ a.e. $t \in [0, 6]$;
9. $p^x(6) + \lambda \bar{x}(6) = \left(-\eta_{12}(6) \frac{\bar{x}_2(6) - \bar{x}_1(6)}{\|\bar{x}_2(6) - \bar{x}_1(6)\|}, \eta_{12}(6) \frac{\bar{x}_2(6) - \bar{x}_1(6)}{\|\bar{x}_2(6) - \bar{x}_1(6)\|} \right)$;
10. $p^u(6) - \left(-\eta_{12}(6) \frac{\bar{x}_2(6) - \bar{x}_1(6)}{\|\bar{x}_2(6) - \bar{x}_1(6)\|}, \eta_{12}(6) \frac{\bar{x}_2(6) - \bar{x}_1(6)}{\|\bar{x}_2(6) - \bar{x}_1(6)\|} \right) \in -2\bar{u}(6) \big(N_{[0,2]}(\|\bar{u}(6)\|) + N_{[1,\infty]}(\|\bar{u}(6)\|) \big)$;
11. $\lambda + \|q^u(0)\| + \|p(6)\| + \|\xi^1\|_{TV} + \|\xi^2\|_{TV} > 0$.

Since the two participants are in contact at the initial time, i.e., $\|\bar{x}_2(0) - \bar{x}_1(0)\| = 6$, they have the same velocity and maintain it till the end of process. Thus the common velocity is constant on $[0, 6]$, which implies that $\bar{a}_i(\cdot) \equiv \bar{a}_i$ on this interval. Hence $\eta_{12}(\cdot) \equiv \eta_{12}$ on $[0, 6]$. It allows us to rewrite (2) as

$$\dot{\bar{x}}_1(t) = \left(3\sqrt{2}\bar{a}_1 - \frac{\sqrt{2}}{2}\eta_{12}, -3\sqrt{2}\bar{a}_1 + \frac{\sqrt{2}}{2}\eta_{12}\right),$$

$$\dot{\bar{x}}_2(t) = \left(\frac{3\sqrt{2}}{2}\bar{a}_2 + \frac{\sqrt{2}}{2}\eta_{12}, -\frac{3\sqrt{2}}{2}\bar{a}_2 - \frac{\sqrt{2}}{2}\eta_{12}\right),$$

which gives us by integration the trajectory formulas

$$\begin{cases} \bar{x}_1(t) = \left(-48 - \frac{6}{\sqrt{2}} + \left(3\sqrt{2}\bar{a}_1 - \frac{\sqrt{2}}{2}\eta_{12}\right)t, 48 + \frac{6}{\sqrt{2}} + \left(-3\sqrt{2}\bar{a}_1 + \frac{\sqrt{2}}{2}\eta_{12}\right)t\right), \\ \bar{x}_2(t) = \left(-48 + \left(\frac{3\sqrt{2}}{2}\bar{a}_2 + \frac{\sqrt{2}}{2}\eta_{12}\right)t, 48 + \left(-\frac{3\sqrt{2}}{2}\bar{a}_2 - \frac{\sqrt{2}}{2}\eta_{12}\right)t\right). \end{cases}$$

We have $\dot{\bar{x}}_1(t) = \dot{\bar{x}}_2(t)$ on $[0, 6]$ due to the same velocity of the participants, which yields

$$\eta_{12} = \frac{6\bar{a}_1 - 3\bar{a}_2}{2}. \tag{9.4}$$

Furthermore, it follows from the optimality conditions in (5) and (8) that

$$\lambda\bar{a}_1 = -3\sqrt{2}q_{11}^x(t) + 3\sqrt{2}q_{12}^x(t), \ \lambda\bar{a}_2 = -\frac{3\sqrt{2}}{2}q_{21}^x(t) + \frac{3\sqrt{2}}{2}q_{22}^x(t) \ \text{on} \ [0, 6]. \tag{9.5}$$

If $\eta_{12} > 0$, we immediately deduce from (4) that

$$-q_{11}^x(t) + q_{12}^x(t) = -q_{21}^x(t) + q_{22}^x(t) \ \text{on} \ [0, 6].$$

Combining the latter with (9.5) shows that $\bar{a}_1 = 2\bar{a}_2$ provided that $\lambda > 0$; otherwise we do not have enough information to proceed. Consider now the following two cases:

Case 1: $\eta_{12} = 0$. Then (9.4) tells us that $\bar{a}_2 = 2\bar{a}_1$ and the cost functional reads as

$$J[\bar{a}_1] = 1311\bar{a}_1^2 - 36(96\sqrt{2} + 6)\bar{a}_1 + \left(48 + \frac{6}{\sqrt{2}}\right)^2 + 48^2$$

while attaining its minimum at $\bar{a}_1 = \dfrac{(96\sqrt{2} + 6)18}{1311} \approx 1.95$. Thus $\bar{a}_2 = 2\bar{a}_1 \approx 3.9$, the minimum cost is $J \approx 66.49$, and the optimal trajectories are calculated by

$$\bar{x}_1(t) = \left(-48 - \frac{6}{\sqrt{2}} + 8.27t, 48 + \frac{6}{\sqrt{2}} - 8.27t\right), \ \bar{x}_2(t) = (-48 + 8.27t, 48 - 8.27t). \tag{9.6}$$

Case 2: $\eta_{12} > 0$. By $\bar{a}_1 = 2\bar{a}_2$ in this case we get $\eta_{12} = \frac{9}{2}\bar{a}_2$. The cost functional reads as

$$J[\bar{a}_2] = 2040\bar{a}_2^2 - 45(96\sqrt{2} + 6)\bar{a}_2 + \left(48 + \frac{6}{\sqrt{2}}\right)^2 + 48^2$$

with its minimum value achieved at $\bar{a}_2 = \dfrac{45(96\sqrt{2} + 6)}{4080} \approx 1.56$. Hence $\bar{a}_1 = 2\bar{a}_2 \approx 3.12$, the minimum cost is $J \approx 45.9$, and the optimal trajectories are given by the same formulas (9.6) as in Case 1. Observe that in both cases we have (9.3) for the corresponding u-control $\bar{u}(\cdot) = (\bar{u}_1(\cdot), \bar{u}_2(\cdot))$ satisfying the constraints $r_1 \leq \|\bar{u}(t)\| \leq r_2$, and thus we obtain a complete solution of the problem under consideration.

References

[1] 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.

[2] 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.

[3] S. Adly, F. Nacry, L. Thibault, Preservation of prox-regularity of sets with applications to constrained optimization, SIAM J. Optim. 26 (2016) 448–473.

[4] C.E. Arround, G. Colombo, A maximum principle of the controlled sweeping process, Set-Valued Var. Anal. 26 (2018), https://doi.org/10.1007/s11228-017-0400-4.

[5] M. Brokate, P. Krejčí, Optimal control of ODE systems involving a rate independent variational inequality, Discrete Contin. Dyn. Syst. Ser. B 18 (2013) 331–348.

[6] A. Canino, On p-convex sets and geodesics, J. Differential Equations 75 (1988) 118-0157.

[7] 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.

[8] 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.

[9] T.H. Cao, B.S. Mordukhovich, Applications of optimal control of a nonconvex sweeping process to optimization of the planar crowd motion model, arXiv:1802.08083, 2018, submitted for publication.

[10] C. Castaing, M.D.P. Monteiro Marques, P. Ranaud de Fitte, Some problems in optimal control governed by the sweeping process, J. Nonlinear Convex Anal. 15 (2014) 1043–1070.

[11] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, P.R. Wolenski, Nonsmooth Analysis and Control Theory, Springer, New York, 1998.

[12] G. Colombo, R. Henrion, N.D. Hoang, B.S. Mordukhovich, Optimal control of the sweeping process, Dyn. Contin. Discrete Impuls. Syst. Ser. B Appl. Algorithms 19 (2012) 117–159.

[13] G. Colombo, R. Henrion, N.D. Hoang, B.S. Mordukhovich, Optimal control of the sweeping process over polyhedral controlled sets, J. Differential Equations 260 (2016) 3397–3447.

[14] 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.

[15] T. Donchev, E. Farkhi, B.S. Mordukhovich, Discrete approximations, relaxation, and optimization of one-sided Lipschitzian differential inclusions in Hilbert spaces, J. Differential Equations 243 (2007) 301–328.

[16] J.F. Edmond, L. Thibault, Relaxation of an optimal control problem involving a perturbed sweeping process, Math. Program. 104 (2005) 347–373.

[17] H. Federer, Curvature measures, Trans. Amer. Math. Soc. 93 (1959) 418–491.

[18] R. Henrion, J. Outrata, T. Surowiec, On the coderivative of normal cone mappings to inequality systems, Nonlinear Anal. 71 (2009) 1213–1226.

[19] P. Krečí, Evolution variational inequalities and multidimensional hysteresis operators, in: P. Drabek, P. Krečí, P. Takac (Eds.), Nonlinear Differential Equations, in: Res. Notes Math., vol. 404, Chapman & Hall/CRC, London, UK, 1999, pp. 47–110.

[20] 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.

[21] B. Maury, J. Venel, A discrete model for crowd motion, ESAIM Math. Model. Numer. Anal. 45 (2011) 145–168.
[22] A. Mielke, Evolution of rate-independent systems, in: C.M. Dafermos, E. Feireis (Eds.), Handbook of Differential Equations: Evolutionary Equations, Vol. II, Elsevier B.V., Amsterdam, 2005, pp. 461–559.
[23] B.S. Mordukhovich, Sensitivity analysis in nonsmooth optimization, in: D.A. Field, V. Komkov (Eds.), Theoretical Aspects of Industrial Design, in: SIAM Proc. Appl Math., vol. 58, 1992, pp. 32–46.
[24] B.S. Mordukhovich, Discrete approximations and refined Euler–Lagrange conditions for differential inclusions, SIAM J. Control Optim. 33 (1995) 882–915.
[25] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation, I: Basic Theory, Springer, Berlin, 2006.
[26] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation, II: Applications, Springer, Berlin, 2006.
[27] B.S. Mordukhovich, R.T. Rockafellar, Second-order subdifferential calculus with applications to tilt stability in optimization, SIAM J. Optim. 22 (2012) 953–986.
[28] J.J. Moreau, Rafle par un convexe variable, I, Sem. d'Anal. Convexe, Montpellier, vol. 1 (1971), Exposé No. 15.
[29] J.J. Moreau, On unilateral constraints, friction and plasticity, in: G. Capriz, G. Stampacchia (Eds.), New Variational Techniques in Mathematical Physics. Proceedings of C.I.M.E. Summer Schools, Cremonese, Rome, 1974, pp. 173–322.
[30] J.J. Moreau, Evolution problem associated with a moving convex set in a Hilbert space, J. Differential Equations 26 (1977) 347–374.
[31] R.A. Poliquin, R.T. Rockafellar, Prox-regular functions in variational analysis, Trans. Amer. Math. Soc. 348 (1996) 1805–1838.
[32] F. Rindler, Approximation of rate-independent optimal control problems, SIAM J. Numer. Anal. 47 (2009) 3884–3909.
[33] R.T. Rockafellar, R.J-B. Wets, Variational Analysis, Springer, Berlin, 1998.
[34] A.H. Siddiqi, P. Manchanda, M. Brokate, On some recent developments concerning Moreau's sweeping process, in: A.H. Siddiqi, M. Kočvara (Eds.), Trends in Industrial and Applied Mathematics, Kluwer, Dordrecht, 2002, pp. 339–354.
[35] L. Thibault, Sweeping process with regular and nonregular sets, J. Differential Equations 193 (2003) 1–26.
[36] A.A. Tolstonogov, Control sweeping process, J. Convex Anal. 23 (2016) 1099–1123.
[37] J. Venel, A numerical scheme for a class of sweeping process, Numer. Math. 118 (2011) 451–484.
[38] R.B. Vinter, Optimal Control, Birkhaüser, Boston, 2000.