Journal of Convex Analysis
Volume 31 (2024), No. 3, 787–825

Pontryagin-Type Maximum Principle for a Controlled Sweeping Process with Nonsmooth and Unbounded Sweeping Set

Chadi Nour*
Dept. of Computer Science and Mathematics, Lebanese American University, Byblos, Lebanon
cnour@lau.edu.lb

Vera Zeidan
Department of Mathematics, Michigan State University, East Lansing, U.S.A.
zeidan@msu.edu

Received: March 8, 2023
Accepted: August 21, 2023

The Pontryagin-type maximum principle derived in a recent article of the authors [*Optimal control of nonconvex sweeping processes with separable endpoints: Nonsmooth maximum principle for local minimizers*, J. Diff. Equations 318 (2022) 113–168] for optimal control problems involving sweeping processes is generalized to the case where the sweeping set C is nonsmooth and not necessarily bounded, namely, C is the intersection of a finite number of sublevel sets of smooth functions.

Keywords: Controlled sweeping process, optimal control, Pontryagin-type maximum principle, intersection of sublevel sets, nonsmooth analysis.

2020 Mathematics Subject Classification: 49K21, 49K15, 49J52.

1. Introduction

Sweeping processes are dynamical systems involving the *normal cone* to a set called *sweeping set*. These systems were introduced in the papers [28, 29, 30] by J. J. Moreau in the context of plasticity and friction theory. Different modifications of this model have since appeared in many applications such as hysteresis, ferromagnetism, electric circuits, phase transitions, crowd motion problems, economics, etc. (see [1] and its references). This naturally motivated launching the subject of optimal control over sweeping processes and deriving *necessary optimality conditions* phrased in terms of *Euler-Lagrange equation* or *Pontryagin-type maximum principle*. These results are mainly established using two different approaches, namely *discrete* approximations (see [5, 6, 7, 8, 9, 14, 15, 16]), and *continuous* approximations (see [3, 17, 20, 32, 33, 34, 40]). Another new method is recently introduced in [23], where, instead of approximating the sweeping process, the authors formulated *standard* optimal control problems having *auxiliary* controls and the sweeping set is considered as *explicit* state constraints. These problems admit the same optimal solution as that of the original problem. As mentioned by the authors in [23, Remark 2.13], the Pontryagin-type necessary conditions derived therein have atypical nondegeneracy condition that requires further analysis.

* Corresponding author.

ISSN 0944-6532 / $ 2.50

Via an innovative *exponential penalization technique* that *approximates* the normal cone, the authors in [17, 18] derived a *smooth* Pontryagin-type maximum principle for *global* minimizers of a controlled sweeping process, in which the sweeping set C is *compact* and is the zero-sublevel set of a $\mathcal{C}^2$-convex function ψ. The surprising feature of this continuous-time approximation technique resides in the fact that, despite that the original sweeping process inherently has a state constraint, namely, $x(t) \in C$, the presence of the penalty term in the formulation of the approximating systems disposes of this state constraint, since the set C turns out to be invariant for this carefully designed approximating systems. This special property of the exponential penalization technique introduced in [17], has demonstrated to be instrumental in [19, 31], where successful numerical algorithms have been developed to efficiently compute *good* approximations of solutions for optimal control problems over sweeping processes via solutions of *standard* optimal control problems in which *no* state-constraint is imposed.

The same exponential penalization technique is used in [32, 40] to generalize the Pontryagin-type maximum principle of [17] in several directions. This includes allowing the compact sweeping set C to be the zero-sublevel set of a $\mathcal{C}^{1,1}$-function that is *not* necessarily convex, a final-state endpoint constraint set to be present, the cost function to depend on both endpoints of the state, strong local minimizers to be considered, and new subdifferentials that are strictly smaller than Clarke and Mordukhovich subdifferentials to be employed.

The goal of this paper is to extend the Pontryagin-type maximum principle of [32] to the case in which the sweeping set is the *intersection* of a finite number of zero-sublevel sets of $\mathcal{C}^{1,1}$-functions ψ_i, $i = 1, \ldots, r$, and hereby, we allow C to be nonsmooth. Moreover, the compactness assumption of C can now be replaced by a weaker condition permitting C to be *unbounded*, and hence, our results here extend the results in [17, 32, 40], not only to the case where the sweeping set is *nonsmooth*, but also *unbounded*, such as, polyhedral, star-shaped, or a set with compact boundaries.

The exponential penalization technique for $r > 1$ is far more complex than that for $r = 1$, which was treated in [17, 32, 40]. In fact, if we attempt here to emulate the case when $r = 1$ by considering C as the zero-sublevel set of the single function $\psi := \max\{\psi_i : i = 1, \ldots, r\}$, then, when $r > 1$ and $\psi_i \in \mathcal{C}^{1,1}$, for all $i \in \{1, \ldots, r\}$, ψ is only guaranteed to be Lipschitz. Hence, unlike the case when $r = 1$, the normal cone to C, N_C, is now given in terms of the subgradient $\partial\psi$, as opposed to $\nabla\psi$, and this set-valued map is not in general Lipschitz. This would lead to approximating the normal cone in the sweeping process by an *exponential penalty* term invoking $\partial\psi$, which produces a differential inclusion to which none of the standard results would apply. Thus, this attempt would not be successful. On the other hand, if we express the normal cone to C at a boundary point x as the cone combination of the corresponding $\nabla\psi_i(x)$, the approximation of $N_C(x)$ in the sweeping process via the exponential penalty technique leads, as in the case $r = 1$, to an ordinary control system, which now involves r *exponential penalty* terms, instead of only one term. The presence of r- penalty terms causes a major obstacle when attempting to prove that C is invariant for the resulting control system, unless one imposes a very *restrictive* assumption such as: $\langle \nabla\psi_i(x), \nabla\psi_j(x)\rangle \geq 0$ in a band around the boundary of C, which excludes sets having acute angles, like triangles, etc.[1]

Therefore, it is clear that in order to deal with the intricacy of the situation for $r > 1$ and to circumvent *both* major obstacles mentioned above, a new idea is required. Our approach in this paper is to approximate the *nonsmooth* max-function ψ, defining the sweeping set C, by a sequence $(\psi_{\gamma_k})_k$ of $\mathcal{C}^{1,1}$-functions, constructed by applying the "epi-multiple log-exponential" operator $\left(\frac{1}{\gamma_k}\star \operatorname{log}\operatorname{exp}\right)$ to the "vecmax" function ψ (see [36, Example 1.30]). This allows us to approximate the nonsmooth set C by a sequence of smooth sets C^{γ_k}, the zero-sublevel sets of ψ_{γ_k}. We employ the exponential penalization technique for ψ_{γ_k} which produces the *same* approximating control system that has r-penalty terms involving $\nabla\psi_i$, $i = 1, ..., r$. For this system, it turns out that C^{γ_k}, instead of C, is actually *invariant* under *no* extra assumption.

This result allowed us to move towards proving a Pontryagin-type maximum principle for $r > 1$ and for also nonsmooth data. However, since the generalized Hessian of the function ψ_{γ_k} is *not* bounded near the given optimal state $\bar{x}$, another complexity surfaces when showing the sequence of adjoint variables p_{γ_k} for the approximating problems has uniform bounded variation. This is resolved by reverting to the form of the approximating dynamic in terms of $(\nabla\psi_i)_{i=1}^r$ and by assuming a *local* condition at the values $\bar{x}(t)$, assumption (A2.3), that invokes the gradients of the corresponding active constraints and that is automatically satisfied when $r = 1$. Note that, unlike [21], we do *not* impose any assumption like $\nabla\psi_i = 0$ on the complement in C of a band around the boundary of C. The absence of this assumption causes the proof to be more challenging.

The layout of the paper is as follows. In the next section, we first display our basic notations, then, we state our optimal control problem (P) over a sweeping process, and we list our hypotheses. In Section 3, we present our main result Theorem 3.1, in which we derive nonsmooth Pontryagin-type maximum principle for *local* minimizers of (P). Section 4 consists of preparatory results that are instrumental for the proof of Theorem 3.1 established in Section 5. To promote continuous flow in the presentation of the results, some of the proofs are provided in the appendix (Section 6), where we also show some auxiliary results that will be used in several places of the paper. An example illustrating the utility of Theorem 3.1 is provided in Section 6.

2. Preliminaries

2.1. Notations

We begin this subsection by briefly presenting the basic notations used in this paper. We use $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ for the Euclidean norm and the usual inner product, respectively. The open unit ball, the closed unit ball, and the unit sphere are denoted by B, $\bar{B}$, and $\mathbb{S}$, respectively. For $x \in \mathbb{R}^n$ and $\rho \geq 0$, the open ball, the closed balls, and the sphere of radius ρ centered at x are written as $B_\rho(x)$, $\bar{B}_\rho(x)$, and $\mathbb{S}_\rho(x)$, respectively. For a set $A \subset \mathbb{R}^n$, $\operatorname{int} A$, $\operatorname{bdry} A$, $\operatorname{cl} A$, $\operatorname{conv} A$, and A^c designate the interior, the boundary, the closure, the convex hull, and the complement of A, respectively.

[1] During the final writing of this paper, we became aware of the manuscript [21] posted on arXiv in which this latter condition is imposed.

The set $A \subset \mathbb{R}^n$ is said to be star-shaped if there exists $a_o \in A$, called center of A, such that the closed interval $[a_o, x] \subset A$ for all $x \in A$. When $A \subset \mathbb{R}^n$ is closed and convex, we denote by h_A the support function of A. For $f\colon \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$ an extended-real-valued function, $\operatorname{dom} f$ is the effective domain of f and $\operatorname{epi} f$ is its epigraph. The Lebesgue space of p-integrable functions $f\colon [a,b] \longrightarrow \mathbb{R}^n$ is denoted by $L^p([a,b];\mathbb{R}^n)$. The norms in $L^p([a,b];\mathbb{R}^n)$ and $L^\infty([a,b];\mathbb{R}^n)$ (or $C([a,b];\mathbb{R}^n)$) are written as $\|\cdot\|_p$ and $\|\cdot\|_\infty$, respectively. The space $W^{1,1}([a,b];\mathbb{R}^n)$ denotes the set of all absolutely continuous functions $x\colon [a,b] \to \mathbb{R}^n$. The set of all functions $f\colon [a,b] \longrightarrow \mathbb{R}^n$ of bounded variations is denoted by $BV([a,b];\mathbb{R}^n))$. The space $C^*([a,b];\mathbb{R})$ denotes the dual of $C([a,b];\mathbb{R})$ equipped with the supremum norm. We denote by $\|\cdot\|_{\text{T.V.}}$ the induced norm on $C^*([a,b];\mathbb{R})$. By Riesz representation theorem, each element in $C^*([a,b];\mathbb{R})$ can be interpreted as an element in the space of finite signed Radon measures on $[a,b]$ equipped with the weak* topology. For the set of $m \times n$-matrix functions on $[a,b]$, we use $\mathcal{M}_{m\times n}([a,b])$. For $A \subset \mathbb{R}^d$ compact, $C(A;\mathbb{R}^n)$ denotes to the set of continuous functions from A to $\mathbb{R}^n$. A vector $\beta := (\beta_i, \dots, \beta_n) \in \mathbb{R}^n$ is said to be positive if $\beta_i > 0$ for $i = 1, \dots, n$.

In what follows, we display some notations from *nonsmooth analysis*. For standard references, see the monographs [10, 12, 26, 36, 38]. Let A be a nonempty and closed subset of $\mathbb{R}^n$, and let $a \in A$. The *proximal*, the *Mordukhovich* (also known as *limiting*), and the *Clarke normal* cones to A at a are denoted by $N_A^P(a)$, $N_A^L(a)$, and $N_A(a)$, respectively. For the *Clarke tangent* cone to A at a, we use $T_A(a)$. Given a *lower semicontinuous* function $f\colon \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$, and $x \in \operatorname{dom} f$, the *proximal*, the *Mordukhovich* (or *limiting*), and the *Clarke subdifferential* of f at x are denoted by $\partial^P f(x)$, $\partial^L f(x)$, and $\partial f(x)$, respectively. Note that if $x \in \operatorname{int}(\operatorname{dom} f)$ and f is Lipschitz near x, [10, Theorem 2.5.1] yields that the Clarke subdifferential of f at x coincides with the *Clarke generalized gradient* of f at x, also denoted here by $\partial f(x)$. If f is $\mathcal{C}^{1,1}$ near $x \in \operatorname{int}(\operatorname{dom} f)$, $\partial^2 f(x)$ denotes the *Clarke generalized Hessian* of f at x. For $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ Lipschitz near $x \in \mathbb{R}^n$, $\partial g(x)$ denotes the *Clarke generalized Jacobian* of g at x.

2.2. Statement of the problem (P), and assumptions

In this paper, we consider the following fixed time Mayer problem

$$(P) \qquad \text{Minimize } g(x(0), x(T)) \text{ over } (x,u) \text{ such that } u(\cdot) \in \mathcal{U} \text{ and}$$
$$\begin{cases} (D) \begin{bmatrix} \dot{x}(t) \in f(t, x(t), u(t)) - \partial\varphi(x(t)), \text{ a.e. } t \in [0,T], \\ x(0) \in C_0 \subset \operatorname{dom}\varphi, \end{bmatrix} \\ x(T) \in C_T, \end{cases}$$

where $T > 0$ is fixed, $g\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$, $f\colon [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$, $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$ is lower semicontinuous, $\partial\varphi$ stands for the Clarke subdifferential of φ, $C := \operatorname{dom}\varphi$ is the intersection of the zero-sublevel sets of a finite sequence of functions $\psi_i\colon \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1, \dots, r$, $C_0 \subset C$ is the initial constraint set, $C_T \subset \mathbb{R}^n$ is the terminal constraint set, and, for $U\colon [0,T] \rightrightarrows \mathbb{R}^m$ a multifunction, the set of control functions $\mathcal{U}$ is defined as

$$\mathcal{U} := \{u\colon [0,T] \longrightarrow \mathbb{R}^m \text{ is measurable and } u(t) \in U(t) \text{ a.e. } t \in [0,T]\}.$$

A pair (x,u) is *admissible* for (P) when $x\colon [0,T] \longrightarrow \mathbb{R}^n$ is absolutely continuous, $u \in \mathcal{U}$, and (x,u) satisfies the *perturbed sweeping process* (D), called the dynamic of (P). An admissible pair $(\bar{x},\bar{u})$ is said to be a *strong local minimizer* if there exists $\delta > 0$ such that

$$g(\bar{x}(T)) \leq g(x(T)),\ \ \forall (x,u) \text{ admissible for } (P) \text{ and satisfying } \|x-\bar{x}\|_\infty \leq \delta.$$

Note that the admissibility of a pair (x,u) yields from (D) that $x(t) \in C$, $\forall t \in [0,T]$.

Now let $(\bar{x},\bar{u})$ be a strong local minimizer of (P). Next, we present our assumptions on the data of (P) needed to derive our Pontryagin-type maximum principle for $(\bar{x},\bar{u})$. Note that not all these assumptions are needed for our intermediate and auxiliary results. Moreover, a global version of (A1) is introduced later in Section 4. Define the sets

$$\mathbb{U} := \bigcup_{t\in[0,T]} U(t) \quad \text{and} \quad C_{\bar{x}} := \bigcup_{t\in[0,T]} C \cap \bar{B}_\delta(\bar{x}(t)).$$

A1: For fixed $(x,u) \in C_{\bar{x}} \times \mathbb{U}$, $f(\cdot,x,u)$ is Lebesgue-measurable; and there exists $M_\ell > 0$ such that, for $t \in [0,T]$ a.e., we have: $(x,u) \mapsto f(t,x,u)$ is continuous on $(C \cap \bar{B}_\delta(\bar{x}(t))) \times U(t)$; for all $u \in U(t)$, $x \mapsto f(t,x,u)$ is M_ℓ-Lipschitz on $C \cap \bar{B}_\delta(\bar{x}(t))$; and $\|f(t,x,u)\| \leq M_\ell$ for all $(x,u) \in (C \cap \bar{B}_\delta(\bar{x}(t))) \times U(t)$.

A2: The set $C := \operatorname{dom}\varphi \neq \emptyset$ is given by

$$C := \bigcap_{i=1}^{r} C_i, \tag{1}$$

where $C_i := \{x \in \mathbb{R}^n : \psi_i(x) \leq 0\}$ and $(\psi_i)_{1\leq i\leq r}$ is a family of functions $\psi_i\colon \mathbb{R}^n \longrightarrow \mathbb{R}$.

A2.1: There exists $\rho > 0$ such that for each i, the function ψ_i is $\mathcal{C}^{1,1}$ on $C + \rho B$.

A2.2: There is a constant $\eta > 0$ such that

$$\Big\|\sum_{i\in\mathcal{I}^0_c} \lambda_i \nabla\psi_i(c)\Big\| > 2\eta, \ \ \forall c \in \{x \in C : \mathcal{I}^0_x \neq \emptyset\},$$

where $\mathcal{I}^0_x := \{i \in \{1,\dots,r\} : \psi_i(x) = 0\}$ and $(\lambda_i)_{i\in\mathcal{I}^0_c}$ is any sequence of nonnegative numbers satisfying $\sum_{i\in\mathcal{I}^0_c} \lambda_i = 1$.

A2.3: There exists a positive vector $\beta = (\beta_1,\dots,\beta_r) \in \mathbb{R}^r$ such that

$$\sum_{i\in\mathcal{I}^0_{\bar{x}(t)},\ i\neq j} \beta_i |\langle \nabla\psi_i(\bar{x}(t)), \nabla\psi_j(\bar{x}(t))\rangle| < \beta_j \|\nabla\psi_j(\bar{x}(t))\|^2$$

for all $t \in I^0(\bar{x})$ and $j \in \mathcal{I}^0_{\bar{x}(t)}$, where

$$I^0(\bar{x}) := \{t \in [0,T] : \bar{x}(t) \in \operatorname{bdry} C\} = \{t \in [0,T] : \mathcal{I}^0_{\bar{x}(t)} \neq \emptyset\}.$$

A2.4: There exist $y_o \in \mathbb{R}^n$ and $R_o > 0$ such that:

(i) For all $c \in (\operatorname{bdry} C) \cap \mathbb{S}_{R_o}(y_o)$, we have $(y_o - c) \notin N_C(c)$.

(ii) $\bar{x}(t) \in B_{R_o}(y_o)$ for all $t \in [0,T]$.

A2.5: The set C has a connected interior.[2]

A3: The function φ is globally Lipschitz on C and $\mathcal{C}^1$ on $\operatorname{int} C$, and the function $\nabla\varphi$ is globally Lipschitz on $\operatorname{int} C$.

A4: The following assumptions on C_0, C_T and U hold:

A4.1: The set $C_0 \subset C$ is nonempty and closed.

A4.2: The set $C_T \subset \mathbb{R}^n$ is nonempty and closed.

A4.3: The graph of $U(\cdot)$ is a $\mathcal{L} \otimes \mathcal{B}$ measurable set, and, for $t \in [0, T]$, $U(t)$ is closed, and bounded uniformly in t, where $\mathcal{B}$ is the Borel σ-field of $\mathbb{R}^m$.

A5: There exists $\tilde{\rho} > 0$ such that g is L_g-Lipschitz on $\tilde{C}_0(\delta) \times \tilde{C}_T(\delta)$, where $\tilde{C}_0(\delta) := \left[\left(C_0 \cap \overline{B}_\delta(\bar{x}(0))\right) + \tilde{\rho}\overline{B}\right] \cap C$ and $\tilde{C}_T(\delta) := \left[\left(C_T \cap \overline{B}_\delta(\bar{x}(T))\right) + \tilde{\rho}\overline{B}\right] \cap C$.

A6: The set $f(t, x, U(t))$ is convex for all $x \in C \cap \overline{B}_\delta(\bar{x}(t))$ and $t \in [0, T]$ a.e.[3]

Remark 2.1. Assumption (A2.4) is *weaker* than the compactness of C. In fact, by Lemma 6.4, assumption (A2.4)(i) is satisfied by an *arbitrary large* R_o when the set C has a compact boundary or is star-shaped (which includes *unbounded* convex and polyhedral sets). Hence, taking R_o large enough, we guarantee that also (A2.4)(ii) is valid by those sets.

Remark 2.2. One can prove that (A2.1)–(A2.2) imply that, for each $x \in \operatorname{bdry} C$, the family of vectors $\{\nabla\psi_i(x)\}_{i\in\mathcal{I}^0_x}$ is *positive-linearly independent.* Furthermore, when C is compact, an argument by contradiction using Lemma 6.2 shows that assumption (A2.2) is actually *equivalent* to saying that $\{\nabla\psi_i(x)\}_{i\in\mathcal{I}^0_x}$ is positive-linearly independent for all $x \in \operatorname{bdry} C$.

Remark 2.3. Under (A2.1) and for $t \in I^0(\bar{x})$ and $\beta := (\beta_1, \ldots, \beta_r) \in \mathbb{R}^r$, denote by $G_\psi(t)$ the *Gramian matrix* of the vectors $\{\nabla\psi_i(\bar{x}(t))\}_{i\in\mathcal{I}^0_{\bar{x}(t)}}$ and by $D_\beta(t)$ the *diagonal matrix* whose diagonal entries are $(\beta_i)_{i\in\mathcal{I}^0_{\bar{x}(t)}}$. It follows that condition (A2.3) is *equivalent* to the existence of a positive vector $\beta = (\beta_1, \ldots, \beta_r) \in \mathbb{R}^r$ such that, for all $t \in I^0(\bar{x})$, the matrix $G_\psi(t)D_\beta(t)$ is *strictly diagonally dominant*, and hence, invertible. Therefore, (A2.1) and (A2.3) imply that for each $t \in I^0(\bar{x})$, the family of vectors $\{\nabla\psi_i(\bar{x}(t))\}_{i\in\mathcal{I}^0_{\bar{x}(t)}}$ is *linearly independent.* It is worth mentioning that a condition related to assumption (A2.3), see [24, Theorem 1.3.1], was first acknowledged in [22] to be effective for the study of sweeping (reflected) processes with *nonsmooth* sweeping sets.

3. Main result

In this section, we display our Pontryagin-type maximum principle for strong local minimizers of (P) whose proof will be given to Section 5. We employ in its statement the following *nonstandard* notions of subdifferentials, which are *strictly smaller* than their standard counterparts:

[2] This assumption is only imposed to obtain the required quasiconvexity of C for the extension function Φ of φ, see Proposition 4.1(ii). Thus, when such an extension is readily available, as is the case when φ is the *indicator* function of C, assumption (A2.5) would be superfluous. For more information about this assumption, see [32, Remark 3.2].

[3] When $C_T = \mathbb{R}^n$, this assumption is *not* required for the main result, i.e., Theorem 3.1.

- $\partial_\ell \varphi$ and $\partial_\ell^2 \varphi$ are the *extended Clarke generalized gradient* and the *extended Clarke generalized Hessian* of φ defined on C, respectively (see [32, Equation (9)–(10)]). Note that if $\partial_\ell \varphi(x)$ is a singleton, then we use the notation ∇_ℓ instead of ∂_ℓ.
- $\partial_\ell^2 \psi$ is the *Clarke generalized Hessian relative* to $\operatorname{int} C$ of ψ (see [32, Eq. (12)]).
- $\partial_\ell^x f(t,\cdot,u)$ is the *extended Clarke generalized Jacobian* of $f(t,\cdot,u)$ defined on $C \cap \bar{B}_\delta(\bar{x}(t))$ (see [32, Equation (11)]).
- $\partial_\ell^L g$ is the *limiting subdifferential* of g *relative* to $\operatorname{int}\big(\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)\big)$ (see [32, Equation (8)]).

For given $x(\cdot) \in W^{1,1}([0,T];C)$, we define the sets

$$I_i(x) := \{t \in [0,T] : x(t) \in \operatorname{int} C_i\} \;\text{ and }\; I_i^0(x) := [0,T] \setminus I_i(x), \;\; \forall i = 1,\dots,r. \quad (2)$$

Theorem 3.1. (Generalized Maximum Principle for (P)) *Let $(\bar{x},\bar{u})$ be a strong local minimizer for (P), and assume that* (A1)–(A6) *hold. Then, there exist an adjoint vector $p \in BV([0,T];\mathbb{R}^n)$, a finite sequence of finite signed Radon measures $(\nu^i)_{i=1}^r$ on $[0,T]$ with ν_i supported in $I_i^0(\bar{x})$ for $i = 1,\dots,r$, a finite sequence of nonnegative functions $(\xi^i)_{i=1}^r \in L^\infty([0,T];\mathbb{R}^+)$ with ξ_i supported in $I_i^0(\bar{x})$ for all $i = 1,\dots,r$, L^∞-functions $\zeta(\cdot)$, $\theta(\cdot)$ and a finite sequence $(\vartheta_i(\cdot))_{i=1}^r$ in $\mathcal{M}_{n\times n}([0,T])$, and $\lambda \geq 0$ satisfying the following*:

(i) (The primal-dual admissible equation)
  (a) $\dot{\bar{x}}(t) = f(t,\bar{x}(t),\bar{u}(t)) - \nabla_\ell \varphi(\bar{x}(t)) - \sum_{i=1}^r \xi^i(t) \nabla \psi_i(\bar{x}(t)), \;\; \forall t \in [0,T]$ a.e.,
  (b) $\psi_i(\bar{x}(t)) \leq 0, \;\; \forall t \in [0,T]$ *and* $\forall i \in \{1,\dots,r\}$;

(ii) (The nontriviality condition) $\|p(T)\| + \lambda = 1$;

(iii) (The adjoint equation) $\forall t \in [0,T]$ a.e.,

$$(\zeta(t),\theta(t)) \in \partial_\ell^x f(t,\bar{x}(t),\bar{u}(t)) \times \partial_\ell^2 \varphi(\bar{x}(t)),$$
$$\vartheta_i(t) \in \partial_\ell^2 \psi_i(\bar{x}(t)) \;\; \textit{for } i = 1,\dots,r,$$

and, for any $z \in C([0,T];\mathbb{R}^n)$, we have

$$\int_{[0,T]} \langle z(t), dp(t)\rangle = \int_0^T \big\langle z(t), \big(\theta(t) - \zeta(t)^\mathsf{T}\big)\, p(t)\big\rangle\, dt$$
$$+ \sum_{i=1}^r \int_0^T \xi^i(t)\, \langle z(t), \vartheta_i(t) p(t)\rangle\, dt + \sum_{i=1}^r \int_{[0,T]} \langle z(t), \nabla \psi_i(\bar{x}(t))\rangle d\nu^i(t),$$

where the left integral is the Riemann–Stieltjes integral of z with respect to p;

(iv) (The complementary slackness conditions) *For $i = 1,\dots,r$, we have*:
  (a) $\xi^i(t) = 0, \;\; \forall t \in I_i(\bar{x})$,
  (b) $\xi^i(t)\langle \nabla \psi_i(\bar{x}(t)), p(t)\rangle = 0, \;\; \forall t \in [0,T]$ a.e.;

(v) (The transversality condition)

$$(p(0), -p(T)) \in \lambda \partial_\ell^L g(\bar{x}(0),\bar{x}(T)) + \big[N_{C_0}^L(\bar{x}(0)) \times N_{C_T}^L(\bar{x}(T))\big];$$

(vi) (The maximization condition)

$$\max_{u\in U}\langle f(t,\bar{x}(t),u),p(t)\rangle \text{ is attained at } \bar{u}(t) \text{ for a.e. } t\in[0,T].$$

Furthermore, if $C_T=\mathbb{R}^n$ *then,* $\lambda=1$ *and the assumption* (A6) *is discarded.*

4. Preparatory results

In this section, we present preparatory results that are fundamental for the proof of Theorem 3.1. In addition, an existence theorem for an optimal solution of (P) is provided in Proposition 4.15.

We assume throughout this section that C is *compact*. The last step in the proof of Theorem 3.1 (Section 5) provides a technique that permits replacing the compactness of C by the weaker assumption (A2.4).

We denote by $\bar{M}_\psi$ a *common upper bound* on C of the finite sequence $(\|\nabla\psi_i(\cdot)\|)_{i=1}^r$, and by $2M_\psi$ a *common Lipschitz constant* of the finite family $\{\nabla\psi_i\}_{i=1}^r$ over the compact set $C+\frac{\rho}{2}\bar{B}$ such that $\bar{M}_\psi\geq 2\eta$ and $M_\psi\geq\frac{4\eta}{\rho}$.

As for the case $r=1$ in [32], (A2.2) and the compactness of C imply the existence of $\varepsilon>0$ such that for each $i\in\{1,\dots,r\}$ we have

$$\big[x\in C \text{ and } \|\nabla\psi_i(x)\|\leq\eta\big] \implies \psi_i(x)<-\varepsilon. \tag{3}$$

4.1. Properties of C, extension of φ, and new notations

We give in the following proposition some important consequences of our assumptions on the set C and the function φ.

Proposition 4.1. (Properties of C − Extension of φ)
Under (A2.1)–(A2.2)*, we have the following:*

(i) *C is amenable,*[4] *$\frac{\eta}{M_\psi}$-prox-regular,*[5] *epi-Lipschitz with* $C=\mathrm{cl}\,(\mathrm{int}\,C)$*,*[6] *and, for all* $x\in\mathrm{bdry}\,C$ *we have*

$$N_C(x)=N_C^P(x)=N_C^L(x)=\left\{\sum_{i\in\mathcal{I}_x^0}\lambda_i\nabla\psi_i(x):\lambda_i\geq 0\right\}\neq\{0\}. \tag{4}$$

Moreover, we have:

$$\mathrm{int}\,C=\bigcap_{i=1}^r\mathrm{int}\,C_i=\bigcap_{i=1}^r\{x\in\mathbb{R}^n:\psi_i(x)<0\}\neq\emptyset$$

and
$$\mathrm{bdry}\,C=C\cap\left(\bigcup_{i=1}^r\mathrm{bdry}\,C_i\right)\neq\emptyset.$$

[4] See [36, Chapter 10].

[5] For $\rho>0$, the closed set $A\subset\mathbb{R}^n$ is said to be *ρ-prox-regular* if for all $a\in A$ and for all unit vector $\zeta\in N_A^P(a)$, we have $\langle\zeta,x-a\rangle\leq\frac{1}{2\rho}\|x-a\|^2$ for all $x\in A$. This latter inequality is known as the *proximal normal inequality*. For more information about prox-regularity, see [35] and references therein.

[6] A closed set $A\subset\mathbb{R}^n$ is said to be *epi-Lipschitz* if for all $a\in A$, the Clarke normal cone of A at a is *pointed*, that is, $N_A(a)\cap -N_A(a)=\{0\}$. For more information about this property, see [10, 12, 36].

(ii) *If also* (A2.5) *holds, then* $\operatorname{int} C$ *is quasiconvex.*[7] *Furthermore, if in addition* (A3) *holds, then there exists a function* $\Phi \in \mathcal{C}^{1,1}(\mathbb{R}^n)$ *such that:*

- Φ *is bounded on* $\mathbb{R}^n$ *and equals to* φ *on* C.
- Φ *and* $\nabla\Phi$ *are globally Lipschitz on* $\mathbb{R}^n$.
- *For all* $x \in C$ *we have*

$$\partial\varphi(x) = \{\nabla\Phi(x)\} + N_C(x). \tag{5}$$

Proof. (i): The amenability of C follows from [36, Example 10.24(d)]. For the prox-regularity of C and the formula (4), use [2, Theorem 9.1] and [13, Corollary 4.15]. Now from (A2.2) and (4), we conclude that $N_C(x)$ is pointed for all $x \in \operatorname{bdry} C$, and hence C is epi-Lipschitz, and thus, $C = \operatorname{cl}(\operatorname{int} C)$. Hence, as $C \neq \emptyset$ and compact, then, $\operatorname{int} C \neq \emptyset$ and $\operatorname{bdry} C \neq \emptyset$. The two formulas of $\operatorname{int} C$ and $\operatorname{bdry} C$ follow directly from C being defined by (1) and from [32, Lemma 3.3].
(ii): See the proof of [40, Proposition 3.1(iii)]. □

Remark 4.2. The extension Φ of φ obtained in Proposition 4.1(ii), satisfies for some $K > 0$,

$$|\Phi(x)| \leq K, \;\; \|\nabla\Phi(x)\| \leq K, \;\text{ and }\; \|\nabla\Phi(x) - \nabla\Phi(y)\| \leq K\|x-y\|, \;\; \forall x,\, y \in \mathbb{R}^n.$$

Equation (5) implies that the sweeping process (D) can now be rephrased in terms of the normal cone to C and the extension Φ of φ as the following

$$(D) \qquad \begin{cases} \dot{x}(t) \in f(t, x(t), u(t)) - \nabla\Phi(x(t)) - N_C(x(t)), \text{ a.e. } t \in [0,T], \\ x(0) \in C_0 \subset C. \end{cases}$$

The reformulation of (D) in Remark 4.2 motivates defining the function f_Φ from $[0,T] \times \mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R}^n$ by

$$f_\Phi(t,x,u) := f(t,x,u) - \nabla\Phi(x), \quad \forall (t,x,u) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m.$$

Clearly $(\mathrm{A1})_{\mathbb{G}}$, (A2.1)–(A2.2), (A2.5), and (A3) imply that the following holds true for f_Φ:

(A1)$_\Phi$ For fixed $(x,u) \in C \times \mathbb{U}$, $f_\Phi(\cdot, x, u)$ is Lebesgue-measurable; and for a.e $t \in [0,T]$ we have: $(x,u) \mapsto f_\Phi(t,x,u)$ is continuous on $C \times U(t)$; for all $u \in U(t)$, $x \mapsto f_\Phi(t,x,u)$ is $\bar{M}$-Lipschitz on C; and $\|f_\Phi(t,x,u)\| \leq \bar{M}$ for all $(x,u) \in C \times U(t)$, where $\bar{M} := M + K$.

The following notations will be used throughout the paper:

- We denote by $(\gamma_k)_k$ a sequence satisfying $\gamma_k > \frac{2\bar{M}}{\eta}$ for all $k \in \mathbb{N}$, and $\gamma_k \to \infty$ as $k \to \infty$. We define the two real sequences $(\alpha_k)_k$ and $(\sigma_k)_k$ by

$$\alpha_k := \frac{\ln\left(\frac{\eta\gamma_k}{2\bar{M}}\right)}{\gamma_k} \quad \text{and} \quad \sigma_k := \frac{r\bar{M}_\psi}{2\eta^2}\left(\frac{\ln(r)}{\gamma_k} + \alpha_k\right), \quad k \in \mathbb{N}. \tag{6}$$

[7] A set $A \subset \mathbb{R}^n$ is *quasiconvex* ([4]) if there exists $c \geq 0$ such that any two points a, b in A can be joined by a polygonal line γ in A satisfying $l(\gamma) \leq c\|a-b\|$, where $l(\gamma)$ denotes the length of γ.

For $(\alpha_k)_k$, we have $\gamma_k e^{-\alpha_k\gamma_k} = \frac{2\bar{M}}{\eta}$, $\alpha_k > 0$ for all $k \in \mathbb{N}$, $\alpha_k \searrow$ and $\alpha_k \longrightarrow 0$. For $(\sigma_k)_k$, we have $\sigma_k > 0$ for all $k \in \mathbb{N}$, $\sigma_k \searrow$ and $\sigma_k \longrightarrow 0$.

- The function $\psi\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is defined by

$$\psi(x) := \max\{\psi_i(x) : i = 1, \dots, r\}, \quad \forall x \in \mathbb{R}^n. \tag{7}$$

Clearly we have that $C = \{x \in \mathbb{R}^n : \psi(x) \le 0\}$.

- For $k \in \mathbb{N}$, we define the function $\psi_{\gamma_k}\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$\psi_{\gamma_k}(x) := \frac{1}{\gamma_k} \ln\left(\sum_{i=1}^{r} e^{\gamma_k \psi_i(x)}\right), \quad \forall x \in \mathbb{R}^n. \tag{8}$$

We also define, for each $k \in \mathbb{N}$,

$$C^{\gamma_k} := \{x \in \mathbb{R}^n : \psi_{\gamma_k}(x) \le 0\} = \Big\{x \in \mathbb{R}^n : \sum_{i=1}^{r} e^{\gamma_k \psi_i(x)} \le 1\Big\}, \text{ and} \tag{9}$$

$$C^{\gamma_k}(k) := \{x \in \mathbb{R}^n : \psi_{\gamma_k}(x) \le -\alpha_k\} = \Big\{x \in \mathbb{R}^n : \sum_{i=1}^{r} e^{\gamma_k \psi_i(x)} \le \frac{2\bar{M}}{\eta\gamma_k}\Big\}. \tag{10}$$

One can easily see that if $r = 1$, then ψ_{γ_k} and C^{γ_k} coincide with ψ and C, respectively.

- For given $x(\cdot) \in W^{1,1}([0,T];C)$, $a \ge 0$, and $y \in C$, we define the sets

$$I(x) := \bigcap_{i=1}^{r} I_i(x) = \{t \in [0,T] : x(t) \in \operatorname{int} C\}, \text{ and hence } I^0(x) = [0,T] \setminus I(x),$$

where $I_i(x)$ is the set defined in (2).

$$\mathcal{I}_y^a := \{i \in \{1, \dots, r\} : -a \le \psi_i(y) \le 0\} \text{ and } I^a(x) := \{t \in [0,T] : \mathcal{I}_{x(t)}^a \ne \emptyset\}. \tag{11}$$

Note that when (A2.1) holds, $I^a(x)$ is compact by Lemma 6.2.

- The approximation dynamic (D_{γ_k}) is defined by

$$(D_{\gamma_k}) \left[\begin{array}{l} \dot{x}(t) = f_\Phi(t, x(t), u(t)) - \displaystyle\sum_{i=1}^{r} \gamma_k e^{\gamma_k \psi_i(x(t))} \nabla\psi_i(x(t)) \text{ a.e. } t \in [0,T], \\ x(0) \in C^{\gamma_k}. \end{array}\right. \tag{12}$$

One can easily verify that the system (D_{γ_k}) can be rewritten using the function ψ_{γ_k} as the following

$$(D_{\gamma_k}) \left[\begin{array}{l} \dot{x}(t) = f_\Phi(t, x(t), u(t)) - \gamma_k e^{\gamma_k \psi_{\gamma_k}(x(t))} \nabla\psi_{\gamma_k}(x(t)) \text{ a.e. } t \in [0,T], \\ x(0) \in C^{\gamma_k}, \end{array}\right. \tag{13}$$

where, by (8),

$$\nabla\psi_{\gamma_k}(x) = \frac{\sum_{i=1}^{r} e^{\gamma_k \psi_i(x)} \nabla\psi_i(x)}{\sum_{i=1}^{r} e^{\gamma_k \psi_i(x)}}. \tag{14}$$

Remark 4.3. (i) Under assumption (A2.1), [10, Proposition 2.3.12] implies that the function ψ defined in (7) is locally Lipschitz on $C+\rho B$, and for all $x\in(C+\rho B)$ we have

$$\partial\psi(x)=\Big\{\sum_{i\in\mathcal{J}_x}\lambda_i\nabla\psi_i(x):\lambda_i\geq 0\ \text{ and }\sum_{i\in\mathcal{J}_x}\lambda_i=1\Big\},\tag{15}$$

where $\mathcal{J}_x:=\{i\in\{1,\dots,r\}:\psi_i(x)=\psi(x)\}$, which is equal to $\mathcal{I}^0_x$ when $\psi(x)=0$. Hence, for $\bar{M}_\psi$ the constant defined before Proposition 4.1, we have

$$\|\zeta\|\leq\bar{M}_\psi,\ \ \forall c\in C\ \text{ and }\forall\zeta\in\partial\psi(c).$$

(ii) Assumption (A2.2) can be rephrased by means of (i) in terms of ψ as follows: There is a constant $\eta>0$ such that

$$\|\zeta\|>2\eta,\ \ \forall c\in\{x\in\mathbb{R}^n:\psi(x)=0\}\ \text{ and }\ \forall\zeta\in\partial\psi(c).$$

4.2. Properties of $(\psi_{\gamma_k})_k$ and $(C^{\gamma_k}(k))_k$

The goal of this subsection is to provide important properties of the functions ψ_{γ_k} and the sets $C^{\gamma_k}(k)$ that play an essential role in the construction of the approximating problems $(\tilde{P}^{\alpha,\beta}_{\gamma_k})$. The proofs of these properties are postponed to Section 6.

For the sequence of functions $(\psi_{\gamma_k})_k$, we have the following.

Proposition 4.4. (Properties of $(\psi_{\gamma_k})_k$)
Under (A2.1)–(A2.2), *the following assertions hold:*

(i) *The sequence $(\psi_{\gamma_k})_k\in\mathcal{C}^{1,1}$ on $C+\rho B$, is monotonically nonincreasing in terms of k, and converges uniformly to ψ. Moreover, for all $k\in\mathbb{N}$, we have*

$$\psi(x)\leq\psi_{\gamma_k}(x)\leq\psi(x)+\frac{\ln(r)}{\gamma_k},\ \ \forall x\in\mathbb{R}^n,\tag{16}$$

and

$$\|\nabla\psi_{\gamma_k}(x)\|\leq\bar{M}_\psi,\ \ \forall x\in C.\tag{17}$$

(ii) *There exist $k_1\in\mathbb{N}$ and $r_1>0$ satisfying $2r_1\leq\rho$, such that for all $k\geq k_1$, for all $x\in\{x\in\mathbb{R}^n:\psi_{\gamma_k}(x)=0\}$, and for all $z\in B_{2r_1}(x)$, we have*

$$\|\nabla\psi_{\gamma_k}(z)\|>2\eta.$$

In particular, for $k\geq k_1$ we have

$$[\psi_{\gamma_k}(x)=0]\Longrightarrow\|\nabla\psi_{\gamma_k}(x)\|>2\eta.\tag{18}$$

(iii) *There exists $k_2\geq k_1$ and $\varepsilon_o>0$ such that for all $k\geq k_2$ we have*

$$\big[x\in C^{\gamma_k}\text{ and }\|\nabla\psi_{\gamma_k}(x)\|\leq\eta\big]\ \Longrightarrow\ \psi_{\gamma_k}(x)<-\varepsilon_o.$$

Remark 4.5. From (16) and the definition of C^{γ_k} in (9), we conclude that $C^{\gamma_k}\subset C$ for all k. On the other hand, it is easy to prove that under the assumptions (A2.1)–(A2.2), and using (9), there exists $k_3\geq k_2$ such that for all $k\geq k_3$, C^{γ_k} is nonempty and compact, with $C^{\gamma_k}\subset\operatorname{int}C$ for $r>1$. Hence, since by Proposition 4.4, ψ_{γ_k} is $C^{1,1}$ on $C^{\gamma_k}+\rho B$ and satisfies (18), we deduce that all the properties satisfied by the set C in [32, Lemma 3.3 & 3.4] are also satisfied by C^{γ_k} for all $k\geq k_3$.

We proceed to present the properties of the sequence of sets $(C^{\gamma_k}(k))_k$. We denote by $2M_{\psi_{\gamma_k}}$ the Lipschitz constant of $\nabla\psi_{\gamma_k}$ over the compact set $C + \frac{\rho}{2}\overline{B}$ satisfying $M_{\psi_{\gamma_k}} \geq \frac{4\eta}{\rho}$.

Proposition 4.6. (Properties of $(C^{\gamma_k}(k))_k$)
Under (A2.1)–(A2.2), *the following assertions hold*:

(i) *For all k, the set $C^{\gamma_k}(k) \subset \operatorname{int} C^{\gamma_k} \subset \operatorname{int} C$ and is compact. Moreover, there exists $k_4 \geq k_3$ such that for $k \geq k_4$, we have*:

  (a) $\operatorname{bdry} C^{\gamma_k}(k) = \{x \in \mathbb{R}^n : \psi_{\gamma_k}(x) = -\alpha_k\} = \left\{x \in \mathbb{R}^n : \sum_{i=1}^{r} e^{\gamma_k \psi_i(x)} = \frac{2\bar{M}}{\eta\gamma_k}\right\} \neq \emptyset.$

  (b) $\operatorname{int} C^{\gamma_k}(k) = \{x \in \mathbb{R}^n : \psi_{\gamma_k}(x) < -\alpha_k\} \neq \emptyset.$

  (c) *$C^{\gamma_k}(k)$ is amenable, epi-Lipschitz with $C^{\gamma_k}(k) = \operatorname{cl}(\operatorname{int} C^{\gamma_k}(k))$, and $\frac{\eta}{2M_{\psi_{\gamma_k}}}$-prox-regular.*

  (d) *For all $x \in \operatorname{bdry} C^{\gamma_k}(k)$ we have*
$$C^{\gamma_k}(k)(x) = N^P_{C^{\gamma_k}(k)}(x) = N^L_{C^{\gamma_k}(k)}(x) = \left\{\lambda \sum_{i=1}^{r} e^{\gamma_k \psi_i(x)} \nabla\psi_i(x) : \lambda \geq 0\right\}.$$

(ii) *The sequence $(C^{\gamma_k}(k))_k$ is a nondecreasing sequence whose Painlevé-Kuratowski limit is C and satisfies*
$$\operatorname{int} C = \bigcup_{k\in\mathbb{N}} \operatorname{int} C^{\gamma_k}(k) = \bigcup_{k\in\mathbb{N}} C^{\gamma_k}(k). \tag{19}$$

(iii) *For $c \in \operatorname{bdry} C$, there exist $\bar{k}_c \geq k_4$, $r_c > 0$, and a vector $d_c \neq 0$ such that*
$$\left(\left[C \cap \overline{B}_{r_c}(c)\right] + \sigma_k \frac{d_c}{\|d_c\|}\right) \subset \operatorname{int} C^{\gamma_k}(k), \quad \forall k \geq \bar{k}_c.$$

In particular, for $k \geq \bar{k}_c$ we have
$$\left(c + \sigma_k \frac{d_c}{\|d_c\|}\right) \in \operatorname{int} C^{\gamma_k}(k). \tag{20}$$

Remark 4.7. From Proposition 4.6(iii), we deduce that for any $c \in C$, there exists a sequence $(c_{\gamma_k})_k$ such that, for k large enough, $c_{\gamma_k} \in \operatorname{int} C^{\gamma_k}(k) \subset \operatorname{int} C^{\gamma_k}$, and $c_{\gamma_k} \longrightarrow c$. Indeed:

(i) For $c \in \operatorname{bdry} C$, we choose $c_{\gamma_k} := c + \sigma_k \frac{d_c}{\|d_c\|}$ for all k. For $k \geq \bar{k}_c$, we have from (20) that $c_{\gamma_k} \in \operatorname{int} C^{\gamma_k}(k)$. Moreover, since $\sigma_k \longrightarrow 0$ and $\frac{d_c}{\|d_c\|}$ is a unit vector, we have $c_{\gamma_k} \longrightarrow c$.

(ii) For $c \in \operatorname{int} C$, equation (19) yields the existence of some $\hat{k}_c \in \mathbb{N}$, such that $c \in \operatorname{int} C^{\gamma_k}(k)$ for all $k \geq \hat{k}_c$. Hence, there exists $\hat{r}_c > 0$ satisfying
$$c \in \overline{B}_{\hat{r}_c}(c) \subset \operatorname{int} C^{\gamma_k}(k), \quad \forall k \geq \hat{k}_c.$$

In this case, we take the sequence $c_{\gamma_k} \equiv c \in \operatorname{int} C^{\gamma_k}(k)$ that converges to c.

Remark 4.8. When combining Proposition 4.4(ii) and Proposition 4.6(ii), it follows that for k large enough, we have
$$[\psi_{\gamma_k}(x) = -\alpha_k] \Longrightarrow \|\nabla\psi_{\gamma_k}(x)\| > \eta.$$

4.3. Connection between (D) and (D_{γ_k}) – Existence of solutions

We begin this subsection with the following lemma that states that the unique solution x_{γ_k} of the Cauchy problem corresponding to the dynamic (D_{γ_k}) with initial condition in C^{γ_k} remains in the set C^{γ_k} for all t, and the sequence $(x_{\gamma_k})_k$ is equicontinuous. Note that since the end-time T was taken to be 1 in [32], when necessary, we shall adjust here the constants' expressions in the results and proofs extracted from [32].

We introduce the following *global* version of (A1):

(A1)$_{\mathbb{G}}$ For fixed $(x,u) \in C \times \mathbb{U}$, $f(\cdot,x,u)$ is Lebesgue-measurable; and there exists $M > 0$ such that for a.e. $t \in [0,T]$ we have: $(x,u) \mapsto f(t,x,u)$ is continuous on $C \times U(t)$; for all $u \in U(t)$, $x \mapsto f(t,x,u)$ is M-Lipschitz on C; and $\|f(t,x,u)\| \le M$ for all $(x,u) \in C \times U(t)$.

Lemma 4.9. (Invariance of C^{γ_k} and uniform bounds) *Let* (A1)$_{\mathbb{G}}$, (A2.1)–(A2.2) *and* (A2.5) *be satisfied. Then for $k \ge k_3$, the system (D_{γ_k}) with $x(0) = c_{\gamma_k} \in C^{\gamma_k}$ and $u_{\gamma_k} \in \mathcal{U}$, has a unique solution x_{γ_k} which belongs to $W^{1,2}([0,T];\mathbb{R}^n)$ and satisfies $x_{\gamma_k}(t) \in C^{\gamma_k}$ for all $t \in [0,T]$. Furthermore, the sequence $(x_{\gamma_k})_k$ satisfies*

$$\|x_{\gamma_k}\|_\infty \le \|c_{\gamma_k}\| + \sqrt{\left(\bar{M}T\right)^2 + 2T} \quad \text{and} \quad \int_0^T \|\dot{x}_{\gamma_k}(t)\|^2 dt \le \bar{M}^2 T + 2,$$

and thus, it is equicontinuous.

Proof. By Proposition 4.4, we have that for each $k \ge k_3$, ψ_{γ_k} satisfies the *same assumptions* satisfied by the function ψ of [32, Lemma 4.1]. Moreover, the system (D_{γ_k}) utilized in [32, Lemma 4.1], in which ψ is replaced by ψ_{γ_k}, coincides with our system (D_{γ_k}), see (13). Therefore, Lemma 4.9 follows immediately from [32, Lemma 4.1], where ψ and C are replaced by ψ_{γ_k} and C^{γ_k}, respectively. We note that in [32, Lemma 4.1], $\|c_{\gamma_k}\|$ is replaced by a bound $\alpha_0 > 0$ of the sequence $(c_{\gamma_k})_k$. □

We denote by $\xi_{\gamma_k}(\cdot)$ the sequence of non-negative continuous functions on $[0,T]$ corresponding to the solution x_{γ_k}, obtained in Lemma 4.9, and defined by

$$\xi_{\gamma_k}(\cdot) := \gamma_k e^{\gamma_k \psi_{\gamma_k}(x_{\gamma_k}(\cdot))} \overset{(8)}{=} \sum_{i=1}^{r} \xi^i_{\gamma_k}(\cdot), \tag{21}$$

where $\xi^i_{\gamma_k}(\cdot) := \gamma_k e^{\gamma_k \psi_i(x_{\gamma_k}(\cdot))}$, $i = 1, \dots, r$.

The following result illustrates the tight link between the solutions of (D_{γ_k}) and those of (D). It is parallel to [32, Theorem 4.1], and it follows using arguments similar to those used in the proofs of [32, Theorem 4.1] and [40, Theorem 4.1]. See also the proof of [20, Theorem 2.2] where a simplified method is used for special setting.

Theorem 4.10. (Convergence of $(\xi_{\gamma_k})_k$ – $(D_{\gamma_k})_k$ approximates (D)) *Assume that* (A1)$_{\mathbb{G}}$, (A2.1)–(A2.2) *and* (A2.5)–(A4.1) *hold. Then we have the following:*

(I). *Let be given a sequence $(c_{\gamma_k})_k$ in C^{γ_k} converging to $x_0 \in C_0$ and a sequence $(u_{\gamma_k})_k$ in $\mathcal{U}$. For $k \ge k_3$, denote by x_{γ_k} the solution of (D_{γ_k}) corresponding to $(x(0),u) = (c_{\gamma_k}, u_{\gamma_k})$ which is obtained via* Lemma 4.9 *and hence, remains in C^{γ_k}.*

Then, the following statements are valid:

(i) *The sequence* $(x_{\gamma_k})_k$ *admits a subsequence, we do not relabel, that converges uniformly to some* $x \in W^{1,2}([0,T];\mathbb{R}^n)$ *whose values are in* C*, and its derivative* $\dot{x}_{\gamma_k}$ *converges weakly in* L^2 *to* $\dot{x}$*. Moreover,*

$$\|x\|_\infty \le \|x_0\| + \sqrt{\left(\bar{M}T\right)^2 + 2T} \quad \text{and} \quad \int_0^T \|\dot{x}(t)\|^2 dt \le \bar{M}^2 T + 2. \tag{22}$$

(ii) *The sequence* $(\|\xi_{\gamma_k}\|_2^2)_k$ *is bounded by a positive constant* M_ξ*, and there exists a subsequence of* $(\gamma_k)_k$*, we do not relabel, such that the sequence of vector functions* $(\xi^1_{\gamma_k}, \dots, \xi^r_{\gamma_k})_k$ *converges weakly in* L^2 *to a nonnegative vector function* $(\xi^1, \dots, \xi^r) \in L^2([0,T],\mathbb{R}^r)$*, and hence,* ξ_{γ_k} *converges weakly in* L^2 *to* $\xi := \sum_{i=1}^r \xi^i$*, with*

$$\xi^i(t) = 0 \ \text{ for all } t \in I_i(x) \text{ and } i \in \{1, \dots r\}, \tag{23}$$

and
$$\xi(t) = 0 \ \text{ for all } t \in I(x). \tag{24}$$

(iii) *If the subsequence* $u_{\gamma_k}(t) \xrightarrow{\text{a.e. } t} u(t)$*, where* $u \in \mathcal{U}$*, then* x *is the unique solution of* (D) *corresponding to* (x_0, u) *and the triplet* $(x, u, (\xi^i)_{i=1}^r)$ *satisfies equations* (25)–(28) *stated below.*

(iv) *If* (A4.3) *holds and* $f(t,x,U(t))$ *is convex for* $x \in C$ *and* $t \in [0,T]$ a.e.*, then there exists* $u \in \mathcal{U}$ *such that the limit function* x *is the unique solution of* (D) *corresponding to* (x_0, u)*. Moreover, we have*

$$\dot{x}(t) = f_\Phi(t, x(t), u(t)) - \sum_{i=1}^r \xi^i(t) \nabla \psi_i(x(t)), \quad t \in [0,T] \text{ a.e.} \tag{25}$$

$$\|\xi^i\|_\infty \le \frac{\bar{M}}{2\eta} \ \text{ for } i = 1, \dots, r, \quad \text{and} \quad \|\xi\|_\infty \le \frac{\bar{M}}{2\eta}. \tag{26}$$

$$\|f_\Phi(t, x(t), u(t)) - \dot{x}(t)\| \le \|f_\Phi(t, x(t), u(t))\|, \quad t \in [0,T] \text{ a.e.} \tag{27}$$

$$\langle f_\Phi(t, x(t), u(t)) - \dot{x}(t), z - x(t) \rangle \le \frac{\bar{M} M_\psi}{2\eta} \|z - x(t)\|^2, \ t \in [0,T] \text{ a.e.}, \forall z \in C. \tag{28}$$

(II). *Conversely, for given* $x_0 \in C_0$ *and* $\bar{u} \in \mathcal{U}$*, the system* (D) *satisfying the condition* $(x(0), u) = (x_0, \bar{u})$ *has a unique solution* $\bar{x}$*, which is the uniform limit of a subsequence of* $(x_{\gamma_k})_k$ *that is not relabeled, where for each* k*,* x_{γ_k} *is the solution of the system* (D_{γ_k}) *corresponding to* $(x(0), u) = (c_{\gamma_k}, \bar{u})$*, and* c_{γ_k} *is the sequence of* Remark 4.7 *that corresponds to* $c = x_0$*. Hence, for* k *sufficiently large,* $x_{\gamma_k}(t) \in C^{\gamma_k}$ *for all* $t \in [0,T]$*. Moreover, for* ξ^i *the* L^2*-weak limit of* $\xi^i_{\gamma_k}$ *obtained in* (ii)*, for* $i = 1, \dots, r$*, the triplet* $(\bar{x}, \bar{u}, (\xi^i)_{i=1}^r)$ *satisfies equations* (22)–(28)*.*

Remark 4.11. Let $u \in \mathcal{U}$ and $x \in W^{1,1}([0,T];\mathbb{R}^n)$ with $x(0) \in C_0$ and $x(t) \in C$ for all $t \in [0,T]$. From the statement and the proof of Theorem 4.10, and from (15), we deduce that under (A1)–(A3), the following assertions are equivalent:

(i) x is a solution for (D) corresponding to the control u.

(ii) There exists a finite sequence of nonnegative measurable functions $(\xi^i)_{i=1}^r$ such that for $i = 1, \dots, r$, $\xi^i(t) = 0$ for all $t \in I_i(x)$, $\|\sum_{i=1}^r \xi^i\|_\infty \le \frac{\bar{M}}{2\eta}$, and $(x, u, (\xi^i)_{i=1}^r)$ satisfies

$$\begin{aligned}\dot{x}(t) &= f_{\Phi}(t, x(t), u(t)) - \sum_{i=1}^{r} \xi^i(t) \nabla \psi_i(x(t)) \\ &= f_{\Phi}(t, x(t), u(t)) - \sum_{i \in \mathcal{I}^0_{x(t)}} \xi^i(t) \nabla \psi_i(x(t)), \quad t \in [0,T] \text{ a.e.}\end{aligned}$$

(iii) There exists nonnegative measurable function ξ such that $\xi(t) = 0$ for all $t \in I(x)$ and

$$\dot{x}(t) \in f_{\Phi}(t, x(t), u(t)) - \xi(t) \partial \psi(x(t)), \quad t \in [0,T] \text{ a.e.}$$

The implication $(iii) \implies (ii)$, is established by taking $\xi_i(t) := \xi(t)\lambda_i(x(t))$ for $i \in \mathcal{I}^0_{x(t)}$, where $(\lambda_i(x(t)))_{i \in \mathcal{I}^0_{x(t)}}$ are the coefficients of the convex combination of $\nabla \psi_i(x(t))$'s obtained via (15) for the element in (iii) belonging to $\partial \psi(x(t))$.

Remark 4.12. Note that when establishing Theorem 4.10(I)(iii)&(iv), the proof that $(x, u, (\xi^i)_{i=1}^r)$ satisfies (25) uses arguments independent of $\xi^i_{\gamma_k}$ being obtained through (21), and hence, that proof is valid for $(\xi^1_{\gamma_k}, \ldots, \xi^r_{\gamma_k})_k$ being any sequence of $L^2([0,T]; \mathbb{R}^r)$-functions that converges weakly in L^2 to $(\xi^1, \ldots, \xi^r)$, as $k \longrightarrow \infty$. Therefore, whenever $(x_j, u_j, (\xi^i_j)_{i=1}^r)_j$ is a sequence solving (25) with x_j converging uniformly to x and $(\xi^1_j, \ldots, \xi^r_j)$ converging weakly in L^2 to $(\xi^1, \ldots, \xi^r)$, we have that $(x, u, (\xi^i)_{i=1}^r)$ satisfies (25) for some $u \in \mathcal{U}$. This function u is the almost everywhere pointwise limit of $(u_j)_j$, whenever this limit exists. Otherwise, u is obtained through the Filippov selection theorem, by assuming that (A4.3) holds and $f(t, x, U(t))$ is convex for $x \in C$ and $t \in [0,T]$ a.e.

The following theorem presents important extra features resulting from starting the solutions of (D_{γ_k}) in Theorem 4.10(I) from the subset $C^{\gamma_k}(k) \subset \operatorname{int} C^{\gamma_k} \subset \operatorname{int} C$ defined in (10).

Theorem 4.13. (Extra properties of $(x_{\gamma_k})_k$ and $(\xi_{\gamma_k})_k$ when $x_{\gamma_k}(0) \in C^{\gamma_k}(k)$) *Assume that* $(\mathrm{A1})_{\mathbb{G}}$, (A2.1)–(A2.2) *and* (A2.5)–(A4.1) *hold. Let $(c_{\gamma_k})_k$ be a sequence such that $c_{\gamma_k} \in C^{\gamma_k}(k)$, for k sufficiently large. Then there exists $k_5 \geq k_4$ such that for all sequences $(u_{\gamma_k})_k$ in $\mathfrak{U}$ and for all $k \geq k_5$, the solution x_{γ_k} of (D_{γ_k}) corresponding to $(c_{\gamma_k}, u_{\gamma_k})$ satisfies*:

(i) $x_{\gamma_k}(t) \in C^{\gamma_k}(k) \subset \operatorname{int} C^{\gamma_k} \subset \operatorname{int} C$ *for all* $t \in [0,T]$.

(ii) $0 \leq \xi^i_{\gamma_k}(t) \leq \xi_{\gamma_k}(t) \leq \frac{2\bar{M}}{\eta}$ *for all* $t \in [0,T]$ *and for* $i = 1, \ldots, r$.

(iii) $\|\dot{x}_{\gamma_k}(t)\| \leq \bar{M} + \frac{2\bar{M}\bar{M}_{\psi}}{\eta}$ *for* a.e. $t \in [0,T]$.

Proof. By Proposition 4.4 and Remark 4.8, we have that, for each $k \geq k_3$, the function ψ_{γ_k} satisfies the *same assumptions* imposed on the function ψ in [32, Theorem 5.1], where the system (D_{γ_k}), in which ψ is replaced by ψ_{γ_k}, coincides with our system (D_{γ_k}). Thus, Theorem 4.13(i)–(iii) follows immediately from [32, Theorem 5.1(i)–(iii)], where ψ and C are replaced by ψ_{γ_k} and C^{γ_k}, respectively. □

As a direct consequence of Theorem 4.13 and parallel to [32, Corollary 5.2], we have the following.

Corollary 4.14. ($\bar{x}_{\gamma_k}$, corresponding to $\bar{x}$, in $C^{\gamma_k}(k)$) *Assume that the conditions* $(A1)_{\mathbb{G}}$*,* (A2.1)–(A2.2)*, and* (A2.5)–(A4.1) *hold. Let* $(\bar{x}, \bar{u})$ *be a solution of* (D)*. Consider* $(\bar{c}_{\gamma_k})_k$ *the sequence obtained from Remark 4.7 for* $c := \bar{x}(0)$*, and* $\bar{x}_{\gamma_k}$ *the solution of* (D_{γ_k}) *corresponding to* $(\bar{c}_{\gamma_k}, \bar{u})$*. Then, there exists* $k_7 \geq k_6$ *such that for all* $k \geq k_7$ *we have:*

(i) $\bar{x}_{\gamma_k}(t) \in C^{\gamma_k}(k) \subset \operatorname{int} C^{\gamma_k} \subset \operatorname{int} C$ *for all* $t \in [0, T]$*.*

(ii) $0 \leq \xi^i_{\gamma_k}(t) \leq \xi_{\gamma_k}(t) \leq \frac{2\bar{M}}{\eta}$ *for all* $t \in [0, T]$ *and for* $i = 1, \ldots, r$*.*

(iii) $\|\dot{\bar{x}}_{\gamma_k}(t)\| \leq \bar{M} + \frac{2\bar{M}\bar{M}_\psi}{\eta}$ *for* a.e. $t \in [0, T]$*.*

4.4. Existence of solutions for (P)

The following is an existence of optimal solution result for the problem (P).

Proposition 4.15. (Existence of optimal solution for (P)) *Assume that* $(A1)_{\mathbb{G}}$*,* (A2.1)–(A2.2) *and* (A2.5)–(A4) *are satisfied,* $g\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *is lower semicontinuous, and* $f(t, x, U(t))$ *is convex for all* $x \in C$ *and* $t \in [0, T]$ a.e. *Then the problem* (P) *has an optimal solution if and only if it has at least one admissible pair* (z_o, u_o) *with* $(z_o(0), z_o(T)) \in \operatorname{dom} g$*.*

Proof. Since any admissible pair (x, u) satisfies $(x(0), x(T)) \in C_0 \times (C_T \cap C)$, then, by the lower semicontinuity of g, the compactness of $C_0 \times (C_T \cap C)$, and the existence of the admissible pair (z_o, u_o) with $(z_o(0), z_o(T)) \in (\operatorname{dom} g) \cap [C_0 \times (C_T \cap C)]$, we have that the $\inf_{(x,u)}(P)$ is finite, and hence, (P) admits a minimizing sequence (x_n, u_n). It follows that x_n satisfies (22) with $x_n(t) \in C$ for all $t \in [0, T]$. Thus, Arzela-Ascoli's theorem produces a subsequence of $(x_n)_n$, we do not relabel, that converges uniformly to an absolutely continuous $\bar{x}$, with $\bar{x}(t) \in C$ for all $t \in [0, T]$. The equivalence between (i) and (ii) in Remark 4.11 produces a sequence of r-tuple nonnegative measurable functions $(\xi^1_n, \ldots, \xi^r_n)_n$ such that, for $i = 1, \ldots, r$, $\xi^i_n(t) = 0$ for all $t \in I_i(x)$ and $\|\xi^i_n\|_\infty \leq \frac{\bar{M}}{2\eta}$, and $(x_n, u_n, (\xi^i_n)^r_{i=1})$ satisfies (25).
Consequently, the sequence $(\xi^1_n, \ldots, \xi^r_n)_n$ admits a subsequence, we do not relabel, that converges weakly in L^2 to $(\xi^1, \ldots, \xi^r)$. Remark 4.12 now asserts the existence of a $\bar{u} \in \mathcal{U}$ such that the pair $(\bar{x}, \bar{u})$ solves (D). Since C_T is closed and $x_n(1) \in C_T$ for all $n \in \mathbb{N}$, it follows that $\bar{x}(T) \in C_T$, and hence, $(\bar{x}, \bar{u})$ is admissible for (P). On the other hand, the lower semicontinuity of g implies

$$g(\bar{x}(0), \bar{x}(T)) \leq \liminf_{n \to \infty} g(x_n(0), x_n(T)) = \inf_{(x,u)}(P).$$

This proves that $(\bar{x}, \bar{u})$ is an optimal solution for (P). □

4.5. Approximating problems for (P)

Let $(\bar{x}, \bar{u})$ is a *strong local minimizer* for (P) with a corresponding $\delta > 0$. The ultimate goal here is to formulate a suitable sequence of *standard* optimal control problems that approximates the problem (P) and whose optimal solutions form a sequence that approximates $(\bar{x}, \bar{u})$. The following lemma states that when (A1) and (A6) are satisfied, the function f can be extended to a function $\tilde{f}$ defined on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^m$ and satisfying $(A1)_{\mathbb{G}}$ with $\tilde{f}(t, x, U(t))$ convex for all $x \in C$ and

$t \in [0,T]$ a.e. The proof of this lemma is not displayed here, as it mimics arguments used at the beginning of [32, Proof of Theorem 6.2], where such an extension is obtained for the case in which $U(\cdot)$ is a *constant* multifunction.

Lemma 4.16. (Extension of f) *Assume that* (A1)–(A2.2) *hold. Then there exists a function* $\tilde{f}\colon [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$ *such that:*

- *For* $t \in [0,T]$ a.e., *we have*

$$\tilde{f}(t,x,u) = f(t,x,u), \;\; \forall (x,u) \in (C \cap \overline{B}_\delta(\bar{x}(t))) \times U(t).$$

- *The function* $\tilde{f}$ *satisfies* $(\mathrm{A1})_{\mathbb{G}}$ *for* $M := 2M_\ell$.
- *If also* (A6) *is satisfied, then* $\tilde{f}(t,x,U(t))$ *is convex for all* $x \in C$ *and* $t \in [0,T]$ a.e.

Now let $\delta_o > 0$ be fixed such that

$$\delta_o \le \begin{cases} \min\left\{\hat{r}_{\bar{x}(0)}, \frac{\delta}{2}\right\} & \text{if } \bar{x}(0) \in \operatorname{int} C, \\ \min\left\{r_{\bar{x}(0)}, \frac{\delta}{2}\right\} & \text{if } \bar{x}(0) \in \operatorname{bdry} C, \end{cases}$$

where $\hat{r}_{\bar{x}(0)}$ and $r_{\bar{x}(0)}$ are, respectively, the constants in Remark 4.7(ii) and Proposition 4.6(iii) corresponding to $c := \bar{x}(0)$. We introduce the following notations:

- The trajectory $\bar{x}_{\gamma_k}$ is the solution of (D_{γ_k}) obtained via Corollary 4.14. That is, it corresponds to $(\bar{c}_{\gamma_k}, \bar{u})$, where $\bar{c}_{\gamma_k}$ is defined via Remark 4.7 for $c := \bar{x}(0)$, and hence, $\bar{c}_{\gamma_k} \in C^{\gamma_k}(k)$ for k sufficiently large.
- The sequence of sets $(C_0^{\gamma_k}(k))_k$, is defined by

$$C_0^{\gamma_k}(k) := \begin{cases} C_0 \cap \overline{B}_{\delta_o}(\bar{x}(0))\,, \;\forall k \in \mathbb{N}, & \text{if } \bar{x}(0) \in \operatorname{int} C, \\ \left[C_0 \cap \overline{B}_{\delta_o}(\bar{x}(0))\right] + \sigma_k \frac{d_{\bar{x}(0)}}{\|d_{\bar{x}(0)}\|}, \;\forall k \in \mathbb{N}, & \text{if } \bar{x}(0) \in \operatorname{bdry} C. \end{cases}$$

- The sequence of sets $(C_T(k))_k$ is defined by

$$C_T(k) := \left[\left(C_T \cap \overline{B}_{\delta_o}(\bar{x}(T))\right) - \bar{x}(T) + \bar{x}_{\gamma_k}(T)\right] \cap C, \quad k \in \mathbb{N}.$$

- The function $L\colon [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}$ is defined by

$$L(t,x) := \max\left\{\|x - \bar{x}(t)\|^2 - \frac{\delta^2}{4}, 0\right\}, \;\; \forall (t,x) \in [0,T] \times C. \tag{29}$$

  It is easy to see that $L(\cdot,x)$ is measurable, and $L(t,\cdot)$ is Lipschitz on C uniformly in $t \in [0,T]$.

A well known optimal control technique (see e.g., [39]) is used to show that a strong local minimizer of an optimal control problem is a global minimizer for another problem, obtained by adding to the objective function the integral of the function L, defined in (29), and by localizing the endpoints constraints. This technique is also employed in [32, Proof of Theorem 6.2] to prove the following lemma, which states that $(\bar{x}, \bar{u})$ is a *global* minimizer for the new problem $(\tilde{P})$.

Lemma 4.17. ($(\bar{x},\bar{u})$ global solution for $(\tilde{P})$) *Assume that* (A1)–(A2.2) *hold, and let $\tilde{f}$ be the extension of f obtained in* Lemma 4.16*. Assume further that* (A4.1)–(A4.2) *are satisfied and g is continuous on $\tilde{C}_0(\delta)\times\tilde{C}_T(\delta)$ for some $\tilde{\rho}>0$. Then, $(\bar{x},\bar{u})$ is a global minimizer for the problem*

$$(\tilde{P})\quad \text{Minimize } g(x(0),x(T))+\tilde{K}\int_0^T L(t,x(t))dt \text{ over } (x,u) \text{ such that } u(\cdot)\in\mathcal{U} \text{ and}$$

$$\begin{cases} (\tilde{D})\ \begin{bmatrix} \dot{x}(t)\in f(t,x(t),u(t))-\partial\varphi(x(t)), \text{ a.e. } t\in[0,T], \\ x(0)\in C_0\cap \overline{B}_{\delta_o}(\bar{x}(0)), \end{bmatrix} \\ x(T)\in C_T\cap\overline{B}_{\delta_o}(\bar{x}(T)), \end{cases}$$

where $\tilde{K}:=\frac{512\,\bar{M}_\ell M_g}{5\,\delta^3}$, $\bar{M}_\ell:=2M_\ell+K$ *and* $M_g:=\max\limits_{\tilde{C}_0(\delta)\times\tilde{C}_T(\delta)} g(x,y)$.

Now we are ready to state our approximation result which follows using arguments similar to those used in the proof of [32, Proposition 6.2].

Proposition 4.18. (Approximating problems for (P)) *Assume that* (A1)–(A2.2), (A2.5)–(A4) *hold, g is continuous on $\tilde{C}_0(\delta)\times\tilde{C}_T(\delta)$ for some $\tilde{\rho}>0$, and $f(t,x,U(t))$ is convex for all $x\in C$ and $t\in[0,T]$* a.e. *Let $\tilde{f}$ be the extension of f in* Lemma 4.16*, and $\tilde{K}$ be the constant in* Lemma 4.17*. Then, for every $\alpha>0$ and $\beta\in(0,1]$, there exist a subsequence of $(\gamma_k)_k$, we do not relabel, and a sequence $(c_{\gamma_k},e_{\gamma_k},u_{\gamma_k})_k$ in $C_0^{\gamma_k}(k)\times C_T(k)\times\mathcal{U}$ such that, for:*

$$\tilde{f}^\beta(t,x,u):=(1-\beta)\tilde{f}(t,x,\bar{u}(t))+\beta\tilde{f}(t,x,u),\ \forall t\in[0,T] \text{ a.e.},\ \forall(x,u)\in\mathbb{R}^n\times U,$$

$$\tilde{f}^\beta_\Phi(t,x,u):=\tilde{f}^\beta(t,x,u)-\nabla\Phi(x),\ \forall t\in[0,T] \text{ a.e.},\ \forall(x,u)\in\mathbb{R}^n\times U,$$

$$\begin{aligned} J(x,u) &:= g(x(0),x(T))+\tilde{K}\int_0^T L(t,x(t))dt \\ &\qquad +\alpha\left(\|u(t)-u_{\gamma_k}(t)\|_1+\|x(0)-c_{\gamma_k}\|+\|x(T)-e_{\gamma_k}\|\right), \end{aligned}$$

and

$$(\tilde{P}^{\alpha,\beta}_{\gamma_k})\quad \text{Minimize } J(x,u) \text{ over } (x,u) \text{ such that } u(\cdot)\in\mathcal{U} \text{ and}$$

$$\begin{cases} (\tilde{D}^\beta_{\gamma_k})\ \begin{bmatrix} \dot{x}(t)=\tilde{f}^\beta_\Phi(t,x(t),u(t))-\sum\limits_{i=1}^r \gamma_k e^{\gamma_k\psi_i(x(t))}\nabla\psi_i(x(t)), \text{ a.e. } t\in[0,T], \\ x(0)\in C_0^{\gamma_k}(k), \end{bmatrix} \\ x(T)\in C_T(k), \end{cases}$$

the pair $(x_{\gamma_k},u_{\gamma_k})$ that solves $(\tilde{D}^\beta_{\gamma_k})$ for $x_{\gamma_k}(0)=c_{\gamma_k}$, is optimal for $(\tilde{P}^{\alpha,\beta}_{\gamma_k})$ and satisfies $x_{\gamma_k}(T)=e_{\gamma_k}$. Moreover, we have

$$u_{\gamma_k}\xrightarrow[L^1]{\text{strongly}}\bar{u},\quad x_{\gamma_k}\xrightarrow{\text{uniformly}}\bar{x},$$

and all the conclusions of Theorem 4.13 *are valid, including $(x_{\gamma_k})_k$ is uniformly Lipschitz and $x_{\gamma_k}(t)\in C^{\gamma_k}(k)\subset\operatorname{int}C^{\gamma_k}\subset\operatorname{int}C$ for all $t\in[0,T]$. Furthermore, for all k sufficiently large, we have*

$$x_{\gamma_k}(0)\in(C_0+\tilde{\rho}B)\cap\operatorname{int}C \quad\textit{and}\quad x_{\gamma_k}(T)\in(C_T+\tilde{\rho}B)\cap\operatorname{int}C.$$

The next proposition is obtained as a direct application of the *nonsmooth* Pontryagin maximum principle (e.g., [39, Theorem 6.2.1]) to each member of the family of approximating problems $(\tilde{P}^{\alpha,\beta}_{\gamma_k})$, obtained in Proposition 4.18. Note that the function $\tilde{f}$ is replaced by f in the statement of the proposition, since we have the following:

- $\tilde{f}(t,x,u) = f(t,x,u)$, for $t \in [0,T]$ a.e. and for all $(x,u) \in (C \cap \overline{B}_\delta(\bar{x}(t))) \times U(t)$.
- The sequence $(x_{\gamma_k})_k$ of Proposition 4.18 belongs to $\operatorname{int} C$ and converges uniformly to $\bar{x}$. Hence, for k large enough,
$$x_{\gamma_k}(t) \in (\operatorname{int} C) \cap B_{\delta_o}(\bar{x}(t)) \subset \operatorname{int}(C \cap \overline{B}_\delta(\bar{x}(t))) \text{ for all } t \in [0,T].$$

Proposition 4.19. (Maximum Principle for the approximating problems)
Let $\alpha > 0$ and $\beta \in (0,1]$ be fixed. Assume that (A1)–(A2.2) *and* (A2.5)–(A6) *hold. Let $(x_{\gamma_k}, u_{\gamma_k})$ be the sequence in* Proposition 4.18 *that is optimal for $(\tilde{P}^{\alpha,\beta}_{\gamma_k})$ with x_{γ_k} converging uniformly to x_{γ_k} and u_{γ_k} converging strongly in L^1 to $\bar{u}$. Then, for $k \in \mathbb{N}$, there exist $p_{\gamma_k} \in W^{1,1}([0,T];\mathbb{R}^n)$ and $\lambda_{\gamma_k} \geq 0$ such that:*

(i) (The nontriviality condition) *For all $k \in \mathbb{N}$, we have*
$$\|p_{\gamma_k}(T)\| + \lambda_{\gamma_k} = 1;$$

(ii) (The adjoint equation) *For $t \in [0,T]$* a.e.,
$$\begin{aligned}\dot{p}_{\gamma_k}(t) \in &-(1-\beta)(\partial^x f(t,x_{\gamma_k}(t),\bar{u}(t)))^{\mathsf{T}} p_{\gamma_k}(t) - \beta(\partial^x f(t,x_{\gamma_k}(t),u_{\gamma_k}(t)))^{\mathsf{T}} p_{\gamma_k}(t) \\ &+ \partial^2 \Phi(x_{\gamma_k}(t)) p_{\gamma_k}(t) + \sum_{i=1}^r \gamma_k e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \partial^2 \psi_i(x_{\gamma_k}(t)) p_{\gamma_k}(t) \\ &+ \sum_{i=1}^r \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \nabla \psi_i(x_{\gamma_k}(t)) \langle \nabla \psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle + \lambda_{\gamma_k} \partial^x L(t, x_{\gamma_k}(t));\end{aligned} \tag{30}$$

(iii) (The transversality equation)
$$\begin{aligned}(p_{\gamma_k}(0), -p_{\gamma_k}(T)) \in\ &\lambda_{\gamma_k} \partial^L g(x_{\gamma_k}(0), x_{\gamma_k}(T)) \\ &+ \left[\left(\alpha\overline{B} + N^L_{C_0^{\gamma_k}(k)}(x_{\gamma_k}(0))\right) \times \left(\alpha\overline{B} + N^L_{C_T(k)}(x_{\gamma_k}(T))\right)\right];\end{aligned}$$

(iv) (The maximization condition)
$$\max_{u \in U} \left\{ \langle f^\beta_\Phi(t, x_{\gamma_k}(t), u), p_{\gamma_k}(t)\rangle - \frac{\alpha\lambda_{\gamma_k}}{\beta}\|u - u_{\gamma_k}(t)\| \right\}$$
is attained at $u_{\gamma_k}(t)$ for $t \in [0,T]$ a.e.

Furthermore, if $C_T = \mathbb{R}^n$, then $\lambda_{\gamma_k} \neq 0$ and is taken to be 1, *and the nontriviality condition* (i) *is eliminated.*

5. Proof of Theorem 3.1

We first prove the theorem under the two temporary additional assumptions: (i) "C is compact" which is *stronger* than (A2.4), see Remark 2.1, and (ii) "The matrix $G_\psi(t)$ is strictly diagonally dominant for all $t \in I^0(\bar{x})$" which is denoted by $(\text{A2.3})_S$ and is *stronger* than (A2.3) since $\beta = (1,\ldots,1)$, see Remark 2.3. The removal of these additional assumptions will constitute the final steps in the proof.

Let $(\bar{x},\bar{u})$ be a strong local minimizer for (P). For each fixed pair of numbers $(\alpha,\beta)\in(0,1]^2:=(0,1]\times(0,1]$, Propositions 4.18 and 4.19 produce a subsequence of $(\gamma_k)_k$, we do not relabel, and corresponding sequences x_{γ_k}, u_{γ_k}, p_{γ_k} and λ_{γ_k} such that:

- For each k, the admissible pair $(x_{\gamma_k},u_{\gamma_k})$ is optimal for $(\tilde{P}^{\alpha,\beta}_{\gamma_k})$.
- u_{γ_k} converges strongly in L^1 to $\bar{u}$, and $u_{\gamma_k}(t)\longrightarrow\bar{u}(t)$ a.e.
- x_{γ_k} converges uniformly to $\bar{x}$.
- All the conclusions of Theorem 4.13 are valid, including $(x_{\gamma_k})_k$ is uniformlyLipschitz and $x_{\gamma_k}(t)\in\operatorname{int}C$ for all $t\in[0,T]$.
- For all k, $x_{\gamma_k}(0)\in(C_0+\tilde{\rho}B)\cap\operatorname{int}C$ and $x_{\gamma_k}(T)\in(C_T+\tilde{\rho}B)\cap\operatorname{int}C$.
- Equations (i)–(iv) of Proposition 4.19 are satisfied.

Since $x_{\gamma_k}(t)\in\operatorname{int}C$ for all $t\in[0,T]$, we have, for $t\in[0,T]$ a.e.,

$$\partial^x f(t,x_{\gamma_k}(t),u_{\gamma_k}(t))=\partial^x_\ell f(t,x_{\gamma_k}(t),u_{\gamma_k}(t)),\quad \partial^2\Phi(x_{\gamma_k}(t))=\partial^2_\ell\varphi(x_{\gamma_k}(t)),$$
$$\partial^x f(t,x_{\gamma_k}(t),\bar{u}(t))=\partial^x_\ell f(t,x_{\gamma_k}(t),\bar{u}(t)),$$
$$\partial^2\psi_i(x_{\gamma_k}(t))=\partial^2_\ell\psi_i(x_{\gamma_k}(t))\ \text{ for }\ i=1,\dots,r,$$

and

$$\partial^L g(x_{\gamma_k}(0),x_{\gamma_k}(T))=\partial^L_\ell g(x_{\gamma_k}(0),x_{\gamma_k}(T)).$$

Hence from (30), it results the existence of sequences ζ_{γ_k}, $\tilde{\zeta}_{\gamma_k}$, θ_{γ_k} and $(\vartheta^i_{\gamma_k})_{i=1}^r$ in $\mathcal{M}_n([0,T])$, and $\omega_{\gamma_k}\colon[0,T]\longrightarrow\mathbb{R}^n$ such that, for a.e. $t\in[0,T]$,

$$(\tilde{\zeta}_{\gamma_k}(t),\zeta_{\gamma_k}(t),\theta_{\gamma_k}(t))\in\partial^x_\ell f(t,x_{\gamma_k}(t),u_{\gamma_k}(t))\times\partial^x_\ell f(t,x_{\gamma_k}(t),\bar{u}(t))\times\partial^2_\ell\varphi(x_{\gamma_k}(t)),$$
$$\vartheta^i_{\gamma_k}(t))\in\partial^2_\ell\psi_i(x_{\gamma_k}(t))\ \text{ for }\ i=1,\dots,r,$$
$$\omega_{\gamma_k}(t)\in\partial^x L(t,x_{\gamma_k}(t))=\operatorname{conv}\{2(x_{\gamma_k}(t)-\bar{x}(t)),0\}=[0,2(x_{\gamma_k}(t)-\bar{x}(t))],\tag{31}$$

and
$$\begin{aligned}\dot{p}_{\gamma_k}(t)=&-(1-\beta)\zeta_{\gamma_k}(t)^\intercal p_{\gamma_k}(t)-\beta\tilde{\zeta}_{\gamma_k}(t)^\intercal p_{\gamma_k}(t)+\theta_{\gamma_k}(t)p_{\gamma_k}(t)\\&+\sum_{i=1}^r\gamma_k e^{\gamma_k\psi_i(x_{\gamma_k}(t))}\vartheta^i_{\gamma_k}(t)p_{\gamma_k}(t)\\&+\sum_{i=1}^r\gamma_k^2 e^{\gamma_k\psi_i(x_{\gamma_k}(t))}\nabla\psi_i(x_{\gamma_k}(t))\langle\nabla\psi_i(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle+\lambda_{\gamma_k}\omega_{\gamma_k}(t).\end{aligned}\tag{32}$$

Note that for each k, the functions $p_{\gamma_k},\dot{p}_{\gamma_k},x_{\gamma_k}$, u_{γ_k} and $\bar{u}$ are measurable on $[0,T]$, and, by [11, Exercise 13.24] and (A1), (A2.1), and (A3), the multifunctions $\partial^x_\ell f(\cdot,x_{\gamma_k}(\cdot),u_{\gamma_k}(\cdot))$, $\partial^x_\ell f(\cdot,x_{\gamma_k}(\cdot),\bar{u}(\cdot))$, $\partial^2_\ell\varphi(x_{\gamma_k}(\cdot))$, $(\partial^2_\ell\psi_i(x_{\gamma_k}(\cdot)))_{i=1}^r$, and, finally, $\partial^x L(\cdot,x_{\gamma_k}(\cdot))$ are measurable on $[0,T]$. Hence, the Filippov measurable selection theorem (see [39, Theorem 2.3.13]) yields that we can assume the measurability of the functions $\zeta_{\gamma_k}(\cdot)$, $\tilde{\zeta}_{\gamma_k}(\cdot)$, $\theta_{\gamma_k}(\cdot)$, $(\vartheta^i_{\gamma_k}(\cdot))_{i=1}^r$, and $\omega_{\gamma_k}(\cdot)$. Moreover, these sequences are uniformly bounded in L^∞, as $\|\zeta_{\gamma_k}\|_\infty\le M_\ell$, $\|\tilde{\zeta}_{\gamma_k}\|_\infty\le M_\ell$, $\|\theta_{\gamma_k}\|_\infty\le K$, $\|\vartheta^i_{\gamma_k}\|_\infty\le 2M_\psi$ for $i=1,\dots,r$, and

$$\|\omega_{\gamma_k}\|_\infty\le\overbrace{2\|x_{\gamma_k}-\bar{x}\|_\infty}^{d_{\gamma_k}}\le M_L,\tag{33}$$

for some $M_L > 0$, which exists from (31) and the uniform convergence of the sequence x_{γ_k} to $\bar{x}$.

The major difference between the proof of this theorem and that of [32, Theorem 6.2] lies in the construction of the adjoint vector p (for fixed $(\alpha, \beta) \in (0,1]^2$). More specifically, the difference is manifested below in the intricacy to prove that in our setting, where $r > 1$, the sequence $(p_{\gamma_k})_k$ enjoys the same properties obtained there for $r = 1$, namely, $(p_{\gamma_k})_k$ is uniformly bounded and has uniform bounded variation. Once we establish these facts for our general setting, the construction of the remaining items ν^i (for each i), ζ, θ, ϑ_i (for each i) and λ, follows arguments similar to those in [32, Proof of Theorem 6.2] when constructing therein ν, ξ, ζ, θ, ϑ and λ, respectively.

We fix $(\alpha, \beta) \in (0,1]^2$. Using (32), we obtain

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|p_{\gamma_k}(t)\|^2 \\
&= \left\langle \left(-(1-\beta)\zeta_{\gamma_k}(t)^\mathsf{T} - \beta\tilde{\zeta}_{\gamma_k}(t)^\mathsf{T} + \theta_{\gamma_k}(t) + \sum_{i=1}^{r} \gamma_k e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \vartheta^i_{\gamma_k}(t) \right) p_{\gamma_k}(t), p_{\gamma_k}(t) \right\rangle \\
&+ \sum_{i=1}^{r} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))} |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle|^2 + \lambda_{\gamma_k}\langle \omega_{\gamma_k}, p_{\gamma_k}\rangle,
\end{aligned}$$

and hence,

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|p_{\gamma_k}(t)\|^2 \\
&\geq \left\langle \left(-(1-\beta)\zeta_{\gamma_k}(t)^\mathsf{T} - \beta\tilde{\zeta}_{\gamma_k}(t)^\mathsf{T} + \theta_{\gamma_k}(t) + \sum_{i=1}^{r} \gamma_k e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \vartheta^i_{\gamma_k}(t) \right) p_{\gamma_k}(t), p_{\gamma_k}(t) \right\rangle \\
&+ \lambda_{\gamma_k}\langle \omega_{\gamma_k}, p_{\gamma_k}\rangle \\
&\geq -\left(M_\ell + K + 2M_\psi \sum_{i=1}^{r} \gamma_k e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \right) \|p_{\gamma_k}(t)\|^2 - 2\lambda_{\gamma_k}\|x_{\gamma_k} - \bar{x}\|_\infty \|p_{\gamma_k}(t)\| \\
&= -\left(M_\ell + K + 2M_\psi \xi_{\gamma_k}(t) \right) \|p_{\gamma_k}(t)\|^2 - \lambda_{\gamma_k} d_{\gamma_k} \|p_{\gamma_k}(t)\|.
\end{aligned}$$

Thus, using Grönwall's Lemma in [37, Lemma 4.1] for $\alpha = \frac{1}{2}$, and the uniform boundedness of $(\|\xi_{\gamma_k}\|_2^2)_k$ proved in Theorem 4.10, we get that

$$\begin{aligned}
\|p_{\gamma_k}(t)\| &\leq e^{(M_\ell + K + 2M_\psi\sqrt{M_\xi})}(\|p_{\gamma_k}(T)\| + \lambda_{\gamma_k} d_{\gamma_k}) \\
&\leq (1 + L_g + M_L) e^{(M_\ell + K + 2M_\psi\sqrt{M_\xi})} := M_1,
\end{aligned} \tag{34}$$

where the last inequality follows from (33), and the nontriviality condition (i) when $C_T \neq \mathbb{R}^n$, and the transversality condition (iii) when $C_T = \mathbb{R}^n$. Hence, (p_{γ_k}) is uniformly bounded by the constant M_1 on $[0, T]$.

We proceed to prove that $(\dot{p}_{\gamma_k})$ is uniformly bounded in L^1. First, we note that the method used in [32], for $r = 1$, cannot be used for the version (13) of (D_{γ_k}) since the generalized Hessian of ψ_{γ_k} is not bounded near $\bar{x}$, nor for the version (12) since here $r \geq 1$.

Hence, a new technique is required. From (32), we have

$$\int_0^T \|\dot{p}_{\gamma_k}(t)\| dt$$
$$\leq \int_0^T \Bigg\| \underbrace{\left(-(1-\beta)\zeta_{\gamma_k}(t)^{\mathsf{T}} - \beta\tilde{\zeta}_{\gamma_k}(t)^{\mathsf{T}} + \theta_{\gamma_k}(t) + \sum_{i=1}^{r} \gamma_k e^{\gamma_k \psi_i(x_{\gamma_k}(t))}\, \vartheta^i_{\gamma_k}(t)\right)}_{\mathcal{B}_{\gamma_k}(t)} p_{\gamma_k}(t) \Bigg\| dt$$
$$+ \sum_{i=1}^{r} \int_0^T \left\| \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \nabla\psi_i(x_{\gamma_k}(t))\, \langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle \right\| dt + \lambda_{\gamma_k} \int_0^T \|\omega_{\gamma_k}(t)\| dt,$$

and hence,

$$\begin{aligned} \int_0^T \|\dot{p}_{\gamma_k}(t)\| dt &\leq \underbrace{\left[T(M_\ell + K) + 2TM_\psi\sqrt{M_\xi}\right]}_{\bar{M}_2} \|p_{\gamma_k}\|_\infty + Td_{\gamma_k} \\ &\quad + \underbrace{\sum_{i=1}^{r} \int_0^T \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))}\, \|\nabla\psi_i(x_{\gamma_k}(t))\|\, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle|\, dt}_{\mathbf{I}} \\ &\leq M_2(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}) + \mathbf{I}, \end{aligned} \tag{35}$$

where $M_2 := \max\{\bar{M}_2, T\}$. In order to prove that $\mathbf{I}$ is uniformly bounded, we write $\mathbf{I} = \mathbf{I}_1 + \mathbf{I}_2$, in which $\bar{a}$ is the positive constant in Lemma 6.3(iii):

- $\mathbf{I}_1 := \displaystyle\sum_{i=1}^{r} \int_{[I^{\bar{a}}(\bar{x})]^c} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))}\, \|\nabla\psi_i(x_{\gamma_k}(t))\|\, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle|\, dt.$
- $\mathbf{I}_2 := \displaystyle\sum_{i=1}^{r} \int_{I^{\bar{a}}(\bar{x})} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))}\, \|\nabla\psi_i(x_{\gamma_k}(t))\|\, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle|\, dt.$

Since x_{γ_k} converges uniformly to $\bar{x}$, then, assumption (A2.1) and $x_{\gamma_k}(t) \in C$ for all $t \in [0,T]$, imply the existence of $\bar{k} \in \mathbb{N}$ such that for $k \geq \bar{k}$,

$$|\psi_i(x_{\gamma_k}(t)) - \psi_i(\bar{x}(t))| \leq \frac{\bar{a}}{2} \text{ for all } t \in [0,T] \text{ and for } i = 1, \ldots, r. \tag{36}$$

Hence, using the definition of $I^{\bar{a}}(\bar{x})$ in (11), it follows that, for $k \geq \bar{k}$,

$$\psi_i(x_{\gamma_k}(t)) \leq -\frac{\bar{a}}{2}, \text{ for all } i = 1, \ldots, r \text{ and for all } t \in [I^{\bar{a}}(\bar{x})]^c. \tag{37}$$

Thus, there exists $\bar{M}_3 > 0$ and $\bar{k}_o \geq \bar{k}$ such that for $k \geq \bar{k}_o$ and for $i = 1, \ldots, r$, we have

$$\int_{[I^{\bar{a}}(\bar{x})]^c} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))}\, \|\nabla\psi_i(x_{\gamma_k}(t))\|\, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle|\, dt \leq \bar{M}_3 \|p_{\gamma_k}\|_\infty.$$

This yields that for $k \geq \bar{k}_o$,

$$\mathbf{I}_1 \leq r\bar{M}_3 \|p_{\gamma_k}\|_\infty := M_3 \|p_{\gamma_k}\|_\infty. \tag{38}$$

Next, the definition of $I^{\bar{a}}(\bar{x})$ and $\mathcal{I}^{\bar{a}}_{\bar{x}(t)}$ given in (11), and equation (36) yield that, for $k \geq \bar{k}_o$,

$$\psi_i(x_{\gamma_k}(t)) \leq -\frac{\bar{a}}{2} \text{ for all } t \in I^{\bar{a}}(\bar{x}) \text{ and for all } i \in [\mathcal{I}^{\bar{a}}_{\bar{x}(t)}]^c. \tag{39}$$

Hence, for $k \geq \bar{k}_o$,

$$\begin{aligned} \mathbf{I}_2 &= \int_{I^{\bar{a}}(\bar{x})} \sum_{i=1}^{r} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \|\nabla\psi_i(x_{\gamma_k}(t))\| \, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle| \, dt \\ &= \int_{I^{\bar{a}}(\bar{x})} \sum_{i\in\mathcal{I}^{\bar{a}}_{\bar{x}(t)}} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \|\nabla\psi_i(x_{\gamma_k}(t))\| \, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle| \, dt \\ &+ \int_{I^{\bar{a}}(\bar{x})} \sum_{i\in[\mathcal{I}^{\bar{a}}_{\bar{x}(t)}]^c} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \|\nabla\psi_i(x_{\gamma_k}(t))\| \, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle| \, dt. \end{aligned}$$

Then, for $k \geq \bar{k}_o$,

$$\begin{aligned} \mathbf{I}_2 &\leq \int_{I^{\bar{a}}(\bar{x})} \sum_{i\in\mathcal{I}^{\bar{a}}_{\bar{x}(t)}} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \|\nabla\psi_i(x_{\gamma_k}(t))\| \, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle| \, dt \\ &+ \int_{I^{\bar{a}}(\bar{x})} \sum_{i\in[\mathcal{I}^{\bar{a}}_{\bar{x}(t)}]^c} \gamma_k^2 \bar{M}_\psi^2 e^{-\gamma_k(\frac{\bar{a}}{2})} \|p_{\gamma_k}\|_\infty dt. \end{aligned}$$

This yields the existence of $M_4 > 0$ and $\bar{k}_1 \geq \bar{k}_o$ such that for $k \geq \bar{k}_1$,

$$\mathbf{I}_2 \leq \int_{I^{\bar{a}}(\bar{x})} \sum_{i\in\mathcal{I}^{\bar{a}}_{\bar{x}(t)}} \gamma_k^2 e^{\gamma_k \psi_i(x_{\gamma_k}(t))} \|\nabla\psi_i(x_{\gamma_k}(t))\| \, |\langle \nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle| \, dt + M_4\|p_{\gamma_k}\|_\infty. \tag{40}$$

Now let $i \in \{1, \ldots, r\}$. Since $\nabla\psi_i$ is Lipschitz, and x_{γ_k} and p_{γ_k} are absolutely continuous, we deduce that the function $|\langle p_{\gamma_k}(\cdot), \nabla\psi_i(x_{\gamma_k}(\cdot))\rangle|$ is absolutely continuous. Thus, for $t \in [0, T]$ a.e., there exists $\bar{\vartheta}^i_{\gamma_k}(t) \in \partial^2\psi_i(x_{\gamma_k}(t))$ such that

$$\begin{aligned} &\frac{d}{dt}\, |\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle| \\ &= \left[\langle \dot{p}_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle + \langle \bar{\vartheta}^i_{\gamma_k}(t)\dot{x}_{\gamma_k}(t), p_{\gamma_k}(t)\rangle\right] \underbrace{\operatorname{sign}(\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle)}_{s^i_{\gamma_k}(t)}. \end{aligned} \tag{41}$$

Now substitute into (41) the term $\dot{p}_{\gamma_k}(t)$ from (32), we obtain that for $t \in [0, T]$ a.e.

$$\begin{aligned} &\sum_{j=1}^{r} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))} \langle \nabla\psi_j(x_{\gamma_k}(t)), \nabla\psi_i(x_{\gamma_k}(t))\rangle \langle \nabla\psi_j(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle \\ &= \frac{d}{dt}|\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle| - s^i_{\gamma_k}(t)\langle \bar{\vartheta}^i_{\gamma_k}(t)\dot{x}_{\gamma_k}(t), p_{\gamma_k}(t)\rangle \\ &\quad - s^i_{\gamma_k}(t)\, \langle \mathcal{B}_{\gamma_k}(t)p_{\gamma_k}(t) + \lambda_{\gamma_k}\omega_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle \,. \end{aligned} \tag{42}$$

As x_{γ_k} converges uniformly to $\bar{x}$, there exists $\bar{k}_2 \geq \bar{k}_1$ such that for $k \geq \bar{k}_2$ we obtain $x_{\gamma_k}(t) \in \bar{B}_{\bar{\rho}}(\bar{x}(t))$ for all $t \in [0, T]$, where $\bar{\rho}$ is the constant of Lemma 6.3(iii).

For $t \in I^{\bar a}(\bar x)$ and $k \geq \bar k_2$, we write

$$\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}}\sum_{j=1}^{r} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\langle\nabla\psi_j(x_{\gamma_k}(t)),\nabla\psi_i(x_{\gamma_k}(t))\rangle\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle$$

$$=\sum_{j=1}^{r}\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\langle\nabla\psi_j(x_{\gamma_k}(t)),\nabla\psi_i(x_{\gamma_k}(t))\rangle\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle$$

$$=\sum_{j\in\mathcal{I}^{\bar a}_{\bar x(t)}}\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\langle\nabla\psi_j(x_{\gamma_k}(t)),\nabla\psi_i(x_{\gamma_k}(t))\rangle\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle$$

$$+\sum_{j\in[\mathcal{I}^{\bar a}_{\bar x(t)}]^c}\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\langle\nabla\psi_j(x_{\gamma_k}(t)),\nabla\psi_i(x_{\gamma_k}(t))\rangle\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle$$

$$=\sum_{j\in\mathcal{I}^{\bar a}_{\bar x(t)}}\Bigg\{\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\Bigg(\|\nabla\psi_j(x_{\gamma_k}(t))\|^2$$

$$+\sum_{\substack{i\in\mathcal{I}^{\bar a}_{\bar x(t)}\\ i\neq j}} s^i_{\gamma_k}(t)s^j_{\gamma_k}(t)\langle\nabla\psi_j(x_{\gamma_k}(t)),\nabla\psi_i(x_{\gamma_k}(t))\rangle\Bigg)|\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle|\Bigg\}$$

$$+\sum_{j\in[\mathcal{I}^{\bar a}_{\bar x(t)}]^c}\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\langle\nabla\psi_j(x_{\gamma_k}(t)),\nabla\psi_i(x_{\gamma_k}(t))\rangle\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle.$$

Hence using Lemma 6.3(iii), we obtain that for $t \in I^{\bar a}(\bar x)$ and $k \geq \bar k_2$,

$$\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}}\sum_{j=1}^{r} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\langle\nabla\psi_j(x_{\gamma_k}(t)),\nabla\psi_i(x_{\gamma_k}(t))\rangle\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle$$

$$\geq(1-\bar b)\sum_{j\in\mathcal{I}^{\bar a}_{\bar x(t)}}\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\|\nabla\psi_j(x_{\gamma_k}(t))\|^2|\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle|$$

$$+\sum_{j\in[\mathcal{I}^{\bar a}_{\bar x(t)}]^c}\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\langle\nabla\psi_j(x_{\gamma_k}(t)),\nabla\psi_i(x_{\gamma_k}(t))\rangle\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle.$$

Combining this latter inequality with the summation over $i \in \mathcal{I}^{\bar a}_{\bar x(t)}$ of (42), and using (39), we deduce that for $t \in I^{\bar a}(\bar x)$ and for $k \geq \bar k_2$,

$$\sum_{j\in\mathcal{I}^{\bar a}_{\bar x(t)}}\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\|\nabla\psi_j(x_{\gamma_k}(t))\|^2|\langle\nabla\psi_j(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle| \tag{43}$$

$$\leq\frac{1}{1-\bar b}\Bigg\{\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}}\Bigg[\frac{d}{dt}|\langle p_{\gamma_k}(t),\nabla\psi_i(x_{\gamma_k}(t))\rangle| - s^i_{\gamma_k}(t)\langle\bar\vartheta^i_{\gamma_k}(t)\dot x_{\gamma_k}(t),p_{\gamma_k}(t)\rangle$$

$$- s^i_{\gamma_k}(t)\left\langle\mathcal{B}_{\gamma_k}(t)p_{\gamma_k}(t)+\lambda_{\gamma_k}\omega_{\gamma_k}(t),\nabla\psi_i(x_{\gamma_k}(t))\right\rangle\Bigg]$$

$$+\sum_{j\in[\mathcal{I}^{\bar a}_{\bar x(t)}]^c}\sum_{i\in\mathcal{I}^{\bar a}_{\bar x(t)}}\bar M^3_\psi\gamma_k^2 e^{-\frac{1}{2}\gamma_k\bar a}\|p_{\gamma_k}\|_\infty\Bigg\}.$$

From (42), we have for $i = 1, \dots, r$, that

$$\begin{aligned}
&\frac{d}{dt}|\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle| \\
&= \sum_{j=1}^{r} s^i_{\gamma_k}(t)\gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\langle\nabla\psi_j(x_{\gamma_k}(t)), \nabla\psi_i(x_{\gamma_k}(t))\rangle\langle\nabla\psi_j(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle \\
&\quad + s^i_{\gamma_k}(t)\langle\bar{\vartheta}^i_{\gamma_k}(t)\dot{x}_{\gamma_k}(t), p_{\gamma_k}(t)\rangle \\
&\quad + s^i_{\gamma_k}(t)\left\langle\mathcal{B}_{\gamma_k}(t)p_{\gamma_k}(t) + \lambda_{\gamma_k}\omega_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\right\rangle.
\end{aligned}$$

This gives using (37), (39) and the uniform boundedness of $(\dot{x}_{\gamma_k})_k$, the existence of $M_5 > 0$ and $\bar{k}_3 \geq \bar{k}_2$ such that for $k \geq \bar{k}_3$, we have

$$\left|\int_{[I^{\bar{a}}(\bar{x})]^c} \sum_{i=1}^{r} \frac{d}{dt}\left|\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle\right| dt\right| \leq M_5(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}), \tag{44}$$

and

$$\left|\int_{I^{\bar{a}}(\bar{x})} \sum_{i\in[\mathcal{I}^{\bar{a}}_{\bar{x}(t)}]^c} \frac{d}{dt}\left|\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle\right| dt\right| \leq M_5(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}). \tag{45}$$

Add to (44) that

$$\left|\int_0^T \sum_{i=1}^{r} \frac{d}{dt}\left|\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle\right| dt\right| \leq 2r\bar{M}_\psi\|p_{\gamma_k}\|_\infty,$$

we deduce the existence of $M_6 > 0$ and $\bar{k}_4 \geq \bar{k}_3$ such that for $k \geq \bar{k}_4$, we have

$$\left|\int_{I^{\bar{a}}(\bar{x})} \sum_{i=1}^{r} \frac{d}{dt}\left|\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle\right| dt\right| \leq M_6(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}).$$

This latter inequality with (45) yield the existence of $M_7 > 0$ and $\bar{k}_5 \geq \bar{k}_4$ such that for $k \geq \bar{k}_5$, we have

$$\left|\int_{I^{\bar{a}}(\bar{x})} \sum_{i\in\mathcal{I}^{\bar{a}}_{\bar{x}(t)}} \frac{d}{dt}\left|\langle p_{\gamma_k}(t), \nabla\psi_i(x_{\gamma_k}(t))\rangle\right| dt\right| \leq M_7(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}). \tag{46}$$

Now, integration both sides of (43) on $I^{\bar{a}}(\bar{x})$, and using (46), we get the existence of $M_8 > 0$ and and $\bar{k}_6 \geq \bar{k}_5$ such that for $k \geq \bar{k}_6$,

$$\begin{aligned}
&\int_{I^{\bar{a}}(\bar{x})} \sum_{j\in\mathcal{I}^{\bar{a}}_{\bar{x}(t)}} \gamma_k^2 e^{\gamma_k\psi_j(x_{\gamma_k}(t))}\|\nabla\psi_j(x_{\gamma_k}(t))\|^2|\langle\nabla\psi_j(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle|dt \\
&\leq M_8(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}).
\end{aligned} \tag{47}$$

Using that x_{γ_k} converges uniformly to $\bar{x}$, and by assuming that $\bar{a} \leq \frac{\varepsilon}{2}$, where ε is the constant of (3), we get the existence of $\bar{k}_7 \geq \bar{k}_6$ such that for $k \geq \bar{k}_7$, $t \in I^{\bar{a}}(\bar{x})$ and $j \in \mathcal{I}^{\bar{a}}_{\bar{x}(t)}$, we have $\psi_j(x_{\gamma_k}(t)) \geq -\varepsilon$, and hence by (3), $\|\nabla\psi_j(x_{\gamma_k}(t)\| > \eta$.

Then, for $M_9 := \frac{M_8}{\eta}$, (47) yields that for $k \geq \bar{k}_7$,

$$\int_{I^{\bar{a}}(\bar{x})} \sum_{j \in \mathcal{I}^{\bar{a}}_{\bar{x}(t)}} \gamma_k^2 e^{\gamma_k \psi_j(x_{\gamma_k}(t))} \|\nabla \psi_j(x_{\gamma_k}(t))\| |\langle \nabla \psi_j(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle| dt \leq M_9(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}).$$

Combining this latter with (40), we conclude that for $k \geq \bar{k}_7$,

$$\mathbf{I}_2 \leq M_9(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}) + M_4 \|p_{\gamma_k}\|_\infty,$$

and hence by (38) and for $M_{10} := M_3 + M_4 + M_9$,

$$\mathbf{I} = \mathbf{I}_1 + \mathbf{I}_2 \leq M_{10}(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}). \tag{48}$$

Therefore, for $k \geq \bar{k}_7$, we have from (33), (34), (35) and (48), that

$$\int_0^T \|\dot{p}_{\gamma_k}(t)\| dt \leq (M_2 + M_{10})(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}) \leq (M_2 + M_{10})(M_1 + M_L).$$

This yields that $(\dot{p}_{\gamma_k})$ is uniformly bounded in L^1, which terminates the proof of Theorem 3.1 under the temporary assumption: C is compact.

Now, we proceed to show that the temporary assumption "C is compact" can be replaced by (A2.4). Let C be unbounded, but satisfies (A2.4) for some y_o and R_o. We introduce an additional constraint to C via $\psi_{r+1} \colon \mathbb{R}^n \longrightarrow \mathbb{R}$, where

$$\psi_{r+1}(x) := \frac{1}{2}\left(\|x - y_o\|^2 - R_o^2\right) \text{ for all } x \in \mathbb{R}^n.$$

Hence, C_{r+1}, the zero-sublevel set of ψ_{r+1}, is $\overline{B}_{R_o}(y_o)$ and it contains $\bar{x}(t)$, for all $t \in [0, T]$, in its interior. Denote by $(\hat{P})$ the problem (P) with the following modifications:

- The *unbounded* sweeping set C is now replaced by the *compact* set

$$\hat{C} := C \cap C_{r+1} = C \cap \overline{B}_{R_o}(y_o) = \bigcap_{i=1}^{r+1} \{x \in \mathbb{R}^n : \psi_i(x) \leq 0\}.$$

- The function φ is replaced by the function $\hat{\varphi} := \varphi + I_{C_{r+1}}$, where $I_{C_{r+1}}$ is the indicator function of the set C_{r+1}. Clearly we have $\operatorname{dom} \hat{\varphi} = \hat{C}$ and (A3) is satisfied $\hat{\varphi}$ on $\hat{C}$. Moreover, since $\varphi = \hat{\varphi}$ on the open set $\operatorname{int} C_{r+1}$, we have

$$\partial\varphi(x) = \partial\hat{\varphi}(x), \quad \forall x \in C \cap \operatorname{int} C_{r+1}. \tag{49}$$

- The set C_0 is replaced by the closed set

$$\hat{C}_0 := C_0 \cap C_{r+1} = C_0 \cap \overline{B}_{R_o}(y_o) \subset C \cap \overline{B}_{R_o}(y_o) = \hat{C} = \operatorname{dom} \hat{\varphi}. \tag{50}$$

We claim that $(\bar{x}, \bar{u})$ is a strong local minimizer for $(\hat{P})$. Indeed, since $(\bar{x}, \bar{u})$ is admissible for (P) and $\bar{x}(t) \in C \cap \operatorname{int} C_{r+1}$ for all $t \in [0, T]$, then (49) and (50) yield that $(\bar{x}, \bar{u})$ is admissible for $(\hat{P})$. For the optimality, let $\hat{\delta} > 0$ satisfying

$$\hat{\delta} < \min\{\delta, R_o - \bar{\varepsilon}\}, \quad \text{where } \bar{\varepsilon} := \max\{\|\bar{x}(t) - y_o\|, \ t \in [0, T]\} < R_o. \tag{51}$$

Let (x,u) be admissible for $(\hat{P})$ such that $\|x-\bar{x}\|_\infty \le \hat{\delta}$. Then, $x(t) \in \hat{C}$ for all $t \in [0,T]$, and by (51), $\|x-\bar{x}\|_\infty \le \delta$ and $\|x(t)-y_o\| < R_o$ for all $t \in [0,T]$. Hence, (49) implies that (x,u) is admissible for (P). Thus, the optimality of $(\bar{x},\bar{u})$ for (P) yields that $g(\bar{x}(0),\bar{x}(T)) \le g(x(0),x(T))$, which shows the optimality of $(\bar{x},\bar{u})$ for $(\hat{P})$. Now, since $\bar{x}(t) \in \operatorname{int} C_{r+1}$ for all $t \in [0,T]$, we have:

- $\hat{I}^0(\bar{x}) := \{t \in [0,T] : \bar{x}(t) \in \operatorname{bdry} \hat{C}\} = \{t \in [0,T] : \bar{x}(t) \in \operatorname{bdry} C\} = I^0(\bar{x})$.
- For $t \in \hat{I}^0(\bar{x})$,

  $\hat{\mathcal{I}}^0_{\bar{x}(t)} := \{i \in \{1,...,r+1\} : \psi_i(\bar{x}(t)) = 0\} = \{i \in \{1,\dots,r\} : \psi_i(\bar{x}(t)) = 0\} = \mathcal{I}^0_{\bar{x}(t)}$.

This yields that $(A2.3)_S$ is the same for the family $(\nabla\psi_i(\bar{x}(\cdot)))_{i=1}^{r+1}$. However, having (A2.2) satisfied by the family $(\nabla\psi_i)_{i=1}^{r+1}$, that now includes $\nabla\psi_{r+1}$, is not automatic, but it holds true under assumption (A2.4)(i), due to Lemma 6.5. On the other hand, since $\hat{C} \subset C$, $\hat{C}_0 \subset C_0$ and $\operatorname{int}\hat{C} = \operatorname{int} C \cap B_{R_o}(y_o)$, we conclude that the data of the problem $(\hat{P})$ satisfies (A2.5), (A5) and (A6), and thus, satisfies all the assumptions (A1)–(A2.2), $(A2.3)_S$ and (A2.5)–(A6), with $\hat{C}$ compact. Therefore, the proof of this theorem, where (A2.4) holds, is completed by applying to the strong local minimizer $(\bar{x},\bar{u})$ of $(\hat{P})$ the version of this theorem already proven for C *compact*, and by noticing the following:

- $\hat{C} \subset C$ yields the set $\hat{C}$ used in the definitions of $\partial_\ell\varphi$, $\partial^2_\ell\varphi$, $\partial^x_\ell f(t,\cdot,u)$, $\partial^2_\ell\psi_i$ and $\partial^L_\ell g$, can be replaced by C.
- $\bar{x}(t) \in \operatorname{int} C_{r+1}$ for all $t \in [0,T]$, implies that $\xi^{r+1} \equiv 0$, $\vartheta_{r+1} \equiv 0$ and $\nu_{r+1} \equiv 0$.
- The local property of the limiting normal cone and $\bar{x}(0) \in \operatorname{int} \overline{B}_{R_o}(x_o)$ (by (A2.4)(ii)), give that

$$N^L_{\hat{C}_0}(\bar{x}(0)) = N^L_{C_0 \cap \overline{B}_{R_o}(x_o)}(\bar{x}(0)) = N^L_{C_0}(\bar{x}(0)).$$

Now, we move to prove that the temporary assumption $(A2.3)_S$ can be replaced by (A2.3). We assume that, instead of $(A2.3)_S$, (A2.3) is satisfied for some positive vector $\beta \in \mathbb{R}^r$. We denote by $(\tilde{P})$ the version of (P) in which the functions ψ_i are replaced by $\tilde{\psi}_i := \beta_i\psi_i$ for $i = 1,\dots,r$. Note that since for $i = 1,\dots,r$, the zero-sublevel set of $\tilde{\psi}_i$ and ψ_i are the same, the two problems $(\tilde{P})$ and (P) *coincides*. Hence, $(\bar{x},\bar{u})$ is a strong local minimizer of $(\tilde{P})$. Clearly the data of $(\tilde{P})$ satisfies (A1) and (A3)–(A6). On the other hand, one can easily see that the assumptions (A2.1) and $(A2.3)_S$ are satisfied by the family $(\tilde{\psi}_i)_{i=1}^r$. Moreover, since $(\psi_i)_{i=1}^r$ satisfies (A2.2) for some $\eta > 0$, it is not difficult to show that $(\tilde{\psi}_i)_{i=1}^r$ also satisfies (A2.2) with the constant $\eta\tilde{b}$ for $\tilde{b} := \min\{\beta_i : i = 1,\dots,r\}$. Therefore, the data of $(\tilde{P})$ satisfies (A1)–(A2.2), $(A2.3)_S$ and (A2.4)–(A6). Thus, by applying to $(\tilde{P})$ and to its strong local minimizer $(\bar{x},\bar{u})$ the already proven version of Theorem 3.1, we get the existence of an adjoint vector $\tilde{p} \in BV([0,T];\mathbb{R}^n)$, a finite sequence of finite signed Radon measures $(\tilde{\nu}^i)_{i=1}^r$ on $[0,T]$ with $\tilde{\nu}_i$ supported in $I^0_i(\bar{x})$ for $i = 1,\dots,r$, a finite sequence of nonnegative functions $(\tilde{\xi}^i)_{i=1}^r \in L^\infty([0,T];\mathbb{R}^+)$ with $\tilde{\xi}_i$ supported in $I^0_i(\bar{x})$ for $i = 1,\dots,r$, L^∞-functions $\tilde{\zeta}(\cdot)$, $\tilde{\theta}(\cdot)$ and a finite sequence $(\tilde{\vartheta}_i(\cdot))_{i=1}^r$ in $\mathcal{M}_{n\times n}([0,T])$, and $\tilde{\lambda} \ge 0$ such that the conditions (i)–(vi) of Theorem 3.1 are satisfied by $\tilde{p}$, $(\tilde{\nu}^i)_{i=1}^r$, $(\tilde{\xi}^i)_{i=1}^r$, $\tilde{\zeta}$, $\tilde{\theta}$, $(\tilde{\vartheta}_i)_{i=1}^r$, $\tilde{\lambda}$, and $(\tilde{\psi}_i)_{i=1}^r$ instead of p, $(\nu^i)_{i=1}^r$, $(\xi^i)_{i=1}^r$, ζ, θ, $(\vartheta_i)_{i=1}^r$, λ, and $(\psi_i)_{i=1}^r$, respectively. Now to terminate the proof of this theorem when (A2.3) is satisfied, it is sufficient to replace in these obtained conditions $\tilde{\psi}_i$ by $\beta_i\psi_i$, and consider $p := \tilde{p}$, $\nu^i := \beta_i\tilde{\nu}^i$, $\xi^i := \beta_i\tilde{\xi}^i$, $\zeta := \tilde{\zeta}$, $\theta := \tilde{\theta}$, $\vartheta_i := \beta_i^{-1}\tilde{\vartheta}_i$, and $\lambda := \tilde{\lambda}$.

For the "Furthermore" part of the theorem, let $C_T = \mathbb{R}^n$. Proposition 4.19 implies that $\lambda = 1$. The convexity assumption of $f(t,x,U)$ for $x \in C \cap \overline{B}_\delta(\bar{x}(t))$ and $t \in [0,T]$, is removed by using the *relaxation* technique of [40, Section 5.2] in the same fashion as in Step 7 of the proof of [40, Theorem 5.1].

The proof of Theorem 3.1 is terminated. □

Remark 5.1. In the nontriviality condition of Theorem 3.1, the presence of $\|p(T)\|$ instead of $\|p\|_\infty$ and $(\|\nu^i\|_{\text{T.V.}})_{i=1}^r$ results from having these two norms bounded above by $\|p(T)\|$. Indeed, if we take $k \longrightarrow \infty$ in (34), we conclude that

$$\|p\|_\infty \le e^{(M_\ell+K+2M_\psi\sqrt{M_\xi})}\|p(T)\| = \frac{M_1}{1+L_g+M_L}\|p(T)\|.$$

On the other hand, using (3) and (48), and by applying on each ψ_i the same argument employed in [40, Equation (82)] we deduce the existence of $M_{10} > 0$ and $\bar{k}_9 \ge \bar{k}_8$ such that for $k \ge \bar{k}_9$ and for $i = 1, \dots, r$, we have

$$\int_0^T \gamma_k^2 e^{\gamma_k\psi_i(x_{\gamma_k}(t))}\,|\langle\nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle|\,dt \le M_{10}(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}).$$

This gives, using (34), that for $k \ge \bar{k}_9$ and for $i = 1, \dots, r$, we have

$$\|\nu^i_{\gamma_k}\|_{\text{T.V.}} \le M_{10}(\|p_{\gamma_k}\|_\infty + d_{\gamma_k}) \le \frac{M_{10}M_1}{1+L_g+M_L}(\|p_{\gamma_k}(T)\| + d_{\gamma_k}) + M_{10}d_{\gamma_k}, \tag{52}$$

where $\nu^i_{\gamma_k}$ be the finite signed Radon measure on $[0,T]$ defined by

$$d\nu^i_{\gamma_k}(t) := \gamma_k\xi^i_{\gamma_k}(t)\langle\nabla\psi_i(x_{\gamma_k}(t)), p_{\gamma_k}(t)\rangle dt.$$

Since for each $i \in \{1, \dots, r\}$, the signed Radon measure ν^i of Theorem 3.1 is the weak* limit of $\nu^i_{\gamma_k}$, see Step 4 of [32, Proof of Theorem 6.1] for more details, we deduce after taking $k \longrightarrow \infty$ in (52) that

$$\|\nu^i\|_{\text{T.V.}} \le \frac{M_{10}M_1}{1+L_g+M_L}\|p(T)\|.$$

6. Appendix

In this section, we present an example to which our Pontryagin-type maximum principle (Theorem 3.1) is applied to obtain an optimal solution. Furthermore, we establish auxiliary results that are used in different places of the paper, and we provide proofs for Propositions 4.4 and 4.6.

6.1. Example

In this example we illustrate how Theorem 3.1 can be applied to find an optimal solution. We consider the following data for (P) (see Figure 1):

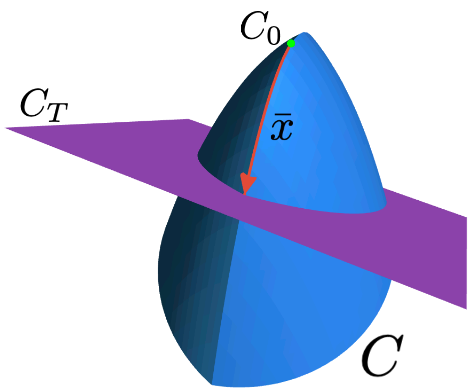

Figure 1: Example 6.1

- The perturbation mapping $f\colon \mathbb{R}^3 \times \mathbb{R} \longrightarrow \mathbb{R}^3$ is defined by

$$f((x_1, x_2, x_3), u) = (4x_1 + u, x_2 - 1, -2x_1 - x_2 + u + 2).$$

- The two functions $\psi_1,\ \psi_2\colon \mathbb{R}^3 \longrightarrow \mathbb{R}$ are defined by

$$\psi_1(x_1, x_2, x_3) := x_1^2 + x_2^2 + x_3 \ \text{ and } \ \psi_2(x_1, x_2, x_3) := x_1^2 + (x_2 - 2)^2 + x_3,$$

  and hence, the set C is the *nonsmooth*, convex and *unbounded* set

$$C = C_1 \cap C_2 := \{(x_1, x_2, x_3) : \psi_1(x_1, x_2, x_3) \leq 0\} \cap \{(x_1, x_2, x_3) : \psi_2(x_1, x_2, x_3) \leq 0\}.$$

- The objective function $g\colon \mathbb{R}^6 \longrightarrow \mathbb{R} \cup \{\infty\}$ is defined by

$$g(x_1, x_2, x_3, x_4, x_5, x_6) := \begin{cases} -x_4^2 - x_6 - 1 & (x_4, x_5, x_6) \in C, \\ \infty & \text{otherwise.} \end{cases}$$

- The function φ is the indicator function of C and $T := \frac{1}{2}$.
- The control multifunction is the constant function $U(t) := [-1, 1]$ for all $t \in [0, \frac{1}{2}]$, $C_0 := \{(0, 1, -1)\}$, and $C_T := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : 8x_1 - 4x_3 - 9 = 0\}$.

One can easily verify that (A1)–(A2.2) and (A2.4)–(A6) are satisfied. Define the curve

$$\Gamma := \{(x_1, x_2, x_3) : x_1^2 + x_3 + 1 = 0 \text{ and } x_2 = 1\} = (\operatorname{bdry} C_1 \cap \operatorname{bdry} C_2) \subset \operatorname{bdry} C.$$

Since $C_0 \subset \Gamma$ and g vanishes on Γ and is strictly positive elsewhere in C, we may seek for (P) a candidate $(\bar{x}, \bar{u})$ for optimality with $\bar{x} := (\bar{x}_1, \bar{x}_2, \bar{x}_3)$ belonging to Γ, if possible, and hence we have

$$\begin{cases} \bar{x}_1^2(t) + \bar{x}_3(t) + 1 = 0 \ \text{ and } \ \bar{x}_2(t) = 1, \ \ \forall t \in [0, \frac{1}{2}] \ \text{ and} \\ \bar{x}(0)^\mathsf{T} = (0, 1, -1) \ \text{ and } \ \bar{x}(\frac{1}{2})^\mathsf{T} \in \{(\frac{1}{2}, 1, -\frac{5}{4}), (-\frac{5}{2}, 1, -\frac{29}{4})\}. \end{cases} \tag{53}$$

Note that (A2.3) is satisfied on Γ for $\beta = (1, 1)$.[8] Then we apply Theorem 3.1 to such a candidate $(\bar{x}, \bar{u})$ and obtain in consequence the existence of an adjoint vector

[8] Note that for $(x_1, x_2, x_3) \in \Gamma$ with $-\frac{\sqrt{3}}{2} < x_1 < \frac{\sqrt{3}}{2}$, we have $\langle \nabla\psi_1(x_1, x_2, x_3), \nabla\psi_2(x_1, x_2, x_3)\rangle = 4x_1^2 - 3 < 0$, and hence, the maximum principle of [21] cannot be applied to this sweeping set C.

$p := (p_1, p_2, p_3) \in BV([0, \frac{1}{2}]; \mathbb{R}^3)$, two finite signed Radon measures ν_1, ν_2 on $\left[0, \frac{1}{2}\right]$, ξ_1, $\xi_2 \in L^\infty([0, \frac{1}{2}]; \mathbb{R}^+)$, and $\lambda \geq 0$, such that when incorporating equations (53) into Theorem 3.1(i)–(vi), we obtain

(a) $\|p(\frac{1}{2})\| + \lambda = 1$.

(b) The admissibility equation holds, that is, for $t \in [0, \frac{1}{2}]$ a.e.,

$$\begin{cases} \dot{\bar{x}}_1(t) = 4\bar{x}_1(t) + \bar{u}(t) - 2\bar{x}_1(t)(\xi_1(t) + \xi_2(t)), \\ 0 = -2(\xi_1(t) - \xi_2(t)), \\ -2\dot{\bar{x}}_1(t)\bar{x}_1(t) = -2\bar{x}_1(t) + \bar{u}(t) + 1 - ((\xi_1(t) + \xi_2(t)). \end{cases}$$

(c) The adjoint equation is satisfied, that is, for $t \in [0, \frac{1}{2}]$,

$$dp(t) = \begin{pmatrix} -4 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} p(t)dt + (\xi_1(t) + \xi_2(t)) \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} p(t)dt$$
$$+ \begin{pmatrix} 2\bar{x}_1(t) \\ 2 \\ 1 \end{pmatrix} d\nu_1 + \begin{pmatrix} 2\bar{x}_1(t) \\ -2 \\ 1 \end{pmatrix} d\nu_2.$$

(d) The complementary slackness condition is valid, that is, for $t \in [0, \frac{1}{2}]$ a.e.,

$$\begin{cases} \xi_1(t)(2p_1(t)\bar{x}_1(t) + 2p_2(t) + p_3(t)) = 0, \\ \xi_2(t)(2p_1(t)\bar{x}_1(t) - 2p_2(t) + p_3(t)) = 0. \end{cases}$$

(e) The transversality condition holds, that is,

$$\begin{cases} -p(\frac{1}{2})^\mathsf{T} \in \lambda\{-1, 0, -1)\} + \{\alpha(2, 0, -1) : \alpha \in \mathbb{R}\} & \text{if } \bar{x}(\frac{1}{2})^\mathsf{T} = (\frac{1}{2}, 1, -\frac{5}{4}), \\ -p(\frac{1}{2})^\mathsf{T} \in \lambda\{(5, 0, -1)\} + \{\alpha(2, 0, -1) : \alpha \in \mathbb{R}\} & \text{if } \bar{x}(\frac{1}{2})^\mathsf{T} = (-\frac{5}{2}, 1, -\frac{29}{4}). \end{cases}$$

(f) $\max\{u(p_1(t) + p_3(t)) : u \in [-1, 1]\}$ is attained at $\bar{u}(t)$ for $t \in [0, \frac{1}{2}]$ a.e.

We temporarily assume that

$$p_1(t) + p_3(t) \geq 0, \quad \forall t \in [0, \tfrac{1}{2}] \text{ a.e.} \tag{54}$$

This gives from (f) that $u(t) = 1$ for $t \in [0, \frac{1}{2}]$ a.e. Now solving the differential equations of (b) and using (53), we obtain that

$$\xi_1(t) = \xi_2(t) = 1 \text{ and } \bar{x}(t)^\mathsf{T} = (t, 1, -1 - t^2), \quad \forall t \in [0, \tfrac{1}{2}].$$

Hence, from (d), we deduce that $p_2(t) = 0$ for $t \in [0, \frac{1}{2}]$ a.e., and

$$2tp_1(t) + p_3(t) = 0, \quad \forall t \in [0, \tfrac{1}{2}] \text{ a.e.} \tag{55}$$

Moreover, the adjoint equation (c) simplifies to the following

$$\begin{cases} dp_1(t) = 2p_3(t)dt + 2td\nu_1 + 2td\nu_2, \\ 0 = p_3(t)dt + 2d\nu_1 - 2d\nu_2, \\ dp_3(t) = d\nu_1 + d\nu_2. \end{cases} \tag{56}$$

Using (a), (55), (e), and (56), a simple calculation gives that

$$\begin{cases} \lambda = \frac{1}{4} \text{ and } p(\frac{1}{2})^{\mathsf{T}} = (\frac{3}{4}, 0, 0), \\ p(t)^{\mathsf{T}} = \left(\frac{3}{4(4t^2+1)}, 0, \frac{-3t}{2(4t^2+1)}\right) \text{ on } [0, \frac{1}{2}), \\ d\nu_1 = \frac{12t^3+24t^2+3t-6}{8(4t^2+1)^2} dt + \frac{3}{16}\delta_{\left\{\frac{1}{2}\right\}} \text{ and } d\nu_2 = \frac{-12t^3+24t^2-3t-6}{8(4t^2+1)^2} dt + \frac{3}{16}\delta_{\left\{\frac{1}{2}\right\}} \text{ on } [0, \frac{1}{2}], \end{cases}$$

where $\delta_{\{a\}}$ denotes the unit measure concentrated on the point a. Note that for all $t \in [0, \frac{1}{2}]$, we have $p_1(t) + p_3(t) \geq 0$, and hence, the temporary assumption (54) is satisfied.

Therefore, the above analysis, realized via Theorem 3.1, produces an admissible pair $(\bar{x}, \bar{u})$, where

$$\bar{x}(t)^{\mathsf{T}} = (t, 1, -1 - t^2) \text{ and } \bar{u}(t) = 1, \quad \forall t \in [0, \tfrac{1}{2}],$$

which is optimal for (P).

6.2. Auxiliary results

Lemma 6.1. *Assume that* (A2.1)–(A2.2) *hold and that* C *is compact. Then for* $c \in \operatorname{bdry} C$, *there exists a vector* $d_c \neq 0$ *such that*

$$\frac{4\eta^2}{\bar{M}_\psi} \leq \|d_c\| \leq r\bar{M}_\psi, \ \ \textit{and} \ \ \left\langle \frac{d_c}{\|d_c\|}, \nabla\psi_i(c) \right\rangle \leq -\frac{4\eta^2}{r\bar{M}_\psi}, \quad \forall i \in \mathcal{I}_c^0.$$

Proof. Let $c \in \operatorname{bdry} C$. For each $i \in \mathcal{I}_c^0$, we denote by $u_i(c)$ and $v_i(c)$ the unique projections of $-\nabla\psi_i(c)$ to $N_C(c)$ and $T_C(c)$, respectively. By Moreau decomposition theorem, see [27], we have

$$-\nabla\psi_i(c) = u_i(c) + v_i(c) \text{ and } \langle u_i(c), v_i(c) \rangle = 0.$$

This yields that $\|\nabla\psi_i(c)\|^2 = \|u_i(c)\|^2 + \|v_i(c)\|^2$, and hence

$$\|v_i(c)\| \leq \|\nabla\psi_i(c)\| \leq \bar{M}_\psi. \tag{57}$$

On the other hand, by Proposition 4.1(ii), we have that $u_i(c) = \sum_{j \in \mathcal{I}_c^0} \lambda_j^i \nabla\psi_j(c)$, where $\lambda_j \geq 0$ for $j \in \mathcal{I}_c^0$. Then by (A2.2),

$$\begin{aligned} \|v_i(c)\| = \|\nabla\psi_i(c) + u_i(c) = & \left\| \left(\sum_{j \in \mathcal{I}_c^0,\ j \neq i} \lambda_j^i \nabla\psi_j(c) \right) + (1 + \lambda_i^i)\nabla\psi_i(c) \right\| \\ & \geq \left(1 + \sum_{j \in \mathcal{I}_c^0} \lambda_j^i \right) 2\eta \geq 2\eta. \end{aligned}$$

Hence, using (57), we obtain that

$$2\eta \leq \|v_i(c)\| \leq \bar{M}_\psi, \quad \forall i \in \mathcal{I}_c^0. \tag{58}$$

It follows that,

$$\langle v_i(c), \nabla\psi_i(c)\rangle = \langle v_i(c), -u_i(c) - v_i(c)\rangle = -\|v_i(c)\|^2 \le -4\eta^2. \tag{59}$$

We define $d_c := \sum_{j\in\mathcal{I}_c^0} v_j(c)$. As $v_j(c)$ belongs to $T_C(c)$ for all $j \in \mathcal{I}_c^0$, we have that $\langle v_j(c), \nabla\psi_i(c)\rangle \le 0$, for all $i, j \in \mathcal{I}_c^0$. This, together with (59) gives that, for all $i \in \mathcal{I}_c^0$ we have

$$\bar{M}_\psi\|d_c\| \ge \langle d_c, -\nabla\psi_i(c)\rangle \ge \langle v_i(c), -\nabla\psi_i(c)\rangle \ge 4\eta^2.$$

Whence, using the right inequality in (58), we have

$$\langle d_c, \nabla\psi_i(c)\rangle \le -4\eta^2,\ \forall i \in \mathcal{I}_c^0, \ \text{ and } \ \frac{4\eta^2}{\bar{M}_\psi} \le \|d_c\| \le r\bar{M}_\psi.$$

This yields that $d_c \neq 0$ and, upon dividing the first inequality in the above equation by $\|d_c\|$, the required result is established. □

Lemma 6.2. *Assume that* (A2.1) *holds. Let $\alpha_n \ge 0$, for all $n \in \mathbb{N}$, with $\alpha_n \longrightarrow \alpha_o$ and let $c_n \in C$ be a sequence such that $\mathcal{I}_{c_n}^{\alpha_n} \neq \emptyset$, for all $n \in \mathbb{N}$, and $c_n \longrightarrow c_o$. Then, $\mathcal{I}_{c_o}^{\alpha_o} \neq \emptyset$ and there exist $\emptyset \neq \mathcal{J}_o \subset \{1,\dots,r\}$ and a subsequence of $(\alpha_n, c_n)_n$ we do not relabel, such that*

$$\mathcal{I}_{c_n}^{\alpha_n} = \mathcal{J}_o \subset \mathcal{I}_{c_o}^{\alpha_o} \text{ for all } n \in \mathbb{N}.$$

In particular, for any continuous function $\bar{x}\colon [0,T] \longrightarrow C$ and for all $a \ge 0$, we have $I^a(\bar{x})$ is closed, and hence compact.

Proof. For each $i \in \{1,\dots,r\}$, we define $\mathcal{N}_i := \{n \in \mathbb{N} : -\alpha_n \le \psi_i(c_n) \le 0\}$. Since $\bigcup_{i=1}^r \mathcal{N}_i = \mathbb{N}$, we have that the set $\mathcal{R} := \{i \in \{1,\dots,r\} : |\mathcal{N}_i| = \infty\}$ is nonempty.
Assume that $\mathcal{R} = \{i_1, i_2, \dots, i_m\}$ where $i_1 < i_2 < \cdots < i_m$.
For each $j \notin \mathcal{R}$ and for each $\ell \in \{1,\dots,m\}$, we eliminate from $\mathcal{N}_{i_\ell}$ all the finite number of indices n for which $-\alpha_n \le \psi_j(c_n) \le 0$, if such indices exist. Thus, we now ensured that

$$\mathcal{I}_{c_n}^{\alpha_n} \subset \mathcal{R} \text{ for all } n \in \bigcup_{\ell=1}^m \mathcal{N}_{i_\ell}.$$

Now to construct the required subsequence and the set $\mathcal{J}_o$, we apply the following algorithm:

(1) Let $\ell = 2$ and $\mathcal{J}_o = \{i_1\}$.

(2) If $\mathcal{N}_{i_1} \cap \mathcal{N}_{i_\ell}$ is an infinite set, then we replace $\mathcal{N}_{i_1}$ by $\mathcal{N}_{i_1} \cap \mathcal{N}_{i_\ell}$, and we add i_2 to $\mathcal{J}_o$. Otherwise, we replace $\mathcal{N}_{i_1}$ by $\mathcal{N}_{i_1} \setminus (\mathcal{N}_{i_1} \cap \mathcal{N}_{i_\ell})$.

(3) We increment ℓ by 1. If $\ell = m+1$, then go to (4). Otherwise, go to (2).

(4) Halt.

At the end of the algorithm, the so-obtained set of indices $\mathcal{N}_{i_1}$ will be an infinite set. Moreover, if we consider $(c_{n_k})_k$ to be the subsequence of $(c_n)_n$ associated to $\mathcal{N}_{i_1}$, then we clearly have that $\mathcal{I}_{c_{n_k}}^{\alpha_{n_k}} = \mathcal{J}_o$ for all k. The continuity of ψ_i for each $i \in \{1,\dots,r\}$, and the convergence of α_n to α_o and c_n to c_o , yield that $\mathcal{J}_o \subset \mathcal{I}_{c_o}^{\alpha_o}$.

We proceed to prove the "in particular" part. Let $\bar{x}\colon [0,T] \longrightarrow C$ be continuous and let $a \geq 0$. For $t_n \in I^a(\bar{x})$ with $t_n \to t_o \in [0,T]$, we consider $\alpha_n = \alpha_o := a$, $c_n := \bar{x}(t_n)$, and $c_o := \bar{x}(t_o)$. By the first part of this lemma, there exists $\mathcal{J}_o \neq \emptyset$ in $\{1,\dots,r\}$ such that, up to a subsequence, $\mathcal{I}^a_{\bar{x}(t_n)} = \mathcal{J}_o \subset \mathcal{I}^a_{\bar{x}(t_o)}$, implying that $t_o \in I^a(\bar{x})$. $\square$

Lemma 6.3. *Assume* (A2.1) *holds, and denote by $G_\psi(t)$ the Gramian matrix of the vectors $\{\nabla\psi_i(\bar{x}(t))\}_{i\in\mathcal{I}^0_{\bar{x}(t)}}$ for all $t \in I^0(\bar{x})$. Then the following assertions are equivalent*:

(i) *For all $t \in I^0(\bar{x})$, $G_\psi(t)$ is strictly diagonally dominant.*

(ii) *There exists $b \in (0,1)$ such that*

$$\sum_{i\in\mathcal{I}^0_{\bar{x}(t)},\ i\neq j} |\langle\nabla\psi_i(\bar{x}(t)), \nabla\psi_j(\bar{x}(t))\rangle| \leq b\|\nabla\psi_j(\bar{x}(t))\|^2, \tag{60}$$

for all $t \in I^0(\bar{x})$ and for all $j \in \mathcal{I}^0_{\bar{x}(t)}$.

(iii) *There exist $\bar{a} > 0$, $\bar{b} \in (0,1)$ and $\bar{\rho} > 0$, such that for all $t \in I^{\bar{a}}(\bar{x})$ and for all $j \in \mathcal{I}^{\bar{a}}_{\bar{x}(t)}$ we have*

$$\sum_{i\in\mathcal{I}^{\bar{a}}_{\bar{x}(t)},\ i\neq j} |\langle\nabla\psi_i(x), \nabla\psi_j(x)\rangle| \leq \bar{b}\|\nabla\psi_j(x)\|^2,\ \ \forall x \in \bar{B}_{\bar{\rho}}(\bar{x}(t)) \cap C.$$

Proof. (i)⇒(ii): For $t \in I^0(\bar{x})$ and $j \in \mathcal{I}^0_{\bar{x}(t)}$, we define

$$b(t,j) := \frac{1}{\|\nabla\psi_j(\bar{x}(t))\|^2} \sum_{i\in\mathcal{I}^0_{\bar{x}(t)},\ i\neq j} |\langle\nabla\psi_i(\bar{x}(t)), \nabla\psi_j(\bar{x}(t))\rangle| < 1. \tag{61}$$

Now for $b := \sup\left\{b(t,j) : t \in I^0(\bar{x}) \text{ and } j \in \mathcal{I}^0_{\bar{x}(t)}\right\}$, we clearly have that (60) holds true. To prove that $b < 1$, it is sufficient to use an argument by contradiction in conjunction with Lemma 6.2 and inequality (61).

(ii)⇒(iii): Follows immediately using an argument by contradiction in conjunction with Lemma 6.2.

(iii)⇒(i): Follows directly since we have $\bar{b} \in (0,1)$, $\mathcal{I}^0_{\bar{x}(t)} \subset \mathcal{I}^{\bar{a}}_{\bar{x}(t)}$ for all $t \in [0,T]$, and $I^0(\bar{x}) \subset I^{\bar{a}}(\bar{x})$. $\square$

Lemma 6.4. *Let $S \subset \mathbb{R}^n$ be a nonempty and closed set. The assumption* (A2.4)(i) *is satisfied by S if one of the following conditions holds*:

(i) bdry S *is compact.*

(ii) *S is star-shaped* (*which includes convex and polyhedral sets*).

Moreover, in the two cases above, the radius R_o can be taken to be arbitrarily large.

Proof. (i): It is sufficient to take y_o any point in $\mathbb{R}^n$ and R_o large enough so that bdry $S \subset B_{R_o}(y_o)$.

(ii): Let y_o be a center of S and let R_o be any positive number. We consider $s \in \text{bdry}\, S \cap (\mathbb{S}_{R_o}(y_o))$. We claim that $y_o - s \notin N_S(s)$. Indeed, if not, then there exists

$\rho_s > 0$ such that for $\zeta_s := \frac{y_o - s}{\|y_o - s\|}$, we have $B_{\rho_s}(s + \rho_s \zeta_s) \subset S^c$. Thus for $s_o := s + \bar{\rho}_s \zeta_s$, where $0 < \bar{\rho}_s < \min\{\rho_s, R_o\}$, we have $s_o \in [s, y_o] \subset S$ and $s_o \in B_{\rho_s}(s + \rho_s \zeta_s) \subset S^c$. This gives the desired contradiction. $\square$

Lemma 6.5. *Assume that* (A2.1)–(A2.2) *and* (A2.4)(i) *hold. Then the assumption* (A2.2) *is satisfied by the family of functions* $(\psi_i)_{i=1}^{r+1}$, *where*

$$\psi_{r+1}(x) := \frac{1}{2}\left(\|x - y_o\|^2 - R_o^2\right), \quad \forall x \in \mathbb{R}^n.$$

Proof. If not, then there exist $(c_n)_n$ in $\mathbb{R}^n$ and $(\lambda_1^n, \ldots, \lambda_{r+1}^n)_n \in [0, \infty)^{r+1}$ such that for each n we have:

- $c_n \in \hat{C} := C \cap C_{r+1} = C \cap \overline{B}_{R_o}(y_o)$ which is a compact set.
- $\hat{\mathcal{I}}^0_{c_n} := \{i \in \{1, \ldots, r+1\} : \psi_i(c_n) = 0\} \neq \emptyset$, and hence $c_n \in \operatorname{bdry} \hat{C}$.
- $\lambda_i^n = 0$ for $i \notin \hat{\mathcal{I}}^0_{c_n}$, and

$$\sum_{i \in \hat{\mathcal{I}}^0_{c_n}} \lambda_i^n = 1 \ \text{ and } \ \Big\| \sum_{i \in \hat{\mathcal{I}}^0_{c_n}} \lambda_i^n \nabla \psi_i(c_n) \Big\| \leq \frac{2}{n}. \tag{62}$$

Since (A2.2) is satisfied, once can easily deduce that for $n \geq \lceil \frac{1}{\eta} \rceil$, we have that $\lambda_{r+1}^n \neq 0$. This yields that $c_n \in (\mathbb{S}_{R_o}(y_o))$, for $n \geq \lceil \frac{1}{\eta} \rceil$. Hence the bounded sequence $(c_n)_n$ admits a subsequence, we do not relabel, that converges to

$$c_o \in C \cap \mathbb{S}_{R_o}(y_o) \subset \operatorname{bdry} \hat{C}.$$

Applying Lemma 6.2, we deduce that $(c_n)_n$ has a subsequence, we do not relabel, that satisfies

$$\exists\, \emptyset \neq \hat{\mathcal{I}}_o \subset \{1, \ldots, r+1\} \text{ such that } \hat{\mathcal{I}}^0_{c_n} = \hat{\mathcal{I}}_o \subset \hat{\mathcal{I}}^0_{c_o} \text{ for all } n \in \mathbb{N}.$$

Hence, using (62), we get that

$$\sum_{i \in \hat{\mathcal{I}}_o} \lambda_i^n = 1 \ \text{ and } \ \Big\| \sum_{i \in \hat{\mathcal{I}}_o} \lambda_i^n \nabla \psi_i(c_n) \Big\| \leq \frac{2}{n}.$$

Now taking $n \longrightarrow \infty$ in this latter and using the boundedness of the sequence $(\lambda_1^n, \ldots, \lambda_{r+1}^n)_n$, we deduce the existence of $(\lambda_1^o, \ldots, \lambda_{r+1}^o) \in [0, \infty)^{r+1}$ such that

$$\sum_{i \in \hat{\mathcal{I}}_o} \lambda_i^o = 1 \ \text{ and } \ \sum_{i \in \hat{\mathcal{I}}_o} \lambda_i^o \nabla \psi_i(c_o) = 0. \tag{63}$$

Case 1: $c_o \in \operatorname{int} C$.

Then, for n sufficiently large, we have $\hat{\mathcal{I}}_o = \hat{\mathcal{I}}^0_{c_n} = \{r+1\}$. This yields, using (63), that $\lambda_{r+1}^o = 1$ and $c_o - y_o = 0$, which contradicts $c_o \in \mathbb{S}_{R_o}(y_o)$.

Case 2: $c_o \in \operatorname{bdry} C$.

We claim that $\lambda_{r+1}^o \neq 0$. Indeed, if not then from (63) and for $\lambda_i^o := 0$ for all $i \in \hat{\mathcal{I}}^0_{c_o} \setminus \hat{\mathcal{I}}_o$, we have

$$\sum_{i \in \hat{\mathcal{I}}^0_{c_o}} \lambda_i^o = 1 \ \text{ and } \ \sum_{i \in \hat{\mathcal{I}}^0_{c_o}} \lambda_i^o \nabla \psi_i(c_o) = 0,$$

which contradicts (A2.2).

Hence, $\lambda^o_{r+1} \neq 0$, and this gives using (63), that

$$y_o - c_o = \sum_{\substack{i \in \hat{\mathcal{I}}_o \\ i \neq r+1}} \frac{\lambda^o_i}{\lambda_{r+1}} \nabla\psi_i(c_o) = \sum_{i \in \hat{\mathcal{I}}^0_{c_o}} \frac{\lambda^o_i}{\lambda_{r+1}} \nabla\psi_i(c_o) \in N^P_C(c_o),$$

where $\lambda^o_i := 0$ for $i \in \hat{\mathcal{I}}^0_{c_o} \setminus \hat{\mathcal{I}}_o$. This contradicts (A2.4)(i). □

6.3. Proofs of Propositions 4.4 and 4.6

Proof of Proposition 4.4. (i): The definition of ψ_{γ_k} in (8) and assumption (A2.1) yield that, for all $k \in \mathbb{N}$, ψ_{γ_k} is $\mathcal{C}^{1,1}$ on $C + \rho B$. The bound in (17) follows from (14) and the definition of $\bar{M}_\psi$. For the remaining properties, see [25, Subsection 2.2].

(ii): If this statement was not true, there would exist a sequence $(\gamma_{k_n})_n$ with $k_n \geq n$, $x_{\gamma_{k_n}} \in \mathbb{R}^n$ with $\psi_{\gamma_{k_n}}(x_{\gamma_{k_n}}) = 0$, and $z_{\gamma_{k_n}} \in B_{\frac{2}{n}}(x_{\gamma_{k_n}})$ such that for all $n > \frac{2}{\rho}$ we have $\|\nabla\psi_{\gamma_{k_n}}(z_{\gamma_{k_n}})\| \leq 2\eta$. Using (16) and the definition of $\psi_{\gamma_{k_n}}$, it follows that

$$\psi(x_{\gamma_{k_n}}) \leq 0 \leq \psi(x_{\gamma_{k_n}}) + \frac{\ln(r)}{\gamma_k} \quad \text{and} \quad \frac{\left\|\sum_{i=1}^r e^{\gamma_{k_n}\psi_i(z_{\gamma_{k_n}})} \nabla\psi_i(z_{\gamma_{k_n}})\right\|}{\sum_{i=1}^r e^{\gamma_{k_n}\psi_i(z_{\gamma_{k_n}})}} \leq 2\eta. \tag{64}$$

Then, $x_{\gamma_{k_n}} \in C$ and hence, there exists a subsequence, we do not relabel, of $(\gamma_{k_n})_n$ along which $(x_{\gamma_{k_n}})_n$ and $(z_{\gamma_{k_n}})_n$ converge to the same element $z_o \in C$. Taking $n \longrightarrow \infty$ in (64) and using the fact that $e^{\gamma_{k_n}\psi_i(z_{\gamma_{k_n}})} \longrightarrow 0$ whenever $\psi_i(z_o) < 0$, we get the existence of a sequence of nonnegative numbers $(\lambda_i)_{i \in \mathcal{I}^0_{z_o}}$ such that

$$\psi(z_o) = 0, \quad \Big\| \sum_{i \in \mathcal{I}^0_{z_o}} \lambda_i \nabla\psi_i(z_o) \Big\| \leq 2\eta \;\text{ and }\; \sum_{i \in \mathcal{I}^0_{z_o}} \lambda_i = 1.$$

This contradicts assumption (A2.2) since $\psi(z_o) = 0$ is equivalent to $\mathcal{I}^0_{z_o} \neq \emptyset$.

(iii): If this statement was not true, there would exist a sequence $(\gamma_{k_n})_n$ with $k_n \geq n$ and $x_{\gamma_{k_n}} \in C^{\gamma_{k_n}} \subset C$ such that

$$\|\nabla\psi_{\gamma_{k_n}}(x_{\gamma_{k_n}})\| \leq \eta \;\text{ and }\; -\frac{1}{n} \leq \psi_{\gamma_{k_n}}(x_{\gamma_{k_n}}) \leq 0. \tag{65}$$

This yields that $x_{\gamma_{k_n}} \in C$ and $\psi(x_{\gamma_{k_n}}) \longrightarrow 0$. Since C is compact, we can assume that $x_{\gamma_{k_n}} \longrightarrow z_o \in C$, and hence $\psi(z_o) = \lim_{n \longrightarrow 0} \psi(x_{\gamma_{k_n}}) = 0$. Now, from (65)$(a)$ and (15), we have

$$\frac{\| \sum_{i=1}^r e^{\gamma_{k_n}\psi_i(x_{\gamma_{k_n}})} \nabla\psi_i(x_{\gamma_{k_n}})\|}{\sum_{i=1}^r e^{\gamma_{k_n}\psi_i(x_{\gamma_{k_n}})}} \leq \eta.$$

Taking $n \longrightarrow \infty$ in this latter and using that $e^{\gamma_{k_n}\psi_i(x_{\gamma_{k_n}})} \longrightarrow 0$ whenever $\psi_i(z_o) < 0$, we get the existence of a sequence of nonnegative numbers $(\lambda_i)_{i \in \mathcal{I}^0_{z_o}}$ such that

$$\Big\| \sum_{i \in \mathcal{I}^0_{z_o}} \lambda_i \nabla\psi_i(z_o) \Big\| \leq \eta \;\text{ and }\; \sum_{i \in \mathcal{I}^0_{z_o}} \lambda_i = 1.$$

This contradicts assumption (A2.2), since $\psi(z_o) = 0$ implies that $\mathcal{I}^0_{z_o} \neq \emptyset$. □

Proof of Proposition 4.6. (i): By Proposition 4.4, we have that for each $k \geq k_3$, ψ_{γ_k} satisfies the *same assumptions* satisfied by the function ψ of [32, Theorem 3.1]. Hence, (i) follows immediately from [32, Theorem 3.1(i)] where ψ is replaced by ψ_{γ_k}.

(ii): First we prove that
$$\operatorname{int} C \subset \bigcup_{k\in\mathbb{N}} \operatorname{int} C^{\gamma_k}(k). \tag{66}$$

Let $x \in \operatorname{int} C$. Since $\psi(x) < 0$ and $\left(\alpha_k + \frac{\ln r}{\gamma_k}\right) \longrightarrow 0$, there exists $k_x \geq k_4$ such that $\alpha_{k_x} + \frac{\ln r}{\gamma_{k_x}} < -\psi(x)$. This yields, using (16), that $\psi_{\gamma_{k_x}}(x) < -\alpha_{k_x}$, and hence

$$x \in \operatorname{int} C^{\gamma_{k_x}}(k_x) \subset \bigcup_{k\in\mathbb{N}} \operatorname{int} C^{\gamma_k}(k).$$

This terminates the proof of (66). Hence,

$$\operatorname{int} C \subset \bigcup_{k\in\mathbb{N}} \operatorname{int} C^{\gamma_k}(k) \subset \bigcup_{k\in\mathbb{N}} C^{\gamma_k}(k) \subset \bigcup_{k\in\mathbb{N}} \operatorname{int} C^{\gamma_k} \subset \operatorname{int} C.$$

Therefore, the equation (19) holds true. Now, since $(\psi_{\gamma_k})_k$ is monotonically nonincreasing in terms of k, and $(\alpha_k)_k$ is a decreasing sequence, we conclude that the sequence $(C^{\gamma_k}(k))_k$ is a nondecreasing sequence. This gives that the Painlevé-Kuratowski limit of the sequence $(C^{\gamma_k}(k))_k$ satisfies

$$\lim_{k\longrightarrow\infty} C^{\gamma_k}(k) = \operatorname{cl}\Big(\bigcup_{k\in\mathbb{N}} C^{\gamma_k}(k)\Big). \tag{67}$$

Upon taking the closure of $\operatorname{int} C$ in (19) and using from Proposition 4.1(i) the fact that $C = \operatorname{cl}(\operatorname{int} C)$, equation (67) yields that the Painlevé-Kuratowski limit of the sequence $(C^{\gamma_k}(k))_k$ is C.

(iii): For $c \in \operatorname{bdry} C$, we set $-2\alpha_c := \max\{\psi_i(c) : i \notin \mathcal{I}_c^0\}$, whenever $\mathcal{I}_c^0 \subsetneq \{1, \dots, r\}$. We have $\alpha_c > 0$ and $\psi_i(c) \leq -2\alpha_c$ for all $i \notin \mathcal{I}_c^0$. The continuity of ψ_i yields the existence of $\bar{r}_1 > 0$ such that $\psi_i(x) < -\alpha_c$, for all $x \in \bar{B}_{2\bar{r}_1}(c)$ and for all $i \notin \mathcal{I}_c^0$. Let d_c be the nonzero vector of Lemma 6.1. Choose $\bar{k}_1 \geq k_3$ large enough so that for all $k \geq \bar{k}_1$ we have $\frac{\ln(r)}{\gamma_k} + \alpha_k \leq \alpha_c$ and $\sigma_k \leq \bar{r}_1$, where α_k and σ_k are the sequences defined in (6). Then, for all $k \geq \bar{k}_1$, and for all $x \in \bar{B}_{\bar{r}_1}(c)$ we have:

- $\left\|x + \sigma_k \frac{d_c}{\|d_c\|} - c\right\| \leq \|x - c\| + \sigma_k \leq 2\bar{r}_1$,
- $\psi_i\left(x + \sigma_k \frac{d_c}{\|d_c\|}\right) < -\alpha_c \leq -\frac{\ln(r)}{\gamma_k} - \alpha_k$ for all $i \notin \mathcal{I}_c^0$.

Therefore,

$$\psi_i\left(x + \sigma_k \frac{d_c}{\|d_c\|}\right) + \frac{\ln(r)}{\gamma_k} < -\alpha_k, \quad \forall i \notin \mathcal{I}_c^0, \quad \forall k \geq \bar{k}_1, \text{ and } \forall x \in \bar{B}_{\bar{r}_1}(c). \tag{68}$$

Now, for $i \in \mathcal{I}_c^0$, by using Lemma 6.1 we get

$$\left\langle \frac{d_c}{\|d_c\|}, \nabla\psi_i(c) \right\rangle \leq -\frac{4\eta^2}{r\bar{M}_\psi}, \quad \forall i \in \mathcal{I}_c^0.$$

The continuity of $\nabla\psi_i$ yields the existence of $0 < r_c \le \bar{r}_1$ such that

$$\left\langle \frac{d_c}{\|d_c\|}, \nabla\psi_i(x) \right\rangle < -\frac{2\eta^2}{r\bar{M}_\psi}, \quad \forall i \in \mathcal{I}_c^0 \text{ and } \forall x \in \overline{B}_{2r_c}(c). \tag{69}$$

We consider $k_c \ge \bar{k}_1$ large enough so that $\sigma_k \le r_c$. Then, for all $k \ge k_c$ and for all $x \in \overline{B}_{r_c}(c)$, we have

$$\left\| x + \sigma_k \frac{d_c}{\|d_c\|} - c \right\| \le \|x - c\| + \sigma_k \le 2r_c.$$

Hence by the mean value theorem applied to ψ_i on $[x + \sigma_k \frac{d_c}{\|d_c\|}, x]$, for $i \in \mathcal{I}_c^0$, for $k \ge k_c$, and for $x \in \overline{B}_{r_c}(c) \cap C$, and using (69) and (6), we obtain the existence of $z \in [x + \sigma_k \frac{d_c}{\|d_c\|}, x] \subset \overline{B}_{2r_c}(c)$ such that

$$\psi_i\left(x + \sigma_k \frac{d_c}{\|d_c\|}\right) = \psi_i(x) + \sigma_k \left\langle \nabla\psi_i(z), \frac{d_c}{\|d_c\|} \right\rangle < 0 - \frac{\ln(r)}{\gamma_k} - \alpha_k.$$

Combining this latter with (68), we conclude that

$$\psi_i\left(x + \sigma_k \frac{d_c}{\|d_c\|}\right) + \frac{\ln(r)}{\gamma_k} < -\alpha_k, \quad \forall i \in \{1, \dots, r\}, \quad \forall k \ge k_c, \text{ and } \forall x \in \overline{B}_{r_c}(c) \cap C.$$

Therefore, using (7) and (16), we deduce that

$$\psi_{\gamma_k}\left(x + \sigma_k \frac{d_c}{\|d_c\|}\right) < -\alpha_k, \quad \forall k \ge k_c \text{ and } \forall x \in \overline{B}_{r_c}(c) \cap C.$$

The proof of (iii) is complete. □

References

[1] L. Adam, J. Outrata: *On optimal control of a sweeping process coupled with an ordinary differential equation*, Discrete Contin. Dyn. Syst. Ser. B 19/9 (2014) 2709–2738.

[2] S. Adly, F. Nacry, L. Thibault: *Discontinuous sweeping process with prox-regular sets*, ESAIM COCV, 23/4 (2017) 1293–1329.

[3] 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.

[4] A. Brudnyi, Y. Brudnyi: *Methods of Geometric Analysis in Extension and Trace Problems. Volume 1*, Monographs in Mathematics Vol. 102, Birkhäuser/Springer, Basel (2012).

[5] T. H. Cao, G. Colombo, B. Mordukhovich, D. Nguyen: *Optimization of fully controlled sweeping processes*, J. Diff. Equations 295 (2021) 138–186.

[6] T. H. Cao, G. Colombo, B. Mordukhovich, D. Nguyen: *Optimization and discrete approximation of sweeping processes with controlled moving sets and perturbations*, J. Diff. Equations 274 (2021) 461–509.

[7] T. H. Cao, B. 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. Mordukhovich: *Optimality conditions for a controlled sweeping process with applications to the crowd motion model*, Discrete Contin. Dyn. Syst. Ser. B 22 (2017) 267–306.

[9] T. H. Cao, B. Mordukhovich: *Optimal control of a nonconvex perturbed sweeping process*, J. Diff. Equations 266 (2019) 1003–1050.

[10] F. H. Clarke: *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York (1983).

[11] F. H. Clarke: *Functional Analysis, Calculus of Variations and Optimal Control*, Graduate Texts in Mathematics 264, Springer, London (2013).

[12] F. H. Clarke, Yu. Ledyaev, R. J. Stern, P. R. Wolenski: *Nonsmooth Analysis and Control Theory*, Graduate Texts in Mathematics 178, Springer, New York (1998).

[13] F. H. Clarke, R. J. Stern, P. R. Wolenski: *Proximal smoothness and the lower-C^2 property*, J. Convex Analysis 2 (1995) 117–144.

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

[15] G. Colombo, R. Henrion, N. D. Hoang, B. S. Mordukhovich: *Optimal control of the sweeping process over polyhedral controlled sets*, J. Diff. Equations 260/4 (2016) 3397–3447.

[16] G. Colombo, B. Mordukhovich, D. Nguyen: *Optimization of a perturbed sweeping process by constrained discontinuous controls*, SIAM J. Control Optimization 58/4 (2020) 2678–2709.

[17] M. d. R. de Pinho, M. M. A. Ferreira, G. V. Smirnov: *Optimal control involving sweeping processes*, Set-Valued Var. Analysis 27/2 (2019) 523–548.

[18] M. d. R. de Pinho, M. M. A. Ferreira, G. V. Smirnov: *Correction to: Optimal control involving sweeping processes*, Set-Valued Var. Analysis 27 (2019) 1025–1027.

[19] M. d. R. de Pinho, M. M. A. Ferreira, G. V. Smirnov: *Optimal control with sweeping processes: Numerical method*, J. Optim. Theory Appl. 185 (2020) 845–858.

[20] M. d. R. de Pinho, M. M. A. Ferreira, G. V. Smirnov: *Necessary conditions for optimal control problems with sweeping systems and end point constraints*, Optimization 71/11 (2021) 3363–3381.

[21] M. d. R. de Pinho, M. M. A. Ferreira, G. V. Smirnov: *A maximum principle for optimal control problems involving sweeping processes with a nonsmooth set*, arXiv: 2301.13620v1 (2023).

[22] J. M. Harrison, M. I. Reiman: *Reflected Brownian motion on an orthant*, Ann. Probab. 9 (1981) 302–308.

[23] C. Hermosilla, M. Palladino: *Optimal control of the sweeping process with a nonsmooth moving set*, SIAM J. Control Optimization 60/5 (2022) 2811–2834.

[24] B. Li: *Generalizations of Diagonal Dominance in Matrix Theory*, Ph. D. Dissertation, The University of Regina (Canada), Order Number: AAINQ30262 (1998).

[25] X.-S. Li, S.-C. Fang: *On the entropic regularization method for solving min-max problems with applications*, Math. Methods Oper. Res. 46 (1997) 119–130.

[26] B. S. Mordukhovich: *Variational Analysis and Generalized Differentiation. I: Basic Theory*, Springer, Berlin (2006).

[27] J. J. Moreau: *Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires*, C. R. Acad. Sci. Paris 255 (1962) 238–240.

[28] J. J. Moreau: *Rafle par un convexe variable I*, Trav. Semin. d'Analyse Convexe, Univ. Montpellier (1971), Exposé 15.

[29] J. J. Moreau, *Rafle par un convexe variable II*, Trav. Semin. d'Analyse Convexe, Univ. Montpellier (1972), Exposé 3.

[30] J. J. Moreau: *Evolution problem associated with a moving convex set in a Hilbert space*, J. Diff. Equations 26 (1977) 347–374.

[31] C. Nour, V. Zeidan: *Numerical solution for a controlled nonconvex sweeping process*, IEEE Control Systems Letters 6 (2022) 1190–1195.

[32] C. Nour, V. Zeidan: *Optimal control of nonconvex sweeping processes with separable endpoints: Nonsmooth maximum principle for local minimizers*, J. Diff. Equations 318 (2022) 113–168.

[33] C. Nour, V. Zeidan: *A control space ensuring the strong convergence of continuous approximation for a controlled sweeping process*, Set-Valued Var. Analysis 31 (2023), art. no. 23.

[34] C. Nour, V. Zeidan: *Nonsmooth optimality criterion for a* $\mathrm{W}^{1,2}$*-controlled sweeping process: Nonautonomous perturbation*, Appl. Set-Valued Anal. Optim. 5 (2023) 193–212.

[35] R. A. Poliquin, R. T. Rockafellar, L. Thibault: *Local differentiability of distance functions*, Trans. Amer. Math. Soc. 352 (2000) 5231–5249.

[36] R. T. Rockafellar, R. J.-B. Wets: *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften 317, Springer, Berlin (1998).

[37] R. E. Showalter: *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, Mathematical Surveys and Monographs Volume 49, American Mathematical Society, Providence (1997).

[38] L. Thibault: *Unilateral Variational Analysis in Banach Spaces*, World Scientific, Singapore (2023).

[39] R. B. Vinter: *Optimal Control*, Systems and Control: Foundations and Applications, Birkhäuser, Boston (2000).

[40] V. Zeidan, C. Nour, H. Saoud: *A nonsmooth maximum principle for a controlled nonconvex sweeping process*, J. Diff. Equations 269 (2020) 9531–9582.