Set-Valued and Variational Analysis (2023) 31:23
https://doi.org/10.1007/s11228-023-00686-z

A Control Space Ensuring the Strong Convergence of Continuous Approximation for a Controlled Sweeping Process

Chadi Nour[1] · **Vera Zeidan**[2]

Received: 28 December 2022 / Accepted: 5 June 2023

Abstract
A controlled sweeping process with prox-regular set, $W^{1,2}$-controls, and separable endpoints constraints is considered in this paper. Existence of optimal solutions is established and *local* optimality conditions are derived via *strong* converging *continuous* approximations, whose state entirely resides in the *interior* of the prox-regular set. Consequently, subdifferentials *smaller* than the standard ones are now employed in the optimality results.

Keywords Controlled sweeping process · Prox-regular sets · Necessary optimality conditions · Local minimizers · Strong convergence · Continuous approximations · Nonsmooth analysis

Mathematics Subject Classification (2010) 49K21 · 49K15 · 49J52

1 Introduction

This paper addresses the following fixed time Mayer-type optimal control problem involving $W^{1,2}$-controlled sweeping systems

$$(P):\ \text{Minimize } g(x(0), x(1))$$
$$\text{over } (x,u) \in AC([0,1];\mathbb{R}^n) \times \mathcal{W} \text{ such that}$$
$$\begin{cases} (D) \begin{bmatrix} \dot{x}(t) \in f(x(t), u(t)) - \partial\varphi(x(t)), \quad \text{a.e. } t \in [0,1], \\ x(0) \in C_0 \subset \operatorname{dom}\varphi, \end{bmatrix} \\ x(1) \in C_1, \end{cases}$$

Communicated by Giovanni Colombo

Chadi Nour and Vera Zeidan contributed equally to this work. This paper is dedicated to our PhD advisor Francis H. Clarke on the occasion of his 75th birthday.

✉ Vera Zeidan
zeidan@msu.edu

Chadi Nour
cnour@lau.edu.lb

1 Department of Computer Science and Mathematics, Lebanese American University, Byblos Campus, P.O. Box 36, Byblos, Lebanon

2 Department of Mathematics, Michigan State University, East Lansing 48824-1027, MI, USA

where, $g\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}\cup\{\infty\}$, $f\colon \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$, $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R}\cup\{\infty\}$, ∂ stands for the Clarke subdifferential, $C := \operatorname{dom}\varphi$ is the zero-sublevel set of a function $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}$, that is, $C = \{x \in \mathbb{R}^n : \psi(x) \leq 0\}$, $C_0 \subset C$, $C_1 \subset \mathbb{R}^n$, and, for $U\colon [0,1] \rightrightarrows \mathbb{R}^m$ a multifunction and $\mathbb{U} := \bigcup_{t\in[0,1]} U(t)$, the set of control functions $\mathcal{W}$ is defined by

$$\mathcal{W} := W^{1,2}([0,1];\mathbb{U}) = \left\{u \in W^{1,2}([0,1];\mathbb{R}^m) : u(t) \in U(t), \quad \forall t \in [0,1]\right\}. \qquad (1)$$

Note that if (x,u) solves (D), it necessarily follows that $x(t) \in C$, $\forall t \in [0,1]$.

A pair (x,u) is *admissible* for (P) if $x\colon [0,1] \longrightarrow \mathbb{R}^n$ is absolutely continuous, $u \in \mathcal{W}$, and (x,u) satisfies the *perturbed* and *controlled sweeping process* (D), called the *dynamic* of (P).

An admissible pair $(\bar{x},\bar{u})$ for (P) is said to be a $W^{1,2}$*-local minimizer* (also known as *intermediate local minimizer of rank* 2) if there exists $\delta > 0$ such that

$$g(\bar{x}(0),\bar{x}(1)) \leq g(x(0),x(1)), \qquad (2)$$

for all (x,u) admissible for (P) with $\|x-\bar{x}\|_\infty \leq \delta$, $\|\dot{x}-\dot{\bar{x}}\|_2^2 \leq \delta$, $\|u-\bar{u}\|_\infty \leq \delta$ and $\|\dot{u}-\dot{\bar{u}}\|_2^2 \leq \delta$. Note that if (2) is satisfied for any admissible pairs (x,u), then $(\bar{x},\bar{u})$ is called a *global minimizer* (or an *optimal solution*) for (P).

J.J. Moreau introduced in [29–31] the model of *sweeping processes* for problems emanating from friction and plasticity theory. Since then, this model and its modified forms have surfaced in many applications not only in physics, but also in engineering, social sciences including economics, etc. (see, e.g., [1] and the references listed therein). This model is distinguished by having in its dynamic the subdifferential of the *indicator function* of C, that is, the normal cone to the set C which is discontinuous and unbounded, and hence, rendering the subject of sweeping processes disjoint from that of the *standard* differential inclusions. Therefore, new approaches are needed to tackle optimal control problems over sweeping processes.

In [3, 17, 18, 20, 34, 39], necessary optimality conditions in the form of a *maximum principle* for optimal control problems involving *measurably*-controlled sweeping processes are derived using *continuous-time* approximations. The continuous approximation employed in [17, 20, 34, 39] is based on replacing the normal cone in (D) by an exponential penalization term leading to the following *standard* control system

$$(D_{\gamma_k}) \quad \dot{x}(t) = f(x(t),u(t)) - \nabla\Phi(x(t)) - \gamma_k e^{\gamma_k\psi(x(t))}\nabla\psi(x(t)), \text{ a.e. } t \in [0,1],$$

where Φ is a smooth extension to $\mathbb{R}^n$ of φ, and $\gamma_k > 0$ with $\gamma_k \longrightarrow \infty$ as $k \longrightarrow \infty$. In these papers, the authors showed that any solution x of the system (D) can be approximated by solutions of (D_{γ_k}) whose velocities converge *weakly* in L^2 to $\dot{x}$. This weak approximation of (D) by (D_{γ_k}) is used in [19, 33] to construct numerical algorithms that solve certain forms of (P), and in [17, 20, 34, 39] to derive necessary optimality conditions via approximating *weakly* the optimal solution of (P) by a sequence of optimal solutions for standard optimal control problems over (D_{γ_k}).

Strong convergence of velocities is well-known to be an essential property for numerical purposes, as it accelerates the convergence of the numerical algorithm, see e.g., [7, 8, 14, 17]. In other words, it is important that the solutions of (P) be *strongly* approximated (in the $W^{1,2}$-norm) by the solutions of approximating problems that are computable via existing numerical algorithms. This question of strong convergence of velocities was previously addressed using *discrete-time* approximations, see for instance, [5–8, 13, 14, 16], where the authors considered optimal control problems involving *various* forms of controlled sweeping processes including the $W^{1,2}$-controls. In [6–8], this approach also served to derive necessary optimality conditions phrased in terms of the weak-Pontryagin-type maximum principle when the control space is $W^{1,2}([0,1];\mathbb{R}^m)$. Therein, these optimality criteria are applied to *real-life* models, whose optimal controls turn out to be $W^{1,2}$.[1]

The main goal of the paper is motivated by the importance of approximating a solution x of the sweeping process (D) by solutions of (D_{γ_k}) whose velocities *strongly* converges to the velocity $\dot{x}$ as described above. We establish the validity of this result when the controls in (D_{γ_k}) are chosen to be *uniformly bounded* in $W^{1,2}$. As a consequence, we approximate a given optimal solution $(\bar{x},\bar{u})$ by a sequence $(x_{\gamma_k},u_{\gamma_k})$ of optimal solutions for standard optimal control problems over (D_{γ_k}) with initial conditions and objective functions carefully formulated to guarantee that (i) the optimal states x_{γ_k} remains entirely in the *interior* of C, see Remark 4.5, and (ii) the optimal controls u_{γ_k} are uniformly bounded in $W^{1,2}$, and hence, the solution velocities $\dot{x}_{\gamma_k}$ *strongly* converges to $\dot{\bar{x}}$. To our knowledge, this is a *first-of-its-kind* result that uses *continuous* approximations, as opposed to discrete approximations, to obtain strong convergence of velocities. Furthermore, necessary optimality conditions are established for $W^{1,2}$-local minimizers of (P) upon taking the limit of the optimality conditions for the corresponding approximating optimal control problems. This latter task requires meticulous analysis.

One may expect that the necessary optimality conditions for (P) could be obtained via a reformulation of the dynamic by considering the state as the pair (x,u) satisfying the sweeping process, and the control to be $v := \dot{u}$, where $v(t)\in\mathbb{R}^m$ a.e., and $u(t)\in U(t)$ for all t, is an explicit state constraint. However, to our knowledge, there is no optimality conditions in the literature for this type of problems.

In the next section, we provide notations and definitions from nonsmooth analysis. A list of assumptions and their analysis are provided in Section 3. In addition, we present some needed results from [34, Sections 4 & 5] including the connection between (D_{γ_k}) and (D) under measurable controls. Section 4 is devoted to (i) showing that (D_{γ_k}) strongly approximates (D) when $W^{1,2}$-bounded controls are utilized, (ii) establishing an existence theorem for an optimal solution of (P), (iii) constructing for (P) a continuous approximating sequence of standard optimal control problems (P_{γ_k}), and (iv) deriving necessary optimality conditions in the form of weak-Pontryagin-type maximum principle for $W^{1,2}$-local minimizers of (P) whose utility is illustrated by an example. To maintain an easy flow of the main results, we postpone the proofs of Theorems 4.1 and 4.7 to Section 5.

2 Preliminaries

2.1 Basic Notations

We denote by $\|\cdot\|$, $\langle\cdot,\cdot\rangle$, the Euclidean norm and the usual inner product, respectively. For $c\in\mathbb{R}^n$ and $r>0$, we define the open (resp. closed) ball centered at c with radius r by $B_r(c) := c + rB$ (resp. $\bar{B}_r(c) := c + r\bar{B}$), where B and $\bar{B}$ denotes the open and the closed unit ball, respectively. For $S\subset\mathbb{R}^n$, the boundary, the interior, the closure, the convex hull, the complement, and the polar of S are denoted by bdry S, int S, cl S, conv S, S^c, and S°, respectively. For $x\in\mathbb{R}^n$ and $S\subset\mathbb{R}^n$, $d(x,S)$ denotes the distance from x to S. The effective

[1] In [22], a new exact penalization technique is introduced to derive Pontryagin-type maximum principle for problem (P), where the control is measurable, φ is the indicator of C, $C_0=\{x_0\}$ and $C_1=\mathbb{R}^n$, and C is the intersection of moving zero sublevel sets of smooth functions. As stated in [22, Remark 2.13], these results have atypical nontriviality condition that requires further study.

domain and the epigraph of an extended-real-valued function $h\colon \mathbb{R}^n \longrightarrow \mathbb{R}\cup\{\infty\}$ are denoted by dom h and epi h, respectively. For a multifunction $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, Gr $F \subset \mathbb{R}^n \times \mathbb{R}^m$ denotes the graph of F. The space $L^p([a,b];\mathbb{R}^n)$ designates the Lebesgue space of p-integrable functions $h\colon [a,b] \longrightarrow \mathbb{R}^n$. We denote by $\|\cdot\|_p$ and $\|\cdot\|_\infty$ the norms of $L^p([a,b];\mathbb{R}^n)$ and $L^\infty([a,b];\mathbb{R}^n)$ (or $C([a,b];\mathbb{R}^n)$), respectively. The set of all $m \times n$-matrix functions on $[a,b]$ is denoted by $\mathcal{M}_{m\times n}([a,b])$. For the set of all absolutely continuous functions from $[a,b]$ to $\mathbb{R}^n$, we use $AC([a,b];\mathbb{R}^n)$. A function $h\colon [a,b] \longrightarrow \mathbb{R}^n$ is said to be a BV-function, if h has a bounded variation, that is, $V_a^b(h) < \infty$, where $V_a^b(h)$ is the total variation of h. The set of all such functions is denoted by $BV([a,b];\mathbb{R}^n)$. We denote by $NBV[a,b]$ the normalized space of BV-functions on $[a,b]$ that consists of those BV-functions h such that $h(a) = 0$ and h is right continuous on (a,b) (see e.g., [27, p.115]). The space $C^*([a,b];\mathbb{R})$ denotes the dual of $C([a,b];\mathbb{R})$ equipped with the supremum norm. The induced norm on $C^*([a,b];\mathbb{R})$ is denoted by $\|\cdot\|_{\text{T.V.}}$. As a consequence of Riesz representation theorem, we can interpret the elements of $C^*([a,b];\mathbb{R})$ as being in $\mathfrak{M}([a,b])$, the set of finite signed Radon measures on $[a,b]$ equipped with the weak* topology. Thereby, to each element of $C^*([a,b];\mathbb{R})$ it corresponds a unique element in $NBV[a,b]$ related through the Stieltjes integral and both elements have the same total variation. The set $C^\oplus(a,b)$ designates the subset of $C^*([a,b];\mathbb{R})$ taking nonnegative values on nonnegative-valued functions in $C([a,b];\mathbb{R})$. For $A \subset \mathbb{R}^d$ compact, the set of continuous functions from A to $\mathbb{R}^n$ is denoted by $C(A;\mathbb{R}^n)$. By $W^{k,p}([a,b];\mathbb{R}^n)$, $k \in \mathbb{N}$ and $p \in [0,+\infty]$, we denote the classical Sobolev space. Note that in this paper, the Sobolev space $W^{1,2}([a,b];\mathbb{R}^n)$ will be considered with the norm $\|x(\cdot)\|_{W^{1,2}} := \|x(\cdot)\|_\infty + \|\dot{x}(\cdot)\|_2$. Hence, the convergence of a sequence x_n strongly in the norm topology of the space $W^{1,2}([a,b];\mathbb{R}^n)$ is equivalent to the uniform convergence of x_n on $[a,b]$ and the strong convergence in L^2 of its derivative $\dot{x}_n$.

2.2 Notions in Nonsmooth Analysis

We begin by listing *standard* notions for normal cones and subdifferentials, and *nonstandard* notions for subdifferentials.We refer the reader to [9, 11, 28, 37], for the standard notions, and to [34, 39], for the nonstandard notions. Let $S \subset \mathbb{R}^n$ be closed, and let $h\colon \mathbb{R}^n \longrightarrow \mathbb{R}\cup\{\infty\}$ and $H\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be two functions such that h is lower semicontinuous.

- For $s \in S$, we denote by $N_S^P(s)$, $N_S^L(s)$, and $N_S(s)$, the *proximal*, the *Mordukhovich* (or *limiting*), and the *Clarke normal cones* to S at s, respectively.
- For $x \in \text{dom}\, h$, we denote by $\partial^P h(x)$, $\partial^L h(x)$, and $\partial h(x)$ the *proximal*, the *Mordukhovich* (or *limiting*), and the *Clarke subdifferential* of h at x, respectively. Note that if h is Lipschitz near x, then the *Clarke generalized gradient* of h at x is also denoted by $\partial h(x)$.
- If h is $\mathcal{C}^{1,1}$ near x, then the *Clarke generalized Hessian* of h at x is denoted by $\partial^2 h(x)$. On the other hand, if H is Lipschitz near $x \in \mathbb{R}^n$, then the *Clarke generalized Jacobian* of H at x is denoted by $\partial H(x)$.
- For $A \subset \text{dom}\, h$ closed with int $A \neq \emptyset$, and $x \in \text{cl}\,(\text{int}\, A)$, we denote by $\partial_\ell^L h(x)$ the *limiting subdifferential* of h *relative* to int A at the point x (see [34, Equation (8)]).
- If dom h is closed with int (dom h) $\neq \emptyset$ and h is locally Lipschitz on int (dom h), then for $x \in \text{cl}\,(\text{int}\,(\text{dom}\, h))$, we denote by $\partial_\ell h(x)$ the *extended Clarke generalized gradient* of h at x (see [34, Equation (9)]).
- If h is $\mathcal{C}^{1,1}$ on int (dom h) and $x \in \text{cl}\,(\text{int}\,(\text{dom}\, h))$, then we denote by $\partial_\ell^2 h(x)$ the *extended Clarke generalized Hessian* of h at x (see [34, Equation (10)]).

- For $A \subset \mathbb{R}^n$ closed with $\operatorname{int} A \neq \emptyset$, if h is $\mathcal{C}^{1,1}$ on an open set containing A, then for $x \in A$, we denote by $\partial^2_\ell h(x)$ the *Clarke generalized Hessian* of h *relative* to $\operatorname{int} S$ at x (see [34, Equation (11)]).
- For $A \subset \mathbb{R}^n$ closed with $\operatorname{int} A \neq \emptyset$, if H is locally Lipschitz on $\operatorname{int} A$, then for $x \in \operatorname{cl}(\operatorname{int} A)$, we denote by $\partial_\ell H(x)$ the *extended Clarke generalized Jacobian* of H at x (see [34, Equation (12)]).

We proceed to introduce three geometric properties, namely, prox-regular, epi-Lipschitz, and quasiconvex, that will be used in different places of the paper. For more information about these properties, see [12, 15, 21, 32, 35] for prox-regularity, [9, 11, 37] for epi-Lipschitz property, and [4] for quasiconvexity.

Definition 2.1 Let $S \subset \mathbb{R}^n$ be a nonempty and closed set.

- Let $r > 0$. We say that S is *r-prox-regular* if for all $s \in S$ and for all ζ unit in $N^P_S(s)$, we have
$$\langle \zeta, x - s\rangle \leq \frac{1}{2r}\|x - s\|^2, \quad \forall x \in S.$$
Note that, in this case, we have $N^P_S(s) = N^L_S(s) = N_S(s)$ for all $s \in S$.
- We say that S is *epi-Lipschitz at $s \in S$* if the Clarke normal cone to S at s is *pointed*, that is, $N_S(s) \cap -N_S(s) = \{0\}$. The set S is said to be *epi-Lipschitz* if it is epi-Lipschitz at s for all $s \in S$. Note that a *convex* set is epi-Lipschitz if and only if it has a nonempty interior.
- The set S is said to be *quasiconvex* if there exists $c \geq 0$ such for any $s_1,\ s_2 \in S$, one can find a polygonal line γ in S joining s_1 to s_2, and satisfying
$$l(\gamma) \leq c\|s_1 - s_2\|, \text{ where } l(\gamma) \text{ is the length of } \gamma.$$
Note that the quasiconvexity of C is an essential property for constructing a special smooth extension to $\mathbb{R}^n$ of the function φ, see Lemma 3.2.

3 Assumptions, Consequences, and known Results

In this section, we introduce assumptions on the data of (P) and we present some of their useful consequences. We also display some needed results from [34, Sections 4 & 5], where the connection between (D_{γ_k}) and (D) under measurable controls is studied. We note that for each result of this paper, we may use a different combination of these assumptions. On the other hand, a local version of (A1), namely, condition $(*)$, is used in Subsections 4.3 and 4.4.

A1: There exist $M > 0$ and $\tilde{\rho} > 0$ such that f is M-Lipschitz on $C \times (\mathbb{U} + \tilde{\rho}\bar{B})$ with $\|f(x,u)\| \leq M$ for all $(x,u) \in C \times (\mathbb{U} + \tilde{\rho}\bar{B})$.

A2: The set $C := \operatorname{dom}\varphi$ is given by $C = \{x \in \mathbb{R}^n : \psi(x) \leq 0\}$, where $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}$.

A2.1: There exists $\rho > 0$ such that ψ is $\mathcal{C}^{1,1}$ on $C + \rho B$.
A2.2: There is a constant $\eta > 0$ such that $\|\nabla\psi(x)\| > 2\eta$ for all $x : \psi(x) = 0$.
A2.3: The function ψ is coercive, that is, $\lim_{\|x\| \longrightarrow \infty} \psi(x) = +\infty$.[2]

[2] This assumption is only needed to get the compactness of C, and then, it can be replaced by the boundedness of C.

A2.4: The set C has a connected interior.[3]

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

A4: For the sets C_0, C_1, and $U(\cdot)$ we have:

A4.1: The set $C_0 \subset C$ is nonempty and closed.
A4.2: The graph of $U(\cdot)$ is a $\mathcal{L} \times \mathcal{B}$ measurable set, and, for $t \in [0, 1]$, $U(t)$ is closed, and bounded uniformly in t.
A4.3: The set $C_1 \subset \mathbb{R}^n$ is nonempty and closed.
A4.4: The multifunction $U(\cdot)$ is lower semicontinuous.

Remark 3.1 For $C \subset \mathbb{R}^n$ defined as the sub-level set of a function ψ, one can show that:

(i) Whenever C is nonempty and compact, ψ is merely $\mathcal{C}^1$ on $C + \rho B$, and (A2.2) holds, then there exists $\varepsilon > 0$ such that

$$x \in C \textit{ and } \|\nabla\psi(x)\| \leq \eta \implies \psi(x) < -\varepsilon. \tag{3}$$

(ii) When ψ is merely $\mathcal{C}^1$ on $C + \rho B$ and satisfies (A2.2)-(A2.3), by [34, Lemma 3.3],

(a) $\operatorname{bdry} C \neq \emptyset$ and $\operatorname{bdry} C = \{x \in \mathbb{R}^n : \psi(x) = 0\}$,
(b) $\operatorname{int} C \neq \emptyset$ and $\operatorname{int} C = \{x \in \mathbb{R}^n : \psi(x) < 0\}$.

The following important properties of the compact set C were obtained in [39, Proposition 3.1], where ψ is assumed to be $\mathcal{C}^{1,1}$ on *all* of $\mathbb{R}^n$. However, a slight modification in the proof of that proposition is performed in [34] to conclude that these properties are actually valid under our assumption (A2.1).

Here and throughout the paper, $\bar{M}_\psi$ denotes an upper bound of $\|\nabla\psi(\cdot)\|$ on the compact set C, and $2M_\psi$ is a *Lipschitz constant* of $\nabla\psi(\cdot)$ over the compact set $C + \frac{\rho}{2}\bar{B}$ chosen large enough so that $M_\psi \geq \frac{4\eta}{\rho}$.

Lemma 3.2 [34, Lemma 3.4] *Under* (A2.1)-(A2.3)*, we have the following*:

(i) *The nonempty set C is compact, amenable* (*in the sense of* [37])*, epi-Lipschitzian, $C = \operatorname{cl}(\operatorname{int} C)$, and C is $\frac{\eta}{M_\psi}$-prox-regular.*
(ii) *For all $x \in \operatorname{bdry} C$ we have $N_C(x) = N_C^P(x) = N_C^L(x) = \{\lambda\nabla\psi(x) : \lambda \geq 0\}$.*
(iii) *If also* (A2.4) *holds, then $\operatorname{int} C$ is quasiconvex. Furthermore, if in addition* (A3) *is satisfied, then there exists a function $\Phi \in \mathcal{C}^1(\mathbb{R}^n)$ such that*:

- *Φ is bounded on $\mathbb{R}^n$, and $\Phi(x) = \varphi(x)$ for all $x \in C$.*
- *Φ and $\nabla\Phi$ are globally Lipschitz on $\mathbb{R}^n$.*
- *For all $x \in C$ we have*

$$\partial\varphi(x) = \{\nabla\Phi(x)\} + N_C(x). \tag{4}$$

Remark 3.3 Lemma 3.2(iii) yields the existence of $K > 0$ and a $\mathcal{C}^1$-extension Φ of φ to $\mathbb{R}^n$ such that (4) is satisfied and

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

[3] Assumption (A2.4) is only needed in Lemma 3.2(iii) to guarantee the quasiconvexity of C used to construct the extension Φ of φ. Hence, when such an extension is trivially accessible, condition (A2.4) would be omitted. This is the case when φ is the *indicator* function of C.

Employing (4), (D) is equivalently phrased in terms of the normal cone to C and the extension Φ of φ, as follows

$$(D)\begin{bmatrix} \dot{x}(t) \in f_\Phi(x(t), u(t)) - N_C(x(t)), & \text{a.e. } t \in [0,1], \\ x(0) \in C_0 \subset C, \end{bmatrix}$$

where $f_\Phi : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$ is defined by

$$f_\Phi(x,u) := f(x,u) - \nabla\Phi(x), \quad \forall (x,u) \in \mathbb{R}^n \times \mathbb{R}^m. \tag{5}$$

Therefore, throughout this paper we will indistinguishably employ the given form of (D) expressed in terms of $\partial\varphi$ and f, or its equivalent form displayed in Remark 3.3 in terms of $N_C(\cdot)$ and the function f_Φ. Note that assumptions (A1)-(A3) imply that, for $\bar{M} := M + K$, f_Φ satisfies the following properties:

(A1)$_\Phi$: The function f_Φ is $\bar{M}$-Lipschitz on $C \times (\mathbb{U} + \tilde{\rho}\bar{B})$ with $\|f_\Phi(x,u)\| \le \bar{M}$ for all $(x,u) \in C \times (\mathbb{U} + \tilde{\rho}\bar{B})$.

We define $\mathcal{U}$ to be

$$\mathcal{U} := \{u : [0,1] \to \mathbb{R}^m : u \text{ is measurable and } u(t) \in U(t),\ t \in [0,1] \text{ a.e.}\}.$$

Remark 3.4 Using [39, Lemma 4.3], it is easy to see that the assumptions (A1)-(A3) and the boundedness of C by some $M_C > 0$ yield that any solution x of (D) corresponding to $(x_0, u) \in C_0 \times \mathcal{U}$ satisfies

$$x(t) \in C, \quad \forall t \in [0,1]; \quad \|x\|_\infty \le M_C; \quad \text{and} \quad \|\dot{x}\|_\infty \le 2\bar{M}. \tag{6}$$

For given $x(\cdot)$: $[0,1] \to \mathbb{R}^n$, we use the following notations throughout this paper: $I^0(x) := \{t \in [0,1] : x(t) \in \text{bdry}\, C\}$ and $I(x) := [0,1] \setminus I^0(x)$.

The next result characterizes the solutions of (D) in terms of the solutions of a standard control system containing an *extra* control ξ that satisfies the mixed control-state *degenerate* constraint, $\xi(t)\psi(x(t)) = 0$. The sufficiency part is straightforward and was used in [39], while the necessary part follows from applying Filippov selection theorem ([38, Theorem 2.3.13]).

Lemma 3.5 [34, 39] *Assume that* (A1)-(A3) *hold. Let* $u \in \mathfrak{U}$ *and* $x \in AC([0,1];\mathbb{R}^n)$ *with* $x(0) \in C_0$ *and* $x(t) \in C$ *for all* $t \in [0,1]$. *Then,* x *is a solution for* (D) *corresponding to the control* u *if and only if there exists a nonnegative measurable function* ξ *supported on* $I^0(x)$ *such that* (x,u,ξ) *satisfies*

$$\dot{x}(t) = f_\Phi(x(t), u(t)) - \xi(t)\nabla\psi(x(t)), \quad t \in [0,1] \text{ a.e.} \tag{7}$$

In this case, the nonnegative function ξ *supported in* $I^0(x)$ *with* (x,u,ξ) *satisfying* (7)*, is unique, belongs to* $L^\infty([0,1];\mathbb{R}^+)$*, and*

$$\begin{cases} \xi(t) = 0 & \textit{for } \ t \in I^-(x), \\ \xi(t) = \frac{\|\dot{x}(t) - f_\Phi(x(t),u(t))\|}{\|\nabla\psi(x(t))\|} \in \left[0, \frac{\bar{M}}{2\eta}\right] & \textit{for } \ t \in I^0(x) \text{ a.e.}, \\ \|\xi\|_\infty \le \frac{\bar{M}}{2\eta}. \end{cases} \tag{8}$$

Throughout the paper we shall employ the following notations, where η and $\bar{M}$ are the constants given in (A2.2) and (A1)$_\Phi$, respectively.

- $(\gamma_k)_k$ is a sequence satisfying

$$\gamma_k > \frac{2\bar{M}}{\eta} \text{ for all } k \in \mathbb{N}, \text{ and } \gamma_k \xrightarrow[k\longrightarrow\infty]{} \infty. \tag{9}$$

- The sequence $(\alpha_k)_k$ is defined by

$$\alpha_k := \frac{\ln\left(\frac{\eta\gamma_k}{2\bar{M}}\right)}{\gamma_k}, \quad k \in \mathbb{N} \tag{10}$$

By (9) and (10), we have that

$$\gamma_k e^{-\alpha_k\gamma_k} = \frac{2\bar{M}}{\eta}, \quad \alpha_k > 0, \quad \alpha_k \searrow \text{ and } \lim_{k\longrightarrow\infty} \alpha_k = 0. \tag{11}$$

- The sequence $(\rho_k)_k$ is defined by $\rho_k := \frac{\alpha_k}{\eta}$ for all $k \in \mathbb{N}$. By (11) we have that $\rho_k > 0$ for all $k \in \mathbb{N}$, $\rho_k \searrow$ and $\lim_{k\longrightarrow\infty} \rho_k = 0$.
- For $k \in \mathbb{N}$, we define the set

$$C(k) := \{x \in C : \psi(x) \leq -\alpha_k\}. \tag{12}$$

The system (D_{γ_k}) is defined as

$$(D_{\gamma_k}) \begin{cases} \dot{x}(t) = f_\Phi(x(t), u(t)) - \gamma_k e^{\gamma_k\psi(x(t))}\nabla\psi(x(t)) \text{ a.e. } t \in [0, 1], \\ x(0) \in C. \end{cases}$$

An important property shown in [34] is the invariance of C for the dynamic (D_{γ_k}), see [34, Lemma 4.1]. This fact is behind disposing of the state constraint in (D_{γ_k}), which represents a good approximation for (D) (see Theorem 3.9 and Corollary 3.12).

Lemma 3.6 ([34], Invariance of C and uniform convergence) *Let* (A1)-(A3) *be satisfied. Then, for each k, the system (D_{γ_k}) with given $x(0) = c_{\gamma_k} \in C$ and $u_{\gamma_k} \in \mathcal{U}$, has a unique solution $x_{\gamma_k} \in W^{1,2}([0, 1]; \mathbb{R}^n)$ such that $x_{\gamma_k}(t) \in C$ for all $t \in [0, 1]$, and, for $\alpha_0 > 0$ a bound of $(c_{\gamma_k})_k$ we have*

$$\|x_{\gamma_k}\|_\infty \leq \alpha_0 + \sqrt{\bar{M}^2 + 2} \quad \textit{and} \quad \int_0^1 \|\dot{x}_{\gamma_k}(t)\|^2 dt \leq \bar{M}^2 + 2. \tag{13}$$

Hence, being equicontinuous and uniformly bounded, $(x_{\gamma_k})_k$ admits a subsequence that converges uniformly to some $x \in W^{1,2}([0, 1]; \mathbb{R}^n)$ whose values are in C and whose derivative $\dot{x}_{\gamma_k}$ converges weakly in L^2 to $\dot{x}$.

The properties of the sets $C(k)$ and the role of the sequence $(\rho_k)_k$ are established in [34, Theorem 3.1 and Remark 3.6]. We enlist here the items that deem important for this paper when constructing the initial constraint set for the approximating problems (P_{γ_k}).

Theorem 3.7 ([34], Properties of $(C(k))_k$) *Under* (A2.1)-(A2.3)*, the following assertions hold*:

(*i*) *For all k, the set $C(k) \subset \operatorname{int} C$ and is compact, and, for k sufficiently large,*

- $\operatorname{bdry} C(k) = \{x \in \mathbb{R}^n : \psi(x) = -\alpha_k\}$ *and* $\operatorname{int} C = \{x \in \mathbb{R}^n : \psi(x) < -\alpha_k\}$;
- $(C(k))_k$ *is a nondecreasing sequence whose Painlevé-Kuratowski limit is C.*

(ii) There exist $r_o > 0$ *and* $\bar{k} \in \mathbb{N}$ *such that*

$$\left[C \cap \bar{B}_{r_o}(c)\right] - \rho_k \frac{\nabla\psi(c)}{\|\nabla\psi(c)\|} \subset \operatorname{int} C(k), \quad \forall k \geq \bar{k} \quad \textit{and} \quad \forall c \in \operatorname{bdry} C. \tag{14}$$

(iii) For $c \in \operatorname{int} C$*, there exist* $\hat{k}_c \in \mathbb{N}$ *and* $\hat{r}_c > 0$ *satisfying*

$$\bar{B}_{\hat{r}_c}(c) \subset \operatorname{int} C(\hat{k}_c) \subset \operatorname{int} C(k), \quad \forall k \geq \hat{k}_c. \tag{15}$$

Remark 3.8 From Theorem 3.7, it follows that for any $c \in C$, there exists a sequence $(c_k)_k$ such that, for k large enough, $c_k \in \operatorname{int} C(k)$ and $c_k \longrightarrow c$. Indeed, for $c \in \operatorname{bdry} C$, take $c_k := c - \rho_k \frac{\nabla\psi(c)}{\|\nabla\psi(c)\|}$ for all k, and for $c \in \operatorname{int} C$, take $c_k = c$ for all k.

The following theorem will be used repeatedly in this paper. It is a special case of [34, Theorem 4.1 & Lemma 4.2]. It provides a sufficient condition for the uniform limit x of the solution x_{γ_k} of (D_{γ_k}) to be a solution of (D), and it connects the multiplier function ξ corresponding to x, via Lemma 3.5, to the positive continuous penalty multiplier ξ_{γ_k}, associated with x_{γ_k} and defined by

$$\xi_{\gamma_k}(\cdot) := \gamma_k e^{\gamma_k \psi(x_{\gamma_k}(\cdot))}. \tag{16}$$

Theorem 3.9 ([34], $(D_{\gamma_k})_k$ & ξ_{γ_k} approximate (D) & ξ) *Assume that* (A1)-(A4.1) *hold. Let* x_{γ_k} *be the solution of* (D_{γ_k}) *corresponding to* $(c_{\gamma_k}, u_{\gamma_k})$*, as in* Lemma 3.6*, and* $x \in W^{1,2}([0,1];\mathbb{R}^n)$ *be its uniform limit. Then, the following statements are valid*:

- (i) *The sequence* $(\xi_{\gamma_k})_k$ *admits a subsequence, we do not relabel, that converges weakly in* L^2 *to a nonnegative function* $\xi \in L^2$ *supported on* $I^0(x)$*.*
- (ii) *If for some* $u \in \mathcal{U}$*, the sequence* $u_{\gamma_k}(t) \xrightarrow{\text{a.e. } t} u(t)$*, then* x *is the unique solution of* (D) *corresponding to* (x_0, u)*, and* (x, u, ξ) *satisfies equations* (7)-(8)*. In particular,* $\xi \in L^\infty([0,1];\mathbb{R}^+)$ *and is supported on* $I^0(x)$*.*

Remark 3.10 Note that when establishing Theorem 3.9(ii) in [34], the arguments used to prove that (x, u, ξ) satisfies (7) are independent of having ξ_{γ_k} defined through (16), and hence, this proof is valid for ξ_{γ_k} being any sequence of L^2-functions converging weakly in L^2 to ξ. Therefore, we have that (x, u, ξ) satisfies (7) whenever $(x_j, u_j, \xi_j)_j$ is a sequence solving (7) with x_j converging uniformly to x, $u_j(t)$ converging pointwise a.e. to $u(t)$, and ξ_j converging weakly in L^2 to ξ.

The following result is extracted from [34, Theorem 5.1], in which more properties are derived. It reveals the significance of initiating in Theorem 3.9 the solutions x_{γ_k} of (D_{γ_k}) from the subset $C(k)$, defined in (12).

Theorem 3.11 ([34], $x_{\gamma_k} \in C(k)$, $\dot{x}_{\gamma_k}$ & ξ_{γ_k} bounded) *Assume* (A1)-(A4.1) *hold. Let* $(c_{\gamma_k})_k$ *be a sequence such that* $c_{\gamma_k} \in C(k)$*, for* k *sufficiently large. Then there exists* $k_o \in \mathbb{N}$ *such that for all sequences* $(u_{\gamma_k})_k$ *in* $\mathcal{U}$ *and for all* $k \geq k_o$*, the solution* x_{γ_k} *of* (D_{γ_k}) *corresponding to* $(c_{\gamma_k}, u_{\gamma_k})$ *satisfies*:

- (i) $x_{\gamma_k}(t) \in C(k) \subset \operatorname{int} C$ *for all* $t \in [0,1]$.
- (ii) $0 \leq \xi_{\gamma_k}(t) \leq \frac{2\bar{M}}{\eta}$ *for all* $t \in [0,1]$.
- (iii) $\|\dot{x}_{\gamma_k}(t)\| \leq \bar{M} + \frac{2\bar{M}\bar{M}_\psi}{\eta}$ *for* a.e. $t \in [0,1]$.

The next result is a simplified version of [34, Corollary 5.1]. It is the converse of Theorem 3.9, as it confirms that any given solution of (D) is approximated by a solution of (D_{γ_k}) that remains in the *interior* of C and enjoys all the properties displayed in Theorem 3.11.

Corollary 3.12 ([34], Solutions of (D) are approximated by sequences in $C(k)$) *Assume that* (A1)-(A4.1) *are satisfied. Let $\bar{x}$ be the solution of (D) corresponding to $(\bar{x}(0), \bar{u}) \in C_0 \times \mathcal{U}$. Consider $(\bar{c}_{\gamma_k})_k$ the sequence in* Remark 3.8 *that converges to $c := \bar{x}(0)$, and $\bar{x}_{\gamma_k}$ the solution of (D_{γ_k}) corresponding to $(\bar{c}_{\gamma_k}, \bar{u})$. Then, there exists $\hat{k}_o \in \mathbb{N}$ such that $\bar{x}_{\gamma_k}$ and its associated $\bar{\xi}_{\gamma_k}$ via* (16) *satisfy the conclusions (i)-(iii) of* Theorem 3.11 *for all $k \geq \hat{k}_o$, and the following holds true: The sequence $\bar{x}_{\gamma_k}$ admits a subsequence, we do not relabel, that converges uniformly to $\bar{x}$, the corresponding subsequence for $\bar{\xi}_{\gamma_k}$ converges weakly in L^2 to some $\bar{\xi} \in L^\infty$, and $(\bar{x}, \bar{u}, \bar{\xi})$ satisfies* (7)-(8)*. That is, $\bar{\xi}$ is the unique function corresponding to $(\bar{x}, \bar{u})$ via* Lemma 3.5.

4 Main Results

This section consists of the main results of this paper, namely, the strong approximation of (D) by (D_{γ_k}) whenever the controls are $W^{1,2}$-bounded (Theorem 4.1 and Corollary 4.2), an existence theorem for an optimal solution of (P) (Theorem 4.4), a strong converging continuous approximation for (P) (Theorem 4.7), and nonsmooth necessary optimality conditions in the form of weak-Pontryagin-type maximum principle (Theorem 4.10).

4.1 (D_{γ_k}) Strongly Approximates (D) with $W^{1,2}$-Controls

The following theorem constitutes the backbone of this paper. It shows that, when the underlying control space is $\mathcal{W}$ (defined in (1)) and $(\|\dot{u}_{\gamma_k}\|_2)_k$ is bounded, the velocities $\dot{x}_{\gamma_k}$ and the functions ξ_{γ_k} corresponding to the approximating sequence x_{γ_k} in Theorem 3.11, converge *strongly* in L^2 to, respectively, $\dot{x}$ and ξ, the functions obtained in Theorem 3.9. The proof of this theorem is postponed to Section 5.

Theorem 4.1 (Strong convergence of the velocity sequence $\dot{x}_{\gamma_k}$) *Let the assumptions* (A1)-(A4.2) *be satisfied. Consider a sequence x_{γ_k} solving (D_{γ_k}) for some $(c_{\gamma_k}, u_{\gamma_k})$, where $c_{\gamma_k} \in C$, $c_{\gamma_k} \longrightarrow x_0 \in C_0$, $u_{\gamma_k} \in \mathcal{W}$, and $(\|\dot{u}_{\gamma_k}\|_2)_k$ is bounded. Denote by (x, ξ) the pair in $W^{1,2} \times L^2$ obtained via* Lemma 3.6 *and* Theorem 3.9(i) *such that a subsequence (not relabeled) of $(x_{\gamma_k}, \xi_{\gamma_k})$ has x_{γ_k} converging uniformly in the set C to x and $(\dot{x}_{\gamma_k}, \xi_{\gamma_k})$ converging weakly in L^2 to $(\dot{x}, \xi)$. Then, the following hold:*

(i) *There exist a subsequence (not relabeled) of u_{γ_k}, and $u \in \mathcal{W}$ such that $(x_{\gamma_k}, u_{\gamma_k})$ converges uniformly to (x, u), and $(\dot{x}_{\gamma_k}, \dot{u}_{\gamma_k}, \xi_{\gamma_k})$ converges weakly in L^2 to $(\dot{x}, \dot{u}, \xi)$. The function x is the unique solution to (D) corresponding to (x_0, u), and (x, u, ξ) satisfies* (7)-(8)*. In particular, $\xi \in L^\infty$ and is supported on $I^0(x)$.*

(ii) *Assume that $c_{\gamma_k} \in C(k)$, for k large. Then, in addition to the conclusions in* Theorem 3.11*, the following holds: The sequence $(\dot{x}_{\gamma_k}, \xi_{\gamma_k})$ is in $W^{1,2}([0, 1]; \mathbb{R}^n) \times W^{2,2}([0, 1]; \mathbb{R}^+)$, has uniform bounded variations, and admits a subsequence, not relabeled, that converges pointwise, and hence, strongly in L^2 to $(\dot{x}, \xi)$, with $\dot{x} \in BV([0, 1]; \mathbb{R}^n)$ and $\xi \in BV([0, 1]; \mathbb{R}^+)$. In this case,* (7)-(8) *hold for all $t \in [0, 1]$, and $x_{\gamma_k} \longrightarrow x$ strongly in the norm topology of $W^{1,2}([0, 1]; \mathbb{R}^n)$.*

Applying Theorem 4.1(ii) to $\bar{c}_{\gamma_k}$, $u_{\gamma_k} := \bar{u}$, $\bar{x}_{\gamma_k}$, and $\bar{\xi}_{\gamma_k}$, the function associated to $\bar{x}_{\gamma_k}$ via (16), we obtain the following corollary that shows how the results in Corollary 3.12 are improved when $W^{1,2}$-controls are utilized.

Corollary 4.2 ($(D_{\gamma_k})_k$ strongly approximates (D)) *If, in addition to the assumptions of* Corollary 3.12*, we have that $\bar{x}$ solves (D) for $\bar{u} \in \mathcal{W}$* (not only in $\mathcal{U}$)*, then $\bar{\xi}_{\gamma_k}$, therein, converges pointwise to $\bar{\xi} \in BV([0,1];\mathbb{R}^+)$ with $\bar{\xi}$ satisfying* (53)*, and $\bar{x}_{\gamma_k}$, therein, converges to $\bar{x}$ strongly in the norm topology of $W^{1,2}([0,1];\mathbb{R}^n)$. Moreover, $(\bar{x},\bar{u},\bar{\xi})$ satisfies* (7)-(8) *for all $t \in [0,1]$, and $\dot{\bar{x}} \in BV([0,1];\mathbb{R}^n)$.*

An important consequence of Corollary 4.2 is the following *compactness result* for the solutions of (D), where the controls are restricted to be in $\mathcal{W}$ and $x(1) \in C_1$. We note that this compactness result will be used in the next subsection to prove the existence of an optimal solution for the problem (P).

Proposition 4.3 (Compact trajectories and controls for (D)) *Assume that* (A1)-(A4.3) *hold. Let $(x_j,u_j)_j$ be a sequence in $W^{1,\infty} \times \mathcal{W}$ satisfying (D) with $x_j(1) \in C_1$, for all $j \in \mathbb{N}$, and $(\|\dot{u}_j\|_2)_j$ be bounded. Consider $(\xi_j)_j$ the corresponding sequence in $L^\infty([0,1];\mathbb{R}^+)$ obtained via* Lemma 3.5*, that is, (x_j,u_j,ξ_j) satisfies* (7)-(8)*, for all j. Then there exist a subsequence of $(x_j,u_j,\xi_j)_j$, we do not relabel, and $(x,u,\xi) \in W^{1,\infty}([0,1];\mathbb{R}^n) \times \mathcal{W} \times L^\infty([0,1];\mathbb{R}^+)$ such that $(x_j,u_j)_j$ converges uniformly to (x,u), $(\dot{x}_j,\xi_j)_j$ now converges pointwise to $(\dot{x},\xi) \in BV([0,1];\mathbb{R}^n) \times BV([0,1];\mathbb{R}^+)$, $\dot{u}_j$ converge weakly in L^2 to $\dot{u}$, and (x,u,ξ) satisfies* (7)-(8) *with $x(1) \in C_1$. In particular, (x,u) is admissible for (P) and $(x_j)_j$ converges to x strongly in the norm topology of $W^{1,2}([0,1];\mathbb{R}^n)$.*

Proof Using (6) in Remark 3.4 for the sequence $(x_j)_j$, the boundedness of $(\|\dot{u}_j\|_2)_j$, that $u_j(t)) \in U(t)$ for all $t \in [0,1]$, and that the sets $U(t)$ are compact and uniformly bounded, by (A4.2), then Arzela-Ascoli's theorem produces a subsequence, we do not relabel, of $(x_j,u_j)_j$, that converges uniformly to an absolutely continuous pair (x,u) with $(x(t),u(t)) \in C \times U(t)$ for all $t \in [0,1]$, and $(\dot{x}_j,\dot{u}_j)_j$ converging weakly in L^2 to $(\dot{x},\dot{u})$. Since, for all $j \in \mathbb{N}$, $x_j(0) \in C_0$ and $x_j(1) \in C_1$, then (A4.1) and (A4.3) yield that $x(0) \in C_0$ and $x(1) \in C_1$. Using Corollary 4.2, we obtain $\|\xi_j\|_\infty \le \frac{\bar{M}}{2\eta}$, $\xi_j \in BV([0,1];\mathbb{R}^+)$, and $V_0^1(\xi_j) \le \tilde{M}_2$, where $\tilde{M}_2$ depends on the uniform bound of $(\|\dot{u}_j\|_2)_j$. By Helly's first theorem, $(\xi_j)_j$ convergence *pointwise* to $\xi \in BV([0,1];\mathbb{R}^+)$. On the other hand, Corollary 4.2 also gives that (7) holds for all $t \in [0,1]$, that is,

$$\dot{x}_j(t) = f_\Phi(x_j(t),u_j(t)) - \xi_j(t)\nabla\psi(x_j(t)), \quad \forall t \in [0,1]. \tag{17}$$

Thus, upon taking the pointwise limit as $j \longrightarrow \infty$ in (17), it follows that $(\dot{x}_j)_j$ converges pointwise to its weak L^2-limit $\dot{x}$, and hence, $\dot{x} \in BV([0,1];\mathbb{R}^n)$. As (x_j,u_j) solves (D), (6) yields that $(\|\dot{x}_j\|_\infty)_j$ is uniformly bounded, and hence, $(\dot{x}_j)_j$ converges to $\dot{x}$ *strongly* in L^2.

We now show that $\xi(t)$ is supported in $I^0(x)$. Let $t \in I(x)$ be fixed, that is, $x(t) \in \operatorname{int} C$. Since $(x_j)_j$ converges uniformly to x, then we can find $\delta_o > 0$ and $j_o \in \mathbb{N}$ such that, for all $s \in (t-\delta, t+\delta) \cap [0,1]$ and for all $j \ge j_o$, we have $x_j(s) \in \operatorname{int} C$, and hence, as ξ_j satisfies (8), $\xi_j(s) = 0$. Thus, $\xi_j(s) \longrightarrow 0$ for $s \in (t-\delta_o, t+\delta_o) \cap [0,1]$, and whence, $\xi(t) = 0$, proving that ξ is supported in $I^0(x)$. Therefore, applying Lemma 3.5 to (x,u,ξ), we conclude that (x,u) solves (D) and (x,u,ξ) satisfies (8). □

4.2 Existence of Optimal Solution for (*P*)

Parallel to [6, 8, Theorems 4.1], where a discretization technique is used, the following existence theorem of an optimal solution for the problem (P) is established based on Corollary 4.2.

Theorem 4.4 (Existence of solution for (P)) *Assume hypotheses* (A1)-(A4.3), $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *is lower semicontinuous, and that a minimizing sequence* (x_j, u_j) *for* (P) *exists such that* $(\|\dot{u}_j\|_2)_j$ *is bounded. Suppose that* (P) *has at least one admissible pair* (y_o, v_o) *with* $(y_o(0), y_o(1)) \in \operatorname{dom} g$*, then the problem* (P) *admits a global optimal solution* $(\bar{x}, \bar{u})$ *such that, along a subsequence, we have*

$$x_j \xrightarrow[W^{1,2}([0,1];\mathbb{R}^n)]{\text{strongly}} \bar{x}, \quad u_j \xrightarrow[C([0,1];\mathbb{R}^m)]{\text{uniformly}} \bar{u}, \quad \textit{and} \quad \dot{u}_j \xrightarrow[L^2([0,1];\mathbb{R}^m)]{\text{weakly}} \dot{\bar{u}}.$$

Proof Given that (P) has an admissible pair (y_o, v_o) with $(y_o(0), y_o(1)) \in \operatorname{dom} g$, then $\inf_{(x,u)}(P) < \infty$. As g is lower semicontinuous and all admissible solutions of (P) satisfy $(x(0), x(1)) \in C_0 \times (C_1 \cap C)$, which is compact, we deduce that $\inf_{(x,u)}(P)$ is finite. On the other hand, being admissible for (P), the minimizing sequence $(x_j, u_j)_j$ satisfies (D) with $x_j(1) \in C_1$. Hence, using that the sequence $(\|\dot{u}_j\|_2)_j$ is bounded, Proposition 4.3 implies the existence of $(\bar{x}, \bar{u}) \in W^{1,\infty}([0,1]; \mathbb{R}^n) \times \mathcal{W}$ satisfying (D) and $\bar{x}(1) \in C_1$, with (x_j, u_j) converges uniformly to $(\bar{x}, \bar{u})$, $(\dot{x}_j)_j$ converges *strongly* in L^2 to $\dot{\bar{x}} \in BV([0,1]; \mathbb{R}^n)$, and $\dot{u}_j$ converges weakly in L^2 to $\dot{\bar{u}}$. Thus, $(\bar{x}, \bar{u})$ is admissible for (P). Owed to the lower semicontinuity of g and to $(\bar{x}, \bar{u})$ being the uniform limit of the minimizing sequence $(x_j, u_j)_j$, the optimality of the pair $(\bar{x}, \bar{u})$ for the problem (P) follows readily. □

4.3 Continuous Approximation for (*P*)

On the journey of seeking for an optimal process $(\bar{x}, \bar{u})$ of (P) a *continuous* approximations consisting of *optimal* solutions for properly-designed standard control problems, it is important that the convergence to $(\bar{x}, \bar{u})$ be *strong* in the norm topology of the considered space, namely, the space $W^{1,2}([0,1]; \mathbb{R}^n) \times \mathcal{W}$. Corollary 4.2 already answered this question for the $W^{1,2}$-strong approximation of a solution $(\bar{x}, \bar{u})$ of (D) by solutions of (D_{γ_k}), in which the *same* control $\bar{u}$ is used. However, $\bar{u}$ may not necessarily be optimal for approximating optimal control problems over (D_{γ_k}).

In this subsection, we approximate the problem (P) by a certain sequence of optimal control problems over (D_{γ_k}) with *special* initial and final state endpoints constraints ($C_0(k) \subset C(k)$ and $C_1(k)$ in a band around C_1), and with an objective function particularly crafted so that an optimal control, u_{γ_k}, exists and has $(\|\dot{u}_{\gamma_k}\|_2)_k$ uniformly bounded, and hence, the *strong* convergence of the optimal state velocities shall be deduced from Theorem 4.1. The necessary optimality conditions for (P) are then established by taking the limit of the optimality conditions for the corresponding approximating problem.

For given $\delta > 0$ and $z \in C([0,1]; \mathbb{R}^s)$, we define the projection on $\mathbb{R}^s$ of the closed δ-*tube* around z by $\bar{\mathbb{B}}_\delta(z) := \bigcup_{t \in [0,1]} \bar{B}_\delta(z(t))$.

Let $(\bar{x}, \bar{u}) \in W^{1,2}([0,1]; \mathbb{R}^n) \times \mathcal{W}$ be a $W^{1,2}$-local minimizer for (P) with associated δ. We fix $\delta_o > 0$ such that

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

where $r_o > 0$ is the constant in Theorem 3.7(ii), and $\hat{r}_{\bar{x}(0)} > 0$ with $\hat{k}_{\bar{x}(0)} \in \mathbb{N}$ are the constants in Theorem 3.7(iii) corresponding to $c := \bar{x}(0)$.

In the remaining part below, we will assume that f satisfies the following *local* version of (A1):

$$\exists\,\tilde{\rho} > 0 \text{ such that } f \text{ is Lipschitz on } [C \cap \bar{\mathbb{B}}_\delta(\bar{x})] \times [(\mathbb{U} + \tilde{\rho}\bar{B}) \cap \bar{\mathbb{B}}_\delta(\bar{u})]. \qquad (*)$$

Note that under the assumption (∗), the function f can be extended to a *globally* Lipschitz function $\tilde{f}\colon \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ by applying [23, Theorem 1] to each component of f. Since in the rest of this section we only consider local optimality notions, then, without loss of generality, we shall use the function f instead of $\tilde{f}$. Hence, when in this section f is assumed to satisfy (∗), it is implied that f also satisfies assumption (A1).

We proceed to suitably-formulate a sequence of approximating problems (P_{γ_k}) and show that its *optimal solutions strongly* converges in $W^{1,2}([0,1];\mathbb{R}^n) \times \mathcal{W}$ to the $W^{1,2}$-local minimizer $(\bar{x}, \bar{u})$ of (P). This naturally requires the domain of the approximating problem (P_{γ_k}) to be in $W^{1,2}([0,1];\mathbb{R}^n) \times \mathcal{W}$. The initial state constraint is taken to be $x(0) \in C_0(k)$, where $C_0(k)$ is the sequence of sets defined by

$$C_0(k) := \begin{cases} C_0 \cap \bar{B}_{\delta_o}(\bar{x}(0))\,, \ \forall k \in \mathbb{N}, & \text{if } \bar{x}(0) \in \operatorname{int} C, \\ \left[C_0 \cap \bar{B}_{\delta_o}(\bar{x}(0))\right] - \rho_k \frac{\nabla\psi(\bar{x}(0))}{\|\nabla\psi(\bar{x}(0))\|}, \ \forall k \in \mathbb{N}, & \text{if } \bar{x}(0) \in \operatorname{bdry} C. \end{cases} \qquad (18)$$

and the final state constraint is $x(1) \in C_1(k)$, where

$$C_1(k) := \left[\left(C_1 \cap \bar{B}_{\delta_o}(\bar{x}(1))\right) - \bar{x}(1) + \bar{x}_{\gamma_k}(1)\right] \cap C, \quad k \in \mathbb{N},$$

in which $\bar{x}_{\gamma_k}$ is the solution of (D_{γ_k}) corresponding to $(\bar{c}_{\gamma_k}, \bar{u})$, where $\bar{c}_{\gamma_k}$ in $C_0(k) \cap \operatorname{int} C(k)$, for k large, and is defined via Remark 3.8 for $c := \bar{x}(0)$, that is,

$$\bar{c}_k := \begin{cases} \bar{x}(0), \ \forall k \in \mathbb{N}, & \text{if } \bar{x}(0) \in \operatorname{int} C, \\ \bar{x}(0) - \rho_k \frac{\nabla\psi(\bar{x}(0))}{\|\nabla\psi(\bar{x}(0))\|}, \ \forall k \in \mathbb{N}, & \text{if } \bar{x}(0) \in \operatorname{bdry} C. \end{cases}$$

Note that $C_0(k)$ and $C_1(k)$ are *closed*, for $k \in \mathbb{N}$. On the other hand, as $\rho_{\gamma_k} \longrightarrow 0$, we have $\bar{c}_{\gamma_k} \longrightarrow \bar{x}(0)$, Corollary 3.12 yields that the sequence $\bar{x}_{\gamma_k}$ converges in C uniformly to $\bar{x}$, and hence, $\bar{x}_{\gamma_k}(1) \longrightarrow \bar{x}(1)$. Add to this that in $C_0(k)$, $\rho_k \longrightarrow 0$, then, for the $\tilde{\rho}$ in (∗), we have that, for k sufficiently large,

$$C_i(k) \subset \underbrace{\left[\left(C_i \cap \bar{B}_\delta(\bar{x}(i))\right) + \tilde{\rho}\bar{B}\right] \cap C}_{\tilde{C}_i(\delta)}, \quad \text{for } \ i = 0, 1, \qquad (19)$$

$$\text{and } \lim_{k\to\infty} C_0(k) = C_0 \cap \bar{B}_{\delta_o}(\bar{x}(0)) \ \& \ \lim_{k\to\infty} C_1(k) = C \cap C_1 \cap \bar{B}_{\delta_o}(\bar{x}(1)). \qquad (20)$$

Remark 4.5 Notice that we can show that, for k large enough, we have $C_0(k) \subset C(k)$, and hence, by Theorem 3.11, any solution of (D_{γ_k}) corresponding to $(c_{\gamma_k}, u_{\gamma_k})$ with $c_{\gamma_k} \in C_0(k)$ and $u_{\gamma_k} \in \mathcal{U}$, satisfies the conditions (i)-(iii) of this theorem. Indeed:

- For $\bar{x}(0) \in \operatorname{int} C$, use that $\delta_o \le \hat{r}_{\bar{x}(0)}$ and (15) we get

$$\bar{B}_{\hat{r}_{\bar{x}(0)}}(\bar{x}(0)) \subset \operatorname{int} C(k) \subset C(k), \quad \forall k \ge \hat{k}_{\bar{x}(0)}.$$

  This gives that

$$C_0(k) := C_0 \cap \bar{B}_{\delta_o}(\bar{x}(0)) \subset \bar{B}_{\hat{r}_{\bar{x}(0)}}(\bar{x}(0)) \subset C(k), \quad \forall k \ge \hat{k}_{\bar{x}(0)}.$$

- For $\bar{x}(0) \in \text{bdry}\, C$, use that $\delta_o \leq r_o$ and $C_0(k)$ is the nonempty set defined by the second equation of (18), to get via (14) that

$$C_0(k) \subset \text{int}\, C(k) \subset C(k), \quad \forall k \geq \bar{k}.$$

Remark 4.6 Using the local property of the normal cones, see [11, Proposition 1.1.5(b)], the cones $N^L_{C_0(k)}(c)$ and $N^L_{C_1(k)}(d)$ can be evaluated in terms of $N^L_{C_0}$ and $N^L_{C_1}$, respectively, as follows

$$N^L_{C_0(k)}(c) = \begin{cases} N^L_{C_0}(c), & \text{if } \bar{x}(0) \in \text{int}\, C, \text{ and} \\ & c \in B_{\delta_o}(\bar{x}(0)), \\ N^L_{C_0}\left(c + \rho_k \frac{\nabla\psi(\bar{x}(0))}{\|\nabla\psi(\bar{x}(0))\|}\right), & \text{if } \bar{x}(0) \in \text{bdry}\, C, \text{ and} \\ & \left(c + \rho_k \frac{\nabla\psi(\bar{x}(0))}{\|\nabla\psi(\bar{x}(0))\|}\right) \in B_{\delta_o}(\bar{x}(0)). \end{cases} \tag{21}$$

$$N^L_{C_1(k)}(d) = N^L_{C_1}(d + \bar{x}(1) - \bar{x}_{\gamma_k}(1)), \quad \forall d \in (\text{int}\, C) \cap B_{\delta_o}(\bar{x}(1)). \tag{22}$$

We introduce the following sequence of approximating problems:

(P_{γ_k}): Minimize

$$J(x, y, z, u) := g(x(0), x(1)) + \tfrac{1}{2}\left(\|u(0) - \bar{u}(0)\|^2 + z(1) + \|x(0) - \bar{x}(0)\|^2\right)$$

over $(x, y, z, u) \in W^{1,2}([0, 1]; \mathbb{R}^n) \times AC([0, 1]; \mathbb{R}) \times AC([0, 1]; \mathbb{R}) \times \mathcal{W}$

such that

$$\begin{cases} (\tilde{D}_{\gamma_k}) \left[\begin{array}{l} \dot{x}(t) = f_\Phi(x(t), u(t)) - \gamma_k e^{\gamma_k \psi(x(t))} \nabla\psi(x(t)), \quad t \in [0, 1] \text{ a.e.,} \\ \dot{y}(t) = \|\dot{x}(t) - \dot{\bar{x}}(t)\|^2, \quad t \in [0, 1] \text{ a.e.,} \\ \dot{z}(t) = \|\dot{u}(t) - \dot{\bar{u}}(t)\|^2, \quad t \in [0, 1] \text{ a.e.,} \\ (x(0), y(0), z(0)) \in C_0(k) \times \{0\} \times \{0\}, \end{array}\right. \\ x(t) \in \bar{B}_\delta(\bar{x}(t)) \text{ and } u(t) \in U(t) \cap \bar{B}_\delta(\bar{u}(t)), \quad \forall t \in [0, 1], \\ (x(1), y(1), z(1)) \in C_1(k) \times [-\delta, \delta] \times [-\delta, \delta]. \end{cases}$$

Note that Lemma 3.6 and the constraints on $u(\cdot)$ confirm that $(P_{\gamma_k})_k$ is actually equivalent to having therein $(x, u) \in AC([0, 1]; \mathbb{R}^n) \times AC([0, 1]; \mathbb{R}^m)$.

Now we are ready to state our *continuous* approximation result, which is parallel to the corresponding result in [6–8, 14], where discrete approximations are used. The proof of this approximation result is presented in Section 5.

Theorem 4.7 ((P_{γ_k}) approximates (P)) *Let $(\bar{x}, \bar{u})$ be a $W^{1,2}$-local minimizer (P) with associated $\bar{\xi} \in L^\infty$ via* Lemma 3.5. *Assume that* (A2)-(A4.3) *hold, and for some $\tilde{\rho} > 0$, f is Lipschitz on $[C \cap \bar{\mathbb{B}}_\delta(\bar{x})] \times [(\mathbb{U} + \tilde{\rho}\bar{B}) \cap \bar{\mathbb{B}}_\delta(\bar{u})]$ and g is continuous on $\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)$. Then for k sufficiently large, the problem (P_{γ_k}) has an optimal solution $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ such that, for ξ_{γ_k} defined in* (16)*, we have, along a subsequence, we do not relabel, that*

$$(x_{\gamma_k}, u_{\gamma_k}) \xrightarrow[W^{1,2}\times\mathcal{W}]{\text{strongly}} (\bar{x}, \bar{u}), \quad (y_{\gamma_k}, z_{\gamma_k}) \xrightarrow[W^{1,1}([0,1];\mathbb{R}^+\times\mathbb{R}^+)]{\text{strongly}} (0, 0), \quad \xi_{\gamma_k} \xrightarrow[L^2([0,1];\mathbb{R}^+)]{\text{strongly}} \bar{\xi},$$

all the conclusions of Theorem 3.11 *hold, including that $x_{\gamma_k}(t) \in \text{int}\, C$ for all $t \in [0, 1]$, and for all k sufficiently large,*

$$x_{\gamma_k}(i) \in \left[\left(C_i \cap \bar{B}_{\delta_o}(\bar{x}(i))\right) + \tilde{\rho}B\right] \cap (\text{int}\, C) \subset \text{int}\, \tilde{C}_i(\delta), \quad \text{for } i = 0, 1.$$

Moreover, $\dot{\bar{x}} \in BV([0,1];\mathbb{R}^n)$, $\bar{\xi} \in BV([0,1];\mathbb{R}^+)$, *and* (7)-(8) *are valid at* $(\bar{x}, \bar{u}, \bar{\xi})$ *for all* $t \in [0,1]$.

We proceed to rewrite the problems (P_{γ_k}) as an optimal control problem with state constraints. Given $(\bar{x}, \bar{u}) \in W^{1,2}([0,1];\mathbb{R}^n) \times \mathcal{W}$ a $W^{1,2}$-local minimizer for (P), for $\bar{v} := \dot{\bar{u}}$, (P_{γ_k}) is reformulated in the following way:

(P_{γ_k}): Minimize
$g(x(0), x(1)) + \frac{1}{2}\left(\|u(0) - \bar{u}(0)\|^2 + z(1) + \|x(0) - \bar{x}(0)\|^2\right)$ over
$(x, y, z, u) \in AC([0,1];\mathbb{R}^n)\times AC([0,1];\mathbb{R})\times AC([0,1];\mathbb{R})\times AC([0,1];\mathbb{R}^m)$
and measurable functions $v\colon [0,1] \longrightarrow \mathbb{R}^m$ such that

$$\begin{cases} \left[\begin{array}{l} \dot{x}(t) = f_{\Phi}(x(t), u(t)) - \gamma_k e^{\gamma_k \psi(x(t))} \nabla\psi(x(t)), \;\; t \in [0,1] \text{ a.e.}, \\ \dot{u}(t) = v(t), \;\; t \in [0,1] \text{ a.e.}, \\ \dot{y}(t) = \| f_{\Phi}(x(t), u(t)) - \gamma_k e^{\gamma_k \psi(x(t))} \nabla\psi(x(t)) - \dot{\bar{x}}(t)\|^2, \;\; t \in [0,1] \text{ a.e.}, \\ \dot{z}(t) = \|v(t) - \bar{v}(t)\|^2, \;\; t \in [0,1] \text{ a.e.}, \end{array}\right. \\ x(t) \in \bar{B}_{\delta}(\bar{x}(t)) \text{ and } u(t) \in U(t) \cap \bar{B}_{\delta}(\bar{u}(t)), \quad \forall t \in [0,1], \\ (x(0), u(0), y(0), z(0)) \in C_0(k) \times \mathbb{R}^m \times \{0\} \times \{0\}, \\ (x(1), u(1), y(1), z(1)) \in C_1(k) \times \mathbb{R}^m \times [-\delta, \delta] \times [-\delta, \delta]. \end{cases}$$

In the following proposition we apply to the above sequence of reformulated problems (P_{γ_k}), the nonsmooth Pontryagin maximum principle for optimal control problems with *multiple* state constraints (see e.g., [38, page 331] and [38, p.332]). For this purpose, (x, y, z, u) is the state function in (P_{γ_k}) and v is the control. Thus, $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ is the optimal state, where $(x_{\gamma_k}, u_{\gamma_k})$ is obtained from Theorem 4.7, $y_{\gamma_k}(t) := \int_0^t \|\dot{x}_{\gamma_k}(s) - \dot{\bar{x}}(s)\|^2\, ds$, $z_{\gamma_k}(t) := \int_0^t \|\dot{u}_{\gamma_k}(s) - \dot{\bar{u}}(s)\|^2\, ds$, and $v_{\gamma_k} = \dot{u}_{\gamma_k}$ is the optimal control. Hence, the function $f(\cdot,\cdot)$ is required to be Lipschitz *near* $(x_{\gamma_k}, u_{\gamma_k})$, which follows from $(*)$, since $x_{\gamma_k}(t) \in \operatorname{int} C$ and $(x_{\gamma_k}, u_{\gamma_k})$ converges uniformly to $(\bar{x}, \bar{u})$ (see Theorem 4.7). Furthermore, as the objective function g must be Lipschitz *near* $(x_{\gamma_k}(0), x_{\gamma_k}(1))$, we introduce the following local assumption on g in which $\tilde{C}_0(\delta)$ and $\tilde{C}_1(\delta)$ are defined in (19):

$$\exists\, \tilde{\rho} > 0 \text{ such that } g \text{ is Lipschitz on } \tilde{C}_0(\delta) \times \tilde{C}_1(\delta).$$

On the other hand, the following *constraint qualification* property (CQ) is required. For a given multifunction $F\colon [0,1] \rightrightarrows \mathbb{R}^m$, with nonempty and closed values, and for $h \in C([0,1]; F)$, that is, $h \in C([0,1];\mathbb{R}^m)$ and satisfies $h(t) \in F(t)$ for all $t \in [0,1]$, we say that $F(\cdot)$ satisfies *the constraint qualification at h* if

(CQ) $\quad \operatorname{conv}(\bar{N}^L_{F(t)}(h(t)))$ is pointed for all $t \in [0,1]$.

Here, $\bar{N}^L_{F(t)}(y)$ stands for the graphical closure at (t, y) of the multifunction $(t, y) \mapsto N^L_{F(t)}(y)$, that is, the graph of $\bar{N}^L_{F(\cdot)}(\cdot)$ is the closure of the graph of $N^L_{F(\cdot)}(\cdot)$.

It is worth noting that in [6–8], where also $W^{1,2}$-controls are employed, the control sets $U(t)$ are assumed to be $\mathbb{R}^m$, for all $t \in [0,1]$, and hence, in this case, $F(\cdot) := U(\cdot) \equiv \mathbb{R}^m$ trivially satisfies (CQ) at any h. For the general case where $F(\cdot) \not\equiv \mathbb{R}^m$, the following remark provides important information about (CQ).

Remark 4.8

(i) Let $F\colon [0,1] \rightrightarrows \mathbb{R}^m$ be a lower semicontinuous multifunction with closed and nonempty values. For $d_F(t,x) := d(x, F(t))$, we have from [26, Proposition 2.3] that for $t \in [0,1]$ and $x \in F(t)$, conv $(\bar{N}^L_{F(t)}(x))$ is pointed if and only if $0 \notin \partial^>_x d_F(t,x)$. The notion of $\partial^>_x g(t,x)$ is introduced for a general function $g(t,x)$ by Clarke in [9, p.121]. For $g(t,x) := d_F(t,x)$, it is shown in [26, Corollary 2.2] that

$$\partial^>_x d_F(t,x) =$$
$$\text{conv}\left\{\zeta : \zeta = \lim_{i\longrightarrow\infty} \zeta_i,\ \|\zeta_i\| = 1,\ \zeta_i \in N^P_{F(t_i)}(x_i) \text{ and } (t_i,x_i) \xrightarrow{\text{Gr}\,F} (t,x)\right\},$$

where $(t_i,x_i) \xrightarrow{\text{Gr}\,F} (t,x)$ signifies that $(t_i,x_i) \longrightarrow (t,x)$ with $x_i \in F(t_i)$ for all i. Therefore, for $h \in C([0,1];F)$, we have that F satisfies the constraint qualification (CQ) at h if and only if $0 \notin \partial^>_x d_F(t,h(t))$ for all $t \in [0,1]$. Note that the multifunction $(t,x) \mapsto \partial^>_x d_F(t,x)$ is uniformly bounded with compact and convex values, and has a closed graph.

(ii) Using the proximal normal inequality, see [11, Proposition 1.1.5(a)], one can easily extend the arguments in the proof of [26, Proposition 2.3(d)], to show that if the lower semicontinuous multifunction F has closed and *r-prox-regular* values, for some $r > 0$, (as opposed to convex), then conv $(\bar{N}^L_{F(t)}(\cdot)) = N^P_{F(t)}(\cdot) = N^L_{F(t)}(\cdot) = N_{F(t)}(\cdot)$, and this cone is pointed at $x \in F(t)$ if and only if $F(t)$ is epi-lipschitz at x, see [9, Theorem 7.3.1] and [37, Exercise 9.42]. Hence, a lower semicontinuous multifunction $F\colon [0,1] \rightrightarrows \mathbb{R}^m$ with values that are closed and r-prox-regular, satisfies the constraint qualification (CQ) at $h \in C([0,1];F)$ if and only if $F(t)$ is epi-Lipschitz at $h(t)$, for all $t \in [0,1]$.

(iii) If $F(t) = F$ for all $t \in [0,1]$, where F is closed, then conv $(\bar{N}^L_{F(t)}(\cdot)) = N_F(\cdot)$, and this cone is pointed at $x \in F$ if and only if F is epi-Lipschitz at x. Hence, a constant multifunction F satisfies the constraint qualification (CQ) at $h \in C([0,1];F)$ if and only if F is epi-Lipschitz at $h(t)$ for all $t \in [0,1]$.

Proposition 4.9 (Maximum Principle for approximating problems (P_{γ_k})) *Let $(\bar{x},\bar{u})$ be a $W^{1,2}$-local minimizer for (P). Assume that* (A2)-(A4) *hold, and for some $\tilde{\rho} > 0$, f is Lipschitz on $[C \cap \bar{\mathbb{B}}_\delta(\bar{x})] \times [(\mathbb{U} + \tilde{\rho}\bar{B}) \cap \bar{\mathbb{B}}_\delta(\bar{u})]$ and g is Lipschitz on $\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)$. Consider the optimal sequence $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k}, v_{\gamma_k})$ for (P_{γ_k}) obtained via* Theorem 4.7. *If for k sufficiently large, $U(\cdot)$ satisfies the constraint qualification* (CQ) *at u_{γ_k}, then for k large enough, there exist $\lambda_{\gamma_k} \geq 0$, $p_{\gamma_k} \in AC([0,1];\mathbb{R}^n)$, $q_{\gamma_k} \in AC([0,1];\mathbb{R}^m)$, $\Omega_{\gamma_k} \in NBV([0,1];\mathbb{R}^m)$, $\mu^o_{\gamma_k} \in C^\oplus([0,1];\mathbb{R}^m)$, and a $\mu^o_{\gamma_k}$-integrable function $\beta_{\gamma_k} : [0,1] \longrightarrow \mathbb{R}^m$ such that $\Omega_{\gamma_k}(t) = \int_{[0,t]} \beta_{\gamma_k}(s)\mu^o_{\gamma_k}(ds)$, for all $t \in (0,1]$, and:*

(i) **(The nontriviality condition)** *For all $k \in \mathbb{N}$, we have*

$$\|p_{\gamma_k}(1)\| + \|q_{\gamma_k}\|_\infty + \|\mu^o_{\gamma_k}\|_{\text{T.V.}} + \lambda_{\gamma_k} = 1;$$

(ii) **(The adjoint equation)** *For* a.e. $t \in [0, 1]$,

$$\begin{pmatrix} \dot{p}_{\gamma_k}(t) \\ \dot{q}_{\gamma_k}(t) \end{pmatrix} \in -\left(\partial^{(x,u)} f_\Phi(t, x_{\gamma_k}(t), u_{\gamma_k}(t))\right)^\mathsf{T} p_{\gamma_k}(t) + \begin{pmatrix} \gamma_k e^{\gamma_k \psi(x_{\gamma_k}(t))} \partial^2 \psi(x_{\gamma_k}(t)) p_{\gamma_k}(t) \\ 0 \end{pmatrix} + \begin{pmatrix} \gamma_k^2 e^{\gamma_k \psi(x_{\gamma_k}(t))} \nabla \psi(x_{\gamma_k}(t)) \langle \nabla \psi(x_{\gamma_k}(t)), p_{\gamma_k}(t) \rangle \\ 0 \end{pmatrix}; \tag{23}$$

(iii) **(The transversality equation)**

$$(p_{\gamma_k}(0), -p_{\gamma_k}(1)) \in \lambda_{\gamma_k} \partial^L g(x_{\gamma_k}(0), x_{\gamma_k}(1)) + \left[\left(\lambda_{\gamma_k}(x_{\gamma_k}(0) - \bar{x}(0)) + N^L_{C_0(k)}(x_{\gamma_k}(0))\right) \times N^L_{C_1(k)}(x_{\gamma_k}(1))\right],$$

$$\textit{and} \;\; q_{\gamma_k}(0) = \lambda_{\gamma_k}(u_{\gamma_k}(0) - \bar{u}(0)), \quad -q_{\gamma_k}(1) = \Omega_{\gamma_k}(1);$$

(iv) **(The maximization condition)** *For* a.e. $t \in [0, 1]$,

$$\max_{v \in \mathbb{R}^m} \left\{ \langle q_{\gamma_k}(t) + \Omega_{\gamma_k}(t), v \rangle - \frac{\lambda_{\gamma_k}}{2} \|v - \dot{\bar{u}}(t)\|^2 \right\} \textit{ is attained at } \dot{u}_{\gamma_k}(t);$$

(v) **(The measure properties)**

$$\operatorname{supp}\{\mu^o_{\gamma_k}\} \subset \left\{t \in [0, 1] : (t, u_{\gamma_k}(t)) \in \operatorname{bdry} \operatorname{Gr}\left[U(t) \cap \bar{B}_\delta(\bar{u}(t))\right]\right\}, \textit{ and}$$

$$\beta_{\gamma_k}(t) \in \partial_u^> d(u_{\gamma_k}(t), U(t) \cap \bar{B}_\delta(\bar{u}(t))) \quad \mu^o_{\gamma_k} \text{ a.e., } \textit{with}$$

$$\partial_u^> d(u_{\gamma_k}(t), U(t) \cap \bar{B}_\delta(\bar{u}(t))) \subset \left[\operatorname{conv} \bar{N}^L_{U(t) \cap \bar{B}_\delta(\bar{u}(t))}(u_{\gamma_k}(t)) \cap \left(\bar{B} \setminus \{0\}\right)\right].$$

4.4 Necessary Optimality Conditions for *(P)*

The main result of this subsection is the following theorem which provides necessary optimality conditions for the $W^{1,2}$-local minimizer, $(\bar{x}, \bar{u})$, of (P).

The following notations are used in the statement of the theorem:

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

Theorem 4.10 (Necessary optimality conditions for (P)) *Let* $(\bar{x}, \bar{u})$ *be a* $W^{1,2}$*-local minimizer for* (P)*. Let* $\bar{\xi} \in L^\infty([0, 1]; \mathbb{R}^+)$ *be the function supported on* $I^0(\bar{x})$ *and associated to* $(\bar{x}, \bar{u})$ *via* Lemma 3.5*. Assume that* (A2)-(A4) *hold,* $U(\cdot)$ *satisfies the constraint qualification* (CQ) *at* $\bar{u}$*, and for some* $\tilde{\rho} > 0$*,* f *is Lipschitz on* $[C \cap \bar{\mathbb{B}}_\delta(\bar{x})] \times [(\mathbb{U} + \tilde{\rho}\bar{B}) \cap \bar{\mathbb{B}}_\delta(\bar{u})]$ *and* g *is Lipschitz on* $\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)$*. Then* $\dot{\bar{x}} \in BV([0, 1]; \mathbb{R}^n)$ *and* $\bar{\xi} \in BV([0, 1]; \mathbb{R}^+)$*, and there exist* $\lambda \geq 0$*, an adjoint vector* $\bar{p} \in BV([0, 1]; \mathbb{R}^n)$*, a finite signed Radon measure* $\bar{\nu}$ *on* $[0, 1]$

supported on $I^0(\bar{x})$, L^∞*-functions* $\bar{\zeta}(\cdot)$, $\bar{\theta}(\cdot)$ *and* $\bar{\vartheta}(\cdot)$ *in* $\mathcal{M}_{n\times n}([0,1])$, *an* L^∞*-function* $\bar{\omega}(\cdot)$ *in* $\mathcal{M}_{n\times m}([0,1])$, *such that for* $t \in [0,1]$ a.e.,

$$\big((\bar{\zeta}(t), \bar{\omega}(t)), \bar{\theta}(t), \bar{\vartheta}(t)\big) \in \partial_\ell^{(x,u)} f(\bar{x}(t), \bar{u}(t)) \times \partial_\ell^2 \varphi(\bar{x}(t)) \times \partial_\ell^2 \psi(\bar{x}(t)),$$

and the following hold:

(i) **(The admissible equation)**

(a) $\dot{\bar{x}}(t) = f(\bar{x}(t), \bar{u}(t)) - \nabla_\ell \varphi(\bar{x}(t)) - \bar{\xi}(t)\nabla\psi(\bar{x}(t)), \quad \forall t \in [0,1],$
(b) $\psi(\bar{x}(t)) \le 0, \quad \forall t \in [0,1];$

(ii) **(The nontriviality condition)**

$$\|\bar{p}(1)\| + \lambda = 1;$$

(iii) **(The adjoint equation)** *For any* $h \in C([0,1];\mathbb{R}^n)$, *we have*

$$\int_{[0,1]} \langle h(t), d\bar{p}(t)\rangle = \int_0^1 \Big\langle h(t), \Big(\bar{\theta}(t) - \bar{\zeta}(t)^\mathsf{T}\Big)\, \bar{p}(t)\Big\rangle\, dt + \int_0^1 \bar{\xi}(t)\,\big\langle h(t), \bar{\vartheta}(t) p(t)\big\rangle\, dt + \int_{[0,1]} \langle h(t), \nabla\psi(\bar{x}(t))\rangle\, d\bar{\nu};$$

(iv) **(The complementary slackness conditions)**

(a) $\bar{\xi}(t) = 0, \quad \forall t \in I(\bar{x}),$
(b) $\bar{\xi}(t)\langle \nabla\psi(\bar{x}(t), \bar{p}(t)\rangle = 0, \quad \forall t \in [0,1]$ a.e.;

(v) **(The transversality equation)**

$$(\bar{p}(0), -\bar{p}(1)) \in \lambda \partial_\ell^L g(\bar{x}(0), \bar{x}(1)) + \big[N_{C_0}^L(\bar{x}(0)) \times\ N_{C_1}^L(\bar{x}(1))\big];$$

(vi) **(The weak maximization condition)**

$$\bar{\omega}(t)^\mathsf{T}\bar{p}(t) \in \operatorname{conv} \bar{N}^L_{U(t)\cap \bar{B}_\delta(\bar{u}(t))}(\bar{u}(t)), \quad t \in [0,1] \text{ a.e.}$$

If in addition there exist $\varepsilon_o > 0$ *and* $r > 0$ *such that* $U(t) \cap \bar{B}_{\varepsilon_o}(\bar{u}(t))$ *is* r*-prox-regular for all* $t \in [0,1]$, *then we have*

$$\max \left\{ \big\langle \bar{\omega}(t)^\mathsf{T}\bar{p}(t), u\big\rangle - \tfrac{\|\bar{\omega}(t)^\mathsf{T}\bar{p}(t)\|}{\min\{\varepsilon_o, 2r\}} \|u - \bar{u}(t)\|^2 : u \in U(t) \right\}$$

is attained at $\bar{u}(t)$ *for* $t \in [0,1]$ a.e.

Furthermore, if $C_1 = \mathbb{R}^n$, *then* $\lambda \ne 0$ *and is taken to be* 1, *and the nontriviality condition* (i) *is eliminated.*

Remark 4.11 Condition (vi) of Theorem 4.10 admits simplified forms when $U(\cdot)$ possesses extra properties:

- If $U(t)$ is *r-prox-regular* for all $t \in [0,1]$, then taking $\varepsilon_o \longrightarrow \infty$, the maximization condition (v) reduces to

$$\max \left\{ \big\langle \bar{\omega}(t)^\mathsf{T}\bar{p}(t), u\big\rangle - \tfrac{\|\bar{\omega}(t)^\mathsf{T}\bar{p}(t)\|}{2r} \|u - \bar{u}(t)\|^2 : u \in U(t) \right\}$$

  is attained at $\bar{u}(t)$ for $t \in [0,1]$ a.e.

- If $U(t) \cap \bar{B}_{\varepsilon_o}(\bar{u}(t))$ is *convex* for all $t \in [0,1]$, then taking $r \longrightarrow \infty$, the maximization condition (v) reduces to

$$\max \left\{ \left\langle \bar{\omega}(t)^{\mathsf{T}} \bar{p}(t), u \right\rangle - \tfrac{\|\bar{\omega}(t)^{\mathsf{T}} \bar{p}(t)\|}{\varepsilon_o} \|u - \bar{u}(t)\|^2 : u \in U(t) \right\}$$

  is attained at $\bar{u}(t)$ for $t \in [0,1]$ a.e.
- If $U(t)$ is *convex* for all $t \in [0,1]$, then taking both $\varepsilon_o \longrightarrow \infty$ and $r \longrightarrow \infty$, the maximization condition (v) reduces to

$$\max \left\{ \left\langle \bar{\omega}(t)^{\mathsf{T}} \bar{p}(t), u \right\rangle : u \in U(t) \right\} \text{ is attained at } \bar{u}(t) \text{ for } t \in [0,1] \text{ a.e.} \tag{24}$$

Proof of Theorem 4.10. Theorem 4.7 produces a subsequence of $(\gamma_k)_k$, we do not relabel, and a corresponding sequence $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})_k$, with associated $(\xi_{\gamma_k})_k$ defined via (16), such that

- For each k, the quadruplet $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ is optimal for (P_{γ_k}).
- $(x_{\gamma_k}, u_{\gamma_k}) \xrightarrow[W^{1,2}\times\mathcal{W}]{\text{strongly}} (\bar{x}, \bar{u})$, $(y_{\gamma_k}, z_{\gamma_k}) \xrightarrow[W^{1,1}([0,1];\mathbb{R}^+\times\mathbb{R}^+)]{\text{strongly}} (0,0)$, $\xi_{\gamma_k} \xrightarrow[L^2([0,1];\mathbb{R}^+)]{\text{strongly}} \bar{\xi}$.
- $\dot{\bar{x}} \in BV([0,1];\mathbb{R}^n)$, $\bar{\xi} \in BV([0,1];\mathbb{R}^+)$, and (7)-(8) are valid at $(\bar{x}, \bar{u}, \bar{\xi})$ for all $t \in [0,1]$.
- All the conclusions of Theorem 3.11 hold, including $(x_{\gamma_k})_k$ is uniformly Lipschitz and $x_{\gamma_k}(t) \in \operatorname{int} C$ for all $t \in [0,1]$.
- For all k, we have

$$x_{\gamma_k}(i) \in \left[\left(C_i \cap \bar{B}_{\delta_o}(\bar{x}(i)) \right) + \tilde{\rho} B \right] \cap (\operatorname{int} C) \subset \operatorname{int} \tilde{C}_i(\delta), \quad \text{for } i = 0, 1.$$

In order to apply Proposition 4.9, we shall show that the constraint qualification (CQ) that holds for $U(\cdot)$ at $\bar{u}$, also holds true at u_{γ_k}, for k large enough. Indeed, if this is false, then, by Remark 4.8(i), there exist an increasing sequence $(k_n)_n$ in $\mathbb{N}$ and a sequence $t_n \in [0,1]$ such that $t_n \longrightarrow t_o \in [0,1]$ and

$$0 \in \partial_u^{>} d_U(t_n, u_{\gamma_{k_n}}(t_n)), \quad \forall n \in \mathbb{N}. \tag{25}$$

The continuity of $\bar{u}$ and the uniform convergence of $u_{\gamma_{k_n}}$ to $\bar{u}$ yield that the sequence $(u_{\gamma_{k_n}}(t_n))_n$ converges to $\bar{u}(t_o)$. Hence, using that the multifunction $(t,x) \mapsto \partial_u^{>} d_U(t,x)$ has closed values and a closed graph, we conclude from (25) that $0 \in \partial_u^{>} d_U(t_o, \bar{u}(t_o))$. This contradicts that the constraint qualification is satisfied by $U(\cdot)$ at $\bar{u}$. Thus, for k sufficiently large, $U(\cdot)$ satisfies the constraint qualification (CQ) at u_{γ_k}.

Hence, by Proposition 4.9, there exist a subsequence of $(\gamma_k)_k$, we do not relabel, and corresponding sequences $p_{\gamma_k}, q_{\gamma_k}$ μ_{γ_k} and λ_{γ_k} satisfying conditions (i)-(v) therein.

Using (23), (5), and that for all $t \in [0,1]$ we have

$$(x_{\gamma_k}(t), u_{\gamma_k}(t)) \in \operatorname{int}\left[\left(C \cap \bar{B}_\delta(\bar{x}(t)) \right) \times \left((U(t) + \tilde{\rho}\bar{B}) \cap \bar{B}_\delta(\bar{u}(t)) \right) \right],$$

we obtain sequences ζ_{γ_k}, θ_{γ_k} and ϑ_{γ_k} in $\mathcal{M}_{n\times n}([0,1])$ and ω_{γ_k} in $\mathcal{M}_{n\times m}([0,1])$ such that, for a.e. $t \in [0,1]$,

$$(\zeta_{\gamma_k}(t), \omega_{\gamma_k}(t)) \in \partial_\ell^{(x,u)} f(x_{\gamma_k}(t), u_{\gamma_k}(t)),$$
$$(\theta_{\gamma_k}(t), \vartheta_{\gamma_k}(t)) \in \partial_\ell^2 \varphi(x_{\gamma_k}(t)) \times \partial_\ell^2 \psi(x_{\gamma_k}(t)),$$

$$\dot{p}_{\gamma_k}(t) = (\theta_{\gamma_k}(t) - \zeta_{\gamma_k}(t))^{\mathsf{T}} p_{\gamma_k}(t) + \gamma_k e^{\gamma_k \psi(x_{\gamma_k}(t))} \vartheta_{\gamma_k}(t)\, p_{\gamma_k}(t) \tag{26}$$
$$+ \gamma_k^2 e^{\gamma_k \psi(x_{\gamma_k}(t))} \nabla\psi(x_{\gamma_k}(t)) \left\langle \nabla\psi(x_{\gamma_k}(t)), p_{\gamma_k}(t) \right\rangle, \text{ and}$$

$$\dot{q}_{\gamma_k}(t) = -(\omega_{\gamma_k}(t))^{\mathsf{T}} p_{\gamma_k}(t). \tag{27}$$

Note that for each k, the functions p_{γ_k}, $\dot{p}_{\gamma_k}$, q_{γ_k}, $\dot{q}_{\gamma_k}$, x_{γ_k}, u_{γ_k} and $\bar{u}$ are measurable on $[0, 1]$, and the multifunctions $\partial_\ell^{(x,u)} f(\cdot, \cdot)$, $\partial_\ell^2 \varphi(\cdot)$, and $\partial_\ell^2 \psi(\cdot)$ are measurable and have closed graphs with nonempty, compact, and convex values. Using (A1), (A2.1), and (A3), the Filippov measurable selection theorem (see [38, Theorem 2.3.13]) yields that we can assume the measurability of the functions $\zeta_{\gamma_k}(\cdot)$, $\theta_{\gamma_k}(\cdot)$, $\vartheta_{\gamma_k}(\cdot)$ and $\omega_{\gamma_k}(\cdot)$. Moreover, these sequences are uniformly bounded in L^∞, as $\|(\zeta_{\gamma_k}, \omega_{\gamma_k})\|_\infty \le M$, $\|\theta_{\gamma_k}\|_\infty \le K$ and $\|\vartheta_{\gamma_k}\|_\infty \le 2M_\psi$.

Step 1. Construction of $\bar{\xi}$, the admissible equation.

From Theorem 4.7, we have that the triplet $(\bar{x}, \bar{y}, \bar{\xi})$ satisfies (7) for all $t \in [0, 1]$. Hence, for all $t \in [0, 1]$ we have

$$\dot{\bar{x}}(t) = f_\Phi(\bar{x}(t), \bar{u}(t)) - \bar{\xi}(t)\nabla\psi(\bar{x}(t)) = f(\bar{x}(t), \bar{u}(t)) - \nabla\Phi(\bar{x}(t)) - \bar{\xi}(t)\nabla\psi(\bar{x}(t)).$$

Since $\nabla\Phi(x) = \partial_\ell \varphi(x) = \nabla_\ell \varphi(x)$ for all $x \in C$, we obtain that

$$\dot{\bar{x}}(t) = f(\bar{x}(t), \bar{u}(t)) - \nabla_\ell \varphi(\bar{x}(t)) - \bar{\xi}(t)\nabla\psi(\bar{x}(t)), \quad \forall t \in [0, 1].$$

On the other hand, since $\bar{x}$ takes values in C, we have $\psi(\bar{x}(t)) \le 0$, $\forall t \in [0, 1]$.

Step 2. Construction of $\bar{p}$, $\bar{\zeta}$, $\bar{\theta}$, $\bar{\vartheta}$, $\bar{\omega}$, $\bar{\nu}$, and the adjoint equation.

For the construction of $\bar{p}$, $\bar{\theta}$, $\bar{\vartheta}$ and $\bar{\nu}$, see Steps 2-4 in the proof of [39, Theorem 5.1]. Note that the uniform boundedness of $p_{\gamma_k}(1)$ established and used in Step 2 of the proof of [39, Theorem 5.1], is easily deduced here from the nontriviality condition of Proposition 4.9. We also note that, similarly to Step 2 of the proof of [39, Theorem 5.1], p_{γ_k} has a uniformly bounded variation, and hence, Helly first theorem implies that p_{γ_k} admits a pointwise convergent subsequence whose limit $\bar{p}$ is also of bounded variation and satisfies, for some $M_1 > 0$, the following

$$\|\bar{p}\|_\infty \le M_1 \|\bar{p}(1)\|. \tag{28}$$

Using Helly second theorem we obtain that for all $h \in C([0, 1]; \mathbb{R}^n)$,

$$\lim_{k\to\infty} \int_{[0,1]} \left\langle h(t), \dot{p}_{\gamma_k}(t) \right\rangle dt = \int_{[0,1]} \langle h(t), d\bar{p}(t) \rangle . \tag{29}$$

Identically to Steps 2-4 in the proof of [39, Theorem 5.1], we also have

$$\int_0^1 \left\langle h(t), \theta_{\gamma_k}(t)\, p_{\gamma_k}(t) \right\rangle dt \longrightarrow \int_0^1 \left\langle h(t), \bar{\theta}(t)\, \bar{p}(t) \right\rangle dt, \tag{30}$$

$$\int_0^1 \xi_{\gamma_k}(t) \left\langle h(t), \vartheta_{\gamma_k}(t)\, p_{\gamma_k}(t) \right\rangle dt \longrightarrow \int_0^1 \bar{\xi}(t) \left\langle h(t), \bar{\vartheta}(t)\, \bar{p}(t) \right\rangle dt, \tag{31}$$

$$\lim_{k\longrightarrow\infty} \int_0^1 \left\langle h(t), \nabla\psi(x_{\gamma_k}(t)) \right\rangle \gamma_k \xi_{\gamma_k}(t) \left\langle \nabla\psi(x_{\gamma_k}(t)), p_{\gamma_k}(t) \right\rangle dt$$
$$= \int_{[0,1]} \langle h(t), \nabla\psi(\bar{x}(t)) \rangle \, d\bar{\nu}(t). \tag{32}$$

We proceed to construct the two functions $\bar{\zeta}$ and $\bar{\omega}$. Note that the construction of ζ done in Step 2 of the proof of [39, Theorem 5.1] cannot be used here, since the *closed graph* hypothesis on the multifunction $(x, u) \mapsto \partial^x f(x, u)$ is required there, but it is not assumed here. As the sequence $(\zeta_{\gamma_k}, \omega_{\gamma_k})_k$ is uniformly bounded in L^∞, it has a subsequence, we do not relabel, that converges weakly in L^1 to some $(\bar{\zeta}, \bar{\omega})$. Using that the multifunction $(x, u) \mapsto \partial_\ell^{(x,u)} f(x, u)$ has closed graph with nonempty, compact and convex values, [10, Theorem 6.39] implies that, for $t \in [0, 1]$ a.e.,

$$(\bar{\zeta}(t), \bar{\omega}(t)) \in \partial_\ell^{(x,u)} f(\bar{x}(t), \bar{u}(t)).$$

Since $(p_{\gamma_k})_k$ is uniformly bounded in L^∞ and converges pointwise to $\bar{p}$, we conclude that

$$\zeta_{\gamma_k}^{\mathsf{T}} p_{\gamma_k} \xrightarrow[L^1]{\text{weakly}} \bar{\zeta}^{\mathsf{T}} \bar{p} \text{ and } \omega_{\gamma_k}^{\mathsf{T}} p_{\gamma_k} \xrightarrow[L^1]{\text{weakly}} \bar{\omega}^{\mathsf{T}} \bar{p}. \tag{33}$$

Hence, for all $h \in C([0, 1]; \mathbb{R}^n)$,

$$\int_0^1 \langle h(t), \zeta_{\gamma_k}(t) p_{\gamma_k}(t) \rangle \, dt \longrightarrow \int_0^1 \langle h(t), \bar{\zeta}(t) \bar{p}(t) \rangle \, dt. \tag{34}$$

Thus, from (26) and (29)-(34), we conclude that the adjoint equation of Theorem 4.10 holds, and it coincides with the adjoint equation of [39, Theorem 5.1].

Step 3. The complementary slackness conditions.

The part (a) follows from the equation (8). The part (b) follows from the uniform boundedness of $\left\| \gamma_k \xi_{\gamma_k}(\cdot) \langle \nabla\psi(x_{\gamma_k}(\cdot)), p_{\gamma_k}(\cdot) \rangle \right\|_1$ established in [39, Equation (97)]. More details can be found in Step 6 of [34, Proof of Theorem 6.1].

Step 4. Construction of λ and the transversality equation.

Form the transversality condition of Proposition 4.9, there exist $\upsilon_{\gamma_k} \in N^L_{C_0(k)}(x_{\gamma_k}(0))$, $\chi_{\gamma_k} \in N^L_{C_1(k)}(x_{\gamma_k}(1))$ and $(a_{\gamma_k}, b_{\gamma_k}) \in \partial^L g(x_{\gamma_k}(0), x_{\gamma_k}(1))$ such that

$$p_{\gamma_k}(0) = \lambda_{\gamma_k} a_{\gamma_k} + \lambda_{\gamma_k}(x_{\gamma_k}(0) - \bar{x}(0)) + \upsilon_{\gamma_k}, \quad -p_{\gamma_k}(1) = \lambda_{\gamma_k} b_{\gamma_k} + \chi_{\gamma_k}, \tag{35}$$

and the following properties hold:

- $\|(a_{\gamma_k}, b_{\gamma_k})\| \leq L_g$, where L_g is the Lipschitz constant of g over $\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)$, and $\|\lambda_{\gamma_k}\| \leq 1$ for all k. The latter inequality gives a subsequence, we do not relabel, such that $\lambda_{\gamma_k} \longrightarrow \lambda \in [0, 1]$.
- Due to Theorem 4.7, we have, for k large enough,

$$(x_{\gamma_k}(0), x_{\gamma_k}(1)) \in \operatorname{int}(\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)), \text{ and } (x_{\gamma_k}(0), x_{\gamma_k}(1)) \longrightarrow (\bar{x}(0), \bar{x}(1)).$$

- We have $p_{\gamma_k}(0) \longrightarrow \bar{p}(0)$ and $p_{\gamma_k}(1) \longrightarrow \bar{p}(1)$.
- Owing to (22), in which $d := x_{\gamma_k}(1) \in \left[C_1(k) \cap (\operatorname{int} C) \cap B_{\delta_o}(\bar{x}(1))\right]$ for k sufficiently large, we have $\chi_{\gamma_k} \in N^L_{C_1(k)}(x_{\gamma_k}(1)) = N^L_{C_1}\left(x_{\gamma_k}(1) + \bar{x}(1) - \bar{x}_{\gamma_k}(1)\right)$, for k large, where, we recall that $\bar{x}_{\gamma_k}(1) \longrightarrow \bar{x}(1)$.
- Owing to (21), in which $c := x_{\gamma_k}(0) \in C_0(k) \cap B_{\delta_o}(\bar{x}(0))$ for k large enough, it follows that:

  (i) If $\bar{x}(0) \in \operatorname{int} C$, then for k sufficiently large

$$\upsilon_{\gamma_k} \in N^L_{C_0(k)}(x_{\gamma_k}(0)) = N^L_{C_0}(x_{\gamma_k}(0)).$$

  (ii) If $\bar{x}(0) \in \operatorname{bdry} C$, using that $x_{\gamma_k}(0) \longrightarrow \bar{x}(0)$ and $\rho_k \longrightarrow 0$, then for k sufficiently large, $\left(x_{\gamma_k}(0) + \rho_k \frac{\nabla\psi(\bar{x}(0))}{\|\nabla\psi(\bar{x}(0))\|}\right) \in B_{\delta_o}(\bar{x}(0))$, and hence,

$$\upsilon_{\gamma_k} \in N^L_{C_0(k)}(x_{\gamma_k}(0)) = N^L_{C_0}\left(x_{\gamma_k}(0) + \rho_k \frac{\nabla\psi(\bar{x}(0))}{\|\nabla\psi(\bar{x}(0))\|}\right) \text{ for } k \text{ large.}$$

Therefore, along a subsequence of $(\gamma_k)_k$, we do not relabel, we have

$$\lambda_{\gamma_k}(a_{\gamma_k}, b_{\gamma_k}) \longrightarrow \lambda(a,b) \in \lambda\partial_\ell^L g(\bar{x}(0), \bar{x}(1)) \;\text{ and }\; \lambda_{\gamma_k}(x_{\gamma_k}(0) - \bar{x}(0)) \longrightarrow 0.$$

Thus, taking the limit as $k \to \infty$ in (35), and using $(p_{\gamma_k}(0), p_{\gamma_k}(1)) \longrightarrow (\bar{p}(0), \bar{p}(1))$, we obtain that $(v_{\gamma_k}, \chi_{\gamma_k})$ must converge to some (v, χ), as all the other terms in (35) converge. The last two bullets, stated above, yield that $v \in N_{C_0}^L(\bar{x}(0))$ and $\chi \in N_{C_1}^L(\bar{x}(1))$. Consequently, the limit of (35) is equivalent to

$$(\bar{p}(0), -\bar{p}(1)) \in \lambda\partial_\ell^L g(\bar{x}(0), \bar{x}(1)) + \left[N_{C_0}^L(\bar{x}(0)) \times \; N_{C_1}^L(\bar{x}(1))\right];$$

This terminates the proof of the transversality equation.

Step 5. The weak maximization condition.

By (27), (33)(b), and the transversality equations of Proposition 4.9, we have that

$$\dot{q}_{\gamma_k} = -(\omega_{\gamma_k})^{\mathsf{T}} p_{\gamma_k} \xrightarrow[L^1]{\text{weakly}} -(\bar{\omega})^{\mathsf{T}}\bar{p} \;\text{ and }\; q_{\gamma_k}(0) = \lambda_{\gamma_k}(u_{\gamma_k}(0) - \bar{u}(0)). \tag{36}$$

The uniform boundedness in L^∞ of the sequences $(p_{\gamma_k})_k$ and $(\omega_{\gamma_k})_k$ give that $(\dot{q}_{\gamma_k})_k$ is uniformly bounded in L^∞, asserting the equicontinuity of $(q_{\gamma_k})_k$. Moreover, the nontriviality condition of Proposition 4.9 gives the uniform boundedness of the sequence $(q_{\gamma_k})_k$. Hence, by Arzelà-Ascoli theorem, the sequence $(q_{\gamma_k})_k$ admits a subsequence, we do not relabel, that converges uniformly to an absolutely continuous function q satisfying $q(0) = 0$ (by (36)(b), where $\lambda_{\gamma_k} \longrightarrow \lambda$ and $u_{\gamma_k}(0) \longrightarrow \bar{u}(0)$ as $k \longrightarrow \infty$). Moreover, up to a subsequence, we also obtain that

$$\dot{q}_{\gamma_k} \xrightarrow[L^1]{\text{weakly}} \dot{q}. \tag{37}$$

Hence, (36)(a) and the uniqueness of the L^1-weak limit yield that

$$\dot{q}(t) = -(\bar{\omega}(t))^{\mathsf{T}}\bar{p}(t), \quad t \in [0,1] \text{ a.e.} \tag{38}$$

We proceed to study the convergence of the sequence of NBV-functions, $(\Omega_{\gamma_k})_k$, obtained in Proposition 4.9. The maximization condition (iv), therein, implies that, for $t \in [0,1]$ a.e.,

$$\Omega_{\gamma_k}(t) = -q_{\gamma_k}(t) + \underbrace{\lambda_{\gamma_k}(\dot{u}_{\gamma_k}(t) - \dot{\bar{u}}(t))}_{\ell_{\gamma_k}(t)}. \tag{39}$$

Without loss of generality, we can assume that (39) is satisfied for all $t \in [0,1]$. In fact, if $\lambda_{\gamma_k} = 0$, using the transversality conditions of Proposition 4.9 and that $\Omega_{\gamma_k} \in NBV[0,1]$, we get that $\Omega_{\gamma_k}(0) = -q_{\gamma_k}(0) = 0$ and $\Omega_{\gamma_k}(1) = -q_{\gamma_k}(1)$, and hence, by the right continuity of Ω_{γ_k} and the continuity of q_{γ_k}, (39) is equivalent to $\Omega_{\gamma_k} \equiv -q_{\gamma_k}$. If, however, $\lambda_{\gamma_k} > 0$, then by modifying the values of $(\dot{u}_{\gamma_k} - \dot{\bar{u}})$ on the set of Lebesgue measure zero, we have (39) satisfied for all $t \in [0,1]$, and hence, $\ell_{\gamma_k} \in BV[0,1]$ is right continuous on $(0,1)$, and satisfies $\ell_{\gamma_k}(0) = q_{\gamma_k}(0)$ and $\ell_{\gamma_k}(1) = 0$. Furthermore, since $\lambda_{\gamma_k} \longrightarrow \lambda$ and $\dot{u}_{\gamma_k}$ strongly converges in L^2 to $\dot{\bar{u}}$, the sequence $(\ell_{\gamma_k})_k$ strongly converges in L^2 to $\ell = 0$.

We claim that $(\Omega_{\gamma_k})_k$, considered as a sequence of continuous linear functionals on $C([0,1]; \mathbb{R}^m)$, admits a subsequence, we do not relabel, that converges weakly* to $-q$. Since Ω_{γ_k} satisfies (39), where the sequence of absolutely continuous functions $(q_{\gamma_k})_k$ converges uniformly to $q \in AC([0,1]; \mathbb{R}^m)$ and, by (37), $(\dot{q}_{\gamma_k})_k$ converges weakly in L^1 to $\dot{q}$, then it is equivalent to show that the BV-sequence $(\ell_{\gamma_k})_k$ converges in $C^*([0,1]; \mathbb{R}^m)$ to 0. The uniform boundedness of the sequence $(\ell_{\gamma_k})_k$ shall follow once we show the uniform boundedness of $(\Omega_{\gamma_k})_k$. For this latter, the nontriviality condition of Proposition 4.9, implies that the sequence $(\mu^o_{\gamma_k})_k$ is uniformly bounded, and hence, it has a subsequence, we do not relabel, that converges weakly* to a $\mu^o \in C^{\oplus}([0,1];\mathbb{R}^m)$. Moreover, by condition (v) of Proposition 4.9 and Remark 4.8(i), $\|\beta_{\gamma_k}(t)\| \leq 1$, except on a set of $\mu^o_{\gamma_k}$-measure zero. Thus, using that

$$\Omega_{\gamma_k}(t) = \int_{[0,t]} \beta_{\gamma_k}(s)\mu^o_{\gamma_k}(ds), \quad \forall t \in (0,1], \quad \text{and} \quad \Omega_{\gamma_k}(0) = 0, \tag{40}$$

we obtain that the sequence $(\Omega_{\gamma_k})_k$ is uniformly bounded, and so is the sequence $(\ell_{\gamma_k})_k$. Hence, to get that the bounded BV-sequence $(\ell_{\gamma_k})_k$ converges weakly* to 0, by the Banach-Steinhaus theorem in [25, p.482], it is sufficient to show that

$$\lim_{k\longrightarrow\infty} \int_0^1 \left\langle h(t), d\ell_{\gamma_k}(t)\right\rangle = 0, \quad \forall h \in C^1([0,1];\mathbb{R}^m).$$

Fix $h \in C^1([0,1];\mathbb{R}^m)$. Using an integration by parts and that $\Omega_{\gamma_k} \in NBV$, we get

$$\begin{aligned}\int_0^1 \langle h(t), d\ell_{\gamma_k}(t)\rangle &= \langle h(1), \ell_{\gamma_k}(1)\rangle - \langle h(0), \ell_{\gamma_k}(0)\rangle - \int_0^1 \left\langle \dot{h}(t), \ell_{\gamma_k}(t)\right\rangle dt \\ &= \left[-\langle h(0), q_{\gamma_k}(0)\rangle - \int_0^1 \left\langle \dot{h}(t), \ell_{\gamma_k}(t)\right\rangle dt\right] \xrightarrow[k\longrightarrow\infty]{} 0,\end{aligned}$$

since $(\ell_{\gamma_k})_k$ strongly converges in L^2 to 0, and $q_{\gamma_k}(0) \longrightarrow q(0) = 0$. This terminates the proof of the claim, that is,

$$\Omega_{\gamma_k} \xrightarrow[C^*([0,1];\mathbb{R}^m)]{\text{weakly*}} -q. \tag{41}$$

By [25, p. 484, #8], we also have that $\Omega_{\gamma_k}(t) \longrightarrow -q(t), \ \forall t \in [0,1]$.

Now define the signed measure $\mu_{\gamma_k}(dt) := \beta_{\gamma_k}(t)\mu^o_{\gamma_k}(dt)$. From (40) we have

$$\Omega_{\gamma_k}(t) = \mu_{\gamma_k}[0,t], \quad \forall t \in (0,1]. \tag{42}$$

Using that $\mu^o_{\gamma_k} \xrightarrow[k\longrightarrow\infty]{\text{weakly*}} \mu^o$, and that Proposition 4.9(v) holds true, then, by applying [38, Proposition 9.2.1] to the following data:

- $A_{\gamma_k}(t) := \partial^>_u d(u_{\gamma_k}(t), U(t) \cap \bar{B}_\delta(\bar{u}(t)))$ for all $t \in [0,1]$,
- $A(t) := \partial^>_u d(\bar{u}(t), U(t) \cap \bar{B}_\delta(\bar{u}(t)))$ for all $t \in [0,1]$,
- $\gamma_{\gamma_k} := \beta_{\gamma_k}$, $\mu_{\gamma_k} := \mu^o_{\gamma_k}$ and $\mu_0 := \mu^o$,

we obtain a Borel measurable function $\beta\colon [0,1] \longrightarrow \mathbb{R}^m$ and $\mu \in C^*([0,1];\mathbb{R}^m)$ such that

$$\mu_{\gamma_k} \xrightarrow[k\longrightarrow\infty]{\text{weakly*}} \mu, \quad \mu(dt) = \beta(t)\mu^o(dt) \text{ and } \beta(t) \in \partial^>_u d(\bar{u}(t), U(t) \cap \bar{B}_\delta(\bar{u}(t))) \quad \mu^o \text{ a.e.}$$

Since u_{γ_k} converges uniformly to $\bar{u}$, and supp $\{\mu^o_{\gamma_k}\}$ satisfies Proposition 4.9(v), we deduce that

$$\text{supp}\,\{\mu^o\} \subset \mathcal{A} := \left\{t \in [0,1] : (t, \bar{u}(t)) \in \text{bdry Gr}\left[U(t) \cap \bar{B}_\delta(\bar{u}(t))\right]\right\}. \tag{43}$$

Adjust $\beta(\cdot)$ on the set of μ^o-measure zero to arrange

$$t \in \mathcal{A} \implies \beta(t) \in \partial^>_u d(\bar{u}(t), U(t) \cap \bar{B}_\delta(\bar{u}(t))),$$

and hence, using [38, Formula (9.17)], we have

$$\beta(t) \in \left[\text{conv}\, \bar{N}^L_{U(t)\cap\bar{B}_\delta(\bar{u}(t))}(\bar{u}(t)) \cap \left(\bar{B} \setminus \{0\}\right)\right], \quad \forall t \in \mathcal{A}. \tag{44}$$

Thus, by (41) and (42), we obtain that $-dq(t) = \mu(dt) = \beta(t)\mu^o(dt)$. Using (38), we arrive to

$$-dq(t) = (\bar{\omega}(t))^{\mathsf{T}}\bar{p}(t)dt = \beta(t)\mu^o(dt). \tag{45}$$

Next, we decompose $\mu^o(dt) = m(t)dt + \mu_s(dt)$ for some *nonnegative* L^1-function $m(\cdot)$ and some nonnegative Borel measure μ_s totally singular with respect to Lebesgue measure. Clearly $m(t) = 0$, for all $t \in \mathcal{A}^c$, and hence, (44) implies that

$$\beta(t)m(t) \in \text{conv}\,\bar{N}^L_{U(t)\cap\bar{B}_\delta(\bar{u}(t))}(\bar{u}(t)), \quad \forall t \in [0,1].$$

Using (45) we get that $(\bar{\omega}(t))^{\mathsf{T}}\bar{p}(t)dt = \beta(t)m(t)dt + \beta(t)\mu_s(dt)$. This gives that $(\bar{\omega}(t))^{\mathsf{T}}\bar{p}(t) = \beta(t)m(t)$, for $t \in [0,1]$ a.e. Therefore,

$$\bar{\omega}(t)^{\mathsf{T}}\bar{p}(t) \in \text{conv}\,\bar{N}^L_{U(t)\cap\bar{B}_\delta(\bar{u}(t))}(\bar{u}(t)), \quad \forall t \in [0,1] \text{ a.e.} \tag{46}$$

In order to show the validity of the "In addition" part of the weak maximization condition, we shall employ the following technical lemma whose proof follows from the local property of the normal cones, the proximal normal inequality ([11, Proposition 1.1.5(a)]), and the fact that the proximal, Mordukhovich, and Clarke normal cones coincide in our setting.

Lemma 4.12 *Let $F\colon [0,1] \rightrightarrows \mathbb{R}^m$ be a lower semicontinuous multifunction with closed and nonempty values and let $h \in C([0,1]; F)$. If there exist $\varepsilon_o > 0$ and $r > 0$ such that $F(t) \cap \bar{B}_{\varepsilon_o}(h(t))$ is r-prox-regular for all $t \in [0,1]$, then for any $\delta > 0$ we have*

$$\text{conv}\,(\bar{N}^L_{F(t)\cap\bar{B}_\delta(h(t))}(h(t))) = N^P_{F(t)\cap\bar{B}_{\varepsilon_o}(h(t))}(h(t)), \quad \forall t \in [0,1].$$

Moreover, for all $t \in [0,1]$ and for all $\zeta \in \text{conv}\,(\bar{N}^L_{F(t)\cap\bar{B}_\delta(h(t))}(h(t)))$, we have

$$\langle \zeta, v - h(t)\rangle \le \tfrac{\|\zeta\|}{\min\{\varepsilon_o, 2r\}}\|v - h(t)\|^2, \quad \forall v \in F(t).$$

Now, for the proof of the "In addition" part, consider $\varepsilon_o > 0$ and $r > 0$ therein such that $U(t) \cap \bar{B}_{\varepsilon_o}(\bar{u}(t))$ is r-prox-regular for all $t \in [0,1]$. From (46) and Lemma 4.12, we obtain that for all $t \in [0,1]$ a.e.

$$\left\langle \bar{\omega}(t)^{\mathsf{T}}\bar{p}(t), u - \bar{u}(t)\right\rangle \le \tfrac{\|\bar{\omega}(t)^{\mathsf{T}}\bar{p}(t)\|}{\min\{\varepsilon_o, 2r\}}\|u - \bar{u}(t)\|^2, \quad \text{for all } u \in U(t).$$

Therefore, for a.e. $t \in [0,1]$,

$$\left\langle \omega(t)^{\mathsf{T}}p(t), u\right\rangle - \tfrac{\|\bar{\omega}(t)^{\mathsf{T}}\bar{p}(t)\|}{\min\{\varepsilon_o, 2r\}}\|u - \bar{u}(t)\|^2 \le \left\langle \bar{\omega}(t)^{\mathsf{T}}\bar{p}(t), \bar{u}(t)\right\rangle, \quad \text{for all } u \in U(t).$$

This terminates the proof of the weak maximization condition.

Step 6. The nontriviality condition.

It is sufficient to prove its equivalent condition: $\|\bar{p}(1)\| + \lambda \neq 0$. Taking the limit as $k \longrightarrow \infty$ in the nontriviality condition of Proposition 4.9, and using the convergence of $\bar{p}_{\gamma_k}(1)$ to $\bar{p}(1)$, the uniform convergence of q_{γ_k} to q, the weak* convergence of $\mu^o_{\gamma_k}$ to μ^o, and the convergence of λ_{γ_k} to λ, we get that

$$1 = \|\bar{p}(1)\| + \|q\|_\infty + \|\mu^o\|_{\text{T.V.}} + \lambda. \tag{47}$$

We argue by contradiction. If $\bar{p}(1) = 0$ and $\lambda = 0$, by (28) we obtain that $\bar{p} = 0$. Hence (38) yields that $\dot{q}(t) = 0$ for a.e. $t \in [0,1]$. This gives that

$$\beta(t)\mu^o(dt) = -dq(t) = 0 \ \text{ and } \ q(t) = q(0) + \int_0^t \dot{q}(\tau)d\tau = 0, \quad \forall t \in [0,1]. \tag{48}$$

Since, by (43) and (45), supp $\{\mu^o\} \subset \mathcal{A}$ and $\beta(t) \neq 0$ for all $t \in \mathcal{A}$, the first equation of (48) yields that $\mu^o = 0$. Therefore, $\|\bar{p}(1)\| + \|q\|_\infty + \|\mu^o\|_{\text{T.V.}} + \lambda = 0$ which contradicts (47). This terminates the proof of the nontriviality condition, and then, the proof of the conditions (i)-(vi) of Theorem 4.10 is completed.

For the "Furthermore" part of the theorem, assume that $C_1 = \mathbb{R}^n$. We need to prove that $\lambda \neq 0$. If not, then from the transversality condition (v), we get that $\bar{p}(1) \in N^L_{C_1}(\bar{x}(1)) = \{0\}$. This contradicts the nontriviality condition (i). □

The following is an example of a problem (P) on which we apply the necessary optimality conditions of Theorem 4.10 to find its optimal solution.

Example 4.13 We consider the following data for (P):

- The perturbation mapping $f : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}^2$ is defined by

$$f((x_1, x_2), u) = (-x_1 - x_2 - u, x_1 - x_2 + u).$$

- The function $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is defined by $\psi(x_1, x_2) := (x_1^2 + x_2^2 - 1)(x_1^2 + x_2^2 - 4)$, and hence, the set C is the nonconvex and compact set

$$C := \{(x_1, x_2) : (x_1^2 + x_2^2 - 1)(x_1^2 + x_2^2 - 4) \leq 0\}.$$

- The objective function $g : \mathbb{R}^4 \longrightarrow \mathbb{R} \cup \{\infty\}$ is defined by

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

- The function φ is the indicator function of C.
- The control multifunction is $U(t) := [0, 1]$ for all $t \in [0, \frac{\pi}{2}]$, $C_0 := \{(1, 0)\}$, and $C_1 := \{(0, x_2) : x_2 \geq 0\}$.

One can easily prove that the assumptions (A2)-(A4) are satisfied. Furthermore, as f is globally Lipschitz on $\mathbb{R}^2 \times \mathbb{R}$, g is globally Lipschitz on $\mathbb{R}^2 \times C$, and U is convex with nonempty interior, we conclude that all the assumptions of Theorem 4.10 are satisfied. Since that g vanishes on the unit circle and is strictly positive elsewhere in C, we may seek for (P) an optimal solution $(\bar{x}, \bar{u})$, if it exists, such that $\bar{x} := (\bar{x}_1, \bar{x}_2)$ belongs to the unit circle, and hence we have

$$\begin{cases} \bar{x}_1^2(t) + \bar{x}_2^2(t) = 1 \; \forall t \in [0, \frac{\pi}{2}]; \text{ and } \; \bar{x}_1(t)\dot{\bar{x}}_1(t) + \bar{x}_2(t)\dot{\bar{x}}_2(t) = 0 \; \text{ a.e.,} \\ \bar{x}(0)^\mathsf{T} = (1, 0) \; \text{ and } \; \bar{x}(\frac{\pi}{2})^\mathsf{T} = (0, 1). \end{cases} \tag{49}$$

Applying Theorem 4.10 to this optimal solution $(\bar{x}, \bar{u})$ and using (24), we get the existence of an adjoint vector $\bar{p} := (\bar{p}_1, \bar{p}_2) \in BV([0, \frac{\pi}{2}]; \mathbb{R}^2)$, a finite signed Radon measure $\bar{\nu}$ on $[0, \frac{\pi}{2}]$, $\bar{\xi} \in L^\infty([0, \frac{\pi}{2}]; \mathbb{R}^+)$, and $\lambda \geq 0$ such that, when incorporating equations (49) into (i)- (vi), these latter simplify to the following

(a) $\|\bar{p}(\frac{\pi}{2})\| + \lambda = 1$.
(b) The admissibility equation holds, that is, for $t \in [0, \frac{\pi}{2}]$ a.e.,

$$\begin{cases} \dot{\bar{x}}_1(t) = -\bar{x}_1(t) - \bar{x}_2(t) - \bar{u} + 6\bar{x}_1(t)\bar{\xi}(t), \\ \dot{\bar{x}}_2(t) = \bar{x}_1(t) - \bar{x}_2(t) + \bar{u} + 6\bar{x}_2(t)\bar{\xi}(t). \end{cases}$$

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

$$d\bar{p}(t) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \bar{p}(t)dt + \bar{\xi}(t) \begin{pmatrix} 8\bar{x}_1^2(t) - 6 & 8\bar{x}_1(t)\bar{x}_2(t) \\ 8\bar{x}_1(t)\bar{x}_2(t) & 8\bar{x}_2^2(t) - 6 \end{pmatrix} \bar{p}(t)dt - 6 \begin{pmatrix} \bar{x}_1(t) \\ \bar{x}_2(t) \end{pmatrix} d\bar{\nu}.$$

(d) The complementary slackness condition is valid, that is,

$$\bar{\xi}(t)(\bar{p}_1(t)\bar{x}_1(t) + \bar{p}_2(t)\bar{x}_2(t)) = 0, \quad t \in \left[0, \tfrac{\pi}{2}\right] \text{ a.e.}$$

(e) The transversality condition holds: $-\bar{p}(\frac{\pi}{2}) \in \lambda\{(0,1)\} + \{(\alpha, 0) \in \mathbb{R}^2 : \alpha \in \mathbb{R}\}$.[4]

(f) $\max\{u(\bar{p}_2(t) - \bar{p}_1(t)) : u \in [0,1]\}$ is attained at $\bar{u}(t)$ for $t \in [0, \frac{\pi}{2}]$ a.e.

Combining (49) and (b), we deduce that

$$\bar{\xi}(t) = \frac{1 + \bar{u}(t)(\bar{x}_1(t) - \bar{x}_2(t))}{6}, \quad \forall t \in [0, \tfrac{\pi}{2}]. \tag{50}$$

On the other hand, employing (d) and (49)(a) in (c), yields that, for $t \in [0, \frac{\pi}{2}]$,

$$\begin{cases} d\bar{p}_1 = (\bar{p}_1(t) - \bar{p}_2(t) - 6\bar{\xi}(t)\bar{p}_1(t))dt - 6\bar{x}_1(t)d\bar{\nu}, \\ d\bar{p}_2 = (\bar{p}_1(t) + \bar{p}_2(t) - 6\bar{\xi}(t)\bar{p}_2(t))dt - 6\bar{x}_2(t)d\bar{\nu}. \end{cases} \tag{51}$$

To benefit from (f), we temporarily assume that

$$\bar{p}_2(t) < \bar{p}_1(t) \text{ for } t \in [0, \tfrac{\pi}{2}] \text{ a.e.} \tag{52}$$

Then $\bar{u} = 0$, which gives using (50) that $\bar{\xi}(t) = \frac{1}{6}$ for all $t \in \left[0, \frac{\pi}{2}\right]$. Solving the two differential equations of (b) and using (49), we conclude that

$$\bar{x}(t)^{\mathsf{T}} = (\cos t, \sin t), \quad \forall t \in [0, \tfrac{\pi}{2}].$$

Using (a), (d), (e), and (51), a simple calculation gives that

$$\begin{cases} \lambda = \frac{3}{8} \text{ and } \bar{p}(\frac{\pi}{2}) = (\frac{1}{2}, -\frac{3}{8}), \\ \bar{p}(t)^{\mathsf{T}} = \frac{1}{2}(\sin t, -\cos t) \text{ on } [0, \frac{\pi}{2}) \text{ and } d\bar{\nu} = \frac{1}{16}\delta_{\{\frac{\pi}{2}\}} \text{ on } [0, \frac{\pi}{2}], \end{cases}$$

where $\delta_{\{a\}}$ denotes the unit measure concentrated on the point a. Note that for all $t \in [0, \frac{\pi}{2}]$, we have $\bar{p}_2(t) < \bar{p}_1(t)$, and hence, the temporary assumption (52) is satisfied.

Therefore, the above analysis, realized via Theorem 4.10, produces an admissible pair $(\bar{x}, \bar{u})$, where

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

which is optimal for (P).

5 Proofs of Theorems 4.1 and 4.7

Proof of Theorem 4.1. (i): Having a uniform bounded derivative in $L^2([0,1]; \mathbb{R}^m)$, the $W^{1,2}$-sequence u_{γ_k} is equicontinuous. Since, by (A4.2), the compact sets $U(t)$ are uniformly

[4] The use of $\partial_\ell^L g$, instead of the usual $\partial^L g$, in our transversality condition yields better information. Indeed, if we use $\partial^L g$, then we obtain in the transversality condition the set $\lambda\{(0, 1-\beta) : \beta \geq 0\}$ instead of the *strictly* smaller set $\lambda\{(0,1)\}$.

bounded, then u_{γ_k} is uniformly bounded in $C([0,1];\mathbb{R}^m)$, and hence, Arzelà-Ascoli theorem asserts that u_{γ_k} admits a subsequence, we do not relabel, that converges uniformly to an absolutely continuous function u with $u(t)\in U(t)$ for all $t\in[0,1]$. As $\dot{u}_{\gamma_k}$ is uniformly bounded in $L^2([0,1];\mathbb{R}^m)$, then, up to a subsequence, it is weakly convergent in L^2. The boundedness of $(u_{\gamma_k}(0))_k$ then yields that the L^2-weak limit of $\dot{u}_{\gamma_k}$ is $\dot{u}$, and whence, $u\in\mathcal{W}$. The fact that x is the unique solution to (D) corresponding to (x_0,u), and the proceeding statements of this part, follow immediately from Theorem 3.9(ii).

(ii): Now, assume that $c_{\gamma_k}\in C(k)$ for $k\ge k_o$, where k_o is the rank in Theorem 3.11.

Let us first show that $(\xi_{\gamma_k})_k$ has uniform bounded variations. Since ψ and $\nabla\psi$ are Lipschitz on C and x_{γ_k} is Lipschitz for $k\ge k_o$, we deduce that, for $k\ge k_o$, the function $\xi_{\gamma_k}(\cdot)\nabla\psi(x_{\gamma_k}(\cdot))$, where ξ_{γ_k} is defined in (16), is Lipschitz continuous on $[0,1]$. Similarly, the Lipschitz property on $C\times(\mathbb{U}+\tilde{\rho}\bar{B})$ of $f(\cdot,\cdot)$ (and then of $f_\Phi(\cdot,\cdot)$) and the fact that $(x_{\gamma_k},u_{\gamma_k})$ is in $W^{1,\infty}\times W^{1,2}$, yield that $f_\Phi(x_{\gamma_k}(\cdot),u_{\gamma_k}(\cdot))$ is in $W^{1,2}([0,1];\mathbb{R}^n)$. Hence,

$$\zeta_{\gamma_k}(t):=\frac{d}{dt}f_\Phi(x_{\gamma_k}(t),u_{\gamma_k}(t)),$$

exists for almost all $t\in[0,1]$. By writing $f_\Phi=(f_\Phi^1,\cdots,f_\Phi^n)^\intercal$, and using that $x_{\gamma_k}(t)\in\operatorname{int}C$ (for all $t\in[0,1]$), and $u_{\gamma_k}(t)\in U(t)\subset\mathbb{U}$ (for $t\in[0,1]$ a.e.), it follows from the proof of [40, Theorem 2.1], that

$$\zeta^i_{\gamma_k}(t)\in\langle\partial f^i_\Phi(x_{\gamma_k}(t),u_{\gamma_k}(t)),(\dot{x}_{\gamma_k}(t),\dot{u}_{\gamma_k}(t))\rangle,\quad t\in[0,1]\text{ a.e.},\ \forall i=1,\cdots,n.$$

Since $(\|\dot{u}_{\gamma_k}\|_2)_k$ is assumed to be bounded, and, by Theorem 3.11, $(\|\dot{x}_{\gamma_k}\|_\infty)_k$ is bounded, then the sequence $(\|\zeta_{\gamma_k}\|_2)_k$ is bounded by some $M_\zeta>0$ that depends on $\bar{M}$, $\bar{M}_\psi$, η, and the bound of $(\|\dot{u}_{\gamma_k}\|_2)_k$.

As $f_\Phi(x_{\gamma_k}(\cdot),u_{\gamma_k}(\cdot))\in W^{1,2}([0,1];\mathbb{R}^n)$, the right hand side of (D_{γ_k}) yields that $\dot{x}_{\gamma_k}$ is in $W^{1,2}([0,1];\mathbb{R}^n)$, and so is the function $|\langle\nabla\psi(x_{\gamma_k}(\cdot)),\dot{x}_{\gamma_k}(\cdot)\rangle|$. This also implies that $\xi_{\gamma_k}\in W^{2,2}([0,1];\mathbb{R}^+)$, due to

$$\dot{\xi}_{\gamma_k}(t)=\gamma_k^2e^{\gamma_k\psi(x_{\gamma_k}(t))}\langle\nabla\psi(x_{\gamma_k}(t)),\dot{x}_{\gamma_k}(t)\rangle.$$

Next, calculating $\ddot{x}_{\gamma_k}$ through (D_{γ_k}) in terms of ζ_{γ_k} and $\dot{x}_{\gamma_k}$, and using the fact that for $h\in AC([0,1];\mathbb{R})$ we have

$$\frac{d}{dt}|h(t)|=\left(\frac{d}{dt}h(t)\right)\operatorname{sign}(h(t))\ \text{ a.e. } t\in(0,1),$$ [5]

it follows that there exist measurable functions $\vartheta^1_{\gamma_k}$ and $\vartheta^2_{\gamma_k}$ whose values at t are in $\partial^2\psi(x_{\gamma_k}(t))$, for almost all $t\in[0,1]$, such that, for $t\in[0,1]$ a.e., we have

$$\begin{aligned}&\frac{d}{dt}|\langle\nabla\psi(x_{\gamma_k}(t)),\dot{x}_{\gamma_k}(t)\rangle|\\&=\left[\left\langle\vartheta^1_{\gamma_k}(t)\dot{x}_{\gamma_k}(t),\dot{x}_{\gamma_k}(t)\right\rangle+\left\langle\nabla\psi(x_{\gamma_k}(t)),\ddot{x}_{\gamma_k}(t)\right\rangle\right]\overbrace{\operatorname{sign}(\langle\nabla\psi(x_{\gamma_k}(t)),\dot{x}_{\gamma_k}(t)\rangle)}^{\alpha(t)}\\&=\left[\left\langle\vartheta^1_{\gamma_k}(t)\dot{x}_{\gamma_k}(t),\dot{x}_{\gamma_k}(t)\right\rangle+\langle\nabla\psi(x_{\gamma_k}(t)),\zeta_{\gamma_k}(t)-\xi_{\gamma_k}(t)\vartheta^2_{\gamma_k}(t)\dot{x}_{\gamma_k}(t)\rangle\right]\alpha(t)\\&\quad-\langle\underbrace{\gamma_k\xi_{\gamma_k}(t)\langle\nabla\psi(x_{\gamma_k}(t)),\dot{x}_{\gamma_k}(t)\rangle}_{\dot{\xi}_{\gamma_k}(t)}\nabla\psi(x_{\gamma_k}(t)),\nabla\psi(x_{\gamma_k}(t))\rangle\alpha(t)\end{aligned}$$

[5] The function sign: $\mathbb{R}\longrightarrow\mathbb{R}$ is defined by: $\operatorname{sign}(x)=\frac{x}{|x|}$ for $x\neq0$, and 0 for $x=0$.

$$[3pt] = \left[\left\langle \vartheta^1_{\gamma_k}(t)\dot{x}_{\gamma_k}(t), \dot{x}_{\gamma_k}(t)\right\rangle + \left\langle \nabla\psi(x_{\gamma_k}(t)), \zeta_{\gamma_k}(t) - \xi_{\gamma_k}(t)\vartheta^2_{\gamma_k}(t)\dot{x}_{\gamma_k}(t)\right\rangle\right]\alpha(t)$$
$$-\gamma_k\xi_{\gamma_k}(t)\,|\langle\nabla\psi(x_{\gamma_k}(t)), \dot{x}_{\gamma_k}(t)\rangle|\;\|\nabla\psi(x_{\gamma_k}(t))\|^2.$$

Integrating both sides on [0, 1] and using the boundedness of $(\|\dot{x}_{\gamma_k}\|_\infty)_k$ and $(\|\xi_{\gamma_k}\|_\infty)_k$ (by Theorem 3.11), and assumption (A2.1), we get the existence of a constant $\tilde{M}_1$ depending on $\bar{M}$, M_ψ, $\bar{M}_\psi$, η, and M_ζ such that

$$\int_0^1 |\dot{\xi}_{\gamma_k}(t)|\|\nabla\psi(x_{\gamma_k}(t))\|^2 dt \le \tilde{M}_1.$$

Using (3) and assumption (A2.2), it follows that

$$\int_0^1 |\dot{\xi}_{\gamma_k}(t)|dt = \int_0^1 \gamma_k^2 e^{\gamma_k\psi(x_{\gamma_k}(t))}|\langle\nabla\psi(x_{\gamma_k}(t)), \dot{x}_{\gamma_k}(t)\rangle|dt$$
$$= \int_{\{t:\|\nabla\psi(x_{\gamma_k}(t))\|\le\eta\}} \gamma_k^2 e^{\gamma_k\psi(x_{\gamma_k}(t))}|\langle\nabla\psi(x_{\gamma_k}(t)), \dot{x}_{\gamma_k}(t)\rangle|dt$$
$$+ \int_{\{t:\|\nabla\psi(x_{\gamma_k}(t))\|>\eta\}} \gamma_k^2 e^{\gamma_k\psi(x_{\gamma_k}(t))}|\langle\nabla\psi(x_{\gamma_k}(t)), \dot{x}_{\gamma_k}(t)\rangle|\frac{\|\nabla\psi(x_{\gamma_k}(t))\|^2}{\|\nabla\psi(x_{\gamma_k}(t))\|^2}dt$$
$$\le \eta\left(\bar{M} + \frac{2\bar{M}\bar{M}_\psi}{\eta}\right)\gamma_k^2 e^{-\gamma_k\varepsilon} + \frac{\tilde{M}_1}{\eta^2} \le \eta\left(\bar{M} + \frac{2\bar{M}\bar{M}_\psi}{\eta}\right) + \frac{\tilde{M}_1}{\eta^2} =: \tilde{M}_2,$$

for k sufficiently large, where $\tilde{M}_2$ depends on the given constants, $\bar{M}$, $\bar{M}_\psi$, M_ψ, η, and on the bound of $(\|\dot{u}_{\gamma_k}\|_2)_k$. Therefore, the sequence ξ_{γ_k} satisfies, for k sufficiently large, $V_0^1(\xi_{\gamma_k}) \le \tilde{M}_2$.

On the other hand, by Theorem 3.11, $\|\xi_{\gamma_k}\|_\infty \le \frac{2\bar{M}}{\eta}$ for all $k \ge k_o$. Hence, by Helly first theorem, ξ_{γ_k} admits a pointwise convergent subsequence, we do not relabel, whose limit is some function $\tilde{\xi} \in BV([0,1];\mathbb{R}^+)$ with $\|\tilde{\xi}\|_\infty \le \frac{2\bar{M}}{\eta}$ and $V_0^1(\tilde{\xi}) \le \tilde{M}_2$. Being pointwise convergent to $\tilde{\xi}$ and uniformly bounded in L^∞, ξ_{γ_k} *strongly* converges in L^2 to $\tilde{\xi}$. However, by part(i) of this theorem, ξ_{γ_k} converges weakly in L^2 to ξ, hence, $\tilde{\xi} = \xi$. Thus, ξ_{γ_k} converges pointwise and strongly in L^2 to ξ, and $\xi \in BV([0,1];\mathbb{R}^+)$ with

$$V_0^1(\xi) \le \tilde{M}_2. \tag{53}$$

As f is M-Lipschitz on $C \times (\mathbb{U} + \tilde{\rho}\bar{B})$, $u \in \mathcal{W}$, $\nabla\psi$ is Lipschitz, and $\xi \in BV$, then equation (7), which is satisfied by (x, u, ξ), now holds for all $t \in [0, 1]$. This yields that (8) is also valid for all $t \in [0, 1]$, and that $\dot{x} \in BV([0,1];\mathbb{R}^n)$.

It remains to show that $\dot{x}_{\gamma_k}$ has uniform bounded variations and converges pointwise and strongly in L^2 to $\dot{x} \in BV([0,1];\mathbb{R}^n)$. Since $\xi_{\gamma_k}(\cdot)\nabla\psi(x_{\gamma_k}(\cdot))$ is Lipschitz, $u_{\gamma_k} \in \mathcal{W}$, and f is M-Lipschitz on $C \times (\mathbb{U} + \tilde{\rho}\bar{B})$, then (D_{γ_k}) holds for all $t \in [0, 1]$, that is,

$$\dot{x}_{\gamma_k}(t) = f_\Phi(x_{\gamma_k}(t), u_{\gamma_k}(t)) - \xi_{\gamma_k}(t)\nabla\psi(x_{\gamma_k}(t)), \quad \forall t \in [0,1].$$

Hence, using part(i) of this theorem, the continuity of $f_\Phi(\cdot,\cdot)$, that the sequence $(x_{\gamma_k}, u_{\gamma_k}, \xi_{\gamma_k})_k$ has uniform bounded variations and converges pointwise to (x, u, ξ), and that (x, u, ξ) satisfies (7) for all $t \in [0, 1]$, we obtain that the sequence $\dot{x}_{\gamma_k}$ is of bounded variations and converges pointwise to $\dot{x} \in BV([0,1];\mathbb{R}^n)$. Since $(\|\dot{x}_{\gamma_k}\|_\infty)_k$ is bounded, we conclude that the sequence $\dot{x}_{\gamma_k}$ also converges *strongly* in L^2 to $\dot{x}$. Therefore, x_{γ_k} converges strongly in the norm topology of $W^{1,2}$ to x. □

***Proof of Theorem* 4.7.** We consider k large enough so that $C_0(k) \subset \tilde{C}_0(\delta)$ and $C_1(k) \subset \tilde{C}_1(\delta)$, see (19). By Corollary 4.2, $\bar{x}_{\gamma_k} \longrightarrow \bar{x}$ *strongly* in $W^{1,2}$, and hence, for k sufficiently large, $\bar{x}_{\gamma_k}(t) \in \bar{B}_\delta(\bar{x}(t))$ for all $t \in [0,1]$, and $\bar{y}_{\gamma_k}(1) \in [-\delta, \delta]$, where

$$\bar{y}_{\gamma_k}(t) := \int_0^t \|\dot{\bar{x}}_{\gamma_k}(s) - \dot{\bar{x}}(s)\|^2 ds.$$

Thus, the triplet state $(\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k}, \bar{z}_{\gamma_k} := 0)$ solves (D_{γ_k}) for $((\bar{c}_k, 0, 0), \bar{u})$, with $\bar{x}_{\gamma_k}(t) \in \bar{B}_\delta(\bar{x}(t))$ and $\bar{u}(t) \in U(t) \cap \bar{B}_\delta(\bar{u}(t))$, for all $t \in [0,1]$, and $(\bar{x}_{\gamma_k}(1), \bar{y}_{\gamma_k}(1), \bar{z}_{\gamma_k}(1) = 0) \in C_1(k) \times [-\delta, \delta] \times [-\delta, \delta]$. Therefore, for k sufficiently large, $(\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k}, 0, \bar{u})$ is an *admissible* quadruplet for (P_{γ_k}). Using the continuity of g on $\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)$ and the definition of $J(x, u, z, u)$, we obtain that $J(x, u, z, u)$ is bounded from below. Hence, for k large enough, $\inf(P_{\gamma_k})$ is *finite*.

Fix k sufficiently large so that $\inf(P_{\gamma_k})$ is *finite*. Let $(x^n_{\gamma_k}, y^n_{\gamma_k}, z^n_{\gamma_k}, u^n_{\gamma_k})_n \in W^{1,2}([0,1]; \mathbb{R}^n) \times AC([0,1]; \mathbb{R}) \times AC([0,1]; \mathbb{R}) \times \mathcal{W}$ be a minimizing sequence for (P_{γ_k}), that is, the sequence is admissible for (P_{γ_k}) and

$$\lim_{n \longrightarrow \infty} J(x^n_{\gamma_k}, y^n_{\gamma_k}, z^n_{\gamma_k}, u^n_{\gamma_k}) = \inf(P_{\gamma_k}). \tag{54}$$

Since for each n, $x^n_{\gamma_k}$ solves (D_{γ_k}) for $(x^n_{\gamma_k}(0), u^n_{\gamma_k})$, and $(x^n_{\gamma_k}(0))_n \in C_0(k) \subset C$, then, by (13), we have that the sequence $(x^n_{\gamma_k})_n$ is uniformly bounded in $C([0,1]; \mathbb{R}^n)$ and the sequence $(\dot{x}^n_{\gamma_k})_n$ is uniformly bounded in L^2. On the other hand, from (A4.2), we have that sets $U(t)$ are compact and uniformly bounded, then, the sequence $(u^n_{\gamma_k})_n$, which is in $\mathcal{W}$, is uniformly bounded in $C([0,1]; \mathbb{R}^m)$. Moreover, its derivative sequence, $(\dot{u}^n_{\gamma_k})_n$, must be uniformly bounded in L^2. Indeed, if this is not true, then there exists a subsequence of $\dot{u}^n_{\gamma_k}$, we do not relabel, such that $\lim_{n \longrightarrow \infty} \|\dot{u}^n_{\gamma_k}\|_2 = \infty$. Using that g is bounded on $\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)$, it follows that

$$\begin{aligned} J(x^n_{\gamma_k}, y^n_{\gamma_k}, z^n_{\gamma_k}, u^n_{\gamma_k}) &\geq \min_{(c_1,c_2) \in \tilde{C}_0(\delta) \times \tilde{C}_1(\delta)} g(c_1, c_2) + \frac{1}{2} z^n_{\gamma_k}(1) \\ &= \min_{(c_1,c_2) \in \tilde{C}_0(\delta) \times \tilde{C}_1(\delta)} g(c_1, c_2) + \frac{1}{2} \|\dot{u}^n_{\gamma_k} - \dot{\bar{u}}\|_2^2 \end{aligned}$$

and hence, $\lim_{n \longrightarrow \infty} J(x^n_{\gamma_k}, y^n_{\gamma_k}, z^n_{\gamma_k}, u^n_{\gamma_k}) = \infty$, contradicting (54). Thus, also $(\dot{u}^n_{\gamma_k})_n$ is uniformly bounded in L^2. Therefore, by Arzelà-Ascoli theorem, along a subsequence (we do not relabel), the sequence $(x^n_{\gamma_k}, u^n_{\gamma_k})_n$ converges uniformly to a pair $(x_{\gamma_k}, u_{\gamma_k})$ and the sequence $(\dot{x}^n_{\gamma_k}, \dot{u}^n_{\gamma_k})_n$ converges weakly in L^2 to the pair $(\dot{x}_{\gamma_k}, \dot{u}_{\gamma_k})$. Hence, $(x_{\gamma_k}, u_{\gamma_k}) \in W^{1,2}([0,1]; \mathbb{R}^n) \times \mathcal{W}$. Moreover, the following two inequalities hold

$$\|\dot{x}_{\gamma_k} - \dot{\bar{x}}\|_2^2 \leq \liminf_{n \longrightarrow \infty} \|\dot{x}^n_{\gamma_k} - \dot{\bar{x}}\|_2^2 \;\text{ and }\; \|\dot{u}_{\gamma_k} - \dot{\bar{u}}\|_2^2 \leq \liminf_{n \longrightarrow \infty} \|\dot{u}^n_{\gamma_k} - \dot{\bar{u}}\|_2^2. \tag{55}$$

Since $C_0(k)$, $C_1(k)$, $\bar{B}_\delta(\bar{x}(t))$ and $U(t) \cap \bar{B}_\delta(\bar{u}(t))$ are closed for all $t \in [0,1]$, and from the uniform convergence, as $n \longrightarrow \infty$, of the sequence $(x^n_{\gamma_k}, u^n_{\gamma_k})$ to $(x_{\gamma_k}, u_{\gamma_k})$, we get that the inclusions $x_{\gamma_k}(0) \in C_0(k)$ and $x_{\gamma_k}(1) \in C_1(k)$, and $x_{\gamma_k}(t) \in \bar{B}_\delta(\bar{x}(t))$, and $u_{\gamma_k}(t) \in U(t) \cap \bar{B}_\delta(\bar{u}(t))$, for all $t \in [0,1]$. To prove that x_{γ_k} is the solution of (D_{γ_k}) corresponding to $(x_{\gamma_k}(0), u_{\gamma_k})$, we first use that $x^n_{\gamma_k}$ is the solution of (D_{γ_k}) for $(x^n_{\gamma_k}(0), u^n_{\gamma_k})$, that is, for $t \in [0,1]$,

$$x^n_{\gamma_k}(t) = x^n_{\gamma_k}(0) + \int_0^t \left[f_\Phi(x^n_{\gamma_k}(s), u^n_{\gamma_k}(s)) - \gamma_k e^{\gamma_k \psi(x^n_{\gamma_k}(s))} \nabla \psi(x^n_{\gamma_k}(s)) \right] ds.$$

Using that $(x^n_{\gamma_k}(t), u^n_{\gamma_k}(t)) \in [C \cap \bar{B}_\delta(\bar{x}(t))] \times [U(t) \cap \bar{B}_\delta(\bar{u}(t))]$, f_Φ is Lipschitz on $[C \cap \bar{\mathbb{B}}_\delta(\bar{x})] \times [(\mathbb{U} + \tilde{\rho}\bar{B}) \cap \bar{\mathbb{B}}_\delta(\bar{u})]$, and $(x^n_{\gamma_k}, u^n_{\gamma_k})$ converges uniformly to $(x_{\gamma_k}, u_{\gamma_k})$, then, upon taking the limit, as $n \longrightarrow \infty$, in the last equation we conclude that $(x_{\gamma_k}, u_{\gamma_k})$ satisfies the same equation, that is,

$$\dot{x}_{\gamma_k}(t) = f_\Phi(x_{\gamma_k}(t), u_{\gamma_k}(t)) - \gamma_k e^{\gamma_k \psi(x_{\gamma_k}(t))} \nabla \psi(x_{\gamma_k}(t)), \quad t \in [0, 1] \text{ a.e.}$$

We define for all $t \in [0, 1]$,

$$y_{\gamma_k}(t) := \int_0^t \|\dot{x}_{\gamma_k}(\tau) - \dot{\bar{x}}(\tau)\|^2 d\tau \;\; \text{and} \;\; z_{\gamma_k}(t) := \int_0^t \|\dot{u}_{\gamma_k}(\tau) - \dot{\bar{u}}(\tau)\|^2 d\tau.$$

Clearly we have:

- $y_{\gamma_k} \in AC([0, 1]; \mathbb{R})$, $\;\dot{y}_{\gamma_k}(t) = \|\dot{x}_{\gamma_k}(t) - \dot{\bar{x}}(t)\|^2$, $\;t \in [0, 1]$ a.e., and $y_{\gamma_k}(0) = 0$.
- $z_{\gamma_k} \in AC([0, 1]; \mathbb{R})$, $\;\dot{z}_{\gamma_k}(t) = \|\dot{u}_{\gamma_k}(t) - \dot{\bar{u}}(t)\|^2$, $\;t \in [0, 1]$ a.e., and $z_{\gamma_k}(0) = 0$.

Moreover, since $\|\dot{x}^n_{\gamma_k} - \dot{\bar{x}}\|_2^2 = y^n_{\gamma_k}(1) \in [-\delta, \delta]$ and $\|\dot{u}^n_{\gamma_k} - \dot{\bar{u}}\|_2^2 = z^n_{\gamma_k}(1) \in [-\delta, \delta]$, the two inequalities of (55) yield that

$$y_{\gamma_k}(1) \in [-\delta, \delta] \;\; \text{and} \;\; z_{\gamma_k}(1) \in [-\delta, \delta]. \tag{56}$$

Hence, $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ is admissible for (P_{γ_k}). Now using (54) and the second inequality of (55), it follows that

$$\begin{aligned}
&\inf(P_{\gamma_k}) = \lim_{n \longrightarrow \infty} J(x^n_{\gamma_k}, y^n_{\gamma_k}, z^n_{\gamma_k}, u^n_{\gamma_k}) \\
&= \lim_{n \longrightarrow \infty} \left(g(x^n_{\gamma_k}(0), x^n_{\gamma_k}(1)) + \frac{1}{2}\left(\|u^n_{\gamma_k}(0) - \bar{u}(0)\|^2 + z^n_{\gamma_k}(1) + \|x^n_{\gamma_k}(0) - \bar{x}(0)\|^2 \right)\right) \\
&= g(x_{\gamma_k}(0), x_{\gamma_k}(1)) + \frac{1}{2}\|u_{\gamma_k}(0) - \bar{u}(0)\|^2 + \frac{1}{2} \liminf_{n \longrightarrow \infty} \|\dot{u}^n_{\gamma_k} - \dot{\bar{u}}\|_2^2 + \frac{1}{2}\|x_{\gamma_k}(0) - \bar{x}(0)\|^2 \\
&\geq g(x_{\gamma_k}(0), x_{\gamma_k}(1)) + \frac{1}{2}\|u_{\gamma_k}(0) - \bar{u}(0)\|^2 + \frac{1}{2}\|\dot{u}_{\gamma_k} - \dot{\bar{u}}\|_2^2 + \frac{1}{2}\|x_{\gamma_k}(0) - \bar{x}(0)\|^2 \\
&= J(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k}).
\end{aligned}$$

Therefore, for each k, large enough, $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ is optimal for (P_{γ_k}).

As Remark 4.5 asserts that, for k large, $C_0(k) \subset C(k) \subset C$, then, Lemma 3.6 and Theorem 3.9(i) yield that the sequence $(x_{\gamma_k}, \xi_{\gamma_k})_k$, where ξ_{γ_k} is given via (16), admits a subsequence, not relabeled, having $(x_{\gamma_k})_k$ converging uniformly to some $x \in W^{1,2}([0, 1]; \mathbb{R}^n)$ with images in C, $(\dot{x}_{\gamma_k}, \xi_{\gamma_k})_k$ converging weakly in L^2 to $(\dot{x}, \xi)$ and ξ supported on $I^0(x)$.

Now, consider the sequence $(u_{\gamma_k})_k$, which is in $\mathcal{W}$. It has a uniformly bounded derivative in L^2. In fact, the admissibility of $(\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k}, 0, \bar{u})$, and the optimality of $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ for (P_{γ_k}), imply that

$$J(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k}) \leq g(\bar{x}_{\gamma_k}(0), \bar{x}_{\gamma_k}(1)) + \frac{1}{2}\|\bar{x}_{\gamma_k}(0) - \bar{x}(0)\|^2. \tag{57}$$

This, together with the continuity of g on $\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)$, the uniform boundedness of the sequences $(x_{\gamma_k})_k$ and $(\bar{x}_{\gamma_k})_k$, and the boundedness of $U(0)$, imply that for some $\hat{M} > 0$ we have that

$$\begin{aligned}
\|\dot{u}_{\gamma_k} - \dot{\bar{u}}\|_2^2 \leq 2 \big(&g(\bar{x}_{\gamma_k}(0), \bar{x}_{\gamma_k}(1)) - g(x_{\gamma_k}(0), x_{\gamma_k}(1)\big) + \|\bar{x}_{\gamma_k}(0) - \bar{x}(0)\|^2 \\
&- \|u_{\gamma_k}(0) - \bar{u}(0)\|^2 - \|x_{\gamma_k}(0) - \bar{x}(0)\|^2 \leq \hat{M}.
\end{aligned}$$

Therefore, $(u_{\gamma_k})_k$ has uniformly bounded derivative in L^2. Now since in addition we have that $x_{\gamma_k}(0) \in C_0(k) \subset C(k)$, we are in a position to apply Theorem 4.1. We obtain a subsequence (not relabeled) of u_{γ_k}, and $u \in \mathcal{W}$ such that $(x_{\gamma_k}, u_{\gamma_k})$ converges uniformly to (x, u), $\dot{u}_{\gamma_k}$ converges weakly in L^2 to $\dot{u}$, all the conclusions of Theorem 3.11 hold including that $x_{\gamma_k}(t) \in \operatorname{int} C$ for all $t \in [0, 1]$, $(\dot{x}_{\gamma_k}, \xi_{\gamma_k})$ converges *strongly* in L^2 to $(\dot{x}, \xi)$, $\dot{x} \in BV([0, 1]; \mathbb{R}^n)$, $\xi \in BV([0, 1]; \mathbb{R}^+)$, and, for all $t \in [0, 1]$, (x, u, ξ) satisfies (7)-(8) and x uniquely solves (D) for u, that is,

$$\begin{cases} \dot{x}(t) = f_\Phi(x(t), u(t)) - \xi(t)\nabla\psi(x(t)) \in f(x(t), u(t)) - \partial\varphi(x(t)), \quad \forall t \in [0, 1], \\ x(0) \in C_0 \cap \bar{B}_{\delta_o}(\bar{x}(0)). \end{cases}$$

Moreover, we have

$$\|\dot{u} - \dot{\bar{u}}\|_2^2 \le \liminf_{k \longrightarrow \infty} \|\dot{u}_{\gamma_k} - \dot{\bar{u}}\|_2^2. \tag{58}$$

We shall show that (x, u) is admissible for (P). Since $\dot{x}_{\gamma_k}$ converges strongly in L^2 to $\dot{x}$, and using (56) and (58), we have:

- $\|\dot{x} - \dot{\bar{x}}\|_2^2 = \lim_{k \longrightarrow \infty} \|\dot{x}_{\gamma_k} - \dot{\bar{x}}\|_2^2 = \lim_{k \longrightarrow \infty} y_{\gamma_k}(1) \overset{(56)}{\in} [-\delta, \delta]$.
- $\|\dot{u} - \dot{\bar{u}}\|_2^2 \overset{(58)}{\le} \liminf_{k \longrightarrow \infty} \|\dot{u}_{\gamma_k} - \dot{\bar{u}}\|_2^2 = \liminf_{k \longrightarrow \infty} z_{\gamma_k}(1) \overset{(56)}{\in} [-\delta, \delta]$.

Hence, $\|\dot{x} - \dot{\bar{x}}\|_2^2 \le \delta$ and $\|\dot{u} - \dot{\bar{u}}\|_2^2 \le \delta$. Since $x_{\gamma_k}(1) \in C_1(k)$, (20)(b) implies that $x(1) \in C_1 \cap \bar{B}_{\delta_o}(\bar{x}(0))$. Furthermore, the two inclusions $x_{\gamma_k}(t) \in \bar{B}_\delta(\bar{x}(t))$ and $u_{\gamma_k}(t) \in U(t) \cap \bar{B}_\delta(\bar{u}(t))$, for all $t \in [0, 1]$, together with the uniform convergence of $(x_{\gamma_k}, u_{\gamma_k})$ to (x, u), give that $x(t) \in \bar{B}_\delta(\bar{x}(t))$ and $u(t) \in U(t) \cap \bar{B}_\delta(\bar{u}(t))$, for all $t \in [0, 1]$. Therefore, (x, u) is admissible for (P). Hence, the optimality of $(\bar{x}, \bar{u})$ for (P) yields that

$$g(\bar{x}(0), \bar{x}(1)) \le g(x(0), x(1)). \tag{59}$$

Now, the uniform convergence of $\bar{x}_{\gamma_k}$ to $\bar{x}$, (57), (59), the continuity of g, and the convergence of $x_{\gamma_k}(0)$ to $x(0)$, imply that

$$u(0) = \bar{u}(0) \quad \text{and} \quad \liminf_{k \longrightarrow \infty} \left(\|\dot{u}_{\gamma_k} - \dot{\bar{u}}\|_2^2 \right) = 0, \quad \text{and} \tag{60}$$

$$x(0) = \bar{x}(0) \quad \text{and} \quad g(\bar{x}(0), \bar{x}(1)) = g(x(0), x(1)). \tag{61}$$

The equality (60) gives the existence of a subsequence of u_{γ_k}, we do not relabel, such that $\dot{u}_{\gamma_k}$ converges *strongly* in L^2 to $\dot{\bar{u}}$. It results that u_{γ_k} converges uniformly to $\bar{u}$, and hence, $u = \bar{u}$. Consequently, u_{γ_k} converges strongly in $\mathcal{W}$ to $\bar{u}$. Moreover, as $u = \bar{u}$, the functions x and $\bar{x}$ solve the dynamic (D) with the same control $\bar{u}$ and initial condition, see (61), hence, by the uniqueness of the solution of (D) we have $x = \bar{x}$. Using Lemma 3.5, we obtain that also $\xi = \bar{\xi}$. Therefore,

$$x_{\gamma_k} \xrightarrow[C([0,1];\mathbb{R}^n)]{\text{uniformly}} \bar{x} \quad \text{and} \quad (\dot{x}_{\gamma_k}, \xi_{\gamma_k}) \xrightarrow[L^2([0,1];\mathbb{R}^n \times \mathbb{R}^+)]{\text{strongly}} (\dot{\bar{x}}, \bar{\xi}).$$

This yields that $(y_{\gamma_k}, z_{\gamma_k}) \longrightarrow (0, 0)$ in the strong topology of $W^{1,1}([0, 1]; \mathbb{R}^+ \times \mathbb{R}^+)$.

Since $x_{\gamma_k}(1) \in \left[\left(C_1 \cap \bar{B}_{\delta_o}(\bar{x}(1))\right) - \bar{x}(1) + \bar{x}_{\gamma_k}(1)\right] \cap (\operatorname{int} C)$ and $\bar{x}_{\gamma_k}(1)$ converges to $\bar{x}(1)$, it follows that $x_{\gamma_k}(1) \in \left[\left(C_1 \cap \bar{B}_{\delta_o}(\bar{x}(1))\right) + \tilde{\rho}B\right] \cap (\operatorname{int} C)$, for k sufficiently large. On the other hand, the definition of $C_0(k)$ and the convergence of ρ_k to 0 yield that, for k large enough, $x_{\gamma_k}(0) \in \left[\left(C_0 \cap \bar{B}_{\delta_o}(\bar{x}(0))\right) + \tilde{\rho}B\right] \cap (\operatorname{int} C)$. This terminates the proof of Theorem 4.7. ◻

Acknowledgements The authors wish to thank the Editor and the anonymous referees for their valuable comments that greatly improved the paper.

References

1. Adam, L., Outrata, J.: On optimal control of a sweeping process coupled with an ordinary differential equation. Discrete Contin. Dyn. Syst. B **19**, 2709–2738 (2014)
2. Aubin, J.-P.: Viability Theory. Modern Birkhäuser Classics. Birkhäuser/Springer, Basel-Boston-Berlin (2009)
3. Brokate, M., Krejčí, P.: Optimal control of ODE systems involving a rate independent variational inequality. Discret. Contin. Dyn. Syst. Ser. B **18**, 331–348 (2013)
4. Brudnyi, A., Brudnyi, Y.: Methods of Geometric Analysis in Extension and Trace Problems, Volume 1, Monographs in Mathematics, 102. Birkhäuser/Springer Basel AG, Basel (2012)
5. Cao, T.H., Colombo, G., Mordukhovich, B., Nguyen, D.: Optimization of Fully Controlled Sweeping Processes. J. Differ. Equ. **295**, 138–186 (2021)
6. Cao, T.H., Mordukhovich, B.: Optimal control of a perturbed sweeping process via discrete approximations. Discret. Contin. Dyn. Syst. Ser. B **21**, 3331–3358 (2016)
7. Cao, T.H., Mordukhovich, B.: Optimality conditions for a controlled sweeping process with applications to the crowd motion model. Disc. Cont. Dyn. Syst. Ser. B **22**, 267–306 (2017)
8. Cao, T.H., Mordukhovich, B.: Optimal control of a nonconvex perturbed sweeping process. J. Differ. Equ. **266**, 1003–1050 (2019)
9. Clarke, F.H.: Optimization and Nonsmooth Analysis. John Wiley, New York (1983)
10. Clarke, F.H.: Functional Analysis, Calculus of Variations and Optimal Control. Graduate Texts in Mathematics, 264. Springer, London (2013)
11. Clarke, F.H., Yu. Ledyaev, Yu., Stern, R.J., Wolenski, P.R.: Nonsmooth Analysis and Control Theory, Graduate Texts in Mathematics, 178, Springer-Verlag, New York (1998)
12. Clarke, F.H., Stern, R.J., Wolenski, P.R.: Proximal smoothness and the lower-C^2 property. J. Convex Anal. **2**, 117–144 (1995)
13. Colombo, G., Henrion, R., Hoang, N.D., Mordukhovich, B.S.: Optimal control of the sweeping process. Dyn. Contin. Discret. Impuls. Syst. Ser. B **19**, 117–159 (2012)
14. Colombo, G., Henrion, R., Hoang, N.D., Mordukhovich, B.S.: Optimal control of the sweeping process over polyhedral controlled sets. J. Differ. Equ. **260**(4), 3397–3447 (2016)
15. Colombo, G., Marigonda, A.: Differentiability properties for a class of non-convex functions. Calc. Var. **25**, 1–31 (2005)
16. Colombo, G., Mordukhovich, B., Nguyen, D.: Optimization of a perturbed sweeping process by constrained discontinuous controls. SIAM J. Control Optim. **58**(4), 2678–2709 (2020)
17. de Pinho, M.d.R., Ferreira, M.M.A., Smirnov, G.V.: Optimal Control Involving Sweeping Processes. Set-Valued Var. Anal. **27**, no. 2, 523–548 (2019)
18. de Pinho, Md.R., Ferreira, M.M.A.: Correction to: Optimal Control Involving Sweeping Processes. Set-Valued Var. Anal. **27**, 1025–1027 (2019)
19. de Pinho, Md.R., Ferreira, M.M.A., Smirnov, G.V.: Optimal Control with Sweeping Processes. Numer. Method. J. Optim. Theory Appl. **185**, 845–858 (2020)
20. de Pinho, Md.R., Ferreira, M.M.A., Smirnov, G.V.: Necessary conditions for optimal control problems with sweeping systems and end point constraints. Optimization **71**, 3363–3381 (2021)
21. Federer, H.: Curvature measures. Trans. Amer. Math. Soc. **93**, 418–491 (1959)
22. Hermosilla, C., Palladino, M.: Optimal Control of the Sweeping Process with a Nonsmooth Moving Set. SIAM J. Control Optim. **60**(5), 2811–2834 (2022)
23. Hiriart-Urruty, J.-B.: Extension of Lipschitz functions. J. Math Anal. Appl. **77**, 539–554 (1980)
24. Hoehener, D.: Variational approach to second-order optimality conditions for control problems with pure state constraints. SIAM J. Control Optim. **50**(3), 1139–1173 (2012)
25. Kadets, V.: A course in functional analysis and measure theory. Translated from the 2006 Russian edition by Andrei Iacob, Universitext. Springer, Cham (2018)
26. Loewen, P.D., Rockafellar, R.T.: The Adjoint Arc in Nonsmooth Optimization. Trans. Am. Math. Soc. **325**(1), 39–72 (1991)
27. Luenberger, D.G: Optimization by vector space methods. John Wiley & Sons, Inc. (1969)
28. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation. I: Basic Theory, Springer, Berlin (2006)

29. Moreau, J.J.: Rafle par un convexe variable, I. Trav. Semin. d'Anal. Convexe, Montpellier 1, Exposé **15**, 36 pp. (1971)
30. Moreau, J.J.: Rafle par un convexe variable, II. Trav. Semin. d'Anal. Convexe, Montpellier 2, Exposé **3**, 43 pp. (1972)
31. Moreau, J.J.: Evolution problem associated with a moving convex set in a Hilbert space. J. Differ. Equ. **26**, 347–374 (1977)
32. Nour, C., Stern, R.J., Takche, J.: Proximal smoothness and the exterior sphere condition. J. Convex Anal. **16**(2), 501–514 (2009)
33. Nour, C., Zeidan, V.: Numerical solution for a controlled nonconvex sweeping process. IEEE Control Syst. Lett. **6**, 1190–1195 (2022)
34. Nour, C., Zeidan, V.: Optimal control of nonconvex sweeping processes with separable endpoints: Nonsmooth maximum principle for local minimizers. J. Differ. Equ. **318**, 113–168 (2022)
35. Poliquin, R.A., Rockafellar, R.T., Thibault, L.: Local differentiability of distance functions. Trans. Am. Math. Soc. **352**, 5231–5249 (2000)
36. Rockafellar, R.T.: Clarke's tangent cones and the boundaries of closed sets in $\mathbb{R}^n$. Nonlinear Anal. Theor. Meth. Appl. **3**, 145–154 (1979)
37. Rockafellar, R.T., Wets, R.J.-B.: Variational analysis. Grundlehren der Mathematischen Wissenschaften, 317, Springer-Verlag, Berlin (1998)
38. Vinter, R.B.: Optimal Control, Birkhäuser. Foundations and Applications, Boston, Systems and Control (2000)
39. Zeidan, V., Nour, C., Saoud, H.: A nonsmooth maximum principle for a controlled nonconvex sweeping process. J. Differ. Equ. **269**(11), 9531–9582 (2021)
40. Zeidan, V.: A modified Hamilton-Jacobi approach in the generalized problem of Bolza. Appl. Math. Optim. **11**, 97–109 (1984)