Appl. Set-Valued Anal. Optim. 5 (2023), No. 2, pp. 193-212
Available online at http://asvao.biemdas.com
https://doi.org/10.23952/asvao.5.2023.2.06

NONSMOOTH OPTIMALITY CRITERION FOR A $W^{1,2}$-CONTROLLED SWEEPING PROCESS: NONAUTONOMOUS PERTURBATION

CHADI NOUR[1], VERA ZEIDAN[2,*]

[1] *Department of Computer Science and Mathematics, Lebanese American University, Byblos Campus, Byblos, Lebanon*
[2] *Department of Mathematics, Michigan State University, East Lansing, MI 48824-1027, USA*

Abstract. We extend to $C\times W^{1,2}$-*local minimizers* and *nonautonomous* perturbation function, the necessary optimality conditions recently derived, via continuous-time approximations, for $W^{1,2}\times W^{1,2}$-*local minimizers* of an optimal control problem governed by a controlled nonconvex sweeping process with *autonomous* perturbation function.
Keywords. Controlled and perturbed nonconvex sweeping process; Continuous-time approximations; Local minimizers; Necessary optimality conditions; Nonsmooth analysis.
2020 Mathematics Subject Classification. 49J52, 49K15, 49J21.

1. INTRODUCTION

Moreau's *sweeping process* is a mathematical model introduced in [1, 2, 3] to describe an elastoplastic mechanical system. Since then, this model appeared in several fields such as mechanics, electrical circuits, engineering, economics, crowd motion control, traffic equilibria, hysteresis phenomena, etc. (see, e.g., [4]). Since the dynamic of this model is an evolution differential inclusions involving an unbounded and discontinuous multifunction, namely, the *normal cone* to a set, the sweeping process model falls outside the scope of standard differential inclusions. Therefore, new techniques are required to address optimal control problems over sweeping processes.

In this paper, we are interested in deriving necessary optimality conditions, phrased in terms of the weak-Pontryagin-type maximum principle, for the optimal control problems governed by $W^{1,2}$-controlled and *nonautonomously* perturbed sweeping processes. The autonomous case was successfully treated in [5, 6, 7] via *discrete approximations*, and in [8] via *continuous approximations*. In these references, the authors considered $W^{1,2}\times W^{1,2}$-*local minimizers*, and obtained, in addition to necessary optimality conditions, the *strong* convergence of velocities, that is, the optimal states of the original problem are *strongly* approximated, in the $W^{1,2}$-norm, by the optimal states of the approximating problems. It is worth to mention that, unlike the case for standard optimal control problems, where the optimal control functions have *weak* regularity properties, applications of optimal control problems over *sweeping processes* have shown to

*Corresponding author.
E-mail addresses: cnour@lau.edu.lb (C. Nour), zeidan@msu.edu (V. Zeidan).
Received May 22, 2023; Accepted June 26, 2023.

possess optimal control functions with *strong* regularity properties, such as, $W^{1,2}$ (see, e.g., [6], [7, Examples 9.1 & 9.2]).

The continuous approximation used in [8] was based on approximating the normal cone by an *exponential penalization* term. This innovative technique was introduced in [9, 10], and further developed in [11, 12, 13, 14, 15]. The *autonomous* assumption on the perturbation function in [8], and also in [5, 6, 7], played a crucial role in the derivation of the necessary optimality conditions using $W^{1,2}$-controls, especially, in the proof of the strong convergence of velocities.

The goal of this paper is to extend the weak Pontryagin maximum principle derived in [8] to the case where the perturbation function is *nonautonomous* and the t-dependence is merely *measurable*. In order to reach this goal, it turns out that the strong convergence of velocities has to be weakened in the following sense: While the convergence of the approximating problems solutions to those of the original problem remains *uniform* for the *state component* and *strong* in $W^{1,2}$ for the *control component*, it is rather *weak** in L^∞ for the state *velocities*. Consequently, our necessary optimality conditions, which coincide with those obtained in [8], are derived for $C\times W^{1,2}$-*local minimizers*. For this class of local minimizers, our formulation here of the approximating problems via the exponential penalization technique is different than that in [8], and so is the proof of existence of optimal solutions for these problems (Lemma 3.1) as well as the proof of Lemma 3.2.

The paper is organized as follows. In the next section, we present our basic notations, define our optimal control problem (P) governed by a $W^{1,2}$-controlled and perturbed sweeping process, list the hypotheses satisfied by the data of (P), and state the main result of the paper, namely, necessary optimality conditions in the form of weak Pontryagin principle for $C\times W^{1,2}$-local minimizers of (P). Section 3 is devoted to the proof of our main result. An example illustrating the utility of our main result is provided in Section 4.

2. Main result

2.1. Basic notations.

The following are the basic notations and definitions used in this paper:

- For the Euclidean norm and the usual inner product, we use $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$, respectively.
- For $y\in\mathbb{R}^n$ and $r>0$, we define the open (resp. closed) ball centered at y with radius r as $B_r(y):=y+rB$ (resp. $\bar{B}_r(y):=y+r\bar{B}$), where B and $\bar{B}$ denotes the open and the closed unit ball, respectively.
- For $C\subset\mathbb{R}^n$, the boundary, the interior, the closure, the convex hull, the complement, and the polar of C are denoted by $\operatorname{bdry} C$, $\operatorname{int} C$, $\operatorname{cl} C$, $\operatorname{conv} C$, C^c, and C°, respectively.
- The distance from a point $x\in\mathbb{R}^n$ to a set $C\subset\mathbb{R}^n$ is denoted by $d(x,C)$.
- For an extended-real-valued function $h\colon\mathbb{R}^n\longrightarrow\mathbb{R}\cup\{\infty\}$, the effective domain of h is $\operatorname{dom} h$, and the epigraph of h is $\operatorname{epi} h$.
- For a multifunction $F\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^m$, we denote by $\operatorname{Gr} F\subset\mathbb{R}^n\times\mathbb{R}^m$ 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. For $C\subset\mathbb{R}^d$ compact, the set of continuous functions from C to $\mathbb{R}^n$ is denoted by $C(C;\mathbb{R}^n)$.

- The set of all $m \times n$-matrix functions on $[a,b]$ is denoted by $\mathcal{M}_{m\times n}([a,b])$.
- A function $h\colon [a,b] \longrightarrow \mathbb{R}^n$ is said to be a BV-function if h has a bounded variation. 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., [16, p.115]).
- The space $\mathrm{C}^*([a,b];\mathbb{R})$ denotes the dual of $\mathrm{C}([a,b];\mathbb{R})$, equipped with the supremum norm. The induced norm on $\mathrm{C}^*([a,b];\mathbb{R})$ is denoted by $\|\cdot\|_{\mathrm{T.V.}}$. As a consequence of Riesz representation theorem, we can interpret the elements of $\mathrm{C}^*([a,b];\mathbb{R})$ 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 $\mathrm{C}^*([a,b];\mathbb{R})$ it corresponds a unique element in $NBV[a,b]$ related through the Stieltjes integral and both elements have the same total variation. The set $\mathrm{C}^{\oplus}(a,b)$ designates the subset of $\mathrm{C}^*([a,b];\mathbb{R})$ taking nonnegative values on nonnegative-valued functions in $\mathrm{C}([a,b];\mathbb{R})$.
- By $\mathrm{W}^{k,p}([a,b];\mathbb{R}^n)$, $k\in\mathbb{N}$ and $p\in[0,+\infty]$, we denote the classical Sobolev space. Hence, the set of all absolutely continuous functions from $[a,b]$ to $\mathbb{R}^n$ is $\mathrm{W}^{1,1}([a,b];\mathbb{R}^n)$. Note that in this paper, the Sobolev space $\mathrm{W}^{1,2}([a,b];\mathbb{R}^m)$ will be considered with the norm $\|u(\cdot)\|_{W^{1,2}} := \|u(\cdot)\|_\infty + \|\dot{u}(\cdot)\|_2$. Hence, the convergence of a sequence u_n strongly in the norm topology of the space $W^{1,2}([a,b];\mathbb{R}^m)$ is equivalent to the uniform convergence of u_n on $[a,b]$ and the strong convergence in L^2 of its derivative $\dot{u}_n$.
- For $C\subset\mathbb{R}^n$ closed and $c\in C$, we denote by $N_C^P(c)$, $N_C^L(c)$, and $N_C(c)$, the *proximal*, the *Mordukhovich* (or *limiting*), and the *Clarke normal cones* to C at c, respectively.
- For $F\colon [a,b] \rightrightarrows \mathbb{R}^m$ a multifunction with closed and nonempty values, $\bar{N}^L_{F(t)}(y)$ 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 $N^L_{F(\cdot)}(\cdot)$.
- For $h\colon \mathbb{R}^n \longrightarrow \mathbb{R}\cup\{\infty\}$ lower semicontinuous and $x\in\operatorname{dom} h$, 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 $h\colon \mathbb{R}^n \longrightarrow \mathbb{R}\cup\{\infty\}$ is $\mathrm{C}^{1,1}$ near $x\in\operatorname{dom} h$, then the *Clarke generalized Hessian* of h at x is denoted by $\partial^2 h(x)$. On the other hand, if $H\colon \mathbb{R}^n\longrightarrow\mathbb{R}^n$ is Lipschitz near $x\in\mathbb{R}^n$, then the *Clarke generalized Jacobian* of H at x is denoted by $\partial H(x)$.
- For $F\colon [a,b]\rightrightarrows\mathbb{R}^m$ a lower semicontinuous multifunction with closed and nonempty values, we define

$$\begin{aligned}&\partial_x^{>} d(x,F(t))\\ &:= \operatorname{conv}\left\{\zeta = \lim_{i\longrightarrow\infty}\zeta_i :\ \|\zeta_i\|=1,\ \zeta_i\in N^P_{F(t_i)}(x_i) \text{ and } (t_i,x_i)\xrightarrow{\operatorname{Gr}F}(t,x)\right\},\end{aligned}$$

where $(t_i,x_i)\xrightarrow{\operatorname{Gr}F}(t,x)$ signifies that $(t_i,x_i)\longrightarrow(t,x)$ with $x_i\in F(t_i)$ for all i. Note that $\partial_x^{>} d(x,F(t))$ coincides with $\partial_x^{>} g(t,x)$ of [17, p.121] for $g(t,x) := d(x,F(t))$, see [18, Corollary 2.2].

- For $r > 0$, a closed and nonempty set $S \subset \mathbb{R}^n$ is said to be r-prox-regular if for all $s \in \operatorname{bdry} S$ and for all $\zeta \in N_S^P(s)$ unit, we have $\langle \zeta, x-s \rangle \leq \frac{1}{2r}\|x-s\|^2$ for all $x \in S$. For more information about prox-regularity, see [19].

2.2. **Statement of problem (P) and hypotheses.** We consider the following fixed time Mayer-type optimal control problem involving $\mathrm{W}^{1,2}$-controlled and perturbed sweeping systems

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

where $g\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$, $f\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$, $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$, ∂ stands for the Clarke subdifferential, $C := \operatorname{dom}\varphi$ is the zero-sublevel set of a function $\psi\colon \mathbb{R}^n \longrightarrow \mathbb{R}$, 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

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

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 *controlled* and *perturbed sweeping process* (D), called the *dynamic* of (P). An admissible pair $(\bar{x},\bar{u})$ for (P) is said to be a $\mathrm{C}{\times}\mathrm{W}^{1,2}$-*local minimizer* if there exists $\delta > 0$ such that

$$g(\bar{x}(0),\bar{x}(1)) \leq g(x(0),x(1)), \tag{2.1}$$

for all (x,u) admissible for (P) with $\|x-\bar{x}\|_\infty \leq \delta$, $\|u-\bar{u}\|_\infty \leq \delta$, and $\|\dot{u}-\dot{\bar{u}}\|_2^2 \leq \delta$. Note that if inequality (2.1) is satisfied by any admissible pairs (x,u), then $(\bar{x},\bar{u})$ is called a *global minimizer* (or an *optimal solution*) for (P).

Let $(\bar{x},\bar{u})$ be a $\mathrm{C}{\times}\mathrm{W}^{1,2}$-local minimizer for (P) with associated δ such that the following hypotheses hold for $\bar{\mathbb{B}}_\delta(\bar{x}) := \bigcup_{t\in[0,1]} \bar{B}_\delta(\bar{x}(t))$:

H1: There exist $\tilde{\rho} > 0$ and $M_\ell > 0$ such that $f(\cdot,x,u)$ is Lebesgue-measurable for $(x,u) \in [C \cap \bar{\mathbb{B}}_\delta(\bar{x})] \times [(\mathbb{U} + \tilde{\rho}\bar{B}) \cap \bar{\mathbb{B}}_\delta(\bar{u})]$; and for a.e. $t \in [0,1]$ we have that: $(x,u) \mapsto f(t,x,u)$ is M_ℓ-Lipschitz on $[C \cap \bar{B}_\delta(\bar{x}(t))] \times \left[\left(U(t) + \tilde{\rho}\bar{B}\right) \cap \bar{B}_\delta(\bar{u}(t))\right]$; and $\|f(t,x,u)\| \leq M_\ell$ for all $(x,u) \in [C \cap \bar{B}_\delta(\bar{x}(t))] \times [U(t) \cap \bar{B}_\delta(\bar{u}(t))]$.

H2: The set $C := \operatorname{dom}\varphi$ is given by $C = \{x \in \mathbb{R}^n : \psi(x) \leq 0\}$, where $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}$.

H2.1: There exists $\rho > 0$ such that ψ is $\mathrm{C}^{1,1}$ on $C + \rho B$.

H2.2: There is a constant $\eta > 0$ such that $\|\nabla\psi(x)\| > 2\eta$ for all x satisfying $\psi(x) = 0$.

H2.3: The function ψ is coercive, that is, $\lim_{\|x\|\longrightarrow\infty} \psi(x) = +\infty$.

H2.4: The set C has a connected interior.

This hypothesis is only required to guarantee that C is compact. Hence, (H2.3) can be replaced by the boundedness of C.

This hypothesis is only imposed to obtain the extension function Φ of φ, see [14, Remark 3.2 & Lemma 3.4(*iii*)]. Thus, when such an extension is readily available, as is the case when φ is the *indicator* function of C, hypothesis (H2.4) is omitted.

H3: The function φ is globally Lipschitz on C and C^1 on $\operatorname{int} C$. Moreover, the function $\nabla\varphi$ is globally Lipschitz on $\operatorname{int} C$.

H4: For the sets C_0, C_1, and $U(\cdot)$ we have that:

- **H4.1**: The set $C_0 \subset C$ is nonempty and closed.
- **H4.2**: 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.
- **H4.3**: The set $C_1 \subset \mathbb{R}^n$ is nonempty and closed.
- **H4.4**: The multifunction $U(\cdot)$ is lower semicontinuous.
- **H4.5**: The multifunction $U(\cdot)$ satisfies the constraint qualification (CQ) at $\bar{u}$, that is,

$$\operatorname{conv}(\bar{N}^L_{U(t)}(\bar{u}(t))) \text{ is pointed } \forall t\in[0,1].$$

H5: There exist $\tilde{\rho}>0$ and $L_g>0$ such that g is L_g-Lipschitz on $\tilde{C}_0(\delta)\times\tilde{C}_1(\delta)$, where

$$\tilde{C}_i(\delta) := \left[(C_i\cap\bar{B}_\delta(\bar{x}(i))) + \tilde{\rho}\bar{B}\right]\cap C, \text{ for } i=0,1.$$

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

Remark 2.1. Since f satisfies (H1), then by [20, Theorem 1] applied to each component, f_i, of $f=(f_1,\cdots,f_n)$, there exists a function $\tilde{f}\colon [0,1]\times\mathbb{R}^n\times\mathbb{R}^m \longrightarrow \mathbb{R}^n$, such that, for almost all $t\in[0,1]$, $\tilde{f}(t,\cdot,\cdot)$ is globally Lipschitz on $\mathbb{R}^n\times\mathbb{R}^m$, and $f(t,x,u)=\tilde{f}(t,x,u)$ for all $(x,u)\in [C\cap\bar{B}_\delta(\bar{x}(t))]\times\left[\left(U(t)+\tilde{\rho}\bar{B}\right)\cap\bar{B}_\delta(\bar{u}(t))\right]$. Moreover, there exists a new constant $M\geq M_\ell$ such that $\tilde{f}$ satisfies the assumption (A1) of [14], in which the constant multifunction U is replaced by $U(\cdot)\cap\bar{B}_\delta(\bar{u}(\cdot))$. Since in this paper we only consider *local* optimality notions, then, without loss of generality, we shall use the function f instead of its extension $\tilde{f}$. In particular, we use that f satisfies the assumption (A1) of [14], and hence, all the results of [14, Sections 3, 4 & 5] are valid. Now, by [14, Lemma 3.4], C is $\frac{\eta}{M_\psi}$-prox-regular, and φ admits a C^1-extension Φ from C to $\mathbb{R}^n$ satisfying $\partial\varphi(x)=\{\nabla\Phi(x)\}+N_C(x)$ for all $x\in C$, with

$$N_C(x)=\{\lambda\nabla\psi(x):\lambda\geq 0\},\ \forall x\in\operatorname{bdry} C, \tag{2.2}$$

and for some $K>0$,

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

This gives that (D) can be equivalently phrased in terms of the normal cone to C and the extension Φ of φ, as follows

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

where $f_\Phi\colon [0,1]\times\mathbb{R}^n\times\mathbb{R}^m\longrightarrow\mathbb{R}^n$ is defined by

$$f_\Phi(t,x,u) := f(t,x,u)-\nabla\Phi(x), \quad \forall(t,x,u)\in[0,1]\times\mathbb{R}^n\times\mathbb{R}^m,$$

and hence, $\|f_\Phi(t,x,u)\|\leq\bar{M}:=M+K$.

The following notations and facts, extracted from [14], will be used throughout the paper.

For more information about the (CQ) property, see [8, Remark 5.2].

- For any (x,u) solution of (D), we have by [14, Equation 50] that
$$\|\dot{x}(t)-f_{\Phi}(t,x(t),u(t))\| \le \|f_{\Phi}(t,x(t),u(t))\| \le \bar{M},\ \ t\in[0,1] \text{ a.e.},$$
and hence, $\|\dot{x}\|_{\infty} \le 2\bar{M}$.
- For given $x(\cdot)\colon [0,1]\to\mathbb{R}^n$, we define
$$I^0(x) := \{t\in[0,1] : x(t)\in \operatorname{bdry} C\} \text{ and } I^{\text{-}}(x) := [0,1]\setminus I^0(x).$$
- We define the set $\mathcal{U}$ by
$$\mathcal{U} := \{u : [0,1]\to\mathbb{R}^m : u \text{ is measurable and } u(t)\in U(t),\ t\in[0,1] \text{ a.e.}\}.$$
- For $(x,u)\in W^{1,1}([0,1];\mathbb{R}^n)\times\mathcal{U}$ with $x(0)\in C_0$ and $x(t)\in C$ for all $t\in[0,1]$, we have from (2.2) and Filippov selection theorem ([21, Theorem 2.3.13]) that 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
$$\dot{x}(t) = f_{\Phi}(t,x(t),u(t)) - \xi(t)\nabla\psi(x(t)),\quad t\in[0,1] \text{ a.e.} \tag{2.3}$$
In this case, the nonnegative function ξ supported in $I^0(x)$ with (x,u,ξ) satisfying equation (2.3), is unique, belongs to $L^{\infty}([0,1];\mathbb{R}^+)$, and
$$\begin{cases} \xi(t)=0 & \text{for } t\in I^{\text{-}}(x),\\ \xi(t)=\frac{\|\dot{x}(t)-f_{\Phi}(t,x(t),u(t))\|}{\|\nabla\psi(x(t))\|}\in\left[0,\frac{\bar{M}}{2\eta}\right] & \text{for } t\in I^0(x) \text{ a.e.},\\ \|\xi\|_{\infty}\le\frac{\bar{M}}{2\eta}. \end{cases} \tag{2.4}$$
- Since $(\bar{x},\bar{u})$ solves the dynamic (D), we denote by $\bar{\xi}$ the corresponding function in $L^{\infty}([0,1];\mathbb{R}^+)$ such that $(\bar{x},\bar{u},\bar{\xi})$ satisfies (2.3) and (2.4).
- Let $(\gamma_k)_k$ be a sequence satisfying
$$\gamma_k > \frac{2\bar{M}}{\eta} \text{ for all } k\in\mathbb{N}, \text{ and } \gamma_k \xrightarrow[k\longrightarrow\infty]{} \infty. \tag{2.5}$$
- The sequence $(\alpha_k)_k$ is defined by
$$\alpha_k := \frac{\ln\left(\frac{\eta\gamma_k}{2\bar{M}}\right)}{\gamma_k},\quad k\in\mathbb{N}. \tag{2.6}$$
By (2.5) and (2.6), we have that
$$\gamma_k e^{-\alpha_k\gamma_k} = \frac{2\bar{M}}{\eta},\ \ \alpha_k>0,\ \ \alpha_k \searrow \text{ and } \lim_{k\longrightarrow\infty}\alpha_k = 0. \tag{2.7}$$
- The sequence $(\rho_k)_k$ is defined by $\rho_k := \frac{\alpha_k}{\eta}$ for all $k\in\mathbb{N}$. By (2.7) we have that $\rho_k>0$ for all $k\in\mathbb{N}$, $\rho_k\searrow$ and $\lim_{k\longrightarrow\infty}\rho_k = 0$.
- For $k\in\mathbb{N}$, the *approximating* system (D_{γ_k}) is defined as
$$(D_{\gamma_k})\ \begin{cases} \dot{x}(t) = f_{\Phi}(t,x(t),u(t)) - \gamma_k e^{\gamma_k\psi(x(t))}\nabla\psi(x(t)) \ \text{ a.e. } t\in[0,1],\\ x(0)\in C. \end{cases}$$

Lemma 4.1 of [14] yields that, 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 $(\|\dot{x}_{\gamma_k}\|_2)_k$ is uniformly bounded. To such a solution x_{γ_k}, we associate

$$\xi_{\gamma_k}(\cdot) := \gamma_k e^{\gamma_k \psi(x_{\gamma_k}(\cdot))}. \tag{2.8}$$

- For $k \in \mathbb{N}$, we define the set $C(k) := \{x \in C : \psi(x) \leq -\alpha_k\} \subset C$.
- The sequence $(\bar{c}_{\gamma_k})_k$ is defined by

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

From [14, Remark 3.6(ii)] we have that, for k sufficiently large, $\bar{c}_{\gamma_k} \in \operatorname{int} C(k)$. Moreover, $\bar{c}_{\gamma_k} \longrightarrow \bar{x}(0)$ as $k \longrightarrow \infty$.
- For each $k \in \mathbb{N}$, we denote by $\bar{x}_{\gamma_k}$ the unique solution of (D_{γ_k}) corresponding to $(\bar{c}_{\gamma_k}, \bar{u})$. Theorem 4.1 of [14] yields that $\bar{x}_{\gamma_k}$ converges in C uniformly to $\bar{x}$.
- We fix $\delta_o > 0$ such that

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

where $r_o > 0$ is the constant in [14, Theorem 3.1(iii)], and $\hat{r}_{\bar{x}(0)} > 0$ with $\hat{k}_{\bar{x}(0)} \in \mathbb{N}$ are the constants in [14, Remark 3.6(ii)] corresponding to $c := \bar{x}(0)$.
- We define the set $C_0(k)$ by

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

One can easily verify that

$$\lim_{k\to\infty} C_0(k) = C_0 \cap \bar{B}_{\delta_o}(\bar{x}(0)) \text{ and } C_0(k) \subset \tilde{C}_0(\delta), \text{ for } k \text{ large enough}. \tag{2.9}$$

Moreover, from [14, Theorem 3.1(iii) & Remark 3.6(ii)], we have that, for k sufficiently large,

$$C_0(k) \subset C(k) \subset C. \tag{2.10}$$

- We define the set $C_1(k)$ by

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

One can easily verify that

$$\lim_{k\to\infty} C_1(k) = C \cap C_1 \cap \bar{B}_{\delta_o}(\bar{x}(1)) \text{ and } C_1(k) \subset \tilde{C}_1(\delta) \text{ for } k \text{ large enough}. \tag{2.11}$$

2.3. **Statement of the main result.** Before presenting the main result of this paper, namely the necessary optimality conditions for a given $C{\times}W^{1,2}$-local minimizer of (P), we establish the following existence of optimal solution theorem for (P), which is parallel to [5, 7, Theorems 4.1] where a discretization technique is used.

Theorem 2.1 (Existence of optimal solution for (P))**.** *Assume that* (H2)-(H4.3) *hold, and that*:

- *For fixed $(x,u) \in C \times \mathbb{U}$, $f(\cdot,x,u)$ is Lebesgue-measurable; and there exists $M > 0$ such that, for* a.e. *$t \in [0,1]$, $(x,u) \mapsto f(t,x,u)$ is continuous on $C \times U(t)$; for all $u \in U(t)$, $x \mapsto f(t,x,u)$ is M-Lipschitz on C; and $\|f(t,x,u)\| \leq M$ for all $(x,u) \in C \times U(t)$.*
- *The function $g\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is lower semicontinuous.*
- *A minimizing sequence (x_j,u_j) for (P) exists such that $(\|\dot{u}_j\|_2)_j$ is bounded.*
- *Problem (P) has at least one admissible pair (y_o,v_o) with $(y_o(0),y_o(1)) \in \operatorname{dom} g$.*

Then problem (P) admits a global optimal solution $(\tilde{x},\tilde{u})$ such that, along a subsequence, we have

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

Proof. The fact that (P) admits an admissible pair (y_o,v_o) with $(y_o(0),y_o(1)) \in \operatorname{dom} g$ yields that $\inf_{(x,u)}(P) < \infty$. Since all admissible solutions (x,u) of (P) have $(x(0),x(1))$ in the compact set $C_0 \times (C_1 \cap C)$, then the lower semicontinuity of g gives that $\inf_{(x,u)}(P)$ is finite. As the minimizing sequence $(x_j,u_j)_j$ solves (D) and C is compact, it follows that the sequences $(\|\dot{x}_j\|_\infty)_j$ and $(\|x_j\|_\infty)_j$ are bounded by $2\bar{M}$ and M_C, respectively.

On the other hand, by hypothesis, we have that $(\|\dot{u}_j\|_2)_j$ is bounded, and by the t-uniform boundedness of $U(t)$ in (H4.2), the sequence $(\|u_j\|_\infty)_j$ is also bounded. Hence, 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 $(\tilde{x},\tilde{u})$ with $(\tilde{x}(t),\tilde{u}(t)) \in C \times U(t)$ for all $t \in [0,1]$, $(\dot{x}_j)_j$ converging to $\dot{\tilde{x}}$ in the weak*-topology of L^∞, and $(\dot{u}_j)_j$ converging weakly in L^2 to $\dot{\tilde{u}}$. As for all $j \in \mathbb{N}$, $(x_j(0),x_j(1)) \in C_0 \times C_1$, then (H4.1) and (H4.3) yield that $(\tilde{x}(0),\tilde{x}(1)) \in C_0 \times C_1$.

To prove that $(\tilde{x},\tilde{u})$ satisfies the sweeping process in (D), we use the equivalence invoking (2.3). Let $(\xi_j)_j$ be the $L^\infty([0,1],\mathbb{R}^+)$- sequence associated via (2.3)-(2.4) to the sequence $(x_j,u_j)_j$ admissible for (D). As (2.4) yields that $(\|\xi_j\|_\infty)_j$ is bounded by $\frac{\bar{M}}{2\eta}$, and $(\xi_j)_j$ admits a subsequence, we do not relabel that weakly* converges in $L^\infty([0,1];\mathbb{R}^+)$ to some $\tilde{\xi} \in L^\infty([0,1];\mathbb{R}^+)$. Using that (x_j,u_j,ξ_j) satisfies (2.3) together with the assumptions on f, and the properties of Φ and ψ, it easily follows upon taking the limit as $j \longrightarrow \infty$ in

$$x_j(t) = x_j(0) + \int_0^t [f_\Phi(s,x_j(s),u_j(s)) - \xi_j(s)\nabla\psi(x_j(s))]\,ds,\;\; \forall t \in [0,1],$$

that also $(\tilde{x},\tilde{u},\tilde{\xi})$ satisfies (2.3). We now show that $\tilde{\xi}$ is supported in $I^0(\tilde{x})$. Let $t \in I^{\text{-}}(\tilde{x})$ be fixed, that is, $\tilde{x}(t) \in \operatorname{int} C$. Since $(x_j)_j$ converges uniformly to $\tilde{x}$, then we can find $\tilde{\delta} > 0$ and $j_o \in \mathbb{N}$ such that for all $s \in (t-\tilde{\delta},t+\tilde{\delta}) \cap [0,1]$. For all $j \geq j_o$, we have $x_j(s) \in \operatorname{int} C$. Hence $\xi_j(s) = 0$, as ξ_j satisfies (2.4). Thus, $\xi_j(s) \longrightarrow 0$ for $s \in (t-\delta_o,t+\delta_o) \cap [0,1]$, and whence, $\tilde{\xi}(t) = 0$, proving that $\tilde{\xi}$ is supported in $I^0(x)$. Thus, $(\tilde{x},\tilde{u})$ solves (D) and $(\tilde{x},\tilde{u},\tilde{\xi})$ satisfies (2.4). Therefore, $(\tilde{x},\tilde{u})$ is admissible for (P). Owed to the lower semicontinuity of g and to $(\tilde{x},\tilde{u})$ being the uniform limit of the minimizing sequence $(x_j,u_j)_j$, the optimality of $(\tilde{x},\tilde{u})$ for (P) follows readily. □

The following theorem, Theorem 2.2, is the main result of this paper. It provides necessary optimality conditions in the form of weak Pontryagin principle for a C×W^{1,2}-local minimizer $(\bar{x},\bar{u})$ in (P). These optimality conditions extend those given in [8, Theorem 4.8], where $(\bar{x},\bar{u})$ is a W^{1,2}×W^{1,2}-local minimizer and the perturbation function is *autonomous*. Note that in the statement of Theorem 2.2, we use the following *nonstandard* notions of subdifferentials, which are *strictly smaller* than their counterparts in standard notions:

- $\partial_\ell \varphi$ and $\partial_\ell^2 \varphi$ defined on C, are the *extended Clarke generalized gradient* and the *extended Clarke generalized Hessian* of φ, respectively (see [14, Equation (9)-(10)]). Note that if $\partial_\ell \varphi(x)$ is a singleton, then we use the notation ∇_ℓ instead of ∂_ℓ.
- $\partial_\ell^{(x,u)} f(t,\cdot,\cdot)$ is the *extended Clarke generalized Jacobian* of $f(t,\cdot,\cdot)$ defined on the set $[C \cap \bar{B}_\delta(\bar{x}(t))] \times \left[(U(t)+\tilde{\rho}\bar{B}) \cap \bar{B}_\delta(\bar{u}(t))\right]$ (see [14, Equation (11)]).
- $\partial_\ell^2 \psi$ is the *Clarke generalized Hessian relative* to $\operatorname{int} C$ of ψ (see [14, Equation (12)]).
- $\partial_\ell^L g$ is the *limiting subdifferential* of g *relative* to $\operatorname{int}\left(\tilde{C}_0(\delta) \times \tilde{C}_1(\delta)\right)$ (see [14, Equation (8)]).

Theorem 2.2 (Necessary optimality conditions)**.** *Let* $(\bar{x},\bar{u})$ *be a* $C{\times}W^{1,2}$*-local minimizer for* (P) *with associated* $\delta > 0$ *at which* (H1)-(H5) *hold. Then, 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* $\mathcal{M}_{n\times n}([0,1])$*, and an* L^∞*-function* $\bar{\omega}(\cdot)$ *in* $\mathcal{M}_{n\times m}([0,1])$*, such that*

$$\left((\bar{\zeta}(t),\bar{\omega}(t)),\bar{\theta}(t),\bar{\vartheta}(t)\right) \in \partial_\ell^{(x,u)} f(t,\bar{x}(t),\bar{u}(t)) \times \partial_\ell^2 \varphi(\bar{x}(t)) \times \partial_\ell^2 \psi(\bar{x}(t)),\ t \in [0,1] \text{ a.e.},$$

and the following holds:

(i) **(The admissible equation)**
 (a) $\dot{\bar{x}}(t) = f(t,\bar{x}(t),\bar{u}(t)) - \nabla_\ell \varphi(\bar{x}(t)) - \bar{\xi}(t)\nabla\psi(\bar{x}(t)),\ \ t \in [0,1]$ a.e.,
 (b) $\psi(\bar{x}(t)) \leq 0,\ \ \forall t \in [0,1]$;

(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*

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

(iv) **(The complementary slackness conditions)**
 (a) $\bar{\xi}(t) = 0,\ \ \forall t \in I^{-}(\bar{x})$,
 (b) $\bar{\xi}(t)\langle\nabla\psi(\bar{x}(t),\bar{p}(t)\rangle = 0,\ \ \forall t \in [0,1]$ a.e.;

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

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

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

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

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

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

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

Remark 2.2. The following are simplified versions of the weak maximization condition of Theorem 2.2 for the special cases: (a) $U(t)$ is r-prox-regular for all $t \in [0,1]$, (b) $U(t) \cap \bar{B}_{\varepsilon_o}(\bar{u}(t))$ is convex for all $t \in [0,1]$, and (c) $U(t)$ is convex for all $t \in [0,1]$.

(a) We take $\varepsilon_o \longrightarrow \infty$ to obtain that, for $t \in [0,1]$ a.e.,

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

(b) We take $r \longrightarrow \infty$ to obtain that, for $t \in [0,1]$ a.e.,

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

(c) We take both $\varepsilon_o \longrightarrow \infty$ and $r \longrightarrow \infty$ to obtain that, for $t \in [0,1]$ a.e.,

$$\max\left\{\left\langle \bar{\omega}(t)^{\mathsf{T}}\bar{p}(t), u\right\rangle : u \in U(t)\right\} \text{ is attained at } \bar{u}(t).$$

3. Proof of the Main Results

The proof of Theorem 2.2 is presented in three steps.

Step 1: Approximating problems for (P). We introduce the following sequence of approximating problems:

(P_{γ_k}): Minimize

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

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

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

Lemma 3.1. *For k sufficiently large, problem (P_{γ_k}) has an optimal solution $(x_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ such that, for ξ_{γ_k} defined in (2.8), we have, along a subsequence, we do not relabel, that*

$$u_{\gamma_k} \xrightarrow[\mathcal{W}]{\text{strongly}} \bar{u}, \quad x_{\gamma_k} \xrightarrow[C([0,1];\mathbb{R}^n)]{\text{uniformly}} \bar{x}, \quad z_{\gamma_k} \xrightarrow[W^{1,1}([0,1];\mathbb{R}^+)]{\text{strongly}} 0, \;\; \textit{and} \;\; (\dot{x}_{\gamma_k}, \xi_{\gamma_k}) \xrightarrow[L^{\infty}([0,1];\mathbb{R}^n\times\mathbb{R}^+)]{\text{weakly*}} (\dot{\bar{x}}, \bar{\xi}).$$

In addition, we have:

(i) $x_{\gamma_k}(t) \in C(k) \subset \operatorname{int} C$, $\forall t \in [0,1]$.

(ii) $0 \le \xi_{\gamma_k}(t) \le \frac{2\bar{M}}{\eta}$, $\forall t \in [0,1]$.

(iii) $\|\dot{x}_{\gamma_k}(t)\| \le \bar{M} + \frac{2\bar{M}\bar{M}_{\psi}}{\eta}$, $\forall t \in [0,1]$ a.e.

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

Proof. By (2.9) and (2.11), let k be large enough so that $C_0(k) \subset \tilde{C}_0(\delta)$ and $C_1(k) \subset \tilde{C}_1(\delta)$. Since $\bar{x}_{\gamma_k} \longrightarrow \bar{x}$ uniformly, then, for k sufficiently large, $\bar{x}_{\gamma_k}(t) \in \bar{B}_{\delta}(\bar{x}(t))$, $\forall t \in [0,1]$. Thus, using that $\bar{c}_{\gamma_k} \in C_0(k)$, for all $k \in \mathbb{N}$, and $\bar{x}(1) \in C_1 \cap \bar{B}_{\delta_o}(\bar{x}(1))$, it follows that, for k large, $(\bar{x}_{\gamma_k}, \bar{z}_{\gamma_k} := 0, \bar{u})$ is an admissible triplet for (P_{γ_k}).

Now, fix k large enough so that $C_0(k)\times C_1(k)\subset \tilde{C}_0(\delta)\times\tilde{C}_1(\delta)$ and $(\bar{x}_{\gamma_k},0,\bar{u})$ is admissible for (P_{γ_k}). Using (H5) and the definition of $J(x,z,u)$, we obtain that $J(x,z,u)$ is bounded from below, and hence, $\inf(P_{\gamma_k})$ is *finite*. Let $(x^n_{\gamma_k},z^n_{\gamma_k},u^n_{\gamma_k})_n\in W^{1,2}([0,1];\mathbb{R}^n)\times W^{1,1}([0,1];\mathbb{R})\times\mathcal{W}$ be a minimizing sequence for (P_{γ_k}), that is, for each $n\in\mathbb{N}$, $(x^n_{\gamma_k},z^n_{\gamma_k},u^n_{\gamma_k})$ is admissible for (P_{γ_k}), and

$$\lim_{n\longrightarrow\infty} J(x^n_{\gamma_k},z^n_{\gamma_k},u^n_{\gamma_k})=\inf(P_{\gamma_k})<\infty. \tag{3.1}$$

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 [14, Lemma 4.1], 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 (H4.2), 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, since we have $z^n_{\gamma_k}(t)=\int_0^t\|\dot{u}^n_{\gamma_k}(\tau)-\dot{\bar{u}}(\tau)\|^2d\tau,\ \forall t\in[0,1]$. Hence,

$$\|\dot{u}^n_{\gamma_k}\|_2\le\|\dot{u}^n_{\gamma_k}-\dot{\bar{u}}\|_2+\|\dot{\bar{u}}\|_2=(z^n_{\gamma_k}(1))^{\frac{1}{2}}+\|\dot{\bar{u}}\|_2\le\sqrt{\delta}+\|\dot{\bar{u}}\|_2. \tag{3.2}$$

Therefore, by Arzelà-Ascoli theorem, along a subsequence (we do not relabel), $(x^n_{\gamma_k},u^n_{\gamma_k})_n$ converges *uniformly* to a pair $(x_{\gamma_k},u_{\gamma_k})$ and $(\dot{x}^n_{\gamma_k},\dot{u}^n_{\gamma_k})_n$ 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,

$$\|\dot{u}_{\gamma_k}-\dot{\bar{u}}\|_2^2\le\liminf_{n\longrightarrow\infty}\|\dot{u}^n_{\gamma_k}-\dot{\bar{u}}\|_2^2. \tag{3.3}$$

Define

$$z_{\gamma_k}(t):=\int_0^t\|\dot{u}_{\gamma_k}(\tau)-\dot{\bar{u}}(\tau)\|^2d\tau,\ \ \forall t\in[0,1]. \tag{3.4}$$

We claim that $(x_{\gamma_k},z_{\gamma_k},u_{\gamma_k})$ is optimal for (P_{γ_k}). First we prove its admissibility. Clearly we have

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

Moreover, since $\|\dot{u}^n_{\gamma_k}-\dot{\bar{u}}\|_2^2=z^n_{\gamma_k}(1)\in[-\delta,\delta]$, (3.3) yields that

$$z_{\gamma_k}(1)\in[-\delta,\delta]. \tag{3.5}$$

The inclusions $(x_{\gamma_k}(0),x_{\gamma_k}(1))\in C_0(k)\times C_1(k)$, 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 $t\in[0,1]$, follow directly from $C_0(k)$, $C_1(k)$, $\bar{B}_\delta(\bar{x}(t))$ and $U(t)\cap\bar{B}_\delta(\bar{u}(t))$ being 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 $(x_{\gamma_k},u_{\gamma_k})$. To prove that x_{γ_k} is the solution of (D_{γ_k}) corresponding to $(x_{\gamma_k}(0),u_{\gamma_k})$, we take the limit, as $n\longrightarrow\infty$, in this integral form of the admissible equation in (D_{γ_k}) for $(x^n_{\gamma_k},u^n_{\gamma_k})$,

$$x^n_{\gamma_k}(t)=x^n_{\gamma_k}(0)+\int_0^t\left[f_\Phi(s,x^n_{\gamma_k}(s),u^n_{\gamma_k}(s))-\gamma_k e^{\gamma_k\psi(x^n_{\gamma_k}(s))}\nabla\psi(x^n_{\gamma_k}(s))\right]ds,\ \ \forall\, t\in[0,1],$$

and use that $(x^n_{\gamma_k}(t),u^n_{\gamma_k}(t))\in[C\cap\bar{B}_\delta(\bar{x}(t))]\times[U(t)\cap\bar{B}_\delta(\bar{u}(t))]$, $(x^n_{\gamma_k},u^n_{\gamma_k})$ converges uniformly to $(x_{\gamma_k},u_{\gamma_k})$, (H1) and (H2.1) hold, and that Φ is C^1, and we conclude that $(x_{\gamma_k},u_{\gamma_k})$ satisfies the same equation, that is,

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

which terminated the proof of the admissibility of $(x_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ for (P_{γ_k}). For its optimality, from (3.1) and the uniform convergence of $(x^n_{\gamma_k}, u^n_{\gamma_k})$ to $(x_{\gamma_k}, u_{\gamma_k})$, it follows that

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

Therefore, for each such k, $(x_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ is an optimal solution to (P_{γ_k}).

For the convergence of $(x_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})_k$ when $k \longrightarrow \infty$, we first note that the sequence $(u_{\gamma_k})_k$ in $\mathcal{W}$ has uniformly bounded derivative in L^2. This follows from using (3.4) and (3.5) to obtain that (3.2) also holds when $(u^n_{\gamma_k}, z^n_{\gamma_k})$ is replaced by $(u_{\gamma_k}, z_{\gamma_k})$. Hence, using the arguments similar to those used above for $(u^n_{\gamma_k})_n$, we obtain the existence of $u \in \mathcal{W}$ such that, along a subsequence not relabled, u_{γ_k} converges uniformly to u, $\dot{u}_{\gamma_k}$ converges weakly in L^2 to $\dot{u}$, and

$$\|\dot{u}-\dot{\bar{u}}\|_2^2 \leq \liminf_{k\longrightarrow\infty}\|\dot{u}_{\gamma_k}-\dot{\bar{u}}\|_2^2. \tag{3.6}$$

On the other hand, by (2.10), we have $C_0(k) \subset C$, for k large. By [14, Theorem 4.1 & Lemma 4.2], the sequence $(x_{\gamma_k}, \xi_{\gamma_k})_k$, where ξ_{γ_k} is given via (2.8), admits a subsequence, not relabled, such that $(x_{\gamma_k})_k$ converges uniformly to some $x \in \mathrm{W}^{1,2}([0,1];\mathbb{R}^n)$ with images in C, $(\dot{x}_{\gamma_k}, \xi_{\gamma_k})_k$ converges weakly in L^2 to $(\dot{x}, \xi)$, and ξ is supported on $I^0(x)$. Furthermore, (x, u, ξ) satisfies (2.3)-(2.4) and x uniquely solves (D) for $(x(0), u)$. Now, as $x_{\gamma_k}(0) \in C_0(k)$, for k large, equation (2.9)(a), implies that $x(0) \in C_0 \cap \bar{B}_{\delta_o}(\bar{x}(0))$. Hence, (x, u, ξ) satisfies

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

Since $x_{\gamma_k}(1) \in C_1(k)$ for k large, equation (2.11)(a) implies that $x(1) \in C_1 \cap \bar{B}_{\delta_o}(\bar{x}(0))$. Furthermore, for all $t \in [0,1]$, from the facts that $x_{\gamma_k}(t) \in \bar{B}_\delta(\bar{x}(t))$ and $u_{\gamma_k}(t) \in U(t) \cap \bar{B}_\delta(\bar{u}(t))$, and that $(x_{\gamma_k}, u_{\gamma_k})$ converges uniformly to (x, u), we obtain $x(t) \in \bar{B}_\delta(\bar{x}(t))$ and $u(t) \in U(t) \cap \bar{B}_\delta(\bar{u}(t))$, for all $t \in [0,1]$. In addition, we have that

$$\|\dot{u}-\dot{\bar{u}}\|_2^2 \overset{(3.6)}{\leq} \liminf_{k\longrightarrow\infty}\|\dot{u}_{\gamma_k}-\dot{\bar{u}}\|_2^2 = \liminf_{k\longrightarrow\infty} z_{\gamma_k}(1) \overset{(3.5)}{\in} [-\delta, \delta],$$

proving that (x, u) is admissible for (P). Thus, by the local optimality of $(\bar{x}, \bar{u})$ for (P), we have that

$$g(\bar{x}(0), \bar{x}(1)) \leq g(x(0), x(1)). \tag{3.7}$$

Now by using the admissibility of $(\bar{x}_{\gamma_k}, 0, \bar{u})$ and the optimality of $(x_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k})$ for (P_{γ_k}), it follows that

$$J(x_{\gamma_k}, z_{\gamma_k}, u_{\gamma_k}) \leq g(\bar{x}_{\gamma_k}(0), \bar{x}_{\gamma_k}(1)) + \frac{1}{2}\|\bar{x}_{\gamma_k}(0)-\bar{x}(0)\|^2. \tag{3.8}$$

Hence, using the uniform convergence of $\bar{x}_{\gamma_k}$ to $\bar{x}$, (3.8), (3.7), the Lipschitz continuity of g, and the uniform convergence of x_{γ_k} to x, we obtain

$$\begin{aligned}
g(x(0),x(1)) &\leq \liminf_{k\longrightarrow\infty}\left(g(x_{\gamma_k}(0),x_{\gamma_k}(1))+\frac{1}{2}\left(\|u_{\gamma_k}(0)-\bar{u}(0)\|^2+\|\dot{u}_{\gamma_k}-\dot{\bar{u}}\|_2^2+\|x_{\gamma_k}(0)-\bar{x}(0)\|^2\right)\right)\\
&= \liminf_{k\longrightarrow\infty} J(x_{\gamma_k},z_{\gamma_k},u_{\gamma_k})\\
&\leq \liminf_{k\longrightarrow\infty}\left(g(\bar{x}_{\gamma_k}(0),\bar{x}_{\gamma_k}(1))+\frac{1}{2}\|\bar{x}_{\gamma_k}(0)-\bar{x}(0)\|^2\right)=g(\bar{x}(0),\bar{x}(1))\leq g(x(0),x(1)).
\end{aligned}$$

Thus

$$u(0)=\bar{u}(0) \ \text{ and } \ \liminf_{k\longrightarrow\infty}\left(\ \|\dot{u}_{\gamma_k}-\dot{\bar{u}}\|_2^2\right)=0, \ \text{ and} \tag{3.9}$$

$$x(0)=\bar{x}(0) \ \text{ and } \ g(\bar{x}(0),\bar{x}(1))=g(x(0),x(1)). \tag{3.10}$$

Equality (3.9) gives the existence of a subsequence of u_{γ_k}, without 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,

$$u_{\gamma_k} \xrightarrow[\mathcal{W}]{\text{strongly}} \bar{u},$$

which yields that $z_{\gamma_k}\longrightarrow 0$ in the strong topology of $W^{1,1}([0,1];\mathbb{R}^+)$. Moreover, as $u=\bar{u}$, then x and $\bar{x}$ solve the dynamic (D) with the same control $\bar{u}$ and initial condition; see (3.10). Hence, by the uniqueness of the solution of (D), we have $x=\bar{x}$. Using (2.4), we also obtain that $\xi=\bar{\xi}$. Therefore,

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

As $x_{\gamma_k}(0)\in C_0(k)$, we have that $x_{\gamma_k}(0)\in C(k)$ for k sufficiently large. Hence using [14, Theorem 5.1], we obtain that the conditions (i)-(iii) of the "In addition" part, hold true, which implies that a subsequence, we do not relabel, of $(\dot{x}_{\gamma_k},\xi_{\gamma_k})$ also converges weakly* in $L^\infty([0,1],\mathbb{R}^{n+1})$ to $(\dot{\bar{x}},\bar{\xi})$. Moreover, 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, $x_{\gamma_k}(0)\in\left[\left(C_0\cap\bar{B}_{\delta_o}(\bar{x}(0))\right)+\tilde{\rho}B\right]\cap(\operatorname{int}C)$. □

Step 2: Maximum principal for the approximation problems. We proceed and rewrite the approximating problems (P_{γ_k}) as a *standard* optimal control problem with state constraints in which the control u, which is in W^{1,2}, is considered as another state variable and its derivative,

$v := \dot{u}$ is the control. For $\bar{v} := \dot{\bar{u}}$, problem (P_{γ_k}) is reformulated in the following manner:

$$\begin{aligned}&(P_{\gamma_k})\colon \text{ Minimize}\\ &\quad J(x,z,u,v) := g(x(0),x(1)) + \tfrac{1}{2}\left(\|u(0)-\bar{u}(0)\|^2 + \|x(0)-\bar{x}(0)\|^2 + z(1)\right)\\ &\quad \text{over } (x,z,u) \in \mathrm{W}^{1,1}([0,1];\mathbb{R}^n)\times \mathrm{W}^{1,1}([0,1];\mathbb{R})\times \mathrm{W}^{1,1}([0,1];\mathbb{R}^m)\\ &\quad \text{and measurable functions } v\colon [0,1]\longrightarrow \mathbb{R}^m \text{ such that}\end{aligned}$$

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

In the following lemma, we apply to the above sequence of reformulated problems (P_{γ_k}), where k is as large as in Lemma 3.1, the nonsmooth Pontryagin maximum principle for standard optimal control problems with *implicit* state constraints (see, e.g., [21, Theorem 9.3.1] and [21, p.332]). For this purpose, (x,z,u) is the state function in (P_{γ_k}) and v is the control. Thus $(x_{\gamma_k},z_{\gamma_k},u_{\gamma_k})$ is the optimal state, where $(x_{\gamma_k},u_{\gamma_k})$ is obtained from Lemma 3.1 and $z_{\gamma_k}(t) := \int_0^t \|\dot{u}_{\gamma_k}(s)-\dot{\bar{u}}(s)\|^2\,ds$, and $v_{\gamma_k} = \dot{u}_{\gamma_k}$ is the optimal control.

Lemma 3.2. *For k large enough, there exist $\lambda_{\gamma_k}\geq 0$, $p_{\gamma_k}\in \mathrm{W}^{1,1}([0,1];\mathbb{R}^n)$, $q_{\gamma_k}\in\mathrm{W}^{1,1}([0,1];\mathbb{R}^m)$, $\Omega_{\gamma_k}\in NBV([0,1];\mathbb{R}^m)$, $\mu^o_{\gamma_k}\in \mathrm{C}^{\oplus}([0,1];\mathbb{R}^m)$, and a $\mu^o_{\gamma_k}$-integrable function $\beta_{\gamma_k}\colon [0,1]\longrightarrow\mathbb{R}^m$ 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*

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

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

$$\begin{aligned}\begin{pmatrix}\dot{p}_{\gamma_k}(t)\\ \dot{q}_{\gamma_k}(t)\end{pmatrix} \in\ & -\ \left(\partial^{(x,u)} f_{\Phi}(t,x_{\gamma_k}(t),u_{\gamma_k}(t))\right)^{\mathsf{T}} p_{\gamma_k}(t)\\ & +\ \begin{pmatrix}\gamma_k e^{\gamma_k\psi(x_{\gamma_k}(t))}\partial^2\psi(x_{\gamma_k}(t))p_{\gamma_k}(t)\\ 0\end{pmatrix}\\ & +\ \begin{pmatrix}\gamma_k^2 e^{\gamma_k\psi(x_{\gamma_k}(t))}\nabla\psi(x_{\gamma_k}(t))\langle\nabla\psi(x_{\gamma_k}(t)),p_{\gamma_k}(t)\rangle\\ 0\end{pmatrix};\end{aligned}\tag{3.11}$$

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

$$(p_{\gamma_k}(0),-p_{\gamma_k}(1))\in$$

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

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. $t \in [0,1]$,

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

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

$$\operatorname{supp}\{\mu^o_{\gamma_k}\}\subset\left\{t\in[0,1]:(t,u_{\gamma_k}(t))\in\operatorname{bdry}\operatorname{Gr}\left[U(t)\cap\bar{B}_\delta(\bar{u}(t))\right]\right\},\ \text{and}$$
$$\beta_{\gamma_k}(t)\in\partial_u^{>}d(u_{\gamma_k}(t),U(t)\cap\bar{B}_\delta(\bar{u}(t)))\quad \mu^o_{\gamma_k}\text{ a.e.},$$
$$\text{with } \partial_u^{>}d(u_{\gamma_k}(t),U(t)\cap\bar{B}_\delta(\bar{u}(t)))\subset\left[\left(\operatorname{conv}\bar{N}^L_{U(t)\cap\bar{B}_\delta(\bar{u}(t))}(u_{\gamma_k}(t))\right)\cap(\bar{B}\setminus\{0\})\right].$$

Proof. We intend to apply to the optimal solution, $\big((x_{\gamma_k},u_{\gamma_k},z_{\gamma_k}),v_{\gamma_k}\big)$, of the reformulated (P_{γ_k}), the *multiple state constraints* maximum principle [21, p.331] in which

$$(h_1(t,x,u),h_2(t,x,u)):=(d(x,\bar{B}_\delta(\bar{x}(t))),d(u,U(t)\cap\bar{B}_\delta(\bar{u}(t)))).$$

First, we 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 [18, Proposition 2.3], 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

$$0\in\partial_u^{>}d(u_{\gamma_{k_n}}(t_n),U(t_n)),\ \ \forall n\in\mathbb{N}. \tag{3.12}$$

The continuity of $\bar{u}$ and the uniform convergence of $u_{\gamma_{k_n}}$ to $\bar{u}$ yield that $(u_{\gamma_{k_n}}(t_n))_n$ converges to $\bar{u}(t_o)$. Hence, using the fact that the multifunction $(t,x)\mapsto\partial_u^{>}d(x,U(t))$ has closed values and a closed graph, we conclude from (3.12) that $0\in\partial_u^{>}d(\bar{u}(t_o),U(t_o))$. 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}.

One can easily prove the lower semicontinuity of the multifunctions $t\mapsto\bar{B}_\delta(\bar{x}(t))$ and $t\mapsto\left[U(t)\cap\bar{B}_\delta(\bar{u}(t))\right]$, and hence, the functions h_1 and h_2 that are Lipschitz in (x,u) are lower semicontinuous in (t,x,u). A simple argument by contradiction that uses the uniform convergence of u_{γ_k} to $\bar{u}$, the local property of the limiting normal cone, and the constraint qualification (CQ) being satisfied by $U(\cdot)$ at u_{γ_k}, yields that, for k sufficiently large, the multifunction $U(\cdot)\cap\bar{B}_\delta(\bar{u}(\cdot))$ satisfies the constraint qualification (CQ) at u_{γ_k}. Since $t\mapsto\bar{B}_\delta(\bar{x}(t))$ is lower semicontinuous and its values are closed, convex, and have nonempty interior, and x_{γ_k} converges uniformly to $\bar{x}$, then we deduce that, for k large enough, $\bar{B}_\delta(\bar{x}(\cdot))$ satisfies the constraint qualification (CQ) at x_{γ_k}. As for all $t\in[0,1]$, $u_{\gamma_k}(t)\in U(t)$, and by Theorem 3.1, $x_{\gamma_k}(t)\in\operatorname{int}C$, then (H1) yields that, for $t\in[0,1]$ a.e., we have $f(t,\cdot,\cdot)$ is M_ℓ-Lipschitz on the *neighborhood* of $(x_{\gamma_k}(t),u_{\gamma_k}(t))$. On the other hand, by Theorem 3.1, we have, for k sufficiently large,

$$\begin{aligned}(x_{\gamma_k}(0),x_{\gamma_k}(1)) &\in \left[\left(C_0\cap\bar{B}_{\delta_o}(\bar{x}(0))\right)+\tilde{\rho}B\right]\cap(\operatorname{int}C)\times\left[\left(C_1\cap\bar{B}_{\delta_o}(\bar{x}(1))\right)+\tilde{\rho}B\right]\cap(\operatorname{int}C)\\ &\subset \operatorname{int}(\tilde{C}_0(\delta)\times\tilde{C}_1(\delta)).\end{aligned}$$

Therefore, the data of (P_{γ_k}) satisfy all the hypotheses of the maximum principle stated in [21, p.331], which is deduced from [21, Theorem 9.3.1], by taking the scalar state constraint function therein to be $h(t,x,u)=\max\{h_1(t,x,u),h_2(t,x,u)\}$. When applying that maximum principle to (P_{γ_k}) at the optimal solution $\big((x_{\gamma_k},z_{\gamma_k},u_{\gamma_k}),v_{\gamma_k}\big)$, we notice that

- As in Step 2 of the proof of [13, Theorem 5.1], the adjoint variable p_{γ_k} corresponding to x_{γ_k} satisfies the adjoint equation that is linear in p_{γ_k} and: there exists $M_1>0$ such that

$$\|p_{\gamma_k}(t)\|\le M_1\|p_{\gamma_k}(1)\|\ \text{ for all } t\in[0,1],$$

which gives that $p_{\gamma_k} = 0$ if and only if $p_{\gamma_k}(1) = 0$. Therefore, in the nontriviality condition, $\|p_{\gamma_k}\|_\infty$ can be replaced by $\|p_{\gamma_k}(1)\|$.

- From Remark (a) on [21, page 330], the set $I(\bar{x})$ in the statement of [21, Theorem 9.3.1], and hence in that of [21, p.331] for h_1 and h_2, can be replaced by
$$\{t \in [0,1] : \partial_x^> h(t,\bar{x}(t)) \neq \emptyset\}. \tag{3.13}$$
Moreover, if $h(t,x) := d(x,F(t))$, where $F\colon [0,1] \rightrightarrows \mathbb{R}^m$ is a lower semicontinuous multifunction with closed and nonempty values, then we obtain from [18, Proposition 2.3(a) & Equation (2.15)] that
$$\{t \in [0,1] : \partial_x^> h(t,\bar{x}(t)) \neq \emptyset\} = \{t \in [0,1] : (t,\bar{x}(t)) \in \operatorname{bdry}\operatorname{Gr} F(t)\}. \tag{3.14}$$
- The measure corresponding to the state constraint "$x(t) \in \bar{B}_\delta(\bar{x}(t))$ for all $t \in [0,1]$" (or equivalently "$h_1(t,x(t),u(t)) \leq 0$") is *null*. This is due to the fact, from Theorem 3.1, that for k sufficiently large, $x_{\gamma_k}(t) \in B_\delta(\bar{x}(t))$ for all $t \in [0,1]$, which gives, for k sufficiently large, that this measure is supported in
$$\begin{aligned} &\{t \in [0,1] : \partial^>_{(x,u)} h_1(t,x_{\gamma_k}(t),u_{\gamma_k}(t)) \neq \emptyset\} \\ \overset{(3.14)}{=} \;& \{t \in [0,1] : (t,x_{\gamma_k}(t)) \in \operatorname{bdry}\operatorname{Gr} B_\delta(\bar{x}(t))\} \\ = \;& \left\{ t \in [0,1] : (t,x_{\gamma_k}(t)) \in \bigcup_{t\in[0,1]} \{t\} \times S_\delta(\bar{x}(t)) \right\} = \emptyset, \end{aligned}$$
where $S_\delta(\bar{x}(t)) := \{x \in \mathbb{R}^n : \|x - \bar{x}(t)\| = \delta\}$.
- The adjoint vector e_{γ_k} corresponding to the optimal state z_{γ_k} is the constant $-\frac{\lambda_{\gamma_k}}{2}$ (where λ_{γ_k} is the cost multiplier). Indeed, since v_{γ_k} converges strongly in L^2 to $\bar{v}$, we have, for k sufficiently large, that $z_{\gamma_k}(1) \in [0,\delta) \subset \operatorname{int}([-\delta,\delta])$. Adding to this that $\dot{e}_{\gamma_k}(t) = 0$ for $t \in [0,1]$ a.e., and using transversality condition, we obtain that
$$e_{\gamma_k}(t) = e_{\gamma_k}(1) \in -\left\{\frac{\lambda_{\gamma_k}}{2}\right\} - N^L_{[-\delta,\delta]}(z_{\gamma_k}(1)) = -\left\{\frac{\lambda_{\gamma_k}}{2}\right\}, \;\; \forall t \in [0,1].$$
Hence, for k sufficiently large, $e_{\gamma_k}(t) = -\frac{\lambda_{\gamma_k}}{2}$ for all $t \in [0,1]$.
- The BV-function associated to the state constraint "$u(t) \in U(t) \cap \bar{B}_\delta(\bar{u}(t))$" (or equivalently "$h_2(t,x(t),u(t)) \leq 0$") in the multiple state maximum principle takes the form $\int_{[0,t)} \beta_{\gamma_k}(t)\mu^o_{\gamma_k}(dt)$, for $t \in [0,1)$ and $\int_{[0,1]} \beta_{\gamma_k}(t)\mu^o_{\gamma_k}(dt)$, for $t = 1$, where, by (3.13) and (3.14),
$$\operatorname{supp}\{\mu^o_{\gamma_k}\} \subset \left\{t \in [0,1] : (t,u_{\gamma_k}(t)) \in \operatorname{bdry}\operatorname{Gr}\left[U(t) \cap \bar{B}_\delta(\bar{u}(t))\right]\right\},$$
$$\beta_{\gamma_k}(t) \in \partial_u^> d(u_{\gamma_k}(t), U(t) \cap \bar{B}_\delta(\bar{u}(t))) \quad \mu^o_{\gamma_k} \text{ a.e.},$$
and, by the (CQ) property and [21, Formula (9.17)],
$$\partial_u^> d(u_{\gamma_k}(t), U(t) \cap \bar{B}_\delta(\bar{u}(t))) \subset \left[\operatorname{conv} \bar{N}^L_{U(t)\cap\bar{B}_\delta(\bar{u}(t))}(u_{\gamma_k}(t)) \cap (\bar{B} \setminus \{0\})\right].$$
However, with a simple normalization procedure (see, e.g., the relevant part in the proof of [22, Theorem 3.4]), we can easily obtain a function $\Omega_{\gamma_k} \in NBV([0,1];\mathbb{R}^m)$ satisfying, together with β_{γ_k} and $\mu^o_{\gamma_k}$, $\Omega_{\gamma_k}(t) = \int_{[0,t]} \beta_{\gamma_k}(t)\mu^o_{\gamma_k}(dt)$, for $t \in (0,1]$, $\Omega_{\gamma_k}(0) = 0$, and the statement of the multiple state maximum principle remains valid with this function Ω_{γ_k}.

Therefore, we obtain the existence of $\lambda_{\gamma_k} \geq 0$, $p_{\gamma_k} \in W^{1,1}([0,1];\mathbb{R}^n)$, $q_{\gamma_k} \in W^{1,1}([0,1];\mathbb{R}^m)$, $\Omega_{\gamma_k} \in NBV([0,1];\mathbb{R}^m)$, $\mu^o_{\gamma_k} \in C^{\oplus}([0,1];\mathbb{R}^m)$, and a Borel measurable function $\beta_{\gamma_k}\colon [0,1] \longrightarrow \mathbb{R}^m$ such that $\Omega_{\gamma_k}(t) = \int_{[0,t]} \beta_{\gamma_k}(s)\mu^o_{\gamma_k}(ds)$ for all $t \in (0,1]$, $\Omega_{\gamma_k}(0) = 0$, and conditions (i)-(v) of this lemma hold. Note that in the adjoint equation (3.11), the values of $f(t,\cdot,\cdot)$ outside the set $[C \cap \bar{B}_\delta(\bar{x}(t))] \times \left[(U(t)+\tilde{\rho}\bar{B}) \cap \bar{B}_\delta(\bar{u}(t))\right]$ are not involved in the calculation of the subdifferential $\partial^{(x,u)} f_\Phi(t, x_{\gamma_k}(t), u_{\gamma_k}(t))$ since $(x_{\gamma_k}(t), u_{\gamma_k}(t))$ belongs to the *interior* of that set, for k sufficiently large and for all $t \in [0,1]$. □

Step 3: Finalizing the proof. By Lemma 3.2, since the conditions (i)-(v) of [8, Proposition 4.7] are valid, it is sufficient to follow the proof of [8, Theorem 4.8] to terminate the proof of Theorem 2.2.

4. Example

In this section, we present an example in which we illustrate how Theorem 2.2 can be used to find an optimal solution when the perturbation function is *nonautonomous*. We consider the problem (P) in which (see Figure 1):

- The *nonautonomous* perturbation mapping $f\colon [0,\frac{\pi}{2}] \times \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}^2$ is defined by
$$f(t,(x_1,x_2),u) = (t - x_1 - x_2 - u, -t + x_1 - x_2 + u).$$
- The function $\psi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ is defined by $\psi(x_1,x_2) := (x_1^2+x_2^2-1)(x_1^2+x_2^2-4)$. Hence, set C is the nonconvex and compact
$$C := \{(x_1,x_2) : (x_1^2+x_2^2-1)(x_1^2+x_2^2-4) \leq 0\}.$$
- The objective function $g\colon \mathbb{R}^4 \longrightarrow \mathbb{R} \cup \{\infty\}$ 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 $U(t) := [t,\pi]$ for all $t \in [0,\frac{\pi}{2}]$.
- The two sets C_0 and C_1 are defined by $C_0 := \{(1,0)\}$ and $C_1 := \{(0,x_2) : x_2 \geq 0\}$.

One can easily verify that hypotheses (H2)-(H4.4) are satisfied. Adding to this that $f(t,\cdot,\cdot)$ is globally Lipschitz on $\mathbb{R}^2 \times \mathbb{R}$, g is globally Lipschitz on $\mathbb{R}^2 \times C$, and $U(t)$ is convex with nonempty interior, we deduce that all the hypotheses of Theorem 2.2 are satisfied. Since g vanishes on the unit circle and is strictly positive elsewhere in C, we may seek for (P) an optimal solution $(\bar{x},\bar{u})$ such that, if possible, $\bar{x} := (\bar{x}_1,\bar{x}_2)$ belongs to the unit circle. Hence we have

$$\begin{cases} \bar{x}_1^2(t) + \bar{x}_2^2(t) = 1, \forall t \in [0,\frac{\pi}{2}]; \text{ and } \bar{x}_1(t)\dot{\bar{x}}_1(t) + \bar{x}_2(t)\dot{\bar{x}}_2(t) = 0, \ \forall t \in [0,\frac{\pi}{2}] \text{ a.e.}, \\ \bar{x}(0)^\intercal = (1,0) \text{ and } \bar{x}(\frac{\pi}{2})^\intercal = (0,1). \end{cases} \tag{4.1}$$

Applying Theorem 2.2 to this optimal solution $(\bar{x},\bar{u})$ and using Remark 2.2(c), we obtain 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 $\left[0,\frac{\pi}{2}\right]$, $\bar{\xi} \in L^\infty([0,\frac{\pi}{2}];\mathbb{R}^+)$, and $\lambda \geq 0$ such that, when incorporating equations (4.1) into (i)-(vi), these latter simplify to the following:

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

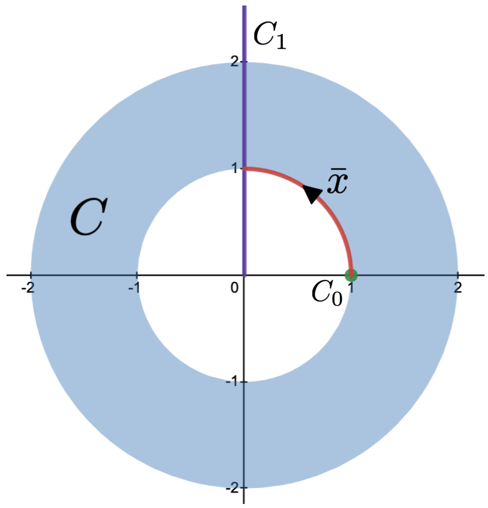

FIGURE 1. Example 4

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

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

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

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

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

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

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

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

From (4.1) combined with (b), we deduce that

$$\bar{\xi}(t) = \frac{1 + (\bar{u}(t) - t)(\bar{x}_1(t) - \bar{x}_2(t))}{6}, \ \forall t \in [0, \tfrac{\pi}{2}]. \tag{4.2}$$

On the other hand, the use of (d) and (4.1) in (c) yields that, for $t \in [0, \frac{\pi}{2}]$,

$$\begin{cases} d\bar{p}_1 = (\bar{p}_1(t) - \bar{p}_2(t) - 6\bar{\xi}(t)\bar{p}_1(t))dt - 6\bar{x}_1(t)d\bar{\nu}, \\ d\bar{p}_2 = (\bar{p}_1(t) + \bar{p}_2(t) - 6\bar{\xi}(t)\bar{p}_2(t))dt - 6\bar{x}_2(t)d\bar{\nu}. \end{cases} \tag{4.3}$$

Now in order to exploit (f), we temporarily assume that

$$\bar{p}_2(t) < \bar{p}_1(t) \text{ for } t \in [0, \tfrac{\pi}{2}] \text{ a.e.,} \tag{4.4}$$

hoping to be able to find $\bar{p}_1$ and $\bar{p}_2$ satisfying this condition. In this case, $\bar{u}(t) = t$ for all $t \in [0, \frac{\pi}{2}]$, which gives using (4.2) that $\bar{\xi}(t) = \frac{1}{6}$ for all $t \in \left[0, \frac{\pi}{2}\right]$. Using these values of $\bar{u}$ and

$\bar{\xi}$, and (4.1), we can solve for $(\bar{x}_1, \bar{x}_2)$ the two differential equations of (b) to obtain that

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

Employing (a), (d), (e), and (4.3), a simple calculation yields that

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

where $\delta_{\{a\}}$ denotes the unit measure concentrated on the point a. Note that, for all $t \in [0, \frac{\pi}{2}]$, we have $\bar{p}_2(t) < \bar{p}_1(t)$. Hence, the temporary assumption (4.4) is satisfied. Therefore, the above analysis, realized via Theorem 2.2, produces an admissible pair $(\bar{x}, \bar{u})$, where

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

which is optimal for (P).

REFERENCES

[1] J.J. Moreau, Rafle par un convexe variable. I, Travaux du Séminaire d'Analyse Convexe, vol. I, Exp. No. 15, U.É.R. de Math., Univ. Sci. Tech. Languedoc, Montpellier (1971), pp. 1-43.

[2] J.J. Moreau, Rafle par un convexe variable. II, Travaux du Séminaire d'Analyse Convexe, vol. II, Exp. No. 3, U.É.R. de Math., Univ. Sci. Tech. Languedoc, Montpellier (1972), pp. 1-36.

[3] J.J. Moreau, Evolution problem associated with a moving convex set in a Hilbert space, J. Differential Equations 26 (1977) 347-374.

[4] L. Adam, J. Outrata, On optimal control of a sweeping process coupled with an ordinary differential equation, Discrete Contin. Dyn. Syst. B 19 (2014), 2709-2738.

[5] T.H. Cao, B. Mordukhovich, Optimal control of a perturbed sweeping process via discrete approximations, Discrete Contin. Dyn. Syst. Ser. B 21 (2016), 3331-3358.

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

[7] T.H. Cao, B. Mordukhovich, Optimal control of a nonconvex perturbed sweeping process, J. Differential Equations 266 (2019), 1003-1050.

[8] C. Nour, V. Zeidan, A Control Space Ensuring the Strong Convergence of Continuous Approximation for a Controlled Sweeping Process, preprint (2023), arXiv:2301.00475.

[9] M.d.R. de Pinho, M.M.A. Ferreira, G.V. Smirnov, Optimal Control Involving Sweeping Processes, Set-Valued Var. Anal. 27 (2019), 523-548

[10] M.d.R. de Pinho, M.M.A. Ferreira, G.V. Smirnov, Correction to: Optimal Control Involving Sweeping Processes, Set-Valued Var. Anal. 27 (2019), 1025-1027.

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

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

[13] V. Zeidan, C. Nour, H. Saoud, A nonsmooth maximum principle for a controlled nonconvex sweeping process, J. Differential Equations 269 (2021), 9531-9582.

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

[15] C. Nour, V. Zeidan, Numerical solution for a controlled nonconvex sweeping process, IEEE Control Sys. Lett. 6 (2022), 1190-1195.

[16] D.G. Luenberger, Optimization by Vector Space Methods, John Wiley & Sons, Inc. 1969.

[17] F.H. Clarke, Optimization and Nonsmooth Analysis, John Wiley, New York, 1983.

[18] P.D. Loewen, R.T. Rockafellar, The adjoint arc in nonsmooth optimization, Trans. Amer. Math. Soc. 325 (1991), 39-72.

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

[20] J.-B. Hiriart-Urruty, Extension of Lipschitz functions, J. Math. Anal. Appl. 77 (1980) 539-554.

[21] R.B. Vinter, Optimal Control, Birkhäuser, Systems and Control: Foundations and Applications, Boston, 2000.

[22] D. Hoehener, Variational approach to second-order optimality conditions for control problems with pure state constraints, SIAM J. Control Optim. 50 (2012), 1139-1173.