NONSMOOTH OPTIMAL CONTROL FOR COUPLED SWEEPING PROCESSES WITH JOINT ENDPOINT CONSTRAINTS UNDER MINIMAL ASSUMPTIONS

By

Samara Chamoun

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Mathematics—Doctor of Philosophy

2025

ABSTRACT

A sweeping process typically refers to a dynamical system represented by a differential inclusion in which the set-valued map is the normal cone to a "nicely" moving closed set called the sweeping set. Although the sweeping process was originally developed for elastoplasticity applications, it has been widely recognized for its application in many other fields, including hysteresis, ferromagnetism, electric circuits, phase transitions, traffic equilibrium, economics, population motion in confined spaces, and other areas of applied sciences and operations research. Due to the nonstandard differential inclusions involved—with unbounded and discontinuous right-hand sides produced by the normal cone—classical results from the literature on differential inclusions are not applicable. In this dissertation, the study of nonsmooth optimal control problems (P) involving a controlled sweeping process with *three* main characteristics is launched. First, the sweeping sets are *nonsmooth*, *time-dependent*, and uniformly prox-regular. Second, the sweeping process is *coupled* with a controlled differential equation. Third, a *joint*-state endpoints constraint set S is present. This general model incorporates various significant controlled submodels, such as a class of second order sweeping processes, and coupled evolution variational inequalities. A full form of the *nonsmooth* Pontryagin maximum principle for *strong* local minimizers in (P) is derived for *bounded or unbounded moving* sweeping sets satisfying *local* constraint qualifications (CQ) *without* any additional restriction. The existence and uniqueness of a Lipschitz solution for the Cauchy problem of our dynamic is established and the existence of an optimal solution for (P) is obtained. Two of the novelties in achieving the first goal are (i) the construction of a problem over *truncated* sweeping sets and *truncated* joint endpoints constraint set preserving the same strong local minimizer of (P) while automatically satisfying (CQ), and (ii) the *complete redesign* of the exponential-penalty approximation technique for problems with moving sweeping sets that *do not require* any assumption on the sets, their corners, or on the gradients of their generators. The utility of the optimality conditions is illustrated with an example.

To my mom, dad, Antonio and Chase.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Vera Zeidan. Thank you for your generosity with time and resources, and for your willingness to spend countless hours working through mathematics with me. I have learned so much from your diligence, hard work, creativity, and thoroughness. You have helped me learn to think outside the box and to trust my mathematical instincts. One phrase of yours will always stay with me: "Everything has a solution"—in mathematics and in life. I hope I've made you proud as your first student.

I am also grateful to my committee members, both past and present—Dr. Jun Kitagawa, Dr. Baisheng Yan, Dr. Gábor Francsics, Dr. Ignacio Uriarte-Tuero and Dr. Matthew Hirn—for their support throughout this journey.

Thank you to Dr. Chadi Nour, who introduced me to the field of control theory and non-smooth analysis, made the connection with Dr. Vera, and encouraged me to apply to the U.S.—even though it wasn't the standard path. I owe much of where I am today to your guidance and encouragement.

Thank you to Dr. Boris Mordukhovich for his encouraging words and for writing letters of recommendation on my behalf.

I am also deeply grateful to all my math teachers in middle school and high school, as well as to the professors at the Lebanese University, the Lebanese American University, and Michigan State University, who ignited my passion for mathematics and shaped my journey as a mathematical scholar. Your guidance has left a lasting impact. I would especially like to thank Dr. Charbel Klaiany, Dr. Rony Touma, Dr. Nader El Khatib, Dr. Carole El Bacha, and Dr. Leila Issa for their support and inspiration along the way.

Graduate school was not only a time of learning and doing mathematics, but also a time when I truly discovered my passion for teaching. I grew into the role of an educator with the help of incredible mentors and leaders who supported me, challenged me, and believed in me.

Andy Krause and Tsveta Sendova, my experience in graduate school would not have been the same without your professional and personal support. What you are doing in the mathematics department is nothing short of amazing. Through your training and the personalized care you offer, you are cultivating a cohort of outstanding and competitive educators. Thank you for all the experiences and opportunities you have given me, for the advice, the resources, the conversations, and for always rooting for me.

Rachael Lund, thank you for being an incredible mentor and leader. I truly enjoyed working with you on the support program and the FAST Fellowship—these were among the most rewarding experiences of my graduate journey and deeply shaped my teaching philosophy. Your kind and compassionate mentorship is something I admire and hope to emulate.

Stefanie Baier, thank you for your fierce and unwavering advocacy for graduate students, and for supporting both my professional and personal aspirations. I am so grateful for the opportunity to be part of the GREAT Office—a space that connected me with an incredible community of educators, both inside and outside MSU. Being part of this office—and the engaging conversations, fun outings, and meaningful friendships it offered—has been one of the highlights of my graduate journey. Thank you to the AMAZING advisory group—Hima Rawal, Gloria Ashaolu, Sewwandi Abeywardana, Arya Gupta, Tianyi (Titi) Kou-Herrema, Seth Hunt, Saviour Kitcher, and Ellen Searle—for your support and friendships. And thank you to the GREAT Fellows—Qi Huang, Amie Musselman, Tara Mesyn and Tianyi (Titi) Kou-Herrema—for sharing this journey with me.

Thank you to the FAST Fellowship community for giving me the opportunity to study the impact of the support program on students' attitudes toward mathematics. I truly enjoyed this work. I am grateful to Dr. Rique Campa, Sevan Chanakian, and the FAST committee and cohort for their support, guidance, and leadership throughout the process.

Thank you to Dr. Jenny Green for being an incredible mentor—always encouraging and supportive of our ideas. It was in your class that I first engaged with the concept of compassionate teaching, an idea that has since taken me so many places. I will continue to share my ideas with you because I know they'll be met with support, curiosity, and care.

Thank you to Dr. Marybeth Heeder for being a powerful example of a kind, compassionate mentor, leader, and researcher in academia. I have deeply appreciated our conversations, your insights, and the resources you've so generously shared with me.

Thank you to Dr. Filomena Nunes for teaching the Women in STEM course, which provided us with the tools and language to advocate for ourselves in STEM fields. Your course offered not only valuable resources, but also a space of encouragement and empowerment.

Thank you to Dr. Amy Martin for connecting me with the Campus Student Success Group. I truly enjoyed engaging with this community and appreciated the opportunity to share resources, ideas, and conversations.

Most importantly, thank you to my students—you have given me some of the most joyful moments of graduate school. Your curiosity, hard work, engagement, kind words, smiles, and vulnerability will always be an inspiration to me. I vividly remember many days when I walked into the classroom feeling tired, and left filled with energy and gratitude. You have my deepest thanks.

To all medical professionals, physicians, mental, spiritual, and physical health providers, and emergency service workers—thank you. Your care, presence, and dedication have made a real difference. I would not be in the same physical and mental headspace without your help. A special thank you to Corey Decker, Elizabeth Malsheske, Dr. Jabir Sawaya, and Brittany Jurek.

Thank you to Dana Carley for always being in my corner. You have stood by me through the toughest and happiest moments of my life. You've seen my vulnerability and insecurities, held my hand through hard times, and helped me come out the other side a better person. I am forever grateful that our paths crossed.

Thank you to Fr. Mike, who helped me reconnect with my spirituality in the U.S. and encouraged me to lean on both God and the arts in moments of uncertainty. Because of you, I started painting.

Thank you to all the unseen helpers—the MSU staff whose work so often goes unrecognized but is deeply appreciated. To the Mathematics Department staff, thank you for your constant support. A special thank you to Taylor Alvarado for always being on top of things and for helping us navigate department requirements on our way to graduation.

Thank you to the Office for International Students and Scholars (OISS), and especially to Ismail Adawe, for your support with the OPT process and other visa-related matters. You've been the best OISS advisor I could have asked for—thank you for your guidance and care.

Thank you to Deanne Arking, who supported me during my time as Vice President for International Affairs with the Council of Graduate Students.

I couldn't have made it through this PhD without the love and support of my friends and family.

To the friends I met in the U.S., you made living away from home not only more bearable, but also more joyful and fun. Thank you for the coffee dates, study sessions, sleepovers, for helping me through tough times, and for celebrating the big moments with me. A heartfelt thank you to Marena Haidar, Charif Yassine, Elliott Haddad, Zuebida (Zee), Jamie Kimble, Nicole Hayes, Valeri Jean-Pierre, Jinting Liang, Shikha Bhutani, Astrid Olave-Herrera, Eloy Moreno Nadales, Rob McConkey, Chamila Malagoda Gamage, Lubashan Pathirana Karunarathna, Hitesh Gakhar, Reshma Menon, Brady Tyburski, Sofia Abreu, Saul Barbosa, Anthony Dickson, Hope Lewis, Sewwandi Abeywardana, Gloria Ashaolu, Arya Gupta, Sevan Chanakian, Ken Bundy, and Hannah Jeffery.

Thank you to Ayesha Farheen for the sleepovers, for being there during the hard times, for opening your home to me, and for being such an incredibly kind and thoughtful human being.

Thank you to Ayesha Bundy for your friendship, for making me feel at home in your space, for the study sessions, the delicious meals, the sense of community, and your constant check-ins and support.

Thank you to Chloe Lewis for being one of the most generous people I know — for sharing resources, helping me navigate the job market, and for all our rich and inspiring teaching conversations.

Thank you to Danika Van Niel for the fun trips and hangouts, for visiting Chicago together for the first time, for traveling to Lebanon and meeting my family, and for being such a supportive and encouraging friend.

Thank you to Hima Rawal for being a trusted friend. I've deeply appreciated the way we've shared our vulnerabilities with one another, without judgment and with so much compassion.

Thank you to Rahaf Ahmad for being a safe space for voicing my anxiety. You were one of the first close friends I made in the U.S., you always show up, and I know I can always count on you.

Thank you to my childhood friend, Perla Alam. We will always be there for each other. Thank you for your constant presence—always being there for me, rooting for me, and walking through life with me. You have been a steady and loving part of my world since childhood.

Thank you to Monica El Khoury and Marina El Khoury for 13 years of friendship. I'm lucky to have you—celebrating my happiness and feeling my sadness as if it were your own. My life wouldn't have been the same without you.

Thank you to my extended family: my grandma, uncles, aunts, cousins, Chase's family, and all the other family members who have supported me along the way. You have given me so much love throughout the years, and I feel incredibly lucky to have so many people in my corner.

Finally, to my family: Mom, Dad, Antonio, and Chase. Simply put, I wouldn't be where I am without you. I have the best support system in you. Mom, thank you for your unconditional love—for asking for nothing but for me to be okay. I have been able to achieve everything I have because of you. You never asked me for anything in return—just to be well—and that kind of love is everything. Dad, thank you for talking things through with me, for being excited about my successes and wanting to celebrate them loudly, and for being there for me during the lows. To my brother Antonio, thank you for being there through the scariest and most anxious moments of my life. You calm me down in a way no one else can. To my fiancé Chase, you lived it all with me—the scary, the bad, the happy, the sad, the anger, the joy, the boredom, and the love. I am the luckiest person to have such a smart, sweet, kind, and supportive partner walking through life by my side. Meeting you as a graduate student will always be the most fortunate part of my PhD. Thank you to the four of you for being my safe place and my home.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

$\|\cdot\|$	Euclidean norm
$\langle\cdot,\cdot\rangle$	Usual inner product
$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{R}_+$	Set of positive real numbers
$B_a(x)$	Open ball centered at x and of radius a
$\bar{B}_a(x)$	Closed ball centered at x and of radius a
B	Open unit ball
$\bar{B}$	Closed unit ball
$\mathscr{M}_{m\times n}[a,b]$	Set of $m \times n$-matrix functions on $[a,b]$
$\mathrm{I}_{r\times r}$	The identity matrix in $\mathscr{M}_{r\times r}$
$\operatorname{int} S$	Interior of a set $S \subset \mathbb{R}^n$
$\operatorname{bdry} S$	Boundary of a set $S \subset \mathbb{R}^n$
$\operatorname{cl}\ S$	Closure of a set $S \subset \mathbb{R}^n$
$\operatorname{conv} S$	Convex hull of a set $S \subset \mathbb{R}^n$
S^c	Complement of a set $S \subset \mathbb{R}^n$
$\operatorname{Gr}\ S(\cdot)$	Graph of the set-valued map $S(\cdot)$
$\sigma(\cdot, S)$	The support of a set $S \subset \mathbb{R}^n$
$d_S(x)$	The distance from x to a set $S \subset \mathbb{R}^n$
$\operatorname{proj}_S(x)$	The closest point or projection of x onto $S \subset \mathbb{R}^n$
$N_S^P(s)$	The proximal normal cone to S at s
$N_S^L(s)$	The Limiting normal cone to S at s
$N_S(s)$	The Clarke normal cone to S at s
$\operatorname{dom} f$	Effective domain of f
$\operatorname{epi} f$	Epigraph of f

Gr f	Graph of f
∇f	Gradient of f
$\partial^P f(x)$	Proximal subdifferential of f at x
$\partial^L f(x)$	Limiting subdifferential of f at x
$\partial f(x)$	Clarke generalized gradient of $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ at x
$\partial^2 f(x)$	Clarke generalized Hessian of f at x
$\partial g(x)$	Clarke generalized Jacobian of $g : \mathbb{R}^n \to \mathbb{R}^n$ at x
$\mathcal{C}([a,b];\mathbb{R}^n)$	Space of continuous functions from $[a,b]$ to $\mathbb{R}^n$
$AC([a,b];\mathbb{R}^n)$	Space of absolutely continuous functions from $[a,b]$ to $\mathbb{R}^n$
$BV([a,b];\mathbb{R}^n)$	Space of bounded variation functions from $[a,b]$ to $\mathbb{R}^n$
$V_a^b(f)$	Total variation of f
$L^p([a,b];\mathbb{R}^n)$	Lebesgue space of p-integrable functions from $[a,b]$ to $\mathbb{R}^n$
$W^{1,p}([a,b];\mathbb{R}^n)$	Sobolev space of continuous functions f whose derivative $\dot{f} \in L^p$
$\mathfrak{M}(S)$	The set of Radon measures on S
$\mathfrak{M}_+(S)$	The set of positive Radon measures on S
$\mathfrak{M}_+^1(S)$	The set of probability measures on S
$\mathcal{C}^*([a,b];\mathbb{R}^n)$	The dual space of $\mathcal{C}([a,b];\mathbb{R}^n)$ equipped with the supremum norm
$\|\cdot\|_{T.V}$	The induced norm on $\mathcal{C}^*([a,b];\mathbb{R}^n)$
$\mathcal{C}^{\oplus}(a,b)$	The set of elements in $\mathcal{C}^*([a,b];\mathbb{R})$ taking non-negative values on nonnegative-valued functions in $\mathcal{C}([a,b];\mathbb{R})$

CHAPTER 1

INTRODUCTION

1.1 Intersectionality of different fields

The research in this thesis centers around different fields of mathematics: control theory, optimization, dynamical systems, nonsmooth analysis, set-valued analysis, and functional analysis.

1.1.1 Control theory

If you are reading this thesis as a non-mathematician or as a mathematician from a different field, this section provides everything you need to know about control theory, including its foundational concepts and its applications in everyday life. A friend shared a fascinating map of mathematics with me (see Figure 1.1), and I invite you to take a moment to explore it and see where control theory fits within the mathematical landscape.

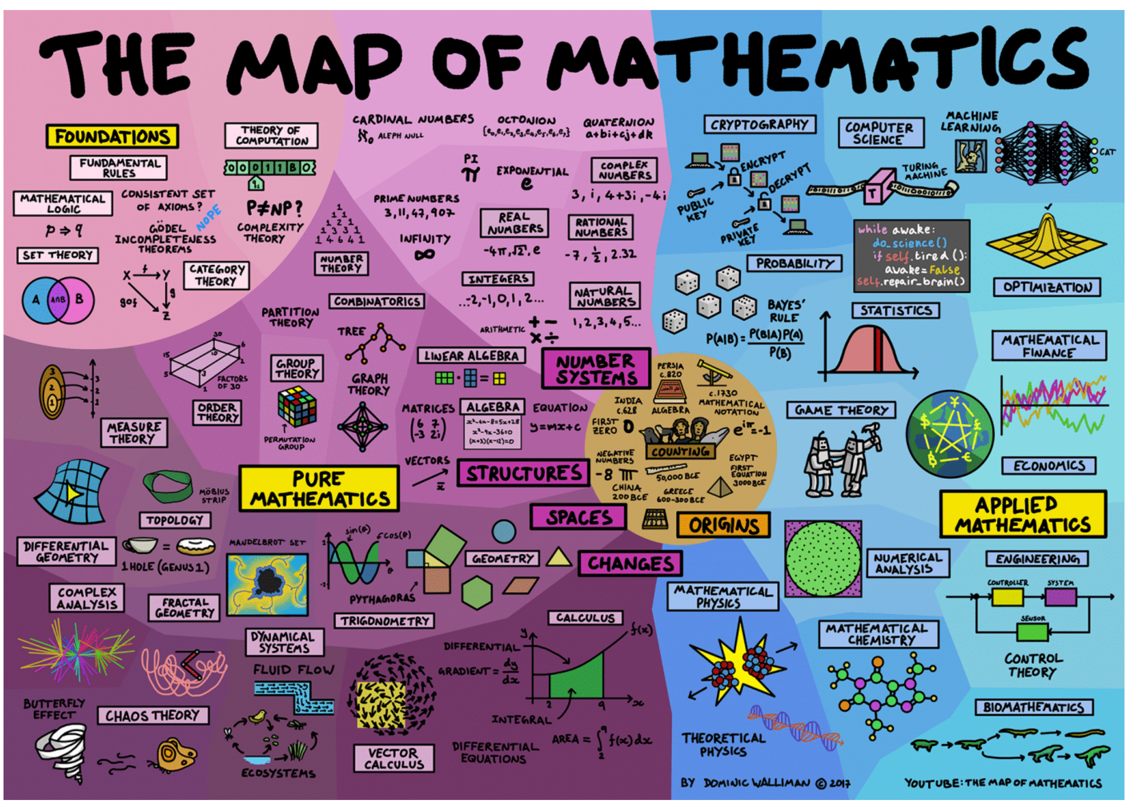

Figure 1.1 Placement of key areas mentioned above on the map of mathematics

Control theory is one of the most interdisciplinary areas of research, serving as a critical intersection between mathematics and engineering. It is a subfield of mathematics that focuses on using feedback to influence the behavior of dynamical systems—whether physical, biological, or otherwise—to achieve specific goals. Before emerging as a distinct field in the late 1950s and early 1960s, control theory was deeply connected to other areas of mathematics, such as calculus of variations and differential equations. Early research often adapted classical theories and techniques from these fields to address control problems, laying the groundwork for the development of modern control theory. I discovered a map of control theory itself, which I invite you to explore as it highlights the different structures and connections within this field (See Figure 1.2).

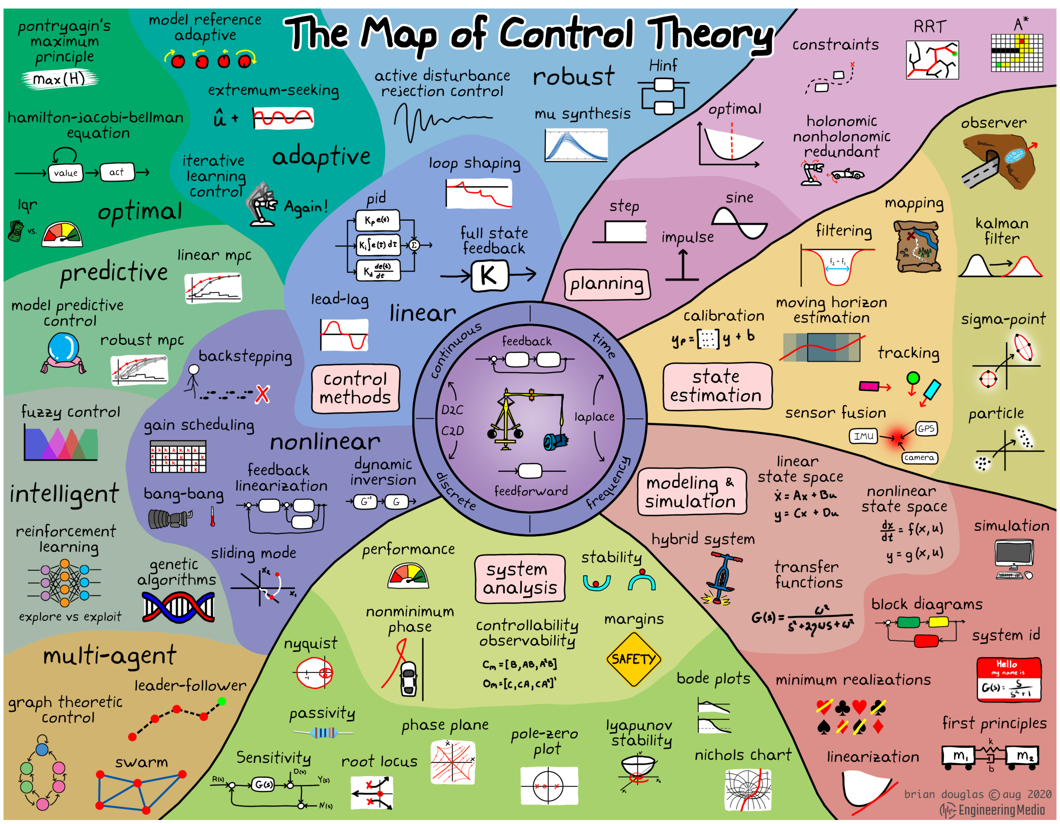

Figure 1.2 Map of control theory

This field can be broadly divided into two branches: linear control systems and nonlinear control systems. While linear control systems are foundational and often easier to analyze, all real-world control systems exhibit nonlinear behavior, making nonlinear control systems more applicable to practical scenarios. Control system theory can contribute to[1]:

- Developing mathematical models to describe system dynamics.
- Simulating and predicting system behavior under various scenarios.
- Analyzing and understanding dynamic interactions within complex systems.
- Filtering and rejecting noise to enhance signal clarity and system accuracy.
- Selecting and designing appropriate hardware to implement control strategies.
- Testing and validating system performance in unpredictable environments.
- Gaining foundational insights into system behavior and functionality.

A controller (see Figure 1.3) operates through different types of feedback loops. As discussed in this article[2] on the difference between **open-loop and closed-loop systems**, an open-loop controller, also known as feedforward, does not use any information about the current state or output of the system to influence its control actions. In contrast, a closed-loop controller, also known as a feedback controller, incorporates feedback into its decision-making process. Closed-loop controllers can be further categorized based on the type of feedback they use: system feedback controllers, which rely on feedback from the internal state of the system, and output feedback controllers, which utilize feedback from the system's output.

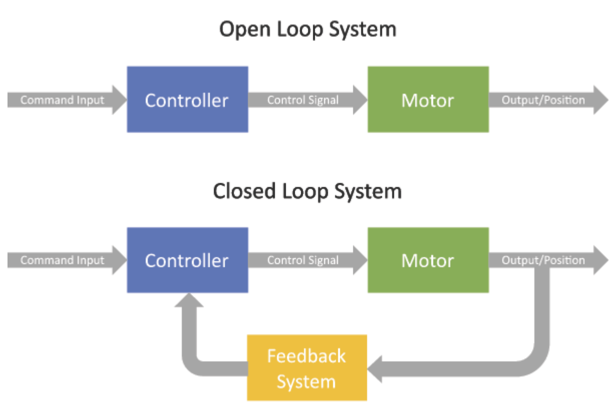

Figure 1.3 Open loop system versus closed loop system [2]

[1]The following was adapted from educational materials presented by Brian Douglas on his YouTube channel.

[2]https://www.ntchip.com/electronics-news/difference-between-open-loop-and-closed-loop.

Open-loop control systems are typically used for simple processes with well-defined input-output relationships. For instance, consider a dishwasher. The objective of the dishwasher (the plant) is to clean dishes (the output). Once the user sets the wash time (the input), the dishwasher will operate for the specified duration, regardless of the actual cleanliness of the dishes. If the dishes were already clean at the start, the dishwasher would still run for the full prescribed time. Similarly, a dryer operates on the same principle. The user sets the drying time (the input), which determines how long the dryer runs. This duration is fixed and unaffected by whether the clothes are already dry.

On the other hand, a closed-loop control system dynamically adjusts its operation based on feedback from its output. The system continuously monitors the output, compares it to the desired outcome, and adjusts the input accordingly to minimize any discrepancies. For example, consider a dryer equipped with a sensor that measures the dryness of the clothes. This sensor provides feedback that is compared to a reference signal representing the desired dryness level (set by the manufacturer or the user). The difference between the measured and desired levels generates an error term, which is sent to a controller. The controller uses this feedback to determine when to shut off the dryer, ensuring the clothes are dried to the desired level (see Figure 1.4).

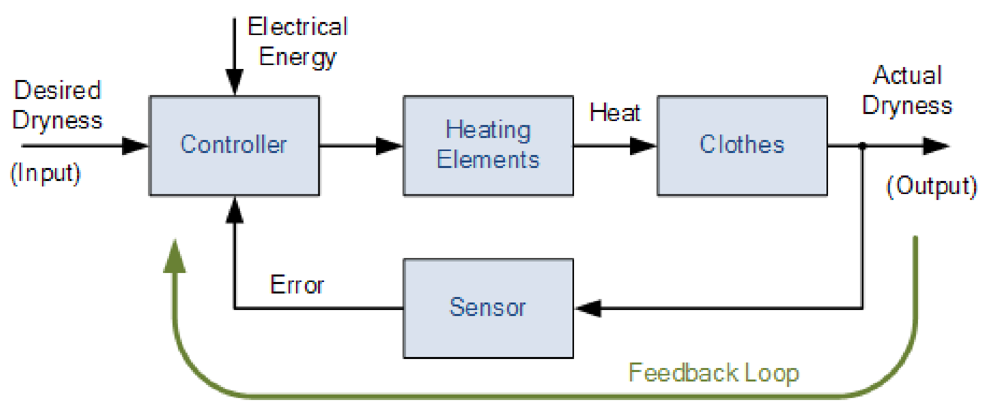

Figure 1.4 Closed loop system in a dryer [2]

Control theory finds extensive **applications across diverse fields**, including biology (e.g. optimal vaccination strategies) , physics (e.g. spacecraft control), engineering (e.g. robotics), economics (economic growth models), medicine (e.g. drug target identification in cancer research), and finance (e.g. risk management).

It is important to note that a mathematical solution to a control problem may not always exist. In the late 1950s, rigorous conditions for existence were established, with controllability being a key criterion, ensuring that some form of control is possible. **Optimal control** focuses on finding a control law for a given system that satisfies a specified optimality criterion. It involves a cost functional, which depends on the state and control variables. An optimal control solution consists of differential equations that describe the evolution of the control variables to minimize the cost function. Such solutions can be derived using Pontryagin's Maximum Principle or by solving the Hamilton-Jacobi-Bellman equation.

1.1.2 Nonsmooth analysis

Nonsmooth analysis, which can be considered a subdomain of nonlinear analysis, refers to differential analysis without the differentiability. It concerns the local description of nondifferentiable functions and sets lacking smooth boundaries, in terms of generalizations of classical concepts of derivatives, normals and tangents. Although this subject has traditional roots, it is only over the last few decades that it has developed rapidly. The reason behind this progress is the acknowledgment of the importance of nondifferential setting, its universal presence and its direct relation with some unusual behaviors such as chaos and catastrophes. It can be viewed, within differential (functional) analysis, as a topic in itself. However, it has also gained a major part in several applications such as optimization and control theory. Among F. Clarke and R. T. Rockafellar, many more such as J. Borwein, A. D. Ioffe, B.Mordukhovich and R. B. Vinter have contributed in its development. The need for nonsmooth analysis in control theory is connected to finding proofs of necessary conditions for optimal control, in particular with the use of Pontryagin Maximum Principle. In general, nonsmooth analysis intervenes when considering nonlinear problems (studying the sensitivity of the problems, deriving necessary conditions or applying sufficient conditions).

1.2 Sweeping process

As mentioned above, optimal control theory involves minimizing an objective function subject to a given control system. The specific system I focus on in this thesis is known as the sweeping process. My work centers on studying the dynamics of the sweeping process and addressing optimal control problems governed by such systems. For readers unfamiliar with the sweeping process, this section provides a brief introduction to its background.

1.2.1 Definition, interpretation, and applications

J.J. Moreau introduced the sweeping process as being a differential inclusion in which the set-valued map is the normal cone to a nicely moving non-empty closed set $C(t)$, called the sweeping set (see [49, 50, 51]). The simplest form of the sweeping process is given by

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

When the set $C(t)$ is convex, $N_{C(t)}$ corresponds to the normal cone of convex analysis. However, when $C(t)$ is non-convex, then it is taken to be uniformly prox-regular, in which case $N_{C(t)}$ is the Clarke normal cone. When we add a perturbation or external force f to (1.1), we call the dynamic a perturbed sweeping process, and when f depends on a control u, we call it a perturbed controlled sweeping process.

To understand what the word "sweeping" means, we can think of a large ring moving while containing a small ball inside. The ring starts moving at $t = 0$, and the movement of the ball depends on how the ring interacts with it. If the ball is not hit by the ring, it remains stationary. However, if the ring hits the ball, the ball is "swept" towards the inside of the ring. The main idea here is that the velocity of the ball must point inwards so that the ball does not escape the ring's bounds (See Figure 1.5).

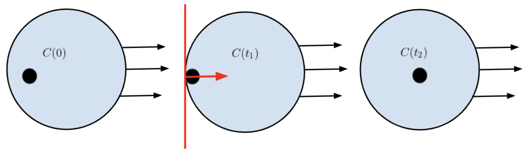

Figure 1.5 Sweeping process interpretation

Sweeping processes have various applications in different fields including elastoplasticity, hysteresis, ferromagnetism, electric circuits, phase transitions, and traffic equilibrium (see, for example, [1, 4, 7, 45, 67]). In the past decade, interest in sweeping processes has grown due to their significant role in emerging applications such as mobile robot models [27], and pedestrian traffic flow models [27]. In these contexts, the primary goal is to efficiently control the state of events by optimizing a specific objective function over the controlled sweeping process.

One of the most fascinating applications of sweeping process is the crowd motion models for emergency evacuation [14, 10]. In case of an emergency evacuation, we want to find the most effective way to leave the room. While we would prefer to move at our "desired" velocity, we need to take into account the direct contact between each other, as well as our contact with different objects and obstacles present in the room. Thus, our "actual" velocity—the closest achievable velocity to our desired one while accounting for direct contact with others—is determined by a sweeping process dynamic.

Figure 1.6 Emergency evacuation

1.2.2 Theoretical results

Due to the unboundedness and discontinuity of the normal cone in in (1.1), standard results involving differential inclusions cannot be used for sweeping processes. Extensive literature exists on the question of existence and *uniqueness* of an *absolutely continuous* or *Lipschitz* solution for the Cauchy problem associated with different forms of the following perturbed controlled sweeping process

$$\dot{x}(t) \in f(t, x(t), u(t)) - N_{C(t)}(x(t)) \text{ a.e. } t \in [0, T], \; x(0) = x_0 \in C(0), \tag{1.2}$$

in which the constraint $x(t) \in C(t)$ is implicit. Initially, such results commonly required the *absolute or Lipschitz continuity* of the set-valued map $C(\cdot)$ (see, e.g., [36]). However, motivated by the need to consider set-valued map $C(\cdot)$ for which these conditions are too strong (see [65]), similar results are derived by merely assuming the same conditions on the *ρ-truncated* set-valued map $C(\cdot) \cap \rho\bar{B}$ (see e.g., [52, 64, 65]). In [41], when $C(t)$ is polyhedral, a constraint qualification is shown to be*sufficient* for those conditions to be satisfied on the *ρ-truncated* polyhedral sets.

Numerous efforts have been made to derive existence theory for *optimal* solutions and/or *optimality* conditions in terms of *Euler-Lagrange* equation or *Pontryagin-type* maximum principle for optimal control problems driven by variants of (1.2). The main approach used to solve different versions of such an optimal control problem is the method of approximation, either *discrete* (see, e.g., [12, 13, 10, 11, 23, 25, 26, 28]), or *continuous* (see, [6, 30, 33, 34, 55, 57, 58, 70]). Our focus in this paper is on the latter, and more specifically, on the *exponential penalty-type*.

1.2.2.1 Selected results for constant sweeping set C.

Work of dePinho et al. in [30, 31]

The exponential penalization technique was first used in [30, 31] to derive existence of solution of (1.2), existence of optimal solution and Pontryagin-type maximum principle for *global* minimizers of a Mayer problem over (1.2), in which:

- f is *smooth* and *convex*,
- C is a *constant compact* set defined as the zero-sublevel set of a C^2-convex function ψ satisfying a constraint qualification on $\mathbb{R}^n$,
- with initial state-constraint set $C_0 \subset C$ and *free* final state.

The *novelty* of this technique resides in approximating $N_C(\cdot)$ by the exponential penalty term $\gamma_k e^{\gamma_k \psi(\cdot)} \nabla \psi(\cdot)$ such that the so-obtained approximating dynamic is a standard control system without state constraints for which C is shown to be *invariant*:

$$\dot{x}(t) = f(t, x(t), u(t)) - \gamma_k e^{\gamma_k \psi(x(t))} \nabla \psi(x(t)) \text{ a.e.}, \quad x(0) = x_0 \in C. \tag{1.3}$$

The absence in (1.3) of the explicit state constraint, $x(t) \in C$, that is implicitly present in (1.2), has also been shown to be instrumental in constructing numerical algorithms for controlled sweeping processes (see [32, 56, 59]).

In summary, the exponential penalization technique works as follows: rather than deriving necessary conditions for optimal solutions of a problem (P), governed by (1.2), directly, we approximate (P) with a sequence of standard optimal control problems (P_k) governed by (1.3). Using existing results, we determine necessary conditions for (P_k), and by analyzing the limit as $k \to \infty$, we then obtain necessary conditions for (P) (see Figure 1.7).

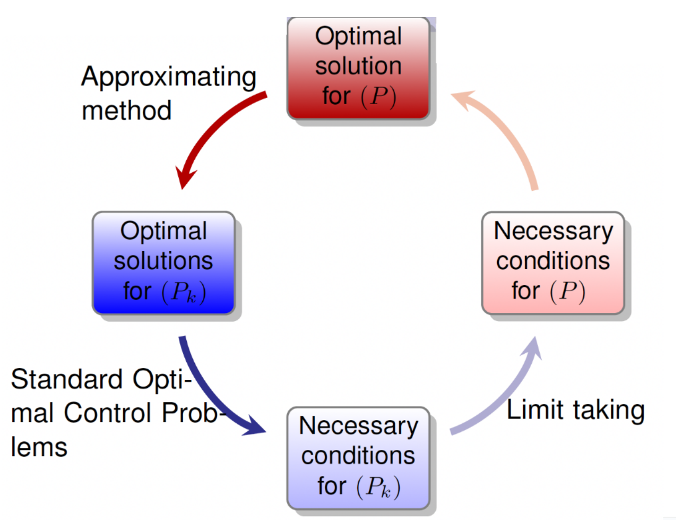

Figure 1.7 Exponential penalization technique

Work of Zeidan et al. in [70, 55, 58]

The domain of applicability of the exponential penalization technique for the results in [30, 31] was later enlarged in [70, 55, 58], to include *strong* local minimizers for controlled sweeping processes having:

- *nonsmooth* perturbation f,
- a *constant* sweeping set C that is *nonsmooth prox-regular* (i.e., C is the intersection of a finite number of zero-sublevel sets of $\mathcal{C}^{1,1}$-*generators* $\psi_1(x), \cdots, \psi_r(x)$ near C), and the functions ψ_i's satisfy a *constraint qualification on* the set C,
- a *final* state constraint set $C_T \subset \mathbb{R}^n$, the cost depends on *both* state-endpoints.

Furthermore, therein:

- the normal cone, N_C, in (1.2) is replaced by a subdifferential, $\partial\varphi$, of a function φ with domain C, and it is shown that such a system is equivalent to (1.2) with a different f. Indeed, the function φ is extended to a function ϕ that is $\mathcal{C}^{1,1}$ on $\mathbb{R}^n$ and that enjoys a globally Lipschitz gradient. Using a formula recently established in [35] for the Clarke subdifferential of an amenable function, the following formula is obtained

  $$\partial\varphi(x) = \{\nabla\phi(x)\} + N_C(x), \quad \forall x \in C.$$

  Using this formula, the dynamic can be rephrased as the original sweeping process

  $$\dot{x}(t) \in f_\phi(t, x(t), u(t)) - N_C(x(t)), \tag{1.4}$$

  where $f_\phi(t, x, u) = f(t, x, u) - \nabla\phi(x)$.
- When $C_T \subsetneq \mathbb{R}^n$, the *convexity* of the sets $f(t, x, U(t))$ is required in [55, 58].
- When C is unbounded, a *restrictive* assumption, (A2.4), is imposed in [58] on the set C and is shown to hold for convex, compact boundary, or polyhedral sets, but *not* for *general* prox-regular sets.
- Note that the *nontriviality* condition in the maximum principle of [55] is simply $\lambda + \|p(T)\| = 1$ and does not invoke the total variation of the measure.

- New subdifferentials are used that are *strictly smaller* than the Clarke and Mordukhovich subdifferentials. The key for this surprising result is the design of an approximating problem whose optimal state remains *entirely in the interior of the set* C.
- Note that in the case when $r > 1$, the invariance of C itself is not always valid, but requires extra *restrictive* hypothesis (see [34]). However, as shown in [58], the invariance of C itself is not essential for the success of this method, as it suffices to establish the invariance of certain *ingenuous* approximations of C from its *interior*, namely, C^{γ_k}, the zero-sublevel set of a *special* single function ψ_{γ_k} approximating $\psi := \max\{\psi_i\}$, and the corresponding $C^{\gamma_k}(k) \subset C^{\gamma_k}$. Furthermore, the *uniform* bounded variation property for the adjoint variables p_{γ_k} of the approximating problems, was cleverly established by employing the *strict diagonal dominance* condition on the Gramian matrix for the gradients of the active constraints at the prescribed optimal solution $\bar{x}$ of the original problem.

1.2.2.2 Selected results for time-dependent sweeping set $C(t)$.

Work of dePinho et. al in [33, 34]

In [33] and later in [34] (independently from [58]), the authors extended their previous *smooth* Pontryagin principle for *global* minimizers in [30], using the exponential penalty-type technique, to the case where:

- the perturbation $f(t, \cdot, u)$ is smooth, $f(t, x, U)$ is convex,
- the sweeping set $C(t)$ is *time-dependent* and *nonsmooth*,
- Gr $C(\cdot)$ is compact,
- the sweeping sets are assumed to have $\mathcal{C}^2$- generators $(\psi_i(t, x))_{i=1}^r$ satisfying a *global* constraint qualification, and a *global diagonal dominance condition* on the Gramian matrix for the gradients of the active constraints is imposed,
- other *demanding* conditions are assumed on the set $C(t)$:
    - $\nabla_x \psi_i(t, \cdot) = 0$ on the complement in $C(t)$ of a uniform band around the boundary

of $C(t)$,

  - $\langle \nabla_x \psi_i(t,x), \nabla_x \psi_j(t,x) \rangle \geq 0$ in a band around the boundary of $C(t)$, that is, all the *corners* of $C(t)$ must have *obtuse* angles. In particular, this last assumption excludes many important sets, including simple ones, like triangles, polytopes or sets with one or more acute angles, etc,

- the initial and final state sets are compact.

- In [33], their exponential penalty technique here deviated from (1.3) by using instead of $\psi(t,x)$, the function $\psi(t,x) - \sigma_k$, where $\sigma_k \searrow 0$, and hence, $C(t)$ is approximated by sets $C^k(t) \supset C(t)$ from the outside and *not* from the interior of $C(t)$.

Work of Hermosilla-Palladino in [42]

In [42], a **different** approach is used to establish a *variant* of *nonsmooth* Pontryagin-type maximum principle for *strong* local minimizers in a controlled sweeping process when:

- the moving set $C(t)$ is, as in [58, 34], nonsmooth and non-convex,
- the set $C(t)$ is uniformly prox-regular,
- the generating functions h_i together with ∇h_i are Lipschitz on a neighborhood of Gr $\bar{x}$, and $(\nabla h_i)_{i=1}^{i=r}$ satisfy a positive linear independence constraint qualification,
- the multifunction $C(\cdot)$ is *Lipschitz* continuous,
- the initial state is *fixed* and the final state is *free*,
- unlike the expected nontriviality condition ($\lambda = 1$ in their case), an *atypical* nondegeneracy condition is obtained which would require further understanding,
- the results involve the standard Clarke and Mordukhovich subdifferentials.

The authors constructed a sequence of standard optimal control problems having auxiliary controls and *explicit* state constraints emanating from the sweeping set, such that all admit the same optimal solution as the original problem.

1.3 Results and outline of this dissertation

1.3.1 Gaps in the literature and answering open questions

We summarize key results from the existing literature in the following comparison tables. These tables will help identify gaps in the past research, that will be addressed once the dissertation results are presented. Table 1.1 and Table 1.2 serve as a foundation for identifying open questions and demonstrating how this dissertation contributes to filling those gaps.

Table 1.1 Comparison of data in [58], [34], and [42]

Data: let $\bar{x}$ the prescribed optimal solution of the original problem			
Reference	**Assumptions on the perturbation f**	**Assumptions on the sweeping set**	**Other assumptions on the data**
[58]	$f(t,\cdot,u)$ *Lipschitz* on a neighborhood of $\bar{x}$	C is *constant nonsmooth prox-regular*, and the generators ψ_i's $C^{1,1}$ on a neighborhood of C and satisfy a *constraint qualification on* the set C	Initial state C_0 is *closed*
[58]	When $C_T \subsetneq \mathbb{R}^n$, the *convexity* of the sets $f(t,x,U(t))$ is required	When C is *unbounded*, a *restrictive* assumption is imposed on the set C and is shown to hold for convex, compact boundary, or polyhedral sets, but *not* for *general* prox-regular sets	Final state C_T is *closed*, and the cost depends on both state-endpoints

Table 1.1 (cont'd)

Reference	**Assumptions on the perturbation** f	**Assumptions on the sweeping set**	**Other assumptions on the data**
		A *strict local diagonal dominance* condition on the Gramian matrix for the gradients of the active constraints is imposed at $\bar{x}$	$U(t)$ is *time-dependent*, *closed* and *uniformly bounded* in t
[34]	$f(t,\cdot,u)$ is $\mathcal{C}^1$	$C(t)$ is *time-dependent, nonsmooth, prox-regular (implied)*, and the generators ψ_i's C^2 and satisfy a *constraint qualification*	Initial state C_0 is *compact*
	$f(t,x,U)$ convex	Gr $C(\cdot)$ is *compact*	Final state C_T is *compact*
		A *global diagonal dominance* condition on the Gramian matrix for the gradients of the active constraints is imposed, $\nabla_x\psi_i(t,\cdot) = 0$ on the complement in $C(t)$ of a uniform band around the boundary of $C(t)$, $\langle \nabla_x\psi_i(t,x), \nabla_x\psi_j(t,x)\rangle \geq 0$ in a band around the boundary of $C(t)$, that is, all the *corners* of $C(t)$ must have *obtuse* angles	U is *constant compact*

Table 1.1 (cont'd)

Reference	**Assumptions on the perturbation** f	**Assumptions on the sweeping set**	**Other assumptions on the data**
[42]	$f(t, \cdot, u)$ *Lipschitz* on a neighborhood of $\bar{x}$	$C(t)$ is *time-dependent, non-smooth, prox-regular*, and the generators h_i's $C^{1,1}$ locally and satisfy a *local constraint qualification*	Initial state is *fixed*
	$f(t, x, U(t))$ *not necessarily convex*	The set-valued map $C(\cdot)$ is *Lipschitz*	Final state is *free* $U(t)$ is *time-dependent* and *not necessarily unifromly bounded in* t.

Table 1.2 Comparison of results in [58], [34] and [42]

Results		
Reference	**Pontryagin's maximum principle**	**Existence results**
[58]	Exponential penalty approximation method *Typical* non-triviality condition Subdifferentials *smaller* than standard subdifferentials are used	Existence solution of the sweeping process, and existence of optimal solution

Table 1.2 (cont'd)

Reference	Pontryagin's maximum principle	Existence results
[34]	Exponential penalty approximation method *Typical* non-triviality condition *Standard* subdifferentials are used	*No* existence results
[42]	Different approximation method *Atypical* non-triviality condition *Standard* subdifferentials are used	*No* existence results

Conclusion I. Therefore, the question of establishing a Pontryagin maximum principle in its *expected* form (i.e., standard nontriviality condition, adjoint equation, transversality condition, and the maximality condition on the Hamiltonian) for optimal control problems over the sweeping process (1.2), remains *open* in each of the following settings:

(i) when the *nonsmooth moving* sweeping sets $C(t)$ are *bounded* and *general* (no restriction);

(ii) when the *general nonsmooth* sweeping sets are *unbounded* (*constant* or *moving*);

(iii) when *joint* state endpoints constraint set is *present*, the *convexity* of $f(t, x, U(t))$ is *absent*, or the *global* constraint qualification is only *local*, for all types of sweeping sets: *smooth*, *nonsmooth*, *constant*, *moving*, *bounded*, or *unbounded*.

In addition to the open problems in **Conclusion I**, new challenges arise when coupling (1.2) with a standard controlled differential equation, and when the joint endpoints constraint is on both states. So, throughout this thesis, we work on the optimal control problem (P), introduced in Chapter 4, governed by the following coupled dynamic (D), where $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^l$, and $u(t) \in U(t)$ a.e.,

$$(D) \begin{cases} \dot{x}(t) \in f(t, x(t), y(t), u(t)) - N_{C(t)}(x(t)), \text{ a.e. } t \in [0, T], \\ \dot{y}(t) = g(t, x(t), y(t), u(t)), \text{ a.e. } t \in [0, T], \end{cases}$$

$$(x(0), y(0), x(T), y(T)) \in S.$$

Our model incorporates different controlled submodels as particular cases:

- *coupled evolution variational inequalities* (see [1], [3], [6]),
- a subclass of *Integro-Differential sweeping processes* of Volterra type (see [5]),
- *second order sweeping processes*, in which the sweeping set is solely time-dependent (see, e.g., [53] for the general setting),
- and *Bolza-type* problems associated to (P).

In other words, optimal control problems governed by either of the four submodels can readily be formulated as a special case of (P) to which all the results of this thesis are applicable.

In [23], necessary conditions in the form of a *weak* maximum principle are derived for a certain form of a *Bolza* problem over a sweeping process. Excluding the *part* of their *integrand* involving $\dot{x}$ that is not covered in our setting, the remaining problem, therein, can be phrased as a *special* form of our problem (P) over a *coupled* sweeping process (D), where the sweeping set is a *constant polyhedron*, and the state endpoints are at most *periodic*.

On the other hand, in [6], a smooth Pontryagin maximum principle in its *expected* form is derived for a special case of our problem (P), namely, where the sweeping set is *constant*, *smooth*, and *strictly convex*, the perturbation f is *linear* in u, the function $g = (g_1, g_2)$ in the *coupled* controlled differential equations has g_1 *linear* in u and g_2 is *quadratic and convex* in u, the initial state is *fixed*, and the final state is *free*. The authors of [6] clearly noted that their method of *standard smooth* penalization does not apply even for the case of a *constant polyhedron* (which is a particular case of our general sweeping sets), and that including an "*additional terminal constraint*", a fortiori joint endpoints constraint, causes issues that are *not* treated therein.

> **Conclusion II.** Therefore, all the problems stated in **Conclusion I** are *open* when replacing the sweeping process (1.2) by (D), even when the sweeping set is *constant polyhedral.*

1.3.2 Findings and results of this thesis

In Chapters 3-4, we resolve all the aforementioned open problems in **Conclusions I** and **II**, while also establishing existence results for solutions to (D) and (P). In Chapter 5, we illustrate these theoretical results with an example and present several models that our findings can help solve.

1.3.2.1 Chapter 3

Chapter 3 is divided into local and global sections.
The **local sections** focus on analyzing the dynamic (D) and the sweeping set $C(t)$, as well as developing and studying a truncated dynamic $(\bar{D})$ and a truncated sweeping set $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ under local assumptions on the data. Two key local results in this chapter are Theorem 3.2.14, which approximates the truncated dynamic $(\bar{D})$ using a sequence of standard control systems $(\bar{D}_{\gamma_k})$, and Corollary 3.2.16, which establishes the existence and uniqueness of Lipschitz solutions to the Cauchy problem associated with $(\bar{D})$.
The main result of the **global section**, Theorem 3.3.7, proves the existence and uniqueness of a *Lipschitz* solution for the *Cauchy problem* corresponding to our dynamic (D) without requiring any *Lipschitz* behavior on the *nonsmooth* moving sets $C(t)$. Instead, we assume Gr $C(\cdot)$ is bounded and the gradients of the active generators are positively linear independent $(A3.2)_G$. Note that this is the first result of its kind for *general nonsmooth moving* sweeping sets, even for system (1.2), that is based on the method of exponential penalty approximation. It is essential for developing a *numerical algorithm* to solve optimal control problems over

such sweeping processes, which is the topic of our forthcoming project.

1.3.2.2 Chapter 4

The only **global** result of Chapter 4 is Theorem 4.1.1, which establishes the existence of a global *optimal solution* for our problem (P) over (D) with joint endpoints constraint set S. This result justifies the pursuit of a Pontryagin maximum principle for an optimal solution of (P).

The main result of Chapter 4, which is **local**, answers collectively all the open questions displayed above and generalizes all previously known results on Pontryagin maximum principle in multiple ways. More specifically, in Theorem 4.2.11 we derive under *minimal* assumptions on the data, a *complete* set of necessary conditions in the form of *nonsmooth* Pontryagin maximum principle for a *strong* local minimizer $((\bar{x}, \bar{y}), \bar{u})$ of the Mayer problem (P) governed by the *coupled* sweeping system (D) together with the *joint* endpoints constraint set S. The *moving* sweeping sets $C(t)$ are *general*, *nonsmooth*, *bounded* or *unbounded*, uniformly *prox-regular*, and defined as the intersection of a finite number of zero sub-level sets of the generators $(\psi_i(t, \cdot))_{i=1}^r$. Note the following.

- The optimal control problems studied in [34, 58] are over (1.2) and not over the general system (D).
- Noteworthy, unlike the result derived in [34] where the sweeping sets are not only assumed to be *bounded*, but satisfy *restrictive* assumptions on their corners (obtuse angles) and on the gradients of their generators ($\nabla_x \psi_i(t, \cdot) = 0$ in a zone in $C(t)$), *no* such *restrictive* assumptions are required in our result over (D), whether the *nonsmooth* moving sweeping sets $C(t)$ are *bounded or unbounded*. While when $C(t) \equiv C$ is a *constant* set, this corner assumption in [34] was removed in [58], its removal is far more intricate when $C(t)$ are *moving sets* (see Section 3.2.2 and Theorem 3.2.14).
- In contrast of the result in [58] established for a *restrictive* class of *constant unbounded* sweeping sets, our result here is valid for *general unbounded, moving*, and *prox-regular* sets that do not necessarily satisfy the restrictive assumption (A2.4) of [58].

- In addition, the *convexity* assumption of the sets $f(t, x, U(t))$ in [34] and [58] is now discarded and not only for the *separable* endpoints case treated therein, but also for general *joint endpoints constraints*.
- Furthermore, as opposed to the *global* constraint qualification on the generators of the sweeping set C in [58] and of $C(t)$ in [34], our constraint qualifications are required to *only* hold at $\bar{x}(t)$ (see (A3.2) and (A3.3)), where (A3.3) is vacuous in the smooth case ($r = 1$).
- Our *nontriviality* condition is simply $\lambda + \|p(T)\| = 1$ and does not invoke the measure corresponding to $x(t) \in C(t)$. This is the *expected* form in a Pontryagin maximum principle for problems over controlled sweeping processes (see [55, 58, 34]).
- In our adjoint inclusion and transversality condition we employ the recently introduced subdifferentials in [55] that are strictly smaller than the Clarke and Mordukhovich subdifferentials.

1.3.2.3 Chapter 5

In this chapter, we provide an example that highlights the significance of our initial model and the practical utility of our results.

1.3.3 Novelty of the methods employed.

There are *three* separate matters to tackle when establishing a Pontryagin maximum principle for a $\bar{\delta}$-minimizer $((\bar{x}, \bar{y}), \bar{u})$ of our problem (P).

1. The **first** matter is the possible *unboundedness* of the moving sweeping sets $C(t)$ and the joint endpoints set S, and the unboundedness of $\mathbb{R}^l$ (the sweeping set for the coupled controlled ODE).
2. The **second**, which is present *even* if $C(t)$ is *bounded* and/or the sweeping process is taken to be (1.2) instead of (D), is that the constraint qualification, (A3.2), on the active generators of $C(t)$ is *only* valid at $\bar{x}$.
3. The **third** is the *absence* of a Pontryagin maximum principle in its expected form for a Mayer problem over (1.2), where the *nonsmooth moving* sweeping sets are *general*

and *bounded*.

The two diagrams in the following pages (see Figures 1.8-1.9) outline the key steps of our approach for the maximum principle, illustrating how we transition from working on the problem (P) to defining a new truncated problem $(\bar{P}_{\delta,\delta})$ to establishing a nonsmooth Pontryagin maximum principle for this truncated problem. We encourage the reader to first examine the two diagrams before proceeding to the following paragraphs. In addition to the techniques shown in the diagrams, we also present the following additional techniques:

- To avoid imposing $\nabla_x \psi_i(t, \cdot) = 0$ on the complement in $C(t)$ of a uniform band around the boundary of $C(t)$, and to establish the uniform bounded variation property of the adjoint variable for the approximating problem, we construct a modified version of ψ_i, $\hat{\psi}_i$, that preserves the original constraint set $C(t)$ and the properties of ψ_i, and whose gradient is zero in certain areas.
- Another useful technique for the uniform bounded variation property of the adjoint variable for the approximating problem is to construct another transformation $\tilde{\psi}_i$ of ψ_i such that the Gramian matrix of the gradients of $\tilde{\psi}_i$ is strictly diagonally dominant, a condition stronger than the local strict diagonal dominance of the Gramian matrix corresponding to the gradients of the active constraints assumed for ψ_i at $\bar{x}$ ((A3.3)). After formulating the max principle in terms of $\hat{\psi}_i$ and $\tilde{\psi}_i$, we then translate the conditions to be formulated in terms of ψ_i.
- To remove the convexity assumption on $(f, g)(t, x, y, U(t))$, we extend the *relaxation* technique from [70] to address: (a) strong local minimizers, (b) time-dependent sweeping sets, $C(t)$, not necessarily moving in an absolutely continuous way, and (c) *general joint state endpoints* constraint set S.

In our case, obtaining the necessary conditions via the penalty-type approximating technique, can be summarized in Figure 1.10. Using our approach to the exponential penalty method without truncating $C(t)$, we prove in Section 3.3.2 the *existence* and *uniqueness* of a Lipschitz solution to the Cauchy problem associated with (D), Theorem 3.3.7. The *existence* of an *optimal* solution for the problem (P), Theorem 4.1.1 employs general results developed in the Appendix.

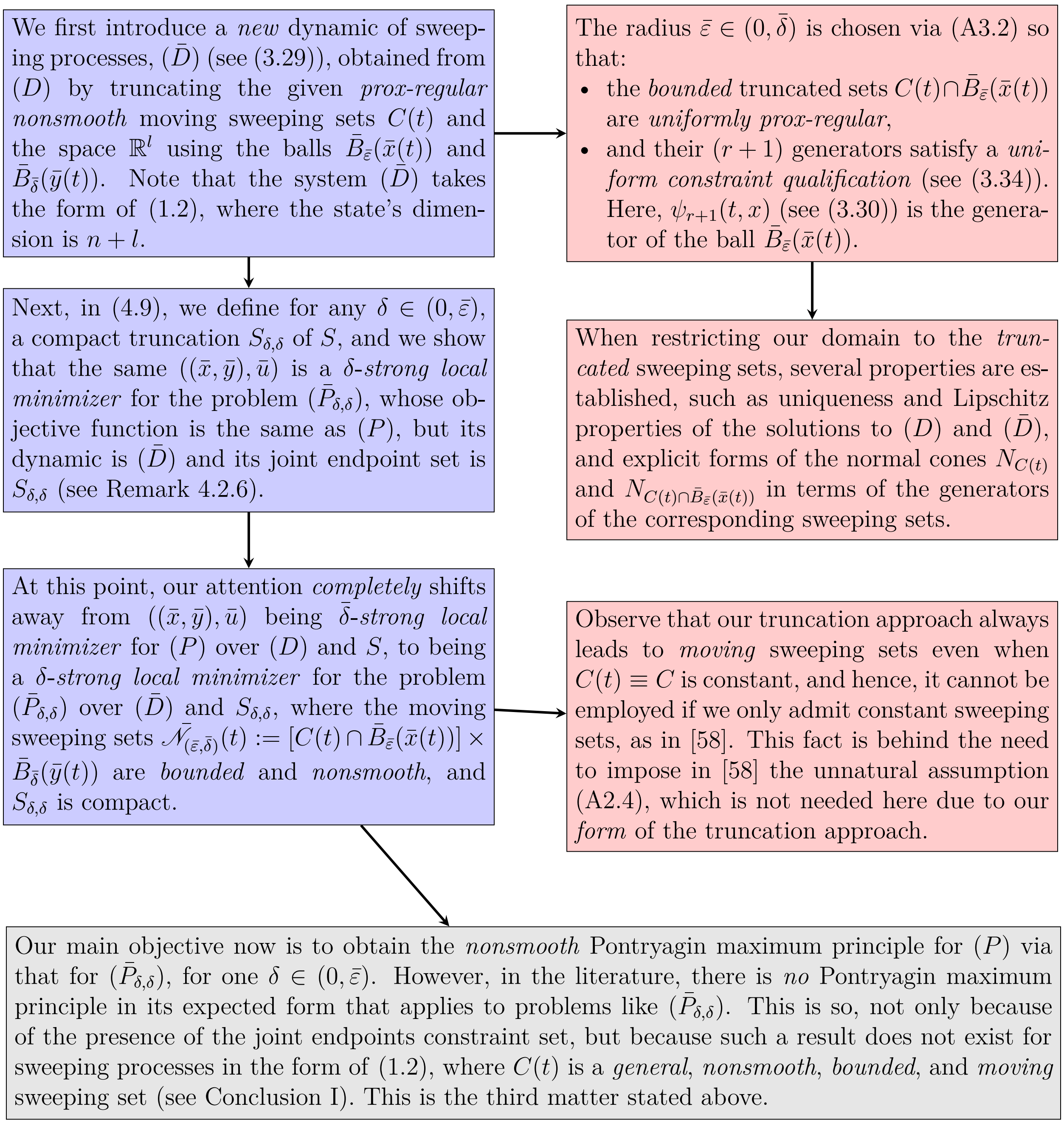

Figure 1.8 Flowchart of addressing the **first** and **second** matters above

To resolve this matter, for a specific choice δ_o of δ, we establish a *nonsmooth* Pontryagin maximum principle for the δ_o-*strong local minimizer* $((\bar{x}, \bar{y}), \bar{u})$ of $(\bar{P}_{\delta_o,\delta_o})$ that does *not* require any demanding conditions like the ones in [34] described prior to Conclusion I.

As such, we derive this result for problems over (1.2) with *general nonsmooth*, *bounded*, and *moving* sweeping sets, including the *joint* state endpoints constraints. This is accomplished by developing a *complete* new design of the method of exponential penalty approximation that drastically differs from the one in [34] and generalizes the one in [58] to the *complex* setting of *time-dependent* sweeping sets and to *joint* endpoint constraints.

We then produce a carefully crafted sequence of *smooth* sets, $(\bar{C}^{\gamma_k}(t) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t)))_k$ (see (3.59)), approximating our *nonsmooth* moving sweeping set $\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t) := [C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))] \times \bar{B}_{\bar{\delta}}(\bar{y}(t))$ from its *interior*, and we show that this smooth sequence (as opposed to the sweeping set itself) is actually *invariant* (see Remark 3.2.15) for our approximating dynamic $(\bar{D}_{\gamma_k})$.

This is different than [34] and far more intricate than [58]. Observe that in [34], the stringent condition on the corners is imposed so that the exponential penalty method works for *nonsmooth* bounded moving sweeping sets in the *same* manner as it did for *smooth* sets in [33], that is, by forcing their sweeping set $C(t)$ itself to be *invariant* for their approximating dynamic.

To obtain the boundedness of the multipliers for the approximated normal cones in our approximating dynamic, we construct another *smooth* approximation $\bar{\mathscr{A}}(t,k)$ of $\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)$, from its *interior*, where $\bar{\mathscr{A}}(t,k) := \bar{C}^{\gamma_k}(t,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t)) \subset \operatorname{int} \bar{C}_{\gamma_k}(t) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t))$. We then show that $\bar{\mathscr{A}}(t,k)$ is *invariant* for *our* approximated dynamic, the multipliers therein are *bounded*, and its solutions *approximate* the solutions to the original dynamic.

Whereas the authors in [34] approximated $C(t)$ from the *outside* by a sequence of invariant sets (as they did for the smooth bounded sets in [33]).

As our sweeping sets and their approximations are *time dependent*, the derivation of these results turns out to be quite involved and challenging in comparison with the constant sweeping set C treated in [58] (see section 3.2.2 and Theorem 3.2.14).

To address the presence of *joint* endpoint constraints, we pick carefully the value δ_o of δ so that we can successfully craft an approximation $S^{\gamma_k}(k)$ of S_{δ_o,δ_o} such that for k large, any solution of $(\bar{D}_{\gamma_k})$ with state endpoints in $S^{\gamma_k}(k)$ remains at all times in the invariant set $\bar{\mathscr{A}}(t,k)$.

At this point, we are ready to design for $(\bar{P}_{\delta_o,\delta_o})$ an approximating problem $(P^{\alpha,\beta}_{\gamma_k})$, defined over $(\bar{D}^{\beta}_{\gamma_k})$ (closely related to $(\bar{D}_{\gamma_k})$) and the joint endpoints constraint set $S^{\gamma_k}(k)$, whose optimal solution converges to $((\bar{x}, \bar{y}), \bar{u})$, see Proposition 4.2.8. Using intricate calculations, we showed that the corresponding adjoint variable sequence is *uniformly* of *bounded variation*, *without* assuming any restrictive condition like $\nabla_x \psi_i(t,\cdot) = 0$ in the zone of $C(t)$. A *careful* analysis of the limit as $k \to \infty$ to the *standard* maximum principle for $(P^{\alpha,\beta}_{\gamma_k})$ leads to our nonsmooth Pontryagin principle.

Figure 1.9 Flowchart of addressing the **third** matter above

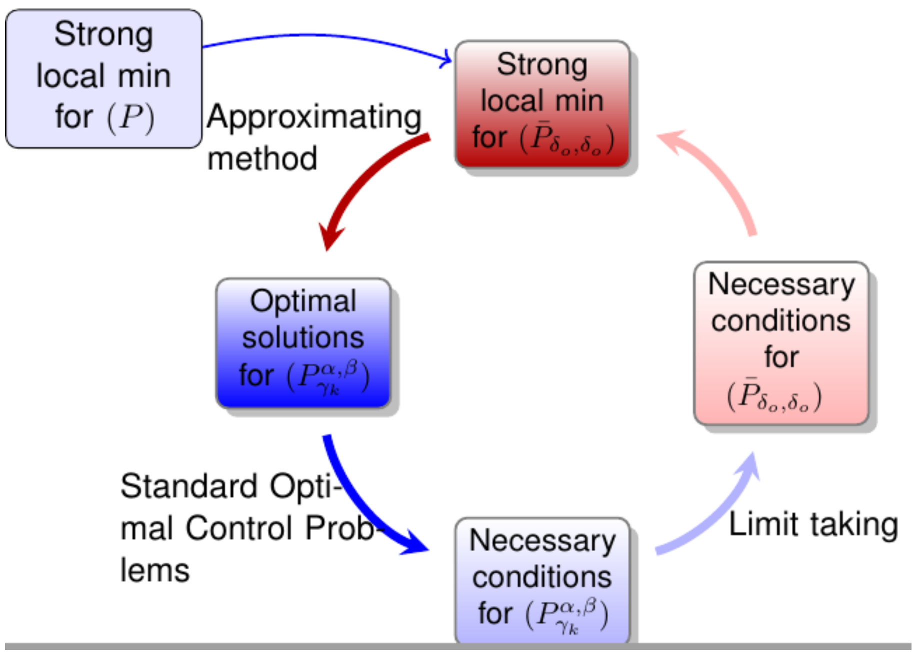

Figure 1.10 Exponential penalization technique in our setting

CHAPTER 2

PRELIMINARIES

In this chapter, we review foundational concepts, definitions, and key theorems from functional analysis, nonsmooth analysis, and control theory as presented in the literature, which we will use throughout this thesis.

2.1 Basic notions and concepts

In the first section, we present the basic notations and concepts used in the thesis.

- We denote by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ the Euclidean norm and the usual inner product, respectively.
- For $x \in \mathbb{R}^n$ and $a > 0$, we denote, respectively, by $B_a(x)$ and $\bar{B}_a(x)$ the open and closed ball centered at x and of radius a. More particularly, B and $\bar{B}$ represent the open unit ball and the closed unit ball, respectively.
- A vector function $f = (f_1, \cdots, f_n) : [0, T] \longrightarrow \mathbb{R}^n$ is said to be positive if f_i is positive for each $i = 1, \cdots, n$.
- $\mathbb{R}_+$ denotes the set of positive real numbers.
- We use $\mathscr{M}_{m\times n}[a, b]$ to indicate the set of $m \times n$-matrix functions on $[a, b]$. For $r \in \mathbb{N}$, we denote the identity matrix in $\mathscr{M}_{r\times r}$ by $\mathrm{I}_{r\times r}$.
- The interior, boundary, closure, convex hull, and complement of a set $S \subset \mathbb{R}^n$ are represented by $\operatorname{int} S$, $\operatorname{bdry} S$, $\operatorname{cl} S$, $\operatorname{conv} S$ and S^c, respectively.
- We note that ∇f of a function f is taken here to be a column vector, that is, the *transpose* of the standard gradient vector.
- For a set valued-map $S(\cdot) : [0, T] \rightsquigarrow \mathbb{R}^n$, $\operatorname{Gr} S(\cdot)$ denotes its graph.

Definition 2.1.1. A matrix $A = (a_{ij})$ of size $n \times n$ is said to be **strictly diagonally dominant** if it satisfies the following condition:

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}| \quad \text{for all } i = 1, 2, \ldots, n.$$

Lemma 2.1.2. By the Levy–Desplanques theorem, any strictly diagonally dominant matrix is nonsingular. Hence, its rows and columns form a basis in $\mathbb{R}^n$.

Limits of sets

Definition 2.1.3 (Limit of sets in the Kuratowski sense). Let $(S_k)_k$ a sequence of nonempty subsets of $\mathbb{R}^n$. We say that $(S_k)_k$ converges in the Kuratowski sense, or simply converges, to S whenever $\liminf_{k\to\infty} S_k = \limsup_{k\to\infty} S_k = S$.

This lemma can be found in [62, Exercise 4.3].

Lemma 2.1.4 (Limits of monotone and sandwiched sequences). We have that:

(a) $\lim_k S_k = \operatorname{cl} \bigcup_{k\in\mathbb{N}} S_k$ whenever $S_k \nearrow$, meaning $S_k \subset S_{k+1} \subset \cdots$;

(b) $\lim_k S_k = \bigcap_{k\in\mathbb{N}} \operatorname{cl} S_k$ whenever $S_k \searrow$, meaning $S_k \supset S_{k+1} \supset \cdots$;

(c) $S_k \to S$ whenever $S_k^1 \subset S_k \subset S_k^2$ with $S_k^1 \to S$ and $S_k^2 \to S$.

This definition can be found in [62, Example 4.13].

Definition 2.1.5 (Pompeiu-Hausdorff distance). For $C, D \subset \mathbb{R}^n$ closed and nonempty, the Pompeiu-Hausdorff distance between C and D is the quantity

$$d_\infty(C, D) := \sup_{x\in\mathbb{R}^n} |d_C(x) - d_D(x)|,$$

where the supremum could equally be taken just over $C \cup D$, yielding the alternative formula

$$d_\infty(C, D) = \inf \left\{ \eta \geq 0 \;\middle|\; C \subset D + \eta B,\ D \subset C + \eta B \right\}.$$

Definition 2.1.6 (Limit of sets in the Hausdorff sense). Let $(S_k)_k$ a sequence of nonempty closed subsets of $\mathbb{R}^n$. We say that $(S_k)_k$ converges with respect to Pompeiu-Hausdorff distance to S when $d_\infty(S_k, S) \to 0$.

Support function of a set

We introduce definitions and results related to the support function of a set S.

Definition 2.1.7 (Support function of a set S). Let $S \subset \mathbb{R}^n$ be nonempty. The support of S is $\sigma(\cdot, S) : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ defined by

$$\sigma(s^*, S) := \sup\{\langle s^*, s\rangle \ \mid \ s \in S\}. \tag{2.1}$$

Lemma 2.1.8. Let S be a closed and convex set of $\mathbb{R}^n$. Then

$$s \in S \iff \langle s^*, s\rangle \le \sigma(s^*, S) \quad \forall s^* \in \mathbb{R}^n. \tag{2.2}$$

Lemma 2.1.9 (Limits of sets and their support functions). Let $(S_k)_k$ a sequence of nonempty compact convex subsets of $\mathbb{R}^n$. Then

$$(S_k)_k \text{ converges to } S, \text{as } k \to \infty \iff \sigma(s^*, S_k) \longrightarrow \sigma(s^*, S), \text{ as } k \to \infty, \quad \forall s^* \in \mathbb{R}^n. \tag{2.3}$$

Lemma 2.1.10 (Hausdorff limits of sets and their support functions). Let $(S_k)_k$ a sequence of nonempty closed convex subsets of $\mathbb{R}^n$. We have, by [63, Theorem 6], that

$$(S_k)_k \xrightarrow[k\to\infty]{\text{Hausdorff}} S \quad \iff \quad \sigma(s^*, S_k) \xrightarrow[k\to\infty]{\text{unif in } s^*} \sigma(s^*, S), \qquad \forall s^* \in \mathbb{R}^n : \|s^*\| \le 1. \tag{2.4}$$

2.2 Nonsmooth analysis

Normal cones: proximal, limiting, and Clarke

Some of the definitions and results in this section are adapted from [17]. For standard references, see the monographs [18, 22, 48, 62, 66].

Proximal normal cone

Let $S \subset \mathbb{R}^n$ a nonempty closed set. For $x \in \mathbb{R}^n$, we recall that the **distance from x to S** is defined by

$$d_S(x) := \inf_{s\in S} \|x - s\|.$$

We can verify that $d_S(\cdot)$ is 1-Lipschitz on $\mathbb{R}^n$, and that there exists at least one point $s \in S$ such that

$$d_S(x) = \|x - s\|.$$

This point s is called **closest point** or **projection** of x onto S. We note that all closest points form a set denoted by $\text{proj}_S(x)$, see Figure 2.1. For $x \in \mathbb{R}^n \setminus S$ and $s \in \text{proj}_S(x)$, we have:

- The vector $x - s$ is called a **proximal normal direction** to S at s.
- For all $t > 0$ any vector $\zeta = t(x - s)$ is called **proximal normal vector** to S at s.

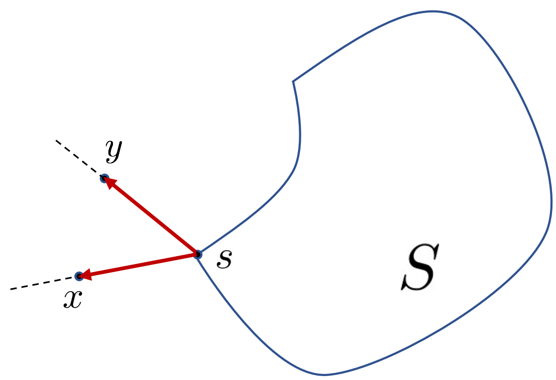

Figure 2.1 Proximal normal cone [1]

Definition 2.2.1. The set of all nonnegative multiple ζ of $x - s$ is called the **proximal normal cone to S at s** and is denoted by $N_S^P(s)$, see Figure 2.1. Thus

$$N_S^P(s) := \{t(x-s) : s \in \text{proj}_S(x) \text{ and } t \geq 0\}.$$

We can also characterize the proximal normal cone analytically and geometrically through the following two representations. Let $s \in S$.

$$\begin{aligned}
\zeta \in N_S^P(s) \quad &\Longleftrightarrow \quad \exists\, \lambda > 0 \text{ such that } \text{proj}_S(s + \lambda\zeta) = \{s\} \\
&\Longleftrightarrow \quad \exists\, \sigma = \sigma(\zeta, s) \geq 0 \text{ s.t. } \langle \zeta, s' - s\rangle \leq \sigma \|s' - s\|^2 \quad \forall s' \in S \\
&\Longleftrightarrow \quad \exists\, \sigma = \sigma(\zeta, s) \geq 0, \eta > 0 \text{ s.t. } \langle \zeta, s' - s\rangle \leq \sigma \|s' - s\|^2 \quad \forall s' \in B\Big(s, \eta\Big) \cap S \\
&\Longleftrightarrow \quad \exists\, r = r(\zeta, s) > 0 \text{ s.t. } B\Big(s + r\frac{\zeta}{\|\zeta\|}; r\Big) \cap S = \emptyset,
\end{aligned}$$

i.e. ζ is realized by an r-sphere (**ball characterization**, see Figure 2.2)

[1]This image was generated by Dr. Chadi Nour.
[2]This image was generated by Dr. Chadi Nour.

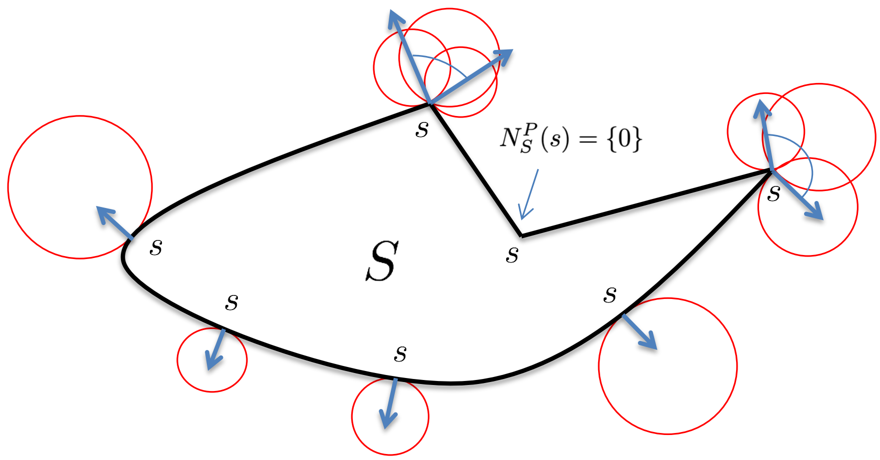

Figure 2.2 Ball characterization [2]

Remark 2.2.2. We have the following:

- $N_S^P(s) = \{0\}$ if $s \in S$ is not the projection of any point $x \notin S$ onto S. Hence, $N_S^P(s) = \{0\}$ when $s \in \text{int}\, S$.
- The proximal normal cone is a convex cone. It is not necessarily open nor closed.

Proposition 2.2.3 (Convex cone). Let $S \subset \mathbb{R}^n$ be a nonempty, closed and convex set. Thus

$$\zeta \in N_S^P(s) \iff \langle \zeta, s' - s \rangle \leq 0 \; \forall s' \in S. \tag{2.5}$$

In this case

- If $s \in \text{bdry}\; S$ then $N_S^P(s) \neq \{0\}$.
- For $s \in \text{bdry}\; S$

$$0 \neq \zeta \in N_S^P(s) \iff \zeta \text{ is realized by an } r\text{-sphere } \forall\, r > 0. \tag{2.6}$$

Lemma 2.2.4 (Local property of limiting normal cone). We deduce from the third equivalence in Definition 2.2.1 that the P-normality is a local property, meaning that the proximal normal cones $N_{S_1}^P(s) = N_{S_2}^P(s)$ if S_1 and S_2 are the same in a neighborhood of s.

Limiting or Mordukhovich normal cone

Definition 2.2.5. The **Limiting or Mordukhovich normal cone** to S at s, $N_S^L(s)$, is defined as

$$N_S^L(s) := \{v \in \mathbb{R}^n : \exists\, s_i \xrightarrow{S} s, \exists\, v_i \longrightarrow v, v_i \in N_S^P(s_i)\},$$

where $s_i \xrightarrow{S} s$ means that $s_i \longrightarrow s$ and $s_i \in S \;\; \forall i$.

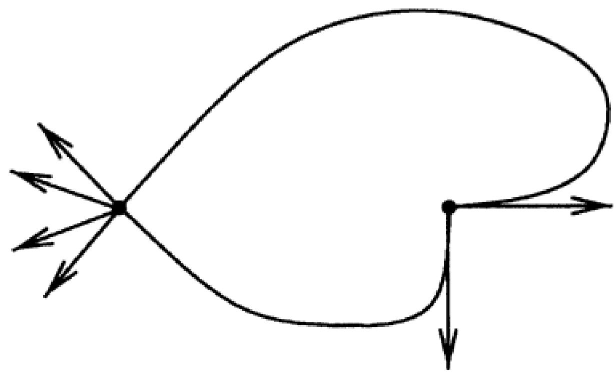

Figure 2.3 Limiting normal cone [3]

Remark 2.2.6. We have the following:

- If $s \in \operatorname{bdry} S$, then $N_S^L(s) \neq \{0\}$.
- The limiting normal cone is a closed cone. It is not necessarily convex.

Clarke normal cone

Definition 2.2.7. The **Clarke normal cone** to S at s, $N_S(s)$, is defined as

$$N_S(s) := \operatorname{conv}\, \{v \in \mathbb{R}^n : \exists\, s_i \xrightarrow{S} s, \exists\, v_i \longrightarrow v, v_i \in N_S^P(s_i)\}.$$

Remark 2.2.8. We have the following:

- The Clarke normal cone is a closed convex cone.

[3]This image is taken from [66].

Lemma 2.2.9 (**Monotonicity of the normal cone operator**). If $S \subseteq C$, then the normal cone satisfies the inclusion:

$$N_S(x) \supseteq N_C(x) \quad \text{for all } x \in S.$$

This means that if S is a subset of C, then any normal vector to C at a point in S is also a normal vector to S.

Definition 2.2.10. We say that a given set of vectors $\{x_i : i = 1, 2, \cdots, k\}$ in X is positively linearly independent if the following implication holds:

$$\sum_{i=1}^{k} \lambda_i x_i = 0, \ \lambda_i \geq 0 \implies \lambda_i = 0 \ \forall i \in \{1, 2, \cdots, k\}.$$

Lemma 2.2.11. [19, Corollary 10.44] Consider a set $S \subset \mathbb{R}^n$, given by

$$S = \{x \in \mathbb{R}^n : f_i(x) \leq 0, \ i = 1, 2, \cdots, k\},$$

where each function f_i is C^1 (locally, at least). Let $x \in S$, and assume $I(x) := \{i : f_i(x) = 0\}$ is a nonempty set, and $\{f_i'(x) : i \in I(x)\}$ is positively linearly independent (we say that the active constraints are positively linear independent). Then,

$$N_S(x) = \{ \sum_{i \in I(x)} \lambda_i f_i'(x) : \lambda_i \geq 0\}.$$

Proximal, Limiting, Clarke subdifferentials

We start some definitions and assumptions on an extended real-valued function f.

Definition 2.2.12. Let $X \subset \mathbb{R}^n$ and $f : X \to (-\infty, \infty]$.

- f is **lower semicontinuous** (lsc) at $x_0 \in X$ iff

$$f(x_0) \leq \liminf_{n \to \infty} f(x_n), \text{ for all } (x_n)_n \in X \text{ with } x_n \to x_0.$$

- f is **upper semicontinuous** (usc) at $x_0 \in X$ iff $-f$ is lsc at x_0.
- The **effective domain** of f is the set

$$\operatorname{dom} f := \{x \in X : f(x) < +\infty\}.$$

- The **epigraph** of f is the subset of $\mathbb{R}^{n+1}$ given by

$$\operatorname{epi} f := \{(x, w) \in X \times \mathbb{R} : x \in \operatorname{dom} f, f(x) \leq w\}.$$

- The **graph** of f is the subset of $\mathbb{R}^{n+1}$ given by

$$\operatorname{Gr} f := \{(x, f(x)) \in X \times \mathbb{R} : x \in \operatorname{dom} f\}.$$

Proximal subgradient

In the following $X \subset \mathbb{R}^n$ is an open set. We have $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ lsc on X such that $X \cap \operatorname{dom} f \neq \emptyset$, and $x \in X \cap \operatorname{dom} f$.

Definition 2.2.13. We denote by $\partial^P f(x)$ the **proximal subdifferential** of f at x. We have

$$\begin{aligned} \zeta \in \partial^P f(x) \quad &\Longleftrightarrow \quad (\zeta, -1) \in N^P_{\operatorname{epi} f}(x, f(x)) \\ &\Longleftrightarrow \quad \exists\, \sigma \geq 0 \text{ and } \eta > 0 \text{ such that for all } y \in B\big(x, \eta\big) \cap X, \\ &\qquad\quad f(y) \geq f(x) + \langle \zeta, y - x \rangle - \sigma \|y - x\|^2 \end{aligned}$$

ζ is said to be a **proximal subgradient** of f at x, see Figure 2.4.

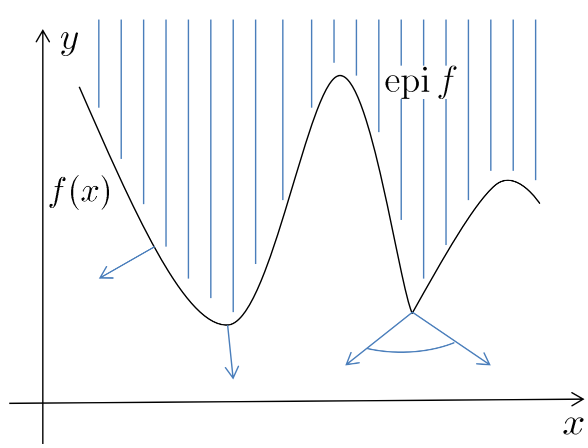

Figure 2.4 Proximal subgradient [4]

[4]This image was generated by Dr. Chadi Nour.

Remark 2.2.14. Note the following properties of the proximal subdifferential.

- $\partial^P f(x)$ is convex, however, it is not necessarily open, closed or nonempty.
- For all $c > 0$, we have $\partial^P(cf)(x) = c\partial^P f(x)$.
- $\partial^P f(x) + \partial^P g(x) \subset \partial^P(f+g)(x)$.

Remark 2.2.15. Let $U \subset X$ be open.

- If f is Gateaux differentiable at $x \in U$, then $\partial^P f(x) \subset \{f'_G(x)\}$.
- If $f \in C^2(U)$, then

$$\partial^P f(x) = \{f'(x)\} \ \forall x \in X.$$

Proposition 2.2.16. We assume that X is as well a convex set. Then f is K-Lipschitz on X iff

$$\|\zeta\| \leq K \quad \forall \zeta \in \partial^P f(x) \quad \forall x \in X.$$

Limiting subgradient

In the following $X \subset \mathbb{R}^n$ is an open set. We have $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ lsc on X such that $X \cap \operatorname{dom} f \neq \emptyset$, and $x \in X \cap \operatorname{dom} f$.

Definition 2.2.17. We denote by $\partial^L f(x)$ the **limiting subdifferential** of f at x. We have

$$\zeta \in \partial^L f(x) \quad \iff \quad (\zeta, -1) \in N^L_{\operatorname{epi} f}(x, f(x)).$$

ζ is said to be a **limiting subgradient** of f at x. Equivalently, we have

$$\partial^L f(x) := \{\lim_{i \to +\infty} \zeta_i : \zeta_i \in \partial^P f(x_i), x_i \xrightarrow{f} x\},$$

where $x_i \xrightarrow{f} x$ means that $x_i \to x$ and $f(x_i) \to f(x)$.

Remark 2.2.18. Note the following properties of the limiting subdifferential.

- $\partial^L f(x)$ is closed for every x, and the multifunction $\partial^L f(\cdot)$ has a closed graph.
- For all $c > 0$, we have $\partial^L(cf)(x) = c\partial^L f(x)$.
- If one f, g is Lipschitz near x, then $\partial^L(f+g)(x) \subset \partial^L f(x) + \partial^L g(x)$.

Clarke generalized gradient, Hessian and Jacobian

Definition 2.2.19. Assume that $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ has $\operatorname{dom} f$ closed with non-empty interior, and that f is locally Lipschitz on $\operatorname{int}(\operatorname{dom} f)$. We denote by $\partial f(x)$ the **Clarke subdifferential or Clarke generalized gradient** of f at $x \in \operatorname{int}(\operatorname{dom} f)$. We have

$$\begin{aligned} \zeta \in \partial f(x) \quad &\iff \quad (\zeta, -1) \in N_{\operatorname{epi} f}(x, f(x)) \\ &\iff \quad \langle \zeta, v \rangle \leq f^{\circ}(x; v) \quad \forall v \in \mathbb{R}^n, \end{aligned}$$

where

$$f^{\circ}(x; v) := \lim \sup_{y \to x,\, h \downarrow 0} \frac{f(y + hv) - f(y)}{h}.$$

Equivalently, we have

$$\partial f(x) \quad = \quad \operatorname{conv} \partial^L f(x),$$

and

$$\partial f(x) \quad = \quad \operatorname{conv} \left\{ \lim_{i \to +\infty} \nabla f(x_i) : x_i \xrightarrow{O} x, \nabla f(x_i) \text{ exists } \forall i \right\},$$

where O is any full-measure subset of $\operatorname{int}(\operatorname{dom} f)$.

Proposition 2.2.20. Take a lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a point $\bar{x} \in \mathbb{R}^n$. Assume that f is Lipschitz continuous on a neighborhood of $\bar{x}$ with Lipschitz constant K. Then:

$$\partial f(\bar{x}) \subset KB.$$

Definition 2.2.21. Assume that f is $\mathcal{C}^{1,1}$ on $\operatorname{int}(\operatorname{dom} f)$. We denote by $\partial^2 f(x)$ the **Clarke generalized Hessian** of f at $x \in \operatorname{int}(\operatorname{dom} f)$. We have

$$\partial^2 f(x) \quad = \quad \operatorname{conv} \left\{ \lim_{i \to +\infty} \nabla^2 f(x_i) : x_i \xrightarrow{O} x, \nabla^2 f(x_i) \text{ exists } \forall i \right\},$$

where O is any full-measure subset of $\operatorname{int}(\operatorname{dom} f)$.

Remark 2.2.22. Notice that

- $\partial f(\cdot)$ and $\partial^2 f(\cdot)$ are locally bounded and measurable multifunctions with closed graph, and their values are nonempty, compact and convex.

Definition 2.2.23. Assume that $g : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz function near $x \in \mathbb{R}^n$, i.e. Lipschitz on a open set Ω containing x. We denote by $\partial g(x)$ the **Clarge generalized Jacobian** of g at x. We have

$$\partial g(x) \;=\; \operatorname{conv}\left\{\lim_{i\to+\infty} Jg(x_i) : x_i \xrightarrow{O} x, Jg(x_i) \text{ exists } \forall i\right\},$$

where O is any full-measure subset of Ω, and J is the Jacobian operator.

Remark 2.2.24. Notice that

- The multifunction $\partial g(\cdot)$ is measurable and has closed graph. Its values are nonempty, convex, and compact in the space of $n \times n$ matrices.

Non-standard notions of subdifferentials

In [70], and later in [55], Zeidan, Nour and Saoud extended the notions of Limiting subdifferential, Clarke generalized gradient, Hessian and Jacobian to nonstandard notions for subdifferentials that are strictly smaller than their standard counterparts.

Extended Limiting subdifferential, Clarke generalized gradient, Hessian and Jacobian

Definition 2.2.25. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a lsc function and $S \subset \operatorname{dom} f$ be a closed set with $\operatorname{int}(\operatorname{dom} f) \neq \emptyset$. For $x \in \operatorname{cl}(\operatorname{int} S)$, we denote by $\partial_l^L f(x)$ to be the **limiting subdifferential of f relative to** $\operatorname{int} S$ at x, and we have:

$$\partial_l^L f(x) := \{\lim_{i\to+\infty} \zeta_i : \zeta_i \in \partial^P f(x_i), x_i \in \operatorname{int} S, x_i \xrightarrow{f} x\}.$$

Remark 2.2.26. We have that

- The multifunction $\partial_l^L f(\cdot)$ has closed graph, and closed values.
- If f Lipschitz on $\operatorname{int} S$, then for any $x \in \operatorname{cl}(\operatorname{int} S)$, $\partial_l^L f(x)$ is nonempty and compact.
- For all $x \in S$, we have $\partial_l^L f(x) \subset \partial^L f(x)$, and equality holds when $x \in \operatorname{int} S$.

Definition 2.2.27. Assume that $f : \mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$ is locally Lipschitz on $\mathrm{int}\,(\mathrm{dom}\, f) \neq \emptyset$. We denote by $\partial_l f(x)$ the **Extended Clarke generalized gradient** of f at $x \in \mathrm{cl}\,(\mathrm{int}\,(\mathrm{dom}\, f))$. We have

$$\partial_l f(x) \;\;=\;\; \mathrm{conv}\left\{\lim_{i\to+\infty} \nabla f(x_i) : x_i \xrightarrow{O} x, \nabla f(x_i) \text{ exists } \forall i\right\},$$

where O is any full-measure subset of $\mathrm{int}\,(\mathrm{dom}\, f)$.

Definition 2.2.28. Assume that $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is $\mathcal{C}^{1,1}$ on $\mathrm{int}\,(\mathrm{dom}\, f) \neq \emptyset$. We denote by $\partial_l^2 f(x)$ the **Extended Clarke generalized Hessian** of f at $x \in \mathrm{cl}\,(\mathrm{int}\,(\mathrm{dom}\, f))$. We have

$$\partial_l^2 f(x) \;\;=\;\; \mathrm{conv}\left\{\lim_{i\to+\infty} \nabla^2 f(x_i) : x_i \xrightarrow{O} x, \nabla^2 f(x_i) \text{ exists } \forall i\right\},$$

where O is any full-measure subset of $\mathrm{int}\,(\mathrm{dom}\, f)$.

Remark 2.2.29. Notice that

- $\partial_l f(\cdot)$ and $\partial_l^2 f(\cdot)$ are measurable multifunctions with closed graph, and their values are nonempty, compact and convex.
- We have $\partial_l f(x) \subset \partial f(x)$ and $\partial_l^2 f(x) \subset \partial^2 f(x)$, with equalities holding when $x \in \mathrm{int}\,(\mathrm{dom}\, f)$.

Definition 2.2.30. Assume that $g : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz function on a closet set $S \subset \mathbb{R}^n$. We denote by $\partial_l g(x)$ the **Extended Clarke generalized Jacobian** of g at $x \in S$ that extends the Clarke generalized Jacobian to the boundary of S . We have

$$\partial_l g(x) \;\;=\;\; \mathrm{conv}\left\{\lim_{i\to+\infty} Jg(x_i) : x_i \xrightarrow{O} x, Jg(x_i) \text{ exists } \forall i\right\},$$

where O is any full-measure subset of $\mathrm{int}\, S$, and J is the Jacobian operator.

Remark 2.2.31. Notice that

- The multifunction $\partial_l g(\cdot)$ is measurable and has closed graph. Its values are nonempty, convex, and compact in the space of $n \times n$ matrices.

- We have $\partial_l g(x) \subset \partial g(x)$, with equality holding when $x \in \operatorname{int} S$.

Definition 2.2.32. Assume that $h : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^{1,1}$ on an open set containing a closet set $S \subset \mathbb{R}^n$. We denote by $\partial_l^2 h(x)$ the **Clarke generalized Hessian of h relative to** int **S** at $x \in S$. We have

$$\partial_l^2 h(x) \quad = \quad \operatorname{conv}\left\{\lim_{i\to+\infty} \nabla^2 h(x_i) : x_i \xrightarrow{O} x, \nabla^2 h(x_i) \text{ exists } \forall i\right\},$$

where O is any full-measure subset of $\operatorname{int} S$.

Remark 2.2.33. Notice that

- $\partial_l^2 h(\cdot)$ is a measurable multifunction with closed graph, and its values are nonempty, compact and covex.
- We have $\partial_l^2 h(x) \subset \partial^2 h(x)$, with equality holding when $x \in \operatorname{int} S$.

Prox-regular sets

We proceed to define the φ_0-convexity property and the prox-regularity of a set. A detailed analysis of this may be found in [21, 29]. For other related properties, we refer the reader to [60, 61, 62, 8, 9] and the references therein.

Definition 2.2.34. Suppose $S \subset \mathbb{R}^n$ is closed. S is said to be φ**-convex**, where φ is taken to be a continuous function from S to $[0, +\infty)$, if

$$\langle \zeta, y - x\rangle \le \varphi(x)\|\zeta\|\|y - x\|^2,$$

for all $x \in \text{bdry } S, y \in S$ and $0 \ne \zeta \in N_S^P(x)$.

Definition 2.2.35. Suppose $S \subset \mathbb{R}^n$ is closed. S is said to be φ_0**-convex** if we can find $\varphi_0 \ge 0$ such that

$$\langle \zeta, y - x\rangle \le \varphi_0\|\zeta\|\|y - x\|^2, \tag{2.7}$$

for all $x \in \text{bdry } S, y \in S$ and $0 \ne \zeta \in N_S^P(x)$.

Remark 2.2.36. We remark that S is φ-convex iff S is φ_0-convex locally.

Definition 2.2.37. Let $S \subset \mathbb{R}^n$ a closed set. We say that S is **r-proximally smooth** or **r-prox-regular** iff there exists $r > 0$ such that for all $x \in \operatorname{bdry} S$ and $\zeta \in N_S^P(x)$

$$\langle \zeta, y - x \rangle \leq \frac{1}{2r} \|\zeta\| \|y - x\|^2 \;\; \forall y \in S.$$

Equivalently, S is r-proximally smooth if and only if for all $x \in$ bdry S and $0 \neq \zeta \in N_S^P(x), \zeta$ is realized by an r-sphere, i.e.

$$B\Big(s + r\frac{\zeta}{\|\zeta\|}; r\Big) \cap S = \emptyset.$$

Remark 2.2.38. Notice the following:

- For $\varphi_0 > 0$, S is φ_0-convex $\iff$ S is $\frac{1}{2\varphi_0}$-prox-regular (or has positive reach with radius $\frac{1}{2\varphi_0}$) (see Figure 2.5).
- S is convex $\iff$ S is 0-convex $\iff$ S is r-prox-regular for all $r > 0$.

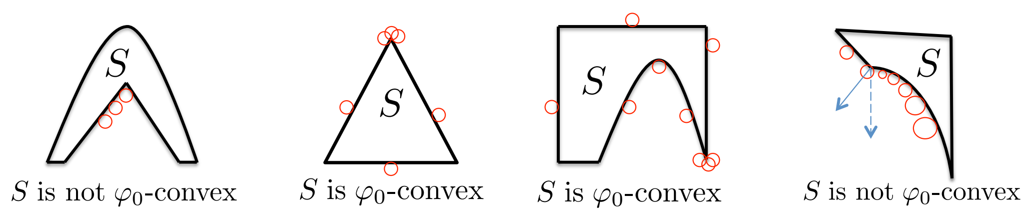

Figure 2.5 φ_0-convexity [5]

Proposition 2.2.39. Let S be r-prox-regular set in $\mathbb{R}^n$, with $r > 0$. Then we have:

(i) [21, Corollary 4.15] For all $x \in S$,

$$N_S(x) = N_S^P(x) = N_S^L(x),$$

and for all $x \in \operatorname{bdry} S$, we have

$$N_S^P(x) \neq \{0\}.$$

(ii) [21, Theorem 4.8] Let $r' \in (0, r)$. Then $\pi_S(\cdot)$ is Lipschitz of rank $\frac{r}{r-r'}$ on $\{u \in \mathbb{R}^n : 0 < d_S(u) < r'\}$, where $\pi_S(\cdot)$ is the projection map into S.

[5]This image was generated by Dr. Chadi Nour.

(*iii*) The normal cone $N_S^P(\cdot)$ is hypomonotone, i.e. for every $x_1, x_2 \in S$, for every $\xi_1 \in N_S^P(x_1)$, $\xi_2 \in N_S^P(x_2)$, ξ_1, ξ_2 unit vectors, we have

$$\langle \xi_2 - \xi_1, x_2 - x_1 \rangle \geq -\frac{1}{r} \|x_2 - x_1\|^2 . \tag{2.8}$$

The following result holds for uniform prox-regular sets $S(t)$. It is an extension of a special case of [55, Lemma 3.2] from a constant compact set S to the case of set-valued maps $S(\cdot)$ with non-compact values. It requires $N_{S(\cdot)}(\cdot)$ to have closed graph.

Lemma 2.2.40. Let $S(\cdot) : [0,T] \rightsquigarrow \mathbb{R}^n$ be such that, for all $t \in [0,T]$, $S(t)$ is nonempty, closed, and uniformly ρ^*-prox-regular, for some $\rho^* > 0$. Let $\bar{x} \in \mathcal{C}([0,T], \mathbb{R}^n)$ with $\bar{x}(t) \in S(t)$ for all $t \in [0,T]$. Assume that for some $\delta \in (0, \rho^*)$, the map $(t,y) \to N_{S(t)}(y)$ has closed graph on the domain $\mathrm{Gr}\left(S(\cdot) \cap \bar{B}_\delta(\bar{x}(\cdot))\right)$.Then, the following holds.

(*i*) Let $t \in [0,T]$, and $y \in S(t) \cap \bar{B}_\delta(\bar{x}(t))$. Then

$$N^P_{S(t)}(y) \cap -N^P_{\bar{B}_\delta(\bar{x}(t))}(y) = \{0\}.$$

(*ii*) There exists $\rho_\delta > 0$ such that for all $t \in [0,T]$ the set $S(t) \cap \bar{B}_\delta(\bar{x}(t))$ is ρ_δ-prox-regular.

(*iii*) For $t \in [0,T]$, $\pi(t,\cdot) := \pi_{\left(S(t)\cap \bar{B}_\delta(\bar{x}(t))\right)}(\cdot)$ is well-defined on $\left(S(t) \cap \bar{B}_\delta(\bar{x}(t))\right) + \rho_\delta B$ and 2-lipschitz on $\left(S(t) \cap \bar{B}_\delta(\bar{x}(t))\right) + \frac{\rho_\delta}{2}\bar{B}$.

Proof. (*i*)-(*ii*): The results are derived by following the proof of [55, Lemma 3.2], where for $t \in [0,T]$, we take $S := S(t)$ and $x := \bar{x}(t)$, and hence, $\mathcal{N}_x$ and ρ_x there , are respectively $\mathcal{N}_t := \mathcal{N}_{\bar{x}(t)}$ and $\rho_t := \rho_{\bar{x}(t)}$. It follows that ρ there is now $\hat{\rho} := \inf\{\rho_t : t \in [0,T]\}$. Following the rest of the proof there, and employing that the map $(t,y) \to N_{S(t)}(y)$ has closed graph on the domain $\mathrm{Gr}\left(S(\cdot) \cap \bar{B}_\delta(\bar{x}(\cdot))\right)$, we conclude that $\hat{\rho} > 0$. Thus, for all $t \in [0,T]$, we deduce that $S(t) \cap \bar{B}_\delta(\bar{x}(t))$ is ρ_δ-prox regular, where $\rho_\delta := \frac{\rho^*\hat{\rho}}{2}$.
(*iii*): It follows from (*ii*) and Proposition 2.2.39(*ii*). ☐

We now present Theorem 9.1 in [2].

Theorem 2.2.41. Let $g_k : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ with $k = 1, \dots, m$ be functions such that, for each $t \in [0,T]$, the set

$$S(t) = \{x \in \mathbb{R}^n : g_1(t,x) \leq 0, \dots, g_m(t,x) \leq 0\} \tag{2.9}$$

is nonempty. Assume that there exists some $\rho \in]0,+\infty]$ such that:

(i) for all $t \in [0,T]$, for all $k \in \{1,\dots,m\}$, $g_k(t,\cdot)$ is of class C^1 on $\{x \in \mathbb{R}^n : d(x,S(t)) < \rho\}$;

(ii) there exists a real $\gamma > 0$ such that, for all $t \in [0,T]$, for all $x \in \operatorname{bdry} S(t)$, for all $y \in \{y \in \mathbb{R}^n : d(y,S(t)) < \rho\}$, for all $k \in \{1,\dots,m\}$ with $g_k(t,x) = 0$,

$$\langle \nabla g_k(t,\cdot)(y) - \nabla g_k(t,\cdot)(x), y - x\rangle \geq -\gamma \|y - x\|^2.$$

Assume also that there is a real $\delta > 0$ such that, for any $(t,x) \in [0,T] \times \mathbb{R}^n$ with $x \in \operatorname{bdry} S(t)$ and any $\zeta \in \operatorname{conv}\{\nabla g_k(t,\cdot)(x) : k \in K(t,x)\}$ where $K(t,x) := \{k \in \{1,\dots,m\} : g_k(t,x) = 0\}$, there exists $v(t,x,\zeta) \in \bar{B}$ satisfying

$$\langle \zeta, v(t,x,\zeta)\rangle \leq -\delta.$$

Then, for all $t \in [0,T]$, the set $S(t)$ is r-prox-regular with $r = \min\left\{\rho, \frac{\delta}{\gamma}\right\}$.

Amenable and Epi-Lipschitzian sets

The following definitions and properties can be found in [22, 62].

Definition 2.2.42. Let $S \subset \mathbb{R}^n$. The set S is **amenable** at one of its points $\bar{x}$ if there exists an open neighborhood V of $\bar{x}$, a $\mathcal{C}^1$ mapping F from V into a space $\mathbb{R}^m$, and a closed, convex set $D \subset \mathbb{R}^m$ such that

$$S \cap V = \{x \in V \mid F(x) \in D\}, \tag{2.10}$$

with

$$\text{the only vector } y \in N_D(F(\bar{x})) \text{ with } \nabla F(\bar{x})^T y = 0 \text{ is } y = 0. \tag{2.11}$$

Lemma 2.2.43. Let $S = \{x \in X \mid F(x) \in D\}$ for closed, convex sets X, D, and a $\mathcal{C}^1$ mapping F. The set S is **amenable** at any of its points $\bar{x}$ where the constraint qualification holds, meaning that the only $y \in N_D(F(\bar{x}))$ with $-\nabla F(\bar{x})^T y \in N_X(\bar{x})$ is $y = 0$.

Definition 2.2.44. For a strictly continuous function $f : \mathbb{R}^n \to \mathbb{R}$, let $S = \{x \mid f(x) \leq \bar{\alpha}\}$ and consider a point $\bar{x}$ of S with $f(\bar{x}) = \bar{\alpha}$. S is said to have an **epi-Lipschitzian boundary** at $\bar{x}$ if

$$0 \notin \operatorname{conv} \partial f(\bar{x}). \tag{2.12}$$

Lemma 2.2.45. Let $S := \{x : f(x) \leq 0\}$, where $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz near x and $0 \in \partial f(x)$. Then, S is epi-Lipschitzian at x.

Lemma 2.2.46. (i) A set $S \subset \mathbb{R}^n$ with boundary point $\bar{x}$ is **epi-Lipschitzian** at $\bar{x}$ if and only if S is locally closed at $\bar{x}$ and the normal cone $N_S(\bar{x})$ is pointed, i.e. $N_S(\bar{x}) \cap -N_S(\bar{x}) = \{0\}$.
(ii) If the set S is epi-Lipschitzian at every $x \in S$, then $S = \operatorname{cl} \operatorname{int} S$.

Lemma 2.2.47. [57, Remark 4.8(ii)] If the lower semicontinuous multifunction F has closed and r-prox-regular values, for some $r > 0$, (as opposed to convex), then

$$\operatorname{conv}\left(\bar{N}^L_{F(t)}(\cdot)\right) = N^P_{F(t)}(\cdot) = N^L_{F(t)}(\cdot) = N_{F(t)}(\cdot),$$

and this cone is pointed at $x \in F(t)$ if and only if $F(t)$ is epi-Lipschitz at x. 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 $N^L_{F(\cdot)}(\cdot)$.

Sub-level sets of a function

The following is adapted from Lemma 3.3-3.4, Theorem 3.1 in [55], and Proposition 3.1 in [70].

Lemma 2.2.48. Let S be a nonempty set given by

$$S := \{x \in \mathbb{R}^n : \psi(x) \leq 0\}, \tag{2.13}$$

where ψ is $\mathcal{C}^{1,1}$ on $S + \rho B$, for some $\rho > 0$, ψ is coercive (i.e. $\lim_{\|x\|\to\infty} \psi(x) = +\infty$) or S bounded, and there is a constant $\eta > 0$ such that

$$\psi(x) = 0 \implies \|\nabla\psi(x)\| > 2\eta.$$

Part I. Let $2M_\psi$ be the Lipschitz constant of $\nabla\psi(\cdot)$ over the compact set $S + \frac{\rho}{2}\bar{B}$ such that $M_\psi \geq \frac{4\eta}{\rho}$. Then,

(i) bdry $S \neq \emptyset$ and bdry $S = \{x \in \mathbb{R}^n : \psi(x) = 0\}$.

(ii) int $S \neq \emptyset$ and int $S = \{x \in \mathbb{R}^n : \psi(x) < 0\}$.

(iii) The nonempty set S is compact, amenable (in the sense of [62]), epi-Lipschitzian,

$$S = \text{cl}(\text{int}\, S), \tag{2.14}$$

and S is $\frac{\eta}{M_\psi}$-prox-regular.

(iv) For all $x \in \text{bdry}\, S$ we have

$$N_S(x) = N_S^P(x) = N_S^L(x) = \{\lambda\nabla\psi(x) : \lambda \geq 0\}. \tag{2.15}$$

Part II. For $k \in \mathbb{N}$, we define the set $S(k)$ by

$$S(k) := \{x \in S : \psi(x) \leq -\alpha_k\}, \tag{2.16}$$

where $(\alpha_k)_k$ the real sequence defined by

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

$M > 0$ positive constant, $(\gamma_k)_k$ a sequence satisfying $\gamma_k > \frac{2M}{\eta}$ for all $k \in \mathbb{N}$, and $\gamma_k \to \infty$ as $k \to \infty$. Then, we have

(i) For all k, the set $S(k) \subset$ int S and is compact,

(ii) bdry $S(k) = \{x \in \mathbb{R}^n : \psi(x) = -\alpha_k\}$ and int $S(k) = \{x \in \mathbb{R}^n : \psi(x) < -\alpha_k\}$ for k sufficiently large,;

(iii) int $S(k)$ is nonempty, $C(k)$ is amenable, epi-Lipschitzian, $\frac{n}{2M_\psi}$-prox-regular, $S(k) =$ cl int $C(k)$, and

$$\forall x \in \text{bdry}\ S(k), \quad N_{S(k)}(x) = N_{S(k)}^P(x) = N_{S(k)}^L(x) = \{\lambda\nabla\psi(x) : \lambda \geq 0\}. \tag{2.17}$$

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

$$\left[S \cap \bar{B}_{r_o}(c)\right] - \rho_k \frac{\nabla\psi(c)}{\|\nabla\psi(c)\|} \subset \operatorname{int} S(k), \quad \forall k \geq \bar{k} \text{ and } \forall c \in \operatorname{bdry} S. \tag{2.18}$$

In particular, we have

$$\left(c - \rho_k \frac{\nabla\psi(c)}{\|\nabla\psi(c)\|}\right) \in \operatorname{int} S(k), \quad \forall k \geq \bar{k} \text{ and } \forall c \in \operatorname{bdry} S. \tag{2.19}$$

Other important results

The following theorem can be found in [66, Theorem 3.3.1].

Theorem 2.2.49 (Ekeland variational principle). Take a complete metric space $(X, d(\cdot,\cdot))$, a lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$, a point $x_0 \in \operatorname{dom} f$, and numbers $\alpha > 0$ and $\lambda > 0$. Assume that

$$f(x_0) \leq \inf_{x \in X} f(x) + \lambda\alpha.$$

Then there exists $\bar{x} \in X$ such that

(i) $f(\bar{x}) \leq f(x_0)$,

(ii) $d(x_0, \bar{x}) \leq \lambda$,

(iii) $f(\bar{x}) \leq f(x) + \alpha d(x, \bar{x})$ for all $x \in X$.

This result can be found in [47].

Lemma 2.2.50. Let

$$F(x) = \max_{1 \leq i \leq m} f_i(x), \quad \text{for } x \in X. \tag{2.20}$$

Define the smooth approximation:

$$F_p(x) = \frac{1}{p} \ln\left(\sum_{i=1}^{m} \exp\left(p f_i(x)\right)\right). \tag{2.21}$$

Then, for $x \in X$, $F_p(x)$ is a monotonically decreasing function in terms of p, and the following inequality holds:

$$F(x) \leq F_p(x) \leq F(x) + \frac{\ln(m)}{p}. \tag{2.22}$$

2.3 Differential equations, set-valued analysis, and control theory

Existence of ODE

Theorem 2.3.1 (Existence and Uniqueness for ODE). Consider the (IVP) system

$$\begin{cases} \dot{x}(t) = f(t,x) \\ x(t_0) = x_0. \end{cases}$$

If f is continuous in (t,x) in a rectangle $D = \{(t,x) : t_0 - \delta < t < t_0 + \delta, x_0 - b < x < x_0 + b\}$, and $f(t,x)$ lipschitz with respect to x on $R = \{(t,x) : t_0 - a < t < t_0 + a, x_0 - b < x < x_0 + b, a < \delta\}$, then the solution in R of the (IVP) exists and shall be unique.

The following is found in [39, Theorem 5.3].

Definition 2.3.2 (Carathéodory function). Suppose D is an open set in $\mathbb{R}^{n+1}$. The function $f : D \to \mathbb{R}^n$ is said to satisfy the Carathéodory conditions on D, if:

- $f(t,x)$ is measurable in t for each fixed x,
- $f(t,x)$ is continuous in x for each fixed t,
- For each compact set $U \subset D$, there exists an integrable function $m_U(t)$ such that

$$|f(t,x)| \leq m_U(t), \quad (t,x) \in U. \tag{2.23}$$

Theorem 2.3.3 (Existence and Uniqueness for ODE). Suppose D is an open set in $\mathbb{R}^{n+1}$. Assume that the function $f : D \to \mathbb{R}^n$ satisfies the Carathéodory conditions on D (see Definition 2.3.2). Additionally, for each compact set $U \subset D$, there exists an integrable function $k_U(t)$ such that

$$|f(t,x) - f(t,y)| \leq k_U(t)|x - y|, \quad (t,x),(t,y) \in U. \tag{2.24}$$

Then, for any $(t_0, x_0) \in U$, there exists a *unique* solution $x(t, t_0, x_0)$ of the initial value problem

$$\dot{x} = f(t,x), \quad x(t_0) = x_0, \tag{2.25}$$

passing through (t_0, x_0). The domain E of definition of $x(t, t_0, x_0)$ in $\mathbb{R}^{n+2}$ is open, and $x(t, t_0, x_0)$ is continuous in E.

Existence of optimal solution for optimal control problems

The following is Theorem 23.10 in [19].

Theorem 2.3.4 (**Existence of optimal solution for optimal control problem).**

$$\begin{cases} \text{Minimize} \quad J(x,u) = \ell(x(a), x(b)) \\ \text{subject to} \quad x'(t) = f(t, x(t), u(t)) \quad \text{a.e.} \\ u(t) \in U(t) \quad \text{a.e.} \\ (t, x(t)) \in Q \quad \forall t \in [a,b], \quad (x(a), x(b)) \in E. \end{cases} \tag{OC1}$$

Let the data of (OC1) satisfy the following hypotheses:

(a) $f(t,x,u)$ is continuous in (x,u) and measurable in t;

(b) $U(\cdot)$ is measurable and compact-valued;

(c) f has linear growth on Q: there is a summable function M such that

$$(t,x) \in Q, \quad u \in U(t) \implies |f(t,x,u)| \le M(t)(1 + |x|);$$

(d) For each $(t,x) \in Q$, the set $f(t,x,U(t))$ is convex;

(e) The sets Q and E are closed, and $\ell : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is lower semicontinuous;

(f) The following set is bounded:

$$\{\alpha \in \mathbb{R}^n : (\alpha, \beta) \in E \text{ for some } \beta \in \mathbb{R}^n\}.$$

Then, if there is at least one admissible process for the problem, it admits a solution.

Other results

We now present Filippov Selection Theorem (see Theorem 2.3.13 in [66]).

Theorem 2.3.5 (**Filippov Selection Theorem).** Let $T > 0$. Consider a nonempty multifunction $\mathcal{X} : [0,T] \rightsquigarrow \mathbb{R}^s$, a function $H : [0,T] \times \mathbb{R}^s \to \mathbb{R}^d$, and a function $v(\cdot) : [0,T] \to \mathbb{R}^d$ satisfying

(i) The set Gr $\mathcal{X}$ is $\mathcal{L} \times \mathcal{B}^s$ measurable;

(*ii*) The function H is $\mathcal{L} \times \mathcal{B}^s$ measurable;

(*iii*) The function $v(\cdot)$ is a measurable function such that $v(t) \in \{H(t,\lambda) : \lambda \in \mathcal{X}(t)\}$ a.e.

Then, there exists a measurable function $\lambda : [0,T] \to \mathbb{R}^s$ such that

$$u(t) \in \mathcal{X}(t) \text{ a.e.}$$

and

$$H(t, \lambda(t)) = v(t) \text{ a.e.}$$

2.4 Functional analysis

We first start by general notations and concepts in functional analysis.

- For $S \subset \mathbb{R}^n$ compact, $\mathcal{C}(S;\mathbb{R}^n)$ denotes to the set of **continuous functions** from S to $\mathbb{R}^n$.
- The class of all functions Lipschitz on S with Lipschitz constant $k \geq 0$ is denoted by $L_k^{\text{ip}}(S)$.
- $\mathcal{C}^{1,1}([a,b];\mathbb{R}^n)$ is the space of continuously differentiable functions f whose derivative $\dot{f}$ is Lipschitz continuous.
- The set of all **absolutely continuous functions** $f \colon [a,b] \longrightarrow \mathbb{R}^n$ is denoted by $AC([a,b];\mathbb{R}^n)$.
  - We say that a function f is absolutely continuous on $[a,b]$ if for every positive number $\varepsilon > 0$, there exists $\delta > 0$, such that whenever a finite sequence of pairwise disjoint sub-intervals (a_i, b_i) of $[a,b]$ with $a_i < b_i$ satisfies $\sum_{i=1}^{N}(b_i - a_i) < \delta$, then
  $$\sum_{i=1}^{N}(f(b_i) - f(a_i)) < \varepsilon.$$
  - Equivalently, we say that f is absolutely continuous if f has a derivative $\dot{f}$ a.e., $\dot{f}$ is Lebesgue integrable, and
  $$f(t) = f(a) + \int_a^t \dot{f}(s)ds, \;\; \forall t \in [a,b].$$
- The set of all functions $f \colon [a,b] \longrightarrow \mathbb{R}^n$ of **bounded variations** is denoted by $BV([a,b];\mathbb{R}^n)$.

- The **total variation** of f is given by

$$V_a^b(f) = \sup_{P \in \mathcal{P}} \sum_{i=0}^{n_P - 1} |f(x_{i+1}) - f(x_i)|,$$

where the supremum is taken over the set

$\mathcal{P} = \{P = \{x_0, \ldots, x_{n_P}\} \mid P$ is a partition of $[a,b]$ satisfying $x_i \leq x_{i+1}$ for $0 \leq i \leq n_P - 1\}$ of all partitions of the interval considered.

- If f is differentiable and its derivative is Riemann-integrable, then its total variation is

$$V_a^b(f) = \int_a^b |f'(x)|\, dx.$$

- For a function f, we say that

$$f \in BV([a,b];\mathbb{R}^n) \quad \Longleftrightarrow \quad V_a^b(f) < +\infty.$$

- The **Lebesgue space of p-integrable functions** $f\colon [a,b] \longrightarrow \mathbb{R}^n$ is denoted by $L^p([a,b];\mathbb{R}^n)$, where the norms in $L^p([a,b];\mathbb{R}^n)$ and $L^\infty([a,b];\mathbb{R}^n)$ (or $\mathcal{C}([a,b];\mathbb{R}^n)$) are written as $\|\cdot\|_p$ and $\|\cdot\|_\infty$, respectively, where for $f \in L^p([a,b];\mathbb{R}^n)$, we have

$$\|f\|_p = \left(\int_a^b |f|^p dx \right)^{\frac{1}{p}},$$

and for $f \in L^\infty([a,b];\mathbb{R}^n)$,

$$\|f\|_\infty = \inf \{ C \geq 0 : |f(x)| \leq C \text{ a.e. } x \in [a,b] \}.$$

- The **Sobolev space** $W^{1,p}([a,b];\mathbb{R}^n)$ denotes the set of continuous functions $f\colon [a,b] \to \mathbb{R}^n$ having $\dot{f} \in L^p([a,b];\mathbb{R}^n)$. More specifically, we have:
  - If $f \in W^{1,1}([a,b];\mathbb{R}^n)$, then f continuous and $\dot{f} \in L^1([a,b];\mathbb{R}^n)$. Hence, $W^{1,1}([a,b];\mathbb{R}^n)$ is the set of all absolutely continuous functions from $[a,b]$ to $\mathbb{R}^n$.
  - If $f \in W^{1,2}([a,b];\mathbb{R}^n)$, then f continuous and $\dot{f} \in L^2([a,b])$. The norm on $W^{1,2}([a,b];\mathbb{R}^n)$ is

$$\|f(\cdot)\|_{W^{1,2}} := \|f(\cdot)\|_\infty + \|\dot{f}(\cdot)\|_2.$$

- Denote $\mathfrak{M}(S)$, $\mathfrak{M}_+(S)$, and $\mathfrak{M}^1_+(S)$ to be, respectively, the set of Radon, positive Radon, and probability measures on S. Note that by Radon measure on S, we mean a finite regular measure on $\mathfrak{B}(S)$, the σ-algebra generated by the Borel subsets of S.
- The space $\mathcal{C}^*([a,b];\mathbb{R}^n)$ denotes the **dual of** $\mathcal{C}([a,b];\mathbb{R}^n)$ equipped with the supremum norm. $\mathcal{C}^*([a,b];\mathbb{R}^n)$ consists of all bounded linear functionals from $\mathcal{C}([a,b];\mathbb{R}^n)$ to $\mathbb{R}$.
    - We denote by $\|\cdot\|_{\text{T.V.}}$ the induced norm on $\mathcal{C}^*([a,b];\mathbb{R}^n)$.
    - For $\nu \in \mathcal{C}^*([a,b];\mathbb{R}^n)$, its support is denoted by $\operatorname{supp}\{\nu\}$, and it is the smallest closed subset $A \subset [a,b]$ with the property that for all relativey open subsets $B \subset [a,b] \setminus A$, we have $\nu(B) = 0$.
    - By Riesz representation theorem, each element in $\mathcal{C}^*([a,b];\mathbb{R})$ can be interpreted as an element in $\mathfrak{M}([a,b])$, the space of finite signed Radon measures on $[a,b]$ equipped with the weak* topology. In other words, every Λ bounded linear functional on $\mathcal{C}([a,b];\mathbb{R})$ is represented as an integral against a finite signed Radon measure ν:
    $$\Lambda(f) = \int_a^b f(x)d\nu(x),$$
    and
    $$\|\Lambda\| = \|\nu\|_{\text{T.V.}}.$$
    - The set of elements in $\mathcal{C}^*([a,b];\mathbb{R})$ taking non-negative values on nonnegative-valued functions in $\mathcal{C}([a,b];\mathbb{R})$ is denoted by $\mathcal{C}^{\oplus}(a,b)$.
    - For $\nu \in \mathcal{C}^{\oplus}(a,b)$, $\|\nu\|_{\text{T.V.}}$, as defined above, coincides with the total variation of ν, i.e.
    $$\|\nu\|_{\text{T.V.}} = \int_{[a,b]} \nu(ds).$$

Important results

We start by Gronwall's Lemma, see [66, Lemma 2.4.4].

Lemma 2.4.1. Take an absolutely continuous function $z : [S,T] \to \mathbb{R}^n$. Assume that there exist nonnegative integrable functions k and v such that

$$\left|\frac{d}{dt}z(t)\right| \leq k(t)|z(t)| + v(t) \quad \text{a.e. } t \in [S,T].$$

Then

$$|z(t)| \leq \exp\left(\int_S^t k(\sigma)d\sigma\right)\left[|z(S)| + \int_S^t \exp\left(-\int_S^\tau k(\sigma)d\sigma\right)v(\tau)d\tau\right]$$

for all $t \in [S,T]$.

This lemma can be found in see [69, equation (3.1)].

Lemma 2.4.2. If the function $W(\cdot,\cdot)$ is lipschitz and $x(\cdot)$ is an absolutely continuous arc, then $W(\cdot, x(\cdot))$ is absolutely continuous, and we have

$$\frac{d}{dt}W(t,x(t)) \in \partial W(t,x(t)).(1,\dot{x}(t)) \text{ a.e.}$$

The following can be found in [43, Theorem 1], and it basically says that a function that is Lipschitz on $S \subset E$ could be extended to the whole space E by preserving a Lipschitz condition.

Theorem 2.4.3. Let $S \subset E$ non-empty. If $f \in L_k^{\text{ip}}(S)$, then $f_{S,k} \in L_k^{\text{ip}}(E)$ and coincides with f on S, where

$$f_{S,k}(x) = \inf_{u \in S}\{f(u) + k\|x-u\|\} \quad \text{for all } x \in E. \tag{2.26}$$

Lemma 2.4.4. Let $S \subset \mathbb{R}^n$ be a compact set, and $f : S \to \mathbb{R} \cup \{\infty\}$ a lower semicontinuous function, and assume there exists $x_0 \in S$ such that $f(x_0) < \infty$. Then, $\inf_{x\in S}$ exists and is finite.

We now present Arzelà–Ascoli theorem.

Theorem 2.4.5 (**Arzelà–Ascoli theorem**)**.** Let $\{f_k\}$ a sequence of continuous functions on $[0,T]$. If $\{f_k\}$ is uniformly bounded and uniformly equicontinuous, then there exists a subsequence of $\{f_k\}$ (we do not relabel) that converges uniformly to a function f.

Theorem 2.4.6 (Helly theorems). Let $\{f_k\}$ be a sequence of bounded variation on $[a,b]$. Assume there is a constant M such that $V_a^b(f_k) \leq M$ and $\|f_k\|_\infty \leq M$ for all k. Then:

(i) **Helly's first theorem.** There is a subsequence of $\{f_k\}$ which converges pointwise everywhere to a function f of bounded variation, with $V_a^b(f) \leq M$ and $\|f\|_\infty \leq M$.

(ii) **Helly's second theorem.** We have:

$$\int_a^b g \, df_k \to \int_a^b g \, df \quad \text{for all } g \in \mathcal{C}([a,b]).$$

Strong convergence, weak convergence, weak* convergence

A significant portion of this section is adapted from Evans lecture notes [37], with further reference to his textbook [38].

Definition 2.4.7. Let $p \in [1,\infty]$. We say that a sequence $\{f_k\}$ **converges strongly** to f in L^p if

$$\|f_k - f\|_p \to 0, \text{ as } k \to \infty.$$

Definition 2.4.8 (When $p \in [1,\infty)$). Let U an open, bounded, smooth subset of $\mathbb{R}^n$, with $n \geq 2$. We assume that $1 \leq p < \infty$, and let q be the conjugate exponent, i.e. $\frac{1}{p} + \frac{1}{q} = 1$, ($q := \infty$ when $p = 1$.) A sequence $\{f_k\} \in L^p(U)$ **converges weakly** to $f \in L^p(U)$, in which case, we write

$$f_k \rightharpoonup f \text{ in } L^p(U),$$

if

$$\int_U f_k g dx \to \int_U f g dx, \quad \forall g \in L^q(U).$$

Definition 2.4.9 (When $p = \infty$). Let U an open, bounded, smooth subset of $\mathbb{R}^n$, with $n \geq 2$. A sequence $\{f_k\} \in L^\infty(U)$ **converges weakly*** to $f \in L^\infty(U)$, in which case, we write

$$f_k \overset{*}{\rightharpoonup} f \text{ in } L^\infty(U),$$

if

$$\int_U f_k g dx \to \int_U f g dx, \quad \forall g \in L^1(U).$$

Theorem 2.4.10 (**Boundedness of weakly converging sequence**)**.** Suppose $1 \leq p < \infty$ and $f_k \rightharpoonup f \quad \text{in } L^p(\Omega) \quad (\overset{*}{\rightharpoonup}$ in $L^\infty(\Omega)$ if $p = \infty$). Then, f_k is bounded in $L^p(\Omega)$ and

$$\|f\|_{L^p(\Omega)} \leq \liminf_{k\to\infty} \|f_k\|_{L^p(\Omega)}.$$

Theorem 2.4.11 (**Weak convergence in L^p**)**.** Suppose $1 < p < \infty$ and the sequence $\{f_k\}_{k\geq 1}$ is bounded in $L^p(U)$. Then there is a subsequence, still denoted by $\{f_k\}_{k\geq 1}$, and a function $f \in L^p(U)$ such that

$$f_k \rightharpoonup f \quad \text{in } L^p(U).$$

If $p = \infty$, the result still holds with $\rightharpoonup$ replaced by $\overset{*}{\rightharpoonup}$.

Theorem 2.4.12. Let $\{f_k\}$ be a sequence of functions that converges pointwise to f and is uniformly bounded in L^∞, i.e., there exists $M > 0$ such that $\|f_k\|_{L^\infty} \leq M$ for all k. Suppose that $\{A_k\}$ is a sequence in L^2 that converges weakly to A in L^2, i.e.,

$$A_k \rightharpoonup A \quad \text{in } L^2.$$

Then, the sequence $\{A_k f_k\}$ converges weakly to Af in L^2, i.e.,

$$A_k f_k \rightharpoonup Af \quad \text{in } L^2.$$

We now prove the following theorem.

Theorem 2.4.13. Let $\{f_k\}_k$ sequence of functions in $W^{1,2}([0,T];\mathbb{R}^n)$ (respectively $W^{1,\infty}([0,T];\mathbb{R}^n)$) such that

$$\|f_k\|_\infty \leq M \quad \text{and} \quad \|\dot{f}_k\|_2 \leq M \text{ (respectively } \|\dot{f}_k\|_\infty \leq M).$$

Then, along a subsequence (we do not relabel), we deduce that there exists a function $f \in W^{1,2}([0,T];\mathbb{R}^n)$ (respectively $f \in W^{1,\infty}([0,T];\mathbb{R}^n)$) such that

$$f_k(\cdot) \xrightarrow{unif} f(\cdot) \quad \text{and} \quad \dot{f}_k(\cdot) \rightharpoonup \dot{f}(\cdot) \text{ weakly in } L^2 \text{ (respectively} \dot{f}_k(\cdot) \overset{*}{\rightharpoonup} \dot{f}(\cdot) \text{ weakly* in } L^\infty).$$

Proof. Let $\varepsilon > 0$ and let $0 \leq t_1, t_2 \leq T$ such that $t_2 - t_1 \leq \frac{\varepsilon^2}{M^2}$ (respectively $\leq \frac{\varepsilon}{M}$). Hence, for every k, we have that

$$\begin{aligned} ||f_k(t_2) - f_k(t_1)|| &= ||\int_{t_1}^{t_2} \dot{f}_k(s)ds|| \\ &\leq \int_{t_1}^{t_2} ||\dot{f}_k(s)||ds \\ &\leq \sqrt{t_2 - t_1}\left(\int_{t_1}^{t_2} ||\dot{f}_k(s)||ds\right)^{\frac{1}{2}} \left(\text{respectively } (t_2 - t_1)||\dot{f}_k||_\infty\right) \\ &\leq \sqrt{t_2 - t_1}\, M \,(\text{respectively } (t_2 - t_1)M) \\ &\leq \varepsilon. \end{aligned}$$

This shows that $(f_k(\cdot))_k$ is equicontinuous. In addition, we have that $(f_k(\cdot))_k$ is uniformly bounded. We deduce from Arzela-Ascoli theorem (Theorem 2.4.5) that along a subsequence of $f_k(\cdot)$ (we do not relabel), we have f_k converges uniformly to an absolutely continuous function f.

Since $(\dot{f}_k(\cdot))_k$ is uniformly bounded in L^2 (respectively in L^∞), we conclude that we can extract a subsequence where f_k converges weakly in L^2 to a limit g (respectively weakly* in L^∞) (see Theorem 2.4.11). Now, for such subsequence (we do not relabel), we have

$$f_k(t) = f_k(0) + \int_0^t \dot{f}_k(s)ds \xrightarrow[k\to\infty]{} f(0) + \int_0^t g(s)ds,$$

we deduce that

$$f(t) = f(0) + \int_0^t g(s)ds,$$

that is, $f(\cdot)$ is absolutely continuous and $\dot{f}(t) = g(t)$ a.e. $t \in [0, T]$. □

The following is Theorem [66, Proposition 9.2.1].

Theorem 2.4.14 (**Convergence of measures**)**.** Take a weak* convergent sequence $\{\mu_i\}$ in $C^\oplus(S, T)$, a sequence of Borel measurable functions $\gamma_i : [S, T] \to \mathbb{R}^n$, and a sequence of closed sets $\{A_i\}$ in $[S, T] \times \mathbb{R}^n$. Take also a closed set A in $[S, T] \times \mathbb{R}^n$, and a measure $\mu \in C^\oplus(S, T)$.

Assume that $A(t)$ is convex for each $t \in \text{dom}\, A(\cdot)$ and that the sets A and $A_1, A_2, \ldots$ are uniformly bounded. Assume further that

$$\limsup_{i \to \infty} A_i \subset A,$$

$$\gamma_i(t) \in A_i(t) \quad \mu_i \text{ a.e. for } i = 1, 2, \ldots$$

and

$$\mu_i \xrightarrow{*} \mu_0 \quad \text{weakly*}.$$

Define $\eta_i \in C^*([S,T]; \mathbb{R}^k)$ by

$$\eta_i(dt) := \gamma_i(t)\mu_i(dt).$$

Then, along a subsequence,

$$\eta_i \xrightarrow{*} \eta_0 \quad \text{weakly*},$$

for some $\eta_0 \in C^*([S,T]; \mathbb{R}^k)$ such that

$$\eta_0(dt) = \gamma_0(t)\mu_0(t),$$

in which γ_0 is a Borel measurable function that satisfies

$$\gamma_0(t) \in A(t) \quad \mu_0 \text{ a.e.}$$

The following is Theorem 6.39 in [19].

Theorem 2.4.15 (Weak-closure theorem). Let $[a,b]$ be an interval in $\mathbb{R}$ and Q a closed subset of $[a,b] \times \mathbb{R}^n$. Let $\Gamma(t,u)$ be a multifunction mapping Q to the closed convex subsets of $\mathbb{R}^n$. We assume that

(a) For each $t \in [a,b]$, the set

$$G(t) = \{(u,z) : (t,u,z) \in Q \times \mathbb{R}^n, z \in \Gamma(t,u)\}$$

is closed and nonempty;

(b) For every measurable function u on $[a, b]$ satisfying $(t, u(t)) \in Q$ a.e. and every $p \in \mathbb{R}^n$, the support function map

$$t \to H_{\Gamma(t,u(t))}(p) = \sup\{\langle p, v\rangle : v \in \Gamma(t, u(t))\}$$

is measurable;

(c) For a summable function k, we have $\Gamma(t, u) \subset B(0, k(t)) \quad \forall (t, u) \in Q$.

Let u_i be a sequence of measurable functions on $[a, b]$ having $(t, u_i(t)) \in Q$ a.e. and converging almost everywhere to u_*, and let $z_i : [a, b] \to \mathbb{R}^n$ be a sequence of functions satisfying $|z_i(t)| \leq k(t)$ a.e. whose components converge weakly in $L^1(a, b)$ to those of z_*. Suppose that, for certain measurable subsets Ω_i of $[a, b]$ satisfying $\lim_{i\to\infty} \operatorname{meas} \Omega_i = b - a$, we have

$$z_i(t) \in \Gamma(t, u_i(t)) + B(0, r_i(t)), \quad t \in \Omega_i \text{ a.e.},$$

where r_i is a sequence of nonnegative functions converging in $L^1(a, b)$ to 0. Then we have in the limit

$$z_*(t) \in \Gamma(t, u_*(t)), \quad t \in [a, b] \text{ a.e.}$$

CHAPTER 3

STUDY OF A COUPLED SWEEPING PROCESS DYNAMIC (D)

In this chapter, we study the following dynamic (D) given by a sweeping process coupled with a differential equation:

$$(D)\begin{cases}\dot{x}(t) \in f(t,x(t),y(t),u(t)) - N_{C(t)}(x(t)), \text{ a.e. } t \in [0,T],\\ \\ \dot{y}(t) = g(t,x(t),y(t),u(t)), \text{ a.e. } t \in [0,T],\end{cases} \tag{3.1}$$

where $T > 0$ is fixed, $f : [0,T] \times \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$, $g : [0,T] \times \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \longrightarrow \mathbb{R}^l$, $C(t)$ is the intersection of the zero-sublevel sets of a finite sequence of functions $\psi_i(t,\cdot)$ where $\psi_i : [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1,\ldots,r$, $N_{C(t)}$ is the *Clarke* normal cone to $C(t)$,
$U(\cdot) : [0,T] \rightsquigarrow \mathbb{R}^m$ is nonempty, closed, and Lebesgue- measurable set-valued map, and the set of control functions $\mathcal{U}$ is defined by

$$\mathcal{U} := \{u : [0,T] \longrightarrow \mathbb{R}^m : u \text{ is measurable and } u(t) \in U(t), \text{ a.e. } t \in [0,T]\}. \tag{3.2}$$

We first introduce the following assumptions on $C(\cdot)$ and $U(\cdot)$ which will be used at different points of the chapter.

(**A1**) **Assumption on $U(\cdot)$:** The measurable set-valued map $U(\cdot)$ has compact images.

(**A2**) **Assumption on $C(\cdot)$:** For $t \in [0,T]$, the set $C(t)$ is *nonempty*, closed, *uniformly* ρ-prox-regular, for some $\rho > 0$, and is given by

$$C(t) := \bigcap_{i=1}^{r} C_i(t), \quad \text{where } C_i(t) := \{x \in \mathbb{R}^n : \psi_i(t,x) \le 0\} \subset \mathbb{R}^n, \tag{3.3}$$

where $(\psi_i)_{1\le i\le r}$ is a family of *continuous* functions $\psi_i : [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}$.

We shall use the following notations. For $x(\cdot) \in \mathcal{C}([0,T];\mathbb{R}^n)$ such that $x(t) \in C(t)$ $\forall t \in [0,T]$, and for $(\tau,z) \in \operatorname{Gr} C(\cdot)$, we define

$$I_i^-(x) := \{t \in [0,T] : x(t) \in \operatorname{int} C_i(t)\} \text{ and } I_i^0(x) := [0,T] \setminus I_i^-(x), \quad \forall i = 1,\ldots,r, \tag{3.4}$$

$$I^-(x) := \bigcap_{i=1}^{r} I_i^-(x) = \{t \in [0,T] : x(t) \in \operatorname{int} C(t)\}, \tag{3.5}$$

$$I^0(x) := \{t \in [0,T] : x(t) \in \operatorname{bdry} C(t)\} = [0,T] \setminus I^-(x) = \{t : \mathcal{I}^0_{(t,x(t))} \neq \emptyset\}, \tag{3.6}$$

$$\text{where } \quad \mathcal{I}^0_{(\tau,z)} := \{i \in \{1,\ldots,r\} : \psi_i(\tau,z) = 0\}. \tag{3.7}$$

We now introduce some local assumptions on $C(\cdot)$, f and g.

For a given pair $(\bar{x},\bar{y})\in\mathcal{C}([0,T];\mathbb{R}^n\times\mathbb{R}^l)$ such that $\bar{x}(t)\in C(t)$ $\forall t\in[0,T]$, and for a constant $\bar{\delta}>0$, we say that the following assumptions hold true at $((\bar{x},\bar{y});\bar{\delta})$ if the corresponding conditions hold true.

(A3) Local assumptions on the functions ψ_i at $(\bar{x};\bar{\delta})$:

(A3.1) There exist $\rho_o>0$ and $L_\psi>0$ such that, for each i, $\nabla_x\psi_i(\cdot,\cdot)$ exists on $\mathrm{Gr}\left(C(\cdot)\cap\bar{B}_{\bar{\delta}}(\bar{x}(\cdot))\right)+\{0\}\times\rho_o B$, and $\psi_i(\cdot,\cdot)$ and $\nabla_x\psi_i(\cdot,\cdot)$ satisfy,

for all (t_1,x_1), $(t_2,x_2)\in\mathrm{Gr}\left(C(\cdot)\cap\bar{B}_{\bar{\delta}}(\bar{x}(\cdot))\right)+\{0\}\times\frac{\rho_o}{2}\bar{B}$,

$$\max\left\{|\psi_i(t_1,x_1)-\psi_i(t_2,x_2)|,\|\nabla_x\psi_i(t_1,x_1)-\nabla_x\psi_i(t_2,x_2)\|\right\}\le L_\psi(\,|t_1-t_2|+\|x_1-x_2\|).$$

(A3.2) For every $t\in I^0(\bar{x})$, the following constraint qualification at $\bar{x}(t)$ holds:

$$\left[\sum_{i\in\mathcal{I}^0_{(t,\bar{x}(t))}}\lambda_i\nabla_x\psi_i(t,\bar{x}(t))=0,\ \text{with each }\lambda_i\ge 0\right]\Longrightarrow\left[\lambda_i=0,\ \forall i\in\mathcal{I}^0_{(t,\bar{x}(t))}\right].$$

For the given $\bar{\delta}$ and for any $a,b>0$, we introduce the following sets

$$\mathbb{C}_{\bar{x}}:=\bigcup_{t\in[0,T]}\left[C(t)\cap\bar{B}_{\bar{\delta}}(\bar{x}(t))\right],\quad \mathbb{B}_{\bar{y}}:=\bigcup_{t\in[0,T]}\bar{B}_{\bar{\delta}}(\bar{y}(t)),\qquad \mathbb{U}:=\bigcup_{t\in[0,T]}U(t),\tag{3.8}$$

$$\bar{\mathscr{N}}_{(a,b)}(t):=\left[C(t)\cap\bar{B}_a(\bar{x}(t))\right]\times\bar{B}_b(\bar{y}(t)),\quad\text{for }\ t\in[0,T].\tag{3.9}$$

(A4) Local assumptions on $h(t,x,y,u):=(f,g)(t,x,y,u)$ at $((\bar{x},\bar{y});\bar{\delta})$:

(A4.1) For $(x,y,u)\in\mathbb{C}_{\bar{x}}\times\mathbb{B}_{\bar{y}}\times\mathbb{U}$, $h(\cdot,x,y,u)$ is Lebesgue-measurable and, for a.e. $t\in[0,T]$, $h(t,\cdot,\cdot,\cdot)$ is continuous on $\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(t)\times U(t)$. There exist $M_h>0$, and $L_h\in L^2([0,T];\mathbb{R}^+)$, such that, for a.e. $t\in[0,T]$, for all (x,y), $(x',y')\in\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(t)$ and $u\in U(t)$,

$$\|h(t,x,y,u)\|\le M_h\ \text{ and }\ \|h(t,x,y,u)-h(t,x',y',u)\|\le L_h(t)\|(x,y)-(x',y')\|.$$

(A4.2) The set $h(t,x,y,U(t))$ is convex for all $(x,y)\in\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(t)$ and $t\in[0,T]$ a.e. [1]

[1]This condition is not needed for Theorem 4.2.11.

3.1 Study of the dynamic (D) under local assumptions

We start by presenting some properties pertaining to the sweeping set $C(t)$ and the sweeping process (D). For the reader's convenience, Table 3.1 at the end of this subsection summarizes all the results presented here.

The following lemma provides an equivalent condition to (A3.2) which allows to obtain the formula for the normal cone to $C(t)$ at points x in $C(t)$ near $\bar{x}(t)$ (Lemma 3.1.3).

Lemma 3.1.1 (Assumption (A3.2)). Let $C(\cdot)$ satisfying $(A2)$ for $\rho > 0$. Consider $\bar{x} \in \mathcal{C}([0,T];\mathbb{R}^n)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$, and $\bar{\delta} > 0$ such that $(A3.1)$ holds at $(\bar{x};\bar{\delta})$. Then, the validity of assumption $(A3.2)$ at $\bar{x}$ is equivalent to the existence of $0 < \varepsilon_o < \bar{\delta}$ and $\eta_o > 0$ such that

$$\left\| \sum_{i \in \mathcal{I}^0_{(t,c)}} \lambda_i \nabla_x \psi_i(t,c) \right\| > 2\eta_o, \quad \forall (t,c) \in \left\{ (\tau,x) \in \mathrm{Gr}\, \left(C(\cdot) \cap \bar{B}_{\varepsilon_o}(\bar{x}(\cdot)\right) : \mathcal{I}^0_{(\tau,x)} \neq \emptyset \right\}, \quad (3.10)$$

where $\mathcal{I}^0_{(\tau,x)}$ is defined in (3.7) and $(\lambda_i)_{i \in \mathcal{I}^0_{(t,c)}}$ is any sequence of nonnegative numbers satisfying $\sum_{i \in \mathcal{I}^0_{(t,c)}} \lambda_i = 1$.

Proof. It suffices to show that (A3.2) implies (3.10). If not, then there exist sequences $t_n \in [0,T]$, $c_n \in C(t_n) \cap \bar{B}_{\frac{1}{n}}(\bar{x}(t_n))$ with $\mathcal{I}^0_{(t_n,c_n)} \neq \emptyset$, and $(\lambda_i^n)_{i \in \mathcal{I}^0_{(t_n,c_n)}}$ with $\sum_{i \in \mathcal{I}^0_{(t_n,c_n)}} \lambda_i^n = 1$ and $\lambda_i^n \geq 0$, for all $i \in \mathcal{I}^0_{(t_n,c_n)}$ and $n \in \mathbb{N}$, such that $\left\| \sum_{i \in \mathcal{I}^0_{(t_n,c_n)}} \lambda_i^n \nabla_x \psi_i(t_n,c_n) \right\| \leq \frac{2}{n}, \forall n \in \mathbb{N}$. As up to a subsequence, $(t_n,c_n) \to (t_o,c_o) := (t_o,\bar{x}(t_o))$, Lemma .0.1 yields the existence of $\emptyset \neq \mathcal{J}_o \subset \{1,\dots,r\}$ and a subsequence of $(t_n,c_n)_n$ we do not relabel, such that $\mathcal{I}^0_{(t_n,c_n)} = \mathcal{J}_o \subset \mathcal{I}^0_{(t_o,c_o)}$ for all $n \in \mathbb{N}$. It follows that

$$\left\| \sum_{i \in \mathcal{J}_o} \lambda_i^n \nabla_x \psi_i(t_n,c_n) \right\| \leq \frac{2}{n}, \quad \sum_{i \in \mathcal{J}_o} \lambda_i^n = 1 \ (\forall n \in \mathbb{N}), \quad \text{and,} \quad \lambda_i^n \geq 0 \ (\forall i \in \mathcal{J}_o, \forall\, n \in \mathbb{N}). \quad (3.11)$$

Hence, after going to a subsequence if necessary, it follows that for all $i \in \mathcal{J}_o$, $\lambda_i^n \to \lambda_i^o \geq 0$ with $\sum_{i \in \mathcal{J}_o} \lambda_i^o = 1$. Upon taking the limit as $n \to \infty$ in (3.11) and by defining $\lambda_i^0 = 0$ for all $i \in \mathcal{I}^0_{(t_o,c_o)} \setminus \mathcal{J}_o$, (A3.1) implies $\sum_{i \in \mathcal{I}^0_{(t_o,c_o)}} \lambda_i^o \nabla_x \psi_i(t_o,c_o) = 0$, which contradicts (A3.2). $\square$

Remark 3.1.2. We can prove that (A3.1) and equation (3.10) imply that for all $(t,x) \in \text{Gr}\left(C(\cdot) \cap \bar{B}_{\varepsilon_o}(\bar{x}(\cdot))\right)$ such that $\mathcal{I}^0_{(t,x)} \neq \emptyset$, the family of vectors $\{\nabla_x \psi_i(t,x)\}_{i \in \mathcal{I}^0_{(t,x)}}$ is positively linearly independent. Indeed, assume there exist $(t,x) \in \text{Gr}\left(C(\cdot) \cap \bar{B}_{\varepsilon_o}(\bar{x}(\cdot))\right)$ such that $\mathcal{I}^0_{(t,x)} \neq \emptyset$, $(\lambda_i)_{i \in \mathcal{I}^0_{(t,x)}} \geq 0$ such that $\sum_{i \in \mathcal{I}^0_{(t,x)}} \lambda_i \nabla_x \psi_i(t,x) = 0$. If there exists i such that $\lambda_i \neq 0$, then $\sum_{i \in \mathcal{I}^0_{(t,x)}} \lambda_i \neq 0$, and we have $\sum_{i \in \mathcal{I}^0_{(t,x)}} \frac{\lambda_i}{\sum_{i \in \mathcal{I}^0_{(t,x)}} \lambda_i} \nabla_x \psi_i(t,x) = 0$. This contradicts (3.10).

Lemma 3.1.3. Let $C(\cdot)$ satisfying $(A2)$ for $\rho > 0$. Consider $\bar{x} \in \mathcal{C}([0,T];\mathbb{R}^n)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$, and $\bar{\delta} > 0$ such that $(A3.1)$ and $(A3.2)$ hold at $(\bar{x};\bar{\delta})$. Let ε_o be the constant from Lemma 3.1.1. Then, we have

$$N_{C(t)}(x) = N^P_{C(t)}(x) = N^L_{C(t)}(x), \quad \forall x \in C(t),$$

and, for all $(t,x) \in \text{Gr}\left(C(\cdot) \cap \bar{B}_{\varepsilon_o}(\bar{x}(\cdot))\right)$,

$$N_{C(t)}(x) = \begin{cases} \left\{ \sum_{i \in \mathcal{I}^0_{(t,x)}} \lambda_i \nabla_x \psi_i(t,x) : \lambda_i \geq 0 \right\} \neq \{0\} & \text{if } x \in \text{bdry } C(t) \\ \{0\} & \text{if } x \in \text{int } C(t). \end{cases} \tag{3.12}$$

Proof. Notice that $C(t)$ is prox-regular. By applying Proposition 2.2.39(i) ([21, Corollary 4.15]), we conclude that the limiting, Clarke and proximal normal cones are all equal to each other. Now, to prove equation (3.12), we apply Lemma 2.2.11 ([19, Corollary 10.44]) and Remark 3.1.2. □

Lemma 3.1.4 (Equivalence). Let $C(\cdot)$ satisfying $(A2)$ for $\rho > 0$. Consider $(\bar{x},\bar{y}) \in \mathcal{C}([0,T];\mathbb{R}^n \times \mathbb{R}^l)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$, and $\bar{\delta} > 0$ such that $(A3.1)$ holds at $(\bar{x};\bar{\delta})$, and $(A4)$ is satisfied by (f,g) at $((\bar{x},\bar{y});\bar{\delta})$. Let $(x,y) \in W^{1,1}([0,T];\mathbb{R}^{n+l})$ be a pair such that $(x(t),y(t)) \in \bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(t)$ $\forall t \in [0,T]$. The following equivalences hold true.

There exists $u \in \mathcal{U}$ such that

$$\begin{cases} \dot{x}(t) \in f(t,x(t),y(t),u(t)) - N_{C(t)}(x(t)) \text{ a.e. } t \in [0,T] \\ \dot{y}(t) = g(t,x(t),y(t),u(t)) \text{ a.e. } t \in [0,T] \end{cases} \tag{3.13}$$

$\overset{(I)}{\Longleftrightarrow}$ There exist $u \in \mathcal{U}$ and $(\lambda_1(\cdot), \cdots, \lambda r(\cdot))$ non-negative measurable functions such that for every $i \in \{1, \cdots, r\}$, $\lambda_i(t) = 0$ for $t \in I_i^{-}(x)$ and

$$\begin{cases} \dot{x}(t) = f(t, x(t), y(t), u(t)) - \sum_{i=1}^{r} \lambda_i(t) \nabla_x \psi_i(t, x(t)) \text{ a.e. } t \in [0, T] \\ \dot{y}(t) = g(t, x(t), y(t), u(t)) \text{ a.e. } t \in [0, T] \end{cases} \tag{3.14}$$

$\overset{(II)}{\Longleftrightarrow}$ There exist $(\lambda_1(\cdot), \cdots, \lambda_r(\cdot))$ non-negative measurable functions such that for every $i \in \{1, \cdots, r\}$, $\lambda_i(t) = 0$ for $t \in I_i^{-}(x)$ and

$$(\dot{x}(t), \dot{y}(t)) \in h(t, x(t), y(t), U(t)) - \left(\sum_{i=1}^{r} \lambda_i(t) \nabla_x \psi_i(t, x(t)), 0\right) \tag{3.15}$$

$\overset{(III)}{\Longleftrightarrow}$ There exist $(\lambda_1(\cdot), \cdots, \lambda_r(\cdot))$ non-negative measurable functions such that for every $i \in \{1, \cdots, r\}$, $\lambda_i(t) = 0$ for $t \in I_i^{-}(x)$ and $\forall z \in \mathbb{R}^n \times \mathbb{R}^l$,

$$\langle z, (\dot{x}(t), \dot{y}(t))\rangle \leq \sigma(z, h(t, x(t), y(t), U(t))) - \langle z, (\sum_{i=1}^{r} \lambda_i(t) \nabla_x \psi_i(t, x(t)), 0)\rangle \text{ a.e.} \tag{3.16}$$

Proof. Equivalences (I) and (II) hold true by applying Filipov Selection Theorem, see Theorem 2.3.5 ([66, Theorem 2.3.13]), and using equation (3.12) for (I). Whereas equivalence (III) holds true by applying the support property in (2.2) on the compact and convex set $S = h(t, x(t), y(t), U(t))$. □

An important consequence of Lemma 3.1.1 and Lemma 3.1.3 is manifested in the following result that establishes the Lipschitz continuity and the uniqueness of the solutions near $(\bar{x}, \bar{y})$ for the Cauchy problem of (D) via its equivalent form. We note that, under global assumptions, the existence of a solution for the Cauchy problem of (D) is given in Theorem 3.3.7, which will be established in Section 3.3.2. First, define μ to be

$$\mu := L_\psi(1 + M_h). \tag{3.17}$$

Lemma 3.1.5. Let $C(\cdot)$ satisfying $(A2)$ for $\rho > 0$. Consider $(\bar{x}, \bar{y}) \in \mathcal{C}([0, T]; \mathbb{R}^n \times \mathbb{R}^l)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0, T]$, and $\bar{\delta} > 0$ such that $(A3.1)$ and $(A3.2)$ hold at $(\bar{x}; \bar{\delta})$, and $(A4.1)$ is satisfied by (f, g) at $((\bar{x}, \bar{y}); \bar{\delta})$. Let $u \in \mathcal{U}$ and $(x_0, y_0) \in \bar{\mathscr{N}}_{(\varepsilon_0, \bar{\delta})}(0)$ be fixed.

Then, a pair $(x, y) \in W^{1,1}([0,T];\mathbb{R}^{n+l})$, such that $(x(t), y(t)) \in \bar{\mathcal{N}}_{(\varepsilon_0,\bar{\delta})}(t)$ $\forall t \in [0,T]$, is a solution of (D) corresponding to $((x_0, y_0), u)$ if and only if there exist measurable functions $(\lambda_1, \cdots, \lambda_r)$ such that, for all $i = 1, \cdots, r$, $\lambda_i(t) = 0$ for $t \in I_i^{-}(x)$, and $((x, y), u)$ together with $(\lambda_1, \cdots, \lambda_r)$ satisfies

$$\begin{cases} \dot{x}(t) = f(t, x(t), y(t), u(t)) - \sum_{i \in \mathcal{I}^0_{(t,x(t))}} \lambda_i(t) \nabla_x \psi_i(t, x(t)) \;\text{ a.e. } t \in [0,T], \\ \dot{y}(t) = g(t, x(t), y(t), u(t)), \;\text{a.e. } t \in [0,T], \\ (x(0), y(0)) = (x_0, y_0). \end{cases} \tag{3.18}$$

Furthermore, we have the following bounds

$$\begin{cases} \|\lambda_i\|_\infty \leq \|\sum_i^r \lambda_i\|_\infty \leq \frac{\mu}{4\eta_0^2}, \;\; \forall i = 1, \cdots, r, \\ \|\dot{x}\|_\infty \leq M_h + \frac{\mu}{4\eta_0^2} L_\psi, \;\; \|\dot{y}\|_\infty \leq M_h. \end{cases} \tag{3.19}$$

Consequently, (x, y) is the unique solution of (D) in $\bar{\mathcal{N}}_{(\varepsilon_0,\bar{\delta})}(\cdot)$ corresponding to $((x_0, y_0), u)$. In particular, if $((\bar{x}, \bar{y}), \bar{u})$ solves (D), then $(\bar{x}, \bar{y})$ is Lipschitz and is the unique solution of (D) corresponding to $((\bar{x}(0), \bar{y}(0)), \bar{u})$.

Proof. The equivalence in the first part of this lemma follows immediately from Filippov selection theorem and the normal cone formula in (3.12) (see Lemma 3.1.4). Now, we proceed to prove the bounds in (3.19). Since for all $i = 1, \cdots, r$, $\psi_i(\cdot, x(\cdot)) \in W^{1,1}$ $\Big(\psi_i(\cdot,\cdot)$ is lipschitz and $x(\cdot)$ is absolutely continuous$\Big)$, then $\frac{d}{dt}\psi_i(t, x(t))$ exists for almost all $t \in [0,T]$. Using assumption $(A3.1)$ and Lemma 2.4.2 (see [69, equation (3.1)]), we deduce that, $\forall i = 1, \cdots, r$,

$$\frac{d}{dt}\psi_i(t, x(t)) \subset \partial^{(t,x)}\psi_i(t, x(t)).(1, \dot{x}(t)).$$

But,

$$\begin{aligned} \partial^{(t,x)}\psi_i(t, x(t)) &= \operatorname{conv}\left\{\lim \nabla_{(t,x)}\psi_i(t_j, x_j) : (t_j, x_j) \xrightarrow{O} (t, x(t))\right\} \\ &= \operatorname{conv}\left\{\lim(\nabla_t \psi_i(t_j, x_j), \nabla_x \psi_i(t_j, x_j)) : (t_j, x_j) \xrightarrow{O} (t, x(t))\right\} \\ &= \hat{\partial}_t \psi_i(t, x(t)) \times \nabla_x \psi_i(t, x(t)), \end{aligned} \tag{3.20}$$

where O is full- measure subset of a neighborhood of $(t, x(t))$, and for $(t, z) \in \mathrm{Gr}(C(\cdot) \cap \bar{B}_{\bar{\delta}}(\bar{x}(\cdot)))$,

$$\hat{\partial}_t \psi_i(t, z) := \mathrm{conv}\{\lim_{j\to\infty} \nabla_t \psi_i(t_j, z_j) : (t_j, z_j) \to (t, z)\}. \tag{3.21}$$

Hence,

$$\frac{d}{dt}\psi_i(t, x(t)) \subset \partial^{(t,x)}\psi_i(t, x(t)).(1, \dot{x}(t)) = \hat{\partial}_t\psi_i(t, x(t)) + \langle \nabla_x \psi_i(t, x(t)), \dot{x}(t)\rangle, \ t \in [0, T] \text{ a.e.},$$

Thus, there exist measurable $\theta_i(\cdot) \in \hat{\partial}_t\psi_i(\cdot, x(\cdot))$ a.e., such that

$$\frac{d}{dt}\psi_i(t, x(t)) = \theta_i(t) + \langle \nabla_x \psi_i(t, x(t)), \dot{x}(t)\rangle \quad \text{a.e. } t \in [0, T], \ \ \forall i = 1, \cdots, r. \tag{3.22}$$

Note that, by (A3.1), we have, for $t \in [0, T]$ a.e., for all $\theta_i(t) \in \hat{\partial}_t\psi_i(t, x(t))$, and for all $i = 1, \cdots, r$,

$$|\theta_i(t) + \langle \nabla_x \psi_i(t, x(t)), f(t, x(t), y(t), u(t))\rangle| \le L_\psi(1 + \|f(t, x(t), y(t), u(t))\|). \tag{3.23}$$

Define in $[0, T]$ the set of full measure:

$$\mathscr{T} := \{t \in (0, T) : \dot{x}(t) \text{ and } \frac{d}{dt}\psi_i(t, x(t)) \text{ exist}, \ \forall i = 1, \cdots, r\}. \tag{3.24}$$

Let $t \in I^{-}(x) \cap \mathscr{T}$. Then, $\mathcal{I}^0_{(t,x(t))} = \emptyset$, and hence, $\forall i = 1, \cdots, r, \ \lambda_i(t) = 0$. This implies that $\dot{x}(t) = f(t, x(t), y(t), u(t))$, and hence $\|\dot{x}(t)\| \le M_h$.

Let $t \in I^0(x) \cap \mathscr{T}$ with $\sum_{i\in\mathcal{I}^0_{(t,x(t))}} \lambda_i(t) \neq 0$; otherwise we join the conclusion of the previous case. Since for all $i \in \mathcal{I}^0_{(t,x(t))}$, we have $\psi_i(t, x(t)) = 0$ and $x(s) \in C(s)$ $\forall s \in [0, T]$, it follows that $\frac{d}{dt}\psi_i(t, x(t)) = 0$, for all $i \in \mathcal{I}^0_{(t,x(t))}$. Hence, for the finite sequence $(\theta_i)_{i=1}^r$ in (3.22), we have

$$0 = \theta_i(t) + \langle \nabla_x \psi_i(t, x(t)), \dot{x}(t)\rangle. \tag{3.25}$$

Multiplying (3.25) by $\lambda_i(t)$, and using the fact that $x(\cdot)$ satisfies the first equation of (3.18), we get that

$$0 = \lambda_i(t)\theta_i(t) + \lambda_i(t)\left\langle \nabla_x\psi_i(t, x(t)), f(t, x(t), y(t), u(t)) - \sum_{j\in\mathcal{I}^0_{(t,x(t))}} \lambda_j(t)\nabla_x\psi_j(t, x(t))\right\rangle \tag{3.26}$$

Summing (3.26) over all $i \in \mathcal{I}^0_{(t,x(t))}$ and using (3.23), we deduce that

$$\begin{aligned}\|\sum_{i\in\mathcal{I}^0_{(t,x(t))}} \lambda_i(t)\nabla_x\psi_i(t,x(t))\|^2 &= \sum_{i\in\mathcal{I}^0_{(t,x(t))}} \lambda_i(t)\left(\theta_i(t) + \langle \nabla_x\psi_i(t,x(t)), f(t,x(t),y(t),u(t))\rangle\right)\\ &\leq L_\psi(1+\|f(t,x(t),y(t),u(t))\|)\sum_{i\in\mathcal{I}^0_{(t,x(t))}} \lambda_i(t).\end{aligned}$$

Hence, utilizing (3.10) on the term on the left hand side, and then dividing by $\sum_{i\in\mathcal{I}^0_{(t,x(t))}} \lambda_i(t) \neq 0$ the last inequality, we deduce from (3.17) that

$$\sum_{i\in\mathcal{I}^0_{(t,x(t))}} \lambda_i(t) \leq \frac{L_\psi}{4\eta_0^2}(1+\|f(t,x(t),y(t),u(t))\|) \overset{(A4.1)}{\leq} \frac{\mu}{4\eta_0^2}. \tag{3.27}$$

Therefore, $\|\sum_{i=1}^r \lambda_i\|_\infty \leq \frac{\mu}{4\eta_0^2}$. Finally, employing (A4.1) for f and g, along with (3.18), the bounds on $\|\dot{x}\|_\infty$ and $\|\dot{y}\|_\infty$ follow.

For the uniqueness, let $X := (x,y)$, $\tilde{X} := (\tilde{x},\tilde{y})$ in $\bar{\mathscr{N}}_{(\varepsilon_0,\bar{\delta})}(\cdot)$ be two solutions of (D) corresponding to $((x_0,y_0),u)$, and let $(\lambda_i)_{i=1}^r, (\tilde{\lambda}_i)_{i=1}^r$ be their corresponding multipliers satisfying (3.18). Using the hypomonoticity of the normal cone to the ρ-prox-regular sets $C(t)$ (see Proposition 2.2.39(iii)), the L_h-Lipschitz property of $h(t,\cdot,\cdot,u(t))$, and the bounds in (3.19) for the multipliers, we deduce that

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}(\|X(t)-\tilde{X}(t)\|^2) = \langle \dot{X}(t)-\dot{\tilde{X}}(t), X(t)-\tilde{X}(t)\rangle\\ \leq\ &(L_h(t) + \frac{\mu}{4\rho\eta_o^2}L_\psi)\|X(t)-\tilde{X}(t)\|^2 := \kappa(t)\|X(t)-\tilde{X}(t)\|^2. \end{aligned}\tag{3.28}$$

Hence using Gronwall's lemma (see Lemma 2.4.1), we deduce that

$$\|X(t)-\tilde{X}(t)\|^2 \leq e^{2\int_0^t \kappa(s)ds}\|X(0)-\tilde{X}(0)\|^2 = 0.$$

Then, $X(t) = \tilde{X}(t) \quad \forall t \in [0,T]$, and the uniqueness is proved. □

Now, we arrive at the table promised earlier, summarizing all the results from this subsection.

Table 3.1 Summary of results from Subsection 3.1 .

Result	Description
Lemma 3.1.1	We provide an equivalent condition to (A3.2) that allows to obtain the formula for the normal cone to $C(t)$ at points x in $C(t)$ near $\bar{x}(t)$.
Remark 3.1.2	We prove that for all $(t,x) \in \operatorname{Gr}\left(C(\cdot) \cap \bar{B}_{\varepsilon_o}(\bar{x}(\cdot)\right)$ such that $\mathcal{I}^0_{(t,x)} \neq \emptyset$, the family of vectors $\{\nabla_x \psi_i(t,x)\}_{i \in \mathcal{I}^0_{(t,x)}}$ is positively linearly independent.
Lemma 3.1.3	We use Lemma 3.1.1 to obtain the formula for the normal cone to $C(t)$ at points x in $C(t)$ near $\bar{x}(t)$.
Lemma 3.1.4	We prove an equivalence between the system (D) and three other systems of equations.
Lemma 3.1.5	We use Lemma 3.1.1 and Lemma 3.1.3 to establish the Lipschitz continuity and the uniqueness of the solutions near $(\bar{x}, \bar{y})$ for the Cauchy problem of (D) via its equivalent form.

3.2 Development and study of a new truncated dynamic $(\bar{D})$ under local assumptions

3.2.1 Preliminary results

To avoid imposing the *boundedness* of $\operatorname{Gr} C(\cdot)$ and a *global* constraint qualification on the sweeping sets $C(t)$ of (D), we shall truncate $C(t)$ by a ball around $\bar{x}(t)$ of a specific radius $\bar{\varepsilon}$ (that will be determined in Remark 3.2.2), so that the uniform prox-regularity of $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ is ensured, its *constraint qualification* is satisfied, and its normal cone explicit formula is valid (see Remark 3.2.2 and Lemmas 3.2.4-3.2.6). After establishing certain properties of the truncated sweeping set $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$, we now turn our focus to the associated truncated dynamic $(\bar{D})$. Our goal is to derive analogous results to those presented in Section 3.1, but now in the context of the truncated sweeping set $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ and the truncated dynamic $(\bar{D})$. See Table 3.2 for summary of the results.

A key element to proving the uniform prox-regularity of the truncated sweeping set $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ is the following lemma, which uses Lemma 3.1.1 to prove the closed graph property of $N_{C(\cdot)}(\cdot)$ in the domain where (3.12) is valid.

Lemma 3.2.1. Let $C(\cdot)$ satisfying $(A2)$ for $\rho > 0$. Consider $\bar{x} \in \mathcal{C}([0,T];\mathbb{R}^n)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$, and $\bar{\delta} > 0$ such that $(A3.1)$ and $(A3.2)$ hold at $(\bar{x};\bar{\delta})$. Then, for ε_o obtained in Lemma 3.1.1, the set-valued map $(t,y) \to N_{C(t)}(y)$ has closed graph on the set $\operatorname{Gr}\Big(C(\cdot) \cap \bar{B}_{\varepsilon_o}(\bar{x}(\cdot))\Big)$.

Proof. Let $v_n \in N_{C(t_n)}(y_n)$ such that $v_n \to v_o$ and $(t_n, y_n) \to (t_o, y_o)$ in $\operatorname{Gr}\Big(C(\cdot) \cap \bar{B}_{\varepsilon_o}(\bar{x}(\cdot))\Big)$. We shall prove that $v_o \in N_{C(t_o)}(y_o)$. If $v_o = 0$ then obviously $v_o \in N_{C(t_o)}(y_o)$. Now, let $v_o \neq 0$, then for n large enough, $v_n \neq 0$, and hence, equation (3.12) implies that $y_n \in \operatorname{bdry} C(t_n)$ and $v_n = \sum_{i\in\mathcal{I}^0_{(t_n,y_n)}} \lambda_i^n \nabla_x \psi_i(t_n, y_n)$ for some $(\lambda_i^n)_i \geq 0$. By Lemma .0.1, we deduce the existence of $\emptyset \neq \mathcal{J}_o \subset \{1, \ldots, r\}$ and a subsequence of $(t_n, y_n)_n$ we do not relabel, such that we have $\mathcal{I}^0_{(t_n,y_n)} = \mathcal{J}_o \subset \mathcal{I}^0_{(t_o,y_o)}$ for all $n \in \mathbb{N}$. Hence, for n large enough, $v_n = \sum_{i\in\mathcal{J}_o} \lambda_i^n \nabla_x \psi_i(t_n, y_n)$ and $\sum_{i\in\mathcal{J}_o} \lambda_i^n > 0$ (since $v_n \neq 0$). Define, for each $i \in \mathcal{J}_o$, the bounded sequence $(\beta_i^n)_n$, where $\beta_i^n := \frac{\lambda_i^n}{\sum_{j\in\mathcal{J}_o} \lambda_j^n} \geq 0$. Since also $\sum_{i\in\mathcal{J}_o} \beta_i^n = 1$ for all n, then for each $i \in \mathcal{J}_o$, along a subsequence (we do not relabel), $\beta_i^n \to \beta_i \geq 0$ with $\sum_{i\in\mathcal{J}_o} \beta_i = 1$. Using $(A3.1)$ and Lemma 3.1.1, we have $0 \neq \sum_{i\in\mathcal{J}_o} \beta_i^n \nabla_x \psi_i(t_n, y_n) \to \sum_{i\in\mathcal{J}_o} \beta_i \nabla_x \psi_i(t_o, y_o) \neq 0$. By writing

$$v_n = \Big(\sum_{j\in\mathcal{J}_o} \lambda_j^n\Big)\Big(\sum_{i\in\mathcal{J}_o} \beta_i^n \nabla_x \psi_i(t_n, y_n)\Big),$$

and using the fact that $v_n \to v_o \neq 0$, we deduce that $\sum_{j\in\mathcal{J}_o} \lambda_j^n$ is convergent to a limit $\beta_o > 0$. Hence, $v_o = \sum_{i\in\mathcal{J}_o} \beta_o \beta_i \nabla_x \psi_i(t_o, y_o)$. Now, define

$$\alpha_i := \begin{cases} \beta_o \beta_i & \text{if } i \in \mathcal{J}_o \\ 0 & \text{if } i \in \mathcal{I}^0_{(t_o,y_o)} \setminus \mathcal{J}_o. \end{cases}$$

Then, $v_o = \sum_{i\in\mathcal{I}^0_{(t_o,y_o)}} \alpha_i \nabla_x \psi_i(t_o, y_o) \in N_{C(t_o)}(y_o)$. □

Combining Lemma 3.2.1 with Lemma 2.2.40 immediately produces a range for $\bar{\varepsilon} > 0$ ensuring the uniform prox-regularity of the truncated sets $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$.

Remark 3.2.2. Let $C(\cdot)$ satisfying $(A2)$ for $\rho > 0$. Consider $\bar{x} \in \mathcal{C}([0,T];\mathbb{R}^n)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$, and $\bar{\delta} > 0$ such that $(A3.1)$ and $(A3.2)$ hold at $(\bar{x};\bar{\delta})$. Then, for $\bar{\varepsilon} \in (0,\rho) \cap (0,\varepsilon_o]$, where ε_o is given in Lemma 3.1.1, there exists $\rho_{\bar{\varepsilon}} > 0$, obtained from Lemma 2.2.40, such that for all $t \in [0,T]$, $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ is $\rho_{\bar{\varepsilon}}$-prox-regular.

Introducing the new truncated sweeping process $(\bar{D})$

Now, our attention shifts from the dynamic (D) to working on the dynamic $(\bar{D})$ obtained from (D) by replacing the sweeping set $C(t)$ by the truncated sweeping set $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$, where $\bar{\varepsilon} \in (0,\rho) \cap (0,\varepsilon_o]$, and by adding $-N_{\bar{B}_{\bar{\delta}}(\bar{y}(t))}$ to the right hand side of the differential equation, which becomes a differential inclusion as a result. Denote by $(\bar{D})$ the aforementioned *truncated* system obtained from (D) by localizing $C(\cdot)$ around $\bar{x}$ and $\mathbb{R}^l$ around $\bar{y}$, that is,

$$(\bar{D}) \begin{cases} \dot{x}(t) \in f(t,x(t),y(t),u(t)) - N_{C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x(t)), \text{ a.e. } t \in [0,T], \\ \dot{y}(t) \in g(t,x(t),y(t),u(t)) - N_{\bar{B}_{\bar{\delta}}(\bar{y}(t))}(y(t)), \text{ a.e. } t \in [0,T]. \end{cases} \tag{3.29}$$

Notice that the truncated sweeping set for x, $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$, is the sub-level set of $\psi_1(t,\cdot), \cdots, \psi_r(t,\cdot)$, and $\psi_{r+1}(t,\cdot)$, where ψ_{r+1} is given by

$$\psi_{r+1}(t,x) = \psi_{r+1}(t,x;\bar{x},\bar{\varepsilon}) := \frac{1}{2}[\|x - \bar{x}(t)\|^2 - \bar{\varepsilon}^2]. \tag{3.30}$$

Therefore, for $C_{r+1}(t) := \bar{B}_{\bar{\varepsilon}}(\bar{x}(t)) = \{x \in \mathbb{R}^n : \psi_{r+1}(t,x) \le 0\}$,

$$C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t)) = C(t) \cap C_{r+1}(t) = \bigcap_{i=1}^{r+1} \{x \in \mathbb{R}^n : \psi_i(t,x) \le 0\},$$

and hence, it is always generated by at least two functions. On the other hand, the truncated sweeping set for y, $\bar{B}_{\bar{\delta}}(\bar{y}(t))$, is generated by a single function $\varphi : [0,T] \times \mathbb{R}^l \longrightarrow \mathbb{R}$, where

$$\varphi(t,y) = \varphi(t,y;\bar{y},\bar{\delta}) := \frac{1}{2}[\|y - \bar{y}(t)\|^2 - \bar{\delta}^2], \tag{3.31}$$

$$\text{i.e.} \quad \bar{B}_{\bar{\delta}}(\bar{y}(t)) = \{y \in \mathbb{R}^l : \varphi(t,y) \le 0\}. \tag{3.32}$$

The following remark shows the relation between pairs that are admissible for (D) and those admissible for $(\bar{D})$.

Remark 3.2.3. We have:

- Any admissible pair $((x, y), u)$ for (D) such that $(x(t), y(t)) \in \bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)$ for all $t \in [0, T]$, is also admissible for $(\bar{D})$. This is due to Lemma 2.2.9.
- On the other hand, any admissible pair $((x, y), u)$ for $(\bar{D})$ such that $(x(t), y(t)) \in \bar{\mathcal{N}}_{(\delta_1,\delta_2)}(t)$ with $\delta_1 < \bar{\varepsilon}$ and $\delta_2 < \bar{\delta}$ is also admissible for (D). This is due to the fact that if $\forall t \in [0, T]$, $(x(t), y(t)) \in B_{\bar{\varepsilon}}(\bar{x}(t)) \times B_{\bar{\delta}}(\bar{y}(t))$, then, using the local property of the proximal normal cone, we have $N^P_{C(t)}(x(t)) = N^P_{C(t)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x(t))$ and $\{0\} = N^P_{\bar{B}_{\bar{\delta}}(\bar{y}(t))}(y(t))$.
- In particular, $((\bar{x}, \bar{y}), \bar{u})$ solves (D) if and only if it solves $(\bar{D})$.

For $x(\cdot) \in \mathcal{C}([0, T]; \mathbb{R}^n)$ such that $x(t) \in C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ $\forall t \in [0, T]$, and $(\tau, z) \in \mathrm{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))\right)$, we define the following sets obtained through adding to those in (3.4)-(3.7) the extra constraint produced by ψ_{r+1}:

$$\begin{aligned}
&I^-_{r+1}(x) := \{t \in [0, T] : x(t) \in B_{\bar{\varepsilon}}(\bar{x}(t))\} \;\text{ and }\; I^0_{r+1}(x) := [0, T] \setminus I^-_{r+1}(x),\\
&\bar{I}^-(x) := \bigcap_{i=1}^{r+1} I^-_i(x) = \{t \in [0, T] : x(t) \in \operatorname{int} C(t) \cap B_{\bar{\varepsilon}}(\bar{x}(t))\}\\
&\qquad\;\; = I^-(x) \cap \{t \in [0, T] : x(t) \in B_{\bar{\varepsilon}}(\bar{x}(t))\},\\
&\bar{I}^0(x) = [0, T] \setminus \bar{I}^-(x) = I^0(x) \cup \{t \in [0, T] : \|x(t) - \bar{x}(t)\| = \bar{\varepsilon}\} = \{t \in [0, T] : \bar{\mathcal{I}}^0_{(t,x(t))} \neq \emptyset\},\\
&\text{where } \bar{\mathcal{I}}^0_{(\tau,z)} := \{i \in \{1, \ldots, r, r + 1\} : \psi_i(\tau, z) = 0\}.
\end{aligned} \tag{3.33}$$

Since $\bar{x}(t) \in B_{\bar{\varepsilon}}(\bar{x}(t))$, then $\psi_{r+1}(t, \bar{x}(t)) < 0$ and hence, $\bar{I}^0(\bar{x}) = I^0(\bar{x})$ and, for $t \in \bar{I}^0(\bar{x})$, $\bar{\mathcal{I}}^0_{(t,\bar{x}(t))} = \mathcal{I}^0_{(t,\bar{x}(t))}$.

The following lemma provides a second condition, (3.34), equivalent to (A3.2) which, unlike (3.10), validates the formula for the normal cone to the uniform prox-regular truncated sets $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$, obtained in Remark 3.2.2, (see Lemma 3.2.6 stated below). Note that since ψ_{r+1}, given by (3.30), is a function of $\bar{\varepsilon}$, this lemma is of a different nature than Lemma 3.1.1. Observe that, for any given $\bar{\varepsilon} > 0$, we have that $\psi_{r+1}(t, x)$ and $\nabla_x \psi_{r+1}(t, x) := x - \bar{x}(t)$ exist and continuous everywhere.

Lemma 3.2.4 (**Assumption (A3.2)**). Let $C(\cdot)$ satisfying $(A2)$ for $\rho > 0$. Consider $\bar{x} \in \mathcal{C}([0,T];\mathbb{R}^n)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$, and $\bar{\delta} > 0$ such that $(A3.1)$ holds at $(\bar{x};\bar{\delta})$. Then, $(A3.2)$ is satisfied at $\bar{x}$ if and only if for $\bar{\varepsilon} \in (0,\rho) \cap (0,\varepsilon_o]$ and its corresponding ψ_{r+1} given by (3.30), there exists $\bar{\eta} \in (0,\eta_0)$ (without loss of generality $\bar{\eta} \leq \frac{\bar{\varepsilon}}{2}$) such that

$$\left\| \sum_{i\in\bar{\mathcal{I}}^0_{(t,c)}} \lambda_i \nabla_x \psi_i(t,c) \right\| > 2\bar{\eta}, \quad \forall (t,c) \in \{(\tau,x) \in \operatorname{Gr}\left(C(\cdot)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))\right) : \bar{\mathcal{I}}^0_{(\tau,x)} \neq \emptyset\}, \tag{3.34}$$

where $(\lambda_i)_{i\in\bar{\mathcal{I}}^0_{(t,c)}}$ is any sequence of nonnegative numbers satisfying $\sum_{i\in\bar{\mathcal{I}}^0_{(t,c)}} \lambda_i = 1$, and $\bar{\mathcal{I}}^0_{(\tau,x)}$ is given by (3.33).

Proof. We only need to show that $(A3.2)$ yields (3.34). For this, assume $(A3.2)$ is valid and let $\bar{\varepsilon} \in (0,\rho)\cap(0,\varepsilon_o]$. From Lemma 3.1.1, it follows that, for any $\bar{\eta} \in (0,\eta_o)$, (3.34) holds for all $(t,c) \in \{(\tau,x) \in \operatorname{Gr}\left(C(\cdot)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))\right) : \bar{\mathcal{I}}^0_{(\tau,x)} \neq \emptyset\}$ such that $(r+1) \notin \bar{\mathcal{I}}^0_{(t,c)}$. It remains to prove that (3.34) is valid for all (t,c) such that $(r+1)$ is necessarily in $\bar{\mathcal{I}}^0_{(t,c)}$, that is, when $\bar{\mathcal{I}}^0_{(t,c)} = \mathcal{I}^0_{(t,c)} \cup \{r+1\}$ and $\lambda_{r+1} \neq 0$. Arguing by contradiction, then there exist sequences $t_n \in [0,T]$, $c_n \in C(t_n)$ with $\|c_n - \bar{x}(t_n)\| = \bar{\varepsilon}$, and $(\lambda_i^n)_{i\in\bar{\mathcal{I}}^0_{(t_n,c_n)}}$ with $\lambda_i^n \geq 0$, for all $i \in \mathcal{I}^0_{(t_n,c_n)}$, $\lambda^n_{r+1} > 0$, and

$$\Big(\sum_{i\in\mathcal{I}^0_{(t_n,c_n)}} \lambda_i^n\Big) + \lambda^n_{r+1} = 1, \tag{3.35}$$

such that $\left\|\sum_{i\in\mathcal{I}^0_{(t_n,c_n)}} \lambda_i^n \nabla_x \psi_i(t_n,c_n) + \lambda^n_{r+1}(c_n - \bar{x}(t_n))\right\| \leq \frac{2}{n}, \quad \forall n \in \mathbb{N}$. Using the compactness of $[0,T]$, (A3.1), and the continuity of $\bar{x}$, it follows that up to subsequences, $t_n \to t_o \in [0,T]$ and $c_n \to c_o \in C(t_o)$ with $\|c_o - \bar{x}(t_o)\| = \bar{\varepsilon}$. Note that $\mathcal{I}^0_{(t_n,c_n)} \neq \emptyset$, since otherwise, (3.35) yields $\lambda^n_{r+1} = 1$, and in this case the above inequality becomes $\|c_n - \bar{x}(t_n)\| \leq \frac{2}{n}$, which is invalid for n large. Thus, by Lemma .0.1, for some $\emptyset \neq \mathcal{J}_o \subset \{1,\ldots,r\}$, $\mathcal{I}^0_{(t_n,c_n)} = \mathcal{J}_o \subset \mathcal{I}^0_{(t_o,c_o)}$, for n large. This implies that, for n large enough,

$$\left\|\sum_{i\in\mathcal{J}_o} \lambda_i^n \nabla_x \psi_i(t_n,c_n) + \lambda^n_{r+1}(c_n - \bar{x}(t_n))\right\| \leq \frac{2}{n}, \tag{3.36}$$

$$\sum_{i\in\mathcal{J}_o} \lambda_i^n + \lambda^n_{r+1} = 1, \quad \lambda^n_{r+1} > 0, \quad \text{and } \lambda_i^n \geq 0 \ \ \forall i \in \mathcal{J}_o.$$

Hence, up to a subsequence, $\lambda_i^n \to \lambda_i^o \geq 0$ for all $i \in \mathcal{J}_o$, and $\lambda_{r+1}^n \to \lambda_{r+1}^o \geq 0$. Upon taking the limit as $n \to \infty$ in (3.36), (A3.1) yields that

$$\sum_{i\in\mathcal{J}_o} \lambda_i^o \nabla_x \psi_i(t_o, c_o) + \lambda_{r+1}^o(c_o - \bar{x}(t_o)) = 0, \quad \sum_{i\in\mathcal{J}_o} \lambda_i^o + \lambda_{r+1}^o = 1, \;\; \lambda_i^o \geq 0 \;\; \forall i \in \mathcal{J}_o \cup \{r+1\}. \tag{3.37}$$

From (3.37) and Lemma 3.1.1 we get that $\lambda_{r+1}^o > 0$. As $\|c_o - \bar{x}(t_o)\| = \bar{\varepsilon}$, (3.37) is translated to saying

$$0 \neq v := \sum_{i\in\mathcal{J}_o} \lambda_i^o \nabla_x \psi_i(t_o, c_o) = -\lambda_{r+1}^o(c_o - \bar{x}(t_o)),$$

and hence, per (3.12), $0 \neq v \in N^P_{C(t_o)}(c_o) \cap -N^P_{\bar{B}_{\bar{\varepsilon}}(\bar{x}(t_o))}(c_o)$. As $\bar{\varepsilon} \in (0, \rho)$, then, this inclusion contradicts Lemma 2.2.40. $\square$

Remark 3.2.5. We can prove that (A3.1) and equation (3.34) imply that for all $(t, x) \in \mathrm{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot)\right)$ such that $\bar{\mathcal{I}}^0_{(t,x)} \neq \emptyset$, the family of vectors $\{\nabla_x \psi_i(t, x)\}_{i\in\bar{\mathcal{I}}^0_{(t,x)}}$ is positively linearly independent.

Important consequences of Lemma 3.2.4 are the following explicit formulae for the normal cone to the truncated sets $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ and for their prox-regularity constant, which shall replace $\rho_{\bar{\varepsilon}}$. Assume without loss of generality that $L_\psi \geq \frac{4\bar{\eta}}{\rho_o}$, where ρ_o is the constant from (A3.1).

Lemma 3.2.6. Let $C(\cdot)$ satisfying $(A2)$ for some $\rho > 0$. Consider $\bar{x} \in \mathcal{C}([0, T]; \mathbb{R}^n)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0, T]$, and $\bar{\delta} > 0$ such that $(A3.1)$ and $(A3.2)$ hold at $(\bar{x}; \bar{\delta})$. Let $\bar{\varepsilon} \in (0, \rho) \cap (0, \varepsilon_o]$ with its corresponding ψ_{r+1}, given by (3.30), and $\bar{\eta}$ from Lemma 3.2.4. Let $\rho_{\bar{\varepsilon}}$ be the uniform prox-regular constant of $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ obtained from Remark 3.2.2. For all $(t, x) \in \mathrm{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot)\right)$,

$$N_{C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x) = N^P_{C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x) = N^L_{C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x),$$

and

$$
N_{C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x) = \begin{cases} \{\sum_{i\in\bar{\mathcal{I}}^0_{(t,x)}} \lambda_i \nabla_x \psi_i(t,x) : \lambda_i \geq 0\} \neq \{0\} & \text{if } x \in \text{bdry}(C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))) \\ \{0\} & \text{if } x \in \text{int}(C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))). \end{cases} \tag{3.38}
$$

Furthermore, $C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ is uniformly $\frac{2\bar{\eta}}{L_\psi}$-prox-regular, $C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ is epi-lipschitzian at every $x \in C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$, and

$$
\text{cl}\left(\text{int}\left(C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))\right)\right) = C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t)). \tag{3.39}
$$

Proof. Since $C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ is prox-regular, we apply Proposition 2.2.39(i) ([21, Corollary 4.15]), and we conclude that the limiting, Clarke and proximal normal cones are all equal to each other. To prove equation (3.38), we apply Lemma 2.2.11 ([19, Corollary 10.44]) and Remark 3.2.5. Now, we prove that $C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ is uniformly $\frac{2\bar{\eta}}{L_\psi}$-prox-regular using Theorem 2.2.41 (see [2, Theorem 9.1]). Indeed, in Theorem 2.2.41, take $m := r+1$, $g_i := \psi_i$, $S(t) := C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$. Notice that $\nabla_x\psi_{r+1}(t,x) = x - \bar{x}(t)$ and condition ($A3.1$) is satisfied, hence conditions (i)-(ii) of Theorem 2.2.41 are satisfied for $\rho := \frac{\rho_o}{2}$, and $\gamma := L_\psi$. Finally, Lemma 3.2.4 implies that the last condition of Theorem 2.2.41 is satisfied by translating [58, Lemma 6.1] to our setting. As a result, for all $t \in [0,T]$, we have $C(t)$ is prox-regular with constant $\min\left\{\frac{\rho_o}{2}, \frac{2\bar{\eta}}{L_\psi}\right\} = \frac{2\bar{\eta}}{L_\psi}$ (since $L_\psi \geq \frac{4\bar{\eta}}{\rho_o}$). To prove that $C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ is epi-lipschitzian for every $x \in C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ and that (3.39) is satisfied, we use Lemma 2.2.46, and equations (3.34)-(3.38). □

We now prove an equivalence between the system $(\bar{D})$ and three other systems.

Lemma 3.2.7 (**Equivalence**). Consider $C(\cdot)$ satisfying (A2) for $\rho > 0$. Consider $(\bar{x},\bar{y}) \in \mathcal{C}([0,T];\mathbb{R}^n\times\mathbb{R}^l)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$. Let $\bar{\delta} > 0$ such that ($A3.1$) and ($A3.2$) hold at $(\bar{x};\bar{\delta})$ and ($A4$) is satisfied by (f,g) at $((\bar{x},\bar{y});\bar{\delta})$, and let $\bar{\varepsilon} \in (0,\rho)\cap(0,\varepsilon_o]$ with its corresponding ψ_{r+1} given by (3.30). Let $(x,y) \in W^{1,1}([0,T];\mathbb{R}^{n+l})$ be a pair such that $(x(t),y(t)) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)$ $\forall t \in [0,T]$. The following equivalences hold true.

There exists $u \in \mathcal{U}$ such that

$$\begin{cases} \dot{x}(t) \in f(t,x(t),y(t),u(t)) - N_{C(t)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x(t)), \text{ a.e. } t \in [0,T], \\ \dot{y}(t) \in g(t,x(t),y(t),u(t)) - N_{\bar{B}_{\bar{\delta}}(\bar{y}(t))}(y(t)), \text{ a.e. } t \in [0,T] \end{cases} \tag{3.40}$$

$\overset{(I)}{\Longleftrightarrow}$ There exist $u \in \mathcal{U}$ and there exist measurable functions $(\lambda_1, \cdots, \lambda_{r+1})$ and ζ such that, $\forall i = 1, \cdots, r+1$, $\lambda_i(t) = 0$ $\forall t \in I_i^-(x)$, $\zeta(t)\varphi(t,y(t)) = 0$ $\forall t \in [0,T]$, and $((x,y),u)$, $(\lambda_i)_{i=1}^{r+1}$, and ζ satisfy

$$\begin{cases} \dot{x}(t) = f(t,x(t),y(t),u(t)) - \sum_{i=1}^{r+1} \lambda_i(t)\nabla_x\psi_i(t,x(t)) \text{ a.e. } t \in [0,T], \\ \dot{y}(t) = g(t,x(t),y(t),u(t)) - \zeta(t)\nabla_y\varphi(t,y(t)), \text{ a.e. } t \in [0,T]. \end{cases} \tag{3.41}$$

$\overset{(II)}{\Longleftrightarrow}$ There exist measurable functions $(\lambda_1, \cdots, \lambda_{r+1})$ and ζ such that, $\forall i = 1, \cdots, r+1$, $\lambda_i(t) = 0$ $\forall t \in I_i^-(x)$, $\zeta(t)\varphi(t,y(t)) = 0$ $\forall t \in [0,T]$, and

$$(\dot{x}(t),\dot{y}(t)) \in h(t,x(t),y(t),U(t)) - \left(\sum_{i=1}^{r+1} \lambda_i(t)\nabla_x\psi_i(t,x(t)), \zeta(t)\nabla_y\varphi(t,y(t))\right) \tag{3.42}$$

$\overset{(III)}{\Longleftrightarrow}$ There exist measurable functions $(\lambda_1, \cdots, \lambda_{r+1})$ and ζ such that, $\forall i = 1, \cdots, r+1$, $\lambda_i(t) = 0$ $\forall t \in I_i^-(x)$, $\zeta(t)\varphi(t,y(t)) = 0$ $\forall t \in [0,T]$, $\forall z \in \mathbb{R}^n \times \mathbb{R}^l$,

$$\langle z, (\dot{x}(t),\dot{y}(t))\rangle \le \sigma(z, h(t,x(t),y(t),U(t))) - \left\langle z, \left(\sum_{i=1}^{r+1} \lambda_i(t)\nabla_x\psi_i(t,x(t)), \zeta(t)\nabla_y\varphi(t,y(t))\right)\right\rangle \text{ a.e.} \tag{3.43}$$

Parallel to Lemma 3.1.5, and based on Lemma 3.2.4 and Lemma 3.2.6, we shall obtain here the Lipschitz continuity and the uniqueness of the solutions of the Cauchy problem corresponding to the *truncated* system $(\bar{D})$, defined in (3.29). We note that the existence of a solution for this more general Cauchy problem is obtained in Corollary 3.2.16.

Lemma 3.2.8. Consider $C(\cdot)$ satisfying (A2) for $\rho > 0$. Consider $(\bar{x},\bar{y}) \in \mathcal{C}([0,T];\mathbb{R}^n \times \mathbb{R}^l)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$ and $(\bar{x}(\cdot),\bar{y}(\cdot))$ is $L_{(\bar{x},\bar{y})}$-Lipschitz on $[0,T]$ for some constant $L_{(\bar{x},\bar{y})} \ge 1$. Let $\bar{\delta} > 0$ such that $(A3.1)$ and $(A3.2)$ hold at $(\bar{x};\bar{\delta})$ and $(A4.1)$ is satisfied by (f,g) at $((\bar{x},\bar{y});\bar{\delta})$, and let $\bar{\varepsilon}\in(0,\rho)\cap(0,\varepsilon_o]$ with its corresponding ψ_{r+1} given by (3.30). Fix $u\in\mathcal{U}$ as well as $(x_0,y_0)\in\mathscr{N}_{(\bar{\varepsilon},\bar{\delta})}(0)$. Then, a pair $(x,y)\in W^{1,1}([0,T];\mathbb{R}^{n+l})$, such that $(x(t),y(t))\in\bar{\mathscr{N}}_{(\varepsilon_0,\bar{\delta})}(t)$ $\forall t\in[0,T]$, solves the system $(\bar{D})$ associated with $((x_0,y_0),u)$ if and only if there exist measurable functions $(\lambda_1,\cdots,\lambda_{r+1})$ and ζ such that, $\forall i=1,\cdots,r+1$, $\lambda_i(t)=0$ $\forall t\in I_i^{\text{-}}(x)$, $\zeta(t)\varphi(t,y(t))=0$ $\forall t\in[0,T]$, and $((x,y),u)$, $(\lambda_i)_{i=1}^{r+1}$, and ζ satisfy

$$\begin{cases}\dot{x}(t)=f(t,x(t),y(t),u(t))-\sum_{i\in\bar{\mathcal{I}}^0_{(t,x(t))}}\lambda_i(t)\nabla_x\psi_i(t,x(t)) \text{ a.e. } t\in[0,T],\\ \dot{y}(t)=g(t,x(t),y(t),u(t))-\zeta(t)\nabla_y\varphi(t,y(t)), \text{ a.e. } t\in[0,T],\\ (x(0),y(0))=(x_0,y_0).\end{cases}\tag{3.44}$$

Furthermore, we have the following bounds

$$\max\{\|\textstyle\sum_{i=1}^{r+1}\lambda_i\|_\infty,\ \|\zeta\|_\infty\}\le\frac{\bar{\mu}}{4\bar{\eta}^2},\quad \max\{\|\dot{x}\|_\infty,\|\dot{y}\|_\infty\}\le M_h+\frac{\bar{\mu}}{4\bar{\eta}^2}\bar{L},\tag{3.45}$$

$$\max\{\|\dot{x}(t)-f(t,x(t),y(t),u(t))\|,\|\dot{y}(t)-g(t,x(t),y(t),u(t))\|\}\le\frac{\bar{\mu}}{4\bar{\eta}^2}\bar{L},\quad t\in[0,T] \text{ a.e.}\tag{3.46}$$

where

$$\bar{L}:=\max\{L_\psi,\bar{\delta}L_{(\bar{x},\bar{y})}\}\ge\bar{\delta},\quad \bar{\mu}:=\bar{L}(1+M_h)\ge\mu.\tag{3.47}$$

Consequently, (x,y) is the unique solution of (3.29) corresponding to $((x_0,y_0),u)$.

Proof. The equivalence follows from Filippov Selection theorem, the normal cone formula in (3.38), and the fact that $N_{\bar{B}_{\bar{\delta}}(\bar{y}(t))}(y)$ equals $\{0\}$ if $\varphi(t,y)<0$, and equals $\{\lambda(y-\bar{y}(t)):\lambda\ge0\}$ if $\varphi(t,y)=0$ (see Lemma 3.2.7).

For the bounds pertaining $\|\sum_{i=1}^{r+1}\lambda_i\|_\infty$ and $\|\dot{\bar{x}}\|_\infty$, we follow the same steps as in proof of Lemma 3.1.5, with the main difference here is that we add an extra constraint to $C(t)$, namely, $\psi_{r+1}(t,x)\le0$. For this reason, it suffices to show that (3.22) and (3.23), where L_ψ is enlarged to $\bar{L}$, are also valid for $i=r+1$, and that the set $\mathscr{T}$ can be modified to take into account the addition of ψ_{r+1}. Once these goals are achieved, the proof follows from that of Lemma 3.1.5, where $\bar{\mathcal{I}}^0_{(t,x(t))}$, Lemma 3.2.4, $\bar{\eta}$, and $\bar{\mu}$ are used instead of $\mathcal{I}^0_{(t,x(t))}$, Lemma 3.1.1, η_0, and μ, respectively.

Note that, by (3.30), $\nabla_x \psi_{r+1}(t,z) = z - \bar{x}(t)$ exists for all $(t,z) \in [0,T] \times \mathbb{R}^n$. Furthermore, as $\bar{x}$ is $L_{(\bar{x},\bar{y})}$-Lipschitz, we have, on Gr $C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))$, that $\psi_{r+1}(\cdot,\cdot)$ is $\bar{\varepsilon} L_{(\bar{x},\bar{y})}$-Lipschitz, $\nabla_x \psi_{r+1}(\cdot,\cdot)$ is bounded by $\bar{\varepsilon} \le \bar{L}$, and $\hat{\partial}_t \psi_{r+1}(\cdot,\cdot)$, defined via (3.21) for $i = r+1$, satisfies

$$\hat{\partial}_t \psi_{r+1}(t,z) = \langle z - \bar{x}(t), -\partial \bar{x}(t) \rangle = \partial_t \psi_{r+1}(t,z), \;\; \forall (t,z) \in \text{Gr } C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot)), \tag{3.48}$$

and hence, $\forall \theta_{r+1} \in \hat{\partial}_t \psi_{r+1}(t,z), |\theta_{r+1}| \le \bar{\varepsilon} L_{(\bar{x},\bar{y})} \le \bar{L}$. Thus, for $t \in [0,T]$ a.e., and for all $\theta_{r+1}(t) \in \hat{\partial}_t \psi_{r+1}(t,x(t))$, we have

$$|\theta_{r+1}(t) + \langle \nabla_x \psi_{r+1}(t,x(t)), f(t,x(t),y(t),u(t)) \rangle \le \bar{L}(1 + \|f(t,x(t),y(t),u(t))\|). \tag{3.49}$$

Therefore, (3.23) and (3.49) yield that (3.23) holds up to $i = r+1$, that is, for $t \in [0,T]$ a.e., for all $\theta_i(t) \in \hat{\partial} \psi_i(t,x(t))$, we have for $i = 1, \cdots, r+1$

$$|\theta_i(t) + \langle \nabla_x \psi_i(t,x(t)), f(t,x(t),y(t),u(t)) \rangle| \le \bar{L}(1 + \|f(t,x(t),y(t),u(t))\|). \tag{3.50}$$

On the other hand, from (3.30), (3.48), and the fact that $\dot{\bar{x}}(t) \in \partial \bar{x}(t)$ a.e., we have

$$\frac{d}{dt} \psi_{r+1}(t,x(t)) = \langle x(t) - \bar{x}(t), -\dot{\bar{x}}(t) + \dot{x}(t) \rangle, \; t \in [0,T] \text{ a.e.}, \tag{3.51}$$

$$= \theta_{r+1}(t) + \langle \nabla_x \psi_{r+1}(t,x(t)), \dot{x}(t) \rangle, \; t \in [0,T] \text{ a.e.}, \tag{3.52}$$

where $\theta_{r+1}(t) = \langle x(t) - \bar{x}(t), -\dot{\bar{x}}(t) \rangle \in \hat{\partial}_t \psi_{r+1}(t,x(t))$ a.e.

Therefore, (3.22) holds up to $i = r+1$, that is, $\forall i$, there is measurable $\theta_i(\cdot) \in \hat{\partial}_t \psi_i(\cdot, x(\cdot))$ a.e., with

$$\frac{d}{dt} \psi_i(t,x(t)) = \theta_i(t) + \langle \nabla_x \psi_i(t,x(t)), \dot{x}(t) \rangle, \text{ a.e., } \forall i = 1, \cdots, r+1. \tag{3.53}$$

Instead of the set $\mathscr{T}$ given in (3.24) in the proof of Lemma 3.1.5, we use the following modified set $\bar{\mathscr{T}}$ that involves $\dot{\bar{x}}$ and on which $\frac{d}{dt}\psi_{r+1}(\cdot, x(\cdot))$ readily exists,

$$\bar{\mathscr{T}} := \{t \in (0,T) : \dot{x}(t), \dot{\bar{x}}(t), \text{ and } \frac{d}{dt} \psi_i(t,x(t)) \text{ exist}, \; \forall i = 1, \cdots, r\}.$$

Therefore, similarly to (3.27) we obtain

$$\sum_{i \in \bar{\mathcal{I}}^0_{(t,x(t))}} \lambda_i(t) \le \frac{\bar{L}}{4\bar{\eta}^2}(1 + \|f(t,x(t),y(t),u(t))\|) \overset{(A4.1)}{\le} \frac{\bar{\mu}}{4\bar{\eta}^2}, \tag{3.54}$$

implying, via first equation of (3.44) and $(A4.1)$, the required bound in (3.45) for $\|\dot{x}\|_\infty$ and the first bound in (3.46).

For the bounds of ζ and $\dot{y}$ in (3.45), we use the full- measure set $\bar{\mathscr{A}} := \{t \in (0,T) : \dot{\bar{y}}(t)$ and $\dot{y}(t)$ exist$\}$. If $t \in \bar{\mathscr{A}}$ and $\varphi(t, y(t)) < 0$, then $\zeta(t) = 0$ and the bound on $\dot{y}$ follows using $(A4.1)$. If $t \in \bar{\mathscr{A}}$ and $\varphi(t, y(t)) = 0$, then $\|y(t) - \bar{y}(t)\| = \bar{\delta}$ and, since $\varphi(\cdot, y(\cdot)) \leq 0$, $\frac{d}{dt}\varphi(t, y(t)) = 0$. Hence, as (3.31) implies that

$$\frac{d}{dt}\varphi(t, y(t)) = \langle y(t) - \bar{y}(t), -\dot{\bar{y}}(t) + \dot{y}(t)\rangle,\ t \in [0,T] \text{ a.e.,} \tag{3.55}$$

then, using $\bar{\eta} < \frac{\bar{\varepsilon}}{2} < \frac{\bar{\delta}}{2}$ (by Lemma 3.2.4), second equation of (3.44), $L_{(\bar{x},\bar{y})} \geq 1$, and $\bar{\delta}L_{(\bar{x},\bar{y})} \leq \bar{L}$ (by (3.47)), we get that for $t \in [0,T]$ a.e.,

$$4\bar{\eta}^2\zeta(t) \leq \bar{\delta}^2\zeta(t) = \langle y(t) - \bar{y}(t), g(t, x(t), y(t), u(t)) - \dot{\bar{y}}(t)\rangle \leq \bar{L}(1 + \|g(t, x(t), y(t), u(t))\|). \tag{3.56}$$

Therefore, by $(A4.1)$ we have, $\|\zeta\|_\infty \leq \frac{\bar{\mu}}{4\bar{\eta}^2}$, which when combined with the second equation of (3.44), yields the bound on $\|\dot{y}\|_\infty$ in (3.45) and the second bound in (3.46).

The uniqueness proof of (x, y) is similar to that in Lemma 3.1.5, where system (D) is replaced by $(\bar{D})$, the ρ-prox-regularily of $C(t)$ is replaced by the $\frac{2\bar{\eta}}{L_\psi}$-prox-regularity of $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ obtained in Lemma 3.2.6, and (3.18)-(3.19), μ, η_o, and L_ψ, are replaced by (3.44)-(3.45), $\bar{\mu}$, $\bar{\eta}$, and $\bar{L}$, respectively. The ∞-prox-regularity of $\bar{B}_{\bar{\delta}}(\bar{y}(t))$ keeps the inequality in (3.28) valid. □

The following table summarizes the results of this subsection.

Table 3.2 Summary of results from Subsection 3.2.1 .

Result	**Description**
Lemma 3.2.1	We use Lemma 3.1.1 to prove the closed graph property of $N_{C(\cdot)}(\cdot)$ in the domain where (3.12) is valid.

Table 3.2 (cont'd)

Result	**Description**
Remark 3.2.2	We use Lemma 3.2.1 with Lemma 2.2.40 to produce a range for $\bar{\varepsilon} > 0$ ensuring the uniform prox-regularity of the truncated sets $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$.
Remark 3.2.3	We show the relation between pairs that are admissible for (D) and those admissible for $(\bar{D})$.
Lemma 3.2.4	We provide a second condition equivalent to (A3.2) which validates the formula for the normal cone to the uniform prox-regular truncated sets $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$.
Remark 3.2.5	We prove that for $(t, x) \in \text{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot)\right)$ such that $\bar{\mathcal{I}}^0_{(t,x)} \neq \emptyset$, the family of vectors $\{\nabla_x \psi_i(t, x)\}_{i \in \bar{\mathcal{I}}^0_{(t,x)}}$ is positively linearly independent.
Lemma 3.2.6	We use Lemma 3.2.4 to derive explicit formulae for the normal cone to the truncated sets $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ and for their prox-regularity constant.
Lemma 3.2.7	We prove an equivalence between the system $(\bar{D})$ and three other systems of equations.
Lemma 3.2.8	We use Lemma 3.2.4 and Lemma 3.2.6 to obtain the Lipschitz continuity and the uniqueness of the solutions of the Cauchy problem corresponding to the *truncated* system $(\bar{D})$, defined in (3.29). This is parallel to Lemma 3.1.5.

3.2.2 Exponential penalty approximation for the system $(\bar{D})$

This section aims to establish the relationship between $(\bar{D})$ and its approximating standard control system $(\bar{D}_{\gamma_k})$, as well as the existence and uniqueness of Lipschitz solutions to the Cauchy problem associated with $(\bar{D})$. Throughout this whole section, we assume $C(\cdot)$ satisfying (A2) for $\rho > 0$, and $(\bar{x}, \bar{y}) \in \mathcal{C}([0, T]; \mathbb{R}^n \times \mathbb{R}^l)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0, T]$, and $(\bar{x}(\cdot), \bar{y}(\cdot))$ is $L_{(\bar{x},\bar{y})}$-Lipschitz on $[0, T]$ for some $L_{(\bar{x},\bar{y})} \geq 1$. Let $\bar{\delta} > 0$ be such that $(A3.1)$

and $(A3.2)$ hold at $(\bar{x};\bar{\delta})$ and $(A4.1)$ is satisfied by (f,g) at $((\bar{x},\bar{y});\bar{\delta})$. Fix $0<\bar{\varepsilon}<\bar{\delta}$, its corresponding ψ_{r+1} given by (3.30), and $\bar{\eta}\in(0,\frac{\bar{\varepsilon}}{2})$, such that $\bar{\varepsilon}, \psi_{r+1}$, and $\bar{\eta}$ satisfy Lemma 3.2.4. Assuming that $L_\psi \geq \frac{4\bar{\eta}}{\rho_o}$, set $\bar{\rho} := \frac{2\bar{\eta}}{L_\psi}$, the prox-regular constant for the sets $C(t)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ from Lemma 3.2.6.

We start by extending the function $h(t,x,\cdot,u)$ from $\bar{B}_{\bar{\delta}}(\bar{y}(t))$ to $\mathbb{R}^l$ so that this extension satisfies for all $y\in\mathbb{R}^l$, $(A4.1)$, and also (A4.2) whenever it is satisfied by h. This extension shall be later used in Theorem 3.2.14.

Remark 3.2.9 (Extension). For $t\in[0,T]$ a.e., $x\in\left[C(t)\cap\bar{B}_{\bar{\delta}}(\bar{x}(t))\right]$, and for $u\in U(t)$, it is possible to extend the function $h(t,x,\cdot,u):=(f,g)(t,x,\cdot,u)$ so that, whenever h satisfies $(A4)$ (including $(A4.2)$), its extension also satisfies $(A4)$ for all $y\in\mathbb{R}^l$. Indeed, the convexity for all $t\in[0,T]$ of $\bar{B}_{\bar{\delta}}(\bar{y}(t))$ yields that $\pi(t,\cdot):=\pi_{\bar{B}_{\bar{\delta}}(\bar{y}(t))}(\cdot)$ is well-defined and 1-Lipschitz on $\mathbb{R}^l$.

Define for a.e. $t\in[0,T]$, and $(x,y,u)\in\left[C(t)\cap\bar{B}_{\bar{\delta}}(\bar{x}(t))\right]\times\mathbb{R}^l\times U(t)$,

$$\bar{h}(t,x,y,u):=h(t,x,\pi(t,y),u).$$

Whenever h satisfies $(A4)$ at $((\bar{x},\bar{y}),\bar{\delta})$, arguments similar to those in [55, Remark 4.1] show that $\bar{h}$ (whose name we keep as h) also satisfies $(A4)$, where $\bar{\mathcal{N}}_{(\bar{\delta},\bar{\delta})}(t)$, which is $\left[C(t)\cap\bar{B}_{\bar{\delta}}(\bar{x}(t))\right]\times\bar{B}_{\bar{\delta}}(\bar{y}(t))$, is now replaced by $\left[C(t)\cap\bar{B}_{\bar{\delta}}(\bar{x}(t))\right]\times\mathbb{R}^l$.

The following notations, which depend on $(\bar{x};\bar{\varepsilon})$ and $(\bar{y};\bar{\delta})$, will be used in the proofs of the results that follow as well as the proof of Theorem 4.2.11. They are instrumental in constructing a dynamic $(\bar{D}_{\gamma_k})$ that approximates $(\bar{D})$ and has rich properties.

- Let $\bar{L}$ and $\bar{\mu}$ be the constants given in (3.47). Define a sequence $(\gamma_k)_k$ such that, for all $k\in\mathbb{N}$, $\gamma_k > \dfrac{2\bar{\mu}}{\bar{\eta}^2}e \;(>\dfrac{e}{\bar{\delta}})$ and $\gamma_k\to\infty$ as $k\longrightarrow\infty$, and the real sequences $(\bar{\alpha}_k)_k, (\bar{\sigma}_k)_k$, and $(\bar{\rho}_k)_k$ by

$$\bar{\alpha}_k := \frac{1}{\gamma_k}\ln\left(\frac{\bar{\eta}^2\gamma_k}{2\bar{\mu}}\right);\; \bar{\sigma}_k := \frac{(r+1)\bar{L}}{2\bar{\eta}^2}\left(\frac{\ln(r+1)}{\gamma_k}+\bar{\alpha}_k\right);\; \bar{\rho}_k := \sqrt{\bar{\delta}^2-2\bar{\alpha}_{\gamma_k}}. \quad (3.57)$$

Our choice of γ_k with the fact that $\bar{\mu} > \bar{\delta} > \bar{\eta}$ yield that $\bar{\delta}^2 > \dfrac{2\ln\left(\gamma_k\bar{\delta}\right)}{\gamma_k} > 2\bar{\alpha}_{\gamma_k}$, and

$$\gamma_k e^{-\gamma_k\bar{\alpha}_k} = \frac{2\bar{\mu}}{\bar{\eta}^2}, \ \ (\bar{\alpha}_k, \bar{\sigma}_k, \bar{\rho}_k) > 0 \ \ \forall\, k \in \mathbb{N}, \ \ \bar{\alpha}_k \searrow 0, \ \bar{\sigma}_k \searrow 0 \ \text{ and } \ \bar{\rho}_k \nearrow \bar{\delta}. \tag{3.58}$$

- For each $t \in [0, T]$ and $k \in \mathbb{N}$, we define the compact sets

$$\bar{C}^{\gamma_k}(t) := \left\{x \in \mathbb{R}^n : \sum_{i=1}^{r+1} e^{\gamma_k\psi_i(t,x)} \le 1\right\} \subset \operatorname{int} C(t) \cap B_{\bar{\varepsilon}}(\bar{x}(t)), \tag{3.59}$$

$$\bar{C}^{\gamma_k}(t,k) := \left\{x \in \mathbb{R}^n : \sum_{i=1}^{r+1} e^{\gamma_k\psi_i(t,x)} \le \frac{2\bar{\mu}}{\bar{\eta}^2\gamma_k} = e^{-\gamma_k\bar{\alpha}_k}\right\} \subset \operatorname{int} \bar{C}^{\gamma_k}(t). \tag{3.60}$$

- For $u \in \mathcal{U}$, the approximation dynamic $(\bar{D}_{\gamma_k})$ of $(\bar{D})$ is defined by

$$(\bar{D}_{\gamma_k}) \begin{cases} \dot{x}(t) = f(t, x(t), y(t), u(t)) - \sum\limits_{i=1}^{r+1} \gamma_k e^{\gamma_k\psi_i(t,x(t))}\nabla_x\psi_i(t, x(t)), \ \text{ a.e. } t \in [0, T], \\ \dot{y}(t) = g(t, x(t), y(t), u(t)) - \gamma_k e^{\gamma_k\varphi(t,y(t))}\nabla_y\varphi(t, y(t)), \text{ a.e. } t \in [0, T]. \end{cases} \tag{3.61}$$

Using Lemma 3.2.4, a translation of [58, equation (8)], and arguments parallel to those used in the proofs of [58, Propositions 4.4 & 4.6] and [59, Proposition 5.3], it is not difficult to derive the following properties for our sets $\bar{C}^{\gamma_k}(t)$ and $\bar{C}^{\gamma_k}(t,k)$, knowing that the sets $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ are $\frac{2\bar{\eta}}{L_\psi}$- prox-regular. Notice, from (3.59) and (3.60), that these sets here are time-dependent, uniformly localized near $\bar{x}(t)$, and are defined not only via $\psi_1, \cdots, \psi_r$ but also via the extra function ψ_{r+1}. For completeness, we provide the adjusted proofs below.

Proposition 3.2.10. The following holds true.

(i) There exist $k_1 \in \mathbb{N}$ and $r_1 \in (0, \frac{\rho_o}{2}]$, such that $\forall k \ge k_1$, $\forall (t, x) \in \{(t, x) \in [0, T] \times \mathbb{R}^n : \sum_{i=1}^{r+1} e^{\gamma_k\psi_i(t,x)} = 1\}$, and $\forall\, (\tau, z) \in B_{2r_1}(t, x)$, we have

$$\left\|\sum_{i=1}^{r+1} e^{\gamma_k\psi_i(\tau,z)}\nabla_x\psi_i(\tau, z)\right\| > 2\bar{\eta}\sum_{i=1}^{r+1} e^{\gamma_k\psi_i(\tau,z)}. \tag{3.62}$$

(ii) There exists $k_2 \ge k_1$ and $\bar{\epsilon}_o > 0$ such that for all $k \ge k_2$ we have

$$\left[x \in \bar{C}^{\gamma_k}(t) \ \ \& \ \|\sum_{i=1}^{r+1} e^{\gamma_k\psi_i(t,x)}\nabla_x\psi_i(t, x)\| \le \bar{\eta}\sum_{i=1}^{r+1} e^{\gamma_k\psi_i(t,x)}\right] \Longrightarrow \left[\sum_{i=1}^{r+1} e^{\gamma_k\psi_i(t,x)} < e^{-\bar{\epsilon}_o\gamma_k}\right]. \tag{3.63}$$

(*iii*) For all $t \in [0,T]$, for all k, $\bar{C}^{\gamma_k}(t) \subset \operatorname{int}\left(C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))\right)$ and $\bar{C}^{\gamma_k}(t,k) \subset \operatorname{int} \bar{C}^{\gamma_k}(t)$, and these sets are uniformly compact. Moreover, there exists $k_3 \in \mathbb{N}$ such that for $k \geq k_3$, for all $t \in [0,T]$, we have

$$\bar{C}^{\gamma_k}(t) = \operatorname{cl}\left(\operatorname{int} \bar{C}^{\gamma_k}(t)\right),$$
$$\bar{C}^{\gamma_k}(t,k) = \operatorname{cl}\left(\operatorname{int} \bar{C}^{\gamma_k}(t,k)\right),$$

$$\operatorname{bdry} \bar{C}^{\gamma_k}(t) := \left\{x \in \mathbb{R}^n : \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t,x)} = 1\right\} \neq \emptyset,$$
$$\operatorname{int} \bar{C}^{\gamma_k}(t) := \left\{x \in \mathbb{R}^n : \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t,x)} < 1\right\} \neq \emptyset,$$
$$\operatorname{bdry} \bar{C}^{\gamma_k}(t,k) := \left\{x \in \mathbb{R}^n : \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t,x)} = \frac{2\bar{\mu}}{\bar{\eta}^2 \gamma_k} = e^{-\gamma_k \bar{\alpha}_k}\right\} \neq \emptyset,$$
$$\operatorname{int} \bar{C}^{\gamma_k}(t,k) := \left\{x \in \mathbb{R}^n : \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t,x)} < \frac{2\bar{\mu}}{\bar{\eta}^2 \gamma_k} = e^{-\gamma_k \bar{\alpha}_k}\right\} \neq \emptyset.$$

Furthermore, $\bar{C}^{\gamma_k}(t)$ and $\bar{C}^{\gamma_k}(t,k)$ are amenable, epi-Lipschitz, and are respectively $\frac{\bar{\eta}}{L_\psi}$- and $\frac{\bar{\eta}}{2L_\psi}$-prox-regular.

(*iv*) $(\bar{C}^{\gamma_k}(t))_k$ and $(\bar{C}^{\gamma_k}(t,k))_k$ are nondecreasing sequences whose Painlevé-Kuratowski limit is $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ and satisfy

$$\operatorname{int}\left(C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))\right) = \bigcup_{k \in \mathbb{N}} \operatorname{int} \bar{C}^{\gamma_k}(t) = \bigcup_{k \in \mathbb{N}} \bar{C}^{\gamma_k}(t) = \bigcup_{k \in \mathbb{N}} \operatorname{int} \bar{C}^{\gamma_k}(t,k) = \bigcup_{k \in \mathbb{N}} \bar{C}^{\gamma_k}(t,k) \tag{3.64}$$

(*v*) For $c \in \operatorname{bdry}\left(C(0) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(0))\right)$, there exist $k_c \geq k_3$, $r_{\mathsf{c}} > 0$, and a vector $d_c \neq 0$ such that

$$\left(\left[\left(C(0) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(0))\right) \cap \bar{B}_{r_c}(c)\right] + \bar{\sigma}_k \frac{d_c}{\|d_c\|}\right) \subset \operatorname{int} \bar{C}^{\gamma_k}(0,k), \quad \forall k \geq k_c. \tag{3.65}$$

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

$$\left(c + \bar{\sigma}_k \frac{d_c}{\|d_c\|}\right) \in \operatorname{int} \bar{C}^{\gamma_k}(0,k). \tag{3.66}$$

Proof. (*i*). If this statement is not true, then there exist $(\gamma_{k_n})_n$ with $k_n \geq n$, $(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \in [0,T] \times \mathbb{R}^n$ with $\sum_{i=1}^{r+1} e^{\gamma_{k_n} \psi_i(t_{\gamma_{k_n}}, x_{\gamma_{k_n}})} = 1$, and $(\tau_{\gamma_{k_n}}, z_{\gamma_{k_n}}) \in B_{\frac{2}{n}}(t_{\gamma_{k_n}}, x_{\gamma_{k_n}})$ such that for all

$n > \frac{2}{\rho_o}$ we have

$$\| \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(\tau_{\gamma_{k_n}}, z_{\gamma_{k_n}})} \nabla_x \psi_i(\tau_{\gamma_{k_n}}, z_{\gamma_{k_n}}) \| \leq 2\bar{\eta} \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(\tau_{\gamma_{k_n}}, z_{\gamma_{k_n}})}. \tag{3.67}$$

Now, let $\bar{\psi}(t,x) := \max_{1 \leq i \leq r+1} \{\psi_i(t,x)\}$. Using Lemma 2.2.50, we have that

$$\bar{\psi}(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \leq \frac{1}{\gamma_{k_n}} \ln \left(\sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t_{\gamma_{k_n}}, x_{\gamma_{k_n}})} \right) = 0 \leq \bar{\psi}(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) + \frac{\ln(r+1)}{\gamma_{k_n}}. \tag{3.68}$$

Using the fact that $\psi_i(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \leq 0$, for all $i = 1, \cdots, r+1$, we deduce that the sequence $(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \in \text{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))\right)$ and hence, there exists a subsequence, we do not relabel, of $(\gamma_{k_n})_n$ along which the sequences $(t_{\gamma_{k_n}}, x_{\gamma_{k_n}})_n$ and $(\tau_{\gamma_{k_n}}, z_{\gamma_{k_n}})_n$ converge to the same element $(t_o, z_o) \in \text{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))\right)$. Taking $n \longrightarrow \infty$ in (3.67)-(3.68) and using the fact that $e^{\gamma_{k_n} \psi_i(\tau_{\gamma_{k_n}}, z_{\gamma_{k_n}})} \longrightarrow 0$ whenever $\psi_i(t_o, z_o) < 0$, we get the existence of a sequence of nonnegative numbers $(\lambda_i)_{i \in \bar{\mathcal{I}}^0_{(t_o,z_o)}}$ such that

$$\bar{\psi}(t_o, z_o) = 0 \quad \text{and} \quad \left\| \sum_{i \in \bar{\mathcal{I}}^0_{(t_o,z_o)}} \lambda_i \nabla_x \psi_i(\tau_o, z_o) \right\| \leq 2\bar{\eta} \;\; \text{with} \sum_{i \in \bar{\mathcal{I}}^0_{(\tau_o,z_o)}} \lambda_i = 1.$$

This contradicts Lemma 3.2.4 since $\bar{\psi}(t_o, z_o) = 0$ is equivalent to $\bar{\mathcal{I}}^0_{(t_o,z_o)} \neq \emptyset$.

(ii). If this statement is not true, there exist $(\gamma_{k_n})_n$ with $k_n \geq n$ and $(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \in [0,T] \times \mathbb{R}^n$ such that

$$e^{-\frac{\gamma_{k_n}}{n}} \leq \sum_{i=1}^{r+1} e^{\gamma_{k_n} \psi_i(t_{\gamma_{k_n}}, x_{\gamma_{k_n}})} \leq 1, \tag{3.69}$$

$$\| \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t_{\gamma_{k_n}}, x_{\gamma_{k_n}})} \nabla_x \psi_i(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \| \leq \bar{\eta} \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t_{\gamma_{k_n}}, x_{\gamma_{k_n}})}. \tag{3.70}$$

This yields that $(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \in \text{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))\right)$. Using (3.68)-(3.69), we deduce that $\bar{\psi}(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \longrightarrow 0$. Since $\text{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))\right)$ is compact, we can assume that $(t_{\gamma_{k_n}}, x_{\gamma_{k_n}}) \longrightarrow (t_o, x_o) \in \text{Gr}\left(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))\right)$, and hence $\bar{\psi}(t_o, z_o) = 0$. Taking $n \longrightarrow \infty$ in (3.70) and using that $e^{\gamma_{k_n} \psi_i(t_{\gamma_{k_n}}, x_{\gamma_{k_n}})} \longrightarrow 0$ whenever $\psi_i(t_o, x_o) < 0$, we get the existence of a sequence of nonnegative numbers $(\lambda_i)_{i \in \bar{\mathcal{I}}^0_{(t_o,x_o)}}$ such that

$$\left\| \sum_{i \in \bar{\mathcal{I}}^0_{(t_o,x_o)}} \lambda_i \nabla_x \psi_i(t_o, z_o) \right\| \leq \bar{\eta} \;\; \text{and} \sum_{i \in \bar{\mathcal{I}}^0_{(t_o,x_o)}} \lambda_i = 1.$$

This contradicts Lemma 3.2.4, since $\bar{\psi}(t_o, z_o) = 0$ implies that $\bar{\mathcal{I}}^0_{(t_o,z_o)} \neq \emptyset$.

(iii). To prove this part, we define for every $k \in \mathbb{N}$, the function $\psi_{\gamma_k} : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ such that

$$\psi_{\gamma_k}(t,x) := \frac{1}{\gamma_k} \ln \left(\sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t,x)} \right).$$

In that case,

$$\nabla_x \psi_{\gamma_k}(t,x) = \frac{\sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t,x)} \nabla_x \psi_i(t,x)}{\sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t,x)}}.$$

Notice that, for each $t \in [0,T]$, for each k, $\psi_{\gamma_k}(t,\cdot)$ is $\mathcal{C}^{1,1}$ on $\bar{C}^{\gamma_k}(t) + \rho_o B$. Translating (i) and applying it to a particular case, we deduce that

$$\text{for every } (t,x) \in \operatorname{bdry} \bar{C}^{\gamma_k}(t), \text{ we have } \|\nabla_x \psi_{\gamma_k}(t,x)\| > 2\bar{\eta}. \tag{3.71}$$

So, in summary, for each $t \in [0,T]$, we apply Lemma 2.2.48 (part **I.**), for $S := \bar{C}^{\gamma_k}(t)$, and $\psi(\cdot) := \psi_{\gamma_k}(t,\cdot)$. This proves all the properties in (iii) pertaining to $\bar{C}^{\gamma_k}(t)$, except the uniform constant for the prox-regularity. To prove that, we follow the same steps to prove the second part of (c) in [59, Proposition 5.3]. Now, to prove the properties pertaining to $\bar{C}^{\gamma_k}(t,k)$, we use Lemma 2.2.48 (part **II.**), for $S(k) := \bar{C}^{\gamma_k}(t,k)$. We use arguments similar to those used in the proof of the second part of (c) in [59, Proposition 5.3] to show that the prox-regular constant of $\bar{C}^{\gamma_k}(t,k)$ is uniform and equal to $\frac{\bar{\eta}}{2L_\psi}$.

(iv). Fix $t \in [0,T]$. Let $x \in \operatorname{int} \Big(C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t)) \Big)$, then $\bar{\psi}(t,x) < 0$. Since $\bar{\alpha}_k \to 0$, $\gamma_k \to \infty$, then there exists $k_x \in \mathbb{N}$, such that for all $k \geq k_x$, we have

$$\bar{\alpha}_{k_x} + \frac{\ln(r+1)}{\gamma_{k_x}} < -\bar{\psi}(t,x).$$

Then, using Lemma 2.2.50, we have that

$$\bar{\psi}(t,x) \leq \frac{1}{\gamma_{k_x}} \ln \left(\sum_{i=1}^{r+1} e^{\gamma_{k_x} \psi_i(t,x)} \right) \leq \bar{\psi}(t,x) + \frac{\ln(r+1)}{\gamma_{k_x}} < -\bar{\alpha}_{k_x}. \tag{3.72}$$

Hence, $\sum_{i=1}^{r+1} e^{\gamma_{k_x} \psi_i(t,x)} < e^{-\gamma_{k_x} \bar{\alpha}_{k_x}}$, and hence $x \in \operatorname{int} \bar{C}^{\gamma_k}(t,k)$. Then,

$$\begin{aligned} \operatorname{int} \Big(C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t)) \Big) \; &\subset \; \bigcup_{k \in \mathbb{N}} \operatorname{int} \bar{C}^{\gamma_k}(t,k) \subset \bigcup_{k \in \mathbb{N}} \bar{C}^{\gamma_k}(t,k) \\ &\subset \; \bigcup_{k \in \mathbb{N}} \operatorname{int} \bar{C}^{\gamma_k}(t) \subset \bigcup_{k \in \mathbb{N}} \bar{C}^{\gamma_k}(t) \subset \operatorname{int} \Big(C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t)) \Big). \end{aligned}$$

This proves that (3.64) is satisfied.

Using Lemma 2.2.50, we notice that for each (t,x), the function $\psi_{\gamma_k}(t,x)$ is non-increasing in k, and hence for each t, the sequence $(\bar{C}^{\gamma_k}(t))_k$ is non-decreasing. As a result, using Lemma 2.1.4, we show that the Painlevé-Kuratowski limit is

$$\lim_{k\to\infty} \bar{C}^{\gamma_k}(t) = \mathrm{cl}\left(\bigcup_{k\in\mathbb{N}} \bar{C}^{\gamma_k}(t)\right).$$

However, using (3.64) and (3.39), we deduce that

$$\mathrm{cl}\left(\bigcup_{k\in\mathbb{N}} \bar{C}^{\gamma_k}(t)\right) = \mathrm{cl\ int}\ \Big(C(t)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))\Big) = C(t)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t)).$$

On the other side, since for each (t,x), the function $\psi_{\gamma_k}(t,x)$ is non-increasing in k and the sequence $\bar{\alpha}_k$ is decreasing, then $(\bar{C}^{\gamma_k}(t,k))_k$ is nondecreasing. Then, using Lemma 2.1.4, we show that the Painlevé-Kuratowski limit is

$$\lim_{k\to\infty} \bar{C}^{\gamma_k}(t,k) = \mathrm{cl}\left(\bigcup_{k\in\mathbb{N}} \bar{C}^{\gamma_k}(t,k)\right) = C(t)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t)).$$

(v). We follow the same steps to prove Proposition 4.6(iii) in [58] replacing C, by $C(0)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(0))$, $\mathcal{I}^0_c$ by $\bar{\mathcal{I}}^0_{(0,c)}$, r by $r+1$, α_k by $\bar{\alpha}_k$, σ_k by $\bar{\sigma}_k$, $\psi_i(\cdot)$ by $\psi_i(0,\cdot)$, $\bar{M}_\psi$ by $\bar{L}$, η by $\bar{\eta}$. $\square$

Remark 3.2.11. We deduce, from Proposition 3.2.10, that for any $c\in C(0)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(0))$, there exists a sequence $(c_{\gamma_k})_k$ such that, for k large enough, $c_{\gamma_k}\in \mathrm{int}\,\bar{C}^{\gamma_k}(0,k)\subset \mathrm{int}\,\bar{C}^{\gamma_k}(0)$, and $c_{\gamma_k}\longrightarrow c$. Indeed:

(i) For $c\in \mathrm{bdry}\ \Big(C(0)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(0))\Big)$, we choose $c_{\gamma_k} := c+\bar{\sigma}_k\frac{d_c}{\|d_c\|}$ for all k. For $k\geq k_c$, we have from (3.66) that $c_{\gamma_k}\in \mathrm{int}\,\bar{C}^{\gamma_k}(0,k)$. Moreover, since $\bar{\sigma}_k\longrightarrow 0$ we have $c_{\gamma_k}\longrightarrow c$.

(ii) For $c\in$ int $\Big(C(0)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(0))\Big)$, Proposition 3.2.10$(iv)$ yields the existence of $\hat{k}_c\in\mathbb{N}$, such that $c\in \mathrm{int}\,\bar{C}^{\gamma_k}(0,k)$ for all $k\geq \hat{k}_c$. Hence, there exists $\hat{r}_c>0$ satisfying

$$c\in \bar{B}_{\hat{r}_c}(c)\subset \mathrm{int}\,\bar{C}^{\gamma_k}(0,k),\quad \forall k\geq \hat{k}_c.$$

In this case, we take the sequence $c_{\gamma_k}\equiv c\in \mathrm{int}\,\bar{C}^{\gamma_k}(0,k)$ that converges to c.

On the other hand, for the ball $\bar{B}_{\bar{\delta}}(\bar{y}(0))$ generated by the single function $\varphi(0,\cdot)$ in (3.31)-(3.32), we have the following property.

Proposition 3.2.12. There exists $k_o \in \mathbb{N}$ such that

$$\bar{B}_{\bar{\delta}}(\bar{y}(0)) \cap \bar{B}_{\frac{\bar{\delta}}{4}}(\mathbf{d}) - \frac{2\bar{\alpha}_k}{\bar{\delta}} \mathscr{V}(\mathbf{d}) \in B_{\bar{\rho}_k}(\bar{y}(0)), \ \ \forall k \geq k_o \ \text{ and } \ \forall \mathbf{d} \in \operatorname{bdry} \bar{B}_{\bar{\delta}}(\bar{y}(0)), \tag{3.73}$$

where $\mathscr{V}(\mathbf{d}) := \frac{\nabla_y \varphi(0,\mathbf{d})}{\|\nabla_y \varphi(0,\mathbf{d})\|} = \frac{\mathbf{d}-\bar{y}(0)}{\bar{\delta}}$.

Proof. This property follows by applying Lemma 2.2.48 (Part **II**)(*iv*) (or Theorem 3.1(*iii*) of [55]) to $S := \bar{B}_{\bar{\delta}}(\bar{y}(0))$, $r_o := \frac{\bar{\delta}}{4}$, and $\eta := \frac{\bar{\delta}}{2}$, and by noting the triangle inequality with $\|\nabla_y \varphi(0,\mathbf{d})\| = \bar{\delta}$ gives

$$\|\nabla_y \varphi(0,z)\| > \frac{\bar{\delta}}{2} \ \text{ and } \ \langle \nabla_y \varphi(0,z), \mathscr{V}(\mathbf{d})\rangle > \frac{\bar{\delta}}{2}, \ \ \forall \mathbf{d} \in \operatorname{bdry} \bar{B}_{\bar{\delta}}(\bar{y}(0)) \ \text{ and } \ \forall z \in B_{\frac{\bar{\delta}}{2}}(\mathbf{d}).$$

□

Parallel to Remark 3.2.11 and using Proposition 3.2.12, we deduce the following.

Remark 3.2.13. For any $\mathbf{d} \in \bar{B}_{\bar{\delta}}(\bar{y}(0))$, there exists a sequence $(d_{\gamma_k})_k$ such that, for k large enough, $d_{\gamma_k} \in \operatorname{int} \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$, and $d_{\gamma_k} \longrightarrow \mathbf{d}$. Indeed:

(*i*) As $\bar{\rho}_k \nearrow \bar{\delta}$, we deduce from (3.73) that for any $\mathbf{d} \in \operatorname{bdry} \bar{B}_{\bar{\delta}}(\bar{y}(0))$, there exists a sequence $(d_{\gamma_k})_k$ such that, for k large enough, $d_{\gamma_k} \in B_{\bar{\rho}_k}(\bar{y}(0)) \subset B_{\bar{\delta}}(\bar{y}(0))$, and $d_{\gamma_k} \longrightarrow \mathbf{d}$.

(*ii*) For $\mathbf{d} \in \operatorname{int} \bar{B}_{\bar{\delta}}(\bar{y}(0))$, there exists $\mathbf{k_d} \in \mathbb{N}$, such that $\mathbf{d} \in \operatorname{int} \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$ for all $k \geq \mathbf{k_d}$. Hence, there exists $\mathbf{r_d} > 0$ satisfying

$$\mathbf{d} \in \bar{B}_{\mathbf{r_d}}(\mathbf{d}) \subset \operatorname{int} \bar{B}_{\bar{\rho}_k}(\bar{y}(0)), \ \ \forall k \geq \mathbf{k_d}.$$

In this case, we take the sequence $d_{\gamma_k} \equiv \mathbf{d} \in \operatorname{int} \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$ that converges to $\mathbf{d}$.

The next theorem is fundamental for the thesis, as it illustrates two key ideas. First, it highlights the invariance for $(\bar{D}_{\gamma_k})$ of $\bar{C}^{\gamma_k}(\cdot,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(\cdot)) \subset \operatorname{int} \bar{C}^{\gamma_k}(\cdot) \times B_{\bar{\delta}}(\bar{y}(\cdot))$. More precisely, for k large, if the initial condition is in $\bar{C}^{\gamma_k}(0,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$, then $(\bar{D}_{\gamma_k})$ has a unique solution which is uniformly Lipschitz and remains in $\bar{C}^{\gamma_k}(t,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t))$ $\forall t \in [0,T]$. This result extends that in [55, 58] in two directions: (i) when the original problem has *coupled* sweeping processes, and (ii) when the sweeping set is *time-dependent* and *localized* near $(\bar{x}, \bar{y})$. Second, it shows that the solution of $(\bar{D}_{\gamma_k})$ uniformly approximates that of $(\bar{D})$.

Theorem 3.2.14. Let $(c_{\gamma_k}, d_{\gamma_k})_k$ be such that $(c_{\gamma_k}, d_{\gamma_k}) \in \bar{C}^{\gamma_k}(0,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$ for every k, and $(c_{\gamma_k}, d_{\gamma_k}) \longrightarrow (x_0, y_0) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(0)$. Let u_{γ_k} be a given sequence in $\mathcal{U}$. The following results hold:

(I). [**Existence of solution to** $(\bar{D}_{\gamma_k})$ **and Invariance**]

For k large enough, the Cauchy problem of the system $(\bar{D}_{\gamma_k})$ corresponding to $(x(0), y(0)) = (c_{\gamma_k}, d_{\gamma_k})$, and $u = u_{\gamma_k}$, has a unique solution $(x_{\gamma_k}, y_{\gamma_k}) \in W^{1,\infty}([0,T]; \mathbb{R}^n \times \mathbb{R}^l)$ such that

$$(x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \bar{C}^{\gamma_k}(t,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t)) \quad \forall t \in [0,T], \tag{3.74}$$

$$\max\{\|\xi_{\gamma_k}\|_\infty, \|\zeta_{\gamma_k}\|_\infty\} \le \frac{2\bar{\mu}}{\bar{\eta}^2}, \quad \max\{\|\dot{x}_{\gamma_k}\|_\infty, \|\dot{y}_{\gamma_k}\|_\infty\} \le M_h + \frac{2\bar{\mu}}{\bar{\eta}^2}\bar{L}, \tag{3.75}$$

where $\xi_{\gamma_k}(\cdot)$ and $\zeta_{\gamma_k}(\cdot)$ are the positive continuous functions on $[0,T]$ corresponding respectively to the solutions x_{γ_k} and y_{γ_k} via the formulae

$$\xi_{\gamma_k}(\cdot) := \sum_{i=1}^{r+1} \xi^i_{\gamma_k}(\cdot); \;\; \xi^i_{\gamma_k}(\cdot) := \gamma_k e^{\gamma_k \psi_i(\cdot, x_{\gamma_k}(\cdot))} \; (i = 1, \dots, r+1); \text{ and } \; \zeta_{\gamma_k}(\cdot) := \gamma_k e^{\gamma_k \varphi(\cdot, y_{\gamma_k}(\cdot))}. \tag{3.76}$$

(II). [**Solution of** $(\bar{D}_{\gamma_k})$ **converges to a unique solution of** $(\bar{D})$]

There exist $(x, y) \in W^{1,\infty}([0,T]; \mathbb{R}^n \times \mathbb{R}^l)$ and $(\xi^1, \cdots, \xi^r, \xi^{r+1}, \zeta) \in L^\infty([0,T]; \mathbb{R}^{r+2}_+)$ such that a subsequence of $((x_{\gamma_k}, y_{\gamma_k}), (\xi^1_{\gamma_k}, \cdots, \xi^r_{\gamma_k}, \xi^{r+1}_{\gamma_k}, \zeta_{\gamma_k}))$ (we do not relabel) satisfies

$$(x_{\gamma_k}, y_{\gamma_k}) \xrightarrow{unif} (x, y), \quad (\dot{x}_{\gamma_k}, \dot{y}_{\gamma_k}) \xrightarrow[in\ L^\infty]{w*} (\dot{x}, \dot{y}), \quad \xi^i_{\gamma_k} \xrightarrow[in\ L^\infty]{w*} \xi^i \; (\forall i), \; \zeta_{\gamma_k} \xrightarrow[in\ L^\infty]{w*} \zeta, \tag{3.77}$$

and ξ_{γ_k} converges weakly* in $L^\infty([0,T]; \mathbb{R}_+)$ to $\xi := \sum_{i=1}^{r+1} \xi^i$. Moreover,

$$(x(t), y(t)) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t) := (C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))) \times \bar{B}_{\bar{\delta}}(\bar{y}(t)) \quad \forall t \in [0,T], \tag{3.78}$$

$$\max\{\|\xi\|_\infty, \|\zeta\|_\infty\} \le \frac{2\bar{\mu}}{\bar{\eta}^2}, \quad \max\{\|\dot{x}\|_\infty, \|\dot{y}\|_\infty\} \le M_h + \frac{2\bar{\mu}}{\bar{\eta}^2}\bar{L}, \tag{3.79}$$

$$\begin{cases} \xi^i(t) = 0 \text{ for all } t \in I^-_i(x), \;\; i \in \{1, \cdots, r, r+1\}, \\ \xi(t) = 0 \text{ for all } t \in \bar{I}^-(x), \;\; \text{and } \; \zeta(t) = 0 \text{ for all } t \text{ such that } \varphi(t, y(t)) < 0. \end{cases} \tag{3.80}$$

If u_{γ_k} admits a subsequence that converges a.e. to some $u \in \mathcal{U}$, or if $(A1)$ and $(A4.2)$ hold, then there exists $u \in \mathcal{U}$ such that (x, y) is the unique solution of $(\bar{D})$ corresponding to $((x_0, y_0), u)$, and, for almost all $t \in [0, T]$,

$$\dot{x}(t) = f(t, x(t), y(t), u(t)) - \sum_{i=1}^{r+1} \xi^i(t) \nabla_x \psi_i(t, x(t)), \tag{3.81}$$

$$= f(t, x(t), y(t), u(t)) - \sum_{i \in \bar{\mathcal{I}}^0{}_{(t,x(t))}} \xi^i(t) \nabla_x \psi_i(t, x(t)), \tag{3.82}$$

$$\dot{y}(t) = g(t, x(t), y(t), u(t)) - \zeta(t) \nabla_y \varphi(t, y(t)). \tag{3.83}$$

Proof. **Part (I).**

Step I.1. A unique solution $(x_{\gamma_k}, y_{\gamma_k})$ of $(\bar{D}_{\gamma_k})$ exists on a certain interval $[0, \hat{T})$.
Recall that in Remark 3.2.9, for $t \in [0, T]$ a.e., $x \in \left[C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))\right]$, and for $u \in U(t)$, we extended $h(t, x, \cdot, u) := (f, g)(t, x, \cdot, u)$ so that $(A4.1)$ holds true for all $y \in \mathbb{R}^l$. Hence, for fixed $k \in \mathbb{N}$ and for $u := u_{\gamma_k}$, the system $(\bar{D}_{\gamma_k})$ is well defined on the set

$$\mathscr{D} := \{(t, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^l : x \in \text{int}\left(C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))\right), \;\; y \in B_{2\bar{\delta}}(\bar{y}(t))\}. \tag{3.84}$$

As $(0, c_{\gamma_k}, d_{\gamma_k}) \in \mathscr{D}$, standard local existence and uniqueness results from ordinary differential equations (see Theorem 2.3.3 or [39, Theorem 5.3]) confirm that for some $T_1 \in (0, T]$, the Cauchy problem $(\bar{D}_{\gamma_k})$ with $(x(0), y(0)) = (c_{\gamma_k}, d_{\gamma_k})$ has a unique solution $(x_{\gamma_k}, y_{\gamma_k}) \in W^{1,1}([0, T_1]; \mathbb{R}^n \times \mathbb{R}^l)$ such that $(s, x_{\gamma_k}(s), y_{\gamma_k}(s)) \in \mathscr{D}$ for all $s \in [0, T_1]$. Set

$$\hat{T} := \sup\{T_1 : \;\; (x, y) \text{ solves } (\bar{D}_{\gamma_k}) \text{ on } [0, T_1] \text{ with } (x(0), y(0)) = (c_{\gamma_k}, d_{\gamma_k}) \text{ and } (t, x(t), y(t)) \in \mathscr{D} \;\; \forall t \in [0, T_1]\}. \tag{3.85}$$

The uniqueness of the solution yields that a solution $(x_{\gamma_k}, y_{\gamma_k})$ of $(\bar{D}_{\gamma_k})$ with $(x(0), y(0)) = (c_{\gamma_k}, d_{\gamma_k})$ exists on the interval $[0, \hat{T})$, and we have $(t, x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \mathscr{D}$, $\forall t \in [0, \hat{T})$.

Step I.2. On $[0, \hat{T}]$, $(x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \bar{C}^{\gamma_k}(t) \times \bar{B}_{\bar{\delta}}(\bar{y}(t))$, and $\hat{T} = T$.
Notice that $x_{\gamma_k}(0) = c_{\gamma_k} \in \bar{C}^{\gamma_k}(0, k) \subset \text{int}\, \bar{C}^{\gamma_k}(0)$ implies that the function $\Delta(\cdot)$ given by $\Delta(\tau) := \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(\tau, x_{\gamma_k}(\tau))} - 1$ has $\Delta(0) < 0$. If for some $t_1 \in (0, \hat{T})$, $\Delta(t_1) = 0$, let $t > t_1$ close enough to t_1 so that $t \in (0, \hat{T})$. Then, from (3.53) and (3.50), we deduce that for

$i = 1, \cdots, r+1$, there exists $\theta^i_{\gamma_k}(\cdot) \in \hat{\partial}_s \psi_i(\cdot, x_{\gamma_k}(\cdot))$ a.e., such that

$$\frac{d}{ds}\psi_i(s, x_{\gamma_k}(s)) = \theta^i_{\gamma_k}(s) + \langle \nabla_x \psi_i(s, x_{\gamma_k}(s)), \dot{x}_{\gamma_k}(s)\rangle, \quad s \in [t_1, t] \text{ a.e.,} \tag{3.86}$$

$$\left|\theta^i_{\gamma_k}(s) + \langle \nabla_x \psi_i(s, x_{\gamma_k}(s)), f(s, x_{\gamma_k}(s), y_{\gamma_k}(s), u_{\gamma_k}(s))\rangle\right| \le \bar{L}(1 + M_h) = \bar{\mu}, \text{ a.e. } s \in [t_1, t]. \tag{3.87}$$

Then, using the first equation of $(\bar{D}_{\gamma_k})$, we obtain

$$\begin{aligned}
\Delta(t) - \Delta(t_1) &= \sum_{i=1}^{r+1} \int_{t_1}^{t} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \frac{d}{ds}\psi_i(s, x_{\gamma_k}(s)) ds \\
&\overset{(3.86)}{=} \int_{t_1}^{t} \Bigg(\sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \left(\theta^i_{\gamma_k}(s) + \langle \nabla_x \psi_i(s, x_{\gamma_k}(s)), f(s, x_{\gamma_k}(s), y_{\gamma_k}(s), u_{\gamma_k}(s))\rangle\right) \\
&\quad - \langle \sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \nabla_x \psi_i(s, x_{\gamma_k}(s)), \sum_{j=1}^{r+1} \gamma_k e^{\gamma_k \psi_j(s, x_{\gamma_k}(s))} \nabla_x \psi_j(s, x_{\gamma_k}(s))\rangle \Bigg) ds \\
&\overset{(3.87)}{\le} \int_{t_1}^{t} \left(\sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \bar{\mu} - \left\| \sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \nabla_x \psi_i(s, x_{\gamma_k}(s)) \right\|^2 \right) ds \qquad (3.88) \\
&\overset{(3.62)}{\le} \int_{t_1}^{t} \left(\sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \left(\bar{\mu} - 4\bar{\eta}^2 \gamma_k \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \right) \right) ds \\
&\le \int_{t_1}^{t} \sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} (\bar{\mu} - 2\bar{\eta}^2 \gamma_k) ds \quad < 0,
\end{aligned}$$

the third and the second to last inequality are due to the fact that we can choose t close enough to t_1 so that, for $s \in [t_1, t]$, $x_{\gamma_k}(s) \in B_{2r_1}(t_1, x_{\gamma_k}(t_1))$ (so we apply (3.62)) and $\sum_{j=1}^{r+1} \gamma_k e^{\gamma_k \psi_j(s, x_{\gamma_k}(s))} > \frac{1}{2}$, and the last inequality follows from $\gamma_k > \frac{\bar{\mu}}{2\bar{\eta}^2}$. This shows that, $\forall t_1 \in (0, \hat{T})$ with $\Delta(t_1) = 0$, $\Delta(t) < 0$ for all $t > t_1$ close enough to t_1. Whence, the continuity of $\Delta(\cdot)$ on $[0, \hat{T})$ and $\Delta(0) < 0$ yield that $\Delta(t) \le 0$ for all $t \in [0, \hat{T})$, that is, $x_{\gamma_k}(t) \in \bar{C}^{\gamma_k}(t) \subset \text{int}\left(C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))\right) \forall t \in [0, \hat{T})$.

On the other hand, as $y_{\gamma_k}(0) = d_{\gamma_k} \in \bar{B}_{\bar{\rho}_k}(\bar{y}(0)) \subset B_{\bar{\delta}}(\bar{y}(0))$, we have $\varphi(0, y_{\gamma_k}(0)) < 0$. If for some $t_1 \in (0, \hat{T})$, $\varphi(t_1, y_{\gamma_k}(t_1)) = 0$, that is, $\|y_{\gamma_k}(t_1) - \bar{y}(t_1)\| = \bar{\delta}$, choose $t > t_1$ close enough to t_1 so that, $\forall s \in [t_1, t]$, $e^{\gamma_k \varphi(s, y_{\gamma_k}(s))} > \frac{1}{2}$ and $\|y_{\gamma_k}(s) - \bar{y}(s)\| > \frac{\bar{\delta}}{2}$. Hence, (3.55), $(\bar{D}_{\gamma_k})$, (A4.1), (3.47), $y_{\gamma_k}(\cdot) \in B_{2\bar{\delta}}(\bar{y}(\cdot))$, and $\bar{\eta} < \frac{\bar{\delta}}{2}$ (by Lemma 3.2.4), yield that, for

$s \in [t_1, t]$ a.e.,

$$\begin{aligned}
\frac{d}{ds}\varphi(s, y_{\gamma_k}(s)) &= \langle y_{\gamma_k}(s) - \bar{y}(s), -\dot{\bar{y}}(s) + g(s, x_{\gamma_k}(s), y_{\gamma_k}(s), u_{\gamma_k}(s))\rangle \\
&\quad -\gamma_k e^{\gamma_k \varphi(s, y_{\gamma_k}(s))}\|y_{\gamma_k}(s) - \bar{y}(s)\|^2 \\
&\leq \|y_{\gamma_k}(s) - \bar{y}(s)\| \; L_{(\bar{x},\bar{y})}(1 + M_h) - \gamma_k e^{\gamma_k \varphi(s, y_{\gamma_k}(s))}\|y_{\gamma_k}(s) - \bar{y}(s)\|^2 \qquad (3.89) \\
&< \;\; 2\bar{\mu} - \frac{\gamma_k}{2}\bar{\eta}^2 < 0,
\end{aligned}$$

the last inequality follows from $\gamma_k > \frac{2\bar{\mu}}{\bar{\eta}^2}e$. Hence, for all t close enough to t_1, we have

$$\varphi(t, y_{\gamma_k}(t)) = \varphi(t, y_{\gamma_k}(t)) - \varphi(t_1, y_{\gamma_k}(t_1)) = \int_{t_1}^{t} \frac{d}{ds}\varphi(s, y_{\gamma_k}(s))ds < 0.$$

This shows that $y_{\gamma_k}(t) \in \bar{B}_{\bar{\delta}}(\bar{y}(t))$ for all $t \in [0, \hat{T})$.

Since for $t \in [0, \hat{T})$, $(t, x_{\gamma_k}(t), y_{\gamma_k}(t))$ remains in the compact set $\mathrm{Gr}\Big(\bar{C}^{\gamma_k}(\cdot) \times \bar{B}_{\bar{\delta}}(\bar{y}(\cdot))\Big)$ then it is possible to extend in this compact set the solution $(x_{\gamma_k}, y_{\gamma_k})$ to the whole interval $[0, \hat{T}]$. If $\hat{T} < T$, the local existence of a solution starting at $\hat{T}$ contradicts the definition of $\hat{T}$, proving that $\hat{T} = T$. This completes Step **I.2**.

Step I.3. Invariance of $\bar{C}^{\gamma_k}(t, k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t))$, i.e., (3.74) is valid.

As $c_{\gamma_k} \in \bar{C}^{\gamma_k}(0, k)$, we have $\sum_{i=1}^{r+1} e^{\gamma_k \psi_i(0, c_{\gamma_k})} \leq \frac{2\bar{\mu}}{\bar{\eta}^2\gamma_k}$. Since $\gamma_k \to \infty$, there exists $k_4 \in \mathbb{N}$ large enough such that for all $k \geq k_4$, we have that

$$\frac{2\bar{\mu}}{\bar{\eta}^2\gamma_k} \geq \max\{e^{-\frac{\gamma_k \bar{\epsilon}_o}{2}}, \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(0, c_{\gamma_k})}\}, \qquad (3.90)$$

where $\bar{\epsilon}_o$ the constant from (3.63). Fix $k \geq k_4$. Let $t_1 \in [0, T)$ such that $\sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t_1, x_{\gamma_k}(t_1))} = \frac{2\bar{\mu}}{\bar{\eta}^2\gamma_k}$. Let $\bar{\epsilon}_k = \min\{\frac{\bar{\epsilon}_o}{2}, \frac{\ln 2}{2\gamma_k}\}$. Using the continuity of x_{γ_k} and $\psi_i(\cdot, \cdot)$, we can choose t close enough to t_1 such that for all $s \in [t_1, t]$,

$$\begin{aligned}
\sum_{i=1}^{r+1} e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} &\geq \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t_1, x_{\gamma_k}(t_1))} e^{-\gamma_k \bar{\epsilon}_k} = \frac{2\bar{\mu}}{\gamma_k \bar{\eta}^2} e^{-\gamma_k \bar{\epsilon}_k} \qquad (3.91) \\
&\overset{(3.90)}{\geq} e^{-\frac{\gamma_k \bar{\epsilon}_o}{2}} e^{-\gamma_k \bar{\epsilon}_k} \geq e^{-\frac{\gamma_k \bar{\epsilon}_o}{2}} e^{-\frac{\gamma_k \bar{\epsilon}_o}{2}} = e^{-\gamma_k \bar{\epsilon}_o}.
\end{aligned}$$

Hence, by Proposition 3.2.10(ii), and the fact that $x_{\gamma_k}(\tau) \in \bar{C}^{\gamma_k}(\tau)$ for all $\tau \in [0, T]$ (see **Step I.2**), we have

$$\|\sum_{i=1}^{r+1} e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \nabla_x \psi_i(s, x_{\gamma_k}(s))\| > \bar{\eta} \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))}, \quad \forall s \in [t_1, t]. \qquad (3.92)$$

Thus, for $\bar{\Delta}(\cdot) := \sum_{j=1}^{r+1} e^{\gamma_k \psi_j(\cdot, x_{\gamma_k}(\cdot))} - \frac{2\bar{\mu}}{\bar{\eta}^2 \gamma_k}$, we have

$$\begin{aligned}
\bar{\Delta}(t) - \bar{\Delta}(t_1) \quad &= \quad \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t, x_{\gamma_k}(t))} - \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(t_1, x_{\gamma_k}(t_1))} \\
&\overset{(3.88)}{\leq} \int_{t_1}^{t} \left(\sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \bar{\mu} - \left\| \sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \nabla_x \psi_i(s, x_{\gamma_k}(s)) \right\|^2 \right) ds \\
&\overset{(3.92)}{\leq} \int_{t_1}^{t} \left(\sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \left(\bar{\mu} - \bar{\eta}^2 \gamma_k \sum_{i=1}^{r+1} e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \right) \right) ds \\
&\overset{(3.91)}{\leq} \int_{t_1}^{t} \sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(s, x_{\gamma_k}(s))} \bar{\mu} \left(1 - 2e^{-\gamma_k \bar{\epsilon}_k} \right) ds < 0,
\end{aligned}$$

the last inequality follows from the definition of $\bar{\epsilon}_k$. This proves that $x_{\gamma_k}(t) \in \bar{C}^{\gamma_k}(t,k)$ for all $t > t_1$ close enough to t_1. Whence, similarly to **Step I.2**, the continuity of $\bar{\Delta}(\cdot)$ and $\bar{\Delta}(0) \leq 0$ imply that $x_{\gamma_k}(t) \in \bar{C}^{\gamma_k}(t,k),\ \forall t \in [0,T]$.

On the other hand, having $y_{\gamma_k}(0) = d_{\gamma_k} \in \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$, where $\bar{\rho}_k$ is given in (3.57), means that $\varphi(0, y_{\gamma_k}(0)) \leq -\bar{\alpha}_k$. Since $\bar{\alpha}_k \to 0$, and $\bar{\alpha}_k > 0$ for all k, then we can find $k_5 \geq k_4$ such that

$$\bar{\alpha}_k \leq \min\left\{ \frac{\bar{\delta}^2}{4}, -\varphi(0, d_{\gamma_k}) \right\} \text{ for all } k \geq k_5.$$

Define $\hat{\epsilon}_k := \min\{\frac{\bar{\delta}^2}{8}, \frac{\ln 2}{2\gamma_k}\}$. If for some $t_1 \in [0,T)$, $\varphi(t_1, y_{\gamma_k}(t_1)) = -\bar{\alpha}_k$, let $t > t_1$ close enough to t_1 such that

$$\varphi(s, y_{\gamma_k}(s)) \geq -\bar{\alpha}_k - \hat{\epsilon}_k, \quad \forall s \in [t_1, t]. \tag{3.93}$$

Then, for all $s \in [t_1, t]$, we have

$$\|y_{\gamma_k}(s) - \bar{y}(s)\|^2 \geq \bar{\delta}^2 - 2\bar{\alpha}_k - 2\hat{\epsilon}_k \geq \frac{\bar{\delta}^2}{4} \geq \bar{\eta}^2. \tag{3.94}$$

Hence, using, respectively, (3.89), $y_{\gamma_k}(\cdot) \in \bar{B}_{\bar{\delta}}(\bar{y}(\cdot))$, (3.47), (3.93), first equation in (3.58), and (3.94), we deduce that

$$\begin{aligned}
\varphi(t, y_{\gamma_k}(t)) - \varphi(t_1, y_{\gamma_k}(t_1)) &= \int_{t_1}^{t} \frac{d}{ds}\varphi(s, y_{\gamma_k}(s))ds \\
&\leq \int_{t_1}^{t} \left(\bar{\mu} - \gamma_k e^{-\gamma_k \bar{\alpha}_k} e^{-\gamma_k \hat{\epsilon}_k} \|y_{\gamma_k}(s) - \bar{y}(s)\|^2\right) ds \\
&\leq \int_{t_1}^{t} \left(\bar{\mu} - \frac{2\bar{\mu}}{\bar{\eta}^2} e^{-\gamma_k \hat{\epsilon}_k} \bar{\eta}^2\right) ds \\
&= \int_{t_1}^{t} \bar{\mu}\left(1 - 2e^{-\gamma_k \hat{\epsilon}_k}\right) ds < 0
\end{aligned}$$

proving that $\varphi(t, y_{\gamma_k}(t)) < -\bar{\alpha}_{\gamma_k}$. Thus, the continuity of $\varphi(\cdot, y_{\gamma_k}(\cdot))$ yields, $y_{\gamma_k}(t) \in \bar{B}_{\bar{\rho}_k}(\bar{y}(t))$ $\forall t \in [0, T]$.

Step I.4. $(x_{\gamma_k}, y_{\gamma_k}, \xi_{\gamma_k}, \zeta_{\gamma_k})$ **satisfy equation (3.75).**

So far, we proved that a solution $(x_{\gamma_k}, y_{\gamma_k})$ of the Cauchy problem of $(\bar{D}_{\gamma_k})$ exists and satisfies (3.74). Hence, the definitions of $\bar{C}^{\gamma_k}(t, k)$ and ξ_{γ_k} given in (3.60) and (3.76), respectively, yield that $\|\xi_{\gamma_k}\|_\infty \leq \frac{2\bar{\mu}}{\bar{\eta}^2}$. On the other hand, the definition of $\bar{B}_{\bar{\rho}_k}(\bar{y}(t))$ yields that $\varphi(t, y_{\gamma_k}(t)) \leq -\bar{\alpha}_k$, and thus, the same bound is immediately obtained for the norm of ζ_{γ_k}, defined in (3.76). Whence, the first inequality in (3.75) is satisfied. Employing this latter in $(\bar{D}_{\gamma_k})$ and then calling on the definition of $\bar{L}$ in (3.47), we obtain that the second inequality in (3.75) is valid.

Part (II).

Step II.1. Existence of $(\xi^1, \cdots, \xi^{r+1}, \zeta)$ **and** (x, y) **satisfying** (3.77)-(3.80)**.**

Using (3.74)-(3.75), it follows that (.1) holds for $\mathcal{R} := r + 1$ and $(x_k, y_k, \xi_k^i, \zeta_k) := (x_{\gamma_k}, y_{\gamma_k}, \xi_{\gamma_k}^i, \zeta_{\gamma_k})$. Hence, by Lemma .0.2$(i)$, there is a subsequence (not relabeled) of $(x_{\gamma_k}, y_{\gamma_k})$, $(\xi_{\gamma_k}^1, \cdots, \xi_{\gamma_k}^{r+1}, \zeta_{\gamma_k})$, that converges, respectively, to some $(x, y) \in W^{1,\infty}([0, T]; \mathbb{R}^{n+l})$, $(\xi^1, \cdots, \xi^{r+1}, \zeta) \in L^\infty([0, T]; \mathbb{R}_+^{r+2})$, such that (3.77) and (3.79) are satisfied. Moreover, (3.78) follows from (3.74), (3.58), and Proposition 3.2.10 (iv).

Now, we show that (3.80) holds. Let $i \in \{1, \cdots, r+1\}$ and $t \in I_i^{\text{-}}(x)$, that is, $\psi_i(t, x(t)) < 0$. Then, by $(A3.1)$ and the uniform convergence of x_{γ_k} to x, there exist $k_t \in \mathbb{N}$, $\alpha_t > 0$, and

$\tau_t > 0$ such that $\forall k \geq k_t$, we have

$$\psi_i(s, x_{\gamma_k}(s)) < -\frac{\alpha_t}{2}, \ \ \forall s \in (t - \tau_t, t + \tau_t) \cap [0, T].$$

Hence, $\xi^i_{\gamma_k}(s) < \gamma_k e^{-\gamma_k \frac{\alpha_t}{2}} \xrightarrow[k \to \infty]{} 0$, uniformly on $(t - \tau_t, t + \tau_t) \cap [0, T]$ and $\xi^i(t) = 0$. Let $t \in \bar{I}^{\text{-}}(x)$, then $t \in I^{\text{-}}_i(x)$ $\forall i \in \{1, \cdots, r+1\}$, and hence, $\xi^i(t) = 0$ $\forall i \in \{1, \cdots, r+1\}$, implying that also $\xi(t) = 0$. Similarly, let $t \in [0, T]$ such that $\varphi(t, y(t)) < 0$. The same arguments now applied to $\varphi(t, y(t))$ yield the existence of $\hat{k}_t \in \mathbb{N}$, $\hat{\alpha}_t > 0$ and $\hat{\tau}_t > 0$ such that, $\forall s \in (t - \hat{\tau}_t, t + \hat{\tau}_t) \cap [0, T]$,

$$\zeta_{\gamma_k}(s) := \gamma_k e^{\gamma_k \varphi(s, y_{\gamma_k}(s))} < \gamma_k e^{\frac{-\gamma_k \hat{\alpha}_t}{2}} \xrightarrow[k \to \infty]{\text{uniformly}} 0, \ \ \text{and hence, } \zeta(t) = 0.$$

Step II.2. Existence of $u \in \mathcal{U}$: $((x, y), u)$ & (ξ^i, ζ) satisfy $(\bar{D})$ and (3.81)-(3.83), (x, y) unique.

Whether u_{γ_k} admits a subsequence that converges to some $u \in \mathcal{U}$, for $t \in [0, T]$ a.e., or assumptions (A1) and (A4.2) are satisfied, apply in each of the two cases the corresponding result in Lemma .0.2(ii) to $Q(\cdot) := C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot))$, $\mathcal{R} := r + 1$, $q_i(\cdot, \cdot) := \psi_i(\cdot, \cdot)$, and to the sequences $((x_k, y_k), u_k) := ((x_{\gamma_k}, y_{\gamma_k}), u_{\gamma_k})$, $\xi^i_k := \xi^i_{\gamma_k}$, and $\zeta_k := \zeta_{\gamma_k}$ in (3.76), and their respective limits (x, y), ξ^i and ζ. Then, there exists $u(\cdot)$ such that $((x, y), u)$, ξ^i $(i = 1, \cdots, r+1)$ and ζ satisfy (3.81)-(3.83). The facts that (x, y) is a solution of $(\bar{D})$ corresponding to $((x_0, y_0), u)$ and is unique follow now directly from Lemma 3.2.8. This completes the proof of this Theorem. □

Remark 3.2.15. Similar arguments to steps **I.2-3** in the proof of Theorem 3.2.14 also show the invariance of the larger sets $\bar{C}^{\gamma_k}(t) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t))$; this means that if $(c_{\gamma_k}, d_{\gamma_k})$ is taken in $\bar{C}^{\gamma_k}(0) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$, then for all $t \in [0, T]$, $(x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \bar{C}^{\gamma_k}(t) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t))$.

The following corollary is an immediate consequence of Theorem 3.2.14, in which we take $u_{\gamma_k} = u$ for all k, and hence, neither (A1) nor (A4.2) is required. It also consists of a *Lipschitz-existence and uniqueness* result for the Cauchy problem of $(\bar{D})$ via the solution of the Cauchy problem of $(\bar{D}_{\gamma_k})$, whose initial condition is carefully chosen.

Corollary 3.2.16. For given $(x_0, y_0) \in \bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(0)$ and $u \in \mathcal{U}$, the system $(\bar{D})$ corresponding to $((x_0, y_0); u)$ has a unique solution (x, y), and hence it is Lipschitz and satisfies (3.44)-(3.46). This solution is the uniform limit of a subsequence (not relabeled) of $(x_{\gamma_k}, y_{\gamma_k})_k$, which is obtained via Theorem 3.2.14 as the solution of $(\bar{D}_{\gamma_k})$ with $((x(0), y(0)); u) = ((c_{\gamma_k}, d_{\gamma_k}); u)$, where c_{γ_k} and d_{γ_k} are the sequences from Remarks 3.2.11 and 3.2.13 corresponding to $c = x_0$ and $\mathbf{d} = y_0$, respectively. Hence, for k sufficiently large, we have that $(x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \bar{C}^{\gamma_k}(t, k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t))\ \forall t \in [0, T]$, $(x_{\gamma_k}, y_{\gamma_k})_k$ is uniformly lipschitz, and all conclusions of Theorem 3.2.14 hold.

We now present the table summarizing the results of Subsection 3.2.2.

Table 3.3 Summary of results from Subsection 3.2.2 .

Result	**Description**
Remark 3.2.9	We extend the function $h(t, x, \cdot, u)$ from $\bar{B}_{\bar{\delta}}(\bar{y}(t))$ to $\mathbb{R}^l$ so that this extension satisfies for all $y \in \mathbb{R}^l$, $(A4.1)$, and also (A4.2) whenever it is satisfied by h. This extension shall be later used in Theorem 3.2.14.
Proposition 3.2.10	We derive properties for the sets $\bar{C}^{\gamma_k}(t)$ and $\bar{C}^{\gamma_k}(t, k)$.
Remark 3.2.11	We approximate any $c \in C(0) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(0))$ by a sequence $c_{\gamma_k} \in \operatorname{int} \bar{C}^{\gamma_k}(0, k) \subset \operatorname{int} \bar{C}^{\gamma_k}(0)$ such that $c_{\gamma_k} \longrightarrow c$.
Proposition 3.2.12	We derive properties for the ball $\bar{B}_{\bar{\delta}}(\bar{y}(0))$ generated by the single function $\varphi(0, \cdot)$.
Remark 3.2.13	We approximate any $\mathbf{d} \in \bar{B}_{\bar{\delta}}(\bar{y}(0))$ by a sequence $d_{\gamma_k} \in \operatorname{int} \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$ such that $d_{\gamma_k} \longrightarrow \mathbf{d}$.
Theorem 3.2.14	We highlight the invariance for $(\bar{D}_{\gamma_k})$ of $\bar{C}^{\gamma_k}(\cdot, k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(\cdot)) \subset \operatorname{int} \bar{C}^{\gamma_k}(\cdot) \times B_{\bar{\delta}}(\bar{y}(\cdot))$, and show that the solution of $(\bar{D}_{\gamma_k})$ uniformly approximates that of $(\bar{D})$.

Table 3.3 (cont'd)

Result	Description
Remark 3.2.15	We highlight the invariance of the larger sets $\bar{C}^{\gamma_k}(t) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t))$.
Corollary 3.2.16	This is an immediate consequence of Theorem 3.2.14 and consists of a *Lipschitz-existence and uniqueness* result for the Cauchy problem of $(\bar{D})$ via the solution of the Cauchy problem of $(\bar{D}_{\gamma_k})$, whose initial condition is carefully chosen.

3.3 Study of the dynamic (D) under global assumptions

We now introduce the following *global* versions of the previous assumptions that shall be used for the global results in this section. For completeness and the reader's convenience, we include them here. $(\mathbf{A3.1})_G$ and $(\mathbf{A4})_G$ are, respectively, assumptions (A3.1) and (A4) when satisfied for $\bar{\delta} = \infty$ (the same constants' labels are kept), that is, $\bar{x}, \bar{y}$ and the balls around them are not involved therein, and $(\mathbf{A3.2})_G$ is a global version of (A3.2) which will imply the uniform prox-regularity of $C(t)$.

$(\mathbf{A3})_G$ Global assumptions on the functions ψ_i:

$(\mathbf{A3.1})_G$ There exist $\rho_o > 0$ and $L_\psi > 0$ such that, for each i, $\nabla_x \psi_i(\cdot,\cdot)$ exists on

Gr $(C(\cdot))$ $+\{0\} \times \rho_o B$, and $\psi_i(\cdot,\cdot)$ and $\nabla_x \psi_i(\cdot,\cdot)$ satisfy, for all

(t_1, x_1), $(t_2, x_2) \in$ Gr $(C(\cdot)) + \{0\} \times \frac{\rho_o}{2}\bar{B}$,

$$\max\{|\psi_i(t_1,x_1) - \psi_i(t_2,x_2)|, \|\nabla_x\psi_i(t_1,x_1) - \nabla_x\psi_i(t_2,x_2)\|\} \le L_\psi(\, |t_1 - t_2| + \|x_1 - x_2\|).$$

$(\mathbf{A3.2})_G$ For every $(t,x) \in$ Gr $C(\cdot)$ with $\mathcal{I}^0_{(t,x)} \neq \emptyset$ we have

$$\left[\sum_{i \in \mathcal{I}^0_{(t,x)}} \lambda_i \nabla_x \psi_i(t,x) = 0, \text{ with each } \lambda_i \ge 0\right] \Longrightarrow \left[\lambda_i = 0, \ \forall i \in \mathcal{I}^0_{(t,x)}\right].$$

$(\mathbf{A4})_G$ Global assumptions on $h(t,x,y,u) := (f,g)(t,x,y,u)$:

$(\mathbf{A4.1})_G$ For $(x,y,u) \in \bigcup_{t\in[0,T]} C(t) \times \mathbb{R}^l \times \mathbb{U}$, $h(\cdot,x,y,u)$ is Lebesgue-measurable and, for a.e. $t \in [0,T]$, $h(t,\cdot,\cdot,\cdot)$ is continuous on $C(t) \times \mathbb{R}^l \times U(t)$. There exist $M_h > 0$, and $L_h \in L^2([0,T];\mathbb{R}^+)$, such that, for a.e. $t \in [0,T]$, for all (x,y), $(x',y') \in C(t) \times \mathbb{R}^l$ and $u \in U(t)$,

$$\|h(t,x,y,u)\| \le M_h \text{ and } \|h(t,x,y,u) - h(t,x',y',u)\| \le L_h(t)\|(x,y) - (x',y')\|.$$

(A4.2)$_G$ The set $h(t,x,y,U(t))$ is convex for all $(x,y) \in C(t) \times \mathbb{R}^l$ and $t \in [0,T]$ a.e. [2]

3.3.1 Preliminary results

The compactness of Gr $C(\cdot)$ assumed in the following lemma allows us to easily imitate the proof of Lemma 3.1.1 and produce the following equivalence between (A3.2)$_G$ and a global version of condition (3.10), namely, (3.95), in which $\bar{x}$ and the localization around it are absent.

Lemma 3.3.1. Assume that $\psi_i(\cdot,\cdot)$ is continuous and, for all $t \in [0,T]$, the set $C(t)$ is nonempty, closed, and given by (3.3). Assume that $(A3.1)_G$ holds and that Gr $C(\cdot)$ is compact. Then $(A3.2)_G$ is equivalent to the existence of a constant $\eta > 0$ such that

$$\left\| \sum_{i \in \mathcal{I}^0_{(t,c)}} \lambda_i \nabla_x \psi_i(t,c) \right\| > 2\eta, \quad \forall (t,c) \in \{(\tau,x) \in \text{Gr } C(\cdot) : \mathcal{I}^0_{(\tau,x)} \neq \emptyset\}, \tag{3.95}$$

where $\mathcal{I}^0_{(\tau,x)}$ is defined in (3.7) and $(\lambda_i)_{i \in \mathcal{I}^0_{(t,c)}}$ is any sequence of nonnegative numbers satisfying $\sum_{i \in \mathcal{I}^0_{(t,c)}} \lambda_i = 1$.

As a consequence of Lemma 3.3.1, we obtain the uniform prox-regularity of $C(t)$, as well as a formula for the normal cone to $C(t)$. For L_ψ the common bound of $\{\|\nabla_x \psi_i(\cdot,\cdot)\|\}_{i=1}^r$ and the common Lipschitz constant of $\{\nabla_x \psi_i(\cdot,\cdot)\}_{i=1}^r$ on the compact set Gr $C(\cdot) + \{0\} \times \frac{\rho_o}{2}\bar{B}$, we assume without loss of generality that $L_\psi \ge \frac{8\eta}{\rho_o}$.

Lemma 3.3.2. Assume that $\psi_i(\cdot,\cdot)$ is continuous and, for all $t \in [0,T]$, the set $C(t)$ is nonempty, closed, and given by (3.3). Assume that $(A3.1)_G$ and $(A3.2)_G$ hold, and that

[2] This condition is only needed for the existence of an optimal solution,Theorem 4.1.1, and not for Theorem 3.3.7.

Gr $C(\cdot)$ is compact. Then, for all $t \in [0,T]$, $C(t)$ is amenable (in the sense of [62]), epi-lipschitzian, $C(t) = \mathrm{cl}\,(\mathrm{int}\, C(t))$, and is *uniformly* $\frac{2\eta}{L_\psi}$-prox-regular. In this global setting, the normal cone formula (3.12) is now valid for all $(t,x) \in$ Gr $C(\cdot)$. In particular,

$$N_{C(t)}(x) = N^P_{C(t)}(x) = N^L_{C(t)}(x) = \left\{ \sum_{i \in \mathcal{I}^0_{(t,x)}} \lambda_i \nabla_x \psi_i(t,x) : \lambda_i \geq 0 \right\} \neq \{0\}, \quad \forall x \in \mathrm{bdry}\, C(t). \tag{3.96}$$

Proof. We use condition (3.95), Lemma 2.2.41 ([2, Theorem 9.1]) (with $\min\{\rho_o, \frac{2\eta}{L_\psi}\} = \frac{2\eta}{L_\psi}$), Lemma 2.2.11, Lemma 2.2.46, and Lemma 2.2.43. □

Remark 3.3.3. Since $C(t)$ is $\frac{2\eta}{L_\psi}$-prox-regular, then each point in $C(t) + \frac{2\eta}{L_\psi}B$ has a unique projection on $C(t)$. Define for a.e. $t \in [0,T]$, and $(x,y,u) \in \left[C(t) + \frac{\eta}{L_\psi}B\right] \times \mathbb{R}^l \times U(t)$,

$$\hat{h}(t,x,y,u) := h(t, \pi_1(t,x), y, u),$$

where $\pi_1(t,\cdot) := \pi_{C(t)}(\cdot)$. Notice that $\hat{h}$ is well-defined, and $\pi_1(t,\cdot)$ is 2-lipschitz on $C(t) + \frac{\eta}{L_\psi}B$ (see Proposition 2.2.39(ii)). This means that the function $\hat{h}$ (which we relabel h) satisfies (A4.1)$_G$, where $C(t)$ is now replaced by $C(t) + \frac{\eta}{L_\psi}B$, and $L_h(t)$ is now replaced by $2L_h(t)$. On the other hand, we note that since $\frac{\eta}{L_\psi} < \frac{\rho_o}{2}$, then $\psi_1, \cdots, \psi_r$ satisfy (A3.1)$_G$ on Gr $(C(\cdot)) + \{0\} \times \frac{\eta}{L_\psi}\bar{B}$.

The following lemma, which requires Gr $C(\cdot)$ bounded, is a global version of Lemma 3.1.5.

Lemma 3.3.4. Assume that $\psi_i(\cdot,\cdot)$ is continuous and, for all $t \in [0,T]$, the set $C(t)$ is nonempty, closed, and given by (3.3). Assume that $(A3.1)_G$, $(A3.2)_G$ and $(A4.1)_G$ hold, and that Gr $C(\cdot)$ is compact. Let $u \in \mathcal{U}$, $(x_0,y_0) \in C(0) \times \mathbb{R}^l$ be fixed, and $(x(\cdot),y(\cdot)) \in W^{1,1}([0,T];\mathbb{R}^n \times \mathbb{R}^l)$ with $(x(0),y(0)) = (x_0,y_0)$ and $x(t) \in C(t)$, $\forall t \in [0,T]$. Then, (x,y) solves (D) corresponding to $((x_0,y_0),u)$ if and only if there exist measurable functions $(\lambda_1(\cdot), \cdots, \lambda_r(\cdot))$ such that, for all $i = 1, \cdots, r$, $\lambda_i(t) = 0$ for $t \in I^-_i(x)$, and $((x,y),u)$

together with $(\lambda_1, \cdots, \lambda_r)$ satisfies

$$\begin{cases} \dot{x}(t) = f(t, x(t), y(t), u(t)) - \sum_{i \in \mathcal{I}^0_{(t,x(t))}} \lambda_i(t) \nabla_x \psi_i(t, x(t)) \ \text{ a.e. } t \in [0, T], \\ \dot{y}(t) = g(t, x(t), y(t), u(t)), \ \text{ a.e. } \ t \in [0, T], \\ (x(0), y(0)) = (x_0, y_0), \end{cases} \tag{3.97}$$

and, we have the following bounds

$$\begin{cases} \|\lambda_i\|_\infty \le \| \sum_i^r \lambda_i \|_\infty \le \frac{\mu}{4\eta^2}, \ \ \forall i = 1, \cdots, r, \\ \|\dot{x}\|_\infty \le M_h + \frac{\mu}{4\eta^2} L_\psi, \ \ \|\dot{y}\|_\infty \le M_h. \end{cases} \tag{3.98}$$

Furthermore, (x, y) is the unique solution of (D) corresponding to $((x_0, y_0), u)$.

Proof. We follow the same proof of Lemma 3.1.5 using the normal cone formula in (3.96) instead of (3.12), Lemma 3.3.1 instead of Lemma 3.1.1, and the prox-regularity constant $\frac{2\eta}{L_\psi}$ provided by Lemma 3.3.2 instead of ρ. □

We now introduce the following notations that are going to be used in our proofs.

- Recall from (3.17) that $\mu := L_\psi(1 + M_h)$. Define a sequence $(\gamma_k)_k$ such that, for all $k \in \mathbb{N}$, $\gamma_k > \dfrac{2\mu}{\eta^2} e$ and $\gamma_k \to \infty$ as $k \longrightarrow \infty$, and the real sequences $(\alpha_k)_k$,and $(\sigma_k)_k$ by
$$\alpha_k := \frac{1}{\gamma_k} \ln \left(\frac{\eta^2 \gamma_k}{2\mu} \right); \ \sigma_k := \frac{r L_\psi}{2\eta^2} \left(\frac{\ln(r)}{\gamma_k} + \alpha_k \right). \tag{3.99}$$
Our choice of γ_k yields that
$$\gamma_k e^{-\gamma_k \alpha_k} = \frac{2\mu}{\eta^2}, \ \ (\alpha_k, \sigma_k) > 0 \ \ \forall\, k \in \mathbb{N}, \ \ \alpha_k \searrow 0, \ \sigma_k \searrow 0. \tag{3.100}$$
- For each $t \in [0, T]$ and $k \in \mathbb{N}$, we define

$$C^{\gamma_k}(t) := \left\{ x \in \mathbb{R}^n : \sum_{i=1}^r e^{\gamma_k \psi_i(t,x)} \le 1 \right\} \subset \operatorname{int} C(t) \text{ for } r > 1, \ \& \ C^{\gamma_k}(t) := C(t) \ \text{ for } r = 1, \tag{3.101}$$

$$C^{\gamma_k}(t, k) := \left\{ x \in \mathbb{R}^n : \sum_{i=1}^r e^{\gamma_k \psi_i(t,x)} \le \frac{2\mu}{\eta^2 \gamma_k} = e^{-\alpha_k \gamma_k} \right\} \subset \operatorname{int} C^{\gamma_k}(t). \tag{3.102}$$

Under the assumptions of Lemma 3.3.2, Proposition 3.2.10 and Remark 3.2.11 hold true in a global setting, that is, the ball around $\bar{x}$ is now omitted in those statements. More precisely, the summations therein are only up to r instead of $r+1$ terms, with $C^{\gamma_k}(t,k)$, $C^{\gamma_k}(t)$, and $C(t)$ replacing $\bar{C}^{\gamma_k}(t,k)$, $\bar{C}^{\gamma_k}(t)$, and $C(t)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$, respectively. For the convenience of our readers, we present those results here.

Proposition 3.3.5. Assume that $\psi_i(\cdot,\cdot)$ is continuous and, for all $t\in[0,T]$, the set $C(t)$ is nonempty, closed, and given by (3.3). Assume that $(A3.1)_G$ and $(A3.2)_G$ hold, and that Gr $C(\cdot)$ is compact. The following holds true.

(*i*) There exist $k_1\in\mathbb{N}$ and $r_1\in(0,\frac{\rho_o}{2}]$, such that $\forall k\geq k_1$, $\forall (t,x)\in\{(t,x)\in[0,T]\times\mathbb{R}^n : \sum_{i=1}^r e^{\gamma_k\psi_i(t,x)}=1\}$, and $\forall\,(\tau,z)\in B_{2r_1}(t,x)$, we have

$$\left\|\sum_{i=1}^r e^{\gamma_k\psi_i(\tau,z)}\nabla_x\psi_i(\tau,z)\right\| > 2\eta\sum_{i=1}^r e^{\gamma_k\psi_i(\tau,z)}. \tag{3.103}$$

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

$$\left[x\in C^{\gamma_k}(t)\ \ \&\ \|\sum_{i=1}^r e^{\gamma_k\psi_i(t,x)}\nabla_x\psi_i(t,x)\|\leq\eta\sum_{i=1}^r e^{\gamma_k\psi_i(t,x)}\right]\Longrightarrow\left[\sum_{i=1}^r e^{\gamma_k\psi_i(t,x)}<e^{-\epsilon_o\gamma_k}\right]. \tag{3.104}$$

(*iii*) For all $t\in[0,T]$, for all k, $C^{\gamma_k}(t)\subset\operatorname{int}C(t)$ for $r>1$ and $C^{\gamma_k}(t,k)\subset\operatorname{int}C^{\gamma_k}(t)$, and these sets are uniformly compact. Moreover, there exists $k_3\in\mathbb{N}$ such that for $k\geq k_3$, these sets are the closure of their interiors, their boundaries and interiors are non-empty, and the formulae for their respective boundaries and interiors are obtained from their own definitions in (3.101) and (3.102) by replacing the inequalities therein by equalities and strict inequalities, respectively. Furthermore, $C^{\gamma_k}(t)$ and $C^{\gamma_k}(t,k)$ are amenable, epi-Lipschitz, and are respectively $\frac{\eta}{L_\psi}$- and $\frac{\eta}{2L_\psi}$-prox-regular.

(*iv*) For every $t\in[0,T]$, $(C^{\gamma_k}(t))_k$ and $(C^{\gamma_k}(t,k))_k$ are nondecreasing sequences whose Painlevé-Kuratowski limit is $C(t)$ and satisfy

$$\operatorname{int}C(t)=\bigcup_{k\in\mathbb{N}}\operatorname{int}C^{\gamma_k}(t)=\bigcup_{k\in\mathbb{N}}C^{\gamma_k}(t)=\bigcup_{k\in\mathbb{N}}\operatorname{int}C^{\gamma_k}(t,k)=\bigcup_{k\in\mathbb{N}}C^{\gamma_k}(t,k). \tag{3.105}$$

(*v*) For $c\in\operatorname{bdry}C(0)$, there exist $k_c\geq k_3$, $r_c>0$, and a vector $d_c\neq 0$ such that

$$\left(\left[C(0)\cap\bar{B}_{r_c}(c)\right]+\sigma_k\frac{d_c}{\|d_c\|}\right)\subset\operatorname{int}C^{\gamma_k}(0,k),\quad\forall k\geq k_c.$$

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

$$\left(c + \sigma_k \frac{d_c}{\|d_c\|}\right) \in \operatorname{int} C^{\gamma_k}(0,k). \tag{3.106}$$

Remark 3.3.6. We deduce, from Proposition 3.3.5, that for any $c \in C(0)$, there exists a sequence $(c_{\gamma_k})_k$ such that, for k large enough, $c_{\gamma_k} \in \operatorname{int} C^{\gamma_k}(0,k) \subset \operatorname{int} C^{\gamma_k}(0)$, and $c_{\gamma_k} \longrightarrow c$. Indeed:

(*i*) For $c \in \operatorname{bdry} C(0)$, we choose $c_{\gamma_k} := c + \sigma_k \frac{d_c}{\|d_c\|}$ for all k. For $k \geq k_c$, we have from (3.106) that $c_{\gamma_k} \in \operatorname{int} C^{\gamma_k}(0,k)$. Moreover, since $\sigma_k \longrightarrow 0$ we have $c_{\gamma_k} \longrightarrow c$.

(*ii*) For $c \in \operatorname{int} C(0)$, Proposition 3.3.5(*iv*) yields the existence of $\hat{k}_c \in \mathbb{N}$, such that $c \in \operatorname{int} C^{\gamma_k}(0,k)$ for all $k \geq \hat{k}_c$. Hence, there exists $\hat{r}_c > 0$ satisfying

$$c \in \bar{B}_{\hat{r}_c}(c) \subset \operatorname{int} C^{\gamma_k}(0,k), \quad \forall k \geq \hat{k}_c.$$

In this case, we take the sequence $c_{\gamma_k} \equiv c \in \operatorname{int} C^{\gamma_k}(0,k)$ that converges to c.

3.3.2 Existence and uniqueness of solution corresponding to (D)

We now prove Theorem 3.3.7, which says that under global assumptions, the Cauchy problem corresponding to (D) has a unique solution that is Lipschitz. Similar to the proof of Corollary 3.2.16, the proof of Theorem 3.3.7 follows closely the arguments used to prove Theorem 3.2.14 after removal of the truncation on $C(t)$ and $\mathbb{R}^l$. However, doing so requires important modifications. For instance, removing the truncation on $C(t)$ in the set $\mathscr{D}$ defined in (3.84), makes it unsuitable for the global setting, and hence, it will have to be redefined (see (3.108)). This discrepancy is due to having $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ always generated by at least *two* functions, and hence, $\bar{C}^{\gamma_k}(t) \subset \operatorname{int} C(t)$ is always valid. While in the global setting, for the case $r = 1$, (3.101) yields that $C^{\gamma_k}(t) = C(t)$ and hence, $\mathscr{D}$ must be modified to include Gr $C(\cdot)$.

Theorem 3.3.7 (**Existence & uniqueness of Lipschitz solutions for** (D))**.** Assume that $\psi_i(\cdot,\cdot)$ continuous, and, for $t \in [0,T]$, $C(t)$ is non-empty, closed, and given by (3.3). Assume that $(A3.1)_G$, $(A3.2)_G$ and $(A4.1)_G$ are satisfied, and that $\operatorname{Gr} C(\cdot)$ is bounded.

Given $(x_0, y_0) \in C(0) \times \mathbb{R}^l$ and $u \in \mathcal{U}$, the Cauchy problem corresponding to (D) and $((x(0), y(0)); u) = ((x_0, y_0); u)$ has a unique solution (x, y), which is Lipschitz and is the uniform limit of a subsequence (not relabeled) of $(x_{\gamma_k}, y_{\gamma_k})_k$, where $(x_{\gamma_k}, y_{\gamma_k})$ is the solution of a standard control system corresponding to u with $x_{\gamma_k}(t) \in \operatorname{int} C(t)$, for all $t \in [0, T]$.

Proof. We denote by M_C the bound of Gr $C(\cdot)$. Consider the Cauchy problem (D) corresponding to $((x(0), y(0)); u) = ((x_0, y_0); u) \in (C(0) \times \mathbb{R}^l) \times \mathcal{U}$. The existence of a solution that is Lipschitz and unique, will be shown by approximating (D) with (D_{γ_k}), defined below as the global version of $(\bar{D}_{\gamma_k})$. Let $c_{\gamma_k} \in C^{\gamma_k}(0, k)$ be the sequence from Remark 3.3.6 corresponding (and converging) to $c = x_0$. We now proceed with the proof by imitating the same steps of the proof of Theorem 3.2.14, in which we employ $(c_{\gamma_k}, d_{\gamma_k}) := (c_{\gamma_k}, y_0)$ and we make the following notable modifications. Using Remark 3.3.3, we can extend $h = (f, g)(t, \cdot, y, u)$ from $C(t)$ to $C(t) + \frac{\eta}{L_\psi} B$ such that h satisfy $(\text{A4.1})_G$, where $C(t)$ is now replaced by $C(t) + \frac{\eta}{L_\psi} B$. For fixed $k \in \mathbb{N}$ large enough, we consider the system (D_{γ_k}) corresponding to $((c_{\gamma_k}, y_0), u)$ to be

$$(D_{\gamma_k}) \begin{cases} \dot{x}(t) = f(t, x(t), y(t), u(t)) - \sum_{i=1}^{r} \gamma_k e^{\gamma_k \psi_i(t, x(t))} \nabla_x \psi_i(t, x(t)), \;\; \text{a.e. } t \in [0, T], \\ \dot{y}(t) = g(t, x(t), y(t), u(t)), \;\text{a.e. } \; t \in [0, T]. \end{cases} \tag{3.107}$$

This system is well defined on the following modified version of the set $\mathscr{D}$, given in (3.84),

$$\mathscr{D}_G := \{(t, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^l : \;\; x \in C(t) + \frac{\eta}{L_\psi} B\}. \tag{3.108}$$

As $(0, c_{\gamma_k}, y_0) \in \mathscr{D}_G$, we follow steps similar to the ones used to reach (3.85), and we deduce that a solution $(x_{\gamma_k}, y_{\gamma_k})$ of (D_{γ_k}) with $(x(0), y(0)) = (c_{\gamma_k}, y_0)$ exists on the interval $[0, \hat{T}_G) \subset [0, T]$, and $(t, x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \mathscr{D}_G$, $\forall t \in [0, \hat{T}_G)$, where $\hat{T}_G$ is given by

$$\hat{T}_G := \sup\{T_1 : \;\; (x, y) \text{ solves } (D_{\gamma_k}) \text{ on } [0, T_1] \text{ with } (x(0), y(0)) = (c_{\gamma_k}, y_0) \\ \text{and } (t, x(t), y(t)) \in \mathscr{D}_G \;\; \forall t \in [0, T_1]\}. \tag{3.109}$$

Simlarily to Step I.2 in the proof of Theorem 3.2.14, we conclude that $x_{\gamma_k}(t) \in C^{\gamma_k}(t)$ for all $t \in [0, \hat{T}_G)$. On the other hand, since φ is absent in (D_{γ_k}) and g is bounded by

M_h (from (A4.1)$_G$), we immediately obtain that, for $t \in [0, \hat{T}_G)$, $y_{\gamma_k}(t) \in M_0\bar{B}$, where $M_0 := \|y_0\| + M_h T$. The boundedness of Gr $C(\cdot)$ by M_C guarantees that the solution remains in the bounded set $M_C\bar{B} \times M_0\bar{B}$ and hence, $\hat{T}_G = T$. By mimicking Step I.3. of the proof of Theorem 3.2.14, we obtain the invariance of $C^{\gamma_k}(t,k)$ in x for (D_{γ_k}), and hence, our solution $(x_{\gamma_k}, y_{\gamma_k})$ for the Cauchy problem (D_{γ_k}) with $((c_{\gamma_k}, y_0), u)$, also satisfies $(x_{\gamma_k}(t), y_{\gamma_k}(t)) \in C^{\gamma_k}(t,k) \times M_0\bar{B}$ for all $t \in [0,T]$. Thus, the definition of $C^{\gamma_k}(t,k)$ and $\xi^i_{\gamma_k}$ $(i = 1, \cdots, r)$ given in (3.102) and (3.76), respectively, and $\xi_{\gamma_k}(\cdot) := \sum_{i=1}^r \xi^i_{\gamma_k}(\cdot)$, yield that $\|\xi_{\gamma_k}\|_\infty \leq \frac{2\mu}{\eta^2}$. Employing this bound of ξ_{γ_k} in (D_{γ_k}), we obtain that $\max\{\|\dot{x}_{\gamma_k}\|_\infty, \|\dot{y}_{\gamma_k}\|_\infty\} \leq M_h + \frac{2\mu}{\eta^2} L_\psi$. It follows that (.1) holds for $\mathcal{R} := r$ and $(x_k, y_k, \xi^i_k, \zeta_k) := (x_{\gamma_k}, y_{\gamma_k}, \xi^i_{\gamma_k}, 0)$. Whence, Lemma .0.2$(i)$ together with Proposition 3.3.5(iv) implies that a subsequence of $(x_{\gamma_k}, y_{\gamma_k})$, and $\xi^i_{\gamma_k}$ $(\forall i = 1, \cdots, r)$ converge respectively to some $(x,y) \in W^{1,\infty}([0,T]; \mathbb{R}^n \times \mathbb{R}^l)$, and $(\xi^1, \cdots, \xi^r) \in L^\infty([0,T]; \mathbb{R}^r_+)$, satisfying

$$(x_{\gamma_k}, y_{\gamma_k}) \xrightarrow{unif} (x,y), \quad (\dot{x}_{\gamma_k}, \dot{y}_{\gamma_k}) \xrightarrow[in\ L^\infty]{w*} (\dot{x}, \dot{y}), \quad \xi^i_{\gamma_k} \xrightarrow[in\ L^\infty]{w*} \xi^i \ (\forall i = 1, \cdots, r),$$

$$(x(t), y(t)) \in C(t) \times M_0\bar{B} \quad \forall t \in [0,T]; \ \max\{\|\dot{x}\|_\infty, \|\dot{y}\|_\infty\} \leq M_h + \frac{2\mu}{\eta^2} L_\psi; \ \|\sum_{i=1}^r \xi^i\|_\infty \leq \frac{2\mu}{\eta^2},$$

$$\xi^i(t) = 0 \text{ for all } t \in I^-_i(x), \ \ i \in \{1, \cdots, r\}, \quad \xi(t) := \sum_{i=1}^r \xi^i = 0 \text{ for all } t \in I^-(x),$$

the validation of the last equations is similar to that for (3.80). We now apply the dominated convergence theorem to (D_{γ_k}) at $((x_{\gamma_k}, y_{\gamma_k}), u_{\gamma_k} := u)$ (as done in the proof of Case 1 in Lemma .0.2 (ii)), and we deduce that $((x,y), u)$ and $\lambda_i = \xi^i$ satisfy (3.97). By means of Lemma 3.3.4, we conclude that (x,y) is the *unique* solution of (D) corresponding to $((x_0, y_0), u)$ and is *Lipschitz*. □

The following table summarizes the results of Section 3.3.

Table 3.4 Summary of results from Section 3.3 .

Result	**Description**
Lemma 3.3.1	We use the compactness of Gr $C(\cdot)$ to produce an equivalence between $(A3.2)_G$ and a global version of condition (3.10), namely, (3.95).
Lemma 3.3.2	We use Lemma 3.3.1 to obtain the uniform prox-regularity of $C(t)$, as well as a formula for the normal cone to $C(t)$.
Remark 3.3.3	We extend h to a function that satisfies $(A4.1)_G$ on $C(t) + \frac{\eta}{L_\psi}B$.
Lemma 3.3.4	We use Lemma 3.3.1 and Lemma 3.3.2 to establish the Lipschitz continuity and the uniqueness of the solutions for the Cauchy problem of (D) via its equivalent form.
Proposition 3.3.5	We derive properties for our sets $C^{\gamma_k}(t)$ and $C^{\gamma_k}(t,k)$.
Remark 3.3.6	We approximate any $c \in C(0)$ by a sequence $c_{\gamma_k} \in \text{int}\, C^{\gamma_k}(0,k) \subset \text{int}\, C^{\gamma_k}(0)$ such that $c_{\gamma_k} \longrightarrow c$.
Theorem 3.3.7	We prove that the Cauchy problem corresponding to (D) has a unique solution that is Lipschitz.

CHAPTER 4

OPTIMAL CONTROL PROBLEM (P) OVER A COUPLED SWEEPING PROCESS DYNAMIC (D)

The aim of this chapter is to derive global existence of optimal solutions and necessary conditions in the form of a maximum principle for a strong local minimizer of the fixed time Mayer problem (P) given by the following:

$$(P)\begin{cases}\text{minimize} \quad J(x(0),y(0),x(T),y(T)) \\ \text{over } ((x,y),u)\in W^{1,1}([0,T],\mathbb{R}^n\times\mathbb{R}^l)\times\mathcal{U} \text{ such that} \\ \quad (D)\begin{cases}\dot{x}(t)\in f(t,x(t),y(t),u(t))-N_{C(t)}(x(t)), \text{ a.e. } t\in[0,T], \\ \dot{y}(t)=g(t,x(t),y(t),u(t)), \text{ a.e. } t\in[0,T],\end{cases} \\ \quad (x(0),y(0),x(T),y(T))\in S, \qquad \textbf{(B.C.)}\end{cases}$$

where $T>0$ is fixed, $J:\mathbb{R}^n\times\mathbb{R}^l\times\mathbb{R}^n\times\mathbb{R}^l\longrightarrow\mathbb{R}\cup\{\infty\}$, $f:[0,T]\times\mathbb{R}^n\times\mathbb{R}^l\times\mathbb{R}^m\longrightarrow\mathbb{R}^n$, $g:[0,T]\times\mathbb{R}^n\times\mathbb{R}^l\times\mathbb{R}^m\longrightarrow\mathbb{R}^l$, $C(t)$ is the intersection of the zero-sublevel sets of a finite sequence of functions $\psi_i(t,\cdot)$ where $\psi_i:[0,T]\times\mathbb{R}^n\longrightarrow\mathbb{R}$, $i=1,\dots,r$, $N_{C(t)}$ is the *Clarke* normal cone to $C(t)$, $S\subset C(0)\times\mathbb{R}^l\times\mathbb{R}^n\times\mathbb{R}^l$ is *closed*, $U(\cdot):[0,T]\rightsquigarrow\mathbb{R}^m$ is nonempty, closed, and Lebesgue- measurable set-valued map, and the set of control functions $\mathcal{U}$ is defined by

$$\mathcal{U}:=\{u:[0,T]\longrightarrow\mathbb{R}^m : u \text{ is measurable and } u(t)\in U(t), \text{ a.e. } t\in[0,T]\}. \tag{4.1}$$

A pair $((x,y),u)$ is *admissible* for (P) if $((x,y),u)\in W^{1,1}([0,T];\mathbb{R}^n\times\mathbb{R}^l)\times\mathcal{U}$ satisfies the dynamic (D) and the boundary conditions **(B.C.)**.

An admissible pair $((\bar{x},\bar{y}),\bar{u})$ is said to be a $\bar{\delta}$-*strong local minimizer* for (P), for some $\bar{\delta}>0$, if for all $((x,y),u)$ admissible for (P) and satisfying $\|(x,y)-(\bar{x},\bar{y})\|_\infty\le\bar{\delta}$, we have

$$J(\bar{x}(0),\bar{y}(0),\bar{x}(T),\bar{y}(T))\le J(x(0),y(0),x(T),y(T)).$$

4.1 Existence of optimal solution for (P) under global assumptions

In this section, we demonstrate the *global* existence of an optimal solution for (P) when the global assumptions are satisfied, see Theorem 4.1.1. Recall assumptions (A1), $(\text{A3.1})_G$, (A3.2)$_G$, and (A4)$_G$ from Chapter 3.

Theorem 4.1.1 (Global existence of optimal solutions for (P)**).** Assume that $(A1)$ holds, $\psi_i(\cdot,\cdot)$ continuous, and, for $t \in [0,T]$, $C(t)$ is non-empty, closed, and given by (3.3). Assume that $(A3.1)_G$, $(A3.2)_G$ and $(A4)_G$ are satisfied, and that $\operatorname{Gr} C(\cdot)$ and $\pi_2(S)$ are bounded, where π_2 is the projection of S into the second component. Let $J : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R} \cup \{\infty\}$ be merely lower semicontinuous. Then, (P) has a global optimal solution if and only if it has at least one admissible pair $((x, y), u)$ with $(x(0), y(0), x(T), y(T)) \in \operatorname{dom} J$.

Proof. Let M_C and $M_{\pi_2(S)}$ be the bounds of Gr $C(\cdot)$ and $\pi_2(S)$, respectively. Observe that $((x, y), u)$ is admissible for (P) is *equivalent* to $((x, y), u)$ solving (D) and

$$(x(0), y(0), x(T), y(T)) \in S_G := S \cap (C(0) \times M_{\pi_2(S)} \times C(T) \times M\bar{B}),$$

where $M := M_{\pi_2(S)} + TM_h$.

Since (P) has an admissible solution with $(x(0), y(0), x(T), y(T)) \in \operatorname{dom} J$, J is lower semicontinuous, and S_G is compact, then the infimum of J over $((x, y), u)$ satisfying (D) and (**B.C.**) exists (see Lemma 2.4.4). Let $((x_n, y_n), u_n)$ be a minimizing sequence for (P). Then, for each $n \in \mathbb{N}$, $((x_n, y_n), u_n)$ satisfies (D), starting at $(x_0, y_0) := (x_n(0), y_n(0))$, and (**B.C.**). Hence, by Lemma 3.3.4, there exists a sequence of $(\lambda_i^n)_{i=1}^r$ such that, for all $i = 1, \cdots, r$, $\lambda_i^n \in L^\infty([0, T]; \mathbb{R}_+)$, $\lambda_i^n = 0$ on $I_i^-(x_n)$, and, $(\lambda_i^n)_{i=1}^r$ with $((x_n, y_n), u_n)$ satisfy (3.97) and the bounds in (3.98). Apply lemma .0.2(i) to $\mathcal{R} := r$, (x_n, y_n), $\xi_n^i := \lambda_i^n$, $\zeta_n := 0$, $M_1 := \max\{M_C, M\}$, $M_2 := M_h + \frac{\mu}{4\eta^2} L_\psi$, and $M_3 := \frac{\mu}{4\eta^2}$, we obtain $(\hat{x}, \hat{y}) \in W^{1,\infty}$, $(\lambda_1, \cdots, \lambda_r) \in L^\infty([0, T]; \mathbb{R}_+^r)$, $\zeta := 0$, and subsequences (not relabeled) of $(x_n, y_n)_n$ and $(\lambda_i^n)_n$ such that $(x_n, y_n) \xrightarrow{unif} (\hat{x}, \hat{y}), (\dot{x}_n, \dot{y}_n) \xrightarrow[in\ L^\infty]{w*} (\dot{\hat{x}}, \dot{\hat{y}}), \lambda_i^n \xrightarrow[in\ L^\infty]{w*} \lambda_i$, for all $i = 1, \cdots, r$, and $(\dot{\hat{x}}, \dot{\hat{y}})$ and $(\lambda_1, \cdots, \lambda_r)$ satisfy the bounds in (.2). Furthermore, we have $(\hat{x}(0), \hat{y}(0), \hat{x}(T), \hat{y}(T)) \in S$.

On the other hand, as $((x_n, y_n), u_n)$ and $(\lambda_i^n)_{i=1}^r$ satisfy (3.97), this means that they solve (.3) for $\zeta := 0$, $q_i := \psi_i$, and $Q(t) := C(t)$. Noting that $(A1)$ and $(A4.2)_G$ hold, then by applying the global version of Lemma .0.2 (ii) (see Remark .0.3), we obtain $\hat{u} \in \mathcal{U}$ such that $((\hat{x}, \hat{y}), \hat{u})$ and $(\lambda_i)_{i=1}^r$ also satisfy (.3), which is (3.97). Thus, to prove that $((\hat{x}, \hat{y}), \hat{u})$ is admissible for

(D), it suffices by the equivalence in Lemma 3.3.4 to show that for all $i = 1, \cdots, r$, λ_i is supported on $I_i^0(\hat{x})$, knowing that for all $i = 1, \cdots, r$, $\lambda_i^n(t) = 0$ for $t \in I_i^-(x_n)$. Fix $t \in I_i^-(\hat{x})$, then, $\psi_i(t, \hat{x}(t)) < 0$. Since x_n converges uniformly to $\hat{x}$ and $\psi_i(\cdot,\cdot)$ is continuous, we can find $\hat{\delta} > 0$ and $\hat{n} \in \mathbb{N}$ such that $\forall s \in (t-\hat{\delta}, t+\hat{\delta}) \cap [0,T]$ and for all $n \geq \hat{n}$, we have $\psi_i(s, x_n(s)) < 0$ and hence $\lambda_i^n(s) = 0$. Thus, as $n \to \infty$, $0 = \lambda_i^n(\cdot) \to 0$ on $(t - \hat{\delta}, t + \hat{\delta}) \cap [0,T]$, and so $\lambda_i(t) = 0$. Therefore, Lemma 3.3.4 yields that $((\hat{x}, \hat{y}), \hat{u})$ is admissible for (D). Using the lower semicontinuinity of J, we deduce that

$$\begin{aligned} J(\hat{x}(0), \hat{y}(0), \hat{x}(T), \hat{y}(T)) &\leq \lim_{n\to\infty} J(x_n(0), y_n(0), x_n(T), y_n(T)) \\ &= \inf_{((x,y),u) \text{ admissible for } (P)} J(x(0), y(0), x(T), y(T)), \end{aligned}$$

showing that $((\hat{x}, \hat{y}), \hat{u})$ is optimal for (P) over all admissible pairs $((x,y),u)$. □

Table 4.1 Summary of results from Section 4.1 .

Result	**Description**
Theorem 4.1.1	We demonstrate the *global* existence of an optimal solution for (P).

4.2 Pontryagin maximum principle for (P) under local assumptions

In this section, we present the maximum principle for the problem (P). We employ a modification of the exponential penalization technique used in [30, 70, 55] for special cases of (P). We first approximate the given optimal solution of (P) with optimal solutions for some approximating problems having *joint*-endpoint constraints, $(P^{\alpha,\beta}_{\gamma_k})$, which are standard optimal control problems involving exponential penalty terms (Proposition 4.2.8). Then, we find necessary conditions for the approximating problems (Proposition 4.2.9), and we finally conclude the necessary conditions for (P) by taking the limit of the necessary conditions for $(P^{\alpha,\beta}_{\gamma_k})$.

For a given pair $(\bar{x}, \bar{y}) \in \mathcal{C}([0,T]; \mathbb{R}^n \times \mathbb{R}^l)$ such that $\bar{x}(t) \in C(t)$ $\forall t \in [0,T]$, and for a constant $\bar{\delta} > 0$, we adopt all the local assumptions introduced at the beginning of Chapter

3 and introduce two additional ones. We say that the following assumptions hold true at $((\bar{x},\bar{y});\bar{\delta})$ if the corresponding conditions hold true.

(**A3.3**) There exists a positive Lipschitz function $\bar{\beta}(\cdot) = (\bar{\beta}_1(\cdot),\cdots,\bar{\beta}_r(\cdot)) : [0,T] \longrightarrow \mathbb{R}^r$ such that

$$\sum_{\substack{j\in\mathcal{I}^0_{(t,\bar{x}(t))}\\ j\neq i}} \bar{\beta}_j(t)|\langle\nabla_x\psi_j(t,\bar{x}(t)),\nabla_x\psi_i(t,\bar{x}(t))\rangle| \; < \bar{\beta}_i(t)\|\nabla_x\psi_i(t,\bar{x}(t))\|^2,\;\; \forall t\in I^0(\bar{x}),\;\; \forall i\in\mathcal{I}^0_{(t,\bar{x}(t))}.$$

For $a>0, b>0$ given, we recall the following set, given in (3.9), by

$$\bar{\mathscr{N}}_{(a,b)}(t) := \Big[C(t)\cap\bar{B}_a(\bar{x}(t))\Big]\times\bar{B}_b(\bar{y}(t)),\quad \text{for}\;\; t\in[0,T],$$

and we introduce a new set

$$\bar{\mathscr{B}}_a := \bar{B}_a(\bar{x}(0))\times\bar{B}_a(\bar{y}(0))\times\bar{B}_a(\bar{x}(T))\times\bar{B}_a(\bar{y}(T)). \tag{4.2}$$

(**A5**) **Local assumption on** J **at** $((\bar{x},\bar{y});\bar{\delta})$**:** There exist $\rho_1>0$ and $L_J>0$ such that J is L_J-Lipschitz on $S(\bar{\delta})$, where

$$S(\bar{\delta}) := \left(\Big[S\cap\bar{\mathscr{B}}_{\bar{\delta}}\Big]+\rho_1\bar{B}\right)\cap\left(\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(0)\times\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(T)\right).$$

4.2.1 Preliminary results

We start the first subsection by presenting consequences of $(A3.3)$ that shall be crucial for our approximating problem and the proof of the maximum principle. For this subsection, let $C(\cdot)$ satisfying $(A2)$ for $\rho>0$. Consider $\bar{x}\in\mathcal{C}([0,T];\mathbb{R}^n)$ with $\bar{x}(t)\in C(t)$ for all $t\in[0,T]$, and $\bar{\delta}>0$ such that $(A3.1)$ holds at $(\bar{x};\bar{\delta})$.

The next remark discusses the significance of (A3.3) in the proof of the maximum principle. In particular, it highlights why it is sufficient to prove the maximum principle (Theorem 4.2.11) under a *stronger* assumption.

Remark 4.2.1 (**Assumption (A3.3)**)**.** Note that when $r=1$, the sets $C(t)$ are smooth and condition $(A3.3)$ trivially holds. Let $r>1$, then the sets $C(t)$ are nonsmooth. In this case, a condition closely related to $(A3.3)$, see [46, Theorem 1.3.1], has been first mentioned in [40] to be useful when sweeping (or reflected) processes over *nonsmooth* sets are studied. For $t \in I^0(\bar{x})$, denote by $\mathcal{G}_\psi(t)$ the Gramian matrix of the vectors $\{\nabla_x \psi_i(t, \bar{x}(t)) : i \in \mathcal{I}^0_{(t,\bar{x}(t))}\}$, i.e. $(\mathcal{G}_\psi(t))_{ij} = \langle \nabla_x \psi_i(t, \bar{x}(t)), \nabla_x \psi_j(t, \bar{x}(t))\rangle$.

- If for all $i \in \mathcal{I}^0_{(t,\bar{x}(t))}$, we have $(A3.3)$ holds for $\bar{\beta}_i(t) \equiv 1$, then the matrix $\mathcal{G}_\psi(t)$ is *strictly diagonally dominant* (see Definition 2.1.1) .
- For the general case, $(A3.3)$ says that for some positive Lipschitz vector function, $\bar{\beta}(\cdot)$, the matrix $\mathcal{G}_\psi(t) D_{\bar{\beta}(t)}(t)$ is strictly diagonally dominant for all $t \in I^0(\bar{x})$, where $D_{\bar{\beta}(t)}(t)$ is the diagonal matrix whose diagonal entries are $(\bar{\beta}_i(t))_{i \in \mathcal{I}^0_{(t,\bar{x}(t))}}$, and $(\mathcal{G}_\psi(t) D_{\bar{\beta}(t)}(t))_{ij} = \bar{\beta}_j(t)\langle \nabla_x \psi_i(t, \bar{x}(t)), \nabla_x \psi_j(t, \bar{x}(t))\rangle$.

Thus,

(i) $(A3.3)$ yields that the vectors $\{\nabla_x \psi_i(t, \bar{x}(t)) : i \in \mathcal{I}^0_{(t,\bar{x}(t))}\}$ are linearly independent, and hence, when $(A3.3)$ is assumed to hold, $(A3.2)$ is automatically satisfied;

(ii) Setting $\tilde{\psi}_i(t, x) := \bar{\beta}_i(t)\psi_i(t, x)$, it easily follows that $C(t)$ is also the zero-sublevel sets of $(\tilde{\psi}_i(t, \cdot))_{i=1}^r$, for $i = 1, \cdots, r$, for some $L_{\tilde{\psi}} > 0$, $\tilde{\psi}_i$ satisfies $(A3.1)$ for all $i = 1, \cdots, r$, and condition $(A3.3)$ is equivalent to saying that for $t \in I^0(\bar{x})$, the Gramian matrix $\mathcal{G}_{\tilde{\psi}}(t)$ of the vectors $\{\nabla_x \tilde{\psi}_i(t, \bar{x}(t)) : i \in \mathcal{I}^0_{(t,\bar{x}(t))}\}$ is strictly diagonally dominant; a fact that shall be used in the proof of the maximum principle;

(iii) From parts $(i)-(ii)$ of this remark, we have $\tilde{\psi}_1, \cdots, \tilde{\psi}_r$ satisfy $(A3.2)$, and hence, (3.34) of Lemma 3.2.4 is valid at $\tilde{\psi}_1, \cdots, \tilde{\psi}_r, \psi_{r+1}$ when replacing $\bar{\eta}$ by $\tilde{\eta} := \bar{\eta}\, \mathbf{b}_{\bar{\beta}}$, where

$$\mathbf{b}_{\bar{\beta}} := \min\left\{1, \min\{\bar{\beta}_i(t) : t \in [0, T],\ i = 1, \cdots, r\}\right\}.$$

Equivalent forms for the strict diagonally dominance of $\mathcal{G}_\psi(t)$ are given in the following lemma.

Lemma 4.2.2. The following assertions are equivalent:

(i) For all $t \in I^0(\bar{x})$, the Gramian matrix $\mathcal{G}_\psi(t)$ of the vectors $\{\nabla_x \psi_i(t, \bar{x}(t)) : i \in \mathcal{I}^0_{(t,\bar{x}(t))}\}$, is strictly diagonally dominant.

(*ii*) There exists $b \in (0,1)$ such that, for all $t \in I^0(\bar{x})$ and for all $i \in \mathcal{I}^0_{(t,\bar{x}(t))}$, we have

$$\sum_{\substack{j \in \mathcal{I}^0_{(t,\bar{x}(t))} \\ j \neq i}} |\langle \nabla_x \psi_j(t,\bar{x}(t)), \nabla_x \psi_i(t,\bar{x}(t)) \rangle| \leq b \|\nabla_x \psi_i(t,\bar{x}(t))\|^2. \tag{4.3}$$

(*iii*) There exist $\bar{c} > 0$, $\bar{b} \in (0,1)$, and $\bar{a} > 0$ such that $\forall (t,x) \in \operatorname{Gr} C(\cdot) \cap \bar{B}_{\bar{c}}(\bar{x}(\cdot))$ with $\mathcal{I}^{\bar{a}}_{(t,x)} \neq \emptyset$, and $\forall\, i \in \mathcal{I}^{\bar{a}}_{(t,x)}$, we have

$$\sum_{\substack{j \in \mathcal{I}^{\bar{a}}_{(t,x)} \\ j \neq i}} |\langle \nabla_x \psi_j(t,x), \nabla_x \psi_i(t,x) \rangle| \leq \bar{b} \|\nabla_x \psi_i(t,x)\|^2, \tag{4.4}$$

$$\text{where } \mathcal{I}^{\bar{a}}_{(t,x)} := \{i \in \{1,\ldots,r\} : -\bar{a} \leq \psi_i(t,x) \leq 0\}. \tag{4.5}$$

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

$$b(t,i) := \frac{1}{\|\nabla_x \psi_i(t,\bar{x}(t))\|^2} \sum_{\substack{j \in \mathcal{I}^0_{(t,\bar{x}(t))} \\ j \neq i}} |\langle \nabla_x \psi_j(t,\bar{x}(t)), \nabla_x \psi_i(t,\bar{x}(t)) \rangle| < 1, \tag{4.6}$$

and set $b := \sup \Big\{ b(t,i) : t \in I^0(\bar{x}) \text{ and } i \in \mathcal{I}^0_{(t,\bar{x}(t))} \Big\}$. Then, (4.3) holds true. To show that $b < 1$, use an argument by contradiction, together with Lemma .0.1 and inequality (4.6).

(*ii*) $\Longrightarrow$ (*iii*): We fix $\bar{b} \in (b,1)$ and we use an argument by contradiction in conjunction with Lemma .0.1.

(*iii*) $\Longrightarrow$ (*i*): Follows directly by taking $\bar{a} := \bar{c} := 0$, $x = \bar{x}(t)$ and using $\bar{b} < 1$. $\square$

The following result, which will be used in the proof of the maximum principle, is an immediate consequence of Lemma 3.2.4 obtained via a simple argument by contradiction and the continuity in (A3.1) of $(\psi_i)_{1 \leq i \leq r}$ and $(\nabla_x \psi_i)_{1 \leq i \leq r}$ on the compact set $\operatorname{Gr} \Big(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot)\Big)$.

Lemma 4.2.3. Let $C(\cdot)$ satisfying ($A2$) for some $\rho > 0$. Consider $\bar{x} \in \mathcal{C}([0,T];\mathbb{R}^n)$ with $\bar{x}(t) \in C(t)$ for all $t \in [0,T]$, and $\bar{\delta} > 0$ such that ($A3.1$) and ($A3.2$) hold at $(\bar{x};\bar{\delta})$. Then, for $\bar{\varepsilon} \in (0,\rho) \cap (0,\varepsilon_o]$, and its corresponding ψ_{r+1} and $\bar{\eta}$ from Lemma 3.2.4, there exists $a_o > 0$ such that for all $i \in \{1,\ldots,r+1\}$ we have

$$\Big[(t,x) \in \operatorname{Gr} \Big(C(\cdot) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot)\Big) \text{ and } \|\nabla_x \psi_i(t,x)\| \leq \bar{\eta}\Big] \Longrightarrow \psi_i(t,x) < -a_o. \tag{4.7}$$

Table 4.2 Summary of results from Section 4.2.1.

Result	Description
Remark 4.2.1	We discuss the significance of (A3.3) in the proof of the maximum principle. In particular, it highlights why it is sufficient to prove the maximum principle (Theorem 4.2.11) under a *stronger* assumption.
Lemma 4.2.2	We provide equivalent forms for the strict diagonally dominance of the Gramian matrix $\mathcal{G}_\psi(t)$ of the vectors $\{\nabla_x \psi_i(t,\bar{x}(t)) : i \in \mathcal{I}^0_{(t,\bar{x}(t))}\}$.
Lemma 4.2.3	We prove that there exists a_o such that $\Big[(t,x) \in \text{Gr}\,\Big(C(\cdot)\cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(\cdot)\Big)$ and $\|\nabla_x\psi_i(t,x)\| \le \bar{\eta}\Big] \implies \psi_i(t,x) < -a_o$.

4.2.2 Study of approximating problems for (P)

Assume that $(A1)$-$(A2)$ are satisfied, and $((\bar{x},\bar{y}),\bar{u})$ is an *admissible solution* for (P) with $\bar{\delta} > 0$ such that $(A3.1)$, $(A3.2)$, $(A4)$ and $(A5)$ are satisfied at $((\bar{x},\bar{y});\bar{\delta})$. Throughout the rest of this chapter, let $\bar{\varepsilon} \in (0,\bar{\delta})$, ψ_{r+1}, $\bar{\eta}$ and $\bar{\rho} = \frac{2\bar{\eta}}{L_\psi}$ be fixed as in Subsection 3.2.2, with $L_\psi \ge \frac{4\bar{\eta}}{\rho_o}$. Let $L_{(\bar{x},\bar{y})}$ denote the Lipschitz constant of $(\bar{x},\bar{y})$, which, by Lemma 3.1.5, is *Lipschitz* and *uniquely* solves (D) corresponding to $((\bar{x}(0),\bar{y}(0)),\bar{u})$. Without loss of generality, we assume $L_{(\bar{x},\bar{y})} \ge 1$. Therefore, all the results of Subsection 3.2.2 are valid for the systems $(\bar{D})$ and $(\bar{D}_{\gamma_k})$, given respectively by (3.29) and (3.61).

For given $\delta \in (0,\bar{\varepsilon}]$, define *the problem* $(\bar{P}_\delta)$ to be the problem (P), in which (D) is replaced by $(\bar{D})$, and S is replaced by S_δ, where

$$S_\delta := S \cap \bar{\mathscr{B}}_\delta, \text{ and } \bar{\mathscr{B}}_\delta \text{ is defined in (4.2).} \tag{4.8}$$

When S_δ is replaced by the following set $S_{\delta,\delta}$,

$$S_{\delta,\delta} = S \cap [\bar{\mathscr{N}}_{(\delta,\delta)}(0) \times \bar{\mathscr{N}}_{(\delta,\delta)}(T)] \subset S(\bar{\delta}) \subset \text{domain of } J, \tag{4.9}$$

the resulting problem is named $(\bar{P}_{\delta,\delta})$.

For clarity and better visualization, we present the problems below in a structured form.

$$(P)\begin{cases}\text{minimize} \quad J(x(0),y(0),x(T),y(T))\\ \text{over } ((x,y),u)\in W^{1,1}([0,T],\mathbb{R}^n\times\mathbb{R}^l)\times\mathcal{U} \ \text{ such that}\\ \quad (D)\begin{cases}\dot{x}(t)\in f(t,x(t),y(t),u(t))-N_{C(t)}(x(t)), \ \text{a.e. } t\in[0,T],\\ \dot{y}(t)=g(t,x(t),y(t),u(t)), \ \text{a.e. } t\in[0,T],\end{cases}\\ \quad (x(0),y(0),x(T),y(T))\in S. \quad \textbf{(B.C.)}\end{cases}$$

$$(\bar{P}_\delta)\begin{cases}\text{minimize} \quad J(x(0),y(0),x(T),y(T))\\ \text{over } ((x,y),u)\in W^{1,1}([0,T],\mathbb{R}^n\times\mathbb{R}^l)\times\mathcal{U} \ \text{ such that}\\ \quad (\bar{D})\begin{cases}\dot{x}(t)\in f(t,x(t),y(t),u(t))-N_{C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x(t)), \ \text{a.e. } t\in[0,T],\\ \dot{y}(t)\in g(t,x(t),y(t),u(t))-N_{\bar{B}_{\bar{\delta}}(\bar{y}(t))}(y(t)), \ \text{a.e. } t\in[0,T].\end{cases}\\ \quad (x(0),y(0),x(T),y(T))\in S_\delta=S\cap\bar{\mathscr{B}}_\delta.\end{cases}$$

$$(\bar{P}_{\delta,\delta})\begin{cases}\text{minimize} \quad J(x(0),y(0),x(T),y(T))\\ \text{over } ((x,y),u)\in W^{1,1}([0,T],\mathbb{R}^n\times\mathbb{R}^l)\times\mathcal{U} \ \text{ such that}\\ \quad (\bar{D})\begin{cases}\dot{x}(t)\in f(t,x(t),y(t),u(t))-N_{C(t)\cap\bar{B}_{\bar{\varepsilon}}(\bar{x}(t))}(x(t)), \ \text{a.e. } t\in[0,T],\\ \dot{y}(t)\in g(t,x(t),y(t),u(t))-N_{\bar{B}_{\bar{\delta}}(\bar{y}(t))}(y(t)), \ \text{a.e. } t\in[0,T].\end{cases}\\ \quad (x(0),y(0),x(T),y(T))\in S_{\delta,\delta}=S\cap[\bar{\mathscr{N}}_{(\delta,\delta)}(0)\times\bar{\mathscr{N}}_{(\delta,\delta)}(T)]\subset S(\bar{\delta}).\end{cases}$$

Notice that $(\bar{P}_{\delta,\delta})$ and $(\bar{P}_\delta)$ have the same sets of admissible and optimal solutions.

The following is an *existence result of an optimal solution* for $(\bar{P}_\delta)$ (and, hence, of $(\bar{P}_{\delta,\delta})$) without requiring $(A5)$.

Theorem 4.2.4 (**Global existence of optimal solution for** $(\bar{P}_\delta)$)**.** Assume that all the aforementioned assumptions in the beginning of this subsection are satisfied except for $(A5)$. Let $J : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R} \cup \{\infty\}$ be merely lower semicontinuous on $S(\bar{\delta})$ with domain of J contains $(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T))$. Then, for any $\delta \in (0, \bar{\varepsilon}]$, $(\bar{P}_\delta)$ has a global optimal solution.

Proof. Fix $\delta \in (0, \bar{\varepsilon}]$. Being admissible for (P), $((\bar{x}, \bar{y}), \bar{u})$ is also admissible for $(\bar{P}_\delta)$, due to Remark 3.2.3 and that $(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)) \in S_\delta$. As any admissible $((x, y), u)$ to $(\bar{P}_\delta)$ also satisfies $(x(0), y(0), x(T), y(T)) \in S_{\delta,\delta} \subset S(\bar{\delta})$, then, the lower semicontinuity of J on $S(\bar{\delta})$ and the compactness of $S_{\delta,\delta}$ yield the infimum of J over $((x, y), u)$ satisfying $(\bar{D})$ and having its states endpoints in S_δ, is finite. Let $((x_n, y_n), u_n)$ be a minimizing sequence for $(\bar{P}_\delta)$. The proof from this point on continues as done in the proof of Theorem 4.1.1, in which (D) and S are now $(\bar{D})$ and S_δ, respectively, and we use Lemma 3.2.8, system (3.44), and the bounds in (3.45) instead of Lemma 3.3.4, system (3.97), and the bounds in (3.98), respectively, and we apply Lemma .0.2 itself, where $\mathcal{R} = r + 1$, $Q(t) := C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ and ζ_n and ζ are present, instead of its global version that was used for $\mathcal{R} := r$, $Q(t) := C(t)$ and $\zeta_n = \zeta = 0$. We deduce the existence of $((\tilde{x}_\delta, \tilde{y}_\delta), \tilde{u}_\delta)$ optimal for $(\bar{P}_\delta)$. □

Remark 4.2.5. Note that Theorem 4.2.4 remains valid if we replace the objective function of $(\bar{P}_\delta)$, $J(x(0), y(0), x(T), y(T))$, by $J(x(0), y(0), x(T), y(T)) + \int_0^T \mathbb{L}(t, x(t), y(t))\, dt$, where $\mathbb{L}$ is a *Carathéodory* function (see Definition 2.3.2) satisfying, for some $\sigma \in L^1([0, T], \mathbb{R}_+)$,

$$|\mathbb{L}(t, x, y)| \leq \sigma(t), \quad \forall (x, y) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t), \text{ and } t \in [0, T]. \tag{4.10}$$

This is so, because for any $u \in \mathcal{U}$, the solution (x, y) of $(\bar{D})$ belongs to the uniformly bounded set valued map $\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(\cdot)$, and $\mathbb{L}$ is a *Carathéodory* function (see Definition 2.3.2) satisfying (4.10), and does not explicitly depend on the control. Indeed, in the proof of Theorem 4.2.4, the existence of a minimizing sequence $((x_n, y_n), u_n)$ for $(\bar{P}_\delta)$, in which this change is implemented, remains valid, and the limit as $n \to \infty$ of the added term is $\int_0^T \mathbb{L}(t, \tilde{x}_\delta(t), \tilde{y}_\delta(t))dt$, by the dominated convergence theorem.

The following remark establishes a connection between a strong local minimizer for (P) and a strong local minimzer for $(\bar{P}_\delta)$.

Remark 4.2.6. Using Theorem 4.2.4 and Remark 3.2.3, we have the following.

(i) If $((\bar{x},\bar{y}),\bar{u})$ is a $\bar{\delta}$-*strong local minimizer* for (P), then, for any $\delta \in (0,\bar{\varepsilon})$, $((\bar{x},\bar{y}),\bar{u})$ is a δ-*strong local minimizer* for $(\bar{P}_\delta)$, and hence for $(\bar{P}_{\delta,\delta})$.

This fact motivates formulating in Proposition 4.2.8 the approximating problem for (P) near $((\bar{x},\bar{y}),\bar{u})$ as being that for $(\bar{P}_{\delta_o,\delta_o})$, where δ_o is chosen strictly less than $\bar{\varepsilon}$. It also plays a key role in step 4 of the proof of Theorem 4.2.11 by relaxing instead of (P), the problem $(\bar{\mathscr{P}})$, which is $(\bar{P}_{\frac{\delta}{2}})$ with extended J and added $\mathbb{L}$.

(ii) Conversely, given $\delta \in (0,\bar{\varepsilon}]$, if $((\bar{x},\bar{y}),\bar{u})$ is a $\hat{\delta}$-strong local minimum for $(\bar{P}_\delta)$ for $\hat{\delta} \in (0,\delta]$, then $((\bar{x},\bar{y}),\bar{u})$ is a $\hat{\delta}$-strong local minimum for (P).

For the rest of the chapter, $((\bar{x},\bar{y}),\bar{u})$ **is taken to be a** $\bar{\delta}$***-strong local minimum*** **for** (P)**.**

We shall employ the following notations.

- If $\bar{x}(0) \in \operatorname{int} C(0)$, then, $\bar{x}(0) \in \operatorname{int}(C(0) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(0)))$, and hence, taking $c := \bar{x}(0)$ in Remark 3.2.11(ii), we deduce that there exist $\hat{k}_{\bar{x}(0)} \in \mathbb{N}$ and $\hat{r}_{\bar{x}(0)} \in (0,\bar{\varepsilon})$, satisfying

$$\bar{x}(0) \in \bar{B}_{\hat{r}_{\bar{x}(0)}}(\bar{x}(0)) \subset \operatorname{int} \bar{C}^{\gamma_k}(0,k), \quad \forall k \geq \hat{k}_{\bar{x}(0)}. \tag{4.11}$$

If $\bar{x}(0) \in \operatorname{bdry} C(0)$, then $\bar{x}(0) \in \operatorname{bdry}(C(0) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(0)))$, and hence, taking $c := \bar{x}(0)$ in Proposition 3.2.10(v), we deduce that there exist a vector $d_{\bar{x}(0)} \neq 0$, $k_{\bar{x}(0)} \geq k_3$, and $r_{\bar{x}(0)} \in (0,\bar{\varepsilon})$, such that

$$\Big(C(0) \cap \bar{B}_{r_{\bar{x}(0)}}(\bar{x}(0))\Big) + \bar{\sigma}_k \frac{d_{\bar{x}(0)}}{\|d_{\bar{x}(0)}\|} \subset \operatorname{int} \bar{C}^{\gamma_k}(0,k), \quad \forall k \geq k_{\bar{x}(0)}. \tag{4.12}$$

- Since $\bar{y}(0) \in \operatorname{int} \bar{B}_{\bar{\delta}}(\bar{y}(0))$, then taking $\mathbf{d} := \bar{y}(0)$ in Remark 3.2.13(ii), we deduce that there exist $\mathbf{k}_{\bar{y}(0)} \in \mathbb{N}$ and $\mathbf{r}_{\bar{y}(0)} > 0$ satisfying

$$\bar{y}(0) \in \bar{B}_{\mathbf{r}_{\bar{y}(0)}}(\bar{y}(0)) \subset \operatorname{int} \bar{B}_{\bar{\rho}_k}(\bar{y}(0)), \quad \forall k \geq \mathbf{k}_{\bar{y}(0)}. \tag{4.13}$$

- Motivated by Remark 4.2.6(i), and equations (4.11)-(4.13), let $\delta_o > 0$ to be the fixed constant

$$\delta_o := \begin{cases} \min\left\{\frac{\bar{\varepsilon}}{2}, \hat{r}_{\bar{x}(0)}, \mathbf{r}_{\bar{y}(0)}\right\} & \text{if } \bar{x}(0) \in \operatorname{int} C(0) \\ \min\left\{\frac{\bar{\varepsilon}}{2}, r_{\bar{x}(0)}, \mathbf{r}_{\bar{y}(0)}\right\} & \text{if } \bar{x}(0) \in \operatorname{bdry} C(0). \end{cases} \tag{4.14}$$

- For $\beta \in (0,1]$, we define for $t \in [0,T]$ a.e., and $(x,y,u) \in \bar{\mathcal{N}}_{(\bar{\delta},\bar{\delta})}(t) \times U(t)$:

$$f^{\beta}(t,x,y,u) := (1-\beta)f(t,x,y,\bar{u}(t)) + \beta f(t,x,y,u),$$
$$g^{\beta}(t,x,y,u) := (1-\beta)g(t,x,y,\bar{u}(t)) + \beta g(t,x,y,u).$$

  Note that also $h^{\beta} := (f^{\beta}, g^{\beta})$ satisfy (A4) as h does, and hence, all the results of Section 3.2 of Chapter 3 hold true for $(\bar{D}^{\beta})$ and $(\bar{D}^{\beta}_{\gamma_k})$, which are respectively obtained from $(\bar{D})$ and $(\bar{D}_{\gamma_k})$ by replacing h by h^{β}. Observe that $h^{\beta}(t,x,y,\bar{u}(t)) = h(t,x,y,\bar{u}(t))$.

- Let $(\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k})$ the solution of $(\bar{D}^{\beta}_{\gamma_k})$ corresponding to $((\bar{c}_{\gamma_k}, \bar{y}(0)), \bar{u})$, where, for k large enough, $\bar{c}_{\gamma_k} \in \operatorname{int} \bar{C}^{\gamma_k}(0,k)$ is the sequence corresponding (and converging) to $c = \bar{x}(0)$ via Remark 3.2.11, namely,

$$\bar{c}_{\gamma_k} = \begin{cases} \bar{x}(0) & \text{if } \bar{x}(0) \in \operatorname{int} C(0) \\ \bar{x}(0) + \bar{\sigma}_k \frac{d_{\bar{x}(0)}}{\|d_{\bar{x}(0)}\|} & \text{if } \bar{x}(0) \in \operatorname{bdry} C(0), \end{cases}$$

  where $d_{\bar{x}(0)}$ is the vector from Proposition 3.2.10(v) corresponding to $\bar{x}(0)$. Then, by Corollary 3.2.16, along a subsequence, $(\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k})$ converges uniformly to $(\bar{x}, \bar{y})$ and satisfies all conclusions of Theorem 3.2.14. In particular, we have that $(\bar{x}_{\gamma_k}(t), \bar{y}_{\gamma_k}(t)) \in \bar{C}^{\gamma_k}(t,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t))$ $\forall t \in [0,T]$ and $(\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k})_k$ is uniformly lipschitz.

- We define for all $k \in \mathbb{N}$

$$S^{\gamma_k}(k) := \begin{cases} [S_{\delta_o} + (0,0,\bar{e}_{\gamma_k}, \bar{\omega}_{\gamma_k})] \cap [\bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(0) \times \bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(T)], & \text{if } \bar{x}(0) \in \operatorname{int} C(0) \\ [S_{\delta_o} + (\bar{\sigma}_k \frac{d_{\bar{x}(0)}}{\|d_{\bar{x}(0)}\|}, 0, \bar{e}_{\gamma_k}, \bar{\omega}_{\gamma_k})] \cap [\bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(0) \times \bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(T)], & \text{if } \bar{x}(0) \in \operatorname{bdry} C(0), \end{cases} \tag{4.15}$$

  where, S_{δ_o} is defined in (4.8), and

$$(\bar{e}_{\gamma_k}, \bar{\omega}_{\gamma_k}) := (\bar{x}_{\gamma_k}(T) - \bar{x}(T), \bar{y}_{\gamma_k}(T) - \bar{y}(T)) \xrightarrow[k\to\infty]{} (0,0).$$

Remark 4.2.7. Our sets $S^{\gamma_k}(k)$ satisfy the following properties:

$$\forall k \in \mathbb{N},\ S^{\gamma_k}(k) \text{ is closed}, \quad \text{and} \quad S^{\gamma_k}(k) \subset S(\bar{\delta}),\ \text{for } k \text{ sufficiently large}, \tag{4.16}$$

$$\{(c,d) : (c,d,e,\omega) \in S^{\gamma_k}(k)\} \subset \operatorname{int} \bar{C}^{\gamma_k}(0,k) \times \operatorname{int} \bar{B}_{\bar{\rho}_k}(\bar{y}(0)) \subset \operatorname{int} \bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(0) \ \text{ for } \ k \text{ large}, \tag{4.17}$$

$$\lim_{k\to\infty} S^{\gamma_k}(k) = S_{\delta_o,\delta_o}, \tag{4.18}$$

$$(\bar{c}_{\gamma_k}, \bar{y}(0), \bar{x}_{\gamma_k}(T), \bar{y}_{\gamma_k}(T)) \in S^{\gamma_k}(k),\ \text{for } k \text{ sufficiently large}. \tag{4.19}$$

Using the local property of limiting normal cone (see Lemma 2.2.4), we know that, for any element $(c,d,e,\omega) \in S^{\gamma_k}(k)$ with $(e,\omega) \in \operatorname{int} \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(T) = (\operatorname{int} C(T) \cap B_{\bar{\varepsilon}}(\bar{x}(T))) \times B_{\bar{\delta}}(\bar{y}(T))$, we have

$$N^L_{S^{\gamma_k}(k)}(c,d,e,\omega) = \begin{cases} N^L_S(c,d,e-\bar{e}_{\gamma_k},\omega-\bar{\omega}_{\gamma_k}) & \text{if } \bar{x}(0) \in \operatorname{int} C(0) \text{ and} \\ & (c,d,e-\bar{e}_{\gamma_k},\omega-\bar{\omega}_{\gamma_k}) \in \operatorname{int} \bar{\mathscr{B}}_{\delta_o} \\ & \\ N^L_S(c-\bar{\sigma}_k\frac{d_{\bar{x}(0)}}{\|d_{\bar{x}(0)}\|},d,e-\bar{e}_{\gamma_k},\omega-\bar{\omega}_{\gamma_k}) & \text{if } \bar{x}(0) \in \operatorname{bdry} C(0) \text{ and} \\ & (c-\bar{\sigma}_k\frac{d_{\bar{x}(0)}}{\|d_{\bar{x}(0)}\|},d,e-\bar{e}_{\gamma_k},\omega-\bar{\omega}_{\gamma_k}) \in \operatorname{int} \bar{\mathscr{B}}_{\delta_o}. \end{cases} \tag{4.20}$$

This next proposition provides a sequence of optimal control problems with specific *joint* endpoint constraints that approximates our initial problem (P) near $((\bar{x},\bar{y}),\bar{u})$, that is, the problem $(\bar{P}_{\delta_o,\delta_o})$.

Proposition 4.2.8 (Approximating problems for (P)). For all $\alpha > 0$ and $\beta \in (0,1]$, there exists a subsequence of $(\gamma_k)_k$ (we do not relabel) and a sequence $(c_{\gamma_k},d_{\gamma_k},e_{\gamma_k},\omega_{\gamma_k};u_{\gamma_k}) \in S^{\gamma_k}(k) \times \mathcal{U}$ such that the associated problem $(P^{\alpha,\beta}_{\gamma_k})$ defined by:

$$\begin{aligned} (P^{\alpha,\beta}_{\gamma_k})\colon\ & \text{Minimize } J(x(0),y(0),x(T),y(T)) \ + \ \alpha\|u-u_{\gamma_k}\|_1 \\ & \quad + \alpha \,\|\,(x(0),y(0),x(T),y(T)) - (c_{\gamma_k},d_{\gamma_k},e_{\gamma_k},\omega_{\gamma_k})\,\|, \\ & \text{over } ((x,y),u) \text{ such that } u(\cdot) \in \mathscr{U} \text{ and} \\ & \begin{cases} (\bar{D}^\beta_{\gamma_k}) \begin{cases} \dot{x}(t) = f^\beta(t,x(t),y(t),u(t)) - \sum_{i=1}^{r+1}\gamma_k e^{\gamma_k\psi_i(t,x(t))}\nabla_x\psi_i(t,x(t)) \text{ a.e. } t\in[0,T], \\ \dot{y}(t) = g^\beta(t,x(t),y(t),u(t)) - \gamma_k e^{\gamma_k\varphi(t,y(t))}\nabla_y\varphi(t,y(t)) \text{ a.e. } t\in[0,T], \end{cases} \\ (x(t),y(t)) \in \bar{B}_{\delta_o}(\bar{x}(t),\bar{y}(t)) \quad \forall t\in[0,T], \qquad \textbf{(S.C)}, \\ (x(0),y(0),x(T),y(T)) \in S^{\gamma_k}(k), \end{cases} \end{aligned}$$

has an optimal solution $((x_{\gamma_k},y_{\gamma_k}),u_{\gamma_k})$ such that

$$(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T)) = (c_{\gamma_k},d_{\gamma_k},e_{\gamma_k},\omega_{\gamma_k})$$

and $(x_{\gamma_k})_k$ and $(y_{\gamma_k})_k$ are uniformly Lipschitz. Moreover,

$$(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T), y_{\gamma_k}(T)) \in (S_{\delta_o} + \rho_1 B) \cap \operatorname{int} \left(\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(0) \times \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(T)\right) \subset \operatorname{int} S(\bar{\delta}), \quad (4.21)$$

$$(x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \bar{C}^{\gamma_k}(t,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(t)), \quad \forall t \in [0,T], \quad (4.22)$$

$$(x_{\gamma_k}, y_{\gamma_k}) \xrightarrow{unif} (\bar{x}, \bar{y}), \quad u_{\gamma_k} \xrightarrow[in\ L^1]{strongly} \bar{u}, \quad \text{and} \quad (\dot{x}_{\gamma_k}, \dot{y}_{\gamma_k}) \xrightarrow[in\ L^\infty]{w*} (\dot{\bar{x}}, \dot{\bar{y}}). \quad (4.23)$$

The functions $\xi^i_{\gamma_k}$ $(i = 1, \cdots, r+1)$ and ζ_{γ_k}, corresponding to x_{γ_k} and y_{γ_k} via (3.76), satisfy (3.75) and there exists $(\xi^1, \cdots, \xi^r) \in L^\infty([0,T], \mathbb{R}^r_+)$ such that

$$\xi^i_{\gamma_k} \xrightarrow[in\ L^\infty]{w*} \xi^i,\ \xi^i = 0 \text{ on } I^-_i(\bar{x})\ (\forall i = 1, \cdots, r),\ \|\sum_{i=1}^r \xi^i\|_\infty \leq \frac{2\bar{\mu}}{\bar{\eta}^2},\ (\gamma_k \xi^{r+1}_{\gamma_k}, \gamma_k \zeta_{\gamma_k}) \xrightarrow{unif} 0, (4.24)$$

and $((\bar{x}, \bar{y}), \bar{u})$ together with $(\xi^1, \cdots, \xi^r)$ satisfies

$$\begin{cases} \dot{\bar{x}}(t) = f(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)) - \sum_{i=1}^r \xi^i(t) \nabla_x \psi_i(t, \bar{x}(t)) \text{ a.e. } t \in [0,T], \\ \dot{\bar{y}}(t) = g(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)) \text{ a.e. } t \in [0,T], \\ \psi_i(t, \bar{x}(t)) \leq 0,\ \forall t \in [0,T], \quad \forall i \in \{1, \cdots, r\}. \end{cases} \quad (4.25)$$

Proof. **Step 1: $(P^{0,\beta}_{\gamma_k})$ admits an optimal solution $((\hat{x}_{\gamma_k}, \hat{y}_{\gamma_k}), \hat{u}_{\gamma_k})$.**

Given that $(\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k})$ is the solution of $(\bar{D}^\beta_{\gamma_k})$ corresponding to $((\bar{c}_{\gamma_k}, \bar{y}(0)), \bar{u})$, the inclusion (4.19) holds, and $(\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k}) \to (\bar{x}, \bar{y})$, then for k large enough, the sequence $((\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k}), \bar{u})$ is admissible for $(P^{\alpha,\beta}_{\gamma_k})$ for every $\alpha \geq 0$ and $\beta \in (0,1]$. In particular, for k large enough, $((\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k}), \bar{u})$ is admissible for $(P^{0,\beta}_{\gamma_k})$. We fix k large enough and $\beta \in (0,1]$. We then apply Theorem 2.3.4 ([19, Theorem 23.10]) to $(P^{0,\beta}_{\gamma_k})$, with $Q = \operatorname{Gr}\left(\bar{B}_{\delta_o}(\bar{x}(\cdot), \bar{y}(\cdot)) \cap C(\cdot) \times \mathbb{R}^l\right)$ and $E = S^{\gamma_k}(k)$. Notice that conditions (a), (b), (c), (d), (e), and (f) of this theorem are satisfied due to the validity of assumptions (A1), (A3.1), (A4), and (A5), along with the properties of $S^{\gamma_k}(k)$. Hence, $(P^{0,\beta}_{\gamma_k})$ admits an optimal solution $((\hat{x}_{\gamma_k}, \hat{y}_{\gamma_k}), \hat{u}_{\gamma_k})$. Using equations (4.16) and (4.18), we deduce that there exists (c, d, e, ω) such that, up to a subsequence

$$(\hat{x}_{\gamma_k}(0), \hat{y}_{\gamma_k}(0), \hat{x}_{\gamma_k}(T), \hat{y}_{\gamma_k}(T)) \longrightarrow (c, d, e, \omega) \in S_{\delta_o,\delta_o}.$$

Step 2: Convergence of $(\hat{x}_{\gamma_k}, \hat{y}_{\gamma_k})$ to an admissible solution for $(\bar{P}_{\delta_o,\delta_o})$ with δ_o distance to $(\bar{x}, \bar{y})$.

As $(\hat{x}_{\gamma_k}(0), \hat{y}_{\gamma_k}(0)) \in \bar{C}^{\gamma_k}(0,k) \times \bar{B}_{\bar{\rho}_k}(\bar{y}(0))$ (see equation (4.17)) and its limit $(c,d) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(0)$, then, applying Theorem 3.2.14(I) to $((\hat{x}_{\gamma_k}(0), \hat{y}_{\gamma_k}(0)), \hat{u}_{\gamma_k})$, we deduce that the resulting unique solution of $(\bar{D}^{\beta}_{\gamma_k})$ is $(\hat{x}_{\gamma_k}, \hat{y}_{\gamma_k})$ and satisfies (3.74)-(3.76), and hence, by $(A1)$, $(A4.2)$, and Theorem 3.2.14(II), there exists $((\hat{x},\hat{y}),u)$, such that along a subsequence of $(\hat{x}_{\gamma_k}, \hat{y}_{\gamma_k})$ (we do not relabel), we have $(\hat{x}_{\gamma_k}, \hat{y}_{\gamma_k}) \xrightarrow{unif} (\hat{x},\hat{y})$, $(\hat{x}(t),\hat{y}(t)) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)$ for all $t \in [0,T]$, and $((\hat{x},\hat{y}),u)$ uniquely solves $(\bar{D}^{\beta})$ starting at (c,d). It follows that $(\hat{x}(T),\hat{y}(T)) = (e,\omega)$. Moreover, as $(\hat{x}_{\gamma_k}, \hat{y}_{\gamma_k})$ satisfies **(S.C)**, then we have $(\hat{x}(t),\hat{y}(t)) \in \bar{B}_{\delta_o}(\bar{x}(t),\bar{y}(t))$ for all $t \in [0,T]$. Using $(A4.2)$ and Filippov Selection Theorem (see Theorem 2.3.5), we can find $\hat{u} \in \mathcal{U}$ such that $((\hat{x},\hat{y}),\hat{u})$ satisfies $(\bar{D})$, and hence $((\hat{x},\hat{y}),\hat{u})$ is admissible for $(\bar{P}_{\delta_o,\delta_o})$ with $\|(\hat{x},\hat{y}) - (\bar{x},\bar{y})\|_\infty \leq \delta_o$.

Step 3: $(P^{\alpha,\beta}_{\gamma_{k_n}})$ **defined by means of** $(c_{\gamma_{k_n}}, d_{\gamma_{k_n}}, e_{\gamma_{k_n}}, \omega_{\gamma_{k_n}}; u_{\gamma_{k_n}})$**, has** $((x_{\gamma_{k_n}}, y_{\gamma_{k_n}}), u_{\gamma_{k_n}})$ **as optimal solution.**

Since $((\bar{x},\bar{y}),\bar{u})$ is a $\bar{\delta}$-strong local minimizer for (P), then, by Remark 4.2.6(i) and $\delta_o < \bar{\varepsilon}$, $((\bar{x},\bar{y}),\bar{u})$ is a δ_o-strong local minimizer for $(\bar{P}_{\delta_o})$ and hence for $(\bar{P}_{\delta_o,\delta_o})$, and hence, we have

$$J(\bar{x}(0),\bar{y}(0),\bar{x}(T),\bar{y}(T)) \leq J(\hat{x}(0),\hat{y}(0),\hat{x}(T),\hat{y}(T)).$$

On the other hand, $((\hat{x}_{\gamma_k}, \hat{y}_{\gamma_k}), \hat{u}_{\gamma_k})$ is an optimal solution for $(P^{0,\beta}_{\gamma_k})$ for which $((\bar{x}_{\gamma_k}, \bar{y}_{\gamma_k}), \bar{u}_{\gamma_k})$ is admissible, we deduce that

$$J(\hat{x}_{\gamma_k}(0), \hat{y}_{\gamma_k}(0), \hat{x}_{\gamma_k}(T), \hat{y}_{\gamma_k}(T)) \leq J(\bar{x}_{\gamma_k}(0), \bar{y}_{\gamma_k}(0), \bar{x}_{\gamma_k}(T), \bar{y}_{\gamma_k}(T)).$$

Combining the above two inequalities and using the continuity of $J(\cdot,\cdot,\cdot,\cdot)$, we deduce that

$$\lim_{k\to\infty} \Big[J(\bar{x}_{\gamma_k}(0), \bar{y}_{\gamma_k}(0), \bar{x}_{\gamma_k}(T), \bar{y}_{\gamma_k}(T)) - J(\hat{x}_{\gamma_k}(0), \hat{y}_{\gamma_k}(0), \hat{x}_{\gamma_k}(T), \hat{y}_{\gamma_k}(T)) \Big] = 0.$$

Thus, for fixed $\alpha > 0$, there exists an increasing sequence $(k_n)_n$ such that $\forall\, n \geq 1$, $\forall\, k_n > n$,

$$J(\bar{x}_{\gamma_{k_n}}(0), \bar{y}_{\gamma_{k_n}}(0), \bar{x}_{\gamma_{k_n}}(T), \bar{y}_{\gamma_{k_n}}(T)) \quad \leq \quad J(\hat{x}_{\gamma_{k_n}}(0), \hat{y}_{\gamma_{k_n}}(0), \hat{x}_{\gamma_{k_n}}(T), \hat{y}_{\gamma_{k_n}}(T)) + \frac{\alpha}{n}.$$

The rest of the proof follows from imitating the proof of [55, Proposition 6.2], and applying Ekeland Variational Principle (Theorem 2.2.49 or [66, Theorem 3.3.1]), to the following version of the data corresponding to our problem:

- $X = \{(c,d,e,\omega;u) \in S^{\gamma_k}(k_n) \times \mathcal{U}$: the unique solution $((x,y),u)$ of $(\bar{D}^{\beta}_{\gamma_{k_n}})$ with $(x(0),y(0)) = (c,d)$ satisfies $(x(T),y(T)) = (e,\omega)$ and $(x(t),y(t)) \in \bar{B}_{\delta_o}(\bar{x}(t),\bar{y}(t))\ \forall t\}$.
- For $(c,d,e,\omega;u),\ (c',d',e',\omega';u') \in X$, we define the distance

$$\mathbb{D}\left((c,d,e,\omega;u),(c',d',e',\omega';u')\right) := \|u-u'\|_{L^1} + \|(c,d,e,\omega)-(c',d',e',\omega')\|.$$

- For $(c,d,e,\omega;u) \in X$, $\mathbb{F}(c,d,e,\omega;u) := J(c,d,e,\omega)$.
- $\alpha := \alpha$ and $\lambda := \frac{1}{n}$.

Notice that $(X,\mathbb{D})$ is a non-empty complete metric space, and $\mathbb{F}$ is continuous on X. Therefore, we deduce the existence of $(c_{\gamma_{k_n}}, d_{\gamma_{k_n}}, e_{\gamma_{k_n}}, \omega_{\gamma_{k_n}}; u_{\gamma_{k_n}}) \in X$ such that, for $(x_{\gamma_{k_n}}, y_{\gamma_{k_n}})$, the solution of $(\bar{D}^{\beta}_{\gamma_{k_n}})$ corresponding to $((c_{\gamma_{k_n}}, d_{\gamma_{k_n}}), u_{\gamma_{k_n}})$, satisfies $(x_{\gamma_{k_n}}(T), y_{\gamma_{k_n}}(T)) = (e_{\gamma_{k_n}}, \omega_{\gamma_{k_n}})$ and $(x_{\gamma_{k_n}}(t), y_{\gamma_{k_n}}(t)) \in \bar{B}_{\delta_o}(\bar{x}(t),\bar{y}(t))\ \forall t$, and the following holds:

$$J(x_{\gamma_{k_n}}(0), y_{\gamma_{k_n}}(0), x_{\gamma_{k_n}}(T), y_{\gamma_{k_n}}(T)) \leq J(\bar{x}_{\gamma_{k_n}}(0), \bar{y}_{\gamma_{k_n}}(0), \bar{x}_{\gamma_{k_n}}(T), \bar{y}_{\gamma_{k_n}}(T)), \tag{4.26}$$

$$\|u_{\gamma_{k_n}} - \bar{u}\|_{L^1} + \left\| \left(c_{\gamma_{k_n}}, d_{\gamma_{k_n}}, e_{\gamma_{k_n}}, \omega_{\gamma_{k_n}}\right) - \left(\bar{c}_{\gamma_{k_n}}, \bar{y}(0), \bar{x}_{\gamma_{k_n}}(T), \bar{y}_{\gamma_{k_n}}(T)\right) \right\| \leq \frac{1}{n}, \tag{4.27}$$

and for all $((c,d,e,\omega);u) \in X$, we have

$$\begin{aligned} J(x_{\gamma_{k_n}}(0), y_{\gamma_{k_n}}(0), x_{\gamma_{k_n}}(T), y_{\gamma_{k_n}}(T)) &\leq J(x(0),y(0),x(T),y(T)) \\ &+\alpha(\|u-u_{\gamma_{k_n}}\|_{L^1} + \left\| (c,d,e,\omega) - \left(c_{\gamma_{k_n}}, d_{\gamma_{k_n}}, e_{\gamma_{k_n}}, \omega_{\gamma_{k_n}}\right) \right\|, \end{aligned} \tag{4.28}$$

where $((x,y),u)$ is the unique solution of $(\bar{D}^{\beta}_{\gamma_{k_n}})$ starting with $(x(0),y(0)) = (c,d)$ and satisfying $(x(T),y(T)) = (e,\omega)$ and $(x(t),y(t)) \in \bar{B}_{\delta_o}(\bar{x}(t),\bar{y}(t))\ \forall\, t \in [0,T]$.

Hence, for n large, the problem $(P^{\alpha,\beta}_{\gamma_{k_n}})$ defined by means of $(c_{\gamma_{k_n}}, d_{\gamma_{k_n}}, e_{\gamma_{k_n}}, \omega_{\gamma_{k_n}}; u_{\gamma_{k_n}})$, has $((x_{\gamma_{k_n}}, y_{\gamma_{k_n}}), u_{\gamma_{k_n}})$ as optimal solution satisfying

$$(x_{\gamma_{k_n}}(0), y_{\gamma_{k_n}}(0), x_{\gamma_{k_n}}(T), y_{\gamma_{k_n}}(T)) = (c_{\gamma_{k_n}}, d_{\gamma_{k_n}}, e_{\gamma_{k_n}}, \omega_{\gamma_{k_n}}) \xrightarrow[n\to\infty]{} (\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)) \in S,$$

$$u_{\gamma_{k_n}} \xrightarrow[L^1]{strongly} \bar{u}, \quad (x_{\gamma_{k_n}}, y_{\gamma_{k_n}}) \xrightarrow[n\to\infty]{unif} (\bar{x},\bar{y}),$$

and all conclusions of Theorem 3.2.14. Hence, (4.23) is valid, and, for $(\xi^i_{\gamma_k})_{i=1}^{r+1}$ and ζ_{γ_k} corresponding to $(x_{\gamma_k}, y_{\gamma_k})$ via (3.76), there exist $(\xi^i)_{i=1}^{r+1}$ and ζ such that (3.75), (3.77), (3.79), (3.80), and (3.81)-(3.83) hold. Notice that, as $\psi_{r+1}(t,\bar{x}(t)) = -\frac{\bar{\varepsilon}^2}{2} < 0$ and $\varphi(t,\bar{y}(t)) = -\frac{\bar{\delta}^2}{2} < 0$ $\forall t \in [0,T]$, we have that $\xi^{r+1} \equiv 0$, $\zeta \equiv 0$, and, for some $\tilde{k} \in \mathbb{N}$, $\psi_{r+1}(t, x_{\gamma_k}(t)) \leq -\frac{\bar{\varepsilon}^2}{4}$ and $\varphi(t, y_{\gamma_k}(t)) \leq -\frac{\bar{\delta}^2}{4}$, $\forall k \geq \tilde{k}$ and $\forall t \in [0,T]$, and hence,

$$\gamma_k \xi^{r+1}_{\gamma_k}(t) \leq \gamma_k^2 e^{-\gamma_k \frac{\bar{\varepsilon}^2}{4}} \quad \text{and} \quad \gamma_k \zeta_{\gamma_k}(t) \leq \gamma_k^2 e^{-\gamma_k \frac{\bar{\delta}^2}{4}}, \;\; \forall k \geq \tilde{k}, \;\; \text{and} \;\; \forall t \in [0,T]. \tag{4.29}$$

That is, $(\gamma_k \xi^{r+1}_{\gamma_k}, \gamma_k \zeta_{\gamma_k}) \xrightarrow{unif} 0$, and thus, (4.24) holds. Furthermore, since $h^\beta(t,\bar{x}(t),\bar{y}(t),\bar{u}(t)) = h(t,\bar{x}(t),\bar{y}(t),\bar{u}(t))$, it follows that $((\bar{x},\bar{y}),\bar{u})$ and $(\xi^1,\cdots,\xi^r)$ satisfy (4.25).

Finally, (4.21) is also valid, due to having $(x_{\gamma_{k_n}}(0), y_{\gamma_{k_n}}(0), x_{\gamma_{k_n}}(T), y_{\gamma_{k_n}}(T)) \in S^{\gamma_k}(k_n)$, $\bar{\sigma}_{k_n} \to 0$, $(\bar{e}_{\gamma_{k_n}}, \bar{\omega}_{\gamma_{k_n}}) \to (0,0)$ (as $n \to \infty$), and $(x_{\gamma_{k_n}}(t), y_{\gamma_{k_n}}(t)) \in \bar{C}^{\gamma_{k_n}}(t,k) \times \bar{B}_{\bar{\rho}_{k_n}}(\bar{y}(t)) \subset \operatorname{int} \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)$. □

The next result is obtained as a direct application of the nonsmooth Pontryagin maximum principle for state constrained problems to each of the approximating problem $(P^{\alpha,\beta}_{\gamma_k})$ defined in Proposition 4.2.8.

Proposition 4.2.9 (Maximum principle for the approximating problems $(P^{\alpha,\beta}_{\gamma_k})$). Let $\alpha > 0$ and $\beta \in (0,1]$ be fixed. Let $((x_{\gamma_k}, y_{\gamma_k}), u_{\gamma_k})$ be the sequence from Proposition 4.2.8 which is optimal for $(P^{\alpha,\beta}_{\gamma_k})$ and satisfying $\lim_{k\to\infty} (x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T), y_{\gamma_k}(T)) = (\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T))$. Then, for each $k \in \mathbb{N}$, there exist $p_{\gamma_k} = (q_{\gamma_k}, v_{\gamma_k}) \in W^{1,1}([0,T];\mathbb{R}^n \times \mathbb{R}^l)$ and a scalar $\lambda_{\gamma_k} \geq 0$ such that

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

$$\|p_{\gamma_k}\|_\infty + \lambda_{\gamma_k} = 1. \tag{4.30}$$

(*ii*) **Transversality equation**

$$\begin{aligned}(p_{\gamma_k}(0), -p_{\gamma_k}(T)) \in \quad & \lambda_{\gamma_k} \partial^L J(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T), y_{\gamma_k}(T)) + \\ & \alpha \bar{B} + N^L_{S^{\gamma_k}(k)}\left(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T), y_{\gamma_k}(T)\right).\end{aligned} \tag{4.31}$$

(*iii*) **Maximization condition**

$$\max_{u \in U(t)} \left\{ \langle (q_{\gamma_k}(t), v_{\gamma_k}(t)), (f,g)(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u) \rangle - \frac{\lambda_{\gamma_k}\alpha}{\beta} \|u - u_{\gamma_k}(t)\| \right\} \tag{4.32}$$

is attained at $u = u_{\gamma_k}(t)$ a.e. $t \in [0,T]$.

(iv) **Adjoint equation** For almost all $t \in [0, T]$,

$$
\begin{aligned}
-\dot{p}_{\gamma_k}(t) = \begin{bmatrix} -\dot{q}_{\gamma_k}(t) \\ -\dot{v}_{\gamma_k}(t) \end{bmatrix} \in\ & (1-\beta)(\partial^{(x,y)} f(t, x_{\gamma_k}(t), y_{\gamma_k}(t), \bar{u}(t)))^T q_{\gamma_k}(t) \\
& +\beta(\partial^{(x,y)} f(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u_{\gamma_k}(t)))^T q_{\gamma_k}(t) \\
& +(1-\beta)(\partial^{(x,y)} g(t, x_{\gamma_k}(t), y_{\gamma_k}(t), \bar{u}(t)))^T v_{\gamma_k}(t) \\
& +\beta(\partial^{(x,y)} g(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u_{\gamma_k}(t)))^T v_{\gamma_k}(t) \\
& - \begin{bmatrix} \left(\partial^x \left(\sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(t, x_{\gamma_k}(t))} \nabla_x \psi_i(t, x_{\gamma_k}(t))\right)\right)^T q_{\gamma_k}(t) \\ \left(\nabla_y \left(\gamma_k e^{\gamma_k \varphi(t, y_{\gamma_k}(t))} \nabla_y \varphi(t, y_{\gamma_k}(t))\right)\right)^T v_{\gamma_k}(t) \end{bmatrix}, \qquad (4.33)
\end{aligned}
$$

where,

$$
\begin{aligned}
\partial^x \left(\sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(t, x_{\gamma_k}(t))} \nabla_x \psi_i(t, x_{\gamma_k}(t))\right) \ \subset\ & \sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(t, x_{\gamma_k}(t))} \partial^{xx} \psi_i(t, x_{\gamma_k}(t)) \\
& + \sum_{i=1}^{r+1} \gamma_k^2 e^{\gamma_k \psi_i(t, x_{\gamma_k}(t))} \nabla_x \psi_i(t, x_{\gamma_k}(t)) \nabla_x \psi_i(t, x_{\gamma_k}(t))^T, \\
\nabla_y \left(\gamma_k e^{\gamma_k \varphi(t, y_{\gamma_k}(t))} \nabla_y \varphi(t, y_{\gamma_k}(t))\right) \ =\ & \gamma_k e^{\gamma_k \varphi(t, y_{\gamma_k}(t))}\ \mathrm{I}_{l \times l} \\
& + \gamma_k^2 e^{\gamma_k \varphi(t, y_{\gamma_k}(t))} \nabla_y \varphi(t, y_{\gamma_k}(t)) \nabla_y \varphi(t, y_{\gamma_k}(t))^T.
\end{aligned}
$$

Proof. As $(P_{\gamma_k}^{\alpha,\beta})$ is a standard optimal control problem with implicit state constraints, we shall apply [66, Theorem 9.3.1 and P.332] for the optimal solution $((x_{\gamma_k}, y_{\gamma_k}), u_{\gamma_k})$ of $(P_{\gamma_k}^{\alpha,\beta})$ obtained in Proposition 4.2.8. The proof is obtained from translating the conditions of [66, Theorem 9.3.1] to our data, and using the standard state augmentation technique.

Step 1. All assumptions of [66, Theorem 9.3.1 and P.332] are satisfied.

Applying the state augmentation technique, our optimal solution is $(x_{\gamma_k}, y_{\gamma_k}, z_{\gamma_k})$, where $(x_{\gamma_k}, y_{\gamma_k})$ is the optimal solution from Proposition 4.2.8, $z_{\gamma_k}(t) := \int_0^t \|u_{\gamma_k}(s) - u_{\gamma_k}(s)\| ds = 0$, and u_{γ_k} is the optimal control.

Assumptions (H1), (H2) and (H3) of [66, Theorem 9.3.1] are satisfied because assumptions (A2), (A3), (A4), and (A5) hold true, $(x_{\gamma_k}, y_{\gamma_k})$ converges uniformly to $(\bar{x}, \bar{y})$ and (4.21) is satisfied. Note that for k large enough, the required constraint qualification (CQ) in [66, Page 332] is satisfied by the multifunction $\bar{B}_{\delta_o}(\bar{x}(\cdot), \bar{y}(\cdot))$ at $(x_{\gamma_k}(t), y_{\gamma_k}(t))$. In other words, we need to show that

1. $\bar{B}_{\delta_o}(\bar{x}(\cdot), \bar{y}(\cdot))$ is lower semicontinuous multifunction,
2. conv $(\bar{N}^L_{\bar{B}_{\delta_o}(\bar{x}(t),\bar{y}(t))}(x_{\gamma_k}(t), y_{\gamma_k}(t)))$ is pointed $\forall t \in [0,T]$, where the graph of $\bar{N}^L_{\bar{B}_{\delta_o}(\bar{x}(\cdot),\bar{y}(\cdot))}(\cdot)$ is defined to be the closure of the graph of $N^L_{\bar{B}_{\delta_o}(\bar{x}(\cdot),\bar{y}(\cdot))}(\cdot)$.

This is due to $(x_{\gamma_k}, y_{\gamma_k})$ converging uniformly to $(\bar{x}, \bar{y})$ and to $\bar{B}_{\delta_o}(\bar{x}(\cdot), \bar{y}(\cdot))$ being lower semicontinuous, with closed, convex, and nonempty interior values (hence epi-Lipschitz), (see Lemma 2.2.47 or [57, Remark 4.8(ii)]).

Step 2: The measure corresponding to the state constraint (S.C) is null.

Notice that the measure $\eta_{\gamma_k} \in C^*([0,T], \mathbb{R}^{n+l})$ corresponding to the state constraint (S.C) produced by [66, Theorem 9.3.1], is actually null. This is due to the fact that its support satisfies

$$\begin{aligned} \text{supp}\,\{\eta_{\gamma_k}\} &\subset \{t \in [0,T] : (t, x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \text{bdry}\,\text{Gr}\ \bar{B}_{\delta_o}(\bar{x}(\cdot), \bar{y}(\cdot))\}, \\ &= \{t \in [0,T] : (t, x_{\gamma_k}(t), y_{\gamma_k}(t)) \in \cup_{t\in[0,T]}\{t\} \times \mathscr{S}_{\delta_o}(\bar{x}(t), \bar{y}(t))\} \\ &= \emptyset, \end{aligned}$$

where $\mathscr{S}_{\delta_o}(\bar{x}(t), \bar{y}(t)) = \{(x,y) : \|(x,y) - (\bar{x}(t), \bar{y}(t))\| = \delta_o\}$. The last equality follows from the uniform convergence to $(\bar{x}, \bar{y})$ of $(x_{\gamma_k}(t), y_{\gamma_k}(t))$, (4.23).

Step 3. Deriving the transversality condition.

Let q_{γ_k}, v_{γ_k} and e_{γ_k} adjoint vectors corresponding to the optimal states x_{γ_k}, y_{γ_k} and z_{γ_k} respectively. We translate equation (iii) in [66, Theorem 9.3.1] to our data. First, notice that

$$e_{\gamma_k}(T) = -\lambda_{\gamma_k}\alpha.$$

In addition, we have, for $p_{\gamma_k} = (q_{\gamma_k}, v_{\gamma_k})$, that

$$\begin{gathered}(p_{\gamma_k}(0), -p_{\gamma_k}(T)) \in \lambda_{\gamma_k}\partial^L J(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T), y_{\gamma_k}(T)) + \\ \alpha\bar{B} + N^L_{S^{\gamma_k}(k)}(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T), y_{\gamma_k}(T)).\end{gathered}$$

Step 4. Deriving the adjoint equation.

We note that the Hamiltonian is given by

$$\begin{aligned} H(t,(x,y,z),(q,v,e),u) &= \langle q, f^{\beta}(t,x,y,u) - \sum_{i=1}^{r+1} \gamma_k e^{\gamma_k \psi_i(t,x)} \nabla_x \psi_i(t,x)\rangle \\ &+ \langle v, g^{\beta}(t,x,y,u) - \gamma_k e^{\gamma_k \varphi(t,y)} \nabla_y \varphi(t,y)\rangle + \langle e, \|u - u_{\gamma_k}(t)\|\rangle. \end{aligned}$$

Using equation (ii) in [66, Theorem 9.3.1], we deduce that equation (4.33) is satisfied, and $\dot{e}_{\gamma_k}(t) = 0$ for $t \in [0,T]$ a.e. Now, we use the transversality condition to deduce that for a.e $t \in [0,T]$, $e_{\gamma_k}(t) = e_{\gamma_k}(T) = -\lambda_{\gamma_k}\alpha$.

Step 5. Deriving the Maximization condition.

Applying equation (iv) in [66, Theorem 9.3.1] to our data, with the fact that $e_{\gamma_k}(t) = -\alpha\lambda_{\gamma_k}$ a.e. $t \in [0,T]$, we deduce that

$$\max_{u \in U(t)} \left\{ \langle q_{\gamma_k}(t), f(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u)\rangle + \langle v_{\gamma_k}(t), g(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u)\rangle - \frac{\lambda_{\gamma_k}\alpha}{\beta}\|u - u_{\gamma_k}(t)\| \right\} \tag{4.34}$$

is attained at $u = u_{\gamma_k}(t)$ a.e. $t \in [0,T]$.

Step 6. Nontriviality condition.

Since η_{γ_k} is null everywhere, we deduce from the nontriviality condition of Theorem 9.3.1 that $(p_{\gamma_k}, e_{\gamma_k}, \lambda_{\gamma_k}) \neq 0$. But $e_{\gamma_k} = -\alpha\lambda_{\gamma_k}$ then the transversality condition translates to

$$\|p_{\gamma_k}\|_{\infty} + \lambda_{\gamma_k} \neq 0.$$

□

Remark 4.2.10. We note the following.

- We will prove in the maximum principle (see equation (4.41)) that there exists $\tilde{M}_p > 0$ such that

$$\|p_{\gamma_k}(t)\| \leq \tilde{M}_p \|p_{\gamma_k}(T)\|, \quad \forall t \in [0,T], \quad \forall k \in \mathbb{N}. \tag{4.35}$$

  Hence, we can replace the nontriviality condition (i) by

$$\|p_{\gamma_k}(T)\| + \lambda_{\gamma_k} = 1. \tag{4.36}$$

This is particularly useful for us when taking the limit of the non-triviality condition in the proof of the maximum principle. As we will see, p_{γ_k} converges pointwise to a function p, allowing us to take the limit in (4.36).

- In addition, if $S = C_0 \times \mathbb{R}^{n+l}$ for a closed $C_0 \subset C(0) \times \mathbb{R}^l$, then $\lambda_{\gamma_k} \neq 0$ and it is taken to be 1 and the nontriviality condition (i) is eliminated. Indeed, if $\lambda_{\gamma_k} = 0$, then using transversality condition (ii), we deduce that $p_{\gamma_k}(T) = 0$. Thus, using equation (4.35), we deduce that p_{γ_k} is null. Hence, $(p_{\gamma_k}, \lambda_{\gamma_k}) = 0$ which contradicts the non-triviality condition.

Table 4.3 Summary of results from Subsection 4.2.2.

Result	**Description**
Theorem 4.2.4	We provide an existence result of an optimal solution for the truncated optimal control problem $(\bar{P}_\delta)$.
Remark 4.2.5	We provide an existence result of an optimal solution for a truncated optimal control problem, which is identical to $(\bar{P}_\delta)$ except for the addition of an integral term involving a *Carathéodory* function in its objective function.
Remark 4.2.6	We establish a connection between a strong local minimizer for (P) and a strong local minimzer for $(\bar{P}_\delta)$.
Remark 4.2.7	We provide properties for the sets $S^{\gamma_k}(k)$.
Proposition 4.2.8	We provide a sequence of optimal control problems with specific *joint* endpoint constraints that approximates our initial problem (P) near $((\bar{x}, \bar{y}), \bar{u})$, that is, the problem $(\bar{P}_{\delta_o,\delta_o})$.
Proposition 4.2.9	We provide necessary conditions to each of the approximating problems $(P^{\alpha,\beta}_{\gamma_k})$ defined in Proposition 4.2.8.
Remark 4.2.10	We provide conditions that could replace the non-triviality condition.

4.2.3 Maximum principle for (P)

The following result provides necessary conditions, in the form of an extended Pontryagin's maximum principle, for a $\bar{\delta}$-strong local minimizer $((\bar{x},\bar{y}),\bar{u})$ for the problem (P). We start by proving the theorem under the temporary assumption (A4.2), and without assuming any uniform bound on the sets $U(t)$ (Step I). In Step II, we show that, when the compact sets $U(t)$ are *uniformly* bounded, the convexity assumption (A4.2) can be removed. First, we introduce the following *nonstandard* notions of subdifferentials that shall be used in Theorem 4.2.11.

- $\partial_\ell^{(x,y)} h(t,\cdot,\cdot,u)$ denotes the *extended Clarke generalized Jacobian* of $h(t,\cdot,\cdot,u)$ that extends from the interior to the boundary of $\mathscr{N}_{(\bar{\delta},\bar{\delta})}(t) := \left[C(t)\cap \bar{B}_{\bar{\delta}}(\bar{x}(t))\right]\times \bar{B}_{\bar{\delta}}(\bar{y}(t))$ the notion of the Clarke generalized Jacobian (see Definition 2.2.30 or [55, Equation(11)]),
- $\partial_\ell^{xx}\psi_i(t,\cdot)$ is the *Clarke generalized Hessian relative to* $\operatorname{int}[C(t)\cap \bar{B}_{\bar{\delta}}(\bar{x}(t))]$ of $\psi_i(t,\cdot)$ (see Definition 2.2.32 or [55, Equation(12)]),
- $\partial_\ell^L J(\cdot,\cdot,\cdot,\cdot)$ is the *limiting subdifferential of* $J(\cdot,\cdot,\cdot,\cdot)$ *relative to* $\operatorname{int} S(\bar{\delta})$ (see Definition 2.2.25 or [55, Equation(8)]).

Theorem 4.2.11 (Generalized Pontryagin principle for (P)). Assume that $(A1)$-$(A2)$ are satisfied. Let $((\bar{x},\bar{y}),\bar{u})$ be a $\bar{\delta}$-strong local minimizer for (P) such that $(A3.1)$, $(A3.3)$, $(A4.1)$ and $(A5)$ are satisfied at $((\bar{x},\bar{y});\bar{\delta})$. Then, whenever $(A4.2)$ holds true, or if sets $U(t)$ are uniformly bounded, there exist an adjoint vector $p=(q,v)$ with $q\in BV([0,T];\mathbb{R}^n)$ and $v\in W^{1,2}([0,T];\mathbb{R}^l)$, finite signed Radon measures $(\nu^i)_{i=1}^r$ on $[0,T]$, nonnegative functions $(\xi^i)_{i=1}^r$ in $L^\infty([0,T];\mathbb{R}^+)$, L^2-measurable functions $\bar{A}(\cdot)$ in $\mathscr{M}_{n\times n}([0,T])$, $\bar{E}(\cdot)$ in $\mathscr{M}_{n\times l}([0,T])$, $\bar{\mathcal{A}}(\cdot)$ in $\mathscr{M}_{l\times n}([0,T])$, and $\bar{\mathcal{E}}(\cdot)$ in $\mathscr{M}_{l\times l}([0,T])$, L^∞-measurable functions $(\vartheta^i(\cdot))_{i=1}^r$ in $\mathscr{M}_{n\times n}([0,T])$, and a scalar $\lambda\geq 0$, satisfying the following:

(i) **Primal-dual admissible equation**

$$\begin{cases} \dot{\bar{x}}(t) = f(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)) - \sum_{i=1}^{r} \xi^i(t) \nabla_x \psi_i(t, \bar{x}(t)) \text{ a.e. } t \in [0, T], \\ \dot{\bar{y}}(t) = g(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)) \text{ a.e. } t \in [0, T], \\ \psi_i(t, \bar{x}(t)) \leq 0, \ \forall t \in [0, T], \quad \forall i \in \{1, \cdots, r\}. \end{cases}$$

(ii) **Non-triviality condition**

$$\lambda + \|p(T)\| = 1.$$

(iii) **Adjoint equations**

For any $z(\cdot) \in \mathcal{C}([0, T], \mathbb{R}^n)$

$$\int_{[0,T]} \langle z(t), dq(t) \rangle = \int_0^T \langle z(t), -\bar{A}(t)^T q(t) \rangle dt + \int_0^T \langle z(t), -\bar{\mathcal{A}}(t)^T v(t) \rangle dt$$
$$+ \sum_{i=1}^{r} \left(\int_0^T \xi^i(t) \langle z(t), \vartheta^i(t) q(t) \rangle dt + \int_0^T \langle z(t), \nabla_x \psi_i(t, \bar{x}(t)) \rangle d\nu^i(t) \right),$$

$$\dot{v}(t) = -\bar{E}(t)^T q(t) - \bar{\mathcal{E}}(t)^T v(t),$$

where for all $t \in [0, T]$ a.e.,

$$(\bar{A}(t), \bar{E}(t)) \in \partial_\ell^{(x,y)} f(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)), \qquad (\bar{\mathcal{A}}(t), \bar{\mathcal{E}}(t)) \in \partial_\ell^{(x,y)} g(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)),$$
$$\vartheta^i(t) \in \partial_\ell^{xx} \psi_i(t, \bar{x}(t)), \text{ for } i = 1, \cdots, r.$$

(iv) **Maximization condition**

$$\max_{u \in U(t)} \left\{ \langle q(t), f(t, \bar{x}(t), \bar{y}(t), u) \rangle + \langle v(t), g(t, \bar{x}(t), \bar{y}(t), u) \rangle \right\}$$

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

(v) **Complementary Slackness condition** For $i = 1, \cdots, r$, we have:

$\xi^i(t) = 0 \ \forall t \in I_i^{-}(\bar{x})$, and $\xi^i(t) \langle \nabla_x \psi_i(t, \bar{x}(t)), q(t) \rangle = 0$ a.e. $t \in [0, T]$.

(vi) **Measures Properties** For $i = 1, \cdots, r$, we have:

$\text{supp}\,\{\nu^i\} \subset I_i^0(\bar{x})$ and the measure $\langle q(t), \nabla_x \psi_i(t, \bar{x}(t)) \rangle d\nu^i(t)$ is nonnegative.

(vii) **Transversality condition**

$$((q,v)(0), -(q,v)(T)) \in \lambda\, \partial_\ell^L J((\bar{x},\bar{y})(0),(\bar{x},\bar{y})(T)) + N_S^L((\bar{x},\bar{y})(0),(\bar{x},\bar{y})(T)).$$

In addition, if $S = C_0 \times \mathbb{R}^{n+l}$, for a closed $C_0 \subset C(0) \times \mathbb{R}^l$, then $\lambda = 1$, and the non-triviality condition is discarded.

Proof. Step I. **Assume for now the temporary assumption (A4.2) holds true.** All the previous results including the consequences in subsection 4.2.2 are valid. In particular, $(\bar{x},\bar{y})$ is $L_{(\bar{x},\bar{y})}$-Lipschitz with $L_{(\bar{x},\bar{y})} \geq 1$. **Assume as well that the additional assumptions, (A3.3)′, is satisfied.**

(A3.3)′ $\forall t \in I^0(\bar{x})$, $\mathcal{G}_\psi(t)$, *the Gramian matrix of the vectors* $\{\nabla_x \psi_i(t,\bar{x}(t)) : i \in \mathcal{I}^0_{(t,\bar{x}(t))}\}$, is *strictly diagonally dominant.*

Since $\{\psi_i\}_{i=1}^r$ satisfy (A3.3)′, then by Lemma 4.2.2, there exist $0 < \bar{a} \leq 2a_o$, $0 < \bar{b} < 1$, and $\bar{c} > 0$ such that (4.4) is satisfied, where a_o is the constant in Lemma 4.2.3.

We begin our proof by introducing the function $\hat{\psi}_i$, which we will work with in place of ψ_i, in order to establish that the function $\hat{q}_{\gamma_k}$ has uniformly bounded variation in Step I.2. After formulating the Pontryagin Maximum Principle in terms of $\hat{\psi}_i$, we will translate the necessary conditions in terms of ψ_i (see Step I.3.4).

Define the following function $\hat{\psi}_i(\cdot,\cdot)$ on the same domain of $\psi_i(\cdot,\cdot)$ as

$$\hat{\psi}_i(t,x) := \begin{cases} \psi_i(t,x) & \text{if } -\frac{\bar{a}}{2} \leq \psi_i(t,x) \leq 0 \quad \text{or} \quad \psi_i(t,x) > 0 \\ s(\psi_i(t,x)) & \text{if } -\bar{a} \leq \psi_i(t,x) < -\frac{\bar{a}}{2} \\ s(-\bar{a}) & \text{if } \psi_i(t,x) < -\bar{a}, \end{cases}$$

where

$$s(z) := -\frac{3}{4}\bar{a} + \frac{1}{\bar{a}}(z+\bar{a})^2, \text{ for } -\bar{a} \leq z \leq -\frac{\bar{a}}{2}.$$

Notice that $s(\cdot)$ is a quadratic function with:

- $s(-\bar{a}) = -\frac{3}{4}\bar{a}$ and $s(-\frac{\bar{a}}{2}) = -\frac{\bar{a}}{2}$.
- $s'(-\bar{a}) = 0$ and $s'(-\frac{\bar{a}}{2}) = 1$.

- $0 \leq s'(z) \leq 1$ for all $-\bar{a} \leq z \leq -\frac{\bar{a}}{2}$.

We also have

$$\nabla_x \hat{\psi}_i(t,x) := \begin{cases} \nabla_x \psi_i(t,x) & \text{if } -\frac{\bar{a}}{2} \leq \psi_i(t,x) \leq 0 \quad \text{or} \quad \psi_i(t,x) > 0 \\ s'(\psi_i(t,x)).\nabla_x \psi_i(t,x) & \text{if } -\bar{a} \leq \psi_i(t,x) < -\frac{\bar{a}}{2} \\ 0 & \text{if } \psi_i(t,x) < -\bar{a}. \end{cases}$$

Notice the following:

- $\{x \in \mathbb{R}^n : \hat{\psi}_i(t,x) \leq 0, \forall i = 1, \cdots, r\} = \{x \in \mathbb{R}^n : \psi_i(t,x) \leq 0, \forall i = 1, \cdots, r\} = C(t)$.
- Since $\{\psi_i\}_{i=1}^r$ satisfy (A3.1) and (A3.2), then $\{\hat{\psi}_i\}_{i=1}^r$ satisfy (A3.1) and (A3.2) with $L_{\hat{\psi}} = L_\psi(1 + \frac{2}{\bar{a}} L_\psi)$ replacing L_ψ.
- All results of Subsection 4.2.2, including Proposition 4.2.8 and Proposition 4.2.9, can now be formulated in terms of $\hat{\psi}_i$ $(i = 1, \cdots, r)$ instead of ψ_i $(i = 1, \cdots, r)$.
- Since $\{\psi_i\}_{i=1}^r$ satisfy (A3.3)$'$, and equation (4.4) is satisfied, we deduce that $\forall (t,x) \in \operatorname{Gr} C(\cdot) \cap \bar{B}_{\bar{c}}(\bar{x}(\cdot))$ with $\mathcal{I}^{\frac{\bar{a}}{2}}_{(t,x)} \neq \emptyset$, and $\forall\, i \in \mathcal{I}^{\frac{\bar{a}}{2}}_{(t,x)}$, we have

$$\sum_{\substack{j \in \mathcal{I}^{\bar{a}}_{(t,x)} \\ j \neq i}} \left| \langle \nabla_x \hat{\psi}_j(t,x), \nabla_x \hat{\psi}_i(t,x) \rangle \right| \leq \bar{b} \|\nabla_x \hat{\psi}_i(t,x)\|^2. \tag{4.37}$$

  This is due to the fact that $\hat{\psi}_i(t,x) = \psi_i(t,x)$ for $i \in \mathcal{I}^{\frac{\bar{a}}{2}}_{(t,x)}$, and $s'(z) \leq 1 \;\; \forall -\bar{a} \leq z \leq -\frac{\bar{a}}{2}$.

Step I.1. Results from Proposition 4.2.8 and Proposition 4.2.9 and formulating the primal-dual admissible equation for fixed (α, β).

Fix $\alpha > 0$ and $\beta \in (0,1]$. Recall from proposition 4.2.8 that there exist a subsequence of $(\gamma_k)_k$ (we do not relabel), an optimal solution $((x_{\gamma_k}, y_{\gamma_k}), u_{\gamma_k})$ for $(P^{\alpha,\beta}_{\gamma_k})$ with corresponding $(\hat{\xi}^1_{\gamma_k}, \cdots, \xi^{r+1}_{\gamma_k}, \zeta_{\gamma_k})$ via (3.76), and $(\hat{\xi}^1, \cdots, \hat{\xi}^r) \in L^\infty([0,T]; \mathbb{R}^r_+)$, such that (4.21)-(4.25) hold and $((\bar{x},\bar{y}),\bar{u})$ together with $(\hat{\xi}^1,\cdots,\hat{\xi}^r)$ satisfies the primal-dual admissible equation

$$\boxed{\begin{cases}\dot{\bar{x}}(t)=f(t,\bar{x}(t),\bar{y}(t),\bar{u}(t))-\sum_{i=1}^r\hat{\xi}^i(t)\nabla_x\hat{\psi}_i(t,\bar{x}(t)) \text{ a.e. } t\in[0,T],\\ \dot{\bar{y}}(t)=g(t,\bar{x}(t),\bar{y}(t),\bar{u}(t)) \text{ a.e. } t\in[0,T],\\ \hat{\psi}_i(t,\bar{x}(t))\le 0,\ \forall t\in[0,T],\quad \forall i\in\{1,\cdots,r\}.\end{cases}} \tag{4.38}$$

Moreover, Proposition 4.2.9 produces $\forall k\in\mathbb{N}$, $\hat{p}_{\gamma_k}=(\hat{q}_{\gamma_k},\hat{v}_{\gamma_k})\in W^{1,1}([0,T];\mathbb{R}^n\times\mathbb{R}^l)$, and $\hat{\lambda}_{\gamma_k}\ge 0$ such that equations (4.30)-(4.33) are valid. For simplicity, the (α,β)-dependency shall only be made visible at the stage when the limit in (α,β) is performed.

Since $(x_{\gamma_k}(t),y_{\gamma_k}(t))\in \operatorname{int}\left(C(t)\cap\bar{B}_{\bar{\delta}}(\bar{x}(t))\right)\times B_{\bar{\delta}}(\bar{y}(t))$ for all $t\in[0,T]$, then

$$\begin{aligned}\partial^{(x,y)}(f,g)(t,x_{\gamma_k}(t),y_{\gamma_k}(t),u_{\gamma_k}(t)) &= \partial_\ell^{(x,y)}(f,g)(t,x_{\gamma_k}(t),y_{\gamma_k}(t),u_{\gamma_k}(t)),\\ \partial^{(x,y)}(f,g)(t,x_{\gamma_k}(t),y_{\gamma_k}(t),\bar{u}(t)) &= \partial_\ell^{(x,y)}(f,g)(t,x_{\gamma_k}(t),y_{\gamma_k}(t),\bar{u}(t))\\ \partial^{xx}\hat{\psi}_i(t,x_{\gamma_k}(t)) &= \partial_\ell^{xx}\ \hat{\psi}_i(t,x_{\gamma_k}(t)) \text{ for } i=1,\cdots,r,\\ \partial^{xx}\psi_{r+1}(t,x_{\gamma_k}(t)) &= \partial_\ell^{xx}\ \psi_{r+1}(t,x_{\gamma_k}(t)).\end{aligned}$$

Also, $(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T))\in \operatorname{int} S(\bar{\delta})$, yields

$$\partial^L J(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T))=\partial_\ell^L J(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T)).$$

Using $(A4.1)$, first equation of (4.23), and Filippov Selection Theorem (Theorem 2.3.5), equation (4.33) yields the existence of measurable $\hat{\bar{A}}_{\gamma_k}(\cdot)$, $\hat{A}_{\gamma_k}(\cdot)$ in $\mathcal{M}_{n\times n}[0,T]$, $\hat{\bar{E}}_{\gamma_k}(\cdot)$, $\hat{E}_{\gamma_k}(\cdot)$ in $\mathcal{M}_{n\times l}[0,T]$, $\hat{\bar{\mathcal{A}}}_{\gamma_k}(\cdot)$, $\hat{\mathcal{A}}_{\gamma_k}(\cdot)$ in $\mathcal{M}_{l\times n}[0,T]$, $\hat{\mathcal{E}}_{\gamma_k}(\cdot)$, $\hat{\bar{\mathcal{E}}}_{\gamma_k}(\cdot)$ in $\mathcal{M}_{l\times l}[0,T]$, $\hat{\vartheta}^i_{\gamma_k}(\cdot)$, $\vartheta^{r+1}_{\gamma_k}(\cdot)$ in $\mathcal{M}_{n\times n}[0,T]$ such that for almost all $t\in[0,T]$,

$$\begin{aligned}&(\hat{\bar{A}}_{\gamma_k},\hat{\bar{E}}_{\gamma_k})(t)\in\partial_\ell^{(x,y)}f(t,x_{\gamma_k}(t),y_{\gamma_k}(t),\bar{u}(t)),\ \ (\hat{A}_{\gamma_k},\hat{E}_{\gamma_k})(t)\in\partial_\ell^{(x,y)}f(t,x_{\gamma_k}(t),y_{\gamma_k}(t),u_{\gamma_k}(t));\\ &(\hat{\bar{\mathcal{A}}}_{\gamma_k},\hat{\bar{\mathcal{E}}}_{\gamma_k})(t)\in\partial_\ell^{(x,y)}g(t,x_{\gamma_k}(t),y_{\gamma_k}(t),\bar{u}(t)),\ \ (\hat{\mathcal{A}}_{\gamma_k},\hat{\mathcal{E}}_{\gamma_k})(t)\in\partial_\ell^{(x,y)}g(t,x_{\gamma_k}(t),y_{\gamma_k}(t),u_{\gamma_k}(t));\\ &\hat{\vartheta}^i_{\gamma_k}(t)\in\partial_\ell^{xx}\hat{\psi}_i(t,x_{\gamma_k}(t)) \text{ for } i=1,\cdots,r,\qquad \vartheta^{r+1}_{\gamma_k}(t)=\mathrm{I}_{n\times n};\\ &\max\left\{\|(\hat{\bar{A}}_{\gamma_k},\hat{\bar{E}}_{\gamma_k})\|_2,\ \|(\hat{A}_{\gamma_k},\hat{E}_{\gamma_k})\|_2,\ \|(\hat{\bar{\mathcal{A}}}_{\gamma_k},\ \hat{\bar{\mathcal{E}}}_{\gamma_k})\|_2,\ \|(\hat{\mathcal{A}}_{\gamma_k},\hat{\mathcal{E}}_{\gamma_k})\|_2\right\}\le\|L_h\|_2;\\ &\|\hat{\vartheta}^i_{\gamma_k}\|_\infty\le L_{\hat{\psi}} \text{ for } i=1,\cdots,r,\qquad \|\vartheta^{r+1}_{\gamma_k}\|_\infty=1,\quad\text{and}\end{aligned}$$

$$\dot{\hat{q}}_{\gamma_k}(t) = \underbrace{-\Big[(1-\beta)\hat{\bar{A}}^T_{\gamma_k}(t) + \beta\hat{A}^T_{\gamma_k}(t)\Big]\hat{q}_{\gamma_k}(t) - \left[(1-\beta)\hat{\bar{\mathcal{A}}}_{\gamma_k}(t)^T + \beta\hat{\mathcal{A}}^T_{\gamma_k}(t)\right]\hat{v}_{\gamma_k}(t)}_{\mathcal{Q}_{\gamma_k(t)}}$$
$$+\underbrace{\sum_{i=1}^{r}\gamma_k e^{\gamma_k\hat{\psi}_i(t,x_{\gamma_k}(t))}\hat{\vartheta}^i_{\gamma_k}(t)\hat{q}_{\gamma_k}(t) + \gamma_k e^{\gamma_k\psi_{r+1}(t,x_{\gamma_k}(t))}\hat{q}_{\gamma_k}(t)}_{\mathcal{X}_{\gamma_k(t)}}$$
$$+\underbrace{\sum_{i=1}^{r}\gamma_k^2 e^{\gamma_k\hat{\psi}_i(t,x_{\gamma_k}(t))}\nabla_x\hat{\psi}_i(t,x_{\gamma_k}(t))\langle\nabla_x\hat{\psi}_i(t,x_{\gamma_k}(t)),\hat{q}_{\gamma_k}(t)\rangle}_{\mathcal{Y}_{\gamma_k}(t)}$$
$$+\underbrace{\gamma_k^2 e^{\gamma_k\psi_{r+1}(t,x_{\gamma_k}(t))}\nabla_x\psi_{r+1}(t,x_{\gamma_k}(t))\langle\nabla_x\psi_{r+1}(t,x_{\gamma_k}(t)),\hat{q}_{\gamma_k}(t)\rangle}_{\mathcal{Z}_{\gamma_k}(t)}, \quad (4.39)$$
$$\dot{\hat{v}}_{\gamma_k}(t) = -\left[(1-\beta)\hat{\bar{E}}_{\gamma_k}(t)^T + \beta\hat{E}_{\gamma_k}(t)^T\right]\hat{q}_{\gamma_k}(t) - \left[(1-\beta)\hat{\bar{\mathcal{E}}}_{\gamma_k}(t)^T + \beta\hat{\mathcal{E}}_{\gamma_k}(t)^T\right]\hat{v}_{\gamma_k}(t)$$
$$+\gamma_k e^{\gamma_k\varphi(t,y_{\gamma_k}(t))}\hat{v}_{\gamma_k}(t) + \gamma_k^2 e^{\gamma_k\varphi(t,y_{\gamma_k}(t))}\nabla_y\varphi(t,y_{\gamma_k}(t))\langle\nabla_y\varphi(t,y_{\gamma_k}(t)),\hat{v}_{\gamma_k}(t)\rangle. \quad (4.40)$$

Step I.2. Uniform boundedness of $\{\hat{p}_{\gamma_k}\}$, $\{\|\dot{\hat{v}}_{\gamma_k}\|_2\}$, and $\{\|\dot{\hat{q}}_{\gamma_k}\|_1\}$.

The proof of this step is a generalization to our general setting of the proof for the corresponding step in [58, Theorem 3.1]. We first start by proving that $\{\hat{p}_{\gamma_k}\}$ is uniformly bounded. We have

$$\frac{1}{2}\frac{d}{dt}\|\hat{p}_{\gamma_k}(t)\|^2 = \langle\hat{q}_{\gamma_k}(t),\dot{\hat{q}}_{\gamma_k}(t)\rangle + \langle\hat{v}_{\gamma_k}(t),\dot{\hat{v}}_{\gamma_k}(t)\rangle$$
$$\overset{(4.39)+(4.40)}{=} \left\langle\hat{q}_{\gamma_k}(t), -[\beta\hat{A}_{\gamma_k}(t)^T + (1-\beta)\hat{\bar{A}}_{\gamma_k}(t)^T]\hat{q}_{\gamma_k}(t) - [\beta\hat{\mathcal{A}}_{\gamma_k}(t)^T + (1-\beta)\hat{\bar{\mathcal{A}}}_{\gamma_k}(t)^T]\hat{v}_{\gamma_k}(t)\right\rangle$$
$$+\sum_{i=1}^{r}\gamma_k e^{\gamma_k\hat{\psi}_i(t,x_{\gamma_k}(t))}\left[\langle\hat{q}_{\gamma_k}(t),\hat{\vartheta}^i_{\gamma_k}(t)\hat{q}_{\gamma_k}(t)\rangle + \underbrace{\gamma_k|\langle\hat{q}_{\gamma_k}(t),\nabla_x\hat{\psi}_i(t,x_{\gamma_k}(t))\rangle|^2}_{\text{positive term}}\right]$$
$$+\underbrace{\gamma_k e^{\gamma_k\psi_{r+1}(t,x_{\gamma_k}(t))}\|\hat{q}_{\gamma_k}(t)\|^2 + \gamma_k^2 e^{\gamma_k\psi_{r+1}(t,x_{\gamma_k}(t))}|\langle\hat{q}_{\gamma_k}(t),\nabla_x\psi_{r+1}(t,x_{\gamma_k}(t))\rangle|^2}_{\text{positive term}}$$
$$+\left\langle\hat{v}_{\gamma_k}(t), -[\beta\hat{E}_{\gamma_k}(t)^T + (1-\beta)\hat{\bar{E}}_{\gamma_k}(t)^T]\hat{q}_{\gamma_k}(t) - [\beta\hat{\mathcal{E}}_{\gamma_k}(t)^T + (1-\beta)\hat{\bar{\mathcal{E}}}_{\gamma_k}(t)^T]\hat{v}_{\gamma_k}(t)\right\rangle$$
$$+\underbrace{\gamma_k e^{\gamma_k\varphi(t,y_{\gamma_k}(t))}\|\hat{v}_{\gamma_k}(t)\|^2 + \gamma_k^2 e^{\gamma_k\varphi(t,y_{\gamma_k}(t))}|\langle\hat{v}_{\gamma_k}(t),\nabla_y\varphi(t,y_{\gamma_k}(t))\rangle|^2}_{\text{positive term}}$$
$$\geq \left[-2L_h(t) - L_{\hat{\psi}}\frac{2\bar{\mu}}{\bar{\eta}^2}\right]\|\hat{p}_{\gamma_k}(t)\|^2 := -L_p(t)\|\hat{p}_{\gamma_k}(t)\|^2,$$

where (3.75) is employed and $L_p(\cdot) \in L^2([0,T],\mathbb{R}_+)$. Using Gronwall's Lemma (Lemma

2.4.1), we deduce that there exists a constant $M_p > 0$ such that

$$\|\hat{p}_{\gamma_k}(t)\| \leq e^{\|L_p(\cdot)\|_1}\|\hat{p}_{\gamma_k}(T)\| \leq M_p, \quad \forall t \in [0,T], \quad \forall k \in \mathbb{N}, \tag{4.41}$$

where the last inequality is due to the uniform boundedness of $\|\hat{p}_{\gamma_k}(T)\|$ obtained from the nontriviality condition (4.36) when S has a general form, and to the transversality condition (4.31), $\hat{\lambda}_{\gamma_k} = 1$, and equation (4.20), when $S = C_0 \times \mathbb{R}^{n+l}$.

We proceed to prove the uniform boundedness of $\{\|\dot{\hat{v}}_{\gamma_k}\|_2\}$ and $\{\|\dot{\hat{q}}_{\gamma_k}\|_1\}$. From (4.40), (4.41), (3.75), and (4.29), there exist $L_v(\cdot) \in L^2([0,T], \mathbb{R}_+)$ and $k_v \in \mathbb{N}$, such that for $k \geq k_v$ we have

$$\begin{aligned} \|\dot{\hat{v}}_{\gamma_k}(t)\| \quad &\leq \quad \| \left[(1-\beta)\dot{\hat{E}}_{\gamma_k}(t)^T + \beta \hat{E}_{\gamma_k}(t)^T\right] \hat{q}_{\gamma_k}(t) + \left[(1-\beta)\dot{\hat{\mathcal{E}}}_{\gamma_k}(t)^T + \beta \hat{\mathcal{E}}_{\gamma_k}(t)^T\right] \hat{v}_{\gamma_k}(t)\| \\ &+ \quad \frac{2\bar{\mu}}{\bar{\eta}^2} M_p + \gamma_k^2 e^{-\gamma_k \frac{\bar{\delta}^2}{4}} \bar{\delta}^2 M_p \leq L_v(t) M_p, \quad \forall t \in [0,T]. \end{aligned}$$

Thus, for all $k \geq k_v$, $(\dot{\hat{v}}_{\gamma_k})$ is uniformly bounded in L^2 by a constant M_v.

We now proceed to prove that $(\dot{\hat{q}}_{\gamma_k})$ is uniformly bounded in L^1. Observe that (4.29) together with (4.22) and (4.41), yields that for some $\bar{k}_1 \in \mathbb{N}$, $\bar{k}_1 \geq k_v$, we have

$$\|\mathcal{Q}_{\gamma_k}(t)\| \leq 2L_h(t)M_p; \quad \|\mathcal{X}_{\gamma_k}(t)\| \leq \frac{2\bar{\mu}}{\bar{\eta}^2} \max\{1, L_{\hat{\psi}}\} M_p; \quad \|\mathcal{Z}_{\gamma_k}(t)\| \leq \gamma_k^2 \, e^{-\gamma_k \frac{\bar{\varepsilon}^2}{4}} \, \bar{\varepsilon}^2 \, M_p. \tag{4.42}$$

Hence, using (4.39) and (4.42), we can see $\{\hat{q}_{\gamma_k}\}$ is of uniformly bounded variation once we prove

$$\int_0^T \|\mathcal{Y}_{\gamma_k}(t)\| dt = \int_0^T \sum_{i=1}^{r} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \|\nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t))\| \left|\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle\right| dt$$

is uniformly bounded.

Denote by

$$\mathcal{I}_k^{\bar{a}} = \mathcal{I}^{\bar{a}}_{(t, x_{\gamma_k}(t))} = \{i \in \{1, \ldots, r\} : -\bar{a} \leq \psi_i(t, x_{\gamma_k}(t)) \leq 0\} \tag{4.43}$$

and define

$$I^{\bar{a}}(x_{\gamma_k}) := \{t \in [0,T] : \mathcal{I}^{\bar{a}}_{(t, x_{\gamma_k}(t))} \neq \emptyset\}. \tag{4.44}$$

Using the definition of $I^{\bar a}(x_{\gamma_k})$, $\mathcal{I}_k^{\bar a}$ and $\mathcal{I}_k^{\frac{\bar a}{2}}$, we deduce that

$$\forall t \in [I^{\bar a}(x_{\gamma_k})]^c,\ \ \forall i = 1, \cdots, r,\ \ \hat\psi_i(t, x_{\gamma_k}(t)) = -\frac{3\bar a}{4},\quad \nabla_x \hat\psi_i(t, x_{\gamma_k}(t)) = 0, \tag{4.45}$$

$$\forall t \in I^{\bar a}(x_{\gamma_k}),\ \ \forall i \in [\mathcal{I}_k^{\bar a}]^c,\ \ \hat\psi_i(t, x_{\gamma_k}(t)) = -\frac{3\bar a}{4},\quad \nabla_x \hat\psi_i(t, x_{\gamma_k}(t)) = 0, \tag{4.46}$$

$$\forall t \in I^{\bar a}(x_{\gamma_k}),\ \forall i \in \mathcal{I}_k^{\frac{\bar a}{2}},\ \ \hat\psi_i(t, x_{\gamma_k}(t)) = \psi_i(t, x_{\gamma_k}(t)),\ \nabla_x \hat\psi_i(t, x_{\gamma_k}(t)) = \nabla_x \psi_i(t, x_{\gamma_k}(t)), \tag{4.47}$$

$$\forall t \in I^{\bar a}(x_{\gamma_k}),\ \ \forall i \in \mathcal{I}_k^{\bar a} \setminus \mathcal{I}_k^{\frac{\bar a}{2}},$$

$$\begin{cases} \hat\psi_i(t, x_{\gamma_k}(t)) < -\frac{\bar a}{2}, \\ \nabla_x \hat\psi_i(t, x_{\gamma_k}(t)) = s'(\psi_i(t, x_{\gamma_k}(t)))\nabla_x \psi_i(t, x_{\gamma_k}(t)). \end{cases} \tag{4.48}$$

As a result of (4.45)-(4.48) and the fact that $0 \le s'(z) \le 1$ for all $-\bar a \le z \le -\frac{\bar a}{2}$, to prove $\int_0^T \|\mathcal{Y}_{\gamma_k}(t)\| dt$ is uniformly bounded, it remains to prove that

$$\mathbf{I}_1 := \int_{I^{\bar a}(x_{\gamma_k})} \sum_{i \in \mathcal{I}_k^{\frac{\bar a}{2}}} \gamma_k^2 e^{\gamma_k \psi_i(t, x_{\gamma_k}(t))} \|\nabla_x \psi_i(t, x_{\gamma_k}(t))\| |\langle \nabla_x \psi_i(t, x_{\gamma_k}(t)), \hat q_{\gamma_k}(t)\rangle| dt \le M_1, \tag{4.49}$$

for a certain constant $M_1 > 0$. For that, it is sufficient to prove that there exists $M_2 > 0$ such that

$$\mathbf{I}_2 := \int_{I^{\bar a}(x_{\gamma_k})} \sum_{i \in \mathcal{I}_k^{\frac{\bar a}{2}}} \gamma_k^2 e^{\gamma_k \psi_i(t, x_{\gamma_k}(t))} \|\nabla_x \psi_i(t, x_{\gamma_k}(t))\|^2 |\langle \nabla_x \psi_i(t, x_{\gamma_k}(t)), \hat q_{\gamma_k}(t)\rangle| dt \le M_2. \tag{4.50}$$

Indeed, for $t \in I^{\bar a}(x_{\gamma_k})$ and $i \in \mathcal{I}_k^{\frac{\bar a}{2}}$, we have $\psi_i(t, x_{\gamma_k}(t)) \ge -\frac{\bar a}{2} \ge -a_o$, and hence the uniform convergence of x_{γ_k} to $\bar x$ and Lemma 4.2.3 yield the existence of $\bar k_2 \in \mathbb{N}$ such that for all $k \ge \bar k_2$, we have $\|\nabla_x \psi_i(t, x_{\gamma_k}(t))\| > \bar\eta$. Thus, if $\mathbf{I}_2$ is uniformly bounded by a constant M_2, then it follows that $\mathbf{I}_1 \le \frac{M_2}{\bar\eta}$, for k large enough.

We proceed to prove that (4.50) holds true. Using Lemma 2.4.2, we first calculate for each $j = 1, \cdots, r$ and $t \in [0, T]$:

$$\begin{aligned} \frac{d}{dt}\left|\langle \hat q_{\gamma_k}(t), \nabla_x \hat\psi_j(t, x_{\gamma_k}(t))\rangle\right| &= \langle \dot{\hat q}_{\gamma_k}(t), \nabla_x \hat\psi_j(t, x_{\gamma_k}(t))\rangle s^j_{\gamma_k}(t) \\ &+ \langle \hat q_{\gamma_k}(t), \Theta^j_{\gamma_k}(t).(1, \dot x_{\gamma_k}(t))\rangle s^j_{\gamma_k}(t), \ \text{ a.e.} \end{aligned} \tag{4.51}$$

where $s^j_{\gamma_k}(t)$ is the sign of $\langle \hat{q}_{\gamma_k}(t), \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t))\rangle$ and $\Theta^j_{\gamma_k}(t) \in \partial^{(t,x)} \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t))$.

Using equation (4.39) in (4.51), we get

$$\begin{aligned}
&\frac{d}{dt}\left|\langle \hat{q}_{\gamma_k}(t), \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t))\rangle\right| = \\
&\langle \mathcal{Q}_{\gamma_k}(t) + \mathcal{X}_{\gamma_k}(t) + \mathcal{Z}_{\gamma_k}(t), \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t))\rangle s^j_{\gamma_k}(t) + \langle \hat{q}_{\gamma_k}(t), \Theta^j_{\gamma_k}(t).(1, \dot{x}_{\gamma_k}(t))\rangle s^j_{\gamma_k}(t) \\
&+ \sum_{i=1}^{r} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i, \nabla_x \hat{\psi}_j\rangle|_{(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle s^j_{\gamma_k}(t) \text{ a.e.} \qquad (4.52)
\end{aligned}$$

Let $t \in I^{\bar{a}}(x_{\gamma_k})$. Summing the previous equality over $j \in \mathcal{I}^{\bar{a}}_k$, we obtain that:

$$\begin{aligned}
\mathbf{J}_1 &:= \sum_{j \in \mathcal{I}^{\bar{a}}_k} \sum_{i=1}^{r} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i, \nabla_x \hat{\psi}_j\rangle|_{(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle s^j_{\gamma_k}(t) \\
&= \sum_{j \in \mathcal{I}^{\bar{a}}_k} \frac{d}{dt}\left|\langle \hat{q}_{\gamma_k}(t), \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t))\rangle\right| - \sum_{j \in \mathcal{I}^{\bar{a}}_k} \langle \mathcal{Q}_{\gamma_k}(t) + \mathcal{X}_{\gamma_k}(t) + \mathcal{Z}_{\gamma_k}(t), \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t))\rangle s^j_{\gamma_k}(t) \\
&- \sum_{j \in \mathcal{I}^{\bar{a}}_k} \langle \hat{q}_{\gamma_k}(t), \Theta^j_{\gamma_k}(t).(1, \dot{x}_{\gamma_k}(t))\rangle s^j_{\gamma_k}(t) \text{ a.e.} \qquad (4.53)
\end{aligned}$$

On the other hand, splitting in the definition of $\mathbf{J}_1$ the summation over i and switching the order of summation between i and j, we have

$$\begin{aligned}
\mathbf{J}_1 &= \sum_{i=1}^{r} \sum_{j \in \mathcal{I}^{\bar{a}}_k} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i, \nabla_x \hat{\psi}_j\rangle|_{(t, x_{\gamma_k}(t))} \ \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle s^j_{\gamma_k}(t) \\
&= \sum_{i \in \mathcal{I}^{\bar{a}}_k} \sum_{j \in \mathcal{I}^{\bar{a}}_k} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i, \nabla_x \hat{\psi}_j\rangle|_{(t, x_{\gamma_k}(t))} \ \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle s^j_{\gamma_k}(t) \\
&= \sum_{i \in \mathcal{I}^{\frac{\bar{a}}{2}}_k} \sum_{j \in \mathcal{I}^{\bar{a}}_k} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i, \nabla_x \hat{\psi}_j\rangle|_{(t, x_{\gamma_k}(t))} \ \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle s^j_{\gamma_k}(t) \\
&+ \sum_{i \in \mathcal{I}^{\bar{a}}_k \setminus \mathcal{I}^{\frac{\bar{a}}{2}}_k} \sum_{j \in \mathcal{I}^{\bar{a}}_k} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i, \nabla_x \hat{\psi}_j\rangle|_{(t, x_{\gamma_k}(t))} \ \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle s^j_{\gamma_k}(t) \\
&= \sum_{i \in \mathcal{I}^{\frac{\bar{a}}{2}}_k} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \Bigg(\|\nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t))\|^2 + \\
&\sum_{\substack{j \in \mathcal{I}^{\bar{a}}_k \\ j \neq i}} s^j_{\gamma_k}(t) s^i_{\gamma_k}(t) \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t))\rangle \Bigg) \ |\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle| \\
&+ \sum_{i \in \mathcal{I}^{\bar{a}}_k \setminus \mathcal{I}^{\frac{\bar{a}}{2}}_k} \sum_{j \in \mathcal{I}^{\bar{a}}_k} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i, \nabla_x \hat{\psi}_j\rangle|_{(t, x_{\gamma_k}(t))} \ \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle s^j_{\gamma_k}(t).
\end{aligned}$$

Using the fact that x_{γ_k} converges uniformly to $\bar{x}$, we deduce from equation (4.37) that, there exists $\bar{k}_3 \in \mathbb{N}$ such that for $k \geq \bar{k}_3$, we have for $i \in \mathcal{I}_k^{\frac{\bar{a}}{2}}$,

$$\sum_{\substack{j \in \mathcal{I}_k^{\bar{a}} \\ j \neq i}} s^j_{\gamma_k}(t) s^i_{\gamma_k}(t) \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t)) \rangle \geq -\bar{b} \|\nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t))\|^2.$$

Then,

$$\begin{aligned}\mathbf{J}_1 \geq (1-\bar{b}) \sum_{i \in \mathcal{I}_k^{\frac{\bar{a}}{2}}} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \|\nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t))\|^2 |\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t) \rangle| \\ + \underbrace{\sum_{i \in \mathcal{I}_k^{\bar{a}} \setminus \mathcal{I}_k^{\frac{\bar{a}}{2}}} \sum_{j \in \mathcal{I}_k^{\bar{a}}} \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle \nabla_x \hat{\psi}_i, \nabla_x \hat{\psi}_j \rangle|_{(t, x_{\gamma_k}(t))} \ \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t) \rangle s^j_{\gamma_k}(t)}_{\mathbf{J}_2}.\end{aligned}$$

Hence, (recalling (4.47)) we have

$$\sum_{i \in \mathcal{I}_k^{\frac{\bar{a}}{2}}} \gamma_k^2 e^{\gamma_k \psi_i(t, x_{\gamma_k}(t))} \|\nabla_x \psi_i(t, x_{\gamma_k}(t))\|^2 |\langle \nabla_x \psi_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t) \rangle| \leq \frac{1}{1-\bar{b}} (\mathbf{J}_1 - \mathbf{J}_2). \tag{4.54}$$

Integrating the last inequality over $I^{\bar{a}}(x_{\gamma_k})$, we deduce from the definition of $\mathbf{I}_2$ that

$$0 \leq \mathbf{I}_2 \leq \frac{1}{1-\bar{b}} \int_{I^{\bar{a}}(x_{\gamma_k})} (\mathbf{J}_1 - \mathbf{J}_2)\, dt \leq \frac{1}{1-\bar{b}} \left| \int_{I^{\bar{a}}(x_{\gamma_k})} \mathbf{J}_1 \, dt \right| + \frac{1}{1-\bar{b}} \int_{I^{\bar{a}}(x_{\gamma_k})} |\mathbf{J}_2| dt.$$

Using (4.48), we deduce that, there exists $\bar{k}_4 \in \mathbb{N}$, there exists constant $M_3 > 0$ such that for all $k \geq \bar{k}_4$, we have

$$\int_{I^{\bar{a}}(x_{\gamma_k})} |\mathbf{J}_2| dt \leq \frac{\gamma_k^2 e^{-\gamma_k \frac{\bar{a}}{2}} L_\psi^3 r^2 M_p T}{1-\bar{b}} \leq M_3. \tag{4.55}$$

Hence,

$$0 \leq \mathbf{I}_2 \leq \frac{1}{1-\bar{b}} \left| \int_{I^{\bar{a}}(x_{\gamma_k})} \mathbf{J}_1 \, dt \right| + M_3, \quad \forall k \geq \bar{k}_4. \tag{4.56}$$

Note that, (4.53) yields that the uniform boundedness of $\left| \int_{I^{\bar{a}}(x_{\gamma_k})} \mathbf{J}_1 \, dt \right|$ is equivalent to that of

$$\Big| \int_{I^{\bar{a}}(x_{\gamma_k})} \sum_{j \in \mathcal{I}_k^{\bar{a}}} \frac{d}{dt} |\langle \hat{q}_{\gamma_k}(t), \nabla_x \hat{\psi}_j(t, x_{\gamma_k}(t)) \rangle| \, dt \Big|,$$

since

$$\Big|\int_{I^{\bar a}(x_{\gamma_k})}\sum_{j\in\mathcal{I}^{\bar a}_k}\langle \hat q_{\gamma_k}(t),\Theta^j_{\gamma_k}(t).(1,\dot x_{\gamma_k}(t))\rangle s^j_{\gamma_k}(t)dt\Big|\le M_pL_\psi(1+M_h+\frac{2\bar\mu}{\bar\eta^2}\bar L)rT,$$

$$\Big|\int_{I^{\bar a}(x_{\gamma_k})}\sum_{j\in\mathcal{I}^{\bar a}_k}\langle \mathcal{Q}_{\gamma_k}(t)+\mathcal{X}_{\gamma_k}(t)+\mathcal{Z}_{\gamma_k}(t),\nabla_x\hat\psi_j(t,x_{\gamma_k}(t))\rangle s^j_{\gamma_k}(t)dt\Big|$$

$$\le\left[(2L_h(t)M_p)+(\frac{2\bar\mu}{\bar\eta^2}\max\{L_\psi,1\}M_p)+(\gamma_k^2e^{-\gamma_k\frac{\bar\varepsilon^2}{4}}\bar\varepsilon^2M_p)\right]L_\psi rT.$$

We now proceed to prove the boundedness of

$$\Big|\int_{I^{\bar a}(x_{\gamma_k})}\sum_{j\in\mathcal{I}^{\bar a}_k}\frac{d}{dt}|\langle \hat q_{\gamma_k}(t),\nabla_x\hat\psi_j(t,x_{\gamma_k}(t))\rangle|\;dt\Big|.$$

Using the Fundamental Theorem of Calculus, we have that

$$\Big|\int_0^T\sum_{j=1}^r\frac{d}{dt}|\langle \hat q_{\gamma_k}(t),\nabla_x\hat\psi_j(t,x_{\gamma_k}(t))\rangle|\;dt\Big|\le 2rL_\psi M_p. \tag{4.57}$$

Using (4.45), (4.46), (4.52), and the uniform boundedness of $(\dot x_{\gamma_k})$, we deduce that that there exists a constant $M_4>0$ such that

$$\Big|\int_{[I^{\bar a}(x_{\gamma_k})]^c}\sum_{j=1}^r\frac{d}{dt}|\langle \hat q_{\gamma_k}(t),\nabla_x\hat\psi_j(t,x_{\gamma_k}(t))\rangle|\;dt\Big|\le M_4, \tag{4.58}$$

$$\Big|\int_{I^{\bar a}(x_{\gamma_k})}\sum_{j\in[\mathcal{I}^{\bar a}_k]^c}\frac{d}{dt}|\langle \hat q_{\gamma_k}(t),\nabla_x\hat\psi_j(t,x_{\gamma_k}(t))\rangle|\;dt\Big|\le M_4. \tag{4.59}$$

Hence, combining (4.57) and (4.58), we conclude that there exists a constant $M_5>0$ such that

$$\Big|\int_{I^{\bar a}(x_{\gamma_k})}\sum_{j=1}^r\frac{d}{dt}|\langle \hat q_{\gamma_k}(t),\nabla_x\hat\psi_j(t,x_{\gamma_k}(t))\rangle|\;dt\Big|\le M_5.$$

This last inequality with (4.59) yield that there exists a constant $M_6>0$ such that

$$\Big|\int_{I^{\bar a}(x_{\gamma_k})}\sum_{j\in\mathcal{I}^{\bar a}_k}\frac{d}{dt}|\langle \hat q_{\gamma_k}(t),\nabla_x\hat\psi_j(t,x_{\gamma_k}(t))\rangle|\;dt\Big|\le M_6.$$

Hence, $\left|\int_{I^{\bar a}(\bar x)}\mathbf{J}_1\;dt\right|$ is uniformly bounded, and by (4.56), $\mathbf{I}_2$ is uniformly bounded. Hence, $\{\|\dot{\hat q}_{\gamma_k}\|_1\}$ uniformly bounded by a constant M_q.

Step I.3. Construction of $p=(q,v)$, $\lambda\ge 0$, ϑ^i (for each i), ν^i (for each i), $\bar A$, A, $\bar{\mathcal{A}}$,

$\mathcal{A}$, $\bar{E}$, $\bar{\mathcal{E}}$, E, $\mathcal{E}$ **for each fixed** (α, β) **satisfying some necessary conditions**.

In Step I.1, we proved the existence of $(\hat{\xi}^i)_{i=1}^r$ in $L^\infty([0,T];\mathbb{R}_+)$ such that condition (i) is satisfied. We now follow steps similar to steps 3-10 in the proof of [55, Theorem 6.1].

Step I.3.1 Construction of $\hat{p} = (\hat{q}, \hat{v})$**.**

From Step I.2, we find that $\hat{q}_{\gamma_k} \in W^{1,1}$ satisfies, for k large enough,

$$\|\hat{q}_{\gamma_k}\|_\infty \le M_p \quad \text{and} \quad V_0^1(\hat{q}_{\gamma_k}) = \|\dot{\hat{q}}_{\gamma_k}\|_1 \le M_q. \tag{4.60}$$

Hence, by Helly first theorem (see Theorem 2.4.6(i)), we deduce that $\hat{q}_{\gamma_k}$ admits a pointwise convergent subsequence, whose limit $\hat{q} \in BV([0,T];\mathbb{R}^n)$, with

$$\|\hat{q}\|_\infty \le M_p \quad \text{and} \quad V_0^1(\hat{q}) \le M_q. \tag{4.61}$$

By Helly second theorem (see Theorem 2.4.6(ii)), we deduce that for any $z \in ([0,T];\mathbb{R}^n)$, we have

$$\lim_{k\to\infty} \int_0^T \langle z(t), \dot{\hat{q}}_{\gamma_k}(t)\rangle dt = \int_{[0,T]} \langle z(t), d\hat{q}(t)\rangle. \tag{4.62}$$

By Step I.2, we also find that $\hat{v}_{\gamma_k} \in W^{1,2}$ satisfies, for k large enough,

$$\|\hat{v}_{\gamma_k}\|_\infty \le M_p \quad \text{and} \quad \|\dot{\hat{v}}_{\gamma_k}\|_2 \le M_v. \tag{4.63}$$

Hence, by Theorems 2.4.10-2.4.13, we deduce that $\hat{v}_{\gamma_k}$ admits a pointwise convergent subsequence to a function $\hat{v}(\cdot) \in W^{1,2}([0,T];\mathbb{R}^l)$ such that

$$\hat{v}_{\gamma_k}(\cdot) \xrightarrow{unif} \hat{v}(\cdot), \qquad \dot{\hat{v}}_{\gamma_k}(\cdot) \xrightarrow[L^2]{w} \dot{\hat{v}}(\cdot),$$
$$\|\hat{v}\|_\infty \le M_p, \qquad \|\dot{\hat{v}}\|_2 \le M_v,$$

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

$$\lim_{k\to\infty} \int_0^T \langle z(t), \dot{\hat{v}}_{\gamma_k}(t)\rangle dt = \int_{[0,T]} \langle z(t), \dot{\hat{v}}(t)\rangle dt. \tag{4.64}$$

Step I.3.2 Construction of $\hat{\bar{A}}$, $\hat{A}$, $\hat{\bar{\mathcal{A}}}$, $\hat{\mathcal{A}}$, $\hat{\bar{E}}$, $\hat{\bar{\mathcal{E}}}$, $\hat{E}$, $\hat{\mathcal{E}}$, $\hat{\vartheta}^i$ **(for** $i = 1, \cdots, r$**),** $\hat{\nu}^i$ **(for** $i = 1, \cdots, r$**) and formulating adjoint equations for fixed** (α, β)**.**

It follows from (4.62), that, for any $z(\cdot) \in \mathcal{C}([0,T],\mathbb{R}^n)$, we have

$$\begin{aligned}\int_{[0,T]} \langle z(t), d\hat{q}(t)\rangle &= \lim_{k\to\infty}\int_0^T \langle z(t), \dot{\hat{q}}_{\gamma_k}(t)\rangle dt\\ &= \lim_{k\to\infty}\int_0^T \langle z(t), \mathcal{Q}_{\gamma_k}(t)\rangle dt + \lim_{k\to\infty}\int_0^T \langle z(t), \mathcal{X}_{\gamma_k}(t)\rangle dt\\ &+ \lim_{k\to\infty}\int_0^T \langle z(t), \mathcal{Y}_{\gamma_k}(t)\rangle dt + \lim_{k\to\infty}\int_0^T \langle z(t), \mathcal{Z}_{\gamma_k}(t)\rangle dt.\end{aligned}$$

We will work on each of these limits above separately. Since

$$\max\left\{\|(\hat{\hat{A}}_{\gamma_k}, \hat{\hat{E}}_{\gamma_k})\|_2,\ \|(\hat{A}_{\gamma_k}, \hat{E}_{\gamma_k})\|_2,\ \|(\hat{\hat{\mathcal{A}}}_{\gamma_k},\ \hat{\hat{\mathcal{E}}}_{\gamma_k})\|_2,\ \|(\hat{\mathcal{A}}_{\gamma_k}, \hat{\mathcal{E}}_{\gamma_k})\|_2\right\} \le \|L_h\|_2$$

then, using Theorem 2.4.11, along a subsequence, we do not relabel, $(\hat{\hat{A}}_{\gamma_k}, \hat{\hat{E}}_{\gamma_k})$, $(\hat{A}_{\gamma_k}, \hat{E}_{\gamma_k})$, $(\hat{\hat{\mathcal{A}}}_{\gamma_k},\ \hat{\hat{\mathcal{E}}}_{\gamma_k})$, $(\hat{\mathcal{A}}_{\gamma_k}, \hat{\mathcal{E}}_{\gamma_k})$ converge weakly in L^2 to some $(\hat{\hat{A}}, \hat{\hat{E}})$, $(\hat{A}, \hat{E})$, $(\hat{\hat{\mathcal{A}}},\ \hat{\hat{\mathcal{E}}})$, $(\hat{\mathcal{A}}, \hat{\mathcal{E}})$ respectively. Using Theorem 2.4.15, we conclude that

$$\begin{aligned}(\hat{\hat{A}}, \hat{\hat{E}})(t) &\in \partial_\ell^{(x,y)} f(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)),\\ (\hat{\hat{\mathcal{A}}}, \hat{\hat{\mathcal{E}}})(t) &\in \partial_\ell^{(x,y)} g(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)).\end{aligned}$$

We also know that $\hat{q}_{\gamma_k}$ and $\hat{v}_{\gamma_k}$ are uniformly bounded in L^∞ and converge pointwise to $\hat{q}(\cdot)$ and $\hat{v}(\cdot)$ respectively. We then conclude using Theorem 2.4.12 that

$$\begin{aligned}\hat{\hat{A}}_{\gamma_k}(t)^T \hat{q}_{\gamma_k}(t) &\xrightarrow[L^2]{\text{weakly}} \hat{\hat{A}}(t)^T \hat{q}(t)\\ \hat{A}_{\gamma_k}(t)^T \hat{q}_{\gamma_k}(t) &\xrightarrow[L^2]{\text{weakly}} \hat{A}(t)^T \hat{q}(t)\\ \hat{\hat{\mathcal{A}}}_{\gamma_k}(t)^T \hat{v}_{\gamma_k}(t) &\xrightarrow[L^2]{\text{weakly}} \hat{\hat{\mathcal{A}}}(t)^T \hat{v}(t)\\ \hat{\mathcal{A}}_{\gamma_k}(t)^T \hat{v}_{\gamma_k}(t) &\xrightarrow[L^2]{\text{weakly}} \hat{\mathcal{A}}(t)^T \hat{v}(t).\end{aligned} \tag{4.65}$$

Then, for any $z(\cdot) \in \mathcal{C}([0,1],\mathbb{R}^n)$, we have

$$\begin{aligned}&\lim_{k\to\infty}\int_0^T \langle z(t), \mathcal{Q}_{\gamma_k}(t)\rangle dt\\ &= \int_0^T \left\langle z(t), -\Big[(1-\beta)\hat{\hat{A}}^T(t) + \beta\hat{A}^T(t)\Big]\hat{q}(t) - \Big[(1-\beta)\hat{\hat{\mathcal{A}}}(t)^T + \beta\hat{\mathcal{A}}^T(t)\Big]\,\hat{v}(t)\right\rangle.\end{aligned}$$

Now, for each i, the sequence of positive and continuous functions $\hat{\xi}^i_{\gamma_k}$ produces a sequence of bounded linear functionals in $C^\oplus(0;T)$ to which it corresponds a sequence of finite positive

Radon measure $\hat{\mu}^i_{\gamma_k} \in \mathfrak{M}_+([0,T])$ such that for all $B \in \mathfrak{B}([0,T])$ and for all $z \in \mathcal{C}([0,T],\mathbb{R})$, we have

$$\hat{\mu}^i_{\gamma_k}(B) = \int_B \hat{\xi}^i_{\gamma_k}(t)\, dt, \qquad \int_{[0,T]} z d\hat{\mu}^i_{\gamma_k} = \int_0^T z(t)\hat{\xi}^i_{\gamma_k}(t)\, dt.$$

Using the fact that $\hat{\xi}^i_{\gamma_k}$ uniformly bounded in L^∞ and converges weakly* in L^∞ to $\hat{\xi}^i$, we conclude from the second equation of (4.66) that $\hat{\mu}^i_{\gamma_k}$ converges weakly* to $\hat{\mu}^i_o$, the element in $\mathfrak{M}_+([0,T])$ corresponding to $\hat{\xi}^i$. Now, using the fact that $\|\hat{\vartheta}^i_{\gamma_k}\|_\infty \leq L_{\hat{\psi}}$ (for $i = 1, \cdots, r$), we apply Theorem 2.4.14 and we follow the same arguments as those used in Step 3 of the proof of Theorem 5.1 in [70] to deduce that there exist $(\hat{\vartheta}^i(\cdot))_{i=1}^r$ such that $\hat{\vartheta}^i(t) \in \partial^{xx}_l \hat{\psi}_i(t, \bar{x}(t))$ a.e. $t \in [0,1]$ and for any $z(\cdot) \in \mathcal{C}([0,1];\mathbb{R}^n)$, we have

$$\lim_{k\to\infty} \int_0^T \sum_{i=1}^r \gamma_k e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle z(t), \hat{\vartheta}^i_{\gamma_k}(t)\hat{q}_{\gamma_k}(t)\rangle dt = \int_0^T \sum_{i=1}^r \hat{\xi}^i(t)\langle z(t), \hat{\vartheta}^i(t)\hat{q}(t)\rangle dt. \tag{4.66}$$

Using (4.24) in which we have $\gamma_k \xi^{r+1}_{\gamma_k} \longrightarrow 0$ *uniformly*, we deduce that, for all $z(\cdot) \in \mathcal{C}([0,T];\mathbb{R}^n)$, we have

$$\lim_{k\to\infty} \int_0^T \langle z(t), \gamma_k e^{\gamma_k \psi_{r+1}(t, x_{\gamma_k}(t))} \hat{q}_{\gamma_k}(t)\rangle dt = 0, \tag{4.67}$$

$$\lim_{k\to\infty} \int_0^T \gamma_k^2 e^{\gamma_k \psi_{r+1}(t, x_{\gamma_k}(t))} \langle \nabla_x \psi_{r+1}(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle \langle z(t), \nabla_x \psi_{r+1}(t, x_{\gamma_k}(t))\rangle dt = 0. \tag{4.68}$$

This means that for any $z(\cdot) \in \mathcal{C}([0,1],\mathbb{R}^n)$, we have

$$\lim_{k\to\infty} \int_0^T \langle z(t), \mathcal{X}_{\gamma_k}(t)\rangle dt = \int_0^T \sum_{i=1}^r \hat{\xi}^i(t)\langle z(t), \hat{\vartheta}^i(t)\hat{q}(t)\rangle dt,$$
$$\lim_{k\to\infty} \int_0^T \langle z(t), \mathcal{Z}_{\gamma_k}(t)\rangle dt = 0.$$

We now work on the last term of our limit taking process:

$$\lim_{k\to\infty} \int_0^T \sum_{i=1}^r \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle z(t), \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t))\rangle \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle dt.$$

Let $\hat{\nu}^i_{\gamma_k}$ the finite signed Radon measure on $[0,T]$, corresponding to the bounded linear functional on $\mathcal{C}([0,1];\mathbb{R})$ defined by $\gamma_k \hat{\xi}^i_{\gamma_k}(t)\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle, t \in [0,1]$, i.e.

$$d\hat{\nu}^i_{\gamma_k}(t) := \gamma_k \hat{\xi}^i_{\gamma_k}(t)\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle dt,\ i = 1, \cdots, r. \tag{4.69}$$

This means that, for all $z \in \mathcal{C}([0,T];\mathbb{R})$, we have

$$\langle \hat{\nu}^i_{\gamma_k}, z\rangle = \int_{[0,T]} z \, d\hat{\nu}^i_{\gamma_k} = \int_0^T z(t)\gamma_k \hat{\xi}^i_{\gamma_k}(t)\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle dt.$$

Using steps similar to step above, we can prove that there exists a constant $M_7 > 0$ such that for k large enough

$$\int_0^T \gamma_k \hat{\xi}^i_{\gamma_k}(t)\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle dt \leq M_7. \tag{4.70}$$

Thus,

$$\|\hat{\nu}^i_{\gamma_k}\|_{T.V.} \leq M_7.$$

Hence, along a subsequence (we do not relabel), the sequence $(\hat{\nu}^i_{\gamma_k})_k$ converges weakly* to a finite signed Radon measure

$\boxed{\hat{\nu}^i \text{ supported in } \{t \in [0,T] : \hat{\psi}_i(t, \bar{x}(t)) = 0\} = \{t \in [0,T] : \psi_i(t, \bar{x}(t)) = 0\} = I^0_i(\bar{x})}$ and

$$\|\hat{\nu}^i\|_{T.V.} \leq M_7.$$

Using Theorem 2.4.14, and the fact that $\nabla_x \hat{\psi}_i(t, x_{\gamma_k})$ is uniformly bounded and converges uniformly to $\nabla_x \hat{\psi}_i(t, \bar{x})$, we deduce that $\nabla_x \hat{\psi}_i(t, x_{\gamma_k})\hat{\nu}^i_{\gamma_k}$ converges weakly* to $\nabla_x \hat{\psi}_i(t, \bar{x})\hat{\nu}^i$, which means that for all $z(\cdot) \in \mathcal{C}([0,1], \mathbb{R}^n)$, we have

$$\begin{aligned}
&\lim_{k\to\infty} \int_0^T \langle z(t), \mathcal{Y}_{\gamma_k}(t)\rangle dt \\
= \;& \lim_{k\to\infty} \int_0^T \sum_{i=1}^r \gamma_k^2 e^{\gamma_k \hat{\psi}_i(t, x_{\gamma_k}(t))} \langle z(t), \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t))\rangle \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle dt \\
= \;& \int_0^T \sum_{i=1}^r \langle z(t), \nabla_x \hat{\psi}_i(t, \bar{x}(t))\rangle d\hat{\nu}^i(t).
\end{aligned}$$

Hence,

$$\boxed{\begin{aligned}
&\int_{[0,T]} \langle z(t), d\hat{q}(t)\rangle \\
&= \int_0^T \left\langle z(t), -\Big[(1-\beta)\hat{\bar{A}}^T(t) + \beta \hat{A}^T(t)\Big]\hat{q}(t) - \Big[(1-\beta)\hat{\bar{\mathcal{A}}}(t)^T + \beta\hat{\mathcal{A}}^T(t)\Big]\,\hat{v}(t)\right\rangle \\
&+ \int_0^T \sum_{i=1}^r \hat{\xi}^i(t)\langle z(t), \hat{\vartheta}^i(t)\hat{q}(t)\rangle dt + \int_0^T \sum_{i=1}^r \langle z(t), \nabla_x \hat{\psi}_i(t, \bar{x}(t))\rangle d\hat{\nu}^i(t).
\end{aligned}} \tag{4.71}$$

Now, notice from (4.64), that for any $z(\cdot) \in \mathcal{C}([0,T];\mathbb{R}^l)$, we have

$$\begin{aligned}
\int_{[0,T]} \langle z(t), \dot{\hat{v}}(t)\rangle dt &= \lim_{k\to\infty} \int_0^T \langle z(t), \dot{\hat{v}}_{\gamma_k}(t)\rangle dt \\
&= \lim_{k\to\infty} \int_0^T \langle z(t), -\left[(1-\beta)\hat{\hat{E}}_{\gamma_k}(t)^T + \beta \hat{E}_{\gamma_k}(t)^T\right] \hat{q}_{\gamma_k}(t)\rangle dt \\
&+ \lim_{k\to\infty} \int_0^T \langle z(t), -\left[(1-\beta)\hat{\hat{\mathcal{E}}}_{\gamma_k}(t)^T + \beta \hat{\mathcal{E}}_{\gamma_k}(t)^T\right] \hat{v}_{\gamma_k}(t)\rangle dt \\
&+ \lim_{k\to\infty} \int_0^T \gamma_k e^{\gamma_k \varphi(t, y_{\gamma_k}(t))} \langle z(t), \hat{v}_{\gamma_k}(t)\rangle dt \\
&+ \lim_{k\to\infty} \int_0^T \gamma_k^2 e^{\gamma_k \varphi(t, y_{\gamma_k}(t))} \langle \nabla_y \varphi(t, y_{\gamma_k}(t)), \hat{v}_{\gamma_k}(t)\rangle \langle z(t), \nabla_y \varphi(t, y_{\gamma_k}(t))\rangle dt.
\end{aligned}$$

Using (4.24), we have $\gamma_k \zeta_{\gamma_k} \longrightarrow 0$ *uniformly.* Hence, for all $z(\cdot) \in \mathcal{C}([0,T];\mathbb{R}^l)$, we have

$$\lim_{k\to\infty} \int_0^T \langle z(t), \gamma_k e^{\gamma_k \varphi(t, y_{\gamma_k}(t))} \hat{v}_{\gamma_k}(t)\rangle dt = 0, \tag{4.72}$$

$$\lim_{k\to\infty} \int_0^T \gamma_k^2 e^{\gamma_k \varphi(t, y_{\gamma_k}(t))} \langle \nabla_y \varphi(t, y_{\gamma_k}(t)), \hat{v}_{\gamma_k}(t)\rangle \langle z(t), \nabla_y \varphi(t, y_{\gamma_k}(t))\rangle dt = 0. \tag{4.73}$$

We also know that $\hat{q}_{\gamma_k}$ and $\hat{v}_{\gamma_k}$ are uniformly bounded in L^∞ and converge pointwise to $\hat{q}(\cdot)$ and $\hat{v}(\cdot)$ respectively. We then conclude that

$$\begin{aligned}
\hat{\hat{E}}_{\gamma_k}(t)^T \hat{q}_{\gamma_k}(t) &\xrightarrow[L^2]{\text{weakly}} \hat{\hat{E}}(t)^T \hat{q}(t) \\
\hat{E}_{\gamma_k}(t)^T \hat{q}_{\gamma_k}(t) &\xrightarrow[L^2]{\text{weakly}} \hat{E}(t)^T \hat{q}(t) \\
\hat{\hat{\mathcal{E}}}_{\gamma_k}(t)^T \hat{v}_{\gamma_k}(t) &\xrightarrow[L^2]{\text{weakly}} \hat{\hat{\mathcal{E}}}(t)^T \hat{v}(t) \\
\hat{\mathcal{E}}_{\gamma_k}(t)^T \hat{v}_{\gamma_k}(t) &\xrightarrow[L^2]{\text{weakly}} \hat{\mathcal{E}}(t)^T \hat{v}(t)
\end{aligned}$$

Hence, we have that

$$\boxed{\begin{aligned}
\int_{[0,T]} \langle z(t), \dot{\hat{v}}(t)\rangle dt = &\int_0^T \langle z(t), -\left[(1-\beta)\hat{\hat{E}}(t)^T + \beta \hat{E}(t)^T\right] \hat{q}(t)\rangle dt \\
+ &\int_0^T \langle z(t), -\left[(1-\beta)\hat{\hat{\mathcal{E}}}(t)^T + \beta \hat{\mathcal{E}}(t)^T\right] \hat{v}(t)\rangle dt.
\end{aligned}} \tag{4.74}$$

Step I.3.3 Formulating non-triviality condition, maximization condition, complementary slackness, measure properties, and transversality condition for fixed (α, β).

For condition (vi), equation (4.69) yields the following

$$\left\langle \hat{q}_{\gamma_k}(t), \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t))\right\rangle d\hat{\nu}^i_{\gamma_k}(t) = \gamma_k \hat{\xi}^i_{\gamma_k}(t) \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t)\rangle^2 \geq 0,$$

and hence, upon taking the limit, we get

$$\boxed{\left\langle \hat{q}(t), \nabla_x \hat{\psi}_i(t, \bar{x}(t)) \right\rangle d\hat{\nu}^i(t) \geq 0.}$$

For condition (ii), since $\hat{\lambda}_{\gamma_k} \in [0,1]$ then, along a subsequence, $\hat{\lambda}_{\gamma_k}$ converges pointwise to a limit $\hat{\lambda} \in [0,1]$. Taking the limit of (4.36), we deduce that

$$\boxed{\hat{\lambda} + \|\hat{p}(T)\| = 1.}$$

For condition (iv), we know by (4.32) that for $t \in [0,T]$, $u \in U(t)$,

$$\begin{aligned} & \langle \hat{q}_{\gamma_k}(t), f(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u) \rangle + \langle \hat{v}_{\gamma_k}(t), g(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u) \rangle - \frac{\hat{\lambda}_{\gamma_k} \alpha}{\beta} \|u - u_{\gamma_k}(t)\| \\ \leq \; & \langle \hat{q}_{\gamma_k}(t), f(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u_{\gamma_k}(t)) \rangle + \langle \hat{v}_{\gamma_k}(t), g(t, x_{\gamma_k}(t), y_{\gamma_k}(t), u_{\gamma_k}(t)) \rangle \text{ a.e. } t \in [0,1]. \end{aligned}$$

Taking the limit when $k \to \infty$ of this last inequality, we conclude that for $t \in [0,T]$, $u \in U(t)$,

$$\begin{aligned} & \langle \hat{q}(t), f(t, \bar{x}(t), \bar{y}(t), u) \rangle + \langle \hat{v}(t), g(t, \bar{x}(t), \bar{y}(t), u) \rangle - \frac{\hat{\lambda} \alpha}{\beta} \|u - \bar{u}(t)\| \\ \leq \; & \langle \hat{q}(t), f(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)) \rangle + \langle \hat{v}(t), g(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)) \rangle \text{ a.e. } t \in [0,1]. \end{aligned}$$

This is equivalent to saying that

$$\boxed{\begin{array}{l} \max\limits_{u \in U(t)} \left\{ \langle \hat{q}(t), f(t, \bar{x}(t), \bar{y}(t), u) \rangle + \langle \hat{v}(t), g(t, \bar{x}(t), \bar{y}(t), u) \rangle - \dfrac{\hat{\lambda} \alpha}{\beta} \|u - \bar{u}(t)\| \right\} \\ \text{is attained at } u = \bar{u}(t) \text{ for a.e. } t \in [0,T]. \end{array}}$$

For condition (v), we have $\hat{\xi}^i \geq 0$ $(i = 1, \cdots, r)$, and $\hat{\xi}^i(t) = 0 \;\; \forall t \in I_i^-(\bar{x})$. We also have using equation (4.70) that

$$\int_0^T \hat{\xi}^i_{\gamma_k}(t) |\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t) \rangle| dt \leq \frac{1}{\gamma_k} \int_0^T \gamma_k \hat{\xi}^i_{\gamma_k}(t) \langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t) \rangle dt \leq \frac{M_7}{\gamma_k}. \tag{4.75}$$

Hence,

$$\lim_{k \to \infty} \int_0^T \hat{\xi}^i_{\gamma_k}(t) |\langle \nabla_x \hat{\psi}_i(t, x_{\gamma_k}(t)), \hat{q}_{\gamma_k}(t) \rangle| = 0.$$

And thus,

$$\int_0^T \hat{\xi}^i(t) |\langle \nabla_x \hat{\psi}_i(t, \bar{x}(t)), \hat{q}(t) \rangle| = 0.$$

We conclude that

$$\boxed{\hat{\xi}^i(t)\langle\nabla_x\hat{\psi}_i(t,\bar{x}(t)),\hat{q}(t)\rangle = 0 \text{ a.e. } t\in[0,1].}$$

Finally, for condition (vii), by equation (4.31), we have that

$$(\hat{p}_{\gamma_k}(0),-\hat{p}_{\gamma_k}(T)) \in \hat{\lambda}_{\gamma_k}\ \partial_l^L J(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T)) + \alpha\bar{B} + N^L_{S^{\gamma_k}(k)}\left(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T)\right). \tag{4.76}$$

This is equivalent to saying that there exist
$(z^1_{\gamma_k},z^2_{\gamma_k},s^1_{\gamma_k},s^2_{\gamma_k}) \in \partial_l^L J(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T))$,
$(w^1_{\gamma_k},w^2_{\gamma_k},m^1_{\gamma_k},m^2_{\gamma_k}) \in N^L_{S^{\gamma_k}(k)}(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T))$, $o_{\gamma_k}\in\bar{B}$ such that

$$(\hat{p}_{\gamma_k}(0),-\hat{p}_{\gamma_k}(T)) = \hat{\lambda}_{\gamma_k}(z^1_{\gamma_k},z^2_{\gamma_k},s^1_{\gamma_k},s^2_{\gamma_k}) + \alpha o_{\gamma_k} + (w^1_{\gamma_k},w^2_{\gamma_k},m^1_{\gamma_k},m^2_{\gamma_k}). \tag{4.77}$$

- As we have seen before, since $\hat{\lambda}_{\gamma_k}\in[0,1]$, then, along a subsequence, $\hat{\lambda}_{\gamma_k}$ converges pointwise to a limit $\hat{\lambda}\in[0,1]$. We also have $\|(z^1_{\gamma_k},z^2_{\gamma_k},s^1_{\gamma_k},s^2_{\gamma_k})\|\le L_g$, then, along a subsequence,

  $$(z^1_{\gamma_k},z^2_{\gamma_k},s^1_{\gamma_k},s^2_{\gamma_k}) \to (z^1,z^2,s^1,s^2).$$

  Since $\partial_l^L J(\cdot,\cdot,\cdot,\cdot)$ has closed graph with nonempty and compact values then, using the fact that $(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T)) \to (\bar{x}(0),\bar{y}(0),\bar{x}(T),\bar{y}(T))$, we get

  $$(z^1,z^2,s^1,s^2)\in\partial_l^L J(\bar{x}(0),\bar{y}(0),\bar{x}(T),\bar{y}(T)).$$

- Since $\|o_{\gamma_k}\|\le 1$, then, along a subsequence, we have that $o_{\gamma_k}\to o\in\bar{B}$.
- We also have $(\hat{p}_{\gamma_k}(0),-\hat{p}_{\gamma_k}(T)) \to (\hat{p}(0),-\hat{p}(T))$.
- We deduce from (4.77) that $(w^1_{\gamma_k},w^2_{\gamma_k},m^1_{\gamma_k},m^2_{\gamma_k})$ must converge to (w^1,w^2,m^1,m^2) respectively.

We now show that $(w^1,w^2,m^1,m^2)\in N^L_S(\bar{x}(0),\bar{y}(0),\bar{x}(T),\bar{y}(T))$. Indeed,

$$(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T)) \in S^{\gamma_k}(k),$$

$$(x_{\gamma_k}(0),y_{\gamma_k}(0),x_{\gamma_k}(T),y_{\gamma_k}(T)) \in (S_{\delta_o}+\rho_1 B)\cap \operatorname{int}\left(\bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(0)\times\bar{\mathcal{N}}_{(\bar{\varepsilon},\bar{\delta})}(T)\right) \subset \operatorname{int} S(\bar{\delta}).$$

We now have two cases:

Case 1: $\bar{x}(0) \in \operatorname{int} C(0)$.

Since $(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T) - \bar{e}_{\gamma_k}, y_{\gamma_k}(T) - \bar{\omega}_{\gamma_k}) \in \operatorname{int} \bar{\mathscr{B}}_{\delta_o}$, then

$$N^L_{S^{\gamma_k}(k)}(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T), y_{\gamma_k}(T)) = N^L_S(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T) - \bar{e}_{\gamma_k}, y_{\gamma_k}(T) - \bar{\omega}_{\gamma_k}).$$

Case 2: $\bar{x}(0) \in \operatorname{bdry} C(0)$.

Since $(x_{\gamma_k}(0) - \bar{\sigma}_k \frac{d_{\bar{x}(0)}}{\|d_{\bar{x}(0)}\|}, y_{\gamma_k}(0), x_{\gamma_k}(T) - \bar{e}_{\gamma_k}, y_{\gamma_k}(T) - \bar{\omega}_{\gamma_k}) \in \operatorname{int} \bar{\mathscr{B}}_{\delta_o}$, then

$$N^L_{S^{\gamma_k}(k)}(x_{\gamma_k}(0), y_{\gamma_k}(0), x_{\gamma_k}(T), y_{\gamma_k}(T)) = N^L_S(x_{\gamma_k}(0) - \bar{\sigma}_k \frac{d_{\bar{x}(0)}}{\|d_{\bar{x}(0)}\|}, y_{\gamma_k}(0), x_{\gamma_k}(T) - \bar{e}_{\gamma_k}, y_{\gamma_k}(T) - \bar{\omega}_{\gamma_k}).$$

In both cases, since $(w^1_{\gamma_k}, w^2_{\gamma_k}, m^1_{\gamma_k}, m^2_{\gamma_k}) \to (w^1, w^2, m^1, m^2)$, and $N^L_S(\cdot)$ has closed values and closed graph, then

$$(w^1, w^2, m^1, m^2) \in N^L_S(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)).$$

Consequently, the limit of (4.76) is

$$\boxed{(\hat{p}(0), -\hat{p}(T)) \in \hat{\lambda}\, \partial^L_l J(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)) + \alpha \bar{B} + N^L_S(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)).}$$

Step I.3.4 Formulating the necessary conditions for each fixed (α, β) in terms of ψ_i.

Notice that $\hat{\nu}^i$ and $\hat{\xi}^i$ are supported in $\{t \in [0, T] : \hat{\psi}_i(t, \bar{x}(t)) = 0\} = \{t \in [0, T] : \psi_i(t, \bar{x}(t)) = 0\} = I^0_i(\bar{x})$, and on this set, $\nabla_x \hat{\psi}_i(t, \bar{x}(t)) = \nabla_x \psi_i(t, \bar{x}(t))$. Hence, all the previous necessary conditions can be formulated in terms of ψ_i by simply taking $q := \hat{q}$, $v := \hat{v}$, $p := \hat{p}$, $\lambda := \hat{\lambda}$, $\bar{A} := \hat{\bar{A}}$, $A := \hat{A}$, $\bar{\mathcal{A}} := \hat{\bar{\mathcal{A}}}$, $\mathcal{A} := \hat{\mathcal{A}}$, $\bar{E} := \hat{\bar{E}}$, $\bar{\mathcal{E}} := \hat{\bar{\mathcal{E}}}$, $E := \hat{E}$, $\mathcal{E} := \hat{\mathcal{E}}$, $\xi^i := \hat{\xi}^i$ (for $i = 1, \cdots, r$), $\vartheta^i := \hat{\vartheta}^i$ (for $i = 1, \cdots, r$) and $\nu^i := \hat{\nu}^i$ (for $i = 1, \cdots, r$).

Step I.4. Taking $\alpha \to 0$.

All the boxed equations above depend on α and β. As the first step, we take the limit $\alpha \to 0$, while keeping β fixed. To explicitly indicate the dependence on α in our notation, we introduce a subscript α_j, where $\alpha_j \in (0, 1]$ and $\alpha_j \to 0$.

First, for each j, $(\xi^1_{\alpha_j}, \cdots, \xi^r_{\alpha_j}) \in L^\infty([0, T], \mathbb{R}^r_+)$ such that

$$\xi^i_{\alpha_j} = 0 \text{ on } I^-_i(\bar{x}) \ (\forall i = 1, \cdots, r), \ \|\sum_{i=1}^r \xi^i_{\alpha_j}\|_\infty \le \frac{2\bar{\mu}}{\bar{\eta}^2}, \tag{4.78}$$

and

$$\begin{cases} \dot{\bar{x}}(t) = f(t,\bar{x}(t),\bar{y}(t),\bar{u}(t)) - \sum_{i=1}^{r} \xi^i_{\alpha_j}(t)\nabla_x\psi_i(t,\bar{x}(t)) \text{ a.e. } t\in[0,T], \\ \dot{\bar{y}}(t) = g(t,\bar{x}(t),\bar{y}(t),\bar{u}(t)) \text{ a.e. } t\in[0,T], \\ \psi_i(t,\bar{x}(t)) \le 0,\ \forall t\in[0,T], \quad \forall i\in\{1,\cdots,r\}. \end{cases} \tag{4.79}$$

Thus, for each $i = 1,\cdots,r$, there exists a subsequence of $\xi^i_{\alpha_j}$ (we do not relabel) that converges weakly* (and hence weakly in L^2) to a non-negative function $\xi^i \in L^\infty([0,T],\mathbb{R})$, with $\xi^i = 0$ on $I^-_i(\bar{x})$. Moreover, using the fact that for each $i\in\{1,\cdots,r\}$,

$$\int_0^T \xi^i_{\alpha_j}(t)\nabla_x\psi_i(t,\bar{x}(t)) \to \int_0^T \xi^i(t)\nabla_x\psi_i(t,\bar{x}(t)),$$

we deduce that condition (i) of our theorem is satisfied (with no dependency on α).

We now show the dependency on α_j in the adjoint equation. For each j, we have

$$\begin{aligned} &\int_{[0,T]} \langle z(t), dq_{\alpha_j}(t)\rangle \\ = \ & \int_0^T \Big\langle z(t), -\Big[(1-\beta)\bar{A}^T_{\alpha_j}(t) + \beta A^T_{\alpha_j}(t)\Big] q_{\alpha_j}(t) - \Big[(1-\beta)\bar{\mathcal{A}}_{\alpha_j}(t)^T + \beta\mathcal{A}^T_{\alpha_j}(t)\Big] v_{\alpha_j}(t)\Big\rangle \\ + \ & \int_0^T \sum_{i=1}^r \xi^i_{\alpha_j}(t)\langle z(t), \vartheta^i_{\alpha_j}(t) q_{\alpha_j}(t)\rangle dt + \int_0^T \sum_{i=1}^r \langle z(t), \nabla_x\psi_i(t,\bar{x}(t))\rangle d\nu^i_{\alpha_j}(t), \end{aligned}$$

$$\begin{aligned} \int_{[0,T]} \langle z(t), \dot{v}_{\alpha_j}(t)\rangle dt &= \int_0^T \langle z(t), -\Big[(1-\beta)\bar{E}_{\alpha_j}(t)^T + \beta E_{\alpha_j}(t)^T\Big] q_{\alpha_j}(t)\rangle dt \\ &+ \int_0^T \langle z(t), -\Big[(1-\beta)\bar{\mathcal{E}}_{\alpha_j}(t)^T + \beta\mathcal{E}_{\alpha_j}(t)^T\Big] v_{\alpha_j}(t)\rangle dt. \end{aligned}$$

Using the results of Steps I.3.1 and I.3.2 with the subscript γ_k being replaced by the subscript α_j, we deduce that there exist a function $q(\cdot)$ of bounded variation, an absolutely continuous function $v(\cdot)$, such that the previous two equations are satisfied with no α_j-dependency.

By step I.3.3, we deduce that $\lambda_{\alpha_j} + \|p_{\alpha_j}(T)\| = 1$. Then, along a subsequence, λ_{α_j} converges to $\lambda\in[0,1]$ and $\lambda + \|p(T)\| = 1$. We also have that

$$\max_{u\in U(t)} \left\{ \langle q_{\alpha_j}(t), f(t,\bar{x}(t),\bar{y}(t),u)\rangle + \langle v_{\alpha_j}(t), g(t,\bar{x}(t),\bar{y}(t),u)\rangle - \frac{\lambda_{\alpha_j}\alpha_j}{\beta}\|u-\bar{u}(t)\| \right\} \tag{4.80}$$

is attained at $u = \bar{u}(t)$ for a.e. $t \in [0,1]$. Hence, taking the limit when $\alpha_j \to 0$, we deduce that

$$\max_{u \in U(t)} \left\{ \langle q(t), f(t, \bar{x}(t), \bar{y}(t), u) \rangle + \langle v(t), g(t, \bar{x}(t), \bar{y}(t), u) \rangle \right\} \tag{4.81}$$

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

As, for each $i = 1, \cdots, r$, $\xi^i_{\alpha_j}$ converges weakly* in L^∞ to ξ^i and $\xi^i_{\alpha_j}(t)\langle \nabla_x \psi_i(t, \bar{x}(t)), q_{\alpha_j}(t)\rangle = 0$ a.e. $t \in [0,T]$, then

$$0 = \lim_{j \to \infty} \int_0^T \xi^i_{\alpha_j}(t) |\langle \nabla_x \psi_i(t, \bar{x}(t)), q_{\alpha_j}(t)\rangle| dt = \int_0^T \xi^i(t) |\langle \nabla_x \psi_i(t, \bar{x}(t)), q(t)\rangle| dt.$$

Hence,

$$\xi^i(t)\langle \nabla_x \psi_i(t, \bar{x}(t)), q(t)\rangle = 0 \text{ a.e. } t \in [0,T].$$

Finally, for the transversality condition, we have

$$(p_{\alpha_j}(0), -p_{\alpha_j}(T)) \in \lambda_{\alpha_j}\, \partial^L_l J(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)) + \alpha_j \bar{B} + N^L_S\left(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)\right).$$

Then, using similar steps used to derive the transversality condition for fixed (α, β) in Step I.3.3, we deduce that

$$(p(0), -p(T)) \in \lambda\, \partial^L_l J(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)) + N^L_S\left(\bar{x}(0), \bar{y}(0), \bar{x}(T), \bar{y}(T)\right).$$

Step I.5. Taking $\beta \to 0$.

In this step, we explicitly indicate the dependence on β in our notation, and we introduce a subscript β_j, where $\beta_j \in (0,1]$ and $\beta_j \to 0$. Deriving all the conditions except for the adjoint equation follows a similar process to Step I.4, replacing the subscript α_j by the subscript β_j. Below, we present our derivation for the adjoint equation. For each j, we have

$$\begin{aligned} & \int_{[0,T]} \langle z(t), dq_{\beta_j}(t)\rangle \\ = \quad & \int_0^T \Big\langle z(t), -\Big[(1-\beta_j)\bar{A}^T_{\beta_j}(t) + \beta_j A^T_{\beta_j}(t)\Big] q_{\beta_j}(t) - \Big[(1-\beta_j)\bar{\mathcal{A}}_{\beta_j}(t)^T + \beta_j \mathcal{A}^T_{\beta_j}(t)\Big] v_{\beta_j}(t) \Big\rangle \\ + \quad & \int_0^T \sum_{i=1}^r \xi^i_{\beta_j}(t) \langle z(t), \vartheta^i_{\beta_j}(t) q_{\beta_j}(t)\rangle dt + \int_0^T \sum_{i=1}^r \langle z(t), \nabla_x \psi_i(t, \bar{x}(t))\rangle d\nu^i_{\beta_j}(t), \end{aligned} \tag{4.82}$$

$$\begin{aligned}\int_{[0,T]} \langle z(t), \dot{v}_{\beta_j}(t)\rangle dt &= \int_0^T \langle z(t), -\left[(1-\beta_j)\bar{E}_{\beta_j}(t)^T + \beta E_{\beta_j}(t)^T\right] q_{\beta_j}(t)\rangle dt \\ &+ \int_0^T \langle z(t), -\left[(1-\beta_j)\bar{\mathcal{E}}_{\beta_j}(t)^T + \beta_j \mathcal{E}_{\beta_j}(t)^T\right] v_{\beta_j}(t)\rangle dt, \quad (4.83)\end{aligned}$$

where

$$\max\left\{\|(\bar{A}_{\beta_j}, \bar{E}_{\beta_j})\|_2,\ \|(A_{\beta_j}, E_{\beta_j})\|_2,\ \|(\bar{\mathcal{A}}_{\beta_j},\ \bar{\mathcal{E}}_{\beta_j})\|_2,\ \|(\mathcal{A}_{\beta_j}, \mathcal{E}_{\beta_j})\|_2\right\} \leq \|L_h\|_2,$$

$$(\bar{A}_{\beta_j}, \bar{E}_{\beta_j})(t) \in \partial_\ell^{(x,y)} f(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)),$$

$$(\bar{\mathcal{A}}_{\beta_j}, \bar{\mathcal{E}}_{\beta_j})(t) \in \partial_\ell^{(x,y)} g(t, \bar{x}(t), \bar{y}(t), \bar{u}(t)),$$

$$\|q_{\beta_j}\|_\infty \leq M_p \quad \text{and} \quad V_0^1(q_{\beta_j}) = \|\dot{q}_{\beta_j}\|_1 \leq M_q,$$

$$\|v_{\beta_j}\|_\infty \leq M_p \quad \text{and} \quad \|\dot{v}_{\beta_j}\|_2 \leq M_v,$$

$$\|\nu^i_{\beta_j}\|_{T.V.} \leq M_7, \quad \text{for } i = 1, \cdots, r,$$

$$\|\vartheta^i_{\beta_j}\|_\infty \leq L_\psi, \quad \text{for } i = 1, \cdots, r.$$

Taking $\beta_j \to 0$, and following steps similar to Steps I.3.1 and I.3.2, we obtain the adjoint equations (condition (iii) of our theorem).

Step II. We have concluded proving the theorem under the temporary assumptions (A4.2) and (A3.3)$'$. **The goal of Step II is to remove those two temporary assumptions.**

Step II.1. Removing assumption (A3.3)$'$.

In this step, we remove (A3.3)$'$, and we simply assume that (A3.3) is satisfied for some $\bar{\beta}(\cdot)$ positive. We use arguments similar to those at the last step of the proof of [58, Theorem 3.1], as well as Remark 4.2.1(ii)-(iii). We first define $\tilde{\psi}_i(t,x) := \bar{\beta}_i(t)\psi_i(t,x)$. Notice that

- $C(t)$ is also the zero-sublevel sets of $(\tilde{\psi}_i(t,\cdot))_{i=1}^r$, for $i = 1, \cdots, r$.
- For some $L_{\tilde{\psi}} > 0$, $\tilde{\psi}_i$ satisfies $(A3.1)$ for all $i = 1, \cdots, r$.
- Condition $(A3.3)$ is equivalent to saying that for $t \in I^0(\bar{x})$, the Gramian matrix $\mathcal{G}_{\tilde{\psi}}(t)$ of the vectors $\{\nabla_x \tilde{\psi}_i(t, \bar{x}(t)) : i \in \mathcal{I}^0_{(t,\bar{x}(t))}\}$ is strictly diagonally dominant.
- $\tilde{\psi}_1, \cdots, \tilde{\psi}_r$ satisfy $(A3.2)$, and hence, (3.34) of Lemma 3.2.4 is valid at $\tilde{\psi}_1, \cdots, \tilde{\psi}_r, \psi_{r+1}$ when replacing $\bar{\eta}$ by $\tilde{\eta} := \bar{\eta}\ \mathbf{b}_{\bar{\beta}}$, where

$$\mathbf{b}_{\bar{\beta}} := \min\left\{1, \min\{\bar{\beta}_i(t) : t \in [0,T],\ i = 1, \cdots, r\}\right\}.$$

We denote by $(\tilde{P})$ the version of (P) in which the functions ψ_i are replaced by $\tilde{\psi}_i$. Note that (P) and $(\tilde{P})$ coincide, and $((\bar{x},\bar{y}),\bar{u})$ is a strong local minimizer for $(\tilde{P})$. Furthermore, the data of $(\tilde{P})$ satisfy the assumptions required for the proven maximum principle (established in Step II.1). Therefore, we apply the proven version of the maximum principle to $(\tilde{P})$, and we get the existence of an adjoint vector $\tilde{p}=(\tilde{q},\tilde{v})$ with $\tilde{q}\in BV([0,T];\mathbb{R}^n)$ and $\tilde{v}\in W^{1,2}([0,T];\mathbb{R}^l)$, finite signed Radon measures $(\tilde{\nu}^i)_{i=1}^r$ on $[0,T]$, nonnegative functions $(\tilde{\xi}^i)_{i=1}^r$ in $L^\infty([0,T];\mathbb{R}^+)$, L^2-measurable functions $\tilde{\bar{A}}(\cdot)$ in $\mathscr{M}_{n\times n}([0,T])$, $\tilde{\bar{E}}(\cdot)$ in $\mathscr{M}_{n\times l}([0,T])$, $\tilde{\bar{\mathcal{A}}}(\cdot)$ in $\mathscr{M}_{l\times n}([0,T])$, and $\tilde{\bar{\mathcal{E}}}(\cdot)$ in $\mathscr{M}_{l\times l}([0,T])$, L^∞-measurable functions $(\tilde{\vartheta}^i(\cdot))_{i=1}^r$ in $\mathscr{M}_{n\times n}([0,T])$, and a scalar $\tilde{\lambda}\geq 0$ that satisfy conditions (i)-(vii). To express those conditions in terms of the original data of (P), we replace $\tilde{\psi}_i(t,x)$ by $\bar{\beta}_i(t)\psi_i(t,x)$, and we take $p:=\tilde{p}, q:=\tilde{q}, v:=\tilde{v}, \xi^i(\cdot):=\bar{\beta}_i(\cdot)\tilde{\xi}^i(\cdot), \lambda:=\tilde{\lambda}$, $\bar{A}(\cdot):=\tilde{\bar{A}}(\cdot)$, $\bar{E}(\cdot):=\tilde{\bar{E}}(\cdot)$, $\bar{\mathcal{A}}(\cdot):=\tilde{\bar{\mathcal{A}}}(\cdot)$, $\bar{\mathcal{E}}(\cdot):=\tilde{\bar{\mathcal{E}}}(\cdot), d\nu^i(\cdot)=\bar{\beta}_i(\cdot)d\tilde{\nu}^i(\cdot)$, and $\vartheta^i(\cdot):=\frac{1}{\bar{\beta}_i(\cdot)}\tilde{\vartheta}^i(\cdot)$.

Step II.2. Removing assumption (A4.2) when the sets $U(t)$ are uniformly bounded. In this step, we remove (A4.2) (so assume h does not satisfy (A4.2)), and we assume that the sets $U(t)$ are uniformly bounded. To remove (A4.2), that is, the convexity assumption of $h(t,x,y,U(t))$ for $(x,y)\in\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(t)$ and $t\in[0,T]$ a.e., we shall extend the *relaxation technique* in [70, Section 5.2], developed for *global* minimizers of Mayer optimal control problems over sweeping processes having *constant* compact sweeping sets and constant control set U, to the case of *strong local* minimizers, the sweeping sets are $\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)$, which are *time-dependent* and *not* necessarily moving in an absolutely continuous way, U is time-dependent, and *joint*-endpoints constraint $S_{\frac{\delta}{2}}$, where $\delta\in(0,\bar{\varepsilon})$ is fixed.

Step II.2.1. $(\bar{X}:=(\bar{x},\bar{y}),\bar{u})$ **is a δ-strong local minimizer for $(\bar{P}_\delta)$ with extended J.** Fix $\delta\in(0,\bar{\varepsilon})$. Using Theorem 2.4.3, there is an L_J-Lipschitz function $\bar{J}:\mathbb{R}^{n+l}\times\mathbb{R}^{n+l}\to\mathbb{R}$ that extends J to $\mathbb{R}^{2(n+l)}$ from $S(\bar{\delta})$. By Remark 4.2.6(i), $((\bar{x},\bar{y}),\bar{u})$ being a $\bar{\delta}$-strong local minimizer for (P), then it is also a δ-strong local minimum for $(\bar{P}_\delta)$ in which we use the extension $\bar{J}$ instead of J.

Step II.2.2. $(\bar{X},\bar{u})$ **is a global minimum for a problem** $(\bar{\mathscr{P}})$**.**

Performing appropriate modifications to the technique presented in the proof of [55, Theorem 6.2], we are then able to formulate the following problem $(\bar{\mathscr{P}})$ associated with $(\bar{P}_\delta)$ for which the same solution $(\bar{X},\bar{u})$ is a *global* minimum:

$$(\bar{\mathscr{P}})\begin{cases}\text{minimize}\quad \bar{J}(X(0),X(T))+\bar{K}\int_0^T \mathcal{L}(t,X(t))\,dt\\ \text{over } X:=(x,y)\in W^{1,1}([0,T],\mathbb{R}^{n+l}),\ u\in\mathcal{U},\ \text{ such that}\\ \qquad(\bar{D})\begin{cases}\dot{X}(t)\in h(t,x(t),y(t),u(t))-N_{\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)}(X(t)),\ \text{a.e. } t\in[0,T],\end{cases}\\ (X(0),X(T))\in S_{\frac{\delta}{2}}=S\cap\bar{\mathscr{B}}_{\frac{\delta}{2}},\end{cases}$$

where $\mathcal{L}:[0,T]\times\mathbb{R}^{n+l}\to\mathbb{R}$ and $\bar{K}>0$ are defined by

$$\mathcal{L}(t,X)=\mathcal{L}(t,x,y):=\max\{\|x-\bar{x}(t)\|^2-\frac{\delta^2}{4},\ \|y-\bar{y}(t)\|^2-\frac{\delta^2}{4},\ 0\}>0, \tag{4.84}$$

$$\bar{K}:=\frac{512\bar{M}_\ell M_J}{5\delta^3},\text{where }\ 2\bar{M}_\ell:=\max\{L_{(\bar{x},\bar{y})},\ M_h+\frac{\bar{\mu}}{4\bar{\eta}^2}\bar{L}\},\quad M_J:=\max_{S(\bar{\delta})}|J(X_1,X_2)|, \tag{4.85}$$

and hence, as $\mathcal{L}(t,\bar{X}(t))\equiv 0$, we deduce $\min(\bar{\mathscr{P}})=J(\bar{X}(0),\bar{X}(T))$.

We now show that $(\bar{X},\bar{u})$ is a $g$lobal minimum for $(\bar{\mathscr{P}})$. Indeed, let (X,u) be admissible for $(\bar{\mathscr{P}})$.

<u>Case 1:</u> $\|X-\bar{X}\|_\infty\le\delta$.

Then, (X,u) being admissible for $(\bar{P}_\delta)$, and $(\bar{X},\bar{u})$ being a δ-$s$trong local minimum for $(\bar{P}_\delta)$, yield that

$$\begin{aligned}\bar{J}(X(0),X(T))\ &+\ \bar{K}\int_0^T\mathcal{L}(t,X(t))\,dt=J(X(0),X(T))+\bar{K}\int_0^T\mathcal{L}(t,X(t))\,dt\\ &\ge\ J(X(0),X(T))\ge J(\bar{X}(0),\bar{X}(T))=\bar{J}(\bar{X}(0),\bar{X}(T))+\bar{K}\int_0^T\mathcal{L}(t,\bar{X}(t))\,dt.\end{aligned}$$

<u>Case 2:</u> $\|X-\bar{X}\|_\infty>\delta$.

Given that $(X(0),X(T))\in S_{\frac{\delta}{2}}$, there exists $\bar{t}\in[0,T]$ such that $\|X(\bar{t})-\bar{X}(\bar{t})\|=\delta$. Using that the function $t\mapsto\|X(t)-\bar{X}(t)\|$ is Lipschitz continuous with Lipschitz constant $4\bar{M}_\ell$ (see equation (3.45)), and the fact that $\|X(0)-\bar{X}(0)\|\le\frac{\delta}{2}$, we get that the Lebesgue measure of

$\Big\{t \in [0,T] : \|X(t) - \bar{X}(t)\| \geq \frac{3\delta}{4}\Big\} \geq \frac{\delta}{16\bar{M}_\ell}$. Hence,

$$\begin{aligned}\bar{J}(X(0),X(T)) &+ \bar{K}\int_0^T \mathcal{L}(t,X(t))\,dt \geq -M_J + \bar{K}\int_0^T \mathcal{L}(t,X(t))\,dt \\ &\geq -M_J + \bar{K}\frac{\delta}{16\bar{M}_\ell}\left(\left(\frac{3\delta}{4}\right)^2 - \frac{\delta^2}{4}\right) \\ &= M_J \geq J(\bar{X}(0),\bar{X}(T)) = \bar{J}(\bar{X}(0),\bar{X}(T)) + \bar{K}\int_0^T \mathcal{L}(t,\bar{X}(t))\,dt.\end{aligned}$$

This proves that $(\bar{X},\bar{u})$ is a global minimum for $(\bar{\mathscr{P}})$.

Step II.2.3. $(\bar{X},\bar{w}) := (\bar{X},((\overbrace{\bar{u},...,\bar{u}}^{n+l+1}),(\overbrace{1,0,...,0}^{n+l+1})))$ **is a global minimum for** $(\tilde{\mathscr{P}})$.

Define the problem $(\tilde{\mathscr{P}})$

$$(\tilde{\mathscr{P}})\begin{cases}\text{minimize} \quad \bar{J}(X(0),X(T)) + \bar{K}\int_0^T \mathcal{L}(t,X(t))\,dt \\ \text{over } X := (x,y) \in W^{1,1}([0,T],\mathbb{R}^{n+l}), \\ \qquad w(\cdot) := \Big((u_0(\cdot),\cdots,u_{n+l}(\cdot)),(\lambda_0(\cdot),\cdots,\lambda_{n+l}(\cdot))\Big) \in \mathscr{W} \text{ such that} \\ \qquad (\tilde{\mathscr{D}})\Big\{\dot{X}(t) \in \tilde{h}(t,X(t),w(t)) - N_{\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)}(X(t)), \text{ a.e. } t \in [0,T], \\ (X(0),X(T)) \in S_{\frac{\delta}{2}},\end{cases}$$

where

$$\tilde{h} : \mathrm{Gr}\,[\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(\cdot) \times (U(\cdot))^{n+l+1}] \times \Lambda \to \mathbb{R}^{n+l} \text{ defined as } \tilde{h}(t,X,w) := \sum_{i=0}^{n+l}\lambda_i h(t,X,u_i), \quad (4.86)$$

$$\Lambda := \Big\{(\lambda_0,\cdots,\lambda_{n+l}) \in \mathbb{R}^{n+l+1} : \lambda_i \geq 0 \text{ for } i = 0,...,n+l \text{ and } \sum_{i=0}^{n+l}\lambda_i = 1\Big\},$$

$$\mathscr{W} := \Big\{w\colon [0,T] \longrightarrow \mathbb{R}^{(m+1)(n+l+1)} \text{ measurable } : \ w(t) \in W(t) := (U(t))^{n+l+1} \times \Lambda \text{ a.e..}\Big\}$$

First, we note the following two facts that are going to be useful for our goal:

- Notice that $\tilde{h}$ satisfies (A4.1), and hence, Corollary 3.2.16 yields that for $X_0 = (x_0,y_0) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(0)$, and for $w \in \mathscr{W}$, $(\tilde{\mathscr{D}})$ has a *unique* solution $X(\cdot)$ corresponding to (X_0,w) which is $(M_h + \frac{\bar{\mu}}{4\bar{\eta}^2}\bar{L})$-Lipschitz and satisfies (3.44)-(3.46).
- Using that $0 < \delta < \bar{\varepsilon} < \bar{\delta}$ and that $\bar{X} := (\bar{x},\bar{y})$ is $L_{(\bar{x},\bar{y})}$-Lipschitz, then the function $\mathcal{L}$, defined in (4.84), is Lipschitz on Gr $\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(\cdot)$ and satisfies

$$\mathcal{L} \equiv 0 \quad \text{on} \quad \mathrm{Gr}\,\bar{\mathscr{N}}_{(\frac{\delta}{2},\frac{\delta}{2})}(\cdot), \quad \text{and} \quad |\mathcal{L}| \leq \ \bar{\delta}^2 \text{ on Gr } \bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(\cdot). \qquad (4.87)$$

Hence, by the convexity of $\tilde{h}(t,X,W(t))$ (so $\tilde{h}$ satisfy (A4)), and by Remark 4.2.5, where $\mathbb{L} := \bar{K}\,\mathcal{L}$, it follows that $\boxed{(\tilde{\mathscr{P}}) \text{ admits a } \textit{global optimal minimizer } (\tilde{X},\tilde{w}).}$

We show that $\boxed{\min(\tilde{\mathscr{P}}) = \min(\bar{\mathscr{P}}) \text{ and, } (\bar{X},\bar{w}) \text{ is optimal for } (\tilde{\mathscr{P}}).}$ Let $\mathbb{U}$ defined in (3.8), the compact set $\mathbb{V} := \text{cl}\,\mathbb{U}$, and

$$\mathscr{R} := \Big\{\sigma\colon [0,T] \to \mathfrak{M}^1_+(\mathbb{V}) \;:\; \sigma \text{ is measurable and } \sigma(t)(U(t)) = 1, \;\; t \in [0,T]\Big\}.$$

This set of relaxed controls satisfies $\mathscr{R} \subset L^1\left([0,T], \mathcal{C}(\mathbb{V};\mathbb{R})\right)^*$, which is endowed with the weak* topology. Each regular control function $u \in \mathscr{U}$ is identified with its associated Dirac relaxed control $\sigma(\cdot) = \delta_{u(\cdot)}$, and thereby $\mathscr{U} \subset \mathscr{R}$ (see e.g., [68]). Define $h_\sigma(t,X)$ and the problem $(\mathscr{P})_r$ by

$$h_\sigma(t,X) := \int_{U(t)} h(t,X,u)\sigma(t)(du), \quad \forall (t,X) \in \text{Gr}\,\bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(\cdot), \quad \sigma \in \mathscr{R},$$

$$(\mathscr{P})_r \begin{cases} \text{minimize} \quad \bar{J}(X(0),X(T)) + \bar{K}\int_0^T \mathcal{L}(t,X(t))\,dt \\ \text{over } X := (x,y) \in W^{1,1}([0,T],\mathbb{R}^{n+l}),\ \sigma \in \mathscr{R}, \text{ such that} \\ \quad (\mathscr{D})_r \begin{cases} \dot{X}(t) \in h_\sigma(t,X(t)) - N_{\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)}(X(t)), \;\; \text{a.e. } t \in [0,T], \end{cases} \\ (X(0),X(T)) \in S_{\frac{\delta}{2}}. \end{cases}$$

Since h satisfies $(A4.1)$ and $\sigma(t)(U(t)) = 1\,(\forall t \in [0,T])$, then $h_\sigma(t,X)$ is *uniformly* bounded by M_h, a Carathéodory function in (t,X), and $L_h(t)$-*Lipschitz* in X, for all t, that is, $h_\sigma(t,X)$ satisfies $(A4.1)$.

Using Corollary 3.2.16 for $X_0 = (x_0,y_0) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(0)$, $\sigma \in \mathcal{R}$, and $(f,g)(t,X,u) = h(t,X,u) := h_\sigma(t,X)$, the Cauchy problem of $(\mathscr{D})_r$ corresponding to (X_0,σ) admits a unique solution which is Lipschitz and satisfies (3.44)-(3.46). It follows that the results in [70, Lemmas 5.1 &5.2] remain valid for the systems $(\tilde{\mathscr{D}})$, and $(\mathscr{D})_r$, defined here, and also for the corresponding $(\mathscr{D})_c$, where

$$(\mathscr{D})_c \begin{cases} \dot{X}(t) \in \text{conv}\, h(t,X(t),U(t)) - N_{\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)}(X(t)), \;\; \text{a.e. } t \in [0,T]. \end{cases}$$

Therefore, for $X := (x, y) \in W^{1,1}([0, T], \mathbb{R}^{n+l})$ and $X(0) \in \bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(0)$, we have

$$\boxed{\begin{aligned}(X, w) \text{ satisfies } (\tilde{\mathscr{D}}),\ \text{for some } w \in \mathscr{W} &\iff (X, \sigma) \text{ satisfies } (\mathscr{D})_r \text{ for some } \sigma \in \mathscr{R}\\ &\iff X \text{ satisfies } (\mathscr{D})_c.\end{aligned}}$$

Furthermore, due to having (3.46) satisfied by the solutions of $(\mathscr{D})_r$ and due to the hypomonotonicity property of the uniform prox-regular sets $\bar{\mathscr{N}}_{(\bar{\varepsilon},\bar{\delta})}(t)$ (which we recall it to be the product of the uniform $\frac{2\bar{\eta}}{L_\psi}$-prox-regular set $C(t) \cap \bar{B}_{\bar{\varepsilon}}(\bar{x}(t))$ with $\bar{B}_{\bar{\delta}}(\bar{y}(t))$), it follows that [36, Theorem 2] (also [15, Proposition 3.5]) is valid. Hence, using that $(\tilde{X}, \tilde{w})$ is optimal for $(\tilde{\mathscr{P}})$, the proof of [70, Proposition 5.2] holds true for our setting, and therefore, as $(\bar{X}, \bar{u})$ is optimal for $(\bar{\mathscr{P}})$, we conclude that

$$\boxed{\min(\mathscr{P})_r = \min(\tilde{\mathscr{P}}) = \min(\bar{\mathscr{P}}) = J(\bar{X}(0), \bar{X}(T)).} \tag{4.88}$$

Now, since $(\bar{X}, \bar{w})$ is admissible for $(\tilde{\mathscr{P}})$ at which the objective value is $J(\bar{X}(0), \bar{X}(T))$, we deduce that $\boxed{(\bar{X}, \bar{w}) \text{ is a } \textit{global} \text{ minimum for } (\tilde{\mathscr{P}}).}$ This terminates proving Key Step 4(c).

Step II.2.4. $((\bar{x}, \bar{y}), \bar{w})$ **is a** $\frac{\delta}{2}$**-strong local minimum for** $(\tilde{P})$ **to which we apply Theorem 4.2.11.**

As $(\bar{X}, \bar{w})$ is a global minimizer for $(\tilde{\mathscr{P}})$, it follows that it is also a $\frac{\delta}{2}$-*strong* local minimum for $(\tilde{\mathscr{P}})$, which, by the first equation of (4.87), has now $\bar{J}(X(0), X(T))$ as objective function. Hence, we conclude that $((\bar{x}, \bar{y}), \bar{w})$ is a $\frac{\delta}{2}$-strong local minimum for the problem $(\tilde{P})$

$$(\tilde{P}) \begin{cases} \text{minimize} \quad \bar{J}(x(0), y(0), x(T), y(T)) \\ \text{over } X := (x, y) \in W^{1,1}([0, T], \mathbb{R}^{n+l}), \\ \qquad w(\cdot) := \Big((u_0(\cdot), \cdots, u_{n+l}(\cdot)), (\lambda_0(\cdot), \cdots, \lambda_{n+l}(\cdot))\Big) \in \mathscr{W} \text{ such that} \\ \qquad (\tilde{D}) \begin{cases} \dot{x}(t) \in \tilde{f}(t, x(t), y(t), w(t)) - N_{C(t)}(x(t)), \text{ a.e. } t \in [0, T], \\ \dot{y}(t) = \tilde{g}(t, x(t), y(t), w(t)), \text{ a.e. } t \in [0, T], \end{cases} \\ (x(0), y(0), x(T), y(T)) \in S_{\frac{\delta}{2}}, \end{cases}$$

where $(\tilde{f}, \tilde{g}) = \tilde{h}$ defined in (4.86), that is,

$$\tilde{f}(t,x,y,w) := \sum_{i=0}^{n+l} \lambda_i f(t,x,y,u_i), \quad \text{and} \quad \tilde{g}(t,x,y,w) := \sum_{i=0}^{n+l} \lambda_i g(t,x,y,u_i).$$

Clearly $(\tilde{P})$ is of the form of (P), where $f(t,x,y,u) := \tilde{f}(t,x,y,w)$, $g(t,x,y,u) := \tilde{g}(t,x,y,w)$, $S := S_{\frac{\delta}{2}}$, $U(t) := W(t)$, and $J := \bar{J}$. Furthermore, the associated $\tilde{h}(t,x,y,u) = (\tilde{f},\tilde{g})(t,x,y,w)$ satisfies that $\tilde{h}(t,x,y,W(t))$ *convex* for each $(t,x,y) \in \operatorname{Gr} \bar{\mathscr{N}}_{(\bar{\delta},\bar{\delta})}(\cdot)$. Thus, assumptions (A1)-(A5) hold at the strong local minimizer $((\bar{x},\bar{y}),\bar{w})$ for $(\tilde{P})$ to which the already proven (i)-(vii) of Theorem 4.2.11 apply. Doing so, and noticing these facts:

- $\bar{J} = J$ on $S(\bar{\delta})$, and hence, $\partial_\ell^L \bar{J}(\bar{x}(0),\bar{y}(0),\bar{x}(T),\bar{y}(T)) = \partial_\ell^L J(\bar{x}(0),\bar{y}(0),\bar{x}(T),\bar{y}(T))$,
- $\tilde{h}(t,\bar{x}(t),\bar{y}(t),\bar{w}(t)) = h(t,\bar{x}(t),\bar{y}(t),\bar{u}(t))$,
- $\partial_\ell^{(x,y)} \tilde{f}(t,\bar{x}(t),\bar{y}(t),\bar{w}(t)) \subset \partial_\ell^{(x,y)} f(t,\bar{x}(t),\bar{y}(t),\bar{u}(t))$,
- $\partial_\ell^{(x,y)} \tilde{g}(t,\bar{x}(t),\bar{y}(t),\bar{w}(t)) \subset \partial_\ell^{(x,y)} g(t,\bar{x}(t),\bar{y}(t),\bar{u}(t))$,
- $\langle \tilde{h}(t,\bar{x}(t),\bar{y}(t),w), p(t)\rangle = \langle h(t,\bar{x}(t),\bar{y}(t),u), p(t)\rangle, \quad \forall w = ((u,...,u),(1,0,...,0)) \in U^{n+l+1} \times \Lambda$,
- $N^L_{S_{\frac{\delta}{2}}}(\bar{x}(0),\bar{y}(0)) = N^L_S(\bar{x}(0),\bar{y}(0))$,

we conclude that Theorem 4.2.11 holds for (P) without assumption $(A4.2)$.

Step II.3 Proof of the "In addition" part of the theorem.

When $S = C_0 \times \mathbb{R}^{n+l}$, for $C_0 \subset C(0) \times \mathbb{R}^l$ closed, Remark 4.2.10 yields that $\lambda = 1$.

This completes the proof of the theorem. □

Table 4.4 Summary of results from Section 4.2.3 .

Result	Description
Theorem 4.2.11	We provide necessary conditions, in the form of an extended Pontryagin's maximum principle, for a $\bar{\delta}$-strong local minimizer $((\bar{x},\bar{y}),\bar{u})$ for the problem (P).

CHAPTER 5

VALIDATING THEORETICAL RESULTS USING AN EXAMPLE

Consider the problem (P) with the following data.

- The perturbation mappings $f\colon [0, \frac{\pi}{2}]\times\mathbb{R}^3\times\mathbb{R}\times\mathbb{R} \longrightarrow \mathbb{R}^3$ and $g\colon [0, \frac{\pi}{2}]\times\mathbb{R}^3\times\mathbb{R}\times\mathbb{R} \longrightarrow \mathbb{R}$ are defined by

$$f\left(t,(x_1,x_2,x_3),y,u\right) = (x_1 - x_2 - u + y^2, x_1 + x_2 + u + y^3, x_3 + t - \pi - 1);$$
$$g\left(t,(x_1,x_2,x_3),y,u\right) = x_1^2 + x_2^2 - 16 + u + y.$$

- The two functions $\psi_1,\ \psi_2\colon [0, \frac{\pi}{2}] \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ are defined by

$$\psi_1(t,x_1,x_2,x_3) := x_1^2 + x_2^2 + \frac{32}{\pi}x_3 + \frac{32}{\pi}t - 48,$$
$$\psi_2(t,x_1,x_2,x_3) := x_1^2 + x_2^2 - \frac{32}{\pi}x_3 - \frac{32}{\pi}t + 16,$$

and hence, for each $t \in [0, \frac{\pi}{2}]$, the set $C(t)$ is the *nonsmooth*, convex and *bounded* set (see Figure 5.1)

$$\begin{aligned} C(t) &= C_1(t) \cap C_2(t) \\ &:= \{(x_1,x_2,x_3) : \psi_1(t,x_1,x_2,x_3) \le 0\} \cap \{(x_1,x_2,x_3) : \psi_2(t,x_1,x_2,x_3) \le 0\}. \end{aligned}$$

- The objective function $J\colon \mathbb{R}^8 \longrightarrow \mathbb{R} \cup \{\infty\}$ is defined by

$$J(x_1,x_2,x_3,y_1,x_4,x_5,x_6,y_2) := \begin{cases} -x_4^2 - x_5^2 + 16 + \left|\frac{\pi}{2} - x_6\right| & (x_4,x_5,x_6) \in C(\frac{\pi}{2}), \\ \infty & \text{Otherwise.} \end{cases}$$

- The control multifunction is the constant $U(t) := [0,1]$ for all $t \in [0, \frac{\pi}{2}]$.
- The set S is given by

$$S := \{(x_1,x_2,x_3,y_1,x_4,x_5,x_6,y_2) \in \mathbb{R}^8 : x_1^2 + x_2^2 = 16, x_3 = \pi, x_2 + x_6^2 = \frac{\pi^2}{4}, \frac{x_1^2}{8} + x_4 = 2, y_1 + x_2^2 = 0\}.$$

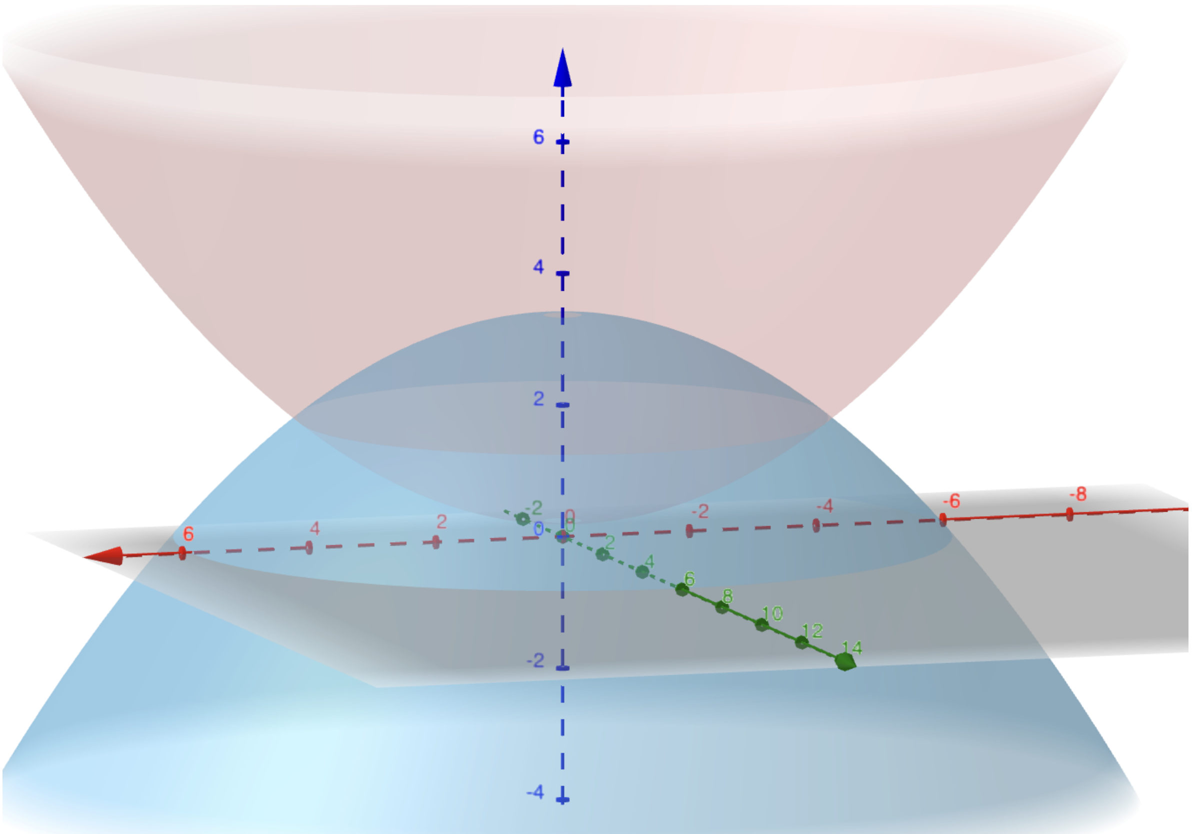

Figure 5.1 The sweeping set $C(t_o)$ at a certain time $t_o \in (0, \frac{\pi}{2})$

Define, for each $t \in [0, \frac{\pi}{2}]$, the curve

$$\Gamma(t) := \Big\{(x_1, x_2, x_3) : x_1^2 + x_2^2 = 16 \text{ and } x_3 = \pi - t\Big\} = (\operatorname{bdry} C_1(t) \cap \operatorname{bdry} C_2(t)) \subset \operatorname{bdry} C(t).$$

Since $S \subset \Gamma(0) \times \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}$ and J vanishes on $\mathbb{R}^3 \times \mathbb{R} \times \Gamma(\frac{\pi}{2}) \times \mathbb{R}$ and is strictly positive elsewhere in $\mathbb{R}^3 \times \mathbb{R} \times C(\frac{\pi}{2}) \times \mathbb{R}$, we may seek for (P) a candidate $((\bar{x}, \bar{y}), \bar{u})$ for optimality with $\bar{x}(t) := (\bar{x}_1(t), \bar{x}_2(t), \bar{x}_3(t))$ belonging to $\Gamma(t)$ for every t, if possible, and hence we have

$$\begin{cases} \bar{x}_1^2(t) + \bar{x}_2^2(t) = 16 \ \text{ and } \ \bar{x}_3(t) = \pi - t \ \forall t \in [0, \frac{\pi}{2}] \ \text{ and} \\ \bar{x}_1(t)\dot{\bar{x}}_1(t) + \bar{x}_2(t)\dot{\bar{x}}_2(t) = 0 \text{ a.e.} \quad \text{and} \\ (\bar{x}(0)^\intercal, \bar{y}(0)^\intercal, \bar{x}(\frac{\pi}{2})^\intercal, \bar{y}(\frac{\pi}{2})^\intercal) \in \{(4, 0, \pi, 0, 0, 4, \frac{\pi}{2}, a), (-4, 0, \pi, 0, 0, 4, \frac{\pi}{2}, b), \\ \qquad (4, 0, \pi, 0, 0, -4, \frac{\pi}{2}, c), (-4, 0, \pi, 0, 0, -4, \frac{\pi}{2}, d); \ a, b, c, d \in \mathbb{R}\}. \end{cases} \tag{5.1}$$

One can readily verify that all assumptions of Theorem 4.2.11 are satisfied for any choice of

$(\bar{x}, \bar{y})$ such that $\bar{x}(t) \in \Gamma(t)$ for all t, with $(A3.3)$ being satisfied for $\bar{\beta} = (1,1)$. Applying[1] Theorem 4.2.11 to such candidate $((\bar{x}, \bar{y}), \bar{u})$, we obtain the existence of an adjoint vector $p = (q, v)$ where $q := (q_1, q_2, q_3) \in BV([0, \frac{\pi}{2}]; \mathbb{R}^3)$, $v \in W^{1,1}([0, \frac{\pi}{2}]; \mathbb{R})$, two finite signed Radon measures ν_1, ν_2 on $\left[0, \frac{\pi}{2}\right]$, ξ_1, $\xi_2 \in L^\infty([0, \frac{\pi}{2}]; \mathbb{R}^+)$, and $\lambda \geq 0$, such that when incorporating equations (5.1) into Theorem 4.2.11(i)-(vii), we obtain

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

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

$$\begin{cases} \dot{\bar{x}}_1(t) = \bar{x}_1(t) - \bar{x}_2(t) - \bar{u}(t) + \bar{y}^2(t) - 2\bar{x}_1(t)(\xi_1(t) + \xi_2(t)), \\ \dot{\bar{x}}_2(t) = \bar{x}_1(t) + \bar{x}_2(t) + \bar{u}(t) + \bar{y}^3(t) - 2\bar{x}_2(t)(\xi_1(t) + \xi_2(t)), \\ \dot{\bar{x}}_3(t) = \bar{x}_3(t) + t - \pi - 1 - \frac{32}{\pi}(\xi_1(t) - \xi_2(t)), \\ \dot{\bar{y}}(t) = \bar{x}_1^2(t) + \bar{x}_2^2(t) - 16 + \bar{u}(t) + \bar{y}(t). \end{cases}$$

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

$$\begin{aligned} dq(t) &= \begin{pmatrix} -1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} q(t)\, dt + \begin{pmatrix} -2\bar{x}_1(t) \\ -2\bar{x}_2(t) \\ 0 \end{pmatrix} v(t) dt \\ &+ (\xi_1(t) + \xi_2(t)) \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} q(t)\, dt + \begin{pmatrix} 2\bar{x}_1(t) \\ 2\bar{x}_2(t) \\ \frac{32}{\pi} \end{pmatrix} d\nu_1 + \begin{pmatrix} 2\bar{x}_1(t) \\ 2\bar{x}_2(t) \\ -\frac{32}{\pi} \end{pmatrix} d\nu_2, \\ \dot{v}(t) &= \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} q(t)\, dt - v(t) \end{aligned}$$

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

$$\begin{cases} \xi_1(t)(2q_1(t)\bar{x}_1(t) + 2q_2(t)\bar{x}_2(t) + \frac{32}{\pi} q_3(t)) = 0, \\ \xi_2(t)(2q_1(t)\bar{x}_1(t) + 2q_2(t)\bar{x}_2(t) - \frac{32}{\pi} q_3(t)) = 0. \end{cases}$$

[1]Note that for $(x_1, x_2, x_3) \in \Gamma(t)$ with $-\frac{\sqrt{3}}{2} < x_1 < \frac{\sqrt{3}}{2}$, we have $\langle \nabla\psi_1(x_1, x_2, x_3), \nabla\psi_2(x_1, x_2, x_3)\rangle = 4x_1^2 - 3 < 0$, and hence, the maximum principle of [34] cannot be applied to this sweeping set $C(t)$.

(e) The transversality condition holds, that is,

$$(q(0), v(0), -q(\tfrac{\pi}{2}), -v(\tfrac{\pi}{2}))^{\mathsf{T}} \in \lambda\{(0,0,0,0,0,-8,\alpha,0) : \alpha \in [-1,1]\}$$
$$+\{(8\alpha_1+\alpha_4, \alpha_3, \alpha_2, \alpha_5, \alpha_4, 0, \alpha_3\pi, 0) : \alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5 \in \mathbb{R}\}$$
$$\text{if } (\bar{x}(0)^{\mathsf{T}}, \bar{y}(0)^{\mathsf{T}}, \bar{x}(\tfrac{\pi}{2})^{\mathsf{T}}, \bar{y}(\tfrac{\pi}{2})^{\mathsf{T}}) \in \{(4,0,\pi,0,0,4,\tfrac{\pi}{2},a) : a \in \mathbb{R}\}.$$

Similarly, we work on deriving transversality conditions for each of the four cases in (5.1).

(f) $\max\{u\,(-q_1(t)+q_2(t)+v(t)) : u \in [0,1]\}$ is attained at $\bar{u}(t)$ for $t \in [0,\frac{\pi}{2}]$ a.e.

We temporarily assume that

$$-q_1(t)+q_2(t)+v(t) < 0, \ \ \forall t \in [0,\tfrac{\pi}{2}] \text{ a.e.} \tag{5.2}$$

This gives from (f) that $\bar{u}(t)=0$ for $t \in [0,\frac{\pi}{2}]$ a.e. Now solving the differential equations of (b) and using (5.1), we obtain that

$$\xi_1(t)=\xi_2(t)=\frac{1}{4},\ \ \bar{x}(t)^{\mathsf{T}} = (4\cos t, 4\sin t, \pi - t)^2, \ \text{ and } \ \bar{y}(t)=0 \ \ \forall t \in [0,\tfrac{\pi}{2}].$$

Hence, from (d), we deduce that $q_3(t)=0$ for $t \in [0,\frac{\pi}{2}]$ a.e., and

$$\cos t\ q_1(t) + \sin t\ q_2(t) = 0, \ \ \forall t \in [0,\tfrac{\pi}{2}] \text{ a.e.,} \tag{5.3}$$

and the adjoint equation (c) simplifies to the following

$$\begin{cases} \dot{v}(t) = -v(t), \\ dq_1(t) = (-q_1(t)-q_2(t))dt - 8\cos t\ v(t)dt + q_1(t)\ dt + 8\cos t\ (d\nu_1 + d\nu_2), \\ dq_2(t) = (q_1(t)-q_2(t))dt - 8\sin t\ v(t)dt + q_2(t)\ dt + 8\sin t\ (d\nu_1 + d\nu_2), \\ dq_3(t) = -q_3(t)\ dt + \frac{32}{\pi}(d\nu_1 - d\nu_2). \end{cases} \tag{5.4}$$

[2]Note that another possible choice for $\bar{x}(\cdot)$ is $\bar{x}(t)^{\mathsf{T}} = (-4\cos t, -4\sin t, \pi - t)$.

Since $v(\frac{\pi}{2}) = 0$ then $v(t) = 0 \; \forall t \in [0, \frac{\pi}{2}]$. Using (a), (5.3), (e), and (5.4), one can get the following

$$\begin{cases} \lambda = \frac{2\pi}{2\pi+\sqrt{1+(16\pi)^2}} \text{ and } A = \frac{1}{2\pi+\sqrt{1+(16\pi)^2}}, \\ q(t)^{\mathsf{T}} = (A\sin t, -A\cos t, 0) \text{ on } [0, \frac{\pi}{2}), \qquad q(\frac{\pi}{2})^{\mathsf{T}} = (A, 16A\pi, 0), \\ d\nu_1 = d\nu_2 = A\pi\delta_{\left\{\frac{\pi}{2}\right\}}, \end{cases}$$

where $\delta_{\{a\}}$ denotes the unit measure concentrated on the point a. Note that for all $t \in [0, \frac{\pi}{2}]$, we have $-q_1(t) + q_3(t) + v(t) < 0$, and hence, the temporary assumption (5.2) is satisfied. Therefore, the above analysis, realized via Theorem 4.2.11, produces an admissible pair $((\bar{x}, \bar{y}), \bar{u})$, where

$$\bar{x}(t)^{\mathsf{T}} = (4\cos t, 4\sin t, \pi - t), \;\; \bar{y}(t) = 0, \;\; \text{and} \;\; \bar{u}(t) = 0, \;\; \forall t \in [0, \tfrac{\pi}{2}],$$

which is optimal for (P).

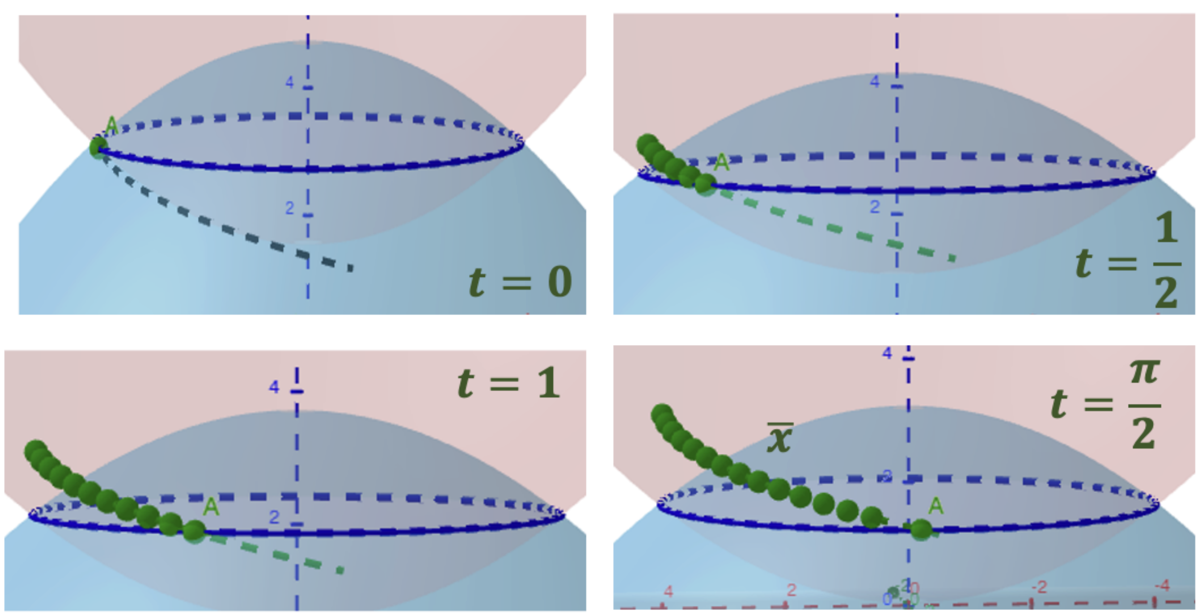

Figure 5.2 The solution $\bar{x}(t)$ (in green) evolving on the set $C(t) = C_1(t) \cap C_2(t)$ over different time instances.

CHAPTER 6

CONCLUSION AND POSSIBLE FUTURE DIRECTIONS

6.1 Conclusion

In this dissertation, we employ the exponential penalty-type approximation method to launch the study of a general model (P) given by:

$$(P)\begin{cases}\text{minimize} \quad J(x(0),y(0),x(T),y(T))\\ \text{over } ((x,y),u)\in W^{1,1}([0,T],\mathbb{R}^n\times\mathbb{R}^l)\times\mathcal{U} \text{ such that}\\ \quad (D)\begin{cases}\dot{x}(t)\in f(t,x(t),y(t),u(t))-N_{C(t)}(x(t)), \text{ a.e. } t\in[0,T],\\ \dot{y}(t)=g(t,x(t),y(t),u(t)), \text{ a.e. } t\in[0,T],\end{cases}\\ \quad (x(0),y(0),x(T),y(T))\in S,\end{cases}$$

where, for $t\in[0,T]$, the set $C(t)$ is defined as the intersection of a finite number of zero sub-level sets of $(\psi_i(t,\cdot))_{i=1}^r$, referred to as generators.

One of the main results of our work, which is global, encompasses the existence and uniqueness of a Lipschitz solution for the Cauchy problem corresponding to our dynamic (D) without requiring any *Lipschitz* property on $C(\cdot)$—a condition commonly required in the literature (see e.g., [36]). Instead, we assume Gr $C(\cdot)$ is bounded and the gradients of the active generators are positively linear independent. Note that this is the first such a result for *general nonsmooth moving* sweeping sets, even for the uncoupled sweeping process, which is based on the method of exponential penalty approximation.

Another global main result encompasses the global existence of optimal solution for our problem (P) under global assumptions. We note that this constitutes the first attempt to prove existence result of optimal solutions for time-dependent *general* sweeping set.

The main local result consists of deriving under *minimal* assumptions on the data, a complete set of necessary conditions in the form of nonsmooth Pontryagin maximum principle for *strong* local minimizers of the problem (P) via developing the exponential penalization technique. Our Pontryagin maximum principle generalizes previously known Pontryagin maximum principle results ([30, 31, 33, 34, 70, 55, 58]). In fact, we establish a Pontryagin maximum principle in its *expected* form (i.e., standard nontriviality condition, adjoint equation, transversality condition, and the maximality condition on the Hamiltonian) for optimal control problems over the sweeping process (1.2) in each of the following settings:

(i) When the *nonsmooth moving* sweeping sets $C(t)$ are *bounded* and *general* (no restriction on the corners);

(ii) When the *general nonsmooth* sweeping sets are *unbounded* (*constant* or *moving*);

(iii) When *joint* state endpoints constraint set is *present*, the *convexity* of $f(t, x, U(t))$ is *absent*, or the *global* constraint qualification is only *local*, for all types of sweeping sets: *smooth*, *nonsmooth*, *constant*, *moving*, *bounded*, or *unbounded*;

(iv) When the sweeping process is coupled with a differential equation.

6.2 Future directions

In this section, we outline several promising future directions that stem from our current work on optimal control problems over sweeping processes. We will focus on five key areas: extending the model to include state constraints, developing a numerical algorithm to solve our model, incorporating control into the sweeping set, exploring the bilateral minimal time function in the context of sweeping processes, and applying these results to real-world scenarios.

Project 1: Adding state constraint

We are currently working on extending the techniques discussed earlier to address problems that include explicit external state constraints: $\omega(t, x(t), y(t)) \leq 0$. This implies that our approximating problems differ from those in Chapter 4 due to the presence of an additional explicit state constraint. This introduces challenges when attempting to prove the boundedness of the adjoint vector for the approximating problem, which subsequently complicates the limit-taking process. It is worth noting that adding a state constraint to the sweeping process has been addressed in the literature, as seen in [44] for example, but only for a special case of our model.

Project 2: Numerical algorithm

We are interested in constructing a numerical algorithm to solve our Mayer problem (P), as in [32, 56, 59]. We plan to expand the domain of applicability of the numerical method to:

- Time-dependent sweeping set $C(t)$,
- Initial state set C_0 instead of fixed x_0,
- Final endpoint C_T instead of free final endpoint.

Project 3: The sweeping set is controlled and is of the form $C(t) + u(t)$

A potential future direction for this work would involve exploring the effects of introducing a control function into the sweeping set. Specifically, one could investigate how our results would change when the sweeping set is defined as $C(t) := C + v(t)$ where $v(\cdot)$ is a control function belonging to $W^{1,2}$.

Project 4: Finding the bilateral minimal time function for the sweeping process

The bilateral minimal time function, introduced by Clarke and Nour in [20], defines $T(\alpha, \beta)$ as the minimum time taken by a trajectory to go from α to β. In my master's thesis, I worked on studying the variational analysis and the sensitivity relations of the bilateral minimal time functions in order to study the regularity of this function for nonlinear control system. The results we obtained, published in [16], extends the main result of [54] where a similar result is obtained for the linear case. We can integrate the study of the bilateral minimal time function with the sweeping process. More specifically, we can study the bilateral minimal time function when the set-valued map that defines the trajectory is given as a sweeping process. This would build on the work done in [24], where the authors have worked on the unilateral minimal time function within the context of sweeping process.

Project 5: Real-life applications of the sweeping process

Another promising future direction involves validating both the numerical and theoretical results of optimal control problems governed by sweeping process using real-life case study models and experimental setups, such as crowd motion models in emergency evacuations, robotics models, marine surface vehicle modeling, and nanoparticle modeling.

BIBLIOGRAPHY

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

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

[3] A. Bensoussan, K. Chandrasekaran, and J. Turi. Optimal control of variational inequalities. *Commun. Inf. Syst.*, 10(4):203–220, 2010.

[4] A. Bergqvist. Magnetic vector hysteresis model with dry friction-like pinning. *Phys. B Condens. Matter*, 233(4):342–347, 1997.

[5] A. Bouach, T. Haddad, and B.S. Mordukhovich. Optimal control of nonconvex integro-differential sweeping processes. *J. Differ. Equ.*, 329:255–317, 2022.

[6] M. Brokate and P. Krejčí. Optimal control of ode systems involving a rate independent variational inequality. *Discrete Contin. Dyn. Syst. B*, 18:331–348, 2013.

[7] M. Brokate and J. Sprekels. *Hysteresis and Phase Transitions*, volume 121. Springer, New York, 1996.

[8] R. J. Stern C. Nour and J. Takche. Proximal smoothness and the exterior sphere condition. *Journal of Convex Analysis*, 16(2):501–514, 2009.

[9] A. Canino. On p-convex sets and geodesics. *Journal of Differential Equations*, 75(1): 118–157, 1998.

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

[11] T.H. Cao and B. Mordukhovich. Optimal control of a nonconvex perturbed sweeping process. *Journal of Differential Equations*, 266:1003–1050, 2019.

[12] T.H. Cao, G. Colombo, B. Mordukhovich, and D. Nguyen. Optimization of fully controlled sweeping processes. *Journal of Differential Equations*, 295:138–186, 2021.

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

[14] T.H. Cao, B.S. Mordukhovich, D. Nguyen, and T. Nguyen. Applications of controlled sweeping processes to nonlinear crowd motion models with obstacles. *IEEE Control*

Systems Letters, 6:740–745, 2021.

[15] C. Castaing, A. Salvadori, and L. Thibault. Functional evolution equations governed by nonconvex sweeping process. *J. Nonlinear Conv. Anal.*, 2:217–241, 2001.

[16] S. Chamoun and C. Nour. A nonlinear φ_0-convexity result for the bilateral minimal time function. *Journal of Convex Analysis*, 28(1):081–102, 2021.

[17] Samara Sarkis Chamoun. Variational analysis and sensitivity relations of the bilateral minimal time function for nonlinear differential inclusions. Master's thesis, Lebanese American University, 2019.

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

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

[20] F. H. Clarke and C. Nour. The hamilton-jacobi equation of minimal time control. *J. Convex Anal.*, 11(2):413–436, 2004.

[21] F. H. Clarke, R. J. Stern, and P. R. Wolenski. Proximal smoothness and the lower-c^2 property. *J. Convex Anal.*, 2:117–144, 1995.

[22] F. H. Clarke, Yu. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory*, volume 178 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[23] G. Colombo and P. Godini. On the optimal control of rate-independent soft crawlers. *Journal de Mathématiques Pures et Appliquées*, 2020.

[24] G. Colombo and M. Palladino. The minimum time function for the controlled moreau's sweeping process. *IAM Journal on Control and Optimization*, 54(4):2036–2062, 2016.

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

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

[27] G. Colombo, B. Mordukhovich, and D. Nguyen. Optimal control of sweeping processes in robotics and traffic flow models. *J. Optim. Theory Appl.*, 182(2):439–472, 2019.

[28] G. Colombo, B. Mordukhovich, and D. Nguyen. Optimization of a perturbed sweeping process by constrained discontinuous controls. *SIAM Journal on Control and Optimiza-*

tion, 58(4):2678–2709, 2020.

[29] Giovanni Colombo and Antonio Marigonda. Differentiability properties for a class of non-convex functions. *Calculus of Variations*, 25:1–31, 2006. doi: 10.1007/s00526-005-0352-7.

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

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

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

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

[34] M.d.R. de Pinho, M.M.A. Ferreira, and G.V. Smirnov. A maximum principle for optimal control problems involving sweeping processes with a nonsmooth set. *J. Optim. Theory Appl.*, 199:273–297, 2023.

[35] D. Drusvyatskiy, A.D. Ioffe, and A.S. Lewis. Clarke subgradients of directionally lipschitzian stratifiable functions. *Mathematics of Operations Research*, 40(2):328–349, 2015.

[36] J.F. Edmond and L. Thibault. Relaxation of an optimal control problem involving a perturbed sweeping process. *Mathematical Programming - Series B*, 104:347–373, 2005.

[37] L. C. Evans. *Weak convergence methods for nonlinear partial differential equations*, volume 74 of *CBMS Regional Conference Series in Mathematics*. Conference Board of the Mathematical Sciences, Washington, DC, 1990.

[38] L. C. Evans. *Partial differential equations.* American Mathematical Society, Providence, RI, 1998.

[39] J.K. Hale. *Ordinary Differential Equations.* Krieger Publishing Company, 2 edition, 1980.

[40] J.M. Harrison and M.I. Reiman. Reflected brownian motion on an orthant. *Annals of Probability*, 9:302–308, 1981.

[41] R. Henrion, A. Jourani, and B. S. Mordukhovich. Controlled polyhedral sweeping processes: Existence, stability, and optimality conditions. *Journal of Differential Equations*,

366:408–443, 2023.

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

[43] J.-B. Hiriart-Urruty. Extension of lipschitz functions. *Journal of Mathematical Analysis and Applications*, 77:539–554, 1980.

[44] N.T. Khalil and F.L. Pereira. A maximum principle for state-constrained optimal sweeping control problems. *IEEE Control Systems Letters*, 7:43–48, 2023.

[45] P. Krejčí and J. Sprekels. Temperature-dependent hysteresis in one-dimensional thermovisco-elastoplasticity. *Applied Mathematics*, 43(3):173–205, 1998.

[46] B. Li. *Generalizations of Diagonal Dominance in Matrix Theory.* PhD thesis, University of Saskatchewan, Regina, 1997.

[47] Xing-Si Li and Shu-Cherng Fang. On the entropic regularization method for solving min-max problems with applications. *Mathematical Methods of Operations Research*, 46:119–130, 1997.

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

[49] J.J. Moreau. Rafle par un convexe variable, i. *Travaux Séminaire d'Analyse Convexe, Montpellier 1*, Exposé 15:36, 1971.

[50] J.J. Moreau. Rafle par un convexe variable, ii. *Travaux Séminaire d'Analyse Convexe, Montpellier 2*, Exposé 3:43, 1972.

[51] J.J. Moreau. Evolution problem associated with a moving convex set in a hilbert space. *Journal of Differential Equations*, 26:347–374, 1977.

[52] F. Nacry and L. Thibault. Regularization of sweeping process: old and new. *Pure and Applied Functional Analysis*, 4(1):59–117, 2019.

[53] F. Nacry, J. Noel, and L. Thibault. On first and second order state-dependent prox-regular sweeping process. *Pure and Applied Functional Analysis*, 6(6):1453–1493, 2021.

[54] C. Nour. Proximal subdifferential of the bilateral minimal time function and some regularity applications. *Journal of Convex Analysis*, 20(4):1095–1112, 2013.

[55] C. Nour and V. Zeidan. Optimal control of nonconvex sweeping processes with separable endpoints: Nonsmooth maximum principle for local minimizers. *Journal of Differential Equations*, 318:113–168, 2022.

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

[57] C. Nour and V. Zeidan. A control space ensuring the strong convergence of continuous approximation for a controlled sweeping process. *Set-Valued and Variational Analysis*, 31(3), 2023.

[58] C. Nour and V. Zeidan. Pontryagin-type maximum principle for a controlled sweeping process with nonsmooth and unbounded sweeping set. *Journal of Convex Analysis*, 31 (3):787–825, 2024.

[59] C. Nour and V. Zeidan. Numerical method for a controlled sweeping process with nonsmooth sweeping set. *Journal of Optimization Theory and Applications*, 203(2): 1385–1412, 2024.

[60] R. A. Poliquin and R. T. Rockafellar. Prox-regular functions in variational analysis. *Transactions of the American Mathematical Society*, 348(5):1805–1838, 1996.

[61] R. T. Rockafellar R. A. Poliquin and L. Thibault. Local differentiability of distance functions. *Transactions of the American Mathematical Society*, 352(11):5231–5249, 2000.

[62] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1998.

[63] G. Salinetti and R. J.-B. Wets. On the convergence of sequences of convex sets in finite dimensions. *SIAM Review*, 21:18–33, 1979.

[64] L. Thibault and J. J. Moreau. Sweeping process with bounded truncated retraction. *Journal of Convex Analysis*, 23(4):1051–1098, 2016.

[65] A. A. Tolstonogov. Polyhedral sweeping processes with unbounded nonconvex-valued perturbation. *Journal of Differential Equations*, 263:7965–7983, 2017.

[66] R. Vinter. *Optimal Control*, volume 1. Springer, 2017.

[67] A. Visintin. *Differential Models of Hysteresis*, volume 111. Springer Berlin Heidelberg, 1994.

[68] J. Warga. *Optimal Control of Differential and Functional Equations*. Academic Press, New York and London, 1972.

[69] V. Zeidan. A modified hamilton-jacobi approach in the generalized problem of bolza. *Appl. Math. Optim*, (11):97–109, 1984.

[70] V. Zeidan, C. Nour, and H. Saoud. A nonsmooth maximum principle for a controlled

nonconvex sweeping process. *Journal of Differential Equations*, 269(11):9531–9582, 2021.

APPENDIX

APPENDIX TO CHAPTERS 3-4

Translating Lemma 6.2 in [58] to our setting gives us the following lemma.

Lemma .0.1. Assume that ψ_i is continuous for all $i = 1, \ldots, r$. Let $\alpha_n \geq 0$, for all $n \in \mathbb{N}$, with $\alpha_n \longrightarrow \alpha_o$ and let $(t_n, c_n) \in \text{Gr } C(\cdot)$ be a sequence such that $\mathcal{I}^{\alpha_n}_{(t_n,c_n)} \neq \emptyset$, for all $n \in \mathbb{N}$, and $(t_n, c_n) \longrightarrow (t_o, c_o)$. Then, $\mathcal{I}^{\alpha_o}_{(t_o,c_o)} \neq \emptyset$ and there exist $\emptyset \neq \mathcal{J}_o \subset \{1, \ldots, r\}$ and a subsequence of $(\alpha_n, t_n, c_n)_n$ we do not relabel, such that

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

In particular, for all $a \geq 0$, for any continuous function $x \colon [0, T] \longrightarrow \mathbb{R}^n$ such that $x(t) \in C(t)$ for all $t \in [0, T]$, we have $I^a(x)$ is closed, and hence compact.

This result shall be used in different places of the thesis.

Lemma .0.2. *(i)* Let $(x_n, y_n) \in W^{1,\infty}([0, T]; \mathbb{R}^{n+d})$, $(\xi^1_n, \cdots, \xi^{\mathcal{R}}_n, \zeta_n) \in L^\infty([0, T], \mathbb{R}^{\mathcal{R}+1}_+)$, be such that, for some positive constants M_1, M_2, M_3 we have, $\forall n \in \mathbb{N}$ and $\forall i \in \{1, \cdots, \mathcal{R}\}$,

$$\|(x_n, y_n)\|_\infty \leq M_1, \quad \|(\dot{x}_n, \dot{y}_n)\|_\infty \leq M_2, \quad \|(\xi^i_n, \zeta_n)\|_\infty \leq M_3. \tag{.1}$$

Then, there exist $(x, y) \in W^{1,\infty}([0, T]; \mathbb{R}^{n+d})$ and $(\xi^1, \cdots, \xi^{\mathcal{R}}, \zeta) \in L^\infty([0, T]; \mathbb{R}^{\mathcal{R}+1}_+)$ such that (x_n, y_n), and $(\xi^1_n, \cdots, \xi^{\mathcal{R}}_n, \zeta_n)$ admit a subsequence (not relabeled) satisfying $\forall i \in \{1, \cdots, \mathcal{R}\}$,

$$\begin{cases} (x_n, y_n) \xrightarrow{unif} (x, y), \quad (\dot{x}_n, \dot{y}_n) \xrightarrow[in\ L^\infty]{w*} (\dot{x}, \dot{y}), \quad (\xi^i_n, \zeta_n) \xrightarrow[in\ L^\infty]{w*} (\xi^i, \zeta), \\ \|(x, y)\|_\infty \leq M_1, \quad \|(\dot{x}, \dot{y})\|_\infty \leq M_2, \quad \|(\xi^i, \zeta)\|_\infty \leq M_3. \end{cases} \tag{.2}$$

(ii) For given $q_i : [0, T] \times \mathbb{R}^n \longrightarrow \mathbb{R}$, let $Q(t) := \cap_{i=1}^{\mathcal{R}} \{x \in \mathbb{R}^n : q_i(t, x) \leq 0\}$ and $(\bar{x}, \bar{y}) \in \mathcal{C}([0, T]; \mathbb{R}^n \times \mathbb{R}^l)$ be such that $\bar{x}(t) \in Q(t)$ $\forall t \in [0, T]$. Assume $(A2)$ is satisfied by $C(t) := Q(t)$, and, for some $\bar{\delta} > 0$, $(A3.1)$ and $(A4.1)$ hold at $((\bar{x}, \bar{y}); \bar{\delta})$ respectively by $\psi_i := q_i$ and $h = (f, g) : [0, T] \times \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \longrightarrow \mathbb{R}^n \times \mathbb{R}^l$. Let (x_n, y_n) and $(\xi_n, \cdots, \xi^{\mathcal{R}}_n, \zeta_n)$ be such

that $(x_n(t), y_n(t)) \in [Q(t) \cap \bar{B}_{\bar{\delta}}(\bar{x}(t))] \times \bar{B}_{\bar{\delta}}(\bar{y}(t))$ $(\forall t \in [0,T])$ and (.1) is satisfied, and let (x, y, ξ^i, ζ) be their corresponding limits via (.2). Consider $u_n \in \mathcal{U}$ such that, for all $n \in \mathbb{N}$, $((x_n, y_n), u_n)$, and $(\xi_n^1, \cdots, \xi_n^{\mathcal{R}}, \zeta_n)$ satisfy

$$\begin{cases} \dot{x}(t) = f(t, x(t), y(t), u(t)) - \sum_{i=1}^{\mathcal{R}} \xi^i(t) \nabla_x q_i(t, x(t)) \text{ a.e. } t \in [0,T], \\ \dot{y}(t) = g(t, x(t), y(t), u(t)) - \zeta(t) \nabla_y \varphi(t, y(t)) \text{ a.e. } t \in [0,T], \end{cases} \tag{.3}$$

where φ is given by (3.31).Then, in either of the following cases, there exists $u \in \mathcal{U}$ such that $((x, y), u)$, and $(\xi^1, \cdots, \xi^{\mathcal{R}}, \zeta)$ also satisfy system (.3).

Case 1. If there exists a subsequence of u_n that converges pointwise a.e. to some $u \in \mathcal{U}$.

Case 2. If $(A1)$ and $(A4.2)$ are satisfied.

Proof. (i): By (.1), the sequence $(x_n, y_n)_n$ is equicontinuous and uniformly bounded. Hence, using Arzela-Ascoli theorem and that $(\dot{x}_n, \dot{y}_n)$ is uniformly bounded in L^∞, it follows that there exists $(x, y) \in W^{1,\infty}([0,T]; \mathbb{R}^{n+d})$ such that along a subsequence (we do not relabel) of (x_n, y_n), we have $(x_n, y_n) \xrightarrow{unif} (x, y)$, $(\dot{x}_n, \dot{y}_n) \xrightarrow[in\ L^\infty]{w*} (\dot{x}, \dot{y})$, with (x, y) and $(\dot{x}, \dot{y})$ satisfy the bounds in (.2) (see Theorem 2.4.13). As $\|(\xi_n^i, \zeta_n)\|_\infty \le M_3$ for all $i = 1, \cdots, \mathcal{R}$ and for all $n \in \mathbb{N}$, some subsequences of $(\xi_n^1, \cdots, \xi_n^{\mathcal{R}}, \zeta_n)$ converge in the weak*-topology to some $(\xi^1, \cdots, \xi^{\mathcal{R}}, \zeta) \in L^\infty$ which satisfy the required bound in (.2) (see Theorem 2.4.11).

(ii) **Case 1.** Let $t \in [0,T)$ lebesgue point of $\dot{x}(\cdot), \dot{y}(\cdot), f(\cdot, x(\cdot), y(\cdot), u(\cdot)), G(\cdot, x(\cdot), y(\cdot), u(\cdot))$, $\xi^i(\cdot)$ for all $i = 1, \cdots, \mathcal{R}$ and ζ, and let $\tau \in (0, T-t)$. Then, (.3) implies

$$\begin{cases} \frac{x_n(t+\tau) - x_n(t)}{\tau} = \frac{1}{\tau} \int_t^{t+\tau} \left[f(s, x_n(s), y_n(s), u_n(s)) - \sum_{i=1}^{\mathcal{R}} \xi_n^i(s) \nabla_x q_i(s, x_n(s)) \right] ds, \\ \frac{y_n(t+\tau) - y_n(t)}{\tau} = \frac{1}{\tau} \int_t^{t+\tau} \left[g(s, x_n(s), y_n(s), u_n(s)) - \zeta_n(s) \nabla_y \varphi(s, y_n(s)) \right] ds. \end{cases} \tag{.4}$$

Using Dominated Convergence Theorem, and taking the limit as $n \to \infty$ of (.4), we deduce that

$$\begin{cases} \frac{x(t+\tau) - x(t)}{\tau} = \frac{1}{\tau} \int_t^{t+\tau} [f(s, x(s), y(s), u(s)) - \sum_{i=1}^{\mathcal{R}} \xi^i(s) \nabla_x q_i(s, x(s))] ds, \\ \frac{y(t+\tau) - y(t)}{\tau} = \frac{1}{\tau} \int_t^{t+\tau} [g(s, x(s), y(s), u(s)) - \zeta(s) \nabla_y \varphi(s, y(s))] ds. \end{cases} \tag{.5}$$

Now, let $\tau \to 0$ in (.5), we get that (.3) is satisfied for every t lebesgue point, hence it hold for a.e. $t \in [0,T]$.

(ii) **Case 2.** For $s \in [0,T]$ a.e., define in $\mathbb{R}^n \times \mathbb{R}^l$ the sets $S_n(s) := h(s, x_n(s), y_n(s), U(s))$ and $S(s) := h(s, x(s), y(s), U(s))$. Using (A1), the continuity of $h(s,\cdot,\cdot,\cdot)$ in (A4.1), and the convexity assumption (A4.2), it follows that $S_n(s)$ and $S(s)$ are nonempty closed convex sets and $S_n(s)$ Hausdorff-converges to $S(s)$. Hence, Filippov Selection Theorem yields that $((x_n, y_n), u_n)$ and (ξ_n^i, ζ_n) satisfying (.3) is equivalent to, $\forall z \in \mathbb{R}^{n+d}$ and $s \in [0,T]$ a.e.,

$$\begin{aligned}\langle z, (\dot{x}_n(s), \dot{y}_n(s))\rangle \leq \quad & \sigma(z, h(s, x_n(s), y_n(s), U(s))) \\ & -\langle z, (\sum_{i=1}^{\mathcal{R}} \xi_n^i(s) \nabla_x q_i(t, x_n(s)), \zeta_n(s) \nabla_y \varphi(s, y_n(s)))\rangle. \qquad (.6)\end{aligned}$$

Furthermore, by (2.4) and the positive homogeneity of $\sigma(\cdot, S_n)$ and $\sigma(\cdot, S)$, we deduce that

$$\sigma(z, h(s, x_n(s), y_n(s), U(s))) \xrightarrow[k\to\infty]{} \sigma(z, h(s, x(s), y(s), U(s))),\ \forall z \in \mathbb{R}^{n+d} \text{ and } \ s \in [0,T] \text{ a.e.},$$

and the bound of h in (A4.1) gives that, for $z \in \mathbb{R}^{n+d}$,

$$|\sigma\left(z, h(s, x_n(s), y_n(s), U(s))\right)| \ \leq \ 2\|z\| M_h.$$

Thus, for $t \in [0,T)$ a lebesgue point of $\xi^i(\cdot), \zeta(\cdot), \dot{x}(\cdot), \dot{y}(\cdot), \sigma(z, h(\cdot, x(\cdot), y(\cdot), U(\cdot)))$, and for $\tau \in (0, T-t)$, when integrating (.6) on $[t, t+\tau]$ and then taking the limit as $n \to \infty$, the Dominated Convergence Theorem yields that, $\forall z = (z_1, z_2) \in \mathbb{R}^n \times \mathbb{R}^l$,

$$\begin{aligned}&\int_t^{t+\tau} \langle z, (\dot{x}(s), \dot{y}(s))\rangle ds \\ &\quad \leq \ \int_t^{t+\tau} [\sigma(z, h(s, x(s), y(s), U(s))) - \langle z, (\sum_{i=1}^{\mathcal{R}} \xi^i(s) \nabla_x q_i(s, x(s)), \zeta(s) \nabla_y \varphi(s, y(s)))\rangle] ds.\end{aligned}$$

Dividing the last equation by τ and taking the limit when $\tau \to 0$, we get that, $\forall z \in \mathbb{R}^n \times \mathbb{R}^l$, and for t lebesgue point,

$$\langle z, (\dot{x}(t), \dot{y}(t))\rangle \leq \sigma(z, h(t, x(t), y(t), U(t))) - \langle z, (\sum_{i=1}^{\mathcal{R}} \xi^i(t) \nabla_x q_i(t, x(t)), \zeta(t) \nabla_y \varphi(t, y(t)))\rangle,$$

and hence this inequality is valid for $t \in [0,T]$ a.e. Therefore, by means of Filipov Selection Theorem, there exists $u \in \mathcal{U}$, such that $((x,y),u)$, and $(\xi^1, \cdots, \xi^{\mathcal{R}}, \zeta)$ satisfy system (.3). □

Remark .0.3. When $\bar{\delta} = \infty$, Lemma .0.2 remains valid with $(\bar{x}, \bar{y})$ and the assumptions involving them are now superfluous. In this case, recall that $(A3.1)$,$(A4.1)$, and $(A4.2)$ are replaced by $(A3.1)_G$, $(A4.1)_G$, and $(A4.2)_G$, respectively.