NONSMOOTH MEAN FIELD GAMES WITH STATE CONSTRAINTS[*]

Saeed Sadeghi Arjmand[1,2,**] and Guilherme Mazanti[2]

Abstract. In this paper, we consider a mean field game model inspired by crowd motion where agents aim to reach a closed set, called target set, in minimal time. Congestion phenomena are modeled through a constraint on the velocity of an agent that depends on the average density of agents around their position. The model is considered in the presence of state constraints: roughly speaking, these constraints may model walls, columns, fences, hedges, or other kinds of obstacles at the boundary of the domain which agents cannot cross. After providing a more detailed description of the model, the paper recalls some previous results on the existence of equilibria for such games and presents the main difficulties that arise due to the presence of state constraints. Our main contribution is to show that equilibria of the game satisfy a system of coupled partial differential equations, known mean field game system, thanks to recent techniques to characterize optimal controls in the presence of state constraints. These techniques not only allow to deal with state constraints but also require very few regularity assumptions on the dynamics of the agents.

1. Introduction

Mean field games (MFGs for short) were first introduced around 2006 by Lasry and Lions [46–48] and independently by Caines, Huang, and Malhamé [38–40], motivated by problems in economics and engineering and based on some previous works on games with infinitely many players, such as those from [4, 5, 43]. MFGs are differential games with a continuum of players, assumed to be rational, indistinguishable, individually negligible, and influenced only by some "average" behavior of other players through a "mean-field type" interaction. Since their introduction, MFGs have been studied both in connection with several applications and from a theoretical point of view, in which the main goals are typically proving the existence of equilibria, characterizing such equilibria as solutions to a system of partial differential equations, called *MFG system*, or studying the connections between MFGs and games with a large (but finite) number of symmetric players. We refer to [1, 18, 20, 22, 23, 33] for more details and further references on mean field games. In this paper, we use the words "players" and "agents" interchangeably to refer to people taking part in the game.

[1] CMLS, École Polytechnique, CNRS, Université Paris-Saclay, 91128, Palaiseau, France.
[2] Université Paris-Saclay, CNRS, CentraleSupélec, Inria, Laboratoire des signaux et systèmes, 91190 Gif-sur-Yvette, France.
[**] Corresponding author: saeed.sadeghi-arjmand@polytechnique.edu

Most works on mean field games consider either first- or second-order mean field games. First-order MFGs, also known as deterministic mean field games, usually assume that players' dynamics are described by a deterministic control system, and their equilibria are characterized by a system of first-order partial differential equations, whereas second-order MFGs, also known as probabilistic mean field games, consider that players' dynamics are determined by a stochastic control system, typically with an additive Brownian motion modeling a random drift, and their equilibria are typically described by a system of second-order partial differential equations.

This paper considers a class of first-order mean field games inspired by crowd motion in which agents are assumed to remain inside of a specific domain while their goal is to arrive at a given target set in minimal time. In order to model congestion, the speed of each agent is constrained by a function depending on the position of the agent and on the distribution of all agents.

Crowd motion has been the subject of a very large number of works in the literature from different points of view, motivated not only by understanding but also by controlling and optimizing the crowd behavior (see, *e.g.*, [27, 32, 35–37, 41, 42, 49, 51–53]). The natural framework for a mean field game modeling crowd motion is to adopt a *macroscopic* modeling of crowd, *i.e.*, to describe the crowd at a given time t as infinitely many agents represented by a measure m_t on the space of possible positions, which evolves according to some conservation law, typically a continuity equation of the form $\partial_t m + \mathrm{div}(mV) = 0$, where V is the velocity field followed by the agents. While most macroscopic crowd motion models consider a given velocity field V constructed from modeling assumptions, the mean field game approach consists instead in considering that each agent will choose their trajectory by solving some optimal control problem, and the velocity field V is a consequence of the optimal choices of the agents.

Up to the authors' knowledge, the first work to be fully dedicated to a mean field game model for crowd motion is [44], which proposes an MFG model for a two-population crowd with trajectories perturbed by additive Brownian motion and considers both their stationary distributions and their evolution on a prescribed time interval. Other works have later proposed MFG models for crowd motion taking into account different characteristics, such as [10], which considers the fast exit of a crowd and proposes a mean field game model which is studied numerically; [21], which is not originally motivated by the modeling of crowd motion but considers a MFG model with a density constraint, which is a natural assumption in some crowd motion models; [8], which presents numerical simulations for some variational mean field games related to crowd motion; or also [26], which provides a generalized MFG model for pedestrians with limited predictive abilities.

The present paper considers the MFG model inspired by crowd motion introduced in [50] (described in details in Sect. 2), which contains two major differences with respect to previous MFGs for crowd motions and also to most MFG models in general. Firstly, the model from [50] assumes that each agent solves an optimal control problem with free final time, and actually that the optimization criterion of each agent is to minimize their arrival time at a certain target set, after which they quit the game. This is in contrast with most of the MFG literature, which usually considers either optimization criteria with a given and known finite final time or infinite-horizon optimization criteria. Secondly, congestion is modeled in [50] by imposing a maximal speed for each agent which depends on their position and the distribution of other agents, while, in most of other MFG works, agents are allowed to choose their speed without constraints, but high speeds and congestion are instead penalized in the cost function. The motivation of [50] to impose a constraint on the speed of agents is to model high-congestion situations in which an agent may be unable to move faster since other agents in front of them may work as a physical barrier which cannot be crossed by simply paying a higher "cost".

The main results of [50] are the existence of equilibria of the proposed MFG model and the characterization of equilibria through an MFG system. Even though existence of equilibria is proved in [50] under rather general assumptions, their characterization through an MFG system is only shown in the case where the target set of the agents is the whole boundary of the compact domain in which they evolve, which avoids the presence of state constraints in the minimal-time optimal control problem solved by each agent. This is a very restrictive assumption for a MFG model for crowd motion, since crowds often evolve in domains with boundaries which cannot be crossed by the agents, such as walls, columns, fences, hedges, or other obstacles. However, from a technical point of view, the major difficulty in analyzing optimal control problems with state constraints is that

their value functions may fail to be semiconcave (see, *e.g.*, [14], Ex. 4.4), the latter property being important in the characterization of optimal controls (see, *e.g.*, [17]), which is a key step in obtaining the MFG system in [50]. Other works, such as [29, 30, 55], have further explored the model from [50] and related models, but none of those works consider the case of MFGs with state constraints.

Studying optimal control problems and mean field games with state constraints turns out to be a challenging problem due to the possible lack of semiconcavity of the value function. The series of papers [11–13] represent an important step in the study of MFGs with state constraints: in those works, the authors prove that, for the optimal control problem they consider, even though the corresponding value function may fail to be semiconcave in the classical sense, *i.e.*, with a *linear* modulus of semiconcavity, the value function is still semiconcave with a *fractional* modulus of semiconcavity, which is sufficient to obtain additional properties of optimal trajectories allowing to characterize optimal controls. The fractional semiconcavity proved in [11] is global in the space variable, which is known not to be the case for some minimal-time optimal control problems. This paper follows a different strategy and proves the required additional properties of optimal trajectories without relying on the semiconcavity of the value function, which has also the additional advantage of requiring fewer regularity properties on the dynamics of the control system satisfied by each agent.

The main contribution of the present paper is thus to characterize optimal controls and deduce the MFG system for the model of [50] under state constraints, which is an important step towards the study of MFG models for crowd motion with more realistic assumptions. A first step in the characterization of optimal controls is to apply Pontryagin Maximum Principle in order to deduce additional information on optimal trajectories. To do so, we adapt the penalization technique used in [15]: we transform the minimal-time optimal control problem with state constraints into a penalized optimal control problem without state constraints and show that, similarly to [15], if the penalization parameter is small enough, optimal trajectories of the penalized problem coincide with optimal trajectories of the original problem (see Thm. 4.8), which allows us to deduce properties of the optimal trajectories of the original optimal control problem by applying Pontryagin Maximum Principle to the penalized problem (see Cor. 4.9).

With those properties of optimal trajectories, we then proceed to the study of optimal controls by following the arguments used in the recent paper [55]: our main results are Theorem 4.11, which provides a boundary condition for the value function of the optimal control problem solved by each agent; Theorem 4.23, which characterizes the optimal control at a given point as the direction of greatest decrease of the value function by adapting the arguments of Theorem 4.14 in [55] to the case with state constraints; and Theorem 5.2, which shows that equilibria of our MFG model satisfy the MFG system.

The present paper is an extended version of [54], which announced most of the results from this paper and provided some details of the strategy of the proofs of the main results. With respect to that reference, this paper works under weaker assumptions on the dynamics of the control system (compare (H4)–(H7) below with (H2)–(H3) of [54]), which renders some proofs more technically involved and require different proof strategies at some points, such as in Lemma 3.2 and for most of the results of Section 4.1. This paper also provides a more detailed introduction to the subject, details proofs that had been omitted or only sketched in [54] (and in particular the results in Sect. 4.3), and provides additional remarks and comments on the results.

Our paper is organized as follows. Some notations and standard definitions are provided in Section 2, together with the precise description of the optimal control problem and the mean field game model considered in this paper, and the list of hypotheses used here. Section 3 presents preliminary results on the optimal control problem and the mean field game, most of which are either easy to prove or already present in the literature. The major new contributions in Section 3 are the proof of Lipschitz continuity of the value function of the optimal control problem under the weaker assumptions (H4) and (H5) as well as the alternative characterization of equilibria from Proposition 3.7, which also holds for other mean field game models (see, *e.g.*, [57], Lem. 3.4).

The main results of this paper are provided in Sections 4 and 5. Section 4.1 uses the strategy of [15] to study an optimal control problem with state constraints through a penalized optimal control problem without state constraints, adapting the techniques of that reference to our setting in order to show that optimal trajectories of the original problem coincide with that of the penalized problem if the penalization parameter is small enough (Thm. 4.8). Section 4.2 provides a boundary condition for the value function of the optimal control problem in

the part of the boundary which is not in the target set. We then provide, in Section 4.3, the characterization of optimal controls as directions of maximal descent of the value function, and we use this characterization in Section 4.4 to introduce the notion of normalized gradient and provide its main properties. The main result concerning the mean field game model of interest is the fact that its equilibria satisfy a system of PDEs, shown in Section 5.

2. NOTATIONS AND DEFINITIONS

2.1. General notations

In this paper, $\mathbb{N}$ denotes the set of positive integers, d is a fixed positive integer, the sets of nonnegative and positive real numbers are denoted respectively by $\mathbb{R}_+$ and $\mathbb{R}_+^*$, $\mathbb{R}^d$ is endowed with the usual Euclidean norm $|\cdot|$, and the unit sphere is denoted by $\mathbb{S}^{d-1}$. For $A \subset \mathbb{R}^d$, $\bar{A}$ denotes its closure, ∂A denotes its boundary, and $\operatorname{co} A$ denotes its convex hull. For any $x \in \mathbb{R}^d$ and $r \geq 0$, $B(x, r)$ (resp., $\bar{B}(x, r)$) denotes[1] the open (resp., closed) ball centered at x and with radius r. We denote this ball simply by B_r (resp., $\bar{B}_r$) if $x = 0$ and by B (resp., $\bar{B}$) if $x = 0$ and $r = 1$, i.e., B (resp., $\bar{B}$) is the open (resp., closed) unit ball centered at the origin.

For two sets A and B, a set-valued map from A to B is a map F that associates, with each $x \in A$, a (possibly empty) set $F(x) \subset B$. We use the notation $F : A \rightrightarrows B$ to indicate that F is a set-valued map from A to B.

Given two metric spaces X and Y and a constant $M > 0$, $\mathbf{C}(X; Y)$, $\operatorname{Lip}(X; Y)$, and $\operatorname{Lip}_M(X; Y)$ denote, respectively, the set of all continuous functions from X to Y, the set of all Lipschitz continuous functions from X to Y, and the subset of $\operatorname{Lip}(X; Y)$ containing only those functions whose Lipschitz constant is at most M. We will often use these spaces to represent time-dependent functions defined for nonnegative times, in which case $X = \mathbb{R}_+$ and, for simplicity, we omit X from the previous notations, writing simply $\mathbf{C}(Y)$, $\operatorname{Lip}(Y)$, and $\operatorname{Lip}_M(Y)$, respectively. When $X \subset \mathbb{R}^k$ and $Y \subset \mathbb{R}^m$ for some positive integers k and m, we also consider the set $\mathbf{C}^1(X; Y)$ of continuously differentiable functions, the set $\mathbf{C}^{1,1}(X; Y)$ of differentiable functions with a Lipschitz continuous differential, and the set $\mathbf{C}_c^\infty(X; Y)$ of infinitely differentiable functions with compact support on X. If $f : X \to Y$ and $Z \subset X$, we write $f \in \mathbf{C}(Z; Y)$ to denote that the restriction of f to Z is continuous (even though f itself may fail to be continuous in the boundary of Z), with similar notations for the other functional spaces defined above.

For compact $A \subset \mathbb{R}^d$, the space $\mathbf{C}(A)$ is assumed to be endowed with the topology of uniform convergence on compact sets, with respect to which $\mathbf{C}(A)$ is a Polish space (see, e.g., [9], Chap. X). For $t \in \mathbb{R}_+$, we denote by $e_t : \mathbf{C}(A) \to A$ the evaluation map at time t, defined by $e_t(\gamma) = \gamma(t)$ for every $\gamma \in \mathbf{C}(A)$.

Recall that, for two metric spaces X and Y endowed with their Borel σ-algebras and a Borel-measurable map $f : X \to Y$, the pushforward of a measure μ on X through f is the measure $f_\# \mu$ on Y defined by $f_\# \mu(B) = \mu(f^{-1}(B))$ for every Borel subset B of Y. Given a Polish space X, the set of all Borel probability measures on X is denoted by $\mathcal{P}(X)$ and is endowed with the topology of weak convergence of measures (which is also sometimes referred to as narrow convergence). The support of a measure $\mu \in \mathcal{P}(X)$ is denoted by $\operatorname{supp}(\mu)$, and is defined as the set of all points $x \in X$ such that $\mu(N_x) > 0$ for every open neighborhood N_x of x. When X is endowed with a complete metric $\mathbf{d}$ with respect to which X is bounded, we endow $\mathcal{P}(X)$ with the Wasserstein distance $\mathbf{W}_1$, defined for $\mu, \nu \in \mathcal{P}(X)$ by

$$\mathbf{W}_1(\mu, \nu) = \inf \left\{ \int_{X \times X} \mathbf{d}(x, y) \, \mathrm{d}\lambda(x, y) \ \middle| \ \lambda \in \Pi(\mu, \nu) \right\},$$

where $\Pi(\mu, \nu) = \{\lambda \in \mathcal{P}(X \times X) \mid \pi_{1\#}\lambda = \mu, \ \pi_{2\#}\lambda = \nu\}$ and $\pi_1, \pi_2 : X \times X \to X$ denote the canonical projections on to the first and second factors of the product $X \times X$, respectively. Recall (see, e.g., [2], Chap. 7 and

---

[1]In accordance with the previous notation, $\bar{B}(x, r)$ is the closure of $B(x, r)$ whenever $r > 0$. However, our notation is slightly ambiguous for $r = 0$: $B(x, 0) = \varnothing$, but $\bar{B}(x, 0) = \{x\}$ is not the closure of $B(x, 0)$. All along the paper, we use the convention that, when referring to balls, the notation $\bar{B}(x, r)$ is the closed ball centered at x and with radius r. The same convention applies to the notation $\bar{B}_r$, and we remark that $\bar{B}$ is unambiguous.

[56], Chap. 5) that $\mathbf{W}_1$ is compatible with the topology of weak convergence in $\mathcal{P}(X)$ and that it admits the dual formulation

$$\mathbf{W}_1(\mu, \nu) = \sup \left\{ \int_X \Phi(x) \, \mathrm{d}(\mu - \nu) \;\middle|\; \Phi \in \mathrm{Lip}_1(X; \mathbb{R}) \right\}.$$

We shall also need in this paper the notion of signed distance to the boundary of a set, whose definition we now provide, together with the classical definition of distance from a point to a set.

Definition 2.1. Let $A \subset \mathbb{R}^d$. If $A \neq \varnothing$, we denote by d_A the Euclidean distance to A, defined by $d_A(x) = \inf_{y \in A} |x - y|$. If $A \neq \varnothing$ and $A \neq \mathbb{R}^d$, the signed distance to the boundary of A, $d_{\partial A}^{\pm} : \mathbb{R}^d \to \mathbb{R}$, is defined by

$$d_{\partial A}^{\pm}(x) = d_A(x) - d_{\mathbb{R}^d \setminus A}(x).$$

Finally, we will need, in Section 4, the notion of normal cones of nonsmooth sets. We recall here the definitions provided in Section 4.2 of [58].

Definition 2.2. Let k be a positive integer, $C \subset \mathbb{R}^k$ be a nonempty closed set, and $x \in C$.

(a) The *proximal normal cone* to C at x is the set $N_C^{\mathrm{P}}(x)$ defined by

$$N_C^{\mathrm{P}}(x) = \{ p \in \mathbb{R}^k \mid \exists M > 0 \text{ such that } p \cdot (y - x) \leq M |y - x|^2 \text{ for every } y \in C \}.$$

(b) The *limiting normal cone* to C at x is the set $N_C(x)$ of vectors $p \in \mathbb{R}^k$ for which there exist a sequence $(x_n)_{n \in \mathbb{N}}$ in C and a sequence $(p_n)_{n \in \mathbb{N}}$ in $\mathbb{R}^k$ such that $x_n \to x$ and $p_n \to p$ as $n \to +\infty$ and $p_n \in N_C^{\mathrm{P}}(x_n)$ for every $n \in \mathbb{N}$.

2.2. The minimal-time optimal control problem

Before introducing the minimal-time mean field game of interest to this paper, let us first consider an auxiliary minimal-time optimal control problem, which is a classical kind of problem in control theory. Let $\Omega \subset \mathbb{R}^d$ be a nonempty open bounded set and $\Gamma \subset \bar{\Omega}$ be a nonempty closed set. We consider the control system

$$\begin{cases} \dot{\gamma}(t) = k(t, \gamma(t)) u(t), & u(t) \in \bar{B}, \\ \gamma(0) = x, \end{cases} \tag{2.1}$$

where $x \in \bar{\Omega}$, $k : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$, and, for $t \geq 0$, $\gamma(t) \in \bar{\Omega}$ is the state and $u(t) \in \bar{B}$ is the control. The function k describes the dynamics of the system and, due to the constraint that $u(t) \in \bar{B}$, $k(t, x)$ can be interpreted as the maximal speed at which an agent at position x at time t can move. The fact of requiring that $\gamma(t) \in \bar{\Omega}$ for every time $t \geq 0$ is a constraint we impose in the state of (2.1).

We are interested in this paper in the optimal control problem consisting in minimizing the time a trajectory of (2.1) takes to reach the set Γ, called the *target set*. This optimal control problem, to which we refer in the sequel as $\mathrm{OCP}(k)$ making explicit its dependence on the dynamics k, is made more precise in the following definition.

Definition 2.3. Let $\Omega \subset \mathbb{R}^d$ be a nonempty open bounded set, $\Gamma \subset \bar{\Omega}$ be a nonempty closed set, and $k : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$.

(a) A curve $\gamma \in \mathrm{Lip}(\bar{\Omega})$ is said to be *admissible* for $\mathrm{OCP}(k)$ if there exists a measurable function $u : \mathbb{R}_+ \to \bar{B}$, called the *control* associated with γ, such that the first equation of (2.1) is satisfied for almost every $t \in \mathbb{R}_+$. The set of all admissible curves for $\mathrm{OCP}(k)$ is denoted by $\mathrm{Adm}(k)$.

(b) Let $t_0 \in \mathbb{R}_+$. The *first exit time* after t_0 of a curve $\gamma \in \mathrm{Lip}(\bar{\Omega})$ is the number $\tau(t_0, \gamma)$ defined by

$$\tau(t_0, \gamma) = \inf\{t \geq 0 \mid \gamma(t + t_0) \in \Gamma\},$$

with the convention that $\tau(t_0, \gamma) = +\infty$ if $\gamma(t + t_0) \notin \Gamma$ for every $t \geq 0$.

(c) Let $t_0 \in \mathbb{R}_+$ and $x_0 \in \bar{\Omega}$. A curve $\gamma \in \mathrm{Lip}(\bar{\Omega})$ is said to be an *optimal trajectory* for OCP(k) from (t_0, x_0) if $\gamma \in \mathrm{Adm}(k)$, $\gamma(t) = x_0$ for every $t \in [0, t_0]$,

$$\tau(t_0, \gamma) = \inf_{\substack{\beta \in \mathrm{Adm}(k) \\ \beta(t_0) = x_0}} \tau(t_0, \beta), \tag{2.2}$$

and $\gamma(t) = \gamma(t_0 + \tau(t_0, \gamma))$ for every $t \in [t_0 + \tau(t_0, \gamma), +\infty)$. In this case, the associated control u of γ is called an *optimal control* associated with γ. The set of all optimal trajectories for OCP(k) from (t_0, x_0) is denoted by $\mathrm{Opt}(k, t_0, x_0)$.

(d) The *value function* of the optimal control problem OCP(k) is the function $\varphi : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$ defined for $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$ by

$$\varphi(t_0, x_0) = \inf_{\substack{\gamma \in \mathrm{Adm}(k) \\ \gamma(t_0) = x_0}} \tau(t_0, \gamma). \tag{2.3}$$

We highlight the fact that, for a curve γ to be admissible, it must not only satisfy the first equation of (2.1) for some control u and almost every time $t \geq 0$, but it must also remain inside $\bar{\Omega}$ for all times.

Remark 2.4. Since we are considering in Definition 2.3(c) optimal trajectories γ starting at time t_0 from x_0 and minimizing the time to reach Γ, these trajectories could have been defined only in the interval $[t_0, t_0 + \tau(t_0, \gamma)]$. However, in order to simplify the notations and deal in this paper only with trajectories and controls defined on $\mathbb{R}_+$, we extend an optimal trajectory to $\mathbb{R}_+$ by requiring it to be constant on $[0, t_0]$ and on $[t_0 + \tau(t_0, \gamma), +\infty)$, as done in Definition 2.3(c).

2.3. The minimal-time mean field game and its equilibria

As in Section 2.2, we fix a nonempty open bounded set $\Omega \subset \mathbb{R}^d$ and a nonempty closed set $\Gamma \subset \bar{\Omega}$. We consider a minimal-time mean field game in which a population of agents evolves on $\bar{\Omega}$ and the goal of each agent is to reach the target set Γ in minimal time. The population is described by a time-dependent probability measure $m_t \in \mathcal{P}(\bar{\Omega})$ for $t \geq 0$, and m_0 is assumed to be known. The trajectory γ of an agent starting its movement at a position $x \in \bar{\Omega}$ is assumed to satisfy the control system

$$\begin{cases} \dot{\gamma}(t) = K(m_t, \gamma(t))u(t), & u(t) \in \bar{B}, \\ \gamma(0) = x, \end{cases} \tag{2.4}$$

where $K : \mathcal{P}(\bar{\Omega}) \times \bar{\Omega} \to \mathbb{R}_+$ and $u : \mathbb{R}_+ \to \bar{B}$ is the control of the agent. Note that (2.4) corresponds to (2.1) with k defined by $k(t, x) = K(m_t, x)$ for $t \geq 0$ and $x \in \bar{\Omega}$. Each agent is assumed to choose their control in order to solve the optimal control problem OCP(k) with k defined from K as before. This mean field game is denoted in the sequel by MFG(K).

The function K models interactions between agents and states that the maximal speed at which an agent at position x at time t can move depends on the position x itself and on the distribution of all agents at time t, m_t. This function can be used to model congestion phenomena in crowd motion by choosing $K(m_t, x)$ to be small when m_t is "large" around x, which means that it is harder for agents to move on more crowded regions.

For instance, K can be chosen as

$$K(\mu, x) = g\left(\int_{\bar{\Omega}} \chi(x-y)\eta(y)\,\mathrm{d}\mu(y)\right), \tag{2.5}$$

where $\chi : \mathbb{R}^d \to \mathbb{R}_+$ is a convolution kernel, $\eta : \bar{\Omega} \to \mathbb{R}_+$ is a weight function in the space $\bar{\Omega}$, and $g : \mathbb{R}_+ \to \mathbb{R}_+$ is a decreasing function. The function χ may represent, for instance, the region around an agent at which they look in order to evaluate local congestion, while η may be a function that is larger in regions difficult to move, such as regions with obstacles or other kinds of difficult terrain, and which can also be used to discount people who already reached the target set, as it was done in [30, 50]. Note, however, that we do *not* assume this specific form for K in this paper.

Note that, since K depends on m_t for all $t \geq 0$, the optimal trajectories taken by the agents depend on m_t. On the other hand, m_t itself describes the evolution of the agents, and hence is determined by their choices of trajectories. We are interested in this paper in *equilibrium* situations, in which, roughly speaking, starting from time evolution of the distribution of agents $m : \mathbb{R}_+ \to \mathcal{P}(\bar{\Omega})$, the optimal trajectories chosen by agents induce an evolution of the initial distribution of agents m_0 that is precisely given by $t \mapsto m_t$.

We provide a more precise notion of equilibrium in what is known as the Lagrangian framework, in which, instead of describing the evolution of agents as a time-dependent measure $m : \mathbb{R}_+ \to \mathcal{P}(\bar{\Omega})$, we rely instead on a measure Q on the set of all possible trajectories $\mathcal{P}(\mathbf{C}(\bar{\Omega}))$. Note that, given a measure $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$, one can obtain the associated time-dependent measure m by setting $m_t = e_{t\#}Q$ for $t \geq 0$. The Lagrangian approach is a classical approach in optimal transport problems (see, *e.g.*, [2, 56]) which has been used to define equilibria of first-order mean field games in some recent works, such as [8, 11, 16, 19, 21, 30, 31, 50, 55].

Definition 2.5. Let $\Omega \subset \mathbb{R}^d$ be a nonempty open bounded set, $\Gamma \subset \bar{\Omega}$ be a nonempty closed set, $K : \mathcal{P}(\bar{\Omega}) \times \bar{\Omega} \to \mathbb{R}_+$, and $m_0 \in \mathcal{P}(\bar{\Omega})$. A measure $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$ is called a *Lagrangian equilibrium* of MFG(K) with initial condition m_0 if $e_{0\#}Q = m_0$ and Q-almost every $\gamma \in \mathbf{C}(\bar{\Omega})$ satisfies $\gamma \in \mathrm{Opt}(k_Q, 0, \gamma(0))$, where $k_Q : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$ is defined for $t \geq 0$ and $x \in \bar{\Omega}$ by $k_Q(t, x) = K(e_{t\#}Q, x)$.

In the sequel of the paper, we refer to Lagrangian equilibria simply as equilibria.

Remark 2.6. Another classical way to describe the evolution of agents in a mean field game, which dates back to [39] and has been used and developed in several other references such as [24, 34, 45], is to fix a probability space $(\mathsf{Z}, \mathcal{Z}, \mathbb{P})$ and describe the motion of agents through a time-dependent random variable in Z, *i.e.*, through a function X defined on $\mathbb{R}_+$ and such that, for every $t \geq 0$, $X(t) : \mathsf{Z} \to \bar{\Omega}$ is measurable. One can then retrieve the time-dependent measure $m_t \in \mathcal{P}(\bar{\Omega})$ as the law of $X(t)$, *i.e.*, $m_t = X(t)_{\#}\mathbb{P}$, but X contains more information than m, since X also carries information on the correlation of the distributions of agents at different times, for instance. This formulation using random variables has the additional advantage of being also adapted to study of other kinds of mean field games, such as second-order mean field games or mean field games of controls.

In addition, if[2] $\mathbb{R}_+ \ni t \mapsto X(t, z) \in \bar{\Omega}$ is continuous for almost every $z \in \mathsf{Z}$ and the function $\Xi : \mathsf{Z} \to \mathbf{C}(\bar{\Omega})$ defined by setting $\Xi(z) = X(\cdot, z)$ for a.e. $z \in \mathsf{Z}$ is measurable, then the probability measure $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$ describing the evolution of agents from a Lagrangian perspective can be retrieved as $Q = \Xi_{\#}\mathbb{P}$.

Conversely, given a measure $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$ describing the evolution of agents, one can consider the probability space $(\mathbf{C}(\bar{\Omega}), \mathcal{B}, Q)$, where $\mathcal{B}$ is the Borel σ-algebra of $\mathbf{C}(\bar{\Omega})$, and in this case the evolution of agents can be described by the time-dependent random variable X defined by $X(t) = e_t$ for $t \geq 0$.

2.4. Hypotheses

Along the paper, we will need several assumptions on Ω, Γ, K, and k, which we collect in this subsection. We start with the following assumptions on the sets Ω and Γ.

---

[2]Here, $X(t, z)$ is a simplified notation for $X(t)(z)$.

(H1) The set $\Omega \subset \mathbb{R}^d$ is nonempty, open, bounded, and connected.

(H2) The set $\Gamma \subset \bar{\Omega}$ is nonempty and closed.

(H3) The boundary $\partial\Omega$ is a compact $\mathbf{C}^{1,1}$ submanifold of $\mathbb{R}^d$ of dimension $d-1$.

Whenever we assume that (H3) holds, we use $\mathbf{n}(x)$ to denote the outward unit normal vector to $\bar{\Omega}$ at the point $x \in \partial\Omega$.

When studying the optimal control problem $\mathrm{OCP}(k)$, we shall need the following assumptions on the function k.

(H4) The function $k : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$ is continuous and there exist positive constants $K_{\min}$ and $K_{\max}$ such that $k(t,x) \in [K_{\min}, K_{\max}]$ for every $(t,x) \in \mathbb{R}_+ \times \bar{\Omega}$.

(H5) The function k is Lipschitz continuous with respect to its second variable, uniformly with respect to the first variable, *i.e.*, there exists $L > 0$ such that, for every $t \in \mathbb{R}_+$ and $x_1, x_2 \in \bar{\Omega}$, we have

$$|k(t,x_1) - k(t,x_2)| \le L|x_1 - x_2|.$$

The counterpart of Hypotheses (H4) and (H5) concerning the function K from the mean field game $\mathrm{MFG}(K)$ are the following.

(H6) The function $K : \mathcal{P}(\bar{\Omega}) \times \bar{\Omega} \to \mathbb{R}_+$ is continuous and there exist positive constants $K_{\min}$ and $K_{\max}$ such that $K(\mu, x) \in [K_{\min}, K_{\max}]$ for every $(\mu, x) \in \mathcal{P}(\bar{\Omega}) \times \bar{\Omega}$.

(H7) The functions K is Lipschitz continuous with respect to its second variable, uniformly with respect to the first variable, *i.e.*, there exists $L > 0$ such that, for every $\mu \in \mathcal{P}(\bar{\Omega})$ and $x_1, x_2 \in \bar{\Omega}$, we have

$$|K(\mu, x_1) - K(\mu, x_2)| \le L|x_1 - x_2|.$$

Note that (H6) and (H7) are satisfied in the particular case where K is chosen as in (2.5) if g is Lipschitz continuous, upper bounded, and lower bounded by a positive constant, χ is Lipschitz continuous, and η is continuous.

3. PRELIMINARY RESULTS

This section presents results on $\mathrm{OCP}(k)$ and $\mathrm{MFG}(K)$ that will be useful in the sequel, most of which are either present in other references or easy to prove using classical techniques. It turns out that the results we will present in this section require fewer regularity assumptions on Ω than (H1) and (H3); namely, we replace here (H3) by the following assumption.

(H3′) There exists $D > 0$ such that, for every $x, y \in \bar{\Omega}$, there exists a curve γ included in $\bar{\Omega}$ connecting x to y and of length at most $D|x-y|$.

Hypothesis (H3′) means that the geodesic distance in $\bar{\Omega}$ is equivalent to the usual Euclidean distance, and it holds in particular when (H1) and (H3) are satisfied.

The first result we present concerns three elementary properties of $\mathrm{OCP}(k)$: existence of optimal trajectories, boundedness of the value function, and the dynamic programming principle. The proof is omitted, since all properties are either easy to prove or classical. Indeed, existence of optimal trajectories can be proved easily using compactness of a minimizing sequence (see Thm. 8.1.4 of [17] for a proof in the autonomous case, *i.e.*, when $k : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$ does not depend on its first variable). Boundedness of the value function follows from the fact that, using (H3′), one can easily construct, for every point $x \in \bar{\Omega}$, an admissible trajectory connecting it to a point in Γ with constant speed $K_{\min}$ and that arrives in Γ in time at most $\frac{D d_\Gamma(x)}{K_{\min}}$, and the continuous function d_Γ is bounded in the compact set $\bar{\Omega}$. Finally, the proof of the dynamic programming principle is classical: the autonomous case can be found, for instance, in Proposition 2.1 of [7] and (8.4) of [17], and the corresponding proofs can be easily adapted to our nonautonomous setting.

Proposition 3.1. *Consider* $\mathrm{OCP}(k)$ *under hypotheses* (H1), (H2), (H3′), (H4), *and* (H5).

(a) *For every* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$, *there exists an optimal trajectory* γ *for* $\mathrm{OCP}(k)$ *from* (t_0, x_0).

(b) *There exists* $T > 0$ *such that, for every* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$, *the value function* φ *satisfies* $\varphi(t_0, x_0) \leq T$.

(c) *For every* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$ *and* $\gamma \in \mathrm{Adm}(k)$ *such that* $\gamma(t_0) = x_0$, *we have, for every* $h \geq 0$,

$$\varphi(t_0 + h, \gamma(t_0 + h)) + h \geq \varphi(t_0, x_0), \tag{3.1}$$

with equality if $\gamma \in \mathrm{Opt}(k, t_0, x_0)$ *and* $h \in [0, \tau(t_0, \gamma)]$. *Conversely, if* $\gamma \in \mathrm{Adm}(k)$ *satisfies* $\gamma(t_0) = x_0$, γ *is constant on* $[0, t_0]$ *and on* $[t_0 + \tau(t_0, \gamma), +\infty)$, *and equality holds in* (3.1) *for every* $h \in [0, \tau(t_0, \gamma)]$, *then* $\gamma \in \mathrm{Opt}(k, t_0, x_0)$.

Our next result deals with Lipschitz continuity of the value function φ. This kind of result is classical for optimal control problems with free final time (see, *e.g.*, Prop. 8.2.5 of [17] for a proof in the autonomous case), and a complete proof for the nonautonomous optimal control problem $\mathrm{OCP}(k)$ was given in Propositions 4.2 and 4.3 of [50]. However, that reference uses the stronger assumption that $k \in \mathrm{Lip}(\mathbb{R}_+ \times \bar{\Omega}; \mathbb{R}_+)$. When the optimal control problem does not have state constraints, this assumption can be relaxed to (H5) by first showing Lipschitz continuity of φ with respect to x, which can be done by adapting the classical proof of Proposition 8.2.5 in [17], and then using the dynamic programming principle to deduce Lipschitz continuity also with respect to t. This strategy was described in Proposition 3.8 of [30] and carried out in details in Lemma 4.7 and Proposition 4.8 of [55], however those proofs rely on the absence of state constraints and cannot be easily generalized to optimal control problems with state constraints. For that reason, we present here a new proof, which is inspired by that of Propositions 4.2 and 4.3 of [50] but uses a technique introduced in the proof of Proposition 3.9 in [30] in order to replace the assumption $k \in \mathrm{Lip}(\mathbb{R}_+ \times \Omega; \mathbb{R}_+)$ by the weaker assumption (H5). As a first step, we prove Lipschitz continuity of φ in space for fixed time.

Lemma 3.2. *Consider the optimal control problem* $\mathrm{OCP}(k)$ *and its value function* φ *and assume that* (H1), (H2), (H3′), (H4), *and* (H5) *are satisfied. Then there exists* $C > 0$ *such that, for every* $t_0 \in \mathbb{R}_+$ *and* $x_0, x_1 \in \bar{\Omega}$, *we have*

$$|\varphi(t_0, x_0) - \varphi(t_0, x_1)| \leq C|x_0 - x_1|. \tag{3.2}$$

Proof. Let $D > 0$ be as in (H3′), $L > 0$ be as in (H5), and $T > 0$ be as in the statement of Proposition 3.1(b). It suffices to show that there exists $C > 0$ such that, for every $t_0 \in \mathbb{R}_+$ and $x_0, x_1 \in \bar{\Omega}$, we have

$$\varphi(t_0, x_1) - \varphi(t_0, x_0) \leq C|x_0 - x_1|, \tag{3.3}$$

since, in this case, (3.2) can be deduced by exchanging the role of x_0 and x_1.

Let $t_0 \in \mathbb{R}_+$ and $x_0, x_1 \in \bar{\Omega}$. Let $\gamma_0 \in \mathrm{Opt}(k, t_0, x_0)$ and denote by u_0 the corresponding optimal control, *i.e.*, $\dot{\gamma}_0(t) = k(t, \gamma_0(t)) u_0(t)$ for a.e. $t \in \mathbb{R}_+$. Let $t_0^* = t_0 + \varphi(t_0, x_0)$ be the time at which γ_0 arrives at the target set Γ.

Let us first describe informally the idea of the proof. We will construct an admissible trajectory $\gamma_1 \in \mathrm{Lip}(\bar{\Omega})$ which remains constant at x_1 in the time interval $[0, t_0]$, then moves from x_1 to x_0 in a time interval $[t_0, t_1]$, and then follows the same path of γ_0, but with a change in the time scale since it starts from x_0 at time $t_1 > t_0$. This trajectory will then arrive at $x_0^* \in \Gamma$, and we will prove that its arrival time at Γ satisfies

$$\tau(t_0, \gamma_1) \leq \varphi(t_0, x_0) + C|x_0 - x_1|, \tag{3.4}$$

which yields the conclusion since $\varphi(t_0, x_1) \leq \tau(t_0, \gamma_1)$. The difficult part of the proof is to perform a suitable change in time scale guaranteeing both that γ_1 is admissible and that its arrival time at Γ satisfies the above inequality.

Applying (H3′) and renormalizing the speed of the curve whose existence is asserted in that hypothesis, we obtain the existence of $t_1 \geq t_0$ and a Lipschitz continuous curve $\sigma : [t_0, t_1] \to \bar{\Omega}$ such that $\sigma(t_0) = x_1$, $\sigma(t_1) = x_0$, $|\dot{\sigma}(t)| = K_{\min}$ for almost every $t \in [t_0, t_1]$, and $t_1 - t_0 \leq \frac{D|x_1 - x_0|}{K_{\min}}$.

Let us now define $\phi : [t_1, +\infty) \to [t_0, +\infty)$ as a solution of the problem

$$\begin{cases} \dot{\phi}(t) = \dfrac{k(t, \gamma_0(\phi(t)))}{k(\phi(t), \gamma_0(\phi(t)))} & \text{for } t \geq t_1, \\ \phi(t_1) = t_0. \end{cases}$$

Note that, since $(t, s) \mapsto \frac{k(t, \gamma_0(s))}{k(s, \gamma_0(s))}$ is continuous (but not necessarily Lipschitz continuous in its second argument), a solution ϕ to the above problem exists and is of class $\mathbf{C}^1$ (but it may not be unique). Moreover, $\dot{\phi}(t) \in \left[\frac{K_{\min}}{K_{\max}}, \frac{K_{\max}}{K_{\min}} \right]$ for every $t \in [t_1, +\infty)$, which implies that $\phi : [t_1, +\infty) \to [t_0, +\infty)$ is increasing and surjective, and hence invertible, and both ϕ and ϕ^{-1} are Lispchitz continuous, with Lipschitz constant $\frac{K_{\max}}{K_{\min}}$. We define $\sigma_1 : [t_1, +\infty) \to \bar{\Omega}$ by $\sigma_1(t) = \gamma_0(\phi(t))$, which, by construction, satisfies $\sigma_1(t_1) = x_0$, $\sigma_1 \in \text{Lip}(\bar{\Omega})$, and $\dot{\sigma}_1(t) = k(t, \sigma_1(t))u_1(t)$ for a.e. $t \in [t_1, +\infty)$, where u_1 is defined for $t \in [t_1, +\infty)$ by $u_1(t) = u_0(\phi(t))$. In particular, since ϕ^{-1} is Lipschitz continuous, u_1 is measurable. We define $t_1^* = \phi^{-1}(t_0^*)$ and remark that $\sigma_1(t_1^*) = \gamma_0(t_0^*) = x_0^* \in \Gamma$.

Finally, we define $\gamma_1 \in \text{Lip}(\bar{\Omega})$ by $\gamma_1(t) = x_1$ for $t \in [0, t_0]$, $\gamma_1(t) = \sigma(t)$ for $t \in [t_0, t_1]$, and $\gamma_1(t) = \sigma_1(t)$ for $t \in [t_1, +\infty)$. By construction, we have $\gamma_1 \in \text{Adm}(k)$ and $\tau(t_0, \gamma_1) \leq t_1^* - t_0$. We are thus only left to show (3.4).

Note that, for every $t \in [t_1, +\infty)$, we have $k(\phi(t), \gamma_0(\phi(t)))\dot{\phi}(t) = k(t, \gamma_0(\phi(t)))$ and thus, by integrating this identity and performing a change of variables, we deduce that, for every $t \geq t_1$,

$$\int_{t_0}^{\phi(t)} k(s, \gamma_0(s)) \, ds = \int_{t_1}^{t} k(s, \gamma_0(\phi(s))) \, ds. \tag{3.5}$$

Let $G : [t_0, +\infty) \to [0, +\infty)$ be defined for $\theta \geq t_0$ by

$$G(\theta) = \int_{t_0}^{\theta} k(s, \gamma_0(s)) \, ds.$$

Then G is differentiable, with $\dot{G}(\theta) = k(\theta, \gamma_0(\theta)) \in [K_{\min}, K_{\max}]$ for every $\theta \geq t_0$. In particular, G is $K_{\max}$-Lipschitz continuous, invertible, and its inverse is $\frac{1}{K_{\min}}$-Lipschitz continuous. Moreover, using (3.5), we have, for every $t \geq t_1$, that

$$G(\phi(t)) = \int_{t_1}^{t} k(s, \gamma_0(\phi(s))) \, ds, \qquad G(t) = \int_{t_0}^{t} k(s, \gamma_0(s)) \, ds.$$

Hence, for every $t \geq t_1$, we have

$$\begin{aligned} |\phi(t) - t| &= \left| G^{-1} \left(\int_{t_1}^{t} k(s, \gamma_0(\phi(s))) \, ds \right) - G^{-1} \left(\int_{t_0}^{t} k(s, \gamma_0(s)) \, ds \right) \right| \\ &\leq \frac{1}{K_{\min}} \left| \int_{t_1}^{t} k(s, \gamma_0(\phi(s))) \, ds - \int_{t_0}^{t} k(s, \gamma_0(s)) \, ds \right| \\ &\leq \frac{1}{K_{\min}} \int_{t_1}^{t} |k(s, \gamma_0(\phi(s))) - k(s, \gamma_0(s))| \, ds + \frac{1}{K_{\min}} \int_{t_0}^{t_1} k(s, \gamma_0(s)) \, ds \\ &\leq \frac{L K_{\max}}{K_{\min}} \int_{t_1}^{t} |\phi(s) - s| \, ds + \frac{K_{\max}}{K_{\min}} (t_1 - t_0). \end{aligned}$$

Thus, by Grönwall's inequality, we have, for every $t \geq t_1$,

$$|\phi(t) - t| \leq (t_1 - t_0)\frac{K_{\max}}{K_{\min}} e^{\frac{LK_{\max}}{K_{\min}}(t-t_1)},$$

which yields, for $t = t_1^*$, that

$$|t_1^* - t_0^*| \leq (t_1 - t_0)\frac{K_{\max}}{K_{\min}} e^{\frac{LK_{\max}}{K_{\min}}(t_1^* - t_1)}.$$

Note that $0 \leq t_1^* - t_1 = \phi^{-1}(t_0^*) - \phi^{-1}(t_0) \leq \frac{K_{\max}}{K_{\min}}(t_0^* - t_0) = \frac{K_{\max}}{K_{\min}}\varphi(t_0, x_0) \leq \frac{TK_{\max}}{K_{\min}}$. Moreover, $t_1^* = t_0 + \tau(t_0, \gamma_1)$ and $t_0^* = t_0 + \varphi(t_0, x_0)$, showing that $|t_1^* - t_0^*| = |\tau(t_0, \gamma_1) - \varphi(t_0, x_0)|$. Hence, we deduce that

$$\tau(t_0, \gamma_1) - \varphi(t_0, x_0) \leq (t_1 - t_0)\frac{K_{\max}}{K_{\min}} \exp\left(\frac{TLK_{\max}^2}{K_{\min}^2}\right).$$

Recalling that $0 \leq t_1 - t_0 \leq \frac{D|x_1 - x_0|}{K_{\min}}$, we finally obtain (3.4) with $C = \frac{DK_{\max}}{K_{\min}^2} \exp\left(\frac{TLK_{\max}^2}{K_{\min}^2}\right)$. □

Now, exactly as in Proposition 4.8 of [55], we can deduce Lipschitz continuity of φ by using Lemma 3.2 and the dynamic programming principle from Proposition 3.1(c). We provide a proof of this fact here for sake of completeness.

Proposition 3.3. *Consider the optimal control problem* $\mathrm{OCP}(k)$ *and its value function* φ *and assume that* (H1), (H2), (H3′), (H4), *and* (H5) *are satisfied. Then there exists* $M > 0$ *such that, for every* $(t_0, x_0), (t_1, x_1) \in \mathbb{R}_+ \times \bar{\Omega}$, *we have*

$$|\varphi(t_0, x_0) - \varphi(t_1, x_1)| \leq M\left(|t_0 - t_1| + |x_0 - x_1|\right).$$

Proof. We denote by $C > 0$ the Lipschitz constant from Lemma 3.2. Let $(t_0, x_0), (t_1, x_1) \in \mathbb{R}_+ \times \bar{\Omega}$ and assume, with no loss of generality, that $t_0 < t_1$. Fix $\gamma_0 \in \mathrm{Opt}(k, t_0, x_0)$ and set $x_0^* = \gamma_0(t_1)$. By Lemma 3.2, we have

$$|\varphi(t_1, x_0^*) - \varphi(t_1, x_1)| \leq C|x_0^* - x_1|. \tag{3.6}$$

If $t_1 \leq t_0 + \varphi(t_0, x_0)$, then, by Proposition 3.1(c), since $\gamma_0 \in \mathrm{Opt}(k, t_0, x_0)$, we have $\varphi(t_1, x_0^*) = \varphi(t_0, x_0) - (t_1 - t_0)$, and thus

$$|\varphi(t_0, x_0) - \varphi(t_1, x_1)| \leq |t_1 - t_0| + C|x_0^* - x_1|. \tag{3.7}$$

Otherwise, we have $t_1 > t_0 + \varphi(t_0, x_0)$, which shows that $x_0^* = \gamma_0(t_1) = \gamma_0(t_0 + \varphi(t_0, x_0)) \in \Gamma$, and thus $\varphi(t_1, x_0^*) = 0$. Combining this with (3.6) and the fact that $\varphi(t_0, x_0) < t_1 - t_0$, we deduce that (3.7) also holds in this case.

Since γ_0 is $K_{\max}$-Lipschitz, we have $|x_0^* - x_0| = |\gamma_0(t_1) - \gamma_0(t_0)| \leq K_{\max}|t_1 - t_0|$. Hence, combining with (3.7), we deduce that

$$|\varphi(t_0, x_0) - \varphi(t_1, x_1)| \leq (CK_{\max} + 1)|t_1 - t_0| + C|x_0 - x_1|,$$

yielding the conclusion. □

The next preliminary result we present is the fact that φ satisfies a Hamilton–Jacobi equation. This kind of result is classical and can be obtained by adapting classical proofs used in the autonomous case, such as those

of Chapter IV, Proposition 2.3 of [7] and Theorem 8.1.8 of [17]. The statement provided here can also be found in Proposition 3.5 of [30] and Theorem 4.1 of [50].

Proposition 3.4. *Consider* $\mathrm{OCP}(k)$ *under hypotheses* (H1), (H2), (H3′), (H4), *and* (H5). *The value function* φ *of* $\mathrm{OCP}(k)$ *satisfies the Hamilton–Jacobi equation*

$$- \partial_t \varphi(t, x) + |\nabla \varphi(t, x)| k(t, x) - 1 = 0 \tag{3.8}$$

in the following sense: φ *is a viscosity subsolution of* (3.8) *in* $\mathbb{R}_+ \times (\Omega \setminus \Gamma)$ *and a viscosity supersolution of* (3.8) *in* $\mathbb{R}_+ \times (\bar{\Omega} \setminus \Gamma)$. *Moreover,* φ *satisfies* $\varphi(t, x) = 0$ *for every* $(t, x) \in \mathbb{R}_+ \times \Gamma$.

We next provide in Proposition 3.5 two properties of φ, the first one providing a lower bound on the rate of change of φ over time at a fixed position, and the second one characterizing the optimal control at points at which φ and the optimal trajectory are differentiable. The first result was shown in Proposition 4.4 of [50], once again under the stronger assumption that $k \in \mathrm{Lip}(\mathbb{R}_+ \times \bar{\Omega}; \mathbb{R}_+)$, but its proof was later refined in Proposition 3.9 of [30] in order to use only assumptions (H4) and (H5). Even though [30] considers only minimal-time mean field games without state constraints, the proof of Proposition 3.9 in [30] remains unchanged when state constraints are present. The second result of Proposition 3.5 follows as a consequence of the first one and Proposition 3.4, as detailed, for instance, in Corollary 4.1 of [50].

Proposition 3.5. *Consider* $\mathrm{OCP}(k)$ *under hypotheses* (H1), (H2), (H3′), (H4), *and* (H5), *and let* φ *be the value function of* $\mathrm{OCP}(k)$.

(a) *There exists* $c > 0$ *such that, for every* $x \in \bar{\Omega}$ *and* $t_0, t_1 \in \mathbb{R}_+$ *with* $t_0 \neq t_1$, *we have*

$$\frac{\varphi(t_1, x) - \varphi(t_0, x)}{t_1 - t_0} \geq c - 1.$$

In particular, if φ *is differentiable at* (t_0, x), *then* $\partial_t \varphi(t_0, x) \geq c - 1$ *and* $|\nabla \varphi(t_0, x)| \geq c$.

(b) *For every* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$, *if* $\gamma \in \mathrm{Opt}(k, t_0, x_0)$, $t \in [t_0, t_0 + \varphi(t_0, x_0))$, *and* φ *is differentiable at* $(t, \gamma(t))$, *then* $|\nabla \varphi(t, \gamma(t))| \neq 0$ *and*

$$\dot{\gamma}(t) = -k(t, \gamma(t)) \frac{\nabla \varphi(t, \gamma(t))}{|\nabla \varphi(t, \gamma(t))|}.$$

We now turn to preliminary results concerning the mean field game $\mathrm{MFG}(K)$. The main result on $\mathrm{MFG}(K)$ we present here is the following, asserting existence of equilibria.

Proposition 3.6. *Consider the mean field game* $\mathrm{MFG}(K)$ *under hypotheses* (H1), (H2), (H3′), (H6), *and* (H7). *Then, for every* $m_0 \in \mathcal{P}(\bar{\Omega})$, *there exists an equilibrium* $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$ *for* $\mathrm{MFG}(K)$ *with initial condition* m_0.

A result similar to Proposition 3.6 was shown in Theorem 5.1 of [50], but requiring the stronger assumption that $K \in \mathrm{Lip}(\mathcal{P}(\bar{\Omega}) \times \bar{\Omega}; \mathbb{R}_+)$ instead of (H7). Proofs using only the weaker assumption (H7) were provided in Theorem 4.4 of [30] and Theorem 5.1 of [55]. Both references consider only mean field games without state constraints, but the only point in those proofs where the absence of state constraints is used is to show Lipschitz continuity of value functions of optimal control problems with functions k satisfying (H4) and (H5). Since Proposition 3.3 above proves this fact also in the presence of state constraints, the proof of Proposition 3.6 can be carried out exactly as in Theorem 4.4 of [30] and Theorem 5.1 of [55], and it is thus omitted here. We also recall that equilibria of $\mathrm{MFG}(K)$ may fail to be unique, as illustrated in Remark 7.1 of [50] in the case of a population without interaction and through a heuristic argument in Remark 5.7 of [55] in the case of a multi-population mean field game.

The definition of equilibrium used here, which follows that provided in [50, 55], requires Q-almost every $\gamma \in \mathbf{C}(\bar{\Omega})$ to be an optimal trajectory for $\mathrm{OCP}(k_Q)$ from $(0, \gamma(0))$. An apparently stronger notion of equilibrium

can be formulated by requiring *every* trajectory in the support of Q to be optimal for the same problem, a notion used, for instance, in [11, 30]. We prove in the next proposition that, for the current model, both notions are equivalent.

Proposition 3.7. *Consider the mean field game* $\mathrm{MFG}(K)$ *under hypotheses* (H1), (H2), (H3'), (H6), *and* (H7), *and let* $m_0 \in \mathcal{P}(\bar{\Omega})$. *A measure* $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$ *is an equilibrium of* $\mathrm{MFG}(K)$ *with initial condition* m_0 *if and only if* $e_{0\#}Q = m_0$ *and every* $\gamma \in \mathrm{supp}(Q)$ *satisfies* $\gamma \in \mathrm{Opt}(k_Q, 0, \gamma(0))$, *where* $k_Q : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$ *is defined for* $t \geq 0$ *and* $x \in \bar{\Omega}$ *by* $k_Q(t, x) = K(e_{t\#}Q, x)$.

Proof. Notice first that $t \mapsto e_{t\#}Q$ is Lipschitz continuous with Lipschitz constant $K_{\max}$, since, for $t_0, t_1 \in \mathbb{R}_+$, we have

$$
\begin{aligned}
\mathbf{W}_1(e_{t_0\#}Q, e_{t_1\#}Q) &= \sup_{\Phi \in \mathrm{Lip}_1(\bar{\Omega};\mathbb{R})} \int_{\bar{\Omega}} \Phi(x) \, \mathrm{d}(e_{t_0\#}Q - e_{t_1\#}Q)(x) \\
&= \sup_{\Phi \in \mathrm{Lip}_1(\bar{\Omega};\mathbb{R})} \int_{\mathbf{C}(\bar{\Omega})} [\Phi(\gamma(t_0)) - \Phi(\gamma(t_1))] \, \mathrm{d}Q(\gamma) \leq K_{\max}|t_0 - t_1|.
\end{aligned}
$$

In particular, k_Q satisfies (H4) and (H5).

Let us assume that $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$ is such that $e_{0\#}Q = m_0$ and every $\gamma \in \mathrm{supp}(Q)$ satisfies $\gamma \in \mathrm{Opt}(k_Q, 0, \gamma(0))$. Since $\mathrm{supp}(Q)$ is a set of full Q measure, we deduce immediately that Q-almost every $\gamma \in \mathbf{C}(\bar{\Omega})$ satisfies $\gamma \in \mathrm{Opt}(k_Q, 0, \gamma(0))$, and hence Q is an equilibrium of $\mathrm{MFG}(K)$ with initial condition m_0.

To prove the converse implication, let us now assume that Q is an equilibrium of $\mathrm{MFG}(K)$ with initial condition m_0. Fix $\gamma \in \mathrm{supp}(Q)$. By definition of support, for every open neighborhood V of γ in $\mathbf{C}(\bar{\Omega})$, we have $Q(V) > 0$, and thus, by assumption, there exists $\gamma_V \in V$ such that $\gamma_V \in \mathrm{Opt}(k_Q, 0, \gamma_V(0))$. In particular, there exists a sequence $(\gamma_n)_{n\in\mathbb{N}}$ such that $\gamma_n \in \mathrm{Opt}(k_Q, 0, \gamma_n(0))$ for every $n \in \mathbb{N}$ and $\gamma_n \to \gamma$ in the topology of $\mathbf{C}(\bar{\Omega})$ (*i.e.*, uniformly on compact subsets of $\mathbb{R}_+$) as $n \to +\infty$.

Since $\gamma_n \in \mathrm{Opt}(k_Q, 0, \gamma_n(0))$, we have in particular $\gamma_n \in \mathrm{Adm}(k_Q)$, which implies that, for every $t_0, t_1 \in \mathbb{R}_+$, we have

$$
|\gamma_n(t_1) - \gamma_n(t_0)| \leq \int_{t_0}^{t_1} k_Q(s, \gamma_n(s)) \, \mathrm{d}s.
$$

Letting $n \to +\infty$, we deduce that

$$
|\gamma(t_1) - \gamma(t_0)| \leq \int_{t_0}^{t_1} k_Q(s, \gamma(s)) \, \mathrm{d}s,
$$

which shows in particular that γ is $K_{\max}$-Lipschitz continuous. Moreover, dividing by $|t_1 - t_0|$ and letting $t_1 \to t_0$, we deduce that $|\dot{\gamma}(t)| \leq k_Q(t, \gamma(t))$ for almost every $t \in \mathbb{R}_+$, showing that $\gamma \in \mathrm{Adm}(k_Q)$.

To prove optimality of γ, notice that $\gamma_n(\varphi_Q(0, \gamma_n(0))) \in \Gamma$, where φ_Q denotes the value function of $\mathrm{OCP}(k_Q)$. Since Γ is closed and, by Proposition 3.3, φ_Q is continuous, we deduce by letting $n \to +\infty$ that $\gamma(\varphi_Q(0, \gamma(0))) \in \Gamma$, showing that $\tau(0, \gamma) \leq \varphi_Q(0, \gamma(0))$. On the other hand, since $\gamma \in \mathrm{Adm}(k_Q)$, we have $\varphi_Q(0, \gamma(0)) \leq \tau(0, \gamma)$ by the definition (2.3) of φ_Q, yielding that $\tau(0, \gamma) = \varphi_Q(0, \gamma(0))$. Thus $\gamma \in \mathrm{Opt}(k_Q, 0, \gamma(0))$, as required. $\square$

4. Further properties of the optimal control problem

We now provide some further results on the optimal control problem $\mathrm{OCP}(k)$. In the references [30, 50, 55], additional properties of the value function and of optimal trajectories, such as $\mathbf{C}^{1,1}$ regularity of optimal trajectories or differentiability of the value function along optimal trajectories, were obtained only for optimal control problems without state constraints. The technique used in those references is to apply the (unconstrained)

Pontryagin Maximum Principle to obtain further properties of optimal trajectories that can then be used to deduce additional results on OCP(k). The main difficulty in generalizing this technique is that, even though versions of Pontryagin Maximum Principle taking into account control systems with state constraints such as (2.1) exist in the literature (see, *e.g.*, [25], Chap. 5), it is a difficult problem to get nice additional properties of OCP(k) from their conclusions.

The strategy we follow in this section relies instead on a penalization technique adapted from [15], which consists in considering an optimal control problem with no state constraints but such that the maximal speed of each agent decays fast to 0 as the agent moves away from $\bar{\Omega}$. It is then possible to show that, if the penalization parameter is small enough, optimal trajectories of the penalized problem never leave $\bar{\Omega}$ and thus they coincide with optimal trajectories of the original problem with state constraints. As consequences, we provide a boundary condition for the Hamilton–Jacobi equation (3.8) and a characterization of optimal controls, the latter being a key ingredient for showing, in Section 5, that equilibria of MFG(K) satisfy an MFG system.

4.1. Penalized optimal control problem

We present in this section the penalized optimal control problem we consider in order to study optimal trajectories of OCP(k). Before turning to the core of this section, let us first recall, in the next proposition, some classical consequences of (H3) on the signed distance to $\partial\Omega$, whose proofs can be found, for instance, in Theorems 5.1 and 5.7 of [28].

Proposition 4.1. *Let $\Omega \subset \mathbb{R}^d$ be a set satisfying* (H3).

(a) *The signed distance $d_{\partial\Omega}^{\pm}$ is Lipschitz continuous on $\mathbb{R}^d$, with Lipschitz constant equal to 1.*
(b) *There exists a neighborhood W of $\partial\Omega$ such that the signed distance to $\partial\Omega$ satisfies $d_{\partial\Omega}^{\pm} \in \mathbf{C}^{1,1}(W; \mathbb{R})$. Moreover, $\left|\nabla d_{\partial\Omega}^{\pm}(x)\right| = 1$ for every $x \in W$.*
(c) *For every $x \in \partial\Omega$, $\nabla d_{\partial\Omega}^{\pm}(x)$ is the outward unit normal to Ω at x.*

Let us now introduced the penalized optimal control problem. We fix Ω, Γ, and k satisfying assumptions (H1)–(H3), (H4), and (H5), and we extend k to a continuous function defined on $\mathbb{R}_+ \times \mathbb{R}^d$ and taking values in $[K_{\min}, K_{\max}]$ such that, for every $t \in \mathbb{R}_+$, $x \mapsto k(t,x)$ is L-Lipschitz continuous on $\mathbb{R}^d$, with L independent of t. Such an extension can be constructed, for instance, by proceeding as in [6].

For $\epsilon > 0$, we define $k_\epsilon : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$ by

$$k_\epsilon(t,x) = k(t,x) \left(1 - \frac{1}{\epsilon} d_\Omega(x) \right)_+ , \tag{4.1}$$

where the notation a_+ is defined by $a_+ = \max(0, a)$ for $a \in \mathbb{R}$. In particular, k_ϵ is continuous, nonnegative, upper bounded by $K_{\max}$, $x \mapsto k_\epsilon(t,x)$ is Lipschitz continuous for every $t \in \mathbb{R}_+$ with a Lipschitz constant $L_\epsilon = L + \frac{K_{\max}}{\epsilon}$ independent of t, $k_\epsilon(t,x) = k(t,x)$ for every $(t,x) \in \mathbb{R}_+ \times \bar{\Omega}$, and $k_\epsilon(t,x) = 0$ if $d_\Omega(x) \geq \epsilon$.

We consider the penalized control system

$$\dot{\gamma}(t) = k_\epsilon(t, \gamma(t))u(t), \qquad t \geq 0, \tag{4.2}$$

where $\gamma(t) \in \mathbb{R}^d$ is the state (which is no longer constrained to remain in $\bar{\Omega}$) and $u(t) \in \bar{B}$ is the control. We denote by OCP$_\epsilon(k_\epsilon)$ the minimal-time optimal control problem of finding, for any $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$, a trajectory $\gamma \in \mathrm{Lip}(\mathbb{R}^d)$ with $\gamma(t_0) = x_0$ and a measurable control $u : \mathbb{R}_+ \to \bar{B}$ satisfying (4.2) and minimizing the time at which γ reaches the target set Γ for the first time after t_0. Similarly to Definition 2.3, we denote by $\mathrm{Adm}_\epsilon(k_\epsilon)$ the set of admissible curves for (4.2), *i.e.*, the set of $\gamma \in \mathrm{Lip}(\mathbb{R}^d)$ such that (4.2) is satisfied for some measurable $u : \mathbb{R}_+ \to \bar{B}$. The definition of the first exit time $\tau(t_0, \gamma)$ from Definition 2.3(b) is extended to curves $\gamma \in \mathrm{Lip}(\mathbb{R}^d)$, and we define optimal trajectories for OCP$_\epsilon(k_\epsilon)$ as in Definition 2.3(c), but requiring in addition that $\tau(t_0, \gamma) < +\infty$ for a trajectory γ to be considered optimal from $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$. The set

of optimal trajectories for $\mathrm{OCP}_\epsilon(k_\epsilon)$ from $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$ is denoted by $\mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$. The value function $\varphi_\epsilon : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+ \cup \{+\infty\}$ of $\mathrm{OCP}_\epsilon(k_\epsilon)$ is defined similarly to Definition 2.3(d).

Standard arguments allow to show that $\mathrm{OCP}_\epsilon(k_\epsilon)$ also satisfies a dynamic programming principle: if $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $\gamma \in \mathrm{Adm}_\epsilon(k_\epsilon)$ is such that $\gamma(t_0) = x_0$, then, for every $h \geq 0$, we have

$$\varphi_\epsilon(t_0 + h, \gamma(t_0 + h)) + h \geq \varphi_\epsilon(t_0, x_0), \tag{4.3}$$

with equality if $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$ and $h \in [0, \varphi_\epsilon(t_0, x_0)]$. However, contrarily to $\mathrm{OCP}(k)$, we may have nonexistence of optimal trajectories for $\mathrm{OCP}_\epsilon(k_\epsilon)$. More precisely, $\mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0) = \varnothing$ if $d_\Omega(x_0) \geq \epsilon$, but one can show by standard arguments that this set is nonempty as soon as $d_\Omega(x_0) < \epsilon$. In addition, $\varphi_\epsilon(t, x) < +\infty$ if and only if $d_\Omega(x) < \epsilon$.

The first result we show for $\mathrm{OCP}_\epsilon(k_\epsilon)$ is the following property of optimal trajectories.

Proposition 4.2. *Consider the optimal control problem $\mathrm{OCP}_\epsilon(k_\epsilon)$ under assumptions (H1)–(H3), (H4), and (H5). Let $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$. Then $d_\Omega(\gamma(t)) < \epsilon$ for every $t \in \mathbb{R}_+$.*

Proof. Denote by $u : \mathbb{R}_+ \to \bar{B}$ an optimal control associated with γ. Assume, to obtain a contradiction, that there exists $t_1 \in \mathbb{R}_+$ such that $d_\Omega(\gamma(t_1)) \geq \epsilon$. In particular, $k_\epsilon(t, \gamma(t_1)) = 0$ for every $t \in \mathbb{R}_+$, and thus the function $\tilde{\gamma} : \mathbb{R}_+ \to \mathbb{R}^d$, defined by $\tilde{\gamma}(t) = \gamma(t_1)$ satisfies $\dot{\tilde{\gamma}}(t) = k_\epsilon(t, \tilde{\gamma}(t))u(t)$ for every $t \in \mathbb{R}_+$. Since the initial value problem consisting of this equation and the initial condition $\tilde{\gamma}(t_1) = \gamma(t_1)$ admits a unique solution, then necessarily $\tilde{\gamma} = \gamma$. However, since $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$, we have $\tau(t_0, \gamma) < +\infty$ and $\gamma(t_0 + \tau(t_0, \gamma)) \in \Gamma \subset \bar{\Omega}$, which contradicts the fact that $d_\Omega(\gamma(t_0 + \tau(t_0, \gamma))) = d_\Omega(\tilde{\gamma}(t_0 + \tau(t_0, \gamma))) = d_\Omega(\gamma(t_1)) \geq \epsilon$. □

We now apply Pontryagin Maximum Principle to optimal trajectories of the penalized problem $\mathrm{OCP}_\epsilon(k_\epsilon)$ starting from a point in $\mathbb{R}^d \setminus \Gamma$.

Proposition 4.3. *Consider the optimal control problem $\mathrm{OCP}_\epsilon(k_\epsilon)$ under assumptions (H1)–(H3), (H4), and (H5), and with k_ϵ defined from k as in (4.1). Let $(t_0, x_0) \in \mathbb{R}_+ \times (\mathbb{R}^d \setminus \Gamma)$, $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$, $T = \varphi_\epsilon(t_0, x_0)$, and $u : \mathbb{R}_+ \to \bar{B}$ be an optimal control associated with γ. Then there exist $\lambda \geq 0$ and an absolutely continuous function $p : [t_0, t_0 + T] \to \mathbb{R}^d$ such that the following assertions hold.*

(a) $\lambda + \max_{t \in [t_0, t_0+T]} |p(t)| > 0$.
(b) *For almost every $t \in [t_0, t_0 + T]$, we have*

$$\dot{p}(t) \in \mathrm{co} \left\{ \zeta \in \mathbb{R}^d \mid (\zeta, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot{\gamma}(t)) \right\}, \tag{4.4}$$

where, for $t \geq 0$, the set $G_\epsilon(t)$ is defined by

$$G_\epsilon(t) = \left\{ (x, v) \in \mathbb{R}^d \times \mathbb{R}^d \mid \exists u \in \bar{B} \text{ such that } k_\epsilon(t, x)u = v \right\}. \tag{4.5}$$

(c) $-p(t_0 + T) \in N_\Gamma(\gamma(t_0 + T))$.
(d) *One has*

$$u(t) = \frac{p(t)}{|p(t)|},$$

almost everywhere on $\{t \in [t_0, t_0 + T] \mid p(t) \neq 0\}$.
(e) $\lambda = k_\epsilon(t_0 + T, \gamma(t_0 + T))|p(t_0 + T)|$.

Before turning to the proof of Proposition 4.3, we remark that, by taking $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$, we implicitly assume that $\mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$ is not empty, and thus $d_\Omega(x_0) < \epsilon$.

Proof. We apply Theorem 8.4.1 of [58] to $\mathrm{OCP}_\epsilon(k_\epsilon)$ and the optimal trajectory γ over the interval $[t_0, t_0 + T]$. Notice, first, that, since $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$, we have in particular $T < +\infty$ and, since $x_0 \in \mathbb{R}^d \setminus \Gamma$, we have $T > 0$. We denote by L_ϵ the Lipschitz constant of $x \mapsto k_\epsilon(t, x)$, and we recall that this constant is independent of $t \in \mathbb{R}_+$.

We first introduce some notation in accordance to the statement of Theorem 8.4.1 of [58]. Let $g : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be defined for $(t_1, x_1, t_2, x_2) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ by $g(t_1, x_1, t_2, x_2) = t_2 - t_1$, define $C = \{(t_1, x_1, t_2, x_2) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \mid t_1 = t_0,\, x_1 = x_0,\, t_2 \geq t_1,\, x_2 \in \Gamma\}$, and let $F_\epsilon : \mathbb{R} \times \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be the set-valued map defined for $(t, x) \in \mathbb{R} \times \mathbb{R}^d$ by $F_\epsilon(t, x) = \{k_\epsilon(\max(t, 0), x)u \mid u \in \bar{B}\}$. Notice that the optimal control problem $\mathrm{OCP}_\epsilon(k_\epsilon)$ starting from the given $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$ can be rephrased as: minimize $g(t_0, \tilde{\gamma}(t_0), t_0 + T, \tilde{\gamma}(t_0 + T))$ over intervals $[t_0, t_0 + T]$ and absolutely continuous functions $\tilde{\gamma} : [t_0, t_0 + T] \to \mathbb{R}^d$ such that $\dot{\tilde{\gamma}}(t) \in F_\epsilon(t, \tilde{\gamma}(t))$ for a.e. $t \in [t_0, t_0 + T]$ and $(t_0, \tilde{\gamma}(t_0), t_0 + T, \tilde{\gamma}(t_0 + T)) \in C$. Remark also that the set $G_\epsilon(t)$ from (b) is the graph of $\mathbb{R}^d \ni x \mapsto F_\epsilon(t, x) \subset \mathbb{R}^d$.

Let us verify the assumptions of Theorem 8.4.1 in [58]. First, the constant $\delta > 0$ from Theorem 8.4.1 in [58] can be chosen arbitrarily. Since g is of class $\mathbf{C}^\infty$ and Γ is closed, assumption (H1) from Theorem 8.4.1 in [58] is verified. The fact that k_ϵ is continuous also allows one to easily verify assumption (H2) from Theorem 8.4.1 in [58]. Using the fact that k_ϵ is L_ϵ-Lipschitz continuous in its second variable uniformly with respect to the first variable, assumption (H3) from Theorem 8.4.1 in [58] is satisfied with $\beta = 0$ and $k_{F_\epsilon}(t) = L_\epsilon$ for every $t \in [t_0, t_0 + T]$. Since t_0 is fixed, assumption (H4) from Theorem 8.4.1 in [58] is not needed. Finally, using once again that k_ϵ is L_ϵ-Lipschitz continuous in its second variable uniformly with respect to the first variable, we deduce that (H5) from Theorem 8.4.1 in [58] holds with $c_1 = K_{\max}$, $k_1 = L_\epsilon$, and with an arbitrary $\delta_1 > 0$. Thus, all assumptions of Theorem 8.4.1 in [58] are verified.

The statement of Theorem 8.4.1 in [58] now asserts the existence of an absolutely continuous $p : [t_0, t_0 + T] \to \mathbb{R}^d$ and real numbers $\lambda \geq 0$, ξ, and η, satisfying (i)–(iv) and (vi) from the statement of Theorem 8.4.1 in [58]. Items (i) and (ii) are exactly (a) and (b) above.

Assertion of Theorem 8.4.1(iii) in [58] states that

$$(-\xi, p(t_0), \eta, -p(t_0 + T)) \in \lambda \partial g(t_0, x_0, t_0 + T, \gamma(t_0 + T)) + N_C(t_0, x_0, t_0 + T, \gamma(t_0 + T)), \qquad (4.6)$$

where ∂g denotes the limiting subdifferential of g (see Def. 4.3.1 in [58] for its definition). Since g is smooth, we have $\partial g(t_0, x_0, t_0 + T, \gamma(t_0 + T)) = \{\nabla g(t_0, x_0, t_0 + T, \gamma(t_0 + T))\} = \{(-1, 0, 1, 0)\}$.

Let us now compute the limiting normal cone of C at $(t_0, x_0, t_0 + T, \gamma(t_0 + T))$. For that purpose, we first compute its proximal normal cone at every $(t_1, x_1, t_2, x_2) \in C$ with $t_2 > t_1$. Let $(s_1, p_1, s_2, p_2) \in N_C^{\mathrm{P}}(t_1, x_1, t_2, x_2)$. By Definition 2.2(a), there exists $M > 0$ such that

$$(s_1, p_1, s_2, p_2) \cdot (t_1' - t_1, x_1' - x_1, t_2' - t_2, x_2' - x_2) \leq M \left[|t_1' - t_1|^2 + |x_1' - x_1|^2 + |t_2' - t_2|^2 + |x_2' - x_2|^2 \right]$$

for every $(t_1', x_1', t_2', x_2') \in C$. By definition of C, we have $t_1 = t_1' = t_0$, $x_1 = x_1' = x_0$, and thus

$$s_2(t_2' - t_2) + p_2 \cdot (x_2' - x_2) \leq M \left[|t_2' - t_2|^2 + |x_2' - x_2|^2 \right].$$

Taking $t_2' = t_2$, we deduce that $p_2 \cdot (x_2' - x_2) \leq M|x_2' - x_2|^2$ for every $x_2' \in \Gamma$, and thus $p_2 \in N_\Gamma^{\mathrm{P}}(x_2)$. Taking $x_2' = x_2$, we deduce that $s_2(t_2' - t_2) \leq M|t_2' - t_2|^2$ for every $t_2' \geq t_1' = t_0$. In particular, for every $\rho > 0$, taking $t_2' = t_2 + \rho$, we have $s_2 \rho \leq M\rho^2$, which yields, since $\rho > 0$ is arbitrary, that $s_2 \leq 0$. On the other hand, taking $t_2' = t_2 - \rho$ for $\rho \in (0, t_2 - t_1]$, we have $-s_2 \rho \leq M\rho^2$, which yields, since $\rho \in (0, t_2 - t_1]$ is arbitrary and $t_2 > t_1$, that $-s_2 \leq 0$, and thus $s_2 = 0$. We have thus shown that[3]

$$N_C^{\mathrm{P}}(t_1, x_1, t_2, x_2) \subset \mathbb{R} \times \mathbb{R}^d \times \{0\} \times N_\Gamma^{\mathrm{P}}(x_2),$$

---

[3]The converse inclusion can be shown by straightforward arguments, but it is not necessary for our proof. Notice also that this inclusion is false in the case $t_2 = t_1$, but this case is not needed in our proof.

It now follows, using Definition 2.2(b), that[4]

$$N_C(t_0, x_0, t_0 + T, \gamma(t_0 + T)) \subset \mathbb{R} \times \mathbb{R}^d \times \{0\} \times N_\Gamma(\gamma(t_0 + T)).$$

In particular, (4.6) imposes no constraints on ξ and $p(t_0)$, and asserts that $\eta = \lambda$ and $-p(t_0 + T) \in N_\Gamma(\gamma(t_0 + T))$, showing (c).

To show (d), notice that Theorem 8.4.1(iv) of [58] states that, for almost every $t \in [t_0, t_0 + T]$, we have $p(t) \cdot \dot{\gamma}(t) \geq p(t) \cdot v$ for every $v \in F_\epsilon(t, \gamma(t))$, which means that, for almost every $t \in [t_0, t_0 + T]$, we have $k_\epsilon(t, \gamma(t))p(t) \cdot u(t) \geq k_\epsilon(t, \gamma(t))p(t) \cdot w$ for every $w \in \bar{B}$. By Proposition 4.2, we have that $d_\Omega(\gamma(t)) < \epsilon$ for every $t \in [t_0, t_0 + T]$, yielding from (4.1) that $k_\epsilon(t, \gamma(t)) > 0$. Hence, for almost every $t \in [t_0, t_0 + T]$, we have $p(t) \cdot u(t) \geq p(t) \cdot w$ for every $w \in \bar{B}$, and this is equivalent to (d).

Finally, define $H_\epsilon : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ for $(t, x, p) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ by $H_\epsilon(t, x, p) = \sup_{v \in F_\epsilon(t,x)} p \cdot v$, and notice that $H_\epsilon(t, x, p) = k_\epsilon(t, x)|p|$. In particular, remarking that H_ϵ is continuous, Theorem 8.4.1(vi) of [58] then asserts that $\eta = k_\epsilon(t_0 + T, \gamma(t_0 + T))|p(t_0 + T)|$, and thus (e) holds since $\eta = \lambda$. $\qquad \square$

The absolutely continuous function p from Proposition 4.3(b) is known as the *costate* associated with the optimal trajectory γ. Even though Proposition 4.3(b) provides the differential inclusion (4.4) for the costate p, this inclusion is in general hard to manipulate. We provide, in the next lemma, the main consequence of this differential inclusion that we will use in the sequel. Recall that, for $r > 0$, $\bar{B}_r$ denotes the closed ball centered at the origin and with radius r.

Lemma 4.4. *Consider the optimal control problem* $\mathrm{OCP}_\epsilon(k_\epsilon)$ *under assumptions* (H1)–(H3), (H4), *and* (H5), *and with* k_ϵ *defined from* k *as in* (4.1). *Let* (t_0, x_0), γ, T, *and* p *be as in the statement of Proposition 4.3 and* $L_\epsilon > 0$ *be the Lipschitz constant of* $x \mapsto k_\epsilon(t, x)$, *which is independent of* t. *Then, for almost every* $t \in [t_0, t_0 + T]$, *we have*

$$\mathrm{co}\left\{ \zeta \in \mathbb{R}^d \mid (\zeta, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot{\gamma}(t)) \right\} \subset \bar{B}_{L_\epsilon|p(t)|}.$$

In particular, for almost every $t \in [t_0, t_0 + T]$, *we have* $|\dot{p}(t)| \leq L_\epsilon|p(t)|$.

Proof. Let G_ϵ be defined as in the statement of Proposition 4.3(b). Since $\bar{B}_{L_\epsilon|p(t)|}$ is convex, the result is proved if we show that, for a.e. $t \in [t_0, t_0 + T]$ and every $\zeta \in \mathbb{R}^d$, if $(\zeta, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot{\gamma}(t))$, then $|\zeta| \leq L_\epsilon|p(t)|$. Note that this inequality is trivial if $\zeta = 0$, and thus we assume from now on that $\zeta \neq 0$.

Fix $t \in [t_0, t_0 + T]$ at which p and γ are differentiable and (4.4) holds. Let $\zeta \in \mathbb{R}^d$ be such that $(\zeta, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot{\gamma}(t))$. By definition of the limiting normal cone of $G_\epsilon(t)$, there exist sequences $(x_n, v_n)_{n \in \mathbb{N}}$ in $G_\epsilon(t)$ and $(\zeta_n, p_n)_{n \in \mathbb{N}}$ in $\mathbb{R}^d \times \mathbb{R}^d$ such that $(x_n, v_n) \to (\gamma(t), \dot{\gamma}(t))$ and $(\zeta_n, p_n) \to (\zeta, p(t))$ as $n \to +\infty$ and $(\zeta_n, p_n) \in N^\mathrm{P}_{G_\epsilon(t)}(x_n, v_n)$ for every $n \in \mathbb{N}$. Since $\zeta \neq 0$, we assume, with no loss of generality, that $\zeta_n \neq 0$ for every $n \in \mathbb{N}$.

Take $n \in \mathbb{N}$. Since $(x_n, v_n) \in G_\epsilon(t)$, it follows from (4.5) that there exists $u_n \in \bar{B}$ such that $v_n = k_\epsilon(t, x_n)u_n$. Since $(\zeta_n, p_n) \in N^\mathrm{P}_{G_\epsilon(t)}(x_n, v_n)$, there exists $M_n > 0$ such that

$$(\zeta_n, p_n) \cdot (y - x_n, k_\epsilon(t, y)u - k_\epsilon(t, x_n)u_n) \leq M_n \left[|y - x_n|^2 + |k_\epsilon(t, y)u - k_\epsilon(t, x_n)u_n|^2 \right] \qquad (4.7)$$

for every $y \in \mathbb{R}^d$ and $u \in \bar{B}$. Let $\alpha_n = ((n+1)(1 + L_\epsilon^2)M_n)^{-1} > 0$ and take $y = x_n + \alpha_n \zeta_n$ and $u = u_n$ in (4.7). Then

$$\alpha_n|\zeta_n|^2 + \left[k_\epsilon(t, x_n + \alpha_n \zeta_n) - k_\epsilon(t, x_n) \right] p_n \cdot u_n \leq M_n \left[\alpha_n^2|\zeta_n|^2 + |k_\epsilon(t, x_n + \alpha_n \zeta_n) - k_\epsilon(t, x_n)|^2 \right].$$

---

[4]Once again, we also have the converse inclusion, but it is not necessary for our proof.

Hence

$$\alpha_n |\zeta_n|^2 \le M_n \alpha_n^2 |\zeta_n|^2 + M_n L_\epsilon^2 \alpha_n^2 |\zeta_n|^2 + \alpha_n L_\epsilon |\zeta_n| |p_n|.$$

Since $\alpha_n > 0$ and $|\zeta_n| > 0$, we deduce that

$$\left(1 - M_n \alpha_n - M_n L_\epsilon^2 \alpha_n\right) |\zeta_n| \le L_\epsilon |p_n|.$$

By definition of α_n, we have $M_n \alpha_n + M_n L_\epsilon^2 \alpha_n = \frac{1}{n+1} < 1$, and thus

$$|\zeta_n| \le \frac{L_\epsilon}{1 - \frac{1}{n+1}} |p_n|.$$

The conclusion follows by letting $n \to +\infty$.   □

Using Lemma 4.4, we can prove additional regularity properties of optimal trajectories and optimal controls for $\mathrm{OCP}_\epsilon(k_\epsilon)$.

Proposition 4.5. *Consider the optimal control problem* $\mathrm{OCP}_\epsilon(k_\epsilon)$ *under assumptions* (H1)–(H3), (H4), *and* (H5), *and with* k_ϵ *defined from* k *as in* (4.1). *Let* (t_0, x_0), γ, T, u, *and* p *be as in the statement of Proposition 4.3 and* $L_\epsilon > 0$ *be the Lipschitz constant of* k_ϵ *with respect to its second variable, which is independent of its first variable. For* $t \in [t_0, t_0 + T]$, *define* $\Xi_\epsilon(t)$ *by*

$$\Xi_\epsilon(t) = \mathrm{co} \left\{ \xi \in \mathbb{R}^d \mid (\xi |p(t)|, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot\gamma(t)) \right\}, \tag{4.8}$$

where $G_\epsilon(t)$ *is given by* (4.5). *Then, up to redefining* u *in a set of Lebesgue measure zero, the following assertions hold.*

(a) *For every* $t \in [t_0, t_0 + T]$, *we have* $|p(t)| \ne 0$ *and* $u(t) = \frac{p(t)}{|p(t)|}$.
(b) *For almost every* $t \in [t_0, t_0 + T]$ *and every* $\xi \in \Xi_\epsilon(t)$, *we have* $|\xi| \le L_\epsilon$.
(c) *We have* $\gamma \in \mathbf{C}^1([t_0, t_0 + T]; \mathbb{R}^d)$, $u \in \mathrm{Lip}_{L_\epsilon}([t_0, t_0 + T]; \mathbb{S}^{d-1})$, *and* (γ, u) *satisfies, for almost every* $t \in [t_0, t_0 + T]$, *the system*

$$\begin{cases} \dot\gamma(t) = k_\epsilon(t, \gamma(t)) u(t) \\ \dot u(t) \in \mathrm{Pr}^\perp_{u(t)} \Xi_\epsilon(t) \end{cases} \tag{4.9}$$

where, for $x \in \mathbb{S}^{d-1}$ *and* $A \subset \mathbb{R}^d$, $\mathrm{Pr}^\perp_x A$ *denotes the projection of the vectors of* A *onto the tangent space of* $\mathbb{S}^{d-1}$ *at* x, *defined by* $\mathrm{Pr}^\perp_x A = \{a - (x \cdot a)x \mid a \in A\}$.

Proof. Let us first prove (a). By Lemma 4.4, we have that $|\dot p(t)| \le |p(t)| L_\epsilon$, implying that, for every $t_1, t_2 \in [t_0, t_0 + T]$, we have

$$|p(t_2)| \le |p(t_1)| + L_\epsilon \int_{\min\{t_1, t_2\}}^{\max\{t_1, t_2\}} |p(s)| \, \mathrm{d}s.$$

Hence, by Grönwall's inequality, for every $t_1, t_2 \in [t_0, t_0 + T]$, we have

$$|p(t_2)| \le |p_\epsilon(t_1)| e^{L_\epsilon |t_2 - t_1|}.$$

If there exists t_1 such that $p(t_1) = 0$, then $p(t) = 0$ for every $t \in [t_0, t_0 + T_\epsilon]$. Letting $\lambda \ge 0$ be as in the statement of Proposition 4.3, we have, by Proposition 4.3(e), that $\lambda = 0$, which is a contradiction according to

Proposition 4.3(a). Therefore $|p(t)| \neq 0$ for every $t \in [t_0, t_0 + T]$. We then deduce from Proposition 4.3(d) that $u(t) = \frac{p(t)}{|p(t)|}$ for almost every $t \in [t_0, t_0 + T]$, and the conclusion for every $t \in [t_0, t_0 + T]$ holds by redefining u in a set of measure zero. Hence (a) is proved.

Note that, for every $t \in [t_0, t_0 + T]$, we have $\xi \in \Xi_\epsilon(t)$ if and only if

$$\frac{\xi}{|p(t)|} \in \mathrm{co} \left\{ \zeta \in \mathbb{R}^d \mid (\zeta, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot\gamma(t)) \right\}.$$

In particular, (b) follows as a consequence of (a) and Lemma 4.4. Notice also that, by Proposition 4.3(b), we have $\dot p(t) \in |p(t)|\Xi_\epsilon(t)$ for a.e. $t \in [t_0, t_0 + T]$.

Let us now prove (c). Note that, by (a), $u(t) \in \mathbb{S}^{d-1}$ for every $t \in [t_0, t_0 + T]$ and, since p is absolutely continuous, u is also absolutely continuous. Let $\xi : [t_0, t_0 + T] \to \mathbb{R}^d$ be a measurable function with $\xi(t) \in \Xi_\epsilon(t)$ and $\dot p(t) = |p(t)|\xi(t)$ for a.e. $t \in [t_0, t_0 + T]$. We compute, for a.e. $t \in [t_0, t_0 + T]$,

$$\dot u(t) = \frac{\dot p(t)|p(t)| - \frac{p(t) \cdot \dot p(t)}{|p(t)|} p(t)}{|p(t)|^2} = \xi(t) - (u(t) \cdot \xi(t))u(t),$$

which yields the differential inclusion for u in system (4.9). By (b), we deduce that $|\dot u(t)| \leq L_\epsilon$ for a.e. $t \in [t_0, t_0 + T]$, yielding that u is L_ϵ-Lipschitz continuous on $[t_0, t_0 + T]$. The first equation in (4.9) is simply (4.2), and its right-hand side is continuous on $[t_0, t_0 + T]$, showing that $\gamma \in \mathbf{C}^1([t_0, t_0 + T]; \mathbb{R}^d)$. □

We will need in the sequel the following technical lemma.

Lemma 4.6. *Consider the optimal control problem* $\mathrm{OCP}_\epsilon(k_\epsilon)$ *under assumptions (H1)–(H3), (H4), and (H5), and with k_ϵ defined from k as in (4.1). Let (t_0, x_0), γ, T, and p be as in the statement of Proposition 4.3 and $L > 0$ be the Lipschitz constant of $\mathbb{R}^d \ni x \mapsto k(t, x) \in [K_{\min}, K_{\max}]$, which is independent of t. For $t \in [t_0, t_0 + T]$, define $\Xi_\epsilon(t)$ as in the statement of Proposition 4.5, i.e.,*

$$\Xi_\epsilon(t) = \mathrm{co} \left\{ \xi \in \mathbb{R}^d \mid (\xi|p(t)|, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot\gamma(t)) \right\},$$

where $G_\epsilon(t)$ is given by (4.5). Finally, let W be a neighborhood of $\partial\Omega$ in which the signed distance $d_{\partial\Omega}^\pm$ belongs to $\mathbf{C}^{1,1}(W; \mathbb{R})$, which exists thanks to Proposition 4.1(b).

For almost every $t \in [t_0, t_0 + T]$, if $\gamma(t) \in W$ and $\gamma(t) \notin \bar\Omega$, then, for every $\xi \in \Xi_\epsilon(t)$, we have

$$\left| \xi \cdot \nabla d_{\partial\Omega}^\pm(\gamma(t)) - \frac{1}{\epsilon} k(t, \gamma(t)) \right| \leq \left(1 - \frac{1}{\epsilon} d_\Omega(\gamma(t)) \right) L. \tag{4.10}$$

Proof. Fix $t \in [t_0, t_0 + T]$ at which γ is differentiable and such that $\gamma(t) \in W$ and $\gamma(t) \notin \bar\Omega$. Since the set of $\xi \in \mathbb{R}^d$ such that (4.10) holds is convex, it suffices to show that, for every $\xi \in \mathbb{R}^d$, if $(\xi|p(t)|, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot\gamma(t))$, then (4.10) holds, where $G_\epsilon(t)$ is given by (4.5). Fix $\xi \in \mathbb{R}^d$ such that $(\xi|p(t)|, p(t)) \in N_{G_\epsilon(t)}(\gamma(t), \dot\gamma(t))$. Let u be an optimal control associated with γ and recall that, by Proposition 4.5, u is Lipschitz continuous.

By definition of $N_{G_\epsilon(t)}(\gamma(t), \dot\gamma(t))$, there exist sequences $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$, $(u_n)_{n\in\mathbb{N}}$ in $\bar B$, $(\zeta_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$, and $(p_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$ such that, as $n \to +\infty$, we have $x_n \to \gamma(t)$, $k_\epsilon(t, x_n)u_n \to k_\epsilon(t, \gamma(t))u(t)$, $\zeta_n \to \xi|p(t)|$, and $p_n \to p(t)$, and in addition $(\zeta_n, p_n) \in N^P_{G_\epsilon(t)}(x_n, k_\epsilon(t, x_n)u_n)$ for every $n \in \mathbb{N}$. By Proposition 4.5(a), we have $|p(t)| \neq 0$, and we then define $\xi_n = \frac{\zeta_n}{|p(t)|}$ for $n \in \mathbb{N}$. Then, as $n \to +\infty$, we have $\xi_n \to \xi$ and, since k_ϵ is continuous and[5] $k_\epsilon(t, \gamma(t)) > 0$, we also have $u_n \to u(t)$.

---

[5]Recall that, by Proposition 4.2, we have $d_\Omega(\gamma(t)) < \epsilon$.

For every $n \in \mathbb{N}$, since $(\zeta_n, p_n) \in N^{\mathrm{P}}_{G_\epsilon(t)}(x_n, k_\epsilon(t, x_n) u_n)$, there exists $M_n > 0$ such that

$$(\xi_n | p(t)|, p_n) \cdot (y - x_n, k_\epsilon(t, y)w - k_\epsilon(t, x_n)u_n) \leq M_n \left[|y - x_n|^2 + |k_\epsilon(t, y)w - k_\epsilon(t, x_n)u_n|^2 \right] \quad (4.11)$$

for every $y \in \mathbb{R}^d$ and $w \in \bar{B}$. Let $(\alpha_n)_{n \in \mathbb{N}}$ be a sequence of positive real numbers such that $\alpha_n \to 0$ and $M_n \alpha_n \to 0$ as $n \to +\infty$. Since $\gamma(t) \in W$, $\gamma(t) \notin \bar{\Omega}$, and $d_\Omega(\gamma(t)) < \epsilon$, for n large enough, we have $x_n \in W$, $x_n \notin \bar{\Omega}$, and $d_\Omega(x_n) < \epsilon$. In particular, by Proposition 4.1(b), $\nabla d^{\pm}_{\partial\Omega}(x_n)$ is well-defined and $|\nabla d^{\pm}_{\partial\Omega}(x_n)| = 1$. For $n \in \mathbb{N}$, apply (4.11) with $y = x_n + \alpha_n \nabla d^{\pm}_{\partial\Omega}(x_n)$ and $w = u_n$. Then

$$\alpha_n |p(t)| \xi_n \cdot \nabla d^{\pm}_{\partial\Omega}(x_n) + p_n \cdot u_n \left[k_\epsilon(t, x_n + \alpha_n \nabla d^{\pm}_{\partial\Omega}(x_n)) - k_\epsilon(t, x_n) \right] \leq M_n \alpha_n^2 (1 + L_\epsilon^2),$$

where L_ϵ is the Lipschitz constant of $\mathbb{R}^d \ni x \mapsto k_\epsilon(t, x) \in \mathbb{R}_+$, which is independent of t. Dividing by α_n, we get

$$|p(t)| \xi_n \cdot \nabla d^{\pm}_{\partial\Omega}(x_n) + p_n \cdot u_n \frac{k_\epsilon(t, x_n + \alpha_n \nabla d^{\pm}_{\partial\Omega}(x_n)) - k_\epsilon(t, x_n)}{\alpha_n} \leq M_n \alpha_n (1 + L_\epsilon^2). \quad (4.12)$$

Since $\nabla d^{\pm}_{\partial\Omega}$ is continuous in W, we have, as $n \to +\infty$, that $\xi_n \cdot \nabla d^{\pm}_{\partial\Omega}(x_n) \to \xi \cdot \nabla d^{\pm}_{\partial\Omega}(\gamma(t))$, $p_n \cdot u_n \to p(t) \cdot u(t) = |p(t)|$, and, by construction, $M_n \alpha_n \to 0$.

Let us denote for simplicity $z_n = x_n + \alpha_n \nabla d^{\pm}_{\partial\Omega}(x_n)$. Since $\alpha_n \to 0$ as $n \to +\infty$ and $\gamma(t) \in W$, $\gamma(t) \notin \bar{\Omega}$, and $d_\Omega(\gamma(t)) < \epsilon$, we deduce that, for n large enough, $x_n \in W$, $z_n \in W$, $x_n \notin \bar{\Omega}$, $z_n \notin \bar{\Omega}$, $d_\Omega(x_n) < \epsilon$, and $d_\Omega(z_n) < \epsilon$ for n large enough. Using those facts and (4.1), we compute

$$\frac{1}{\alpha_n} \left[k_\epsilon(t, z_n) - k_\epsilon(t, x_n) \right] = \frac{1}{\alpha_n} \left[k(t, z_n) \left(1 - \frac{1}{\epsilon} d^{\pm}_{\partial\Omega}(z_n) \right) - k(t, x_n) \left(1 - \frac{1}{\epsilon} d^{\pm}_{\partial\Omega}(x_n) \right) \right]$$

$$= \frac{1}{\epsilon} k(t, z_n) \frac{d^{\pm}_{\partial\Omega}(x_n) - d^{\pm}_{\partial\Omega}(z_n)}{\alpha_n} + \beta_n \left(1 - \frac{1}{\epsilon} d^{\pm}_{\partial\Omega}(x_n) \right),$$

where

$$\beta_n = \frac{k(t, z_n) - k(t, x_n)}{\alpha_n}.$$

Note that, since k is L-Lipschitz continuous in its second argument, we have $|\beta_n| \leq L$ for n large enough. In particular, up to extracting a subsequence, there exists $\beta_* \in [-L, L]$ such that $\beta_n \to \beta_*$ as $n \to +\infty$. Moreover, since $d^{\pm}_{\partial\Omega} \in \mathbf{C}^{1,1}(W; \mathbb{R})$, we have $\frac{d^{\pm}_{\partial\Omega}(x_n) - d^{\pm}_{\partial\Omega}(z_n)}{\alpha_n} \to -\left| \nabla d^{\pm}_{\partial\Omega}(\gamma(t)) \right|^2 = -1$ as $n \to +\infty$. Hence, letting $n \to +\infty$ in (4.12) and dividing by $|p(t)|$, we deduce that

$$\xi \cdot \nabla d^{\pm}_{\partial\Omega}(\gamma(t)) - \frac{1}{\epsilon} k(t, \gamma(t)) + \beta_* \left(1 - \frac{1}{\epsilon} d^{\pm}_{\partial\Omega}(\gamma(t)) \right) \leq 0. \quad (4.13)$$

Performing the same computations as above from (4.11) but now taking $y = x_n - \alpha_n \nabla d^{\pm}_{\partial\Omega}(x_n)$ and $w = u_n$, we get that

$$-\xi \cdot \nabla d^{\pm}_{\partial\Omega}(\gamma(t)) + \frac{1}{\epsilon} k(t, \gamma(t)) + \tilde{\beta}_* \left(1 - \frac{1}{\epsilon} d^{\pm}_{\partial\Omega}(\gamma(t)) \right) \leq 0$$

for some $\tilde{\beta}_* \in [-L, L]$. Together with (4.13), this yields (4.10), as required. $\qquad\square$

Proposition 4.2 stated that, for every $\epsilon > 0$, optimal trajectories for $\mathrm{OCP}_\epsilon(k_\epsilon)$ always remain in an ϵ neighborhood of $\bar{\Omega}$. Using the previous results, we can now be more precise and show that, if $\epsilon > 0$ is small enough, optimal trajectories for $\mathrm{OCP}_\epsilon(k_\epsilon)$ starting at $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$ never leave $\bar{\Omega}$.

Proposition 4.7. *Consider the optimal control problem* $\mathrm{OCP}_\epsilon(k_\epsilon)$ *under assumptions* (H1)–(H3), (H4), *and* (H5) *and* k_ϵ *defined from* k *as in* (4.1). *There exists* $\epsilon_0 > 0$ *such that, for every* $\epsilon \in (0, \epsilon_0)$, $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$, *and* $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$, *we have* $d_\Omega(\gamma(t)) \le d_\Omega(x_0)$ *for every* $t \in \mathbb{R}_+$. *In particular, if* $x_0 \in \bar{\Omega}$, *then* $\gamma(t) \in \bar{\Omega}$ *for every* $t \in \mathbb{R}_+$.

Proof. Take $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$. Notice first that, if $x_0 \in \Gamma$, then $\mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$ is equal to the singleton containing only the trajectory that remains in x_0 for all times. We thus assume, for the rest of the proof, that $x_0 \in \bar{\Omega} \setminus \Gamma$.

Let W be as in the statement of Proposition 4.1(b) and take $\epsilon_* > 0$ such that $\{x \in \mathbb{R}^d \mid d_{\partial\Omega}(x) < \epsilon_*\} \subset W$. Let C be a Lipschitz constant of $\nabla d_{\partial\Omega}^\pm$ on W and L, $K_{\max}$, and $K_{\min}$ be as in (H4) and (H5). Let $\epsilon_0 = \min\left\{\epsilon_*, \frac{K_{\min}}{L+CK_{\max}}\right\}$ and take $\epsilon \in (0, \epsilon_0)$.

Fix $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$ and let $T = \varphi_\epsilon(t_0, x_0)$. Recall that, by Proposition 4.2, $d_\Omega(\gamma(t)) < \epsilon$ for every $t \in \mathbb{R}_+$. Let $p : [t_0, t_0 + T] \to \mathbb{R}^d$ be the costate associated with γ, whose existence is asserted in Proposition 4.3. By Propositions 4.3(b) and 4.5(a), there exists a measurable function $\xi : [t_0, t_0 + T] \to \mathbb{R}^d$ such that $\xi(t) \in \Xi_\epsilon(t)$ and $\dot{p}(t) = |p(t)|\xi(t)$ for a.e. $t \in [t_0, t_0 + T]$, where $\Xi_\epsilon(t)$ is defined in (4.8).

By definition of optimal trajectories, we have that $\gamma(t) = x_0$ for every $t \in [0, t_0]$ and $\gamma(t) = \gamma(t_0 + T) \in \Gamma \subset \bar{\Omega}$ for every $t \in [t_0 + T, +\infty)$, and hence $d_\Omega(\gamma(t)) \le d_\Omega(x_0)$ for every $t \in [0, t_0] \cup [t_0 + T, +\infty)$. Assume, to obtain a contradiction, that there exist $a, b \in [t_0, t_0 + T]$ such that $a < b$, $d_\Omega(\gamma(t)) > d_\Omega(x_0)$ for $t \in (a, b)$ and $d_\Omega(\gamma(t)) = d_\Omega(x_0)$ for $t \in \{a, b\}$. In particular, $d_\Omega(\gamma(t)) > 0$ for $t \in (a, b)$, implying that $\gamma(t) \notin \bar{\Omega}$ for $t \in (a, b)$ and $\gamma(t) \notin \Omega$ for $t \in [a, b]$.

Recall that, by Proposition 4.2, $d_\Omega(\gamma(t)) < \epsilon$ for every $t \in \mathbb{R}_+$, and thus, in particular, $\gamma(t) \in W$ for every $t \in [a, b]$. Hence, by Propositions 4.1 and 4.5(c), the map $t \mapsto d_{\partial\Omega}^\pm(\gamma(t))$ is differentiable on $(t_0, t_0 + T)$, $d_{\partial\Omega}^\pm(\gamma(t)) > d_\Omega(x_0)$ for $t \in (a, b)$, and $d_{\partial\Omega}^\pm(\gamma(t)) = d_\Omega(x_0)$ for $t \in \{a, b\}$. Thus, its derivative is nonnegative at a and nonpositive at b, *i.e.*,

$$\dot{\gamma}(a) \cdot \nabla d_{\partial\Omega}^\pm(\gamma(a)) \ge 0 \qquad \text{and} \qquad \dot{\gamma}(b) \cdot \nabla d_{\partial\Omega}^\pm(\gamma(b)) \le 0.$$

Since, by Propositions 4.2 and 4.5, we have $\dot{\gamma}(t) = k_\epsilon(t, \gamma(t))\frac{p(t)}{|p(t)|}$ and $\frac{k_\epsilon(t, \gamma(t))}{|p(t)|} > 0$ for every $t \in [t_0, t_0 + T]$, we deduce that

$$p(a) \cdot \nabla d_{\partial\Omega}^\pm(\gamma(a)) \ge 0 \qquad \text{and} \qquad p(b) \cdot \nabla d_{\partial\Omega}^\pm(\gamma(b)) \le 0. \qquad (4.14)$$

Consider now the map $\alpha : t \mapsto p(t) \cdot \nabla d_{\partial\Omega}^\pm(\gamma(t))$. Since $\gamma(t) \in W$ for every $t \in [a, b]$ and $d_{\partial\Omega}^\pm \in \mathbf{C}^{1,1}(W; \mathbb{R})$, α is absolutely continuous on $[a, b]$ and, for a.e. $t \in (a, b)$, we have, using Lemma 4.6 and the fact that $\nabla d_{\partial\Omega}^\pm \circ \gamma$ is $CK_{\max}$-Lipschitz continuous in $[a, b]$, that

$$
\begin{aligned}
\dot{\alpha}(t) &= \dot{p}(t) \cdot \nabla d_{\partial\Omega}^\pm(\gamma(t)) + p(t) \cdot \frac{\mathrm{d}\big[\nabla d_{\partial\Omega}^\pm \circ \gamma\big]}{\mathrm{d}t}(t) \\
&= |p(t)|\xi(t) \cdot \nabla d_{\partial\Omega}^\pm(\gamma(t)) + p(t) \cdot \frac{\mathrm{d}\big[\nabla d_{\partial\Omega}^\pm \circ \gamma\big]}{\mathrm{d}t}(t) \\
&\ge \frac{1}{\epsilon}|p(t)|k(t, \gamma(t)) - \left(1 - \frac{1}{\epsilon}d_\Omega(\gamma(t))\right)L|p(t)| - CK_{\max}|p(t)| \\
&\ge |p(t)|\left(\frac{K_{\min}}{\epsilon} - L - CK_{\max}\right) > 0,
\end{aligned}
$$

which contradicts (4.14). □

We are now in position to show the main result of this section.

Theorem 4.8. *Consider the optimal control problem* $\mathrm{OCP}(k)$ *under assumptions* (H1)–(H3), (H4), *and* (H5), *as well as the problem* $\mathrm{OCP}_\epsilon(k_\epsilon)$ *with* k_ϵ *defined from* k *as in* (4.1). *There exists* $\epsilon_0 > 0$ *such that, for every* $\epsilon \in (0, \epsilon_0)$ *and* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$, *we have* $\mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0) = \mathrm{Opt}(k, t_0, x_0)$ *and* $\varphi_\epsilon(t_0, x_0) = \varphi(t_0, x_0)$.

Proof. Let $\epsilon_0 > 0$ be as in the statement of Proposition 4.7, $\epsilon \in (0, \epsilon_0)$, and $(t_0, x_0) \in \mathbb{R} \times \bar{\Omega}$. Let $\gamma_\epsilon \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$ and $\gamma \in \mathrm{Opt}(k, t_0, x_0)$, and denote $T_\epsilon = \varphi_\epsilon(t_0, x_0)$ and $T = \varphi(t_0, x_0)$. Clearly, $\gamma \in \mathrm{Adm}_\epsilon(k_\epsilon)$, and thus $T_\epsilon \leq T$. On the other hand, thanks to Proposition 4.7, $\gamma_\epsilon \in \mathrm{Adm}(k)$, which leads us to $T \leq T_\epsilon$. Hence $T = T_\epsilon$, which shows that we have both $\gamma_\epsilon \in \mathrm{Opt}(k, t_0, x_0)$ and $\gamma \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$, as required. □

Thanks to Theorem 4.8, all properties established in this section for optimal trajectories of $\mathrm{OCP}_\epsilon(k_\epsilon)$ also hold for optimal trajectories of $\mathrm{OCP}(k)$. The next statement, whose proof is an immediate consequence of Proposition 4.5(c) and Theorem 4.8, gathers the properties of optimal trajectories of $\mathrm{OCP}(k)$ that will be used in the sequel.

Corollary 4.9. *Consider the optimal control problem* $\mathrm{OCP}(k)$ *under assumptions* (H1)–(H3), (H4), *and* (H5). *Let* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$, $\gamma \in \mathrm{Opt}(k, t_0, x_0)$, $T = \varphi(t_0, x_0)$, *and* $u : \mathbb{R}_+ \to \bar{B}$ *be an optimal control associated with* γ. *Then, up to redefining* u *in a set of Lebesgue measure zero, we have* $\gamma \in \mathbf{C}^1([t_0, t_0 + T]; \bar{\Omega})$ *and* $u \in \mathrm{Lip}([t_0, t_0 + T]; \mathbb{S}^{d-1})$.

Before concluding this section, we also provide the following result on the local Lipschitz continuity of φ_ϵ, which can be obtained as a consequence of Proposition 4.7. For its statement, we introduce, for $\epsilon > 0$, the set

$$\Omega_\epsilon = \{x \in \mathbb{R}^d \mid d_\Omega(x) < \epsilon\}.$$

Proposition 4.10. *Consider the optimal control problem* $\mathrm{OCP}_\epsilon(k_\epsilon)$ *and its value function* φ_ϵ *under assumptions* (H1)–(H3), (H4), *and* (H5) *and* k_ϵ *defined from* k *as in* (4.1). *Then there exists* $\epsilon_0 > 0$ *such that, for every* $\epsilon \in (0, \epsilon_0)$, *the value function* $\varphi_\epsilon : \mathbb{R}_+ \times \Omega_\epsilon \to \mathbb{R}_+$ *of* $\mathrm{OCP}_\epsilon(k_\epsilon)$ *is locally Lipschitz continuous. More precisely, for every* $\eta \in [0, \epsilon)$, φ_ϵ *is Lipschitz continuous on* $\mathbb{R}_+ \times \bar{\Omega}_\eta$.

Proof. Let $\epsilon_0 > 0$ be as in the statement of Proposition 4.7 and fix $\epsilon \in (0, \epsilon_0)$ and $\eta \in [0, \epsilon)$. We consider the optimal control problem $\widehat{\mathrm{OCP}}$ defined as follows: given $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}_\eta$, minimize $\tau(t_0, \gamma)$ over all Lipschitz continuous curves $\gamma : \mathbb{R}_+ \to \bar{\Omega}_\eta$ satisfying (4.2) for some measurable function $u : \mathbb{R}_+ \to \bar{B}$ and $\gamma(t_0) = x_0$. We denote by $\widehat{\varphi} : \mathbb{R}_+ \times \bar{\Omega}_\eta \to \mathbb{R}_+$ the value function of $\widehat{\mathrm{OCP}}$ and by $\widehat{\mathrm{Opt}}(k_\epsilon, t_0, x_0)$ the set of optimal trajectories for (t_0, x_0), with the convention that $\gamma \in \widehat{\mathrm{Opt}}(k_\epsilon, t_0, x_0)$ must satisfy $\gamma(t) = x_0$ for $t \in [0, t_0]$ and $\gamma(t) = \gamma(t_0 + \widehat{\varphi}(t_0, x_0))$ for $t \in [t_0 + \widehat{\varphi}(t_0, x_0), +\infty)$.

Note that $\widehat{\mathrm{OCP}}$ is a minimal-time optimal control problem with dynamics defined by k_ϵ and with the state constraint $\gamma(t) \in \bar{\Omega}_\eta$ for every $t \geq 0$. In particular, $\widehat{\mathrm{OCP}}$ coincides with the optimal control problem $\mathrm{OCP}(k)$ if one replaces Ω by Ω_η and k by the restriction of k_ϵ to $\mathbb{R}_+ \times \bar{\Omega}_\eta$, and assumptions (H1)–(H3), (H4), and (H5) are satisfied for $\widehat{\mathrm{OCP}}$, with $K_{\min}$ in (H4) replaced by $\widehat{K}_{\min} = \left(1 - \frac{\eta}{\epsilon}\right) K_{\min}$ and L in (H5) replaced by $\widehat{L} = L_\epsilon = L + \frac{K_{\max}}{\epsilon}$. Hence, all properties of Section 3 apply to $\widehat{\mathrm{OCP}}$ and thus, by Proposition 3.3, $\widehat{\varphi}$ is Lipschitz continuous on $\mathbb{R}_+ \times \bar{\Omega}_\eta$.

In order to conclude, it suffices to show that φ_ϵ and $\widehat{\varphi}$ coincide on this set. For this purpose, we proceed as in the proof of Theorem 4.8. Let $(t_0, x_0) \in \mathbb{R} \times \bar{\Omega}_\eta$. Let $\gamma_\epsilon \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$ and $\widehat{\gamma} \in \widehat{\mathrm{Opt}}(k_\epsilon, t_0, x_0)$, and denote $T_\epsilon = \varphi_\epsilon(t_0, x_0)$ and $\widehat{T} = \widehat{\varphi}(t_0, x_0)$. Clearly, $\widehat{\gamma} \in \mathrm{Adm}_\epsilon(k_\epsilon)$, and thus $T_\epsilon \leq \widehat{T}$. On the other hand, thanks to Proposition 4.7, we have $d_\Omega(\gamma_\epsilon(t)) \leq d_\Omega(x_0) \leq \eta$ for every $t \geq 0$, showing that γ_ϵ takes values in $\bar{\Omega}_\eta$. Hence γ_ϵ is an admissible trajectory for the optimal control problem $\widehat{\mathrm{OCP}}$, showing that $\widehat{T} \leq T_\epsilon$. Hence $\widehat{T} = T_\epsilon$, as required. □

4.2. Boundary condition of the value function

Proposition 3.4 in Section 3 asserts that φ is a viscosity solution of the Hamilton–Jacobi equation (3.8) and also provides its boundary condition on $\mathbb{R}_+ \times \Gamma$. The only information provided in Proposition 3.4 on φ on the part of the boundary $\mathbb{R}_+ \times (\partial\Omega \setminus \Gamma)$ is that φ is a viscosity supersolution there. The main result of this section, Theorem 4.11, provides additional information on φ on $\mathbb{R}_+ \times (\partial\Omega \setminus \Gamma)$. We recall that, given $x \in \partial\Omega$, $\mathbf{n}(x)$ denotes the outward unit normal vector to $\bar{\Omega}$ at x.

Theorem 4.11. *Consider the optimal control problem* OCP(k) *and its value function* φ *under assumptions* (H1)–(H3)*,* (H4)*, and* (H5)*. Then* φ *satisfies* $\nabla\varphi(t,x) \cdot \mathbf{n}(x) \geq 0$ *in the viscosity supersolution sense for every* $(t,x) \in \mathbb{R}_+ \times (\partial\Omega \setminus \Gamma)$.

Proof. We consider in this proof that k is extended to a continuous function, still denoted by k, defined in $[-1,+\infty) \times \mathbb{R}^d$ which is Lipschitz continuous with respect to its second variable, uniformly with respect to the first one. Note that, up to considering the optimal control problem OCP$(k(\,\cdot\, - 1,\cdot))$, all of the previous results on minimal-time optimal control problems apply to k with the time domain $\mathbb{R}_+$ replaced by $[-1,+\infty)$

Fix $(t_0, x_0) \in \mathbb{R}_+ \times (\partial\Omega \setminus \Gamma)$ and let $\xi \in \mathbf{C}^1(V;\mathbb{R})$ be such that $\xi(t_0,x_0) = \varphi(t_0,x_0)$ and $\xi(t,x) \leq \varphi(t,x)$ for every $(t,x) \in V$, where V is a neighborhood of (t_0,x_0) in $[-1,+\infty) \times \bar{\Omega}$. Assume, to obtain a contradiction, that $\nabla\xi(t_0,x_0) \cdot \mathbf{n}(x_0) < 0$. In particular, we have $\nabla\xi(t_0,x_0) \neq 0$.

Let $\gamma \in \mathrm{Opt}(k,t_0,x_0)$, u be an optimal control for γ, and $T = \varphi(t_0,x_0)$. Since $x_0 \notin \Gamma$, we have $T > 0$ and, by Corollary 4.9, we have $\gamma \in \mathbf{C}^1([t_0,t_0+T];\bar{\Omega})$ and $u \in \mathrm{Lip}([t_0,t_0+T];\mathbb{S}^{d-1})$.

Let $\sigma : [-1,+\infty) \to \mathbb{R}^d$ denote the solution of the initial value problem

$$
\begin{cases}
\dot{\sigma}(t) = -k(t,\sigma(t))\dfrac{\nabla\xi(t_0,x_0)}{|\nabla\xi(t_0,x_0)|}, \\[2mm]
\sigma(t_0) = x_0.
\end{cases}
$$

Since k is continuous, by construction, $\sigma \in \mathbf{C}^1([-1,+\infty),\mathbb{R}^d)$. We claim that there exists $\epsilon \in (0,1]$ such that $\sigma(t) \in \bar{\Omega}$ and $(t,\sigma(t)) \in V$ for every $t \in [t_0-\epsilon,t_0]$. Indeed, let W be as in Proposition 4.1(b) and notice that, since $x_0 \in \partial\Omega$, there exists $\epsilon_0 \in (0,1]$ such that $\sigma(t) \in W$ for every $t \in [t_0-\epsilon_0,t_0+\epsilon_0]$. Moreover, up to reducing ϵ_0, we also have $(t,\sigma(t)) \in V$ for every $t \in [t_0-\epsilon_0,t_0+\epsilon_0]$. Consider the map $\beta : t \mapsto d_{\partial\Omega}^{\pm}(\sigma(t))$, which is Lipschitz continuous on $[t_0-\epsilon_0,t_0+\epsilon_0]$. By differentiating, we have, for a.e. $t \in [t_0-\epsilon_0,t_0+\epsilon_0]$,

$$
\dot{\beta}(t) = \dot{\sigma}(t) \cdot \nabla d_{\partial\Omega}^{\pm}(\sigma(t)) = -k(t,\sigma(t))\nabla d_{\partial\Omega}^{\pm}(\sigma(t)) \cdot \frac{\nabla\xi(t_0,x_0)}{|\nabla\xi(t_0,x_0)|}. \tag{4.15}
$$

Recalling that $\mathbf{n}(x_0) = \nabla d_{\partial\Omega}^{\pm}(x_0)$ and using the continuity of $\nabla d_{\partial\Omega}$, we deduce from the fact that $\nabla\xi(t_0,x_0) \cdot \mathbf{n}(x_0) < 0$ that there exists $\epsilon \in (0,\epsilon_0]$ such that $\nabla d_{\partial\Omega}^{\pm}(\sigma(t)) \cdot \nabla\xi(t_0,x_0) < 0$ for every $t \in [t_0-\epsilon,t_0+\epsilon]$. Hence, by (4.15), we have $\dot{\beta}(t) > 0$ for a.e. $t \in [t_0-\epsilon,t_0+\epsilon]$. Since $\beta(t_0) = 0$, we thus have $\beta(t) \leq 0$ for $t \in [t_0-\epsilon,t_0]$, which implies that $\sigma(t) \in \bar{\Omega}$ for every $t \in [t_0-\epsilon,t_0]$. Since $\epsilon \in (0,\epsilon_0]$, we also have in particular that $(t,\sigma(t)) \in V$ for every $t \in [t_0-\epsilon,t_0]$.

Define $\tilde{\gamma} : [-1,+\infty) \to \bar{\Omega}$ by

$$
\tilde{\gamma}(t) = \begin{cases}
\sigma(t_0-\epsilon), & \text{if } t \in [-1,t_0-\epsilon), \\
\sigma(t), & \text{if } t \in [t_0-\epsilon,t_0), \\
\gamma(t), & \text{if } t \geq t_0.
\end{cases}
$$

Hence $\tilde{\gamma}$ is an admissible trajectory for k and thus, by Proposition 3.1(c), for every $h \in [0,\epsilon]$, we have $\varphi(t_0,x_0) \geq \varphi(t_0-h,\tilde{\gamma}(t_0-h)) - h$, and thus $\xi(t_0,x_0) \geq \xi(t_0-h,\tilde{\gamma}(t_0-h)) - h$. Notice also that, for $h \in [0,\epsilon]$, we have

$\tilde{\gamma}(t_0 - h) = \sigma(t_0 - h)$, yielding that $\xi(t_0, x_0) \geq \xi(t_0 - h, \sigma(t_0 - h)) - h$. Since ξ and σ are $\mathbf{C}^1$ functions, for $h \in [0, \epsilon]$, we have

$$\xi(t_0 - h, \sigma(t_0 - h)) = \xi(t_0, x_0) - h\partial_t\xi(t_0, x_0) - h\nabla\xi(t_0, x_0) \cdot \dot{\sigma}(t_0) + o(h).$$

Inserting this fact in the inequality $\xi(t_0, x_0) \geq \xi(t_0 - h, \sigma(t_0 - h)) - h$ for $h \in (0, \epsilon]$, using the definition of σ, dividing by h and letting $h \to 0^+$, we deduce that

$$- \partial_t\xi(t_0, x_0) + k(t_0, x_0)|\nabla\xi(t_0, x_0)| - 1 \leq 0. \tag{4.16}$$

Since $\gamma \in \mathrm{Opt}(k, t_0, x_0)$, we have, by Proposition 3.1(c), that $\varphi(t_0, x_0) = \varphi(t_0 + h, \gamma(t_0 + h)) + h$ for every $h \in [0, T]$, and thus $\xi(t_0, x_0) \geq \xi(t_0 + h, \gamma(t_0 + h)) + h$. We also have, for $h \in [0, T]$,

$$\xi(t_0 + h, \gamma(t_0 + h)) = \xi(t_0, x_0) + h\partial_t\xi(t_0, x_0) + h\nabla\xi(t_0, x_0) \cdot \dot{\gamma}(t_0^+) + o(h),$$

where we use the fact that $\gamma \in \mathbf{C}^1([t_0, t_0 + T]; \bar{\Omega})$ and $\dot{\gamma}(t_0^+)$ denotes the limit of $\dot{\gamma}(t)$ as $t \to t_0^+$. Using the fact that γ satisfies the first equation of (2.1), we deduce as before, dividing by h and letting $h \to 0^+$, that

$$\partial_t\xi(t_0, x_0) + k(t_0, x_0)\nabla\xi(t_0, x_0) \cdot u(t_0) + 1 \leq 0.$$

Adding with (4.16), we deduce that $\nabla\xi(t_0, x_0) \cdot u(t_0) + |\nabla\xi(t_0, x_0)| \leq 0$ and, since $u(t_0) \in \bar{B}$, this implies that $u(t_0) = -\frac{\nabla\xi(t_0, x_0)}{|\nabla\xi(t_0, x_0)|}$.

Consider now the function $\alpha : t \mapsto d^{\pm}_{\partial\Omega}(\gamma(t))$ defined on $[t_0, t_0 + T]$. Since $\gamma(t_0) = x_0 \in \partial\Omega$, there exists $\delta_0 \in (0, T]$ such that $\gamma(t) \in W$ for every $t \in [t_0, t_0 + \delta_0]$, and thus α is $\mathbf{C}^1$ on $[t_0, t_0 + \delta_0]$, with

$$\dot{\alpha}(t) = \nabla d^{\pm}_{\partial\Omega}(\gamma(t)) \cdot \dot{\gamma}(t) = k(t, \gamma(t))\nabla d^{\pm}_{\partial\Omega}(\gamma(t)) \cdot u(t).$$

In particular, $\dot{\alpha}(t_0) = -k(t_0, x_0)\mathbf{n}(x_0) \cdot \frac{\nabla\xi(t_0, x_0)}{|\nabla\xi(t_0, x_0)|} > 0$, showing that there exists $\delta \in (0, \delta_0]$ such that $\alpha(t) > \alpha(t_0) = 0$ for every $t \in (t_0, t_0 + \delta]$. From the definition of α, this means that $\gamma(t) \notin \bar{\Omega}$ for $t \in (t_0, t_0 + \delta]$, which is a contradiction since $\gamma \in \mathrm{Opt}(k, t_0, x_0)$. This contradiction establishes that $\nabla\xi(t_0, x_0) \cdot \mathbf{n}(x_0) \geq 0$, as required. $\square$

Remark 4.12. Under the assumptions of Theorem 4.11, if $(t_0, x_0) \in \mathbb{R}_+ \times (\partial\Omega \setminus \Gamma)$ and φ is differentiable at (t_0, x_0), then the inequality $\nabla\varphi(t_0, x_0) \cdot \mathbf{n}(x_0) \geq 0$ holds in the classical sense.

Proposition 3.4 asserts that the value function φ is a viscosity solution of the Hamilton–Jacobi equation (3.8) in $\mathbb{R}_+ \times (\Omega \setminus \Gamma)$ and thus, as a consequence of standard properties of viscosity solutions, (3.8) is satisfied in the classical sense at all points $(t, x) \in \mathbb{R}_+ \times (\Omega \setminus \Gamma)$ at which φ is differentiable. Thanks to Theorem 4.11, we can prove that (3.8) is actually satisfied at all points $(t, x) \in \mathbb{R}_+ \times (\bar{\Omega} \setminus \Gamma)$ at which φ is differentiable.

Proposition 4.13. *Consider the optimal control problem* $\mathrm{OCP}(k)$ *and its value function* φ *under assumptions* (H1)–(H3), (H4), *and* (H5). *Let* $(t_0, x_0) \in \mathbb{R}_+ \times (\bar{\Omega} \setminus \Gamma)$ *be such that* φ *is differentiable at* (t_0, x_0). *Then*

$$-\partial_t\varphi(t_0, x_0) + |\nabla\varphi(t_0, x_0)|k(t_0, x_0) - 1 = 0.$$

Proof. If $x_0 \in \Omega \setminus \Gamma$, this fact is a consequence of classical properties of viscosity solutions, and so we assume in the sequel that $x_0 \in \partial\Omega \setminus \Gamma$. By Remark 4.12, we have that $\nabla\varphi(t_0, x_0) \cdot \mathbf{n}(x_0) \geq 0$. Moreover, by Proposition 3.5(a), we have $\nabla\varphi(t_0, x_0) \neq 0$. Since, by Proposition 3.4, φ is a viscosity supersolution of (3.8) in $\mathbb{R}_+ \times (\bar{\Omega} \setminus \Gamma)$, we have in particular that

$$-\partial_t\varphi(t_0, x_0) + |\nabla\varphi(t_0, x_0)|k(t_0, x_0) - 1 \geq 0.$$

We are thus left to show the converse inequality.

Let D be the set of unit vectors pointing to the inside of $\bar{\Gamma}$ at x_0, *i.e.*, D is the set of $u_0 \in \mathbb{S}^{d-1}$ for which there exists $h_0 > 0$ such that $x_0 + h u_0 \in \bar{\Omega}$ for every $h \in [0, h_0]$. Notice that $\{u_0 \in \mathbb{S}^{d-1} \mid u_0 \cdot \mathbf{n}(x_0) < 0\} \subset D \subset \{u_0 \in \mathbb{S}^{d-1} \mid u_0 \cdot \mathbf{n}(x_0) \leq 0\}$ and thus, in particular, $\bar{D} = \{u_0 \in \mathbb{S}^{d-1} \mid u_0 \cdot \mathbf{n}(x_0) \leq 0\}$. Since $\nabla\varphi(t_0, x_0) \cdot \mathbf{n}(x_0) \geq 0$, we deduce that $-\frac{\nabla\varphi(t_0,x_0)}{|\nabla\varphi(t_0,x_0)|} \in \bar{D}$.

Let $u \in D$ and $\sigma \in \mathbf{C}^1(\mathbb{R}_+; \mathbb{R}^d)$ be the unique solution of

$$
\begin{cases}
\dot{\sigma}(t) = k(t, \sigma(t))u \\
\sigma(t_0) = x_0.
\end{cases}
\tag{4.17}
$$

Since $u \in D$, there exists $h_0 > 0$ such that $\sigma(t) \in \bar{\Omega}$ for every $t \in [t_0, t_0 + h_0]$. Let $\gamma \in \mathrm{Adm}(k)$ be defined by $\gamma(t) = x_0$ for $t \in [0, t_0]$, $\gamma(t) = \sigma(t)$ for $t \in [t_0, t_0 + h_0]$, and $\gamma(t) = \sigma(t_0 + h_0)$ for $t \geq t_0 + h_0$. By Proposition 3.1(c), we have, for every $h \in [0, h_0]$,

$$
\varphi(t_0, x_0) \leq h + \varphi(t_0 + h, \sigma(t_0 + h)).
\tag{4.18}
$$

One the other hand, notice that $\sigma(t_0 + h) = x_0 + h k(t_0, x_0)u + o(h)$ as $h \to 0$. Since φ is differentiable at (t_0, x_0) and Lipschitz continuous, we deduce that, as $h \to 0^+$,

$$
\varphi(t_0 + h, \sigma(t_0 + h)) - \varphi(t_0, x_0) = h \partial_t \varphi(t_0, x_0) + h k(t_0, x_0) \nabla\varphi(t_0, x_0) \cdot u + o(h).
$$

Dividing the above expression by h, using (4.18) and letting $h \to 0^+$, we get

$$
-1 \leq \partial_t \varphi(t_0, x_0) + k(t_0, x_0) \nabla\varphi(t_0, x_0) \cdot u.
$$

Since this holds for every $u \in D$ and the right-hand side of the above inequality is continuous in u, we get that

$$
\partial_t \varphi(t_0, x_0) + k(t_0, x_0) \inf_{u \in \bar{D}} \nabla\varphi(t_0, x_0) \cdot u + 1 \geq 0.
$$

Since $-\frac{\nabla\varphi(t_0,x_0)}{|\nabla\varphi(t_0,x_0)|} \in \bar{D}$, the above infimum is attained at $u = -\frac{\nabla\varphi(t_0,x_0)}{|\nabla\varphi(t_0,x_0)|}$, yielding that

$$
-\partial_t \varphi(t_0, x_0) + |\nabla\varphi(t_0, x_0)| k(t_0, x_0) - 1 \leq 0,
$$

as required. $\qquad\square$

Remark 4.14. At the light of Proposition 4.13, a natural question is whether φ is also a viscosity subsolution of (3.8) in $\mathbb{R}_+ \times (\bar{\Omega} \setminus \Gamma)$, since this would imply the result of Proposition 4.13. However, this may fail to be the case, as illustrated by the following example. Consider the case $d = 1$, $\Omega = (0, 1)$, $\Gamma = \{0\}$, and $k : \mathbb{R}_+ \times [0, 1] \to \mathbb{R}_+$ given by $k(t, x) = 1$ for every $(t, x) \in \mathbb{R}_+ \times [0, 1]$. One easily computes that $\varphi(t, x) = x$ for every $(t, x) \in \mathbb{R}_+ \times [0, 1]$. For every $\alpha \in (-\infty, 1]$, the function $\xi : \mathbb{R}_+ \times [0, 1]$ defined for $(t, x) \in \mathbb{R}_+ \times [0, 1]$ by $\xi(t, x) = 1 + \alpha(x - 1)$ satisfies $\xi(t, 1) = \varphi(t, 1) = 1$ and $\xi(t, x) \geq \varphi(t, x)$ for every $(t, x) \in \mathbb{R}_+ \times [0, 1]$. However, for $\alpha < -1$ and $t \in \mathbb{R}_+$, we have

$$
-\partial_t \xi(t, 1) + |\partial_x \xi(t, 1)| k(t, 1) - 1 = |\alpha| - 1 > 0,
$$

showing that φ cannot be a viscosity subsolution of (3.8) at $(t, 1)$.

4.3. Characterization of optimal controls

We now turn to the problem of characterizing the optimal control $u : \mathbb{R}_+ \to \bar{B}$ associated with an optimal trajectory $\gamma \in \mathrm{Opt}(k, t_0, x_0)$. By Proposition 3.5(b), the optimal control u can be written as $u(t) = -\frac{\nabla\varphi(t,\gamma(t))}{|\nabla\varphi(t,\gamma(t))|}$ for every $t \in [t_0, t_0 + \varphi(t_0, x_0))$ such that φ is differentiable at $(t, \gamma(t))$. For optimal control problems without state constraints, one can typically prove that φ is a semiconcave function (see, *e.g.*, [17], Thm. 8.2.7) and use this fact and properties of superdifferentials of semiconcave functions to deduce that φ is indeed differentiable along optimal trajectories, except possibly at their starting and ending times (see, *e.g.*, [17], Thm. 8.4.6).

Since value functions of optimal control problems with state constraints may fail to be semiconcave, we propose in this section an alternative way to characterize the optimal control. In particular, we are also able to provide a characterization of the optimal control under weaker assumptions: semiconcavity results usually require the dynamics of the system to be $\mathbf{C}^{1,1}$ with respect to the state, as in Theorem 8.2.7 of [17], but the Lipschitz continuity assumption (H5) is actually sufficient for our strategy. The approach we follow here is an adaptation to the case with state constraints of the characterization of optimal controls from [55]. We start by introducing the set of optimal directions at a point.

Definition 4.15. Consider the optimal control problem $\mathrm{OCP}(k)$ and assume that (H1)–(H3), (H4), and (H5) hold. Let $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$. We define the set $\mathcal{U}(t_0, x_0)$ of *optimal directions* at (t_0, x_0) as the set of $u_0 \in \mathbb{S}^{d-1}$ for which there exists $\gamma \in \mathrm{Opt}(k, t_0, x_0)$ such that the corresponding optimal control $u \in \mathrm{Lip}([t_0, t_0 + \varphi(t_0, x_0)], \mathbb{S}^{d-1})$ associated with γ satisfies $u(t_0) = u_0$.

Notice that, by Proposition 3.1(a) and Corollary 4.9, the set $\mathcal{U}(t_0, x_0)$ is non-empty whenever $x_0 \in \bar{\Omega} \setminus \Gamma$. We now show that $\mathcal{U}(t_0, x_0)$ is singleton along optimal trajectories, except possibly at their initial and final points. This result was already presented in Proposition 4.11 of [50] for optimal control problems without state constraints and with the stronger assumption that $k \in \mathrm{Lip}(\mathbb{R}_+ \times \bar{\Omega}; \mathbb{R}_+)$, but the proof presented in that reference also applies to our setting. We provide the proof here for completeness and also since it is an easy and interesting consequence of Corollary 4.9.

Proposition 4.16. *Consider the optimal control problem $\mathrm{OCP}(k)$ and its value function φ and assume that* (H1)–(H3), (H4), *and* (H5) *hold. Let $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$ and $\gamma \in \mathrm{Opt}(k, t_0, x_0)$. Then, for every $t \in (t_0, t_0 + \varphi(t_0, x_0))$, the set $\mathcal{U}(t, \gamma(t))$ contains exactly one element.*

Proof. Let $u : \mathbb{R}_+ \to \bar{B}$ be the optimal control associated with γ, $t_1 \in (t_0, t_0 + \varphi(t_0, x_0))$, and $x_1 = \gamma(t_1)$. Since $x_1 \in \bar{\Omega} \setminus \Gamma$, we have $\mathcal{U}(t_1, x_1) \neq \varnothing$. Moreover, by Proposition 3.1(c), we have $\varphi(t_1, x_1) + t_1 - t_0 = \varphi(t_0, x_0)$ and, noticing that $\tau(t_1, \gamma) = t_0 + \varphi(t_0, x_0) - t_1 = \varphi(t_1, x_1)$, we deduce that the restriction of γ to $[t_1, t_0 + \varphi(t_0, x_0)]$ (extended by constant values to $\mathbb{R}_+$) is an optimal trajectory for $\mathrm{OCP}(k)$ from (t_1, x_1), and thus $u(t_1) \in \mathcal{U}(t_1, x_1)$.

We conclude the proof by showing that, for every $u_1 \in \mathcal{U}(t_1, x_1)$, we have $u_1 = u(t_1)$. For that purpose, take $u_1 \in \mathcal{U}(t_1, x_1)$. By definition of $\mathcal{U}(t_1, x_1)$, there exists $\tilde{\gamma} \in \mathrm{Opt}(k, t_1, x_1)$ such that its associated control $\tilde{u}$ satisfies $\tilde{u}(t_1) = u_1$. Consider the trajectory $\hat{\gamma} : \mathbb{R}_+ \to \bar{\Omega}$ defined by $\hat{\gamma}(t) = \gamma(t)$ for $t \leq t_1$ and $\hat{\gamma}(t) = \tilde{\gamma}(t)$ for $t \geq t_1$, and notice that its associated control $\hat{u}$ satisfies $\hat{u}(t) = u(t)$ for $t < t_1$ and $\hat{u}(t) = \tilde{u}(t)$ for $t > t_1$. Then $\hat{\gamma} \in \mathrm{Adm}(k)$ and, by construction, $\tau(t_0, \hat{\gamma}) = t_1 - t_0 + \tau(t_1, \tilde{\gamma}) = t_1 - t_0 + \varphi(t_1, x_1) = \varphi(t_0, x_0)$, which yields that $\hat{\gamma} \in \mathrm{Opt}(k, t_0, x_0)$. In particular, by Corollary 4.9, we have $\hat{u} \in \mathrm{Lip}([t_0, t_0 + \varphi(t_0, x_0)]; \mathbb{S}^{d-1})$, and thus $\hat{u}(t_1) = \lim_{t \to t_1^-} \hat{u}(t) = u(t_1)$ and $\hat{u}(t_1) = \lim_{t \to t_1^+} \hat{u}(t) = \tilde{u}(t_1) = u_1$, yielding that $u_1 = u(t_1)$, as required. $\square$

Similarly to [55], the goal of this section is to characterize $\mathcal{U}(t_0, x_0)$ as the set of directions of maximal descent of the value function φ of $\mathrm{OCP}(k)$. For that purpose, we first introduce the notion of descent rate of φ along a given direction.

Definition 4.17. Consider the optimal control problem $\mathrm{OCP}(k)$ and its value function φ and assume that (H1)–(H3), (H4), and (H5) hold. Let $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$.

(a) The set of *inward pointing directions* at x_0 is the set $\text{In}(x_0)$ of vectors $u_0 \in \mathbb{S}^{d-1}$ for which there exists $h_0 > 0$ such that $x_0 + hu_0 \in \bar{\Omega}$ for every $h \in [0, h_0]$.

(b) The *descent rate* of φ at (t_0, x_0) is the function $\Delta_{(t_0,x_0)} : \text{In}(x_0) \to \mathbb{R}$ defined for $u_0 \in \text{In}(x_0)$ by

$$\Delta_{(t_0,x_0)}(u_0) = \limsup_{h \to 0^+} \frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h}.$$

Note that $\text{In}(x_0) = \mathbb{S}^{d-1}$ if $x_0 \in \Omega$, while, thanks to (H3), for $x_0 \in \partial\Omega$, we have

$$\{u_0 \in \mathbb{S}^{d-1} \mid u_0 \cdot \mathbf{n}(x_0) < 0\} \subset \text{In}(x_0) \subset \{u_0 \in \mathbb{S}^{d-1} \mid u_0 \cdot \mathbf{n}(x_0) \leq 0\}.$$

In general, both inclusions may be strict. As a consequence of those inclusions, we have

$$\overline{\text{In}}(x_0) = \{u_0 \in \mathbb{S}^{d-1} \mid u_0 \cdot \mathbf{n}(x_0) \leq 0\}.$$

The next proposition provides important properties of the descent rate function.

Proposition 4.18. *Consider the optimal control problem* $\text{OCP}(k)$ *and assume that* (H1)–(H3), (H4), *and* (H5) *hold. Let* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$.

(a) *For every* $u_0 \in \text{In}(x_0)$, *we have* $\Delta_{(t_0,x_0)}(u_0) \geq -1$.

(b) *The function* $\Delta_{(t_0,x_0)}$ *is Lipschitz continuous on* $\text{In}(x_0)$, *with a Lipschitz constant depending only on* $K_{\max}$ *and the Lipschitz constant of* φ, *and thus independent of* (t_0, x_0).

Proof. To show (a), take $u_0 \in \text{In}(x_0)$ and let $\sigma \in \mathbf{C}^1(\mathbb{R}_+; \mathbb{R}^d)$ be the unique solution of

$$\begin{cases} \dot{\sigma}(t) = k(t, \sigma(t))u_0, \\ \sigma(t_0) = x_0. \end{cases}$$

Since $u_0 \in \text{In}(x_0)$, there exists $t_* > t_0$ such that $\sigma(t) \in \bar{\Omega}$ for every $t \in [t_0, t_*]$. Define $\gamma \in \text{Adm}(k)$ by

$$\gamma(t) = \begin{cases} x_0, & \text{if } t \in [0, t_0], \\ \sigma(t), & \text{if } t \in [t_0, t_*], \\ \sigma(t_*), & \text{if } t \geq t_*. \end{cases}$$

Applying Proposition 3.1(c) to γ, we deduce that $\varphi(t_0 + h, \gamma(t_0 + h)) + h \geq \varphi(t_0, x_0)$ for every $h \geq 0$, and thus, for $h \in (0, t_* - t_0]$, we have

$$\frac{\varphi(t_0 + h, \sigma(t_0 + h)) - \varphi(t_0, x_0)}{h} \geq -1.$$

Since $\sigma(t_0 + h) = x_0 + hk(t_0, x_0)u_0 + o(h)$ as $h \to 0^+$ and φ is Lipschitz continuous, we deduce that $\varphi(t_0 + h, \sigma(t_0 + h)) = \varphi(t_0 + h, x_0 + hk(t_0, x_0)u_0) + o(h)$, yielding that, as $h \to 0^+$,

$$\frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} \geq -1 + o(1), \tag{4.19}$$

and the conclusion of (a) follows from the definition of $\Delta_{(t_0,x_0)}$.

In order to prove (b), take $u_1, u_2 \in \mathrm{In}(x_0)$ and denote by $C > 0$ the Lipschitz constant of φ. For $h > 0$ small enough, we have

$$
\begin{aligned}
\frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_1) - \varphi(t_0, x_0)}{h} &= \frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_2) - \varphi(t_0, x_0)}{h} \\
&+ \frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_1) - \varphi(t_0 + h, x_0 + hk(t_0, x_0)u_2)}{h} \\
&\leq \frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_2) - \varphi(t_0, x_0)}{h} + CK_{\max}|u_1 - u_2|.
\end{aligned}
$$

Taking the $\limsup$ as $h \to 0^+$, we deduce that $\Delta_{(t_0, x_0)}(u_1) \leq \Delta_{(t_0, x_0)}(u_2) + CK_{\max}|u_1 - u_2|$. Since this holds for every $u_1, u_2 \in \mathrm{In}(x_0)$, we obtain that

$$
|\Delta_{(t_0, x_0)}(u_1) - \Delta_{(t_0, x_0)}(u_2)| \leq CK_{\max}|u_1 - u_2| \qquad \text{for every } u_1, u_2 \in \mathrm{In}(x_0),
$$

as required. $\square$

As a consequence of Proposition 4.18(b), $\Delta_{(t_0, x_0)} : \mathrm{In}(x_0) \to \mathbb{R}$ can be extended in a unique way to a Lipschitz continuous function defined on $\overline{\mathrm{In}}(x_0)$. In the sequel, by a slight abuse of notation, we use $\Delta_{(t_0, x_0)}$ to denote this Lipschitz continuous extension. Note that this extension still satisfies the assertions of Proposition 4.18.

We are now in position to provide the definition of the set of directions of maximal descent of the value function.

Definition 4.19. Consider the optimal control problem $\mathrm{OCP}(k)$ and its value function φ and assume that (H1)–(H3), (H4), and (H5) hold. Let $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$. We define the set $\mathcal{W}(t_0, x_0)$ of *directions of maximal descent of φ* at (t_0, x_0) by

$$
\mathcal{W}(t_0, x_0) = \Delta_{(t_0, x_0)}^{-1}(\{-1\}) = \{u_0 \in \overline{\mathrm{In}}(x_0) \mid \Delta_{(t_0, x_0)}(u_0) = -1\}.
$$

Notice that the term *maximal descent* is motivated by Proposition 4.18(a), since we are considering the elements $u_0 \in \overline{\mathrm{In}}(x_0)$ reaching the lower bound -1 on $\Delta_{(t_0, x_0)}$.

Remark 4.20. If $x_0 \in \Omega$ (or, more generally, if the set $\mathrm{In}(x_0)$ is closed), then $\mathcal{W}(t_0, x_0)$ is simply the set of $u_0 \in \mathbb{S}^{d-1}$ such that

$$
\lim_{h \to 0^+} \frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} = -1, \tag{4.20}
$$

which is the definition of $\mathcal{W}(t_0, x_0)$ provided previously in Definition 4.11 of [55] for optimal control problems without state constraints. Indeed, notice that, if $\mathrm{In}(x_0)$ is closed, then $u_0 \in \mathcal{W}(t_0, x_0)$ if and only if

$$
\limsup_{h \to 0^+} \frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} = -1. \tag{4.21}
$$

On the other hand, (4.19) also yields that

$$
\liminf_{h \to 0^+} \frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u) - \varphi(t_0, x_0)}{h} \geq -1
$$

for every $u \in \mathrm{In}(x_0)$, and hence (4.20) is equivalent to (4.21).

Our next result shows that, at the points (t_0, x_0) where φ is differentiable, $\mathcal{W}(t_0, x_0)$ contains a unique direction of maximal descent which, as one might expect, is equal to $-\frac{\nabla\varphi(t_0, x_0)}{|\nabla\varphi(t_0, x_0)|}$. This was already shown in Proposition 4.13 of [55] for optimal control problems without state constraints, and that proof also carries over to the present case thanks to Proposition 4.13. For sake of completeness, and since this proof is quite elementary, we provide it here.

Proposition 4.21. *Consider the optimal control problem* OCP(k) *and its value function* φ *and assume that* (H1)–(H3), (H4), *and* (H5) *hold. Let* $(t_0, x_0) \in \mathbb{R}_+ \times (\bar{\Omega} \setminus \Gamma)$ *be such that* φ *is differentiable at* (t_0, x_0). *Then*

$$\mathcal{W}(t_0, x_0) = \left\{ -\frac{\nabla\varphi(t_0, x_0)}{|\nabla\varphi(t_0, x_0)|} \right\}. \tag{4.22}$$

Proof. Since φ is differentiable at (t_0, x_0), we have, for every $u \in \mathrm{In}(x_0)$,

$$\Delta_{(t_0, x_0)}(u) = \partial_t\varphi(t_0, x_0) + k(t_0, x_0)\nabla\varphi(t_0, x_0) \cdot u. \tag{4.23}$$

By continuity, the above equality also holds for every $u \in \overline{\mathrm{In}}(x_0)$. We also have, by Proposition 4.13, that

$$-\partial_t\varphi(t_0, x_0) + k(t_0, x_0)|\nabla\varphi(t_0, x_0)| - 1 = 0. \tag{4.24}$$

Moreover, recall that $k(t_0, x_0) > 0$ and, by Proposition 3.5(a), $\nabla\varphi(t_0, x_0) \neq 0$. If $u_0 \in \mathcal{W}(t_0, x_0)$, then $\Delta_{(t_0, x_0)}(u_0) = -1$ and, combining with (4.23) and (4.24), we deduce that

$$k(t_0, x_0)\left[|\nabla\varphi(t_0, x_0)| + \nabla\varphi(t_0, x_0) \cdot u_0\right] = 0,$$

which yields $u_0 = -\frac{\nabla\varphi(t_0, x_0)}{|\nabla\varphi(t_0, x_0)|}$ since $u_0 \in \mathbb{S}^{d-1}$. Conversely, defining $u_0 = -\frac{\nabla\varphi(t_0, x_0)}{|\nabla\varphi(t_0, x_0)|}$, Remark 4.12 ensures that $u_0 \in \overline{\mathrm{In}}(x_0)$ and it is immediate to compute, using (4.23) and (4.24), that $\Delta_{(t_0, x_0)}(u_0) = -1$, showing that $u_0 \in \mathcal{W}(t_0, x_0)$. $\square$

The reason why we go through the definition of descent rate $\Delta_{(t_0, x_0)}$ in order to define $\mathcal{W}(t_0, x_0)$ instead of the more direct definition provided in Definition 4.11 of [55] is that, if $x_0 \in \partial\Omega$ and $u_0 \in \mathbb{S}^{d-1}$, one might have $x_0 + hk(t_0, x_0)u_0 \notin \bar{\Omega}$ for every $h > 0$ small enough. Since φ is defined only in the set $\mathbb{R}_+ \times \bar{\Omega}$, this means that the term $\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_0)$ is not well-defined for any $h > 0$ small enough, and thus the limit in the left-hand side of (4.20) does not make sense. An alternative approach, however, is to replace φ in (4.20) by the value function φ_ϵ of the penalized optimal control problem OCP$_\epsilon(k_\epsilon)$ defined in Section 4.1. This is the subject of our next definition, which coincides with Definition 4.11 of [55] but with φ is replaced by φ_ϵ whenever its argument might fail to belong to its domain $\mathbb{R}_+ \times \bar{\Omega}$.

Definition 4.22. Consider the optimal control problems OCP(k) and OCP$_\epsilon(k_\epsilon)$ and their respective value functions φ and φ_ϵ and assume that (H1)–(H3), (H4), and (H5) hold. Let $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$. For $\epsilon > 0$, we define the set $\mathcal{W}_\epsilon(t_0, x_0)$ of *directions of maximal descent of* φ_ϵ *at* (t_0, x_0) by

$$\mathcal{W}_\epsilon(t_0, x_0) = \left\{ u_0 \in \mathbb{S}^{d-1} \;\middle|\; \lim_{h \to 0^+} \frac{\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} = -1 \right\}.$$

Our next result concerns the relation between optimal directions (*i.e.*, elements of $\mathcal{U}(t_0, x_0)$) and directions of maximal descent of φ (*i.e.*, elements of $\mathcal{W}(t_0, x_0)$), and asserts that both notions actually coincide, and that they also coincide with $\mathcal{W}_\epsilon(t_0, x_0)$ for $\epsilon > 0$ small enough. The fact that $\mathcal{U}(t_0, x_0) = \mathcal{W}(t_0, x_0)$ was already established in Theorem 4.14 of [55] for minimal-time optimal control problems without state constraints, and

the proof we provide here follows the same lines, but passes through the optimal control problem $\mathrm{OCP}_\epsilon(k_\epsilon)$ in order to handle the state constraints appropriately.

Theorem 4.23. *Consider the optimal control problems* $\mathrm{OCP}(k)$ *and* $\mathrm{OCP}_\epsilon(k_\epsilon)$ *and assume that* (H1)–(H3), (H4), *and* (H5) *hold. There exists* $\epsilon_0 > 0$ *such that, for every* $\epsilon \in (0, \epsilon_0)$ *and* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$, *we have* $\mathcal{U}(t_0, x_0) = \mathcal{W}(t_0, x_0) = \mathcal{W}_\epsilon(t_0, x_0)$.

Proof. We first remark that, if $x_0 \in \Gamma$, then $\mathcal{U}(t_0, x_0) = \mathcal{W}(t_0, x_0) = \mathcal{W}_\epsilon(t_0, x_0) = \varnothing$, and so we are only left to consider the case $x_0 \in \bar{\Omega} \setminus \Gamma$. We let $\epsilon_0 > 0$ be as in the statement of Propositions 4.7 and 4.10 and Theorem 4.8, and we fix $\epsilon \in (0, \epsilon_0)$. We proceed by proving that $\mathcal{U}(t_0, x_0) \subset \mathcal{W}(t_0, x_0) \subset \mathcal{W}_\epsilon(t_0, x_0) \subset \mathcal{U}(t_0, x_0)$.

Part I: Proof of the inclusion $\mathcal{U}(t_0, x_0) \subset \mathcal{W}(t_0, x_0)$. Let $u_0 \in \mathcal{U}(t_0, x_0)$ and take $\gamma \in \mathrm{Opt}(k, t_0, x_0)$ and $u : \mathbb{R}_+ \to \bar{B}$ an optimal control associated with γ such that $u \in \mathrm{Lip}([t_0, t_0 + \varphi(t_0, x_0)]; \mathbb{S}^{d-1})$ and $u(t_0) = u_0$. By Proposition 3.1(c), we have, for every $h \in (0, \varphi(t_0, x_0)]$, that

$$\frac{\varphi(t_0 + h, \gamma(t_0 + h)) - \varphi(t_0, x_0)}{h} = -1. \tag{4.25}$$

We claim that $u_0 \in \overline{\mathrm{In}}(x_0)$. Indeed, this is trivial if $x_0 \in \Omega$, and, if $x_0 \in \partial\Omega$, since γ takes values in $\bar{\Omega}$ and $\dot{\gamma}(t_0) = k(t_0, x_0)u_0$, one can easily check that $u_0 \cdot \mathbf{n}(x_0) \le 0$, implying that $u_0 \in \overline{\mathrm{In}}(x_0)$.

Let $(u_n)_{n\in\mathbb{N}}$ be a sequence in $\mathrm{In}(x_0)$ such that $u_n \to u_0$ as $n \to +\infty$. Fix $n \in \mathbb{N}$ and notice that, as $h \to 0^+$, we have

$$\begin{aligned}
\gamma(t_0 + h) &= x_0 + hk(t_0, x_0)u_0 + o(h) \\
&= x_0 + hk(t_0, x_0)u_n + hk(t_0, x_0)(u_0 - u_n) + o(h).
\end{aligned}$$

Recall that φ is Lipschitz continuous (Prop. 3.3) and denote by $M > 0$ its Lipschitz constant. Using the fact that $x_0 + hk(t_0, x_0)u_n \in \bar{\Omega}$ for $h > 0$ small enough, we deduce that, as $h \to 0^+$,

$$|\varphi(t_0 + h, \gamma(t_0 + h)) - \varphi(t_0 + h, x_0 + hk(t_0, x_0)u_n)| \le hMK_{\max}|u_n - u_0| + o(h).$$

Combining with (4.25) and letting $h \to 0$, we deduce that, for every $n \in \mathbb{N}$,

$$\Delta_{(t_0, x_0)}(u_n) = \limsup_{h\to 0^+} \frac{\varphi(t_0 + h, x_0 + hk(t_0, x_0)u_n) - \varphi(t_0, x_0)}{h} \le -1 + MK_{\max}|u_n - u_0|.$$

Hence, letting $n \to +\infty$ and using Proposition 4.18, we deduce that $\Delta_{(t_0, x_0)}(u_0) = -1$, showing that $u_0 \in \mathcal{W}(t_0, x_0)$, as required.

Part II: Proof of the inclusion $\mathcal{W}(t_0, x_0) \subset \mathcal{W}_\epsilon(t_0, x_0)$. Let $u_0 \in \mathcal{W}(t_0, x_0)$ and consider a sequence $(u_n)_{n\in\mathbb{N}}$ in $\mathrm{In}(x_0)$ such that $u_n \to u_0$ as $n \to +\infty$. By Proposition 4.10, φ_ϵ is Lipschitz continuous on $\mathbb{R}_+ \times \bar{\Omega}_{\epsilon/2}$, and we denote by $C_\epsilon > 0$ a Lipschitz constant of this map. Notice that, for $h \in \left[0, \frac{\epsilon}{2K_{\max}}\right]$, we have $d_\Omega(x_0 + hk(t_0, x_0)u_n) \le \frac{\epsilon}{2}$ for every $n \in \mathbb{N} \cup \{0\}$. Hence, for $h \in \left(0, \frac{\epsilon}{2K_{\max}}\right]$ and $n \in \mathbb{N}$, we have

$$\frac{\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} \le \frac{\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_n) - \varphi(t_0, x_0)}{h} + C_\epsilon K_{\max}|u_n - u_0|.$$

For $n \in \mathbb{N}$, since $u_n \in \mathrm{In}(x_0)$, we have $x_0 + hk(t_0, x_0)u_n \in \bar{\Omega}$ for $h > 0$ small enough, and thus, by Theorem 4.8, we have $\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_n) = \varphi(t_0 + h, x_0 + hk(t_0, x_0)u_n)$ for h small enough. Letting $h \to 0^+$, we

thus deduce that

$$\limsup_{h \to 0^+} \frac{\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} \leq \Delta_{(t_0, x_0)}(u_n) + C_\epsilon K_{\max}|u_n - u_0|.$$

Taking now the limit as $n \to +\infty$, we obtain that

$$\limsup_{h \to 0^+} \frac{\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} \leq -1. \tag{4.26}$$

Let $\gamma : \mathbb{R}_+ \to \mathbb{R}^d$ be defined by $\gamma(t) = x_0$ for $t \in [0, t_0]$ and as the solution of the differential equation $\dot{\gamma}(t) = k_\epsilon(t, \gamma(t))u_0$ with initial condition $\gamma(t_0) = x_0$ for $t \geq t_0$. Then $\gamma \in \mathrm{Adm}_\epsilon(k_\epsilon)$ and, by the dynamic programming principle (4.3) for $\mathrm{OCP}_\epsilon(k_\epsilon)$, and using also Theorem 4.8, we have, for $h \geq 0$,

$$\frac{\varphi_\epsilon(t_0 + h, \gamma(t_0 + h)) - \varphi(t_0, x_0)}{h} \geq -1.$$

Using that $\gamma(t_0 + h) = x_0 + hk(t_0, x_0)u_0 + o(h)$ as $h \to 0^+$, $\gamma(t_0 + h) \in \bar{\Omega}_{\epsilon/2}$ for $h \geq 0$ small enough, and that φ_ϵ is Lipschitz continuous in $\mathbb{R}_+ \times \bar{\Omega}_{\epsilon/2}$, we deduce that, as $h \to 0^+$,

$$\frac{\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} \geq -1 + o(1).$$

Together with (4.26), this shows that

$$\lim_{h \to 0^+} \frac{\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_0) - \varphi(t_0, x_0)}{h} = -1,$$

and thus $u_0 \in \mathcal{W}_\epsilon(t_0, x_0)$, as required.

Part III: Proof of the inclusion $\mathcal{W}_\epsilon(t_0, x_0) \subset \mathcal{U}(t_0, x_0)$. Let $u_0 \in \mathcal{W}_\epsilon(t_0, x_0)$ and $h > 0$, which is implicitly always assumed to be small enough. Then, by definition of $\mathcal{W}_\epsilon(t_0, x_0)$, we have, as $h \to 0^+$,

$$\varphi_\epsilon(t_0 + h, x_0 + hk(t_0, x_0)u_0) = \varphi(t_0, x_0) - h + o(h). \tag{4.27}$$

Define $\gamma_0 : [t_0, t_0 + h] \to \mathbb{R}^d$ by

$$\begin{cases} \dot{\gamma}_0(t) = k_\epsilon(t, \gamma_0(t))u_0, \\ \gamma_0(t_0) = x_0. \end{cases} \tag{4.28}$$

Let $x_1^h = \gamma_0(t_0 + h)$ and $t_1^h = t_0 + h$. Since Γ is closed and $x_0 \notin \Gamma$, we have $x_1^h \notin \Gamma$ for $h > 0$ small enough. Let $\gamma_1^h \in \mathrm{Opt}_\epsilon(k_\epsilon, t_1^h, x_1^h)$ and u_1^h be the optimal control associated with γ_1^h, which satisfies $u_1^h \in \mathrm{Lip}([t_1^h, t_1^h + \varphi_\epsilon(t_1^h, x_1^h)]; \mathbb{S}^{d-1})$ by Proposition 4.5. Set $\bar{u}_1^h = u_1^h(t_1^h) \in \mathbb{S}^{d-1}$ and define $\bar{\gamma}_1^h : [t_1^h, t_1^h + h] \to \mathbb{R}^d$ by

$$\begin{cases} \dot{\bar{\gamma}}_1^h(t) = k_\epsilon(t, \bar{\gamma}_1^h(t))\bar{u}_1^h \\ \bar{\gamma}_1^h(t_1^h) = x_1^h. \end{cases} \tag{4.29}$$

Let us also set $t_2^h = t_1^h + h$, $x_2^h = \gamma_1^h(t_2^h)$ and $\bar{x}_2^h = \bar{\gamma}_1^h(t_2^h)$. We split the sequel of the proof in two cases.

Case 1. We assume in this case that $\lim_{h \to 0^+} \bar{u}_1^h = u_0$. Let $\hat{u}_1^h \in \mathrm{Lip}(\mathbb{S}^{d-1})$ be defined by $\hat{u}_1^h(t) = \bar{u}_1^h$ for $t \in [0, t_1^h]$, $\hat{u}_1^h(t) = u_1^h(t)$ for $t \in [t_1^h, t_1^h + \varphi_\epsilon(t_1^h, x_1^h)]$, and $\hat{u}_1^h(t) = u_1^h(t_1^h + \varphi_\epsilon(t_1^h, x_1^h))$ for $t \geq t_1^h + \varphi_\epsilon(t_1^h, x_1^h)$. Since γ_1^h and $\hat{u}_1^h$
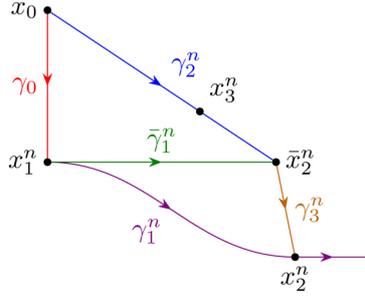
FIGURE 1. Illustration of the constructions used in the proof of Theorem 4.23.

are Lipschitz continuous and their Lipschitz constants do not depend on h (see Prop. 4.5), one deduces from Arzelà–Ascoli Theorem that there exist a positive sequence $(h_n)_{n \in \mathbb{N}}$ converging to 0 as $n \to +\infty$ and elements $\gamma^* \in \mathrm{Lip}_{K_{\max}}(\mathbb{R}^d)$ and $u^* \in \mathrm{Lip}(\mathbb{S}^{d-1})$ such that $\gamma_1^{h_n} \to \gamma^*$ and $\hat{u}_1^{h_n} \to u^*$ as $n \to +\infty$, uniformly on compact time intervals. Since $\gamma_1^h \in \mathrm{Opt}_\epsilon(k_\epsilon, t_1^h, x_1^h)$ for $h > 0$ and $t_1^h \to t_0$ and $x_1^h \to x_0$ as $h \to 0^+$, one can easily show, using the continuity of φ_ϵ, that $\gamma^* \in \mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0)$ and its corresponding optimal control coincides with u^* on $[t_0, t_0 + \varphi(t_0, x_0)]$. Moreover, since $x_0 \in \bar{\Omega}$, we have from Theorem 4.8 that $\mathrm{Opt}_\epsilon(k_\epsilon, t_0, x_0) = \mathrm{Opt}(k, t_0, x_0)$. In addition,

$$u^*(t_0) = \lim_{n \to +\infty} \hat{u}_1^{h_n}(t_1^{h_n}) = \lim_{n \to +\infty} \bar{u}_1^{h_n} = u_0,$$

which implies that $u_0 \in \mathcal{U}(t_0, x_0)$, as required.

Case 2. We now consider the case where $(\bar{u}_1^h)_{h>0}$ does not converge to u_0 as $h \to 0^+$, and we prove that this case is not possible. Let $\epsilon > 0$ and $(h_n)_{n \in \mathbb{N}}$ be a positive sequence such that $h_n \to 0$ as $n \to +\infty$ and $|\bar{u}_1^{h_n} - u_0| \geq \epsilon$ for every $n \in \mathbb{N}$. For simplicity, we set $t_1^{h_n} = t_1^n$, $x_1^{h_n} = x_1^n$, and similarly for all other variables whose upper index is h_n. In order to clarify the constructions used in this case, we illustrate them in Figure 1.

Notice that $|x_1^n - x_0| \leq K_{\max} h_n$, $|\gamma_1^n(t) - x_1^n| \leq K_{\max} h_n$ for every $t \in [t_1^n, t_1^n + h_n]$, and $|\bar{x}_2^n - x_1^n| \leq K_{\max} h_n$. Hence, for all $t \in [t_1^n, t_1^n + h_n]$, we have $x_1^n, \gamma_1^n(t), x_2^n, \bar{x}_2^n \in \bar{B}(x_0, 2K_{\max} h_n)$. Since $x_0 \in \bar{\Omega}$, for n large enough, this ball is included in $\bar{\Omega}_{\epsilon/2} = \{x \in \mathbb{R}^d \mid d_\Omega(x) \leq \frac{\epsilon}{2}\}$, and in particular $k_\epsilon(t, x) \geq \frac{1}{2} K_{\min}$ for every $(t, x) \in \mathbb{R}_+ \times \bar{B}(x_0, 2K_{\max} h_n)$.

Integrating (4.28) on $[t_0, t_1^n]$, we get

$$x_1^n - x_0 = \int_{t_0}^{t_1^n} k_\epsilon(s, \gamma_0(s)) \, \mathrm{d}s \, u_0,$$

and, proceeding similarly for (4.29), we get

$$\bar{x}_2^n - x_1^n = \int_{t_1^n}^{t_2^n} k_\epsilon(s, \bar{\gamma}_1^n(s)) \, \mathrm{d}s \, \bar{u}_1^n.$$

Denote the integrals in the right-hand side of the above equalities by I_0^n and I_1^n, respectively. We have

$$\begin{aligned}
|\bar{x}_2^n - x_0|^2 &= (I_0^n u_0 + I_1^n \bar{u}_1^n) \cdot (I_0^n u_0 + I_1^n \bar{u}_1^n) \\
&= (I_0^n)^2 + (I_1^n)^2 + 2 I_0^n I_1^n u_0 \cdot \bar{u}_1^n \\
&= |x_1^n - x_0|^2 + |\bar{x}_2^n - x_1^n|^2 + 2 I_0^n I_1^n u_0 \cdot \bar{u}_1^n.
\end{aligned}$$

We know that $|\bar{u}_1^n - u_0| \geq \epsilon$, which leads us to observe that there exists $\alpha \in (0,1)$ such that $u_0 \cdot \bar{u}_1^n < \alpha$ for every $n \in \mathbb{N}$. Thus

$$|\bar{x}_2^n - x_0|^2 < |x_1^n - x_0|^2 + |\bar{x}_2^n - x_1^n|^2 + 2\alpha I_0^n I_1^n.$$

Define

$$\rho := \sqrt{1 - (1-\alpha)\frac{K_{\min}^2}{8K_{\max}^2}},$$

then obviously $\rho < 1$ and

$$
\begin{aligned}
|\bar{x}_2^n - x_0|^2 &< (|x_1^n - x_0| + |\bar{x}_2^n - x_1^n|)^2 - 2(1-\alpha)I_0^n I_1^n \\
&= \left(1 - (1-\alpha)\frac{2I_0^n I_1^n}{(I_0^n + I_1^n)^2}\right)(|x_1^n - x_0| + |\bar{x}_2^n - x_1^n|)^2 \\
&\leq \rho^2 \left(|x_1^n - x_0| + |\bar{x}_2^n - x_1^n|\right)^2,
\end{aligned}
\tag{4.30}
$$

where we use that $I_i^n \in \left[\frac{1}{2}hK_{\min}, hK_{\max}\right]$ for $i \in \{1,2\}$. Let $u_2^n = \frac{\bar{x}_2^n - x_0}{|\bar{x}_2^n - x_0|}$ (with the convention $u_2^n = 0$ if $\bar{x}_2^n = x_0$) and define $\gamma_2^n : [t_0, t_0 + \tau^n] \to \mathbb{R}^d$ by

$$
\begin{cases}
\dot{\gamma}_2^n(t) = k_\epsilon(t, \gamma_2^n(t))u_2^n \\
\gamma_2^n(t_0) = x_0,
\end{cases}
\tag{4.31}
$$

where $\tau^n \geq 0$ is chosen so that[6] $\gamma_2^n(t_0 + \tau^n) = \bar{x}_2^n$.

Claim 4.24. *As $n \to +\infty$, we have $\tau^n \leq 2\rho h_n + o(h_n)$.*

Proof. Note that we have nothing to prove in the case $\bar{x}_2^n = x_0$, and hence we assume $\bar{x}_2^n \neq x_0$ in the sequel. If $|\bar{x}_2^n - x_0| \leq \rho|x_1^n - x_0|$, we let $x_3^n = \bar{x}_2^n$, otherwise we choose x_3^n as the unique point in the segment $(x_0, \bar{x}_2^n)$ such that $|x_3^n - x_0| = \rho|x_1^n - x_0|$. In both cases, we have $|x_3^n - x_0| = \bar{\rho}|x_1^n - x_0|$ for some $\bar{\rho} \leq \rho$. Let τ_1^n be the time that γ_2^n takes to reach the point x_3^n, i.e., $\gamma_2^n(t_0 + \tau_1^n) = x_3^n$. (Note that $\tau_1^n = \tau^n$ in the case $|\bar{x}_2^n - x_0| \leq \rho|x_1^n - x_0|$.)

We first show that $\tau_1^n \leq \rho h_n + o(h_n)$. To obtain that, we observe, by integrating (4.28) and (4.31) and doing a change of variables, that

$$
\begin{aligned}
\int_{t_0}^{t_0 + \tau_1^n} k_\epsilon(s, \gamma_2^n(s))\,\mathrm{d}s &= |x_3^n - x_0| = \bar{\rho}|x_1^n - x_0| = \bar{\rho}\int_{t_0}^{t_0+h_n} k_\epsilon(s, \gamma_0(s))\,\mathrm{d}s \\
&= \int_{t_0}^{t_0+\bar{\rho}h_n} k_\epsilon\left(t_0 + \frac{s - t_0}{\bar{\rho}}, \gamma_0\left(t_0 + \frac{s - t_0}{\bar{\rho}}\right)\right)\mathrm{d}s \\
&= \int_{t_0}^{t_0+\bar{\rho}h_n} k_\epsilon(s, \gamma_2^n(s))\,\mathrm{d}s \\
&\quad + \int_{t_0}^{t_0+\bar{\rho}h_n}\left[k_\epsilon\left(t_0 + \frac{s-t_0}{\bar{\rho}}, \gamma_0\left(t_0 + \frac{s-t_0}{\bar{\rho}}\right)\right) - k_\epsilon(s, \gamma_2^n(s))\right]\mathrm{d}s.
\end{aligned}
\tag{4.32}
$$

---

[6]This is possible since, for n large enough, all points in the segment from x_0 to $\bar{x}_2^n$ lie within $\Omega_{\epsilon/2}$, on which k_ϵ is lower bounded by $\frac{1}{2}K_{\min}$.

Let us show that

$$\lim_{n \to +\infty} \frac{1}{h_n} \int_{t_0}^{t_0+\bar\rho h_n} \left[k_\epsilon \left(t_0 + \frac{s-t_0}{\bar\rho}, \gamma_0 \left(t_0 + \frac{s-t_0}{\bar\rho} \right) \right) - k_\epsilon(s, \gamma_2^n(s)) \right] \mathrm{d}s = 0 \qquad (4.33)$$

Let $\delta > 0$. Since k_ϵ is continuous in (t_0, x_0), there exists $\eta > 0$ such that $|k_\epsilon(t, x) - k_\epsilon(t_0, x_0)| < \delta$ for every $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ satisfying $|t - t_0| < \eta$ and $|x - x_0| < \eta$. Since $h_n \to 0$ as $n \to +\infty$, there exists $N \in \mathbb{N}$ such that, for every $n \geq N$, we have $h_n < \eta$ and $K_{\max} h_n < \eta$. Noticing that, for every $s \in [t_0, t_0 + \bar\rho h_n]$, we have $\left| \frac{s-t_0}{\bar\rho} \right| \leq h_n < \eta$, $\left| \gamma_0 \left(t_0 + \frac{s-t_0}{\bar\rho} \right) - x_0 \right| \leq K_{\max} h_n < \eta$, $|s - t_0| \leq \bar\rho h_n < \eta$, and $|\gamma_2^n(s) - x_0| \leq \bar\rho K_{\max} h_n < \eta$, we deduce that, for $n \geq N$, we have, for every $s \in [t_0, t_0 + \bar\rho h_n]$,

$$k_\epsilon(t_0, x_0) - \delta < k_\epsilon \left(t_0 + \frac{s-t_0}{\bar\rho}, \gamma_0 \left(t_0 + \frac{s-t_0}{\bar\rho} \right) \right) < k_\epsilon(t_0, x_0) + \delta,$$
$$k_\epsilon(t_0, x_0) - \delta < k_\epsilon(s, \gamma_2^n(s)) < k_\epsilon(t_0, x_0) + \delta.$$

Subtracting those inequalities, integrating on s in $[t_0, t_0 + \bar\rho h_n]$, and dividing by h_n, we deduce that

$$\frac{1}{h_n} \left| \int_{t_0}^{t_0+\bar\rho h_n} \left[k_\epsilon \left(t_0 + \frac{s-t_0}{\bar\rho}, \gamma_0 \left(t_0 + \frac{s-t_0}{\bar\rho} \right) \right) - k_\epsilon(s, \gamma_2^n(s)) \right] \mathrm{d}s \right| < 2\bar\rho\delta,$$

concluding the proof of (4.33). Now, (4.32) and (4.33) imply that

$$\int_{t_0}^{t_0+\tau_1^n} k_\epsilon(s, \gamma_2^n(s)) \, \mathrm{d}s = \int_{t_0}^{t_0+\bar\rho h_n} k_\epsilon(s, \gamma_2^n(s)) \, \mathrm{d}s + o(h_n). \qquad (4.34)$$

Define $F : [0, \tau^n] \to \mathbb{R}_+$ by $F(t) = \int_{t_0}^{t_0+t} k_\epsilon(s, \gamma_2^n(s)) \, \mathrm{d}s$, then obviously F is continuous and increasing, which implies that F^{-1} is well-defined on the range of F. Since $\dot F(t) = k(t, \gamma_2^n(t))$, F is $K_{\max}$-Lipschitz continuous and, since $\frac{\mathrm{d}}{\mathrm{d}t} F^{-1}(t) = \frac{1}{\dot F(F^{-1}(t))}$, we also deduce that F^{-1} is $\frac{2}{K_{\min}}$-Lipschitz continuous. Therefore, by (4.34), we deduce that

$$\tau_1^n = F^{-1}(F(\bar\rho h_n) + o(h_n)) = \bar\rho h_n + o(h_n) \leq \rho h_n + o(h_n).$$

This concludes the proof of the claim in the case $|\bar x_2^n - x_0| \leq \rho |x_1^n - x_0|$, since $\tau_1^n = \tau^n$ in that case.

Otherwise, we have $\bar\rho = \rho$ and $|x_3^n - x_0| = \rho |x_1^n - x_0|$, and thus, from (4.30), we get

$$|\bar x_2^n - x_0| < \rho(|x_1^n - x_0| + |\bar x_2^n - x_1^n|) = |x_3^n - x_0| + \rho |\bar x_2^n - x_1^n|.$$

On the other hand, since x_3^n belongs to the segment $(x_0, \bar x_2^n)$, we have $|\bar x_2^n - x_0| = |\bar x_2^n - x_3^n| + |x_3^n - x_0|$, hence the inequality $|\bar x_2^n - x_3^n| \leq \rho |\bar x_2^n - x_1^n|$ holds. Suppose τ_2^n is the time the trajectory γ_2^n takes to go from x_3^n to $\bar x_2^n$, i.e., $\gamma_2^n(t_0 + \tau_1^n + \tau_2^n) = \bar x_2^n$, and note that $\tau^n = \tau_1^n + \tau_2^n$. As before, we compare the times between $|\bar x_2^n - x_3^n|$

and $|\bar{x}_2^n - x_1^n|$. Let $\beta \leq \rho$ be such that $|\bar{x}_2^n - x_3^n| = \beta|\bar{x}_2^n - x_1^n|$. Proceeding similarly to (4.32), we get

$$\int_0^{\tau_2^n} k_\epsilon(s + t_0 + \tau_1^n, \gamma_2^n(s + t_0 + \tau_1^n))\,\mathrm{d}s = |\bar{x}_2^n - x_3^n|$$

$$= \beta|\bar{x}_2^n - x_1^n| = \beta \int_{t_1^n}^{t_2^n} k_\epsilon(s, \bar{\gamma}_1^n(s))\,\mathrm{d}s$$

$$= \int_0^{\beta h_n} k_\epsilon\left(\frac{s}{\beta} + t_0 + h_n, \bar{\gamma}_1^n\left(\frac{s}{\beta} + t_0 + h_n\right)\right)\,\mathrm{d}s$$

$$= \int_0^{\beta h_n} k_\epsilon(s + t_0 + \tau_1^n, \gamma_2^n(s + t_0 + \tau_1^n))\,\mathrm{d}s$$

$$+ \int_0^{\beta h_n} \left[k_\epsilon\left(\frac{s}{\beta} + t_0 + h_n, \bar{\gamma}_1^n\left(\frac{s}{\beta} + t_0 + h_n\right)\right) - k_\epsilon(s + t_0 + \tau_1^n, \gamma_2^n(s + t_0 + \tau_1^n))\right]\,\mathrm{d}s.$$

Proceeding as in the proof of (4.33), we can show that the last integral in the above expression is an $o(h_n)$ as $n \to +\infty$, and thus

$$\int_{t_0}^{t_0 + \tau_2^n} k_\epsilon(s + \tau_1^n, \gamma_2^n(s + \tau_1^n))\,\mathrm{d}s = \int_{t_0}^{t_0 + \beta h_n} k_\epsilon(s + \tau_1^n, \gamma_2^n(s + \tau_1^n))\,\mathrm{d}s + o(h_n).$$

Defining $F(t) = \int_{t_0}^{t_0+t} k_\epsilon(s + \tau_1^n, \gamma_2^n(s + \tau_1^n))\,\mathrm{d}s$ and arguing similarly to above, we deduce that $\tau_2^n = \beta h_n + o(h_n)$. Therefore the time τ^n to reach $\bar{x}_2^n$ from x_0 satisfies

$$\tau^n = (\rho + \beta)h_n + o(h_n) \leq 2\rho h_n + o(h_n). \qquad \square$$

Let us now compare the trajectories $\bar{\gamma}_1^n$ and γ_1^n on $[t_1^n, t_2^n]$. Let $\delta_1^n(t) = \gamma_1^n(t) - \bar{\gamma}_1^n(t)$. Hence, from the ODEs satisfied by the trajectories $\bar{\gamma}_1^n$ and γ_1^n, we have

$$\delta_1^n(t) = \int_{t_1^n}^t \left[k_\epsilon(s, \gamma_1^n(s))u_1^n(s) - k_\epsilon(s, \bar{\gamma}_1^n(s))\bar{u}_1^n\right]\,\mathrm{d}s$$

$$= \int_{t_1^n}^t \left[k_\epsilon(s, \gamma_1^n(s)) - k_\epsilon(s, \bar{\gamma}_1^n(s))\right]u_1^n(s)\,\mathrm{d}s + \int_{t_1^n}^t k_\epsilon(s, \bar{\gamma}_1^n(s))(u_1^n(s) - \bar{u}_1^n)\,\mathrm{d}s.$$

Since u_1^n is the optimal control associated with γ_1^n, by Proposition 4.5, it is Lipschitz continuous in $[t_1^n, t_1^n + \varphi_\epsilon(t_1^n, x_1^n)]$ and its Lipschitz constant is independent of n. Therefore, denoting by $L_\epsilon > 0$ the Lipschitz constant of k_ϵ with respect to its second variable (which is independent of the first variable) and $C > 0$ the Lipschitz constant of u_1^n, we have

$$|\delta_1^n(t)| \leq L_\epsilon \int_{t_1^n}^t |\delta_1^n(s)|\,\mathrm{d}s + CK_{\max} \int_{t_1^n}^t |s - t_1^n|\,\mathrm{d}s,$$

and hence, by using Grönwall's inequality,

$$|\delta_1^n(t)| \leq CK_{\max}\frac{(t - t_1^n)^2}{2}e^{L_\epsilon(t - t_1^n)}.$$

In particular, if we set $t = t_1^n + h_n$, then

$$|x_2^n - \bar{x}_2^n| \le CK_{\max} \frac{h_n^2}{2} e^{L_\epsilon h_n} = O(h_n^2).$$

Let $u_3^n = \frac{x_2^n - \bar{x}_2^n}{|x_2^n - \bar{x}_2^n|}$ (with the convention $x_3^n = 0$ if $x_2^n = \bar{x}_2^n$) and γ_3^n be the solution of

$$\begin{cases} \dot{\gamma}_3^n(t) = k_\epsilon(t, \gamma_3^n(t))u_3^n \\ \gamma_3^n(t_0 + \tau^n) = \bar{x}_2^n. \end{cases} \tag{4.35}$$

Using the lower bound $\frac{1}{2}K_{\min}$ on k_ϵ and the fact that $|x_2^n - \bar{x}_2^n| = O(h_n^2)$, one can easily deduce that the time σ^n from $\bar{x}_2^n$ to x_2^n along γ_3^n (i.e., the value $\sigma^n > 0$ such that $\gamma_3^n(t_0 + \tau^n + \sigma^n) = x_2^n$) satisfies $\sigma^n = O(h_n^2)$.

We have thus constructed two ways to go from x_0 to x_2^n. The first one is to choose the path containing x_0, x_1^n, and x_2^n, which corresponds to the concatenation of the trajectories γ_0 on $[t_0, t_1^n]$ and γ_1^n on $[t_1^n, t_2^n]$, and the second one is the path containing x_0, $\bar{x}_2^n$, and x_2^n, which corresponds to the concatenation of the trajectories γ_2^n on $[t_0, t_0 + \tau^n]$ and γ_3^n on $[t_0 + \tau^n, t_0 + \tau^n + \sigma^n]$. Letting T_1^n and T_2^n be the times for going from x_0 to x_2^n along these two paths, respectively, we have, by construction and the claim, that $T_1^n = 2h_n$ and $T_2^n = \tau^n + \sigma^n \le 2\rho h_n + o(h_n)$. Hence, since $\rho < 1$, we have, for n large enough, that $T_2^n < T_1^n$.

By definition of γ_0, we have $x_1^n = \gamma_0(t_1^n) = x_0 + h_n k(t_0, x_0)u_0 + o(h_n)$ and, using (4.27) and the fact that φ_ϵ is Lipschitz continuous (Prop. 4.10), we deduce that

$$\varphi(t_0, x_0) = \varphi_\epsilon(t_1^n, x_1^n) + h_n + o(h_n).$$

Moreover, since $\gamma_1^n \in \mathrm{Opt}_\epsilon(k_\epsilon, t_1^n, x_1^n)$, we have from (4.3) that $\varphi_\epsilon(t_1^n, x_1^n) = \varphi_\epsilon(t_2^n, x_2^n) + h_n$, yielding that

$$\varphi(t_0, x_0) = \varphi(t_2^n, x_2^n) + T_1^n + o(h_n).$$

On the other hand, since the path from x_0 to x_2^n going through $\bar{x}_2^n$ is an admissible trajectory for k_ϵ, we have, by (4.3), that $\varphi(t_0, x_0) \le T_2^n + \varphi_\epsilon(t_0 + T_2^n, x_2^n)$. Hence

$$\varphi_\epsilon(t_2^n, x_2^n) + T_1^n + o(h_n) \le T_2^n + \varphi_\epsilon(t_0 + T_2^n, x_2^n). \tag{4.36}$$

We also know that $t_0 + T_2^n < t_0 + T_1^n = t_2^n$ for n large enough. Therefore[7], by Proposition 3.5(a), there exists a constant $c > 0$ such that

$$\varphi_\epsilon(t_2^n, x_2^n) > \varphi_\epsilon(t_0 + T_2^n, x_2^n) + (c-1)(t_2^n - t_0 - T_2^n) = \varphi_\epsilon(t_0 + T_2^n, x_2^n) + (c-1)(T_1^n - T_2^n),$$

and, using (4.36), we get $(c-1)(T_1^n - T_2^n) + T_1^n + o(h_n) \le T_2^n$, which leads to

$$2h_n + o(h_n) = T_1^n + o(h_n) \le T_2^n \le 2\rho h_n + o(h_n).$$

Divide above inequality by h_n to observe that

$$2 + o(1) \le 2\rho + o(1).$$

---

[7]Strictly speaking, Proposition 3.5(a) only applies to φ, and not to φ_ϵ. However, one can still get its conclusion by arguing as follows. Let $\widehat{\mathrm{OCP}}$ be defined as in the proof of Proposition 4.10 with $\eta = \frac{\epsilon}{2}$ and consider its value function $\widehat{\varphi}$. By the arguments provided in that proof, $\widehat{\varphi}$ satisfies Proposition 3.5(a) and $\widehat{\varphi}$ and φ_ϵ coincide in $\bar{\Omega}_{\epsilon/2}$. The conclusion then follows since all points involved here belong to $\bar{\Omega}_{\epsilon/2}$ for n large enough. Note that the constant $c > 0$ depends on ϵ.

Finally by letting $n \to +\infty$, we conclude that $\rho \geq 1$, which is a contradiction. Therefore Case 2 will never happen and this ends the proof. $\qquad\square$

We conclude this section with a technical result showing that, for $x \in \partial\Omega \setminus \Gamma$, if all directions of $\mathcal{W}(t, x)$ point to the inside of the domain, then no optimal trajectories starting at time 0 are close to x at time t. For that purpose, we introduce.

Proposition 4.25. *Consider the optimal control problem* $\mathrm{OCP}(k)$ *and assume that* (H1)–(H3), (H4), *and* (H5) *hold. Let* $x \in \partial\Omega \setminus \Gamma$, $t > 0$, *and assume that there exists* $w \in \mathcal{W}(t, x)$ *such that* $w \cdot \mathbf{n}(x) < 0$. *Then there exists a neighborhood* N *of* x *in* $\bar{\Omega}$ *such that, for every* $x_0 \in \bar{\Omega}$ *and* $\gamma \in \mathrm{Opt}(k, 0, x_0)$, *we have* $\gamma(t) \notin N$.

Proof. Assume, to obtain a contradiction, that there exist sequences $(x_{0,n})_{n \in \mathbb{N}}$ and $(\gamma_n)_{n \in \mathbb{N}}$ with $x_{0,n} \in \bar{\Omega}$ and $\gamma_n \in \mathrm{Opt}(k, 0, x_{0,n})$ for every $n \in \mathbb{N}$ and such that $\gamma_n(t) \to x$ as $n \to +\infty$. Since $\bar{\Omega}$ is compact and $(\gamma_n)_{n \in \mathbb{N}}$ is a sequence of functions which are $K_{\max}$-Lipschitz continuous, applying Arzelà–Ascoli Theorem, we deduce that there exist $x_0 \in \bar{\Omega}$ and $\gamma \in \mathrm{Lip}_{K_{\max}}(\bar{\Omega})$ such that, up to extracting subsequences (which we still denote by $(x_{0,n})_{n \in \mathbb{N}}$ and $(\gamma_n)_{n \in \mathbb{N}}$ for simplicity), we have, as $n \to +\infty$, that $x_{0,n} \to x_0$ and $\gamma_n \to \gamma$ uniformly on compact time intervals. By straightforward arguments based on the continuity of the value function, we deduce that $\gamma \in \mathrm{Opt}(k, 0, x_0)$, and in addition we have $\gamma(t) = x$.

Recall that, by Corollary 4.9, we have $\gamma \in \mathbf{C}^1([0, T]; \bar{\Omega})$, where $T = \varphi(0, x_0)$ and φ is the value function of $\mathrm{OCP}(k)$. Since $t > 0$ and $x \notin \Gamma$, we have $t \in (0, T)$, and thus γ is differentiable at t. Moreover, by Proposition 4.16 and Theorem 4.23, the set $\mathcal{W}(t, x)$ contains exactly one element, which we denote by $w_0 \in \mathbb{S}^{d-1}$, and thus, by assumption, we have $w_0 \cdot \mathbf{n}(x) < 0$. By Definition 4.15 and Theorem 4.23, we deduce also that $\dot{\gamma}(t) = k(t, x)w_0$.

Let $\alpha : h \mapsto d^{\pm}_{\partial\Omega}(\gamma(t + h))$ be defined on an open neighborhood of 0 in $\mathbb{R}$. Then $\alpha(0) = d^{\pm}_{\partial\Omega}(x) = 0$ and $\dot{\alpha}(0) = k(t, x)\mathbf{n}(x) \cdot w_0 < 0$. In particular, there exists $h_0 \in [-t, 0)$ such that, for every $h \in [h_0, 0)$, we have $\alpha(h) > 0$, and thus $\gamma(t + h) \notin \bar{\Omega}$, which contradicts the fact that γ takes values in $\bar{\Omega}$. This contradiction yields the conclusion. $\qquad\square$

4.4. The normalized gradient

Now that Theorem 4.23 characterizes the set of optimal directions $\mathcal{U}(t_0, x_0)$ as the set $\mathcal{W}(t_0, x_0)$ of directions of maximal descent of φ, we provide the following definition, which is motivated by Proposition 4.21.

Definition 4.26. Consider the optimal control problem $\mathrm{OCP}(k)$ and its value function φ and assume that (H1)–(H3), (H4), and (H5) hold. Let $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$. If $\mathcal{W}(t_0, x_0)$ contains exactly one element, we denote this element by $-\widehat{\nabla\varphi}(t_0, x_0)$, and call $\widehat{\nabla\varphi}(t_0, x_0)$ the *normalized gradient* of φ at (t_0, x_0).

As a consequence of Proposition 4.16 and Theorem 4.23, we immediately obtain the following characterization of optimal controls.

Corollary 4.27. *Consider the optimal control problem* $\mathrm{OCP}(k)$ *and its value function* φ *and assume that* (H1)–(H3), (H4), *and* (H5) *hold. Let* $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$, $\gamma \in \mathrm{Opt}(k, t_0, x_0)$, *and* $T = \varphi(t_0, x_0)$. *Then for every* $t \in (t_0, t_0 + T)$, φ *admits a normalized gradient at* $(t, \gamma(t))$ *and*

$$\dot{\gamma}(t) = -k(t, \gamma(t))\widehat{\nabla\varphi}(t, \gamma(t)).$$

Combining the above result with Corollary 4.9, we deduce that, for every optimal trajectory γ, the map $t \mapsto \widehat{\nabla\varphi}(t, \gamma(t))$ is Lipschitz continuous as long as $\gamma(t) \in \bar{\Omega} \setminus \Gamma$ and t is larger than the initial time of γ. However, this provides no information on the regularity of $(t, x) \mapsto \widehat{\nabla\varphi}(t, x)$. We are interested in proving the continuity of this map on its domain of definition. For that purpose, we first prove the upper semi-continuity of the set-valued map $\mathcal{U}$ (we refer to Definition 1.4.1 of [3] for the definition of upper semi-continuity for set-valued maps).

Proposition 4.28. *Consider the optimal control problem* OCP(k) *and assume that* (H1)–(H3)*,* (H4)*, and* (H5) *hold. Let* $\mathcal{U} : \mathbb{R}_+ \times \bar{\Omega} \rightrightarrows \mathbb{S}^{d-1}$ *be the set valued map introduced in Definition 4.15. Then* $\mathcal{U}$ *is upper semi-continuous.*

Proof. Since $\mathbb{S}^{d-1}$ is a compact set, it suffices to show that $\mathcal{U}$ has a closed graph (see, *e.g.*, [3], Prop. 1.4.8). Let $(t_n, x_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}_+ \times \bar{\Omega}$ converging as $n \to +\infty$ to some $(t_0, x_0) \in \mathbb{R}_+ \times \bar{\Omega}$, $(\bar{u}_n)_{n \in \mathbb{N}}$ be a sequence such that $\bar{u}_n \in \mathcal{U}(t_n, x_n)$ for every $n \in \mathbb{N}$ and $\bar{u}_n \to \bar{u}_0$ as $n \to +\infty$ for some $\bar{u}_0 \in \mathbb{S}^{d-1}$. Since $\bar{u}_n \in \mathcal{U}(t_n, x_n)$ for every $n \in \mathbb{N}$, there exists a sequence $(\gamma_n)_{n \in \mathbb{N}}$ of optimal trajectories with $\gamma_n \in \mathrm{Opt}(k, t_n, x_n)$ for every $n \in \mathbb{N}$ and a corresponding sequence of the associated optimal controls $(u_n)_{n \in \mathbb{N}}$ with $u_n \in \mathrm{Lip}([t_n, t_n + \varphi(t_n, x_n)]; \mathbb{S}^{d-1})$ and $u_n(t_n) = \bar{u}_n$ for every $n \in \mathbb{N}$. From Corollary 4.9, up to modifying u_n outside of the interval $[t_n, t_n + \varphi(t_n, x_n)]$, the sequences $(\gamma_n)_{n \in \mathbb{N}}$ and $(u_n)_{n \in \mathbb{N}}$ are sequences of Lipschitz continuous functions with Lipschitz constants independent of n, and thus, by Arzelà–Ascoli Theorem, one finds elements $\gamma^* \in \mathrm{Lip}(\bar{\Omega})$ and $u^* \in \mathrm{Lip}(\mathbb{S}^{d-1})$ such that, up to a subsequence, $\gamma_n \to \gamma^*$ and $u_n \to u^*$ uniformly on compact time intervals. In particular,

$$u^*(t_0) = \lim_{n \to \infty} u_n(t_n) = \lim_{n \to \infty} \bar{u}_n = \bar{u}_0$$

and

$$\gamma^*(t_0) = \lim_{n \to \infty} \gamma_n(t_n) = \lim_{n \to \infty} x_n = x_0.$$

By using the dynamic programming principle from Proposition 3.1(c) and the Lipschtiz continuity of the value function from Proposition 3.3, one observes that the restriction of u^* to $[t_0, t_0 + \varphi(t_0, x_0)]$ is the optimal control corresponding to the optimal trajectory γ^*. Hence $\bar{u}_0 = u^*(t_0) \in \mathcal{U}(t_0, x_0)$, concluding the proof that $\mathcal{U}$ has a closed graph. $\square$

On the set of points where a set-valued map is single-valued, upper semi-continuity coincides with standard continuity of single-valued functions. As an immediate consequence of this fact, Proposition 4.28, and Theorem 4.23, we have the following result.

Corollary 4.29. *Consider the optimal control problem* OCP(k) *and its value function* φ *and assume that* (H1)–(H3)*,* (H4)*, and* (H5) *hold. Let* $\mathcal{W} : \mathbb{R}_+ \times \bar{\Omega} \rightrightarrows \mathbb{S}^{d-1}$ *be the set valued map introduced in Definition 4.19 and* $\widehat{\nabla\varphi}$ *be the normalized gradient of* φ*. Then* $\mathcal{W}$ *is upper semi-continuous and* $\widehat{\nabla\varphi}$ *is a continuous function on the set where it is defined.*

5. THE MFG SYSTEM

We now turn to the characterization of equilibria $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$ of a mean field game MFG(K) through a system of PDEs, known as the MFG system, consisting of a continuity equation on the density of agents $t \mapsto m_t$, defined through the relation $m_t = e_{t\#}Q$, coupled with the Hamilton–Jacobi equation (3.8) on the value function φ of the optimal control problem OCP(k), where k is defined from K and Q by setting $k(t, x) = K(m_t, x)$. The main difficulty in this characterization lies within the continuity equation for m_t, and more precisely on the characterization of the corresponding velocity field. Corollary 4.27 suggests that such velocity field should be given by the opposite of the normalized gradient of φ multiplied by the dynamics k.

Corollary 4.29 states the continuity of the normalized gradient $\widehat{\nabla\varphi}$ on the set of points of $\mathbb{R}_+ \times \bar{\Omega}$ where it is defined. We will now show that, for the purposes of studying the equilibria of mean field games, this set is quite "large". More precisely, consider the mean field game MFG(K) and an equilibrium $Q \in \mathcal{P}(\mathbf{C}(\bar{\Omega}))$ for this game. We will prove that, for every fixed $t > 0$, the set of points $x \in \bar{\Omega} \setminus \Gamma$ at which $\widehat{\nabla\varphi}$ does not exist has m_t measure zero, where m_t is the evaluation at time t of the equilibrium measure Q.

For that purpose, let us introduce the set

$$\Upsilon = \left\{ (t, x) \in \mathbb{R}_+^* \times (\bar{\Omega} \setminus \Gamma) \mid \exists t_0 \in [0, t), \exists x_0 \in \bar{\Omega}, \exists \gamma \in \mathrm{Opt}(k, t_0, x_0) \text{ s.t. } \gamma(t) = x \right\}. \tag{5.1}$$

In other words, Υ contains all points $(t, x) \in \mathbb{R}_+^* \times (\bar{\Omega} \setminus \Gamma)$ which are strictly between the starting and the final points of an optimal trajectory. In particular, it follows from Corollary 4.27 that $\widehat{\nabla \varphi}(t, x)$ exists for every $(t, x) \in \Upsilon$, and, by Corollary 4.29, $\widehat{\nabla \varphi}$ is continuous in Υ. We also introduce, for $t > 0$, the set

$$\Upsilon_t = \left\{ x \in \bar{\Omega} \setminus \Gamma \mid (t, x) \in \Upsilon \right\}. \tag{5.2}$$

Proposition 5.1. *Consider the mean field game* $\mathrm{MFG}(K)$ *under the assumptions* (H1)–(H3), (H6), *and* (H7). *Let Q be an equilibrium for* $\mathrm{MFG}(K)$, *set $m_t = e_{t\#}Q$ for $t \geq 0$, define $k : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$ by $k(t, x) = K(m_t, x)$, consider the optimal control problem* $\mathrm{OCP}(k)$, *and let Υ and Υ_t be defined as in* (5.1) *and* (5.2), *respectively. Then for every $t > 0$, we have $m_t(\bar{\Omega} \setminus (\Gamma \cup \Upsilon_t)) = 0$.*

Proof. Let $\mathbf{Opt} = \bigcup_{x_0 \in \bar{\Omega}} \mathrm{Opt}(k, 0, x_0)$. Since Q is an equilibrium of $\mathrm{MFG}(K)$, then $Q(\mathbf{Opt}) = 1$. For every $t > 0$, from the definition of Υ, one has that $\{\gamma \in \mathbf{Opt} \mid \gamma(t) \in \bar{\Omega} \setminus (\Gamma \cup \Upsilon_t)\} = \varnothing$, and then $m_t(\bar{\Omega} \setminus (\Gamma \cup \Upsilon_t)) = Q\Big(\{\gamma \in \mathbf{Opt} \mid \gamma(t) \in \bar{\Omega} \setminus (\Gamma \cup \Upsilon_t)\}\Big) = Q(\varnothing) = 0.$ $\qquad\square$

We are now ready to provide our main result concerning the MFG system of $\mathrm{MFG}(K)$.

Theorem 5.2. *Consider the mean field game* $\mathrm{MFG}(K)$ *under assumptions* (H1)–(H3), (H6), *and* (H7). *Let $m_0 \in \mathcal{P}(\bar{\Omega})$, Q be an equilibrium of* $\mathrm{MFG}(K)$ *with initial condition m_0, $m_t \in \mathcal{P}(\bar{\Omega})$ be defined for $t > 0$ by $m_t = e_{t\#}Q$, $k : \mathbb{R}_+ \times \bar{\Omega} \to \mathbb{R}_+$ be defined from Q and K by $k(t, x) = K(e_{t\#}Q, x)$, φ be the value function of* $\mathrm{OCP}(k)$, *$\mathcal{W}$ be the set-valued map provided in Definition 4.19, and $\widehat{\nabla \varphi}$ be the normalized gradient of φ from Definition 4.26. For $t > 0$, set*

$$\partial \Omega_t^- = \{x \in \partial \Omega \setminus \Gamma \mid \exists w \in \mathcal{W}(t, x) \text{ such that } w \cdot \mathbf{n}(x) < 0\}.$$

Then (m, φ) solves the MFG system

$$\begin{cases} \partial_t m_t(x) - \mathrm{div}\left(m_t(x) K(m_t, x) \widehat{\nabla \varphi}(t, x) \right) = 0, & \text{in } \mathbb{R}_+^* \times (\bar{\Omega} \setminus \Gamma), \\ -\partial_t \varphi(t, x) + |\nabla \varphi(t, x)| K(m_t, x) - 1 = 0, & \text{in } \mathbb{R}_+ \times (\bar{\Omega} \setminus \Gamma), \\ m_t(x) = 0, & \text{for } t > 0 \text{ and } x \in \partial \Omega_t^-, \\ \varphi(t, x) = 0, & \text{on } \mathbb{R}_+ \times \Gamma, \\ \nabla \varphi(t, x) \cdot \mathbf{n}(x) \geq 0, & \text{on } \mathbb{R}_+ \times (\partial \Omega \setminus \Gamma), \\ m_t = m_0, & \text{in } \{0\} \times \bar{\Omega}, \end{cases} \tag{5.3}$$

where the first equation is satisfied in the sense of distributions, the second and fifth equations are satisfied in the viscosity senses of Proposition 3.4 and Theorem 4.11, respectively, and the third equation is satisfied in the following sense: for every $t > 0$ and $x \in \partial \Omega_t^-$, there exists a neighborhood N of x such that $m_t(N) = 0$.

Proof. Notice first that, since K satisfies (H6) and (H7), then k satisfies (H4) and (H5). The Hamilton–Jacobi equation on φ and its boundary conditions then follow immediately from Proposition 3.4 and Theorem 4.11. The initial condition on m_t follows from its definition.

Let us prove that m_t satisfies the continuity equation in (5.3). First, thanks to Corollary 4.29, $\widehat{\nabla \varphi}$ is continuous on the set Υ defined in (5.1). Let $\xi \in \mathbf{C}_c^\infty(\mathbb{R}_+^* \times (\bar{\Omega} \setminus \Gamma); \mathbb{R})$ be a test function. Take $\gamma \in \mathrm{Opt}(k, 0, x_0)$ and let $T = \varphi(0, x_0)$. By Corollary 4.27, one has $\dot{\gamma}(t) = -K(m_t, \gamma(t)) \widehat{\nabla \varphi}(t, \gamma(t))$ for every $t \in (0, T)$. Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\xi(t, \gamma(t))\Big) = \partial_t \xi(t, \gamma(t)) - \nabla_x \xi(t, \gamma(t)) \cdot \widehat{\nabla \varphi}(t, \gamma(t)) \, K(m_t, \gamma(t)). \tag{5.4}$$

Let us denote the set of all optimal trajectories by $\mathbf{Opt}$, *i.e.*, $\mathbf{Opt} = \bigcup_{x_0 \in \bar{\Omega}} \mathrm{Opt}(k, 0, x)$. Thanks to the continuity of the right-hand side of (5.4) on $\mathbb{R}_+^* \times \mathbf{Opt}$, one can integrate to observe that

$$
\int_0^\infty \int_{\mathbf{Opt}} \frac{\mathrm{d}}{\mathrm{d}t} \Big(\xi(t, \gamma(t)) \Big) \, \mathrm{d}Q(\gamma) \, \mathrm{d}t
$$
$$
= \int_0^\infty \int_{\mathbf{Opt}} \partial_t \, \xi(t, \gamma(t)) \, \mathrm{d}Q(\gamma) \, \mathrm{d}t
$$
$$
- \int_0^\infty \int_{\mathbf{Opt}} \nabla_x \, \xi(t, \gamma(t)) \cdot \widehat{\nabla \varphi}(t, \gamma(t)) \, K(m_t, \gamma(t)) \, \mathrm{d}Q(\gamma) \, \mathrm{d}t.
$$

Since ξ is compactly supported, the left-hand side of the above equality is zero. Hence, by using the Proposition 5.1 and the relation between m_t and Q, one concludes that

$$
\int_0^\infty \int_{\bar{\Omega}} \partial_t \xi(t, x) \, \mathrm{d}m_t(x) \, \mathrm{d}t = \int_0^\infty \int_{\bar{\Omega}} \nabla_x \, \xi(t, x) \cdot \widehat{\nabla \varphi}(t, x) \, K(m_t, x) \, \mathrm{d}m_t(x) \, \mathrm{d}t,
$$

which is precisely the weak formulation of the continuity equation in (5.3).

Finally, let us prove the boundary condition on m_t. Let $t > 0$ and $x \in \partial\Omega_t^-$. Then, using the definition of m_t and Proposition 4.25, there exists a neighborhood N of x such that

$$
m_t(N) = Q\left(\{\gamma \in \mathbf{Opt} \mid \gamma(t) \in N\}\right) = Q(\varnothing) = 0,
$$

as required. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 5.3. Note that both the continuity and the Hamilton–Jacobi equations in (5.3) are satisfied in the spatial domain $\bar{\Omega} \setminus \Gamma$ and, in particular, they are satisfied up to the boundary, excluding the target set Γ. The velocity field at a point $(t, x) \in \mathbb{R}_+^* \times (\partial\Omega \setminus \Gamma)$ in the boundary is thus $-K(m_t, x)\widehat{\nabla \varphi}(t, x)$, *i.e.*, the same as in the interior. Note also that the fifth equation in (5.3) shows that such a velocity field either points towards the inside of the domain or is tangent to it. In particular, contrarily to the mean field game system obtained in [13] for mean field games with state constraints, fixed final time, and more general cost functions, no correction to the velocity field is needed in the boundary of the domain in the present model due to the particular structure of the optimization criterion solved by each agent.

References

[1] Y. Achdou, P. Cardaliaguet, F. Delarue, A. Porretta and F. Santambrogio, Mean field games, vol. 2281 of *Lecture Notes in Mathematics*. Springer (2020).

[2] L. Ambrosio, N. Gigli and G. Savaré, Gradient flows in metric spaces and in the space of probability measures. *Birkhäuser Verlag, Basel* (2005) viii+333.

[3] J.-P. Aubin and H. Frankowska, Set-valued analysis, *Systems & Control: Foundations & Applications*, vol. 2. Birkhäuser Boston, Inc., Boston, MA (1990).

[4] R.J. Aumann, Markets with a continuum of traders. *Econometrica* **32** (1964) 39–50.

[5] R.J. Aumann and L.S. Shapley, Values of non-atomic games. Princeton University Press, Princeton, N.J. (1974), a Rand Corporation Research Study.

[6] S. Banach, Wstęp do teorii funkcji rzeczywistych, Monografie Matematyczne. Tom XVII.], Polskie Towarzystwo Matematyczne, Warszawa-Wrocław (1951).

[7] M. Bardi and I. Capuzzo-Dolcetta, Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA (1997).

[8] J.-D. Benamou, G. Carlier and F. Santambrogio, Variational mean field games, in Active particles. Vol. 1. Advances in theory, models, and applications. Model. Simul. Sci. Eng. Technol. Birkhäuser/Springer, Cham (2017) 141–171.

[9] N. Bourbaki, Topologie Générale. Chapitres 5 à 10, Éléments de Mathématique. Springer (2007).

[10] M. Burger, M. Di Francesco, P.A. Markowich and M.-T. Wolfram, On a mean field game optimal control approach modeling fast exit scenarios in human crowds, in 52nd IEEE Conference on Decision and Control, IEEE (2013).

[11] P. Cannarsa and R. Capuani, Existence and uniqueness for mean field games with state constraints, in PDE models for multi-agent phenomena. *Springer INdAM Ser.*, vol. 28. Springer, Cham (2018) 49–71.

[12] P. Cannarsa, R. Capuani and P. Cardaliaguet, $C^{1,1}$-smoothness of constrained solutions in the calculus of variations with application to mean field games. *Math. Eng.* **1** (2019) 174–203.

[13] P. Cannarsa, R. Capuani and P. Cardaliaguet, Mean field games with state constraints: from mild to pointwise solutions of the PDE system, *Calc. Var. Partial Differ. Equ.* **60** (2021) Paper No. 108, 33.

[14] P. Cannarsa and M. Castelpietra, Lipschitz continuity and local semiconcavity for exit time problems with state constraints. *J. Differ. Equ.* **245** (2008) 616–636.

[15] P. Cannarsa, M. Castelpietra and P. Cardaliaguet, Regularity properties of attainable sets under state constraints, in Geometric control and nonsmooth analysis. *Ser. Adv. Math. Appl. Sci.*, vol. 76, 120–135, World Sci. Publ., Hackensack, NJ (2008).

[16] P. Cannarsa and C. Mendico, Mild and weak solutions of mean field game problems for linear control systems. *Minimax Theory Appl.* **5** (2020) 221–250.

[17] P. Cannarsa and C. Sinestrari, Semiconcave functions, Hamilton-Jacobi equations, and optimal control. Vol. 58 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA (2004).

[18] P. Cardaliaguet, Notes on mean field games, https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf.

[19] P. Cardaliaguet, Weak solutions for first order mean field games with local coupling, in Analysis and geometry in control theory and its applications. *Springer INdAM Ser.*, vol. 11. Springer, Cham (2015) 111–158.

[20] P. Cardaliaguet, F. Delarue, J.-M. Lasry and P.-L. Lions, The master equation and the convergence problem in mean field games, vol. 201 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ (2019).

[21] P. Cardaliaguet, A.R. Mészáros and F. Santambrogio, First order mean field games with density constraints: pressure equals price. *SIAM J. Control Optim.* **54** (2016) 2672–2709.

[22] R. Carmona and F. Delarue, Probabilistic theory of mean field games with applications. I. vol. 83 of *Probability Theory and Stochastic Modelling*. Springer, Cham (2018).

[23] R. Carmona and F. Delarue, Probabilistic theory of mean field games with applications. II. Vol. 84 of *Probability Theory and Stochastic Modelling*. Springer, Cham (2018).

[24] R. Carmona and D. Lacker, A probabilistic weak formulation of mean field games and applications. *Ann. Appl. Probab.* **25** (2015) 1189–1231.

[25] F.H. Clarke, Optimization and nonsmooth analysis. *Classics in Applied Mathematics*, vol. 5, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1990).

[26] E. Cristiani, A. De Santo and M. Menci, A Generalized Mean-Field Game Model for the Dynamics of Pedestrians with Limited Predictive Abilities (2021). arXiv:2108.00086.

[27] E. Cristiani, B. Piccoli and A. Tosin, Multiscale modeling of granular flows with application to crowd dynamics. *Multiscale Model. Simul.* **9** (2011) 155–182.

[28] M.C. Delfour and J.-P. Zolésio, Shape analysis via oriented distance functions. *J. Funct. Anal.* **123** (1994) 129–201.

[29] R. Ducasse, G. Mazanti and F. Santambrogio, Second order local minimal-time mean field games. *NoDEA Nonlinear Differential Equations Appl.* **29** (2022) Paper No. 37, 32 pp.

[30] S. Dweik and G. Mazanti, Sharp semi-concavity in a non-autonomous control problem and L^p estimates in an optimal-exit MFG. *NoDEA Nonlinear Differ. Equ. Appl.* **27** (2020) Paper No. 11, 59 pp.

[31] M. Fischer and F.J. Silva, On the asymptotic nature of first order mean field games. *Appl. Math. Optim.* **84** (2021) 2327–2357.

[32] L. Gibelli and N. Bellomo (editors), Volume 1 of Crowd Dynamics. Springer International Publishing (2018).

[33] D.A. Gomes, E.A. Pimentel and V. Voskanyan, Regularity theory for mean-field game systems, SpringerBriefs in Mathematics, Springer, [Cham] (2016).

[34] D.A. Gomes and V.K. Voskanyan, Extended deterministic mean-field games. *SIAM J. Control Optim.* **54** (2016) 1030–1055.

[35] D. Helbing and P. Molnár, Social force model for pedestrian dynamics. *Phys. Rev. E* **51** (1995) 4282–4286.

[36] D. Helbing, I. Farkas, and T. Vicsek, Simulating dynamical features of escape panic. *Nature* **407** (2000) 487–490.

[37] L.F. Henderson, The statistics of crowd fluids. *Nature* **229** (1971) 381–383.

[38] M. Huang, P.E. Caines and R.P. Malhamé, Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions, in 42nd IEEE Conference on Decision and Control, 2003. Proceedings, vol. 1. IEEE (2003) 98–103.

[39] M. Huang, P.E. Caines and R.P. Malhamé, Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria. *IEEE Trans. Automat. Control* **52** (2007) 1560–1571.

[40] M. Huang, R.P. Malhamé and P.E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6** (2006) 221–251.

[41] R.L. Hughes, A continuum theory for the flow of pedestrians. *Transp. Res. B: Methodolog.* **36** (2002) 507–535.

[42] R.L. Hughes, The flow of human crowds, in Annual review of fluid mechanics. Vol. 35. *Annu. Rev. Fluid Mech.*, vol. 35, 169–182, Annual Reviews, Palo Alto, CA (2003).

[43] B. Jovanovic and R.W. Rosenthal, Anonymous sequential games. *J. Math. Econom.* **17** (1988) 77–87.

[44] A. Lachapelle and M.-T. Wolfram, On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transp. Res. Part B* **45** (2011) 1572–1589.

[45] D. Lacker, Mean field games via controlled martingale problems: existence of Markovian equilibria. *Stoch. Process. Appl.* **125** (2015) 2856–2894.

[46] J.-M. Lasry and P.-L. Lions, Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris* **343** (2006) 619–625.

[47] J.-M. Lasry and P.-L. Lions, Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris* **343** (2006) 679–684.

[48] J.-M. Lasry and P.-L. Lions, Mean field games. *Jpn. J. Math.* **2** (2007) 229–260.

[49] B. Maury and S. Faure, Crowds in equations, Advanced Textbooks in Mathematics. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ (2019).

[50] G. Mazanti and F. Santambrogio, Minimal-time mean field games. *Math. Models Methods Appl. Sci.* **29** (2019) 1413–1464.

[51] A. Muntean and F. Toschi (editors), Collective dynamics from bacteria to crowds. *CISM International Centre for Mechanical Sciences. Courses and Lectures*, vol. 553. Springer, Vienna (2014).

[52] B. Piccoli and A. Tosin, Time-evolving measures and macroscopic modeling of pedestrian flow. *Arch. Ration. Mech. Anal.* **199** (2011) 707–738.

[53] M.D. Rosini, Macroscopic models for vehicular flows and crowd dynamics: theory and applications, Understanding Complex Systems, Springer, Heidelberg (2013).

[54] S. Sadeghi Arjmand and G. Mazanti, On the characterization of equilibria of nonsmooth minimal-time mean field games with state constraints, in 2021 60th IEEE Conference on Decision and Control (CDC) (2021) 5300–5305.

[55] S. Sadeghi Arjmand and G. Mazanti, Multipopulation minimal-time mean field games. *SIAM J. Control Optim.* **60** (2022) 1942–1969.

[56] F. Santambrogio, Optimal transport for applied mathematicians. *Progress in Nonlinear Differential Equations and their Applications*, vol. 87. Birkhäuser/Springer, Cham (2015).

[57] F. Santambrogio and W. Shim, A Cucker-Smale inspired deterministic mean field game with velocity interactions. *SIAM J. Control Optim.* **59** (2021) 4155–4187.

[58] R. Vinter, Optimal control, Modern Birkhäuser Classics. Birkhäuser Boston, Ltd., Boston, MA (2010).