SHARP SEMI-CONCAVITY IN A NON-AUTONOMOUS CONTROL PROBLEM AND L^p ESTIMATES IN AN OPTIMAL-EXIT MFG

SAMER DWEIK AND GUILHERME MAZANTI

Abstract. This paper studies a mean field game inspired by crowd motion in which agents evolve in a compact domain and want to reach its boundary minimizing the sum of their travel time and a given boundary cost. Interactions between agents occur through their dynamic, which depends on the distribution of all agents.

We start by considering the associated optimal control problem, showing that semi-concavity in space of the corresponding value function can be obtained by requiring as time regularity only a lower Lipschitz bound on the dynamics. We also prove differentiability of the value function along optimal trajectories under extra regularity assumptions.

We then provide a Lagrangian formulation for our mean field game and use classical techniques to prove existence of equilibria, which are shown to satisfy a MFG system. Our main result, which relies on the semi-concavity of the value function, states that an absolutely continuous initial distribution of agents with an L^p density gives rise to an absolutely continuous distribution of agents at all positive times with a uniform bound on its L^p norm. This is also used to prove existence of equilibria under fewer regularity assumptions on the dynamics thanks to a limit argument.

Contents

Date: May 1, 2019.

2010 *Mathematics Subject Classification.* 91A13, 93C15, 49J15, 49N70, 35B65.

Key words and phrases. Mean field games, non-autonomous optimal control, semi-concavity of the value function, L^p estimate, MFG system.

This work was partially supported by a public grant as part of the “Investissement d’avenir” project, reference ANR-11-LABX-0056-LMH, LabEx LMH, PGMO project VarPDEMFG. The first author was also partially supported by the by the French ANR project “GEOMETRYA”, reference ANR-12-BS01-0014, and by the Région Ile-de-France. The second author was also partially supported by the French ANR project “MFG”, reference ANR-16-CE40-0015-01, and by the Hadamard Mathematics LabEx (LMH) through the grant number ANR-11-LABX-0056-LMH in the “Investissement d’avenir” project.

1. Introduction

Mean field games (MFGs for short) are differential games with a continuum of rational players, assumed to be indistinguishable, individually neglectable, and influenced only by the average behavior of other players through a mean-field type interaction. Introduced independently around 2006 by Jean-Michel Lasry and Pierre-Louis Lions [54–56] and by Peter E. Caines, Minyi Huang, and Roland P. Malhamé [46–48], these models have since been studied from several perspectives, including approximation of games with a large number of players by MFGs [21, 52], numerical methods for approximating MFG equilibria [1, 2, 27, 41], games with large time horizon [18, 23], variational mean field games [9, 24, 59, 61], games on graphs or networks [13, 14, 37, 42], or the characterization of equilibria using the master equation [10, 22, 28]. We refer to [19, 38, 43] for more details and further references on mean field games. The words "player" and "agent" are used interchangeably in this paper to refer to those taking part in a game.

This paper continues the analysis of the mean field game model introduced in [58], which considers players evolving in a compact domain $\Omega \subset \mathbb{R}^d$, their goal being to reach the boundary $\partial\Omega$. The distribution of players at time $t \geq 0$ is described by a Borel probability measure $\rho_t \in \mathcal{P}(\Omega)$ and, as in [58], we assume that the interaction between players occur through their dynamics, the trajectory $\gamma : [0, +\infty) \to \Omega$ of a given player being described by the control system $\gamma'(t) = k(\rho_t, \gamma(t))u(t)$, where the control $u : [0, +\infty) \to \mathbb{R}^d$ satisfies $|u(t)| \leq 1$ for every $t \geq 0$ and the function $k : \mathcal{P}(\Omega) \times \Omega \to [0, +\infty)$ describes the maximal speed $k(\mu, x)$ an agent may have when their position is x and agents are distributed according to μ. A player chooses their control u in order to minimize the sum of their travel time to $\partial\Omega$ with a boundary cost $g(z)$ on their arrival position $z \in \partial\Omega$, which is a generalization of the time-minimization criterion of [58].

The above mean field game is proposed as a simple model for crowd motion, in which the crowd, modeled macroscopically by the measures $\rho_t \in \mathcal{P}(\Omega)$, evolves in Ω and wishes to leave this domain through its boundary $\partial\Omega$ while optimizing exit time and position. Crowd motion has been extensively studied from a mathematical point of view, with a wide range of models being used to describe crowd behavior, ranging from Maxwell–Boltzmann models [45], particle systems [44], granular media [35], scalar conservation laws [29], time-varying measures [60], or models based on gradient flows [57]. Several works also address the question of controlling crowd behavior [4, 31]. Mean field games have already been used to model crowd motion, for instance in [9, 12, 24, 53], however these models differ from ours since they consider a fixed final time, identical for all agents, and no constraints on the control, which is instead penalized on the cost function. As detailed in [58], the model we consider here is also closely related to Hughes' model for crowd motion [49, 50] and our notion of equilibrium is related to the standard notion of Wardrop equilibria in non-atomic congestion games [25, 26, 30, 64].

The function k in our model is intended to represent congestion, i.e., the difficulty of moving in high-density areas. Several mean field games with congestion have been previously considered [3, 34, 36, 39, 42, 61], their common feature being to model congestion as a penalization in the cost function of each agent when passing through crowded regions. The penalization term is usually chosen as a negative power of the density, which introduces a singularity in the Hamilton–Jacobi equation of the corresponding optimal control problem. Our model considers instead that, in some crowd motion situations, an agent may not be able to move faster by simply paying some additional cost, since the congestion provoked by other agents may work as a physical barrier for the agent to increase their speed. Hence, we model congestion as a constraint on the maximal speed, given by k. Singularities on the Hamilton–Jacobi equation are avoided by assuming that k is upper bounded.

In order to properly model congestion, k should compute $k(\mu, x)$ by evaluating μ at or around x and giving as a result some non-increasing function of this evaluation, meaning that the maximal speed of an agent is a non-increasing function of some "average density" around x. This is the case, for instance, when k is given by

$$k(\mu, x) = V\left(\int_\Omega \chi(x - y)\psi(y)\,\mathrm{d}\mu(y)\right), \tag{1.1}$$

where $\chi : \mathbb{R}^d \to [0, +\infty)$ is a convolution kernel, $\psi : \mathbb{R}^d \to [0, +\infty)$ may serve as a weight on Ω or as a cut-off function to discount some part of Ω, and the non-increasing function $V : [0, +\infty) \to [0, +\infty)$ provides the maximal speed in terms of the average density computed by the integral. Even though the results of this paper do not assume a particular form for k, we make use of (1.1) to justify some of our assumptions in Section 4 and we also verify that our main results apply when k is given by (1.1) and V, χ, and ψ satisfy suitable regularity assumptions.

We are interested in describing equilibria of the above mean field game, which correspond, roughly speaking, to evolutions $t \mapsto \rho_t$ for which almost every agent satisfies their optimization criterion. Contrarily to the classical approach for mean field games consisting on describing equilibria in terms of ρ_t, we adopt here a Lagrangian approach, which amounts to describing the motion of agents as a measure on the set of all possible trajectories. This classical approach in optimal transport [7, 62, 63] has been used in some recent works on mean field games [9, 15, 20, 24, 58].

In order to analyze the above mean field game, we start by considering the corresponding optimal control problem. Assuming that k is a given function depending on time instead of the measure ρ_t, we start by obtaining some properties of optimal trajectories using classical optimal control techniques, such as Pontryagin Maximum Principle. We then prove our two main results for the value function φ of this optimal control problem: if the time derivative of k is lower bounded, then φ is semi-concave in space (Theorem 3.23) and, if k is $C^{1,1}$, then φ is differentiable along optimal trajectories (Theorem 3.31).

After this preliminary study of the optimal control problem, we turn to the analysis of the mean field game itself. We start by proving existence of equilibria in a Lagrangian setting and obtaining the corresponding MFG system of PDEs on ρ_t and the value function φ using arguments similar to [58]. We then prove that an absolutely continuous initial distribution of agents with an L^p density gives rise to an absolutely continuous distribution of agents at all positive times with a uniform bound on its L^p norm (Theorem 4.13) and use this result and a limit procedure to obtain existence of equilibria and the corresponding MFG system for a less regular model to which the arguments of [58] do not apply (Theorem 4.17).

The paper is organized as follows. Useful notations and definitions used throughout the paper are provided in Section 2. Section 3 considers the optimal control problem corresponding to our mean field game, remarking first on Sections 3.1 and 3.2 that classical results for autonomous systems can be easily extended to a time-dependent framework with very few assumptions on the time regularity of the dynamics, before proving our main results in Sections 3.3 and 3.4. The mean field game is finally considered in Section 4, with existence of equilibria and the MFG system being considered in Section 4.1, before the L^p estimates of Section 4.2 and the ensuing results for a less regular mean field game in Section 4.3.

2. Notations and definitions

Let us set the main notation used in this paper. We let $\mathbb{R}^+ = [0, +\infty)$ and denote the usual Euclidean scalar product and norm in $\mathbb{R}^d$ by $x \cdot y$ and $|x|$, respectively, for $x, y \in \mathbb{R}^d$. The closure, interior, convex hull, and diameter of a set $A \subset \mathbb{R}^d$ are denoted by $\overline{A}$, $\mathring{A}$, $\operatorname{conv} A$, and $\operatorname{diam}(A)$, respectively, with $\mathring{A}$ also denoted by $\operatorname{int} A$. The open and closed Euclidean balls of center x and radius r in $\mathbb{R}^d$ are denoted respectively by $B(x, r)$ and $\bar{B}(x, r)$.

For $A \subset B$, the function $\mathbb{1}_A : B \to \{0, 1\}$ denotes the characteristic function of A, i.e., $\mathbb{1}_A(x) = 1$ if and only if $x \in A$. Given two sets A, B, the notation $f : A \rightrightarrows B$ indicates that f is a set-valued map from A to B, i.e., f maps a point $x \in A$ to a subset $f(x) \subset B$. The maps $\Pi_t : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $\Pi_x : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ denote the canonical projections into the factors of the product $\mathbb{R} \times \mathbb{R}^d$.

If f is a function defined on (a subset of) $\mathbb{R} \times \mathbb{R}^d$, we use the notations Df, $\partial_t f$, and ∇f to denote its differential with respect to, respectively, both variables, its first variable, and its second variable. Similar notations are used for related notions, such as super and subdifferentials.

For a given metric space X, $\mathcal{M}(X)$ denotes the set of all Borel nonnegative measures on X endowed with the topology of weak convergence of measures, and $\mathcal{P}(X) \subset \mathcal{M}(X)$ denotes the subset of probability measures. The support of a measure $\eta \in \mathcal{M}(X)$ is denoted by $\operatorname{spt}(\eta)$. For $\eta \in \mathcal{M}(X)$ and $Y \subset X$ a Borel set, we denote by $\eta|_Y \in \mathcal{M}(Y)$ the restriction of η to Y. When $X \subset \mathbb{R}^d$ is Borel and $\eta \in \mathcal{M}(X)$ is absolutely continuous with respect to the Lebesgue measure, we use the same notation η for its density.

Let $\Omega \subset \mathbb{R}^d$ be a compact domain. We denote by $C(\mathbb{R}^+, \Omega)$ the space of all continuous curves from $\mathbb{R}^+$ to Ω, equipped with the topology of uniform convergence on compact sets, with respect to which $C(\mathbb{R}^+, \Omega)$ is a Polish space (see, for instance, [11, Chapter X]). Whenever needed, we fix a metric d on $C(\mathbb{R}^+, \Omega)$. Recall that, if $L > 0$ is fixed, the set of all L-Lipschitz continuous curves in $C(\mathbb{R}^+, \Omega)$ is compact thanks to Arzelà–Ascoli Theorem.

We recall that, for $B \subset \mathbb{R}^d$, a function $u : B \to \mathbb{R}$ is called *semi-concave* if it is continuous in B and there exists $C \geq 0$ such that, for every $x, h \in \mathbb{R}^d$ with $[x - h, x + h] \subset B$, one has

$$u(x - h) + u(x + h) - 2u(x) \leq C|h|^2.$$

The constant C is called a *semi-concavity constant* for u. When $A \subset \mathbb{R}$, $B \subset \mathbb{R}^d$, and $\varphi : A \times B \to \mathbb{R}$, we say that φ is semi-concave with respect to x, uniformly in t, if $x \mapsto \varphi(t, x)$ is semi-concave for every $t \in A$ with a semi-concavity constant independent of t.

We shall also need in this paper the classical notions of generalized gradients and some of their elementary properties, which we now recall, following the presentation from [17, Chapter 3].

Definition 2.1. Let $A \subset \mathbb{R}$, $B \subset \mathbb{R}^d$, $\varphi : A \times B \to \mathbb{R}$, and $(t, x) \in A \times B$. The sets

$$D^+\varphi(t, x) := \left\{(h, p) \in \mathbb{R} \times \mathbb{R}^d \,:\, \limsup_{(s,y) \to (t,x)} \frac{\varphi(s, y) - \varphi(t, x) - (h, p) \cdot (s - t, y - x)}{|(s - t, y - x)|} \leq 0\right\},$$

$$D^-\varphi(t, x) := \left\{(h, p) \in \mathbb{R} \times \mathbb{R}^d \,:\, \liminf_{(s,y) \to (t,x)} \frac{\varphi(s, y) - \varphi(t, x) - (h, p) \cdot (s - t, y - x)}{|(s - t, y - x)|} \geq 0\right\},$$

are called, respectively, the *superdifferential* and *subdifferential* of φ at (t, x). We also define the superdifferential and subdifferential of φ with respect to x by

$$\nabla^+\varphi(t, x) := \left\{p \in \mathbb{R}^d \,:\, \limsup_{y \to x} \frac{\varphi(t, y) - \varphi(t, x) - p \cdot (y - x)}{|y - x|} \leq 0\right\},$$

$$\nabla^- \varphi(t,x) := \left\{ p \in \mathbb{R}^d \,:\, \liminf_{y \to x} \frac{\varphi(t,y) - \varphi(t,x) - p \cdot (y-x)}{|y-x|} \geq 0 \right\},$$

respectively. Finally, we say that a vector $(h,p) \in \mathbb{R} \times \mathbb{R}^d$ is a *reachable gradient* of φ at $(t,x) \in A \times B$ if there is a sequence $\{(t_k,x_k)\}_k$ in $A \times B$ such that φ is differentiable at (t_k,x_k) for every $k \in \mathbb{N}$, and

$$\lim_{k\to\infty} (t_k,x_k) = (t,x), \qquad \lim_{k\to\infty} D\varphi(t_k,x_k) = (h,p).$$

The set of all reachable gradients of φ at (t,x) is denoted by $D^\star\varphi(t,x)$.

As a simple consequence of the definitions of $D^+\varphi$, $D^-\varphi$, $\nabla^+\varphi$, and $\nabla^-\varphi$, we have the inclusions $\Pi_x(D^+\varphi(t,x)) \subset \nabla^+\varphi(t,x)$ and $\Pi_x(D^-\varphi(t,x)) \subset \nabla^-\varphi(t,x)$. Moreover, if φ is Lipschitz continuous, then $D^\star\varphi(t,x)$ is a compact set: it is closed by definition and it is bounded since φ is Lipschitz. From Rademacher's theorem it follows that $D^\star\varphi(t,x) \neq \emptyset$ for every $(t,x) \in \overline{\operatorname{int}(A \times B)}$. We gather in the next proposition some classical additional properties for semi-concave functions (see, e.g., [17, Propositions 3.3.1 and 3.3.4, Theorem 3.3.6, and Lemma 3.3.16]).

Proposition 2.2. *Let $A \subset \mathbb{R}$, $B \subset \mathbb{R}^d$, $\varphi : A \times B \to \mathbb{R}$ be semi-concave, and $(t,x) \in \operatorname{int}(A \times B)$. Then*

(a) *$D^\star\varphi(t,x) \subset \partial D^+\varphi(t,x)$;*
(b) *$D^+\varphi(t,x) \neq \emptyset$;*
(c) *if $D^+\varphi(t,x)$ is a singleton, then φ is differentiable at (t,x);*
(d) *$D^+\varphi(t,x) = \operatorname{conv} D^\star\varphi(t,x)$;*
(e) *$\Pi_x(D^+\varphi(t,x)) = \nabla^+\varphi(t,x)$;*
(f) *if $C > 0$ is a semiconcavity constant for φ, a vector $p \in \mathbb{R}^d$ belongs to $\nabla^+\varphi(t,x)$ if and only if*

$$\varphi(t,y) - \varphi(t,x) - p \cdot (y-x) \leq C|y-x|^2$$

for every $y \in B$ such that $[x,y] \subset B$.

3. Exit-time optimal control problem

As a preliminary step for the study of our mean field game model, we consider in this section the optimal control problem solved by each agent of the game. We assume that each agent is submitted to a non-autonomous control system, the time-dependence of the dynamic being a consequence of the interaction between agents. The optimization criterion takes into account the time to reach a certain target set, considered as an *exit*, and a cost on the position at which the agent reaches the exit. For this reason, our optimal control problem is qualified as "exit-time". In our setting, all agents evolve in a given compact set $\Omega \subset \mathbb{R}^d$ and the exit is assumed to be $\partial\Omega$. Notice that the particular case where the cost on the exit position is identically zero corresponds to the problem of reaching the target set in minimal time.

We start the section by providing a precise definition of our optimal control problem and recalling some well-known facts, in particular concerning its value function, while also exploiting the consequences of the first order optimality conditions from Pontryagin Maximum Principle. We then turn to the two main results of this section. The first one concerns the semi-concavity of the value function with respect to the space variable x under weak assumptions on the smoothness of the dynamic with respect to time. Our second main result shows that the value function is differentiable along optimal trajectories, except possibly their endpoints.

3.1. **Definition, existence and first properties.** We consider control systems whose state equation is of the form

$$\begin{cases} \gamma'(t) = k(t, \gamma(t))u(t), & \text{for a.e. } t \geq t_0, \\ \gamma(t_0) = x_0, \end{cases} \tag{3.1}$$

where $\gamma(t) \in \mathbb{R}^d$ is the state, the continuous function $k : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^+$ is called the *dynamic* of the system, $t_0 \in \mathbb{R}^+$, $x_0 \in \mathbb{R}^d$, and $u : [t_0, \infty) \to \bar{B}(0, 1)$ is a measurable function (which is called a *control*).

We list some basic assumptions on the dynamic k:

$$0 < k_{\min} := \inf k \leq k_{\max} := \sup k < +\infty, \tag{H1}$$

$$\exists L_x > 0 \text{ such that } |k(t, x_1) - k(t, x_2)| \leq L_x |x_1 - x_2| \quad \text{for all } x_1, x_2 \in \mathbb{R}^d \text{ and } t \in \mathbb{R}^+. \tag{H2}$$

Notice that (H2) ensures the existence of a unique global solution to the state equation (3.1) for any choice of t_0, x_0 and u. We denote the solution of (3.1) by $\gamma^{t_0,x_0,u}$ and we call it an *(admissible) trajectory* of the system, corresponding to the initial condition $\gamma(t_0) = x_0$ and to the control u.

Let Ω be a compact domain in $\mathbb{R}^d$: for a given trajectory $\gamma = \gamma^{t_0,x_0,u}$ of (3.1), we set

$$\tau^{t_0,x_0,u} = \inf\{\tau \geq 0 \,:\, \gamma^{t_0,x_0,u}(t_0 + \tau) \in \partial\Omega\},$$

with the convention $\tau^{t_0,x_0,u} = +\infty$ if $\gamma^{t_0,x_0,u}(t_0 + \tau) \notin \partial\Omega$ for all $\tau \in \mathbb{R}^+$. This means that we consider $\partial\Omega$ as the target set. We call $\tau^{t_0,x_0,u}$ the *exit time* of the trajectory. If $\tau^{t_0,x_0,u} < +\infty$, we set for simplicity

$$\gamma_\tau^{t_0,x_0,u} := \gamma^{t_0,x_0,u}(t_0 + \tau^{t_0,x_0,u})$$

to denote the point where the trajectory reaches the target $\partial\Omega$. As $k_{\min} > 0$, one can see easily that, for every $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$, there is always some control u such that $\tau^{t_0,x_0,u} < +\infty$.

An optimal control problem consists of choosing the control strategy u in the state equation (3.1) in order to minimize a given functional. Let $g : \partial\Omega \to \mathbb{R}^+$ be a given continuous function. For every $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$, we minimize the cost

$$\tau^{t_0,x_0,u} + g(\gamma_\tau^{t_0,x_0,u}) \tag{3.2}$$

among all controls u. A control u and the corresponding trajectory $\gamma^{t_0,x_0,u}$ are called *optimal* for the point x_0 at time t_0 if u minimizes (3.2). Remark that optimal controls $u : [t_0, \infty) \to \bar{B}(0, 1)$ are arbitrary for $t > t_0 + \tau^{t_0,x_0,u}$ and so, as a convention and unless otherwise stated, we choose $u(t) = 0$ for $t > t_0 + \tau^{t_0,x_0,u}$. Now, suppose that

$$\exists \lambda \in \left(0, \frac{1}{k_{\max}}\right) \text{ s.t. } |g(x) - g(y)| \leq \lambda |x - y| \quad \text{for all } x, y \in \partial\Omega. \tag{H3}$$

This is a standard assumption in exit-time optimal control problems with boundary costs (see, e.g., [17, (8.6) and Remark 8.1.5] and [33]), its importance being the following property.

Lemma 3.1. *Let $g : \partial\Omega \to \mathbb{R}^+$ satisfy (H3) and $\gamma : \mathbb{R}^+ \to \Omega$ be $k_{\max}$-Lipschitz. If $t_1, t_2 \in \mathbb{R}^+$ are such that $t_1 < t_2$ and $\gamma(t_1), \gamma(t_2) \in \partial\Omega$, then*

$$t_1 + g(\gamma(t_1)) < t_2 + g(\gamma(t_2)).$$

Proof. We have

$$g(\gamma(t_1)) - g(\gamma(t_2)) \leq \lambda |\gamma(t_1) - \gamma(t_2)| \leq \lambda k_{\max}(t_2 - t_1) < t_2 - t_1,$$

yielding the result. $\square$

Under assumptions (H1), (H2) and (H3), we have the following existence result.

Proposition 3.2. *For every $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$, there exists an optimal control u for the cost (3.2).*

Proof. Let $(u_n)_n$ be a minimizing sequence. We set for simplicity $\tau_n := \tau^{t_0,x_0,u_n}$, $\gamma_n := \gamma^{t_0,x_0,u_n}$ and $z_n := \gamma^{t_0,x_0,u_n}_{\tau_n}$. It is clear that, up to extracting subsequences, $\tau_n \to \bar\tau$ and $\gamma_n \to \gamma$, where $\gamma = \gamma^{t_0,x_0,u}$ is an admissible trajectory. On the other hand, we have

$$z_n \to \gamma(t_0 + \bar\tau),$$

which implies that $\gamma(t_0 + \bar\tau) \in \partial\Omega$ and $\tau := \tau^{t_0,x_0,u} \leq \bar\tau$. Yet, it is not possible to have $\tau < \bar\tau$. Indeed, since $(u_n)_n$ is a minimizing sequence, one has

$$\lim_n \tau_n + g(z_n) = \bar\tau + g(\gamma(t_0 + \bar\tau)) \leq \tau + g(\gamma(t_0 + \tau)),$$

which is a contradiction thanks to Lemma 3.1. Thus, we have $\tau = \bar\tau$ and this completes the proof that γ is an optimal trajectory and u is the associated optimal control. □

We note that the condition (H3) is crucial for this result. Without this condition, one should replace the cost in (3.2) by $\inf\{t + g(\gamma(t_0 + t)) \,:\, \gamma(t_0 + t) \in \partial\Omega\}$.

Another easily obtained property is that the restriction of an optimal control is still optimal.

Proposition 3.3. *Let $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$, u be an optimal control for x_0, at time t_0, $\gamma = \gamma^{t_0,x_0,u}$, and $\tau_0 = \tau^{t_0,x_0,u}$. Then, for every $t \in [t_0, t_0 + \tau_0)$, $u|_{[t,t_0+\tau_0]}$ is an optimal control for $\gamma(t)$, at time t.*

Proof. Let $\bar u = u|_{[t,t_0+\tau_0]}$ and notice that $\gamma^{t_0,x_0,u}_\tau = \gamma^{t,\gamma(t),\bar u}_\tau$ and $\tau^{t_0,x_0,u} = \tau^{t,\gamma(t),\bar u} + t - t_0$. If $\bar u$ is not optimal for $\gamma(t)$, at time t, one can find a control v for $\gamma(t)$, at time t, such that

$$\tau^{t,\gamma(t),\bar u} + g(\gamma^{t,\gamma(t),\bar u}_\tau) > \tau^{t,\gamma(t),v} + g(\gamma^{t,\gamma(t),v}_\tau).$$

By concatenating $u|_{[t_0,t]}$ with v, one obtains a control $\tilde u$ for x_0, at time t_0, such that

$$\begin{aligned}\tau^{t_0,x_0,\tilde u} + g(\gamma^{t_0,x_0,\tilde u}_\tau) &= t - t_0 + \tau^{t,\gamma(t),v} + g(\gamma^{t,\gamma(t),v}_\tau)\\ &< t - t_0 + \tau^{t,\gamma(t),\bar u} + g(\gamma^{t,\gamma(t),\bar u}_\tau) = \tau^{t_0,x_0,u} + g(\gamma^{t_0,x_0,u}_\tau),\end{aligned}$$

contradicting the optimality of u. □

The value function $\varphi : \mathbb{R}^+ \times \Omega \to \mathbb{R}^+$ of the above optimal control problem is defined by

$$\varphi(t, x) = \min\{\tau^{t,x,u} + g(\gamma^{t,x,u}_\tau) \,:\, u \text{ is a control}\}, \quad t \in \mathbb{R}^+,\ x \in \Omega. \tag{3.3}$$

The first important fact is that the value function φ satisfies the so-called *dynamic programming principle* stated in the next lemma, which can be proved by standard techniques in optimal control (see, for instance, [17, (8.4)]):

Lemma 3.4. *For any $t_0 \in \mathbb{R}^+$, $x_0 \in \Omega$ and any control $u : [t_0, \infty) \to \bar B(0, 1)$, we have*

$$\varphi(t_0, x_0) \leq t - t_0 + \varphi(t, \gamma^{t_0,x_0,u}(t)), \quad \textit{for all } t \in [t_0, t_0 + \tau^{t_0,x_0,u}],$$

with equality if u is optimal.

One can also use standard techniques in optimal control, similar to those, e.g., in [17, Theorem 8.1.8], to show that the value function φ is a viscosity solution of a suitable partial differential equation.

Proposition 3.5. *The value function φ is a viscosity solution of the following Hamilton–Jacobi equation*

$$-\partial_t\varphi(t,x)+k(t,x)|\nabla\varphi(t,x)|-1=0,\quad (t,x)\in\mathbb{R}^+\times\mathring{\Omega}. \tag{3.4}$$

Moreover, one has $\varphi(t,x)=g(x)$ for every $(t,x)\in\mathbb{R}^+\times\partial\Omega$.

Our next result shows that, if we consider points along optimal trajectories different from the endpoints, we can prove that the elements of $D^+\varphi$ also satisfy (3.4).

Proposition 3.6. *Let $\gamma:[t_0,t_0+\tau_0]\to\Omega$ be an optimal trajectory for x_0, at time t_0, where $\tau_0=\tau^{t_0,x_0,u}$ and u is the associated optimal control. Then, for every $t\in(t_0,t_0+\tau_0)$, we have*

$$-p_t+k(t,\gamma(t))|p_x|-1=0,\quad \textit{for all } (p_t,p_x)\in D^+\varphi(t,\gamma(t)).$$

Proof. Let us take $t\in(t_0,t_0+\tau_0)$. Since φ is a viscosity subsolution of (3.4), we have

$$-p_t+k(t,\gamma(t))|p_x|-1\le 0,\quad \text{for all } (p_t,p_x)\in D^+\varphi(t,\gamma(t)).$$

So, it suffices to prove that the converse inequality also holds. First, we observe that, by the dynamic programming principle,

$$\varphi(t,\gamma(t))=\varphi(t-h,\gamma(t-h))-h,\quad 0\le h\le t-t_0.$$

On the other hand, if $(p_t,p_x)\in D^+\varphi(t,\gamma(t))$, we have

$$\varphi(t-h,\gamma(t-h))-\varphi(t,\gamma(t))\le -p_t h-p_x\cdot(\gamma(t)-\gamma(t-h))+o(h).$$

Therefore, we find that

$$\begin{aligned}0&\le -p_t h-p_x\cdot(\gamma(t)-\gamma(t-h))-h+o(h)\\&=-p_t h-\int_{t-h}^{t}k(s,\gamma(s))p_x\cdot u(s)\,\mathrm{d}s-h+o(h)\\&\le h(-p_t+\sup_{s\in[t-h,t]}k(s,\gamma(s))|p_x|-1)+o(h),\end{aligned}$$

which yields the conclusion. ☐

We now want to provide an upper bound on the optimal exit time $\tau^{t_0,x_0,u}$, where u is an optimal control for x_0, at time t_0. To do so, we compare $\tau^{t_0,x_0,u}$ with the minimal time needed to reach $\partial\Omega$ from x_0 at time t_0. Let us introduce the *minimal-time function* $T:\mathbb{R}^+\times\Omega\to\mathbb{R}^+$ defined by

$$T(t,x)=\inf\{\tau^{t,x,u}\ :\ u \text{ is a control}\},\quad (t,x)\in\mathbb{R}^+\times\Omega, \tag{3.5}$$

which corresponds to taking $g=0$ in (3.2). An optimal control u for the optimization problem of $T(t,x)$ is called a *minimal-time control*. We then have the following result.

Proposition 3.7. *For every $(t,x)\in\mathbb{R}^+\times\Omega$, one has $T(t,x)\le k_{\min}^{-1}d(x,\partial\Omega)$. Moreover, if u is an optimal control for (3.2) for x at time t, then one has*

$$\tau^{t,x,u}\le\frac{1+\lambda k_{\max}}{1-\lambda k_{\max}}T(t,x).$$

In particular,

$$\tau^{t,x,u}\le\frac{k_{\min}^{-1}(1+\lambda k_{\max})}{1-\lambda k_{\max}}d(x,\partial\Omega).$$

Proof. Let $\gamma : [0, d(x, \partial\Omega)] \to \Omega$ be a segment from x to the closest point from x on the boundary $\partial\Omega$ with $|\gamma'| = 1$. Set $\widetilde{\gamma}(s) = \gamma(k_{\min}(s-t))$, for all $s \in [t, t + k_{\min}^{-1} d(x, \partial\Omega)]$. It is clear that there is a control $\bar{v}$ such that $\widetilde{\gamma} = \gamma^{t,x,\bar{v}}$. Hence, by definition of T, one has $T(t,x) \leq \tau^{t,x,\bar{v}} = k_{\min}^{-1} d(x, \partial\Omega)$.

Now, let u be an optimal control for (3.2) and v be a minimal-time control for (t,x). Then, we have

$$\tau^{t,x,u} + g(\gamma_\tau^{t,x,u}) \leq \tau^{t,x,v} + g(\gamma_\tau^{t,x,v})$$

and so,

$$\begin{aligned}\tau^{t,x,u} &\leq \tau^{t,x,v} + \lambda|\gamma_\tau^{t,x,v} - \gamma_\tau^{t,x,u}| \leq \tau^{t,x,v} + \lambda\big(|\gamma_\tau^{t,x,v} - x| + |\gamma_\tau^{t,x,u} - x|\big) \\ &\leq \tau^{t,x,v} + \lambda k_{\max}(\tau^{t,x,v} + \tau^{t,x,u}).\end{aligned}$$

Consequently, we get

$$\tau^{t,x,u} \leq \frac{1 + \lambda k_{\max}}{1 - \lambda k_{\max}} \tau^{t,x,v} = \frac{1 + \lambda k_{\max}}{1 - \lambda k_{\max}} T(t,x). \qquad \square$$

Now, we want to give a result about the Lipschitz continuity of the value function φ.

Proposition 3.8. *Let our system satisfy properties* (H1), (H2), & (H3). *Then the value function φ is Lipschitz continuous in $\mathbb{R}^+ \times \Omega$.*

Proof. Following the same lines of the proof of Proposition 8.2.5 in [17], one can check that the value function φ is Lipschitz w.r.t. x as soon as k is Lipschitz in x with a Lipschitz constant $c > 0$ independent of t. Now, to prove Lipschitz continuity with respect to t, take $t_0 \in \mathbb{R}^+$, $x \in \Omega$ and let u be an optimal control for x, at time t_0. Let $\delta \in (0, \tau^{t_0,x,u})$. Then, using Lemma 3.4, one has

$$\begin{aligned}&|\varphi(t_0 + \delta, x) - \varphi(t_0, x)| \\ &\leq |\varphi(t_0+\delta, x) - \varphi(t_0 + \delta, \gamma^{t_0,x,u}(t_0+\delta))| + |\varphi(t_0+\delta, \gamma^{t_0,x,u}(t_0+\delta)) - \varphi(t_0,x)| \\ &\leq c|\gamma^{t_0,x,u}(t_0+\delta) - x| + \delta \\ &\leq (1 + ck_{\max})\delta. \qquad \square\end{aligned}$$

The last preliminary result we present in this subsection provides a lower bound on the variation in time of φ.

Proposition 3.9. *Assume that* (H1), (H2), *and* (H3) *hold. Then there exists $c > 0$ depending only on $k_{\min}$, $k_{\max}$, $\operatorname{diam}(\Omega)$, λ, and L_x such that, for every $x \in \Omega$ and $t_0, t_1 \in \mathbb{R}^+$ with $t_0 \neq t_1$,*

$$\frac{\varphi(t_1, x) - \varphi(t_0, x)}{t_1 - t_0} \geq c - 1. \tag{3.6}$$

Proof. Suppose, without loss of generality, that $t_0 < t_1$. Let γ_1 be an optimal trajectory for x, at time t_1, and u_1 be the associated optimal control. Define $\phi : [t_0, +\infty) \to [t_1, +\infty)$ as a function satisfying

$$\begin{cases} \phi'(t) = \dfrac{k(t, \gamma_1(\phi(t)))}{k(\phi(t), \gamma_1(\phi(t)))}, \\ \phi(t_0) = t_1. \end{cases} \tag{3.7}$$

Notice that, since k is only continuous with respect to its first variable, ϕ is not unique a priori. Set $\gamma_0(t) = \gamma_1(\phi(t))$ for all $t \geq t_0$. By construction of ϕ, it is clear that there is a control u_0 such that $\gamma_0 = \gamma^{t_0,x,u_0}$ (more precisely, $u_0(t) = u_1(\phi(t))$ for $t \geq t_0$). Moreover, we have $\tau_0 := \tau^{t_0,x,u_0} = \phi^{-1}(t_1 + \tau_1) - t_0$, where $\tau_1 := \tau^{t_1,x,u_1}$. So, $\phi(t_0 + \tau_0) = t_1 + \tau_1$ and

$\phi(t_0 + \tau_0) + g(\gamma_0(t_0 + \tau_0)) = t_1 + \varphi(t_1, x)$. On the other hand, from (3.7), it is easy to see that, for all $t, \bar{t} \geq t_0$, one has

$$\int_{\phi(t)}^{\phi(\bar{t})} k(s, \gamma_1(s)) \,\mathrm{d}s = \int_t^{\bar{t}} k(s, \gamma_1(\phi(s))) \,\mathrm{d}s.$$

Now, set

$$G(\theta) = \int_{\theta}^{\phi(\bar{t})} k(s, \gamma_1(s)) \,\mathrm{d}s, \quad \forall \theta \in \mathbb{R}^+,$$

where we extend γ_1 to $\mathbb{R}^+$ by setting $\gamma_1(s) = \gamma_1(t_1)$ for $s \in [0, t_1)$. Then, using that G is bi-Lipschitz, we have

$$\begin{aligned}
|\phi(t) - t| &= \left|G^{-1}\left(\int_t^{\bar{t}} k(s, \gamma_1(\phi(s))) \,\mathrm{d}s\right) - G^{-1}\left(\int_t^{\phi(\bar{t})} k(s, \gamma_1(s)) \,\mathrm{d}s\right)\right| \\
&\leq C\left|\int_t^{\bar{t}} k(s, \gamma_1(\phi(s))) \,\mathrm{d}s - \int_t^{\phi(\bar{t})} k(s, \gamma_1(s)) \,\mathrm{d}s\right| \\
&\leq C\left(|\phi(\bar{t}) - \bar{t}| + \int_t^{\bar{t}} |k(s, \gamma_1(\phi(s))) - k(s, \gamma_1(s))| \,\mathrm{d}s\right) \\
&\leq C|\phi(\bar{t}) - \bar{t}| + C\int_t^{\bar{t}} |\phi(s) - s| \,\mathrm{d}s,
\end{aligned}$$

where $C > 0$ denotes a constant depending only on $k_{\min}$, $k_{\max}$, and L_x, whose value may change from one line to the other. Using the fact that $\phi(t_0) = t_1 > t_0$, we infer that $\phi(t) > t$ for all $t \geq t_0$. Now, if $\bar{t} = t_0 + \tau_0$, we get, using Gronwall's inequality, that

$$\phi(t) - t \leq C e^{C|t_0 + \tau_0 - t|}(\phi(t_0 + \tau_0) - (t_0 + \tau_0)).$$

Setting $t = t_0$, one has

$$c(t_1 - t_0) \leq \phi(t_0 + \tau_0) - (t_0 + \tau_0) = t_1 + \varphi(t_1, x) - g(\gamma_0(t_0 + \tau_0)) - t_0 - \tau_0,$$

where we use Proposition 3.7 to provide an upper bound on τ_0 and $c > 0$ only depends on $k_{\min}$, $k_{\max}$, $\operatorname{diam}(\Omega)$, λ, and L_x. Then

$$(c - 1)(t_1 - t_0) \leq \varphi(t_1, x) - g(\gamma_0(t_0 + \tau_0)) - \tau_0 = \varphi(t_1, x) - \varphi(t_0, x),$$

as required. $\square$

Remark 3.10. The analogue of Proposition 3.9 was already proved in [58, Proposition 4.5] for the minimal-time function T. Even though the proof of [58] could be easily adapted to our setting, it would require Lipschitz continuity of k with respect to t. Our proof refines that of [58] and does not require such an assuption. The fact that c does not depend on any Lipschitz behavior of k with respect to t will be a key property for the results in Sections 4.2 and 4.3.

Proposition 3.9 yields a lower bound on the time derivative of the value function φ, which can be used to obtain information on the gradient of φ thanks to the Hamilton–Jacobi equation (3.4).

Corollary 3.11. *There exists $c > 0$ (which only depends on $k_{\min}$, $k_{\max}$, $\operatorname{diam}(\Omega)$, λ, and L_x) such that $\partial_t \varphi(t, x) \geq c - 1$ and $|\nabla \varphi(t, x)| \geq c$ for all $(t, x) \in \mathbb{R}^+ \times \Omega$ where φ is differentiable.*

3.2. **Pontryagin Maximum Principle and its consequences.** In this subsection, we use the necessary optimality conditions of Pontryagin Maximum Principle to obtain further properties of optimal trajectories and the value function φ. In addition to (H1), (H2), and (H3), we also assume that

$$\partial\Omega \text{ is of class } C^{1,1}, \tag{H4}$$

$$\nabla k \in C(\mathbb{R}^+ \times \Omega), \tag{H5}$$

$$g \in C^1(\partial\Omega). \tag{H6}$$

In order to state a version of Pontryagin Maximum Principle, we start with a preliminary result (see [17, Lemma 8.4.2]).

Lemma 3.12. *Given $z \in \partial\Omega$, let $\mathbf{n}$ be the outer normal to $\partial\Omega$ at z. Then, for every $t \in \mathbb{R}^+$, there exists a unique $\mu > 0$ such that $k(t,z)|\nabla g(z) - \mu\mathbf{n}| - 1 = 0$.*

We are now ready to state Pontryagin Maximum Principle for this control problem.

Proposition 3.13. *Let properties (H1), (H2), (H3), (H4), (H5), and (H6) hold, let $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$ and let $\bar{u}$ be an optimal control for x_0, at time t_0. Set for simplicity*

$$\gamma := \gamma^{t_0,x_0,\bar{u}}, \qquad \tau_0 := \tau^{t_0,x_0,\bar{u}}, \qquad z := \gamma_{\tau}^{t_0,x_0,\bar{u}},$$

and denote by $\mathbf{n}$ the outer normal to $\partial\Omega$ at z. Let $\mu > 0$ be such that $k(t_0+\tau_0, z)|\nabla g(z) - \mu\mathbf{n}| - 1 = 0$ (μ is uniquely determined by the previous lemma). Let $p : [t_0, t_0+\tau_0] \to \mathbb{R}^d$ be the solution to the system

$$\begin{cases} p'(t) = -\nabla k(t,\gamma(t))\bar{u}(t) \cdot p(t), \\ p(t_0+\tau_0) = \nabla g(z) - \mu\mathbf{n}. \end{cases} \tag{3.8}$$

Then, for a.e. $t \in [t_0, t_0+\tau_0]$,

$$-p(t) \cdot \bar{u}(t) = \max_{u \in \bar{B}(0,1)} -p(t) \cdot u.$$

We refer the reader to [17, Lemma 8.4.2 and Theorem 8.4.3] for proofs of the above results. Even though the proofs in [17] only consider the case of autonomous dynamics, their extension to our non-autonomous setting is straightforward.

As a consequence of Proposition 3.13, we get the following.

Proposition 3.14. *Let $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$ and $\bar{u}$, γ, τ_0, and p be as in the statement of Proposition 3.13. Then p is non-zero in $[t_0, t_0+\tau_0]$, $\bar{u}(t) = -\frac{p(t)}{|p(t)|}$ for every $t \in [t_0, t_0+\tau_0]$, $\bar{u}$ is L_x-Lipschitz continuous on $[t_0, t_0+\tau_0]$ and satisfies*

$$\bar{u}'(t) = -\nabla k(t,\gamma(t)) + \bar{u}(t) \cdot \nabla k(t,\gamma(t))\bar{u}(t) \quad \textit{for a.e. } t \in [t_0, t_0+\tau_0],$$

and γ is C^1 on $[t_0, t_0+\tau_0]$. Moreover, $\gamma \in C^{1,1}([t_0, t_0+\tau_0], \Omega)$ as soon as k is Lipschitz in t.

Remark 3.15. Similar results have been obtained for the minimal-time problem (3.5) in [58, Lemma 4.13 and Corollary 4.14]. Proposition 3.14 can be obtained using analogous arguments. Since the proof is short, we provide it here for the reader's convenience.

Proof. From (3.8), one obtains that, for every $t \in [t_0, t_0+\tau_0]$, $|p'(t)| \leq L_x |p(t)|$. Hence, we get

$$|p(t)| \leq |p(s)| e^{L_x |t-s|}$$

for every $s, t \in [t_0, t_0+\tau_0]$. Thus, if there exists $s \in [t_0, t_0+\tau_0]$ such that $p(s) = 0$, one concludes that $p(t) = 0$ for every $t \in [t_0, t_0+\tau_0]$, which is a contradiction as $p(t_0+\tau_0) = \nabla g(\gamma(t_0+\tau_0)) - \mu \mathbf{n} \neq 0$, thanks to Lemma 3.12. Since $p(t) \neq 0$ for every $t \in [t_0, t_0+\tau_0]$, it follows immediately from Proposition 3.13 that $\bar{u}(t) = -\frac{p(t)}{|p(t)|}$. Hence,

$$\begin{aligned}\bar{u}'(t) &= -\frac{|p(t)| p'(t) - \frac{p(t)\cdot p'(t)}{|p(t)|} p(t)}{|p(t)|^2} \\ &= -\nabla k(t, \gamma(t)) + \bar{u}(t) \cdot \nabla k(t, \gamma(t)) \bar{u}(t).\end{aligned}$$

In particular, $\bar{u}$ is L_x-Lipschitz continuous. The conclusions on γ follow immediately from (3.1). ◻

From now on, we suppose also that

(H7) $\exists L_{xx} > 0$ such that $|\nabla k(t, x_0) - \nabla k(t, x_1)| \leq L_{xx} |x_0 - x_1| \quad$ for all $x_0, x_1 \in \Omega,\ t \in \mathbb{R}^+$.

Then, under the assumptions of Proposition 3.14 and (H7), $(\gamma, \bar{u})$ is the unique solution on $[t_0, t_0+\tau_0]$ of

$$\begin{cases} \gamma'(t) = k(t, \gamma(t)) u(t), \\ u'(t) = -\nabla k(t, \gamma(t)) + u(t) \cdot \nabla k(t, \gamma(t)) u(t), \\ \gamma(t_0) = x_0, \\ u(t_0) = \bar{u}(t_0). \end{cases} \tag{3.9}$$

Now, let us introduce the following lemma, which shows that the uniform limit of optimal trajectories is an optimal trajectory.

Lemma 3.16. *Assume that (H1)—(H6) hold. Let $(t_n, x_n)_n$ be a sequence in $\mathbb{R}^+ \times \Omega$ such that $t_n \to t$ and $x_n \to x$. For each n, let γ_n be an optimal trajectory for x_n, at time t_n, and u_n be the associated optimal control. Then, up to extracting subsequences, there exist γ and u such that $\gamma_n \to \gamma$ and $u_n \to u$ uniformly, where γ is an optimal trajectory for x, at time t, and u is its associated optimal control.*

Proof. The fact that $\gamma_n \to \gamma$ and $u_n \to u$ follows immediately using Proposition 3.14. Yet, for every n, we have

$$\gamma_n'(s) = k(s, \gamma_n(s)) u_n(s), \quad \text{for a.e. } s > t_n.$$

Passing to the limit as $n \to +\infty$, we get

$$\gamma'(s) = k(s, \gamma(s)) u(s), \quad \text{for a.e. } s > t. \tag{3.10}$$

Moreover, $\gamma_n(t_n) = x_n$ implies, at the limit, that $\gamma(t) = x$. Hence, γ is an admissible trajectory for x, at time t, and u is its associated control. Now, set $\tau_n = \tau^{t_n, x_n, u_n}$ and $z_n = \gamma_n(t_n + \tau_n) \in \partial\Omega$. From Proposition 3.7, we have that, up to extracting subsequences, $\tau_n \to \bar{\tau}$ for some $\bar{\tau} \in \mathbb{R}^+$. In addition, $z_n \to \gamma(t+\bar{\tau}) \in \partial\Omega$. Hence, $\tau := \tau^{t,x,u} \leq \bar{\tau}$ and, one has

$$\varphi(t_n, x_n) = \tau_n + g(z_n) \to \bar{\tau} + g(\gamma(t+\bar{\tau})).$$

Yet, by the continuity of the value function φ, we infer that

$$\varphi(t, x) = \bar{\tau} + g(\gamma(t+\bar{\tau})) \leq \tau + g(\gamma(t+\tau))$$

and then, thanks to Lemma 3.1, $\bar{\tau} = \tau$. ◻

On the other hand, we have the following result about the uniqueness of optimal control at any interior point of an optimal trajectory.

Proposition 3.17. *Assume that (H1)—(H7) hold. Let γ be an optimal trajectory for x_0 at time t_0, and set $\tau_0 = \tau^{t_0,x_0,u}$, where u is the associated optimal control. Then, for every $t \in (t_0, t_0+\tau_0)$, u is the unique optimal control for $\gamma(t)$, at time t.*

Proof. Fix $t \in (t_0, t_0+\tau_0)$ and let v be an optimal control for $x := \gamma(t)$, at time t. Set

$$\widetilde{u}(s) = \begin{cases} u(s), & \text{if } s < t, \\ v(s), & \text{if } s \geq t. \end{cases}$$

Then $\widetilde{u}$ is an optimal control for x_0, at time t_0. Indeed, using the optimality of v, we have $\varphi(t_0,x_0) \leq \tau^{t_0,x_0,\widetilde{u}} + g(\gamma_\tau^{t_0,x_0,\widetilde{u}}) = t - t_0 + \varphi(t,x)$. On the other hand, since u is optimal, one obtains from Lemma 3.4 that $\varphi(t_0,x_0) = t - t_0 + \varphi(t,x)$. Then $\varphi(t_0,x_0) = \tau^{t_0,x_0,\widetilde{u}} + g(\gamma_\tau^{t_0,x_0,\widetilde{u}})$, and so the control $\widetilde{u}$ is optimal. Hence, by Proposition 3.14, $\widetilde{u}$ is continuous, which proves that $u(t) = v(t) := q$. The fact that $u(s) = v(s)$, for all $s \geq t$, follows from the uniqueness of solutions to the system (3.9) with initial conditions $\gamma(t) = x$ and $u(t) = q$. $\square$

Given an optimal trajectory γ for x_0 at time t_0, we will say that p is a *dual arc* associated with γ if it satisfies the properties of Proposition 3.13, that is, if it solves (3.8). Our next result states that the dual arc p is included in the superdifferential of the value function φ with respect to x, $\nabla^+\varphi$.

Proposition 3.18. *Under the assumptions of Proposition 3.13, the arc p solution of (3.8) satisfies*

$$p(t) \in \nabla^+\varphi(t,\gamma(t)), \quad \textit{for all } t \in [t_0, t_0+\tau_0).$$

The proof of Proposition 3.18 can be obtained by easily adapting the proof of [17, Theorem 8.4.4] to our non-autonomous setting, and is omitted here for simplicity. Similarly, one can obtain an analogous property for the subdifferential by an immediate adaptation of the techniques from [17, Theorem 7.3.4].

Proposition 3.19. *Let u be an optimal control for $(t_0,x_0) \in \mathbb{R}^+ \times \Omega$ and γ be its associated optimal trajectory. Let $p : [t_0, t_0+\tau_0] \to \mathbb{R}^d$ be any solution of the adjoint equation*

$$p'(t) = -\nabla k(t,\gamma(t))u(t) \cdot p(t), \quad t \in [t_0, t_0+\tau_0], \tag{3.11}$$

where $\tau_0 = \tau^{t_0,x_0,u}$. Suppose that $p(t_0) \in \nabla^-\varphi(t_0,x_0)$, then $p(t) \in \nabla^-\varphi(t,\gamma(t))$, for all $t \in [t_0, t_0+\tau_0)$.

As a consequence of the previous results, we can show that the existence of $\nabla\varphi$ at some point (t_0,x_0) is sufficient to ensure uniqueness of the optimal trajectory for x_0, at time t_0.

Proposition 3.20. *Let $(t_0,x_0) \in \mathbb{R}^+ \times \Omega$ and assume that $\nabla\varphi(t_0,x_0)$ exists. Then there exists a unique trajectory γ which is optimal for x_0, at time t_0.*

Proof. Assume that γ_1, γ_2 are optimal trajectories for x_0, at time t_0, and denote the respective optimal controls by u_1, u_2. For $i \in \{1,2\}$, write $\tau_i = \tau^{t_0,x_0,u_i}$ and let $p_i : [t_0, t_0+\tau_i] \to \mathbb{R}^d$ be a dual arc associated with γ_i. By Proposition 3.14, p_i is non-zero and u_i is Lipschitz continuous on $[t_0, t_0+\tau_i]$, with $u_i(t) = -\frac{p_i(t)}{|p_i(t)|}$ for every $t \in [t_0, t_0+\tau_i]$. By Proposition 3.18, $p_i(t) \in \nabla^+\varphi(t,\gamma_i(t))$ for every $t \in [t_0, t_0+\tau_i)$. In particular, since $\nabla\varphi(t_0,x_0)$ exists, one has $p_1(t_0) = p_2(t_0) = \nabla\varphi(t_0,x_0)$, yielding that $\nabla\varphi(t_0,x_0) \neq 0$ and $u_1(t_0) = u_2(t_0) = \frac{\nabla\varphi(t_0,x_0)}{|\nabla\varphi(t_0,x_0)|}$. This means that both (γ_1,u_1) and (γ_2,u_2) solve (3.9) with the same initial conditions $\gamma_1(t_0) =$

$\gamma_2(t_0) = x_0$ and $u_1(t_0) = u_2(t_0) = \frac{\nabla\varphi(t_0,x_0)}{|\nabla\varphi(t_0,x_0)|}$, yielding, by uniqueness of the solutions of (3.9), that $\gamma_1 = \gamma_2$. $\square$

To conclude this subsection, we prove that, when optimal trajectories are close enough to the boundary, they always move towards the boundary, in the sense that the scalar product between the direction of the trajectory and some normal direction is lower bounded by a positive constant. To do so, we make use of the *signed distance* to $\partial\Omega$, which is the function $d^{\pm} : \mathbb{R}^d \to \mathbb{R}$ defined for $x \in \mathbb{R}^d$ by

$$d^{\pm}(x) = \begin{cases} d(x, \partial\Omega), & \text{if } x \notin \Omega, \\ -d(x, \partial\Omega), & \text{otherwise.} \end{cases} \tag{3.12}$$

Recall that, thanks to (H4), $d^{\pm}$ is 1-Lipschitz on $\mathbb{R}^d$, $C^{1,1}$ in a neighborhood of the boundary, and, if $x \in \partial\Omega$, then $\nabla d^{\pm}(x)$ is the outer normal to $\partial\Omega$ at x (see, e.g., [32]).

Proposition 3.21. *There exist $c > 0$ (depending only on $k_{\min}$, $k_{\max}$, and λ) and $\delta > 0$ (depending only on $k_{\min}$, $k_{\max}$, λ, $\operatorname{diam}(\Omega)$, and the curvature of $\partial\Omega$) such that, for every $(t_0, x_0) \in \mathbb{R}^+ \times \mathring{\Omega}$, if u is an optimal control for x_0, at time t_0, $\gamma := \gamma^{t_0,x_0,u}$ is the corresponding optimal trajectory, and $\tau_0 := \tau^{t_0,x_0,u}$, then, for every $t \in [t_0, t_0 + \tau_0]$ such that $d(\gamma(t), \partial\Omega) \leq \delta$, one has*

$$\nabla d^{\pm}(\gamma(t)) \cdot u(t) \geq c. \tag{3.13}$$

In particular, if $d(x_0, \partial\Omega) \leq \delta$ and $\nabla\varphi(t_0, x_0)$ exists, then $\nabla\varphi(t_0, x_0) \neq 0$ and

$$-\nabla d^{\pm}(x_0) \cdot \frac{\nabla\varphi(t_0, x_0)}{|\nabla\varphi(t_0, x_0)|} \geq c. \tag{3.14}$$

Proof. Let $p : [t_0, t_0 + \tau_0] \to \mathbb{R}^d$ be a dual arc associated with γ. By Proposition 3.14, p is nonzero and u is Lipschitz continuous on $[t_0, t_0 + \tau_0]$, with $u(t) = -\frac{p(t)}{|p(t)|}$ for every $t \in [t_0, t_0 + \tau_0]$. In particular, $u(t_0 + \tau_0) = \frac{\mu\mathbf{n} - \nabla g(z)}{|\mu\mathbf{n} - \nabla g(z)|}$, where z, $\mathbf{n}$, and μ are as in the statement of Proposition 3.13.

We first prove (3.13) at the final time $t_0 + \tau_0$. Recalling that $k(t_0 + \tau_0, z)|\nabla g(z) - \mu\mathbf{n}| = 1$, one has

$$\frac{1}{k(t_0 + \tau_0, z)^2} = |\nabla g(z) - \mu\mathbf{n}|^2 = |\nabla g(z)|^2 - 2\mu\nabla g(z) \cdot \mathbf{n} + \mu^2,$$

and thus

$$2\mu(\mu - \nabla g(z) \cdot \mathbf{n}) = \frac{1}{k(t_0 + \tau_0, z)^2} - |\nabla g(z)|^2 + \mu^2. \tag{3.15}$$

On the other hand, one also has that

$$\frac{1}{k(t_0 + \tau_0, z)} = |\nabla g(z) - \mu\mathbf{n}| \geq \mu - |\nabla g(z)|,$$

and thus

$$\mu \leq \frac{1}{k(t_0 + \tau_0, z)} + |\nabla g(z)|.$$

Combining this with (3.15), one gets that

$$\begin{aligned}\mu - \nabla g(z)\cdot \mathbf{n} &= \frac{\frac{1}{k(t_0+\tau_0,z)^2} - |\nabla g(z)|^2 + \mu^2}{2\mu} > \frac{\frac{1}{k(t_0+\tau_0,z)^2} - |\nabla g(z)|^2}{2\left(\frac{1}{k(t_0+\tau_0,z)} + |\nabla g(z)|\right)} \\ &= \frac{1}{2}\left(\frac{1}{k(t_0+\tau_0,z)} - |\nabla g(z)|\right) \geq \frac{1}{2}\left(\frac{1}{k_{\max}} - \lambda\right) > 0.\end{aligned}$$

Hence, recalling that $|\mu\mathbf{n} - \nabla g(z)| = \frac{1}{k(t_0+\tau_0,z)} \leq \frac{1}{k_{\min}}$, one obtains that

$$\nabla d^{\pm}(z)\cdot u(t_0+\tau_0) = \mathbf{n}\cdot \frac{\mu\mathbf{n} - \nabla g(z)}{|\mu\mathbf{n} - \nabla g(z)|} = \frac{\mu - \nabla g(z)\cdot\mathbf{n}}{|\mu\mathbf{n} - \nabla g(z)|} \geq \frac{k_{\min}}{2}\left(\frac{1}{k_{\max}} - \lambda\right), \tag{3.16}$$

which corresponds to (3.13) at the final time $t_0+\tau_0$.

Now, let $\delta_0 > 0$ be such that $d^{\pm}$ is $C^{1,1}$ on the set $\{x \in \mathbb{R}^d : d(x,\partial\Omega) \leq \delta_0\}$ and $L_d > 0$ be a Lipschitz constant for $\nabla d^{\pm}$ on this set. By Proposition 3.14, u is L_x-Lipschitz on $[t_0, t_0+\tau_0]$. Take

$$\begin{aligned} c &= \frac{k_{\min}}{4}\left(\frac{1}{k_{\max}} - \lambda\right), \\ \delta &= \min\left\{\delta_0, \frac{k_{\min}^2(1-\lambda k_{\max})^2}{4k_{\max}(1+\lambda k_{\max})(L_d k_{\max} + L_x)}\right\}.\end{aligned}$$

Let $t \in [t_0, t_0+\tau_0)$ be such that $d(\gamma(t),\partial\Omega) \leq \delta$. By Proposition 3.3, $u|_{[t,t_0+\tau_0]}$ is an optimal control for $\gamma(t)$, at time t, and thus, by Proposition 3.7, one obtains that

$$t_0 + \tau_0 - t = \tau^{t,\gamma(t),u|_{[t,t_0+\tau_0]}} \leq \frac{(1+\lambda k_{\max})\delta}{(1-\lambda k_{\max})k_{\min}} \leq \frac{k_{\min}(1-\lambda k_{\max})}{4k_{\max}(L_d k_{\max} + L_x)}.$$

Hence, by the previous inequality and (3.16), one has

$$\begin{aligned}\nabla d^{\pm}(\gamma(t))\cdot u(t) &= \nabla d^{\pm}(z)\cdot u(t_0+\tau_0) + \big(\nabla d^{\pm}(\gamma(t)) - \nabla d^{\pm}(z)\big)\cdot u(t) \\ &\qquad + \nabla d^{\pm}(z)\cdot(u(t) - u(t_0+\tau_0)) \\ &\geq \frac{k_{\min}}{2}\left(\frac{1}{k_{\max}} - \lambda\right) - L_d|\gamma(t) - z| - L_x|t_0+\tau_0 - t| \\ &\geq \frac{k_{\min}}{2}\left(\frac{1}{k_{\max}} - \lambda\right) - (L_d k_{\max} + L_x)(t_0+\tau_0 - t) \\ &\geq \frac{k_{\min}}{4}\left(\frac{1}{k_{\max}} - \lambda\right) = c,\end{aligned}$$

concluding the proof of (3.13).

Concerning the last part of the statement, notice that, as a consequence of Proposition 3.18 and the fact that $\nabla\varphi(t_0,x_0)$ exists, one deduces that $p(t_0) = \nabla\varphi(t_0,x_0)$, yielding that $\nabla\varphi(t_0,x_0) \neq 0$ and $u(t_0) = -\frac{\nabla\varphi(t_0,x_0)}{|\nabla\varphi(t_0,x_0)|}$. Hence (3.14) follows from (3.13). □

3.3. **Sharp semi-concavity.** In this subsection, we investigate the hypotheses under which the value function φ of our exit-time optimal control problem is semi-concave with respect to x. A semi-concavity result for autonomous exit-time optimal control problems is provided in [17, Theorem 8.2.7] and, up to performing a classical state augmentation technique to regard (3.1) as an autonomous system (which consists of considering $z(t) = (t,\gamma(t))$ as the state), one can readily obtain the semi-concavity of φ with respect to (t,x) provided that $k \in C^{1,1}(\mathbb{R}^+ \times \Omega)$.

By looking at the proof of [17, Theorem 8.2.7], one can also notice that immediate adaptations of the proof allow one to obtain semi-concavity of φ with respect to x as soon as k is $C^{1,1}$ with respect to x and Lipschitz continuous in t. It turns out that, in our setting, we can refine the proof of [17, Theorem 8.2.7] to show that semi-concavity of φ with respect to x can be obtained under a weaker assumption on the behavior of k with respect to t, namely that $\partial_t k$ is lower bounded. This is the main result of this subsection, proved in Theorem 3.23.

We note that semi-concavity of φ is related not only to the regularity of k, but also to the smoothness of the target $\partial\Omega$. We also make use of the fact that the distance function $d(\cdot, \overline{\mathbb{R}^d \setminus \Omega})$ is semi-concave in Ω, which is a consequence of (H4) (or more generally, a uniform exterior ball condition on Ω). Notice that this distance function coincides with the value function φ in the particular case $k \equiv 1$ and $g \equiv 0$, justifying the importance of its properties in the proof of Theorem 3.23.

We first introduce the following estimates on the trajectories, which will be repeatedly used in our analysis.

Proposition 3.22. *Assume that (H2) and (H7) hold and let $t_0, t \in \mathbb{R}^+$. Then there exists $c > 0$, depending only on $t - t_0$, L_x, and L_{xx}, such that, for every $x_0, x_1 \in \Omega$ and every control $u : [t_0, \infty) \to \bar{B}(0, 1)$, one has*

$$|\gamma^{t_0,x_0,u}(t) - \gamma^{t_0,x_1,u}(t)| \leq c|x_0 - x_1|$$

and

$$\left|\gamma^{t_0,x_0,u}(t) + \gamma^{t_0,x_1,u}(t) - 2\gamma^{t_0,\frac{x_0+x_1}{2},u}(t)\right| \leq c|x_0 - x_1|^2.$$

Proposition 3.22 can be proved exactly as in [17, Lemma 7.1.2] and thus its proof is omitted here.

To prove the semi-concavity of φ, we need to assume that (H1), (H2), (H3), (H4), (H5), (H6), and (H7) are satisfied. In addition, we suppose that there exists a constant $\ell_t > 0$ such that, for every $x \in \Omega$, $t \mapsto k(t, x)$ is absolutely continuous and, almost everywhere in $t \in \mathbb{R}^+$,

$$\partial_t k \geq -\ell_t. \tag{H8}$$

Moreover, we assume that

$$g \text{ is semi-concave on } \partial\Omega. \tag{H9}$$

Then, we have the following result.

Theorem 3.23. *The value function φ is semi-concave w.r.t. x, and its semi-concavity constant depends only on λ, $k_{\min}$, $k_{\max}$, κ, L_x, L_{xx}, M, and ℓ_t, where κ is a bound on the curvatures of $\partial\Omega$ and M is the semi-concavity constant of g.*

Proof. Along this proof, c is used to denote positive constants depending only on λ, $k_{\min}$, $k_{\max}$, κ, L_x, L_{xx}, M, and ℓ_t, and the value of these constants may change from one expression to another. Some parts of this proof, in particular Case 1 and the first arguments in Case 2, are treated exactly as in the corresponding parts of the proof of [17, Theorem 8.2.7], and we only detail them here for the sake of completeness.

Let $(t_0, x) \in \mathbb{R}^+ \times \Omega$. For simplicity of exposition, we suppose that $t_0 = 0$. Let $h \in \mathbb{R}^d$ be such that $x - h, x + h \in \Omega$ and u be an optimal control for x, at time 0. We consider the trajectories $\gamma^{0,x,u}$, $\gamma^{0,x-h,u}$, and $\gamma^{0,x+h,u}$, and split the proof into cases according to which of these trajectories arrives first at $\partial\Omega$.

- *Case 1:* $\tau_0 := \tau^{0,x,u} \leq \min\{\tau^{0,x-h,u}, \tau^{0,x+h,u}\}$.

  Since u is optimal for x, at time 0, it follows from Lemma 3.4) that

$$\varphi(0,x-h) + \varphi(0,x+h) - 2\varphi(0,x) \leq \varphi(\tau_0, x^-) + \varphi(\tau_0, x^+) - 2g(\gamma_\tau^{0,x,u}), \tag{3.17}$$

where

$$x^+ := \gamma^{0,x+h,u}(\tau_0) \qquad \text{and} \qquad x^- := \gamma^{0,x-h,u}(\tau_0).$$

Let u^+, u^- be two optimal controls for x^+ and x^-, at time τ_0, respectively, and define $y^\pm = \gamma_\tau^{\tau_0,x^\pm,u^\pm}$ and $\tau^\pm := \tau^{\tau_0,x^\pm,u^\pm}$. Then

$$\varphi(\tau_0, x^-) + \varphi(\tau_0, x^+) - 2g(\gamma_\tau^{0,x,u}) = \tau^- + g(y^-) + \tau^+ + g(y^+) - 2g(\gamma_\tau^{0,x,u}). \tag{3.18}$$

Yet, by Proposition 3.7, we have

$$\tau^\pm \leq cd(x^\pm, \overline{\mathbb{R}^d \setminus \Omega}). \tag{3.19}$$

As the distance function $d(\cdot, \overline{\mathbb{R}^d \setminus \Omega})$ is 1-Lipschitz, semi-concave in $\bar{\Omega}$, and its semi-concavity constant is bounded by κ, and taking into account that $\gamma_\tau^{0,x,u} \in \partial\Omega$, we obtain that

(3.20)

$$\begin{aligned} d(x^+, \overline{\mathbb{R}^d \setminus \Omega}) + d(x^-, \overline{\mathbb{R}^d \setminus \Omega}) &= d(x^+, \overline{\mathbb{R}^d \setminus \Omega}) + d(x^-, \overline{\mathbb{R}^d \setminus \Omega}) - 2d\left(\frac{x^+ + x^-}{2}, \overline{\mathbb{R}^d \setminus \Omega}\right) \\ &\quad + 2\left(d\left(\frac{x^+ + x^-}{2}, \overline{\mathbb{R}^d \setminus \Omega}\right) - d(\gamma_\tau^{0,x,u}, \overline{\mathbb{R}^d \setminus \Omega})\right) \\ &\leq c|x^+ - x^-|^2 + |x^+ + x^- - 2\gamma_\tau^{0,x,u}| \leq c|h|^2, \end{aligned}$$

where the last inequality follows from Proposition 3.22. On the other hand, from the assumptions on g, we have

(3.21)

$$\begin{aligned} g(y^+) + g(y^-) - 2g(\gamma_\tau^{0,x,u}) &= g(y^+) + g(y^-) - 2g\left(\frac{y^+ + y^-}{2}\right) + 2\left(g\left(\frac{y^+ + y^-}{2}\right) - g(\gamma_\tau^{0,x,u})\right) \\ &\leq c\Big(|y^+ - y^-|^2 + |y^+ + y^- - 2\gamma_\tau^{0,x,u}|\Big). \end{aligned}$$

Yet,

$$|y^+ - y^-| \leq |y^+ - x^+| + |x^+ - x^-| + |x^- - y^-| \leq |y^+ - x^+| + |x^- - y^-| + c|h|.$$

In addition, we have

$$|y^\pm - x^\pm| = \left|\int_{\tau_0}^{\tau_0 + \tau^\pm} k(s, \gamma^{\tau_0,x^\pm,u^\pm}(s)) u^\pm(s)\, \mathrm{d}s\right| \leq k_{\max}\tau^\pm \leq c|h|^2,$$

which implies that

$$|y^+ - y^-| \leq c|h|. \tag{3.22}$$

For the second term in (3.21), we have

$$|y^+ + y^- - 2\gamma_\tau^{0,x,u}| \leq |y^+ - x^+| + |x^+ + x^- - 2\gamma_\tau^{0,x,u}| + |x^- - y^-| \leq c|h|^2. \tag{3.23}$$

Consequently, inserting (3.22) and (3.23) into (3.21) and combining this with (3.17), (3.18), (3.19), and (3.20), we conclude that

$$\varphi(0,x-h) + \varphi(0,x+h) - 2\varphi(0,x) \leq c|h|^2.$$

- *Case 2:* $\tau_0 := \tau^{0,x-h,u} \leq \min\{\tau^{0,x,u}, \tau^{0,x+h,u}\}$.

  It suffices to treat this case to conclude the proof, since the other remaining case $\tau^{0,x+h,u} \leq \min\{\tau^{0,x,u}, \tau^{0,x-h,u}\}$ is identical up to exchanging h and $-h$. Let

$$x_0 = \gamma^{0,x-h,u}(\tau_0), \qquad x_1 = \gamma^{0,x,u}(\tau_0), \qquad x_2 = \gamma^{0,x+h,u}(\tau_0).$$

By Lemma 3.4, we have

$$\varphi(0, x-h) + \varphi(0, x+h) - 2\varphi(0, x) \leq \varphi(\tau_0, x_2) - 2\varphi(\tau_0, x_1) + g(x_0), \tag{3.24}$$

By Proposition 3.3, u is also an optimal control starting from x_1, at time τ_0, with $\tau_1 := \tau^{\tau_0,x_1,u} = \tau^{0,x,u} - \tau_0$. As $x_0 \in \partial\Omega$, then, by Propositions 3.7 and 3.22, we get that

$$\tau_1 \leq c d(x_1, \overline{\mathbb{R}^d \setminus \Omega}) \leq c|x_1 - x_0| \leq c|h|. \tag{3.25}$$

Let $u^\star$ be the control defined for $t \geq \tau_0$ by $u^\star(t) := u(\frac{t+\tau_0}{2})$ and consider the trajectory $\gamma^{\tau_0,x_2,u^\star}$. We split the remainder of the proof into two cases requiring separate analyses.

- *Case 2(a):* $\tau_1 < \frac{\tau^{\tau_0,x_2,u^\star}}{2}$.

  By Lemma 3.4,

$$\varphi(\tau_0, x_2) - 2\varphi(\tau_0, x_1) + g(x_0) \leq \varphi(\tau_0 + 2\tau_1, z_2) + g(x_0) - 2g(z_1), \tag{3.26}$$

where $z_1 = \gamma_\tau^{\tau_0,x_1,u} \in \partial\Omega$ and $z_2 = \gamma^{\tau_0,x_2,u^\star}(\tau_0 + 2\tau_1)$. Let v be an optimal control for z_2, at time $\tau_0 + 2\tau_1$, and set $w_2 = \gamma_\tau^{\tau_0+2\tau_1,z_2,v}$. Then, by Proposition 3.7, we have

$$\begin{aligned}
\varphi(\tau_0 + 2\tau_1, z_2) + g(x_0) - 2g(z_1) &= \tau^{\tau_0+2\tau_1,z_2,v} + g(w_2) + g(x_0) - 2g(z_1) \\
&\leq c d(z_2, \overline{\mathbb{R}^d \setminus \Omega}) + g(w_2) + g(x_0) - 2g(z_1) \\
&= c d(z_2, \overline{\mathbb{R}^d \setminus \Omega}) + g(w_2) + g(x_0) \\
&\quad - 2g\left(\frac{x_0 + w_2}{2}\right) + 2\left(g\left(\frac{x_0 + w_2}{2}\right) - g(z_1)\right).
\end{aligned}$$

From (H3) & (H9), we infer that

$$\varphi(\tau_0 + 2\tau_1, z_2) + g(x_0) - 2g(z_1) \leq c\Big[d(z_2, \overline{\mathbb{R}^d \setminus \Omega}) + |w_2 - x_0|^2 + |x_0 + w_2 - 2z_1|\Big]. \tag{3.27}$$

Yet, using Proposition 3.22, we have

$$|w_2 - x_0| \leq |w_2 - z_2| + |z_2 - x_2| + |x_2 - x_0| \leq |w_2 - z_2| + |z_2 - x_2| + c|h|. \tag{3.28}$$

In addition, by Proposition 3.7, one has

$$\begin{aligned}
|w_2 - z_2| &= \left|\int_{\tau_0+2\tau_1}^{\tau_0+2\tau_1+\tau^{\tau_0+2\tau_1,z_2,v}} k\Big(s, \gamma^{\tau_0+2\tau_1,z_2,v}(s)\Big) v(s) \, \mathrm{d}s\right| \\
&\leq k_{\max} \tau^{\tau_0+2\tau_1,z_2,v} \leq c d(z_2, \overline{\mathbb{R}^d \setminus \Omega}).
\end{aligned} \tag{3.29}$$

In the same way, we have, using (3.25), that

$$|z_2 - x_2| = \left|\int_{\tau_0}^{\tau_0+2\tau_1} k\Big(s, \gamma^{\tau_0,x_2,u^\star}(s)\Big) u^\star(s) \, \mathrm{d}s\right| \leq 2k_{\max}\tau_1 \leq c|h|. \tag{3.30}$$

Moreover,

$$|x_0 + w_2 - 2z_1| \leq |x_0 + z_2 - 2z_1| + |w_2 - z_2|. \tag{3.31}$$

Hence, inserting (3.29) and (3.30) into (3.28), and again (3.29) into (3.31), it follows from (3.24), (3.26), and (3.27) that the proof of Case 2(a) is completed if one shows that

$$d(z_2, \overline{\mathbb{R}^d \setminus \Omega}) + |x_0 + z_2 - 2z_1| \leq c|h|^2. \tag{3.32}$$

Note that

$$d(z_2, \overline{\mathbb{R}^d \setminus \Omega}) \leq |z_2 - 2z_1 + x_0| + d(2z_1 - x_0, \overline{\mathbb{R}^d \setminus \Omega}). \tag{3.33}$$

Yet,

$$d(2z_1 - x_0, \overline{\mathbb{R}^d \setminus \Omega}) = d(2z_1 - x_0, \overline{\mathbb{R}^d \setminus \Omega}) + d(x_0, \overline{\mathbb{R}^d \setminus \Omega}) - 2d(z_1, \overline{\mathbb{R}^d \setminus \Omega}),$$

as $x_0, z_1 \in \partial\Omega$. Hence, by the semi-concavity of the distance function $d(\cdot, \overline{\mathbb{R}^d \setminus \Omega})$ in $\bar{\Omega}$,

$$d(2z_1 - x_0, \overline{\mathbb{R}^d \setminus \Omega}) \leq c|z_1 - x_0|^2.$$

Now, using Proposition 3.22, we have

$$|z_1 - x_0| \leq |z_1 - x_1| + |x_1 - x_0| \leq c|h| \tag{3.34}$$

since, by (3.25), we have

$$|z_1 - x_1| = \left| \int_{\tau_0}^{\tau_0+\tau_1} k\Big(s, \gamma^{\tau_0, x_1, u}(s)\Big) u(s) \, \mathrm{d}s \right| \leq k_{\max} \tau_1 \leq c|h|. \tag{3.35}$$

Then $d(2z_1 - x_0, \overline{\mathbb{R}^d \setminus \Omega}) \leq c|h|^2$. Hence, by (3.33), in order to prove (3.32), it suffices to show that

$$|z_2 - 2z_1 + x_0| \leq c|h|^2. \tag{3.36}$$

Let $\mathbf{n}$ be the unit outward normal vector at z_1 and let $w := u(\tau_0 + \tau_1) = -\frac{\nabla g(z_1) - \mu \mathbf{n}}{|\nabla g(z_1) - \mu \mathbf{n}|}$ be the unit optimal control vector at z_1, at time $\tau_0 + \tau_1$ (where μ is the unique constant so that $k(\tau_0 + \tau_1, z_1)|\nabla g(z_1) - \mu \mathbf{n}| = 1$; see Lemma 3.12). If $d = 1$, then there exists $\alpha \in \mathbb{R}$ such that $2z_1 - x_0 - z_2 = \alpha \mathbf{n}$. Otherwise, for $d \geq 2$, notice that, by Proposition 3.21, $\mathbf{n} \cdot w \geq c > 0$, which shows that $\mathbf{n}$ and w are not orthogonal, and thus there exists a unit vector e orthogonal to w such that

$$2z_1 - x_0 - z_2 = \alpha \mathbf{n} + \beta e. \tag{3.37}$$

We also write (3.37) when $d = 1$ using the convention $e = 0$ for this case. Notice that $|\mathbf{n}|^2 \geq |\mathbf{n} \cdot w|^2 + |\mathbf{n} \cdot e|^2$, and thus

$$1 - |\mathbf{n} \cdot e|^2 \geq c. \tag{3.38}$$

We have

$$\begin{aligned} (2z_1 - x_0 - z_2) \cdot \mathbf{n} &= \alpha + \beta e \cdot \mathbf{n}, \\ (2z_1 - x_0 - z_2) \cdot e &= \alpha e \cdot \mathbf{n} + \beta. \end{aligned}$$

Then,

$$\begin{pmatrix} (2z_1 - x_0 - z_2) \cdot \mathbf{n} \\ (2z_1 - x_0 - z_2) \cdot e \end{pmatrix} = \begin{pmatrix} 1 & e \cdot \mathbf{n} \\ e \cdot \mathbf{n} & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

or equivalently,

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 & e \cdot \mathbf{n} \\ e \cdot \mathbf{n} & 1 \end{pmatrix}^{-1} \begin{pmatrix} (2z_1 - x_0 - z_2) \cdot \mathbf{n} \\ (2z_1 - x_0 - z_2) \cdot e \end{pmatrix}.$$

Thus

$$|2z_1 - x_0 - z_2| \leq |\alpha| + |\beta| \leq \frac{2}{1 - (e \cdot \mathbf{n})^2} \Big(|(2z_1 - x_0 - z_2) \cdot \mathbf{n}| + |(2z_1 - x_0 - z_2) \cdot e| \Big), \tag{3.39}$$

where the denominator can be estimated thanks to (3.38). We note that

$$\begin{aligned} & z_2 - 2z_1 + x_0 \\ &= x_0 + x_2 - 2x_1 + \int_{\tau_0}^{\tau_0+2\tau_1} k(s, \gamma^{\tau_0,x_2,u^\star}(s)) u^\star(s)\,\mathrm{d}s - 2\int_{\tau_0}^{\tau_0+\tau_1} k(s, \gamma^{\tau_0,x_1,u}(s)) u(s)\,\mathrm{d}s \\ &= x_0 + x_2 - 2x_1 + \int_{\tau_0}^{\tau_0+2\tau_1} k(s, \gamma^{\tau_0,x_2,u^\star}(s)) u\left(\frac{s+\tau_0}{2}\right)\mathrm{d}s - 2\int_{\tau_0}^{\tau_0+\tau_1} k(s, \gamma^{\tau_0,x_1,u}(s)) u(s)\,\mathrm{d}s \\ &= x_0 + x_2 - 2x_1 + 2\int_{\tau_0}^{\tau_0+\tau_1} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s))\Big) u(s)\,\mathrm{d}s. \end{aligned}$$

Hence,

$$\begin{aligned} (z_2 - 2z_1 + x_0)\cdot e &= (x_0 + x_2 - 2x_1)\cdot e \\ &\quad + 2\int_{\tau_0}^{\tau_0+\tau_1} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s))\Big) u(s)\cdot e\,\mathrm{d}s. \end{aligned} \tag{3.40}$$

Yet, from Proposition 3.22, we have

$$|(x_0 + x_2 - 2x_1)\cdot e| \leq |x_0 + x_2 - 2x_1| \leq c|h|^2.$$

To estimate the second term in (3.40), we first observe that, since u is L_x-Lipschitz continuous by Proposition 3.14, we have, for all $s \in [\tau_0, \tau_0 + \tau_1]$,

$$|u(s) - w| = |u(s) - u(\tau_0 + \tau_1)| \leq c(\tau_0 + \tau_1 - s) \leq c|h| \tag{3.41}$$

using (3.25). This implies that

$$|u(s)\cdot e| = |(u(s) - w)\cdot e| \leq c|h|.$$

Hence, using again (3.25), we get

$$\begin{aligned} &\left|\int_{\tau_0}^{\tau_0+\tau_1} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s))\Big) u(s)\cdot e\,\mathrm{d}s\right| \\ &\qquad\qquad \leq 2k_{\max}\int_{\tau_0}^{\tau_0+\tau_1} |u(s)\cdot e|\,\mathrm{d}s \leq c|h|\tau_1 \leq c|h|^2. \end{aligned}$$

Consequently,

$$|(z_2 - 2z_1 + x_0)\cdot e| \leq c|h|^2.$$

To complete the proof of (3.36), it now suffices, by (3.39), to show that

$$|(2z_1 - x_0 - z_2)\cdot \mathbf{n}| \leq c|h|^2.$$

We have

$$(2z_1 - x_0 - z_2)\cdot \mathbf{n} = (z_1 - x_0)\cdot \mathbf{n} + (z_1 - z_2)\cdot \mathbf{n}.$$

Let $d^\pm$ be defined by (3.12) and recall that $d^\pm$ is $C^{1,1}$ in a neighborhood of $\partial\Omega$. Hence, we have

$$d^\pm(x_0) = d^\pm(z_1) + \nabla d^\pm(z_1)\cdot(x_0 - z_1) + O(|z_1 - x_0|^2).$$

As $x_0, z_1 \in \partial\Omega$ and $\nabla d^\pm(z_1) = \mathbf{n}$, we get

$$(x_0 - z_1)\cdot \mathbf{n} = O(|z_1 - x_0|^2).$$

Yet, by (3.34), $|z_1 - x_0| \leq c|h|$. Then

$$|(z_1 - x_0)\cdot \mathbf{n}| \leq c|h|^2.$$

Moreover, notice that

$$|z_2 - z_1| \le |z_2 - x_2| + |x_2 - x_1| + |x_1 - z_1| \le c|h| \tag{3.42}$$

by (3.30), Proposition 3.22, and (3.35). We have

$$d^{\pm}(z_2) = d^{\pm}(z_1) + \nabla d^{\pm}(z_1) \cdot (z_2 - z_1) + O(|z_2 - z_1|^2).$$

As $z_2 \in \Omega$ and $z_1 \in \partial\Omega$, we get

$$-\mathbf{n} \cdot (z_2 - z_1) + O(|z_2 - z_1|^2) \ge 0.$$

Then (3.42) implies that

$$(z_1 - z_2) \cdot \mathbf{n} \ge -c|h|^2.$$

Consequently,

$$|(2z_1 - x_0 - z_2) \cdot \mathbf{n}| \le (2z_1 - x_0 - z_2) \cdot \mathbf{n} + c|h|^2.$$

To complete the proof of Case 2(a), we are now left to prove that

$$(2z_1 - x_0 - z_2) \cdot \mathbf{n} \le c|h|^2.$$

As in (3.40),

$$\begin{aligned}(2z_1 - x_0 - z_2) \cdot \mathbf{n} = &-(x_0 + x_2 - 2x_1) \cdot \mathbf{n} \\ &- 2\int_{\tau_0}^{\tau_0+\tau_1} \Big(k(2s - \tau_0, \gamma^{\tau_0,x_2,u^\star}(2s - \tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s)) \Big) u(s) \cdot \mathbf{n} \, \mathrm{d}s.\end{aligned}$$

From Proposition 3.22, we get again that

$$-(x_0 + x_2 - 2x_1) \cdot \mathbf{n} \le |x_0 + x_2 - 2x_1| \le c|h|^2.$$

For the second term, we have

$$\begin{aligned}&-\int_{\tau_0}^{\tau_0+\tau_1} \Big(k(2s - \tau_0, \gamma^{\tau_0,x_2,u^\star}(2s - \tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s)) \Big) u(s) \cdot \mathbf{n} \, \mathrm{d}s \\ &= -\int_{\tau_0}^{\tau_0+\tau_1} \Big(k(2s - \tau_0, \gamma^{\tau_0,x_2,u^\star}(2s - \tau_0)) - k(s, \gamma^{\tau_0,x_2,u^\star}(2s - \tau_0)) \Big) u(s) \cdot \mathbf{n} \, \mathrm{d}s \\ &\qquad - \int_{\tau_0}^{\tau_0+\tau_1} \Big(k(s, \gamma^{\tau_0,x_2,u^\star}(2s - \tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s)) \Big) u(s) \cdot \mathbf{n} \, \mathrm{d}s.\end{aligned} \tag{3.43}$$

From (H2), we have

$$\begin{aligned}&\left| \int_{\tau_0}^{\tau_0+\tau_1} \Big(k(s, \gamma^{\tau_0,x_2,u^\star}(2s - \tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s)) \Big) u(s) \cdot \mathbf{n} \, \mathrm{d}s \right| \\ &\qquad \le c \int_{\tau_0}^{\tau_0+\tau_1} |\gamma^{\tau_0,x_2,u^\star}(2s - \tau_0) - \gamma^{\tau_0,x_1,u}(s)| \, \mathrm{d}s.\end{aligned} \tag{3.44}$$

Yet, by Proposition 3.22, one has

$$\begin{aligned}&\left| \gamma^{\tau_0,x_2,u^\star}(2s - \tau_0) - \gamma^{\tau_0,x_1,u}(s) \right| \\ &= \left| x_2 + \int_{\tau_0}^{2s-\tau_0} k(t, \gamma^{\tau_0,x_2,u^\star}(t)) u^\star(t) \, \mathrm{d}t - x_1 - \int_{\tau_0}^{s} k(t, \gamma^{\tau_0,x_1,u}(t)) u(t) \, \mathrm{d}t \right| \\ &\le |x_2 - x_1| + \int_{\tau_0}^{2s-\tau_0} k(t, \gamma^{\tau_0,x_2,u^\star}(t)) \, \mathrm{d}t + \int_{\tau_0}^{s} k(t, \gamma^{\tau_0,x_1,u}(t)) \, \mathrm{d}t \\ &\le c|h| + 3k_{\max}(s - \tau_0).\end{aligned} \tag{3.45}$$

Hence, we get by (3.25) that

$$\left|\int_{\tau_0}^{\tau_0+\tau_1}\Big(k(s,\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s,\gamma^{\tau_0,x_1,u}(s))\Big)u(s)\cdot\mathbf{n}\,\mathrm{d}s\right| \leq c\int_{\tau_0}^{\tau_0+\tau_1}(|h|+(s-\tau_0))\,\mathrm{d}s \leq c|h|^2.$$

We are left to consider the first term of the right-hand side of (3.43). Recalling that $w\cdot\mathbf{n}\geq c$ by Proposition 3.21, we finally obtain, using (H8), (3.25), and (3.41), that

$$\begin{aligned}
&-\int_{\tau_0}^{\tau_0+\tau_1}\Big(k(2s-\tau_0,\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s,\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0))\Big)u(s)\cdot\mathbf{n}\,\mathrm{d}s\\
&\quad= -\int_{\tau_0}^{\tau_0+\tau_1}\Big(k(2s-\tau_0,\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s,\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0))\Big)w\cdot\mathbf{n}\,\mathrm{d}s\\
&\quad-\int_{\tau_0}^{\tau_0+\tau_1}\Big(k(2s-\tau_0,\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s,\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0))\Big)(u(s)-w)\cdot\mathbf{n}\,\mathrm{d}s\\
&\quad\leq \int_{\tau_0}^{\tau_0+\tau_1}\int_s^{2s-\tau_0} -k_t(t,\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0))w\cdot\mathbf{n}\,\mathrm{d}t\,\mathrm{d}s + c|h|^2 \leq c(\tau_1^2+|h|^2)\leq c|h|^2.
\end{aligned}$$

- *Case 2(b):* $\tau_2 := \tau^{\tau_0,x_2,u^\star} \leq 2\tau_1$.

  Set

$$z_1 := \gamma_\tau^{\tau_0,x_1,u},\, z_2 := \gamma_\tau^{\tau_0,x_2,u^\star} \in \partial\Omega.$$

Recall that, by (3.24), it suffices to estimate $\varphi(\tau_0,x_2)-2\varphi(\tau_0,x_1)+g(x_0)$. Using Lemma 3.4 and (H9), we have

$$\begin{aligned}
&\varphi(\tau_0,x_2)-2\varphi(\tau_0,x_1)+g(x_0)\\
&\leq \tau_2+g(z_2)-2\tau_1-2g(z_1)+g(x_0)\\
&= \tau_2-2\tau_1+2\left(g\left(\frac{x_0+z_2}{2}\right)-g(z_1)\right)+g(z_2)+g(x_0)-2g\left(\frac{x_0+z_2}{2}\right)\\
&\leq \tau_2-2\tau_1+2\left(g\left(\frac{x_0+z_2}{2}\right)-g(z_1)\right)+c|z_2-x_0|^2.
\end{aligned}\tag{3.46}$$

Using Proposition 3.22, we obtain that

$$|z_2-x_0|\leq|z_2-x_2|+|x_2-x_0|\leq|z_2-x_2|+c|h|.\tag{3.47}$$

Yet, using (3.25),

$$|z_2-x_2| = \left|\int_{\tau_0}^{\tau_0+\tau_2}k(s,\gamma^{\tau_0,x_2,u^\star}(s))u^\star(s)\,\mathrm{d}s\right|\leq 2k_{\max}\tau_1\leq c|h|.\tag{3.48}$$

On the other hand, using (H9), we have

$$g\left(\frac{x_0+z_2}{2}\right)-g(z_1)\leq\frac{1}{2}\nabla g(z_1)\cdot(x_0+z_2-2z_1)+O(|x_0+z_2-2z_1|^2).\tag{3.49}$$

But it is clear that

$$|x_0 + z_2 - 2z_1| \leq |x_0 + x_2 - 2x_1| + 2\int_{\tau_0}^{\tau_0+\tau_1} |k(s, \gamma^{\tau_0,x_1,u}(s))u(s)| \,\mathrm{d}s$$
$$+ \int_{\tau_0}^{\tau_0+\tau_2} |k(s, \gamma^{\tau_0,x_2,u^\star}(s))u^\star(s)| \,\mathrm{d}s,$$

which implies, using Proposition 3.22 and (3.25), that

$$|x_0 + z_2 - 2z_1| \leq c|h|^2 + 4k_{\max}\tau_1 \leq c|h|. \tag{3.50}$$

So, inserting (3.48) into (3.47) and (3.50) into (3.49), we conclude from (3.24) and (3.46) that the proof of Case 2(b) is completed if one shows that

$$\tau_2 - 2\tau_1 + \nabla g(z_1) \cdot (x_0 + z_2 - 2z_1) \leq c|h|^2.$$

Let $\mathbf{n}$ be the unit outward normal vector at z_1. If $d = 1$, there exists $\alpha \in [-\lambda, \lambda]$ such that $\nabla g(z_1) = \alpha\mathbf{n}$. Otherwise, for $d \geq 2$, there exist a unit vector e orthogonal to $\mathbf{n}$ and $\alpha, \beta \in \mathbb{R}$ such that

$$\nabla g(z_1) = \alpha\mathbf{n} + \beta e. \tag{3.51}$$

We write (3.51) also when $d = 1$ by setting $e = 0$ in this case. Notice that $\alpha^2 + \beta^2 = |\nabla g(z_1)|^2 \leq \lambda^2$. We have

$$\begin{aligned}\nabla g(z_1) \cdot (x_0 + z_2 - 2z_1) &= (\alpha\mathbf{n} + \beta e) \cdot (x_0 + z_2 - 2z_1)\\ &= \alpha\mathbf{n} \cdot (x_0 + z_2 - 2z_1) + \beta e \cdot (x_0 + z_2 - 2z_1).\end{aligned}$$

Similarly to (3.34) and (3.42) from Case 2(a), one can show that

$$|x_0 - z_1| + |z_1 - z_2| \leq c|h|.$$

From (H4) and the fact that $x_0, z_1, z_2 \in \partial\Omega$, we infer that

$$\alpha\mathbf{n} \cdot (x_0 + z_2 - 2z_1) \leq c|h|^2.$$

We are now left to prove that

$$\tau_2 - 2\tau_1 + \beta e \cdot (x_0 + z_2 - 2z_1) \leq c|h|^2.$$

Set

$$z := \gamma^{\tau_0,x_1,u}\left(\tau_0 + \frac{\tau_2}{2}\right).$$

Then, one has

$$\tau_2 - 2\tau_1 + \beta e \cdot (x_0 + z_2 - 2z_1) = \tau_2 - 2\tau_1 + \beta e \cdot (x_0 + z_2 - 2z) + 2\beta e \cdot (z - z_1).$$

Let us observe that

$$|z - z_1| = \left|\int_{\tau_0+\frac{\tau_2}{2}}^{\tau_0+\tau_1} k(s, \gamma^{\tau_0,x_1,u}(s))u(s) \,\mathrm{d}s\right| \leq k_{\max}\left(\tau_1 - \frac{\tau_2}{2}\right).$$

Using $k_{\max}|\beta| \leq k_{\max}\lambda < 1$, we infer that

$$\tau_2 - 2\tau_1 + \beta e \cdot (x_0 + z_2 - 2z) + 2\beta e \cdot (z - z_1) \leq \beta e \cdot (x_0 + z_2 - 2z).$$

So, the aim, now, is to prove that

$$\beta e \cdot (x_0 + z_2 - 2z) \leq c|h|^2. \tag{3.52}$$

Let us observe that

$$
\begin{aligned}
(3.53)\quad & x_0 + z_2 - 2z \\
&= x_0 + x_2 - 2x_1 - 2\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} k(s, \gamma^{\tau_0,x_1,u}(s))u(s)\,\mathrm{d}s + \int_{\tau_0}^{\tau_0+\tau_2} k(s, \gamma^{\tau_0,x_2,u^\star}(s))u^\star(s)\,\mathrm{d}s \\
&= x_0 + x_2 - 2x_1 - 2\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} k(s, \gamma^{\tau_0,x_1,u}(s))u(s)\,\mathrm{d}s + \int_{\tau_0}^{\tau_0+\tau_2} k(s, \gamma^{\tau_0,x_2,u^\star}(s))u\left(\frac{s+\tau_0}{2}\right)\mathrm{d}s \\
&= x_0 + x_2 - 2x_1 + 2\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s))\Big)u(s)\,\mathrm{d}s \\
&= x_0 + x_2 - 2x_1 + 2\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0))\Big)u(s)\,\mathrm{d}s \\
&\qquad + 2\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s))\Big)u(s)\,\mathrm{d}s.
\end{aligned}
$$

Recall that, by Proposition (3.22),

$$
|x_2 + x_0 - 2x_1| \leq c|h|^2. \tag{3.54}
$$

From (H2) and proceeding as in (3.44) and (3.45), we infer that

$$
\begin{aligned}
&\left|\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s))\Big)u(s)\,\mathrm{d}s\right| \\
&\qquad\qquad \leq c\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \left|\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0) - \gamma^{\tau_0,x_1,u}(s)\right|\mathrm{d}s,
\end{aligned}
$$

and, by Proposition 3.22,

$$
\begin{aligned}
&\left|\gamma^{\tau_0,x_2,u^\star}(2s-\tau_0) - \gamma^{\tau_0,x_1,u}(s)\right| \\
&= \left|x_2 + \int_{\tau_0}^{2s-\tau_0} k(t, \gamma^{\tau_0,x_2,u^\star}(t))u^\star(t)\,\mathrm{d}t - x_1 - \int_{\tau_0}^{s} k(t, \gamma^{\tau_0,x_1,u}(t))u(t)\,\mathrm{d}t\right| \\
&\leq |x_2 - x_1| + \left|\int_{\tau_0}^{2s-\tau_0} k(t, \gamma^{\tau_0,x_2,u^\star}(t))u^\star(t)\,\mathrm{d}t\right| + \left|\int_{\tau_0}^{s} k(t, \gamma^{\tau_0,x_1,u}(t))u(t)\,\mathrm{d}t\right| \\
&\leq c|h| + 3k_{\max}(s-\tau_0).
\end{aligned}
$$

Consequently, we get

$$
\left|\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_1,u}(s))\Big)u(s)\,\mathrm{d}s\right| \leq c|h|^2. \tag{3.55}
$$

On the other hand, using the fact from Proposition 3.14 that u is L_x-Lipschitz continuous, we get that

$$
(3.56)\quad \int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0))\Big)u(s)\,\mathrm{d}s
$$

$$
\begin{aligned}
&= \int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) \Big) u(\tau_0+\tau_1)\,\mathrm{d}s \\
&+ \int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) \Big) (u(s)-u(\tau_0+\tau_1))\,\mathrm{d}s \\
&\leq \int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) \Big) u(\tau_0+\tau_1)\,\mathrm{d}s + c|h|^2.
\end{aligned}
$$

We recall that

$$
u(\tau_0+\tau_1) = -\frac{\nabla g(z_1) - \mu \mathbf{n}}{|\nabla g(z_1) - \mu \mathbf{n}|} \qquad \text{and} \qquad k(\tau_0+\tau_1, z_1)|\nabla g(z_1) - \mu \mathbf{n}| = 1,
$$

and so, using (H8) and (3.25), we get

(3.57)

$$
\begin{aligned}
&\int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) \Big) u(\tau_0+\tau_1)\cdot \beta e\,\mathrm{d}s \\
&= -k(\tau_0+\tau_1, z_1)\beta^2 \int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \Big(k(2s-\tau_0, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) - k(s, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0)) \Big)\,\mathrm{d}s \\
&= k(\tau_0+\tau_1, z_1)\beta^2 \int_{\tau_0}^{\tau_0+\frac{\tau_2}{2}} \int_s^{2s-\tau_0} -k_t(t, \gamma^{\tau_0,x_2,u^\star}(2s-\tau_0))\,\mathrm{d}t\,\mathrm{d}s \leq c|h|^2.
\end{aligned}
$$

We then obtain (3.52) by combining (3.53), (3.55), (3.56), and (3.57). □

We finish this section by a remark on the importance of assuming (H8).

Remark 3.24. One can give an example showing that a lower bound on the derivative of the dynamic k with respect to t is a sharp condition to obtain semi-concavity of φ. To see that, let Ω be the unit ball in $\mathbb{R}^d$. Let κ be a differentiable real function with $0 < \kappa_{\min} \leq \kappa \leq \kappa_{\max} < +\infty$. Set $k(t,x) := \kappa(t)$, for every $(t,x) \in \mathbb{R}^+ \times \Omega$. For a given $x \in \Omega$, the optimal trajectory for x, at time 0, will be given by

$$
\gamma'(s) = k(s, \gamma(s)) e(x) = \kappa(s) e(x),
$$

where $e(x) := x/|x|$. Let φ be the value function associated with this optimal control problem. We observe easily that

$$
\int_0^{\varphi(0,x)} \kappa(s)\,\mathrm{d}s = 1 - |x|.
$$

Now, set

$$
G(T) := \int_0^T \kappa(s)\,\mathrm{d}s, \quad \text{for all } T \geq 0
$$

and $H := G^{-1}$. This yields that

$$
\varphi(0,x) = H(1-|x|).
$$

Consequently, we have

$$
\nabla^2 \varphi(0,x) = H''(1-|x|) e(x) \otimes e(x) - \frac{H'(1-|x|)}{|x|} (I - e(x)\otimes e(x)),
$$

where

$$
H' = \frac{1}{\kappa} \circ H \qquad \text{and} \qquad H'' = -\frac{\kappa'}{\kappa^3} \circ H.
$$

This shows that $\nabla^2 \varphi$ cannot be bounded from above unless κ' is bounded from below.

3.4. **Differentiability of the value function.** We prove in this section an extra result on our exit-time optimal control problem, namely that the value function φ is differentiable along optimal trajectories. This kind of result is classical (see [17]) and one can even obtain more (for instance, in [16] smoothness of the value function in a neighborhood of optimal trajectories is proven under some suitable conditions). Yet, most of the literature is concerned with the autonomous case (see in particular [17]), which motivates us to provide a detailed proof here dealing with the subtleties of our non-autonomous setting.

The results of this subsection will be of use in Section 4.1 in order to obtain the continuity equation in (4.7). They require a result stronger than Theorem 3.23, namely the semi-concavity of φ on both variables (t, x), and not only on x. We then need the stronger assumption that

$$k \in C^{1,1}(\mathbb{R}^+ \times \Omega). \tag{H10}$$

Remark 3.25. The sharper assumption (H8) will be of use in Section 4.3 when studying a less regular MFG model. Its study is carried out by an approximation procedure, with approximated dynamics $k_\varepsilon \in C^{1,1}(\mathbb{R}^+ \times \Omega)$ for $\varepsilon > 0$ but with no uniform bounds on their $C^{1,1}$ behavior, except for an uniform lower bound on $\partial_t k_\varepsilon$, which is the motivation for (H8). Since the differentiability of φ along optimal trajectories plays no particular role in this approximation procedure, one may assume the stronger assumption (H10) for the purposes of this section.

Our first result concerns the semi-concavity of φ on (t, x).

Proposition 3.26. *Under assumptions (H1), (H3), (H4), (H9), and (H10), the value function φ is semi-concave on $\mathbb{R}^+ \times \Omega$.*

Proof. We apply the classical semi-concavity result from [17, Theorem 8.2.7] to the augmented system $z' = \widetilde{k}(z, u)$, where $z = (t, \gamma)$ and $\widetilde{k}$ is given by $\widetilde{k}(z, u) = (1, k(z)u)$. □

As a consequence of the semi-concavity of φ on $\mathbb{R}^+ \times \Omega$ and the standard properties of semi-concave functions recalled in Proposition 2.2, one obtains the following result.

Proposition 3.27. *Let $c > 0$ be the constant from Corollary 3.11. Let $(t_0, x_0) \in \mathbb{R}^+ \times \mathring{\Omega}$ and assume that $\nabla\varphi(t_0, x_0)$ exists. Then $|\nabla\varphi(t_0, x_0)| \geq c$.*

Notice that this improves the result of Corollary 3.11 concerning $\nabla\varphi$, since one does not assume differentiability of φ on (t_0, x_0) in the statement of Proposition 3.27, but only the existence of $\nabla\varphi$.

Proof. By Proposition 3.8, φ is Lipschitz continuous on $\mathbb{R}^+ \times \Omega$, and hence it is differentiable almost everywhere. Then there exists sequences $(t_n)_{n\in\mathbb{N}^*}$ in $\mathbb{R}^+$ and $(x_n)_{n\in\mathbb{N}^*}$ in Ω such that φ is differentiable at (t_n, x_n) for every $n \in \mathbb{N}^*$ and $t_n \to t_0$ and $x_n \to x_0$ as $n \to \infty$. In particular, one has $|\nabla\varphi(t_n, x_n)| \geq c$ for every $n \in \mathbb{N}$.

Let $p_n = D\varphi(t_n, x_n)$. Since φ is Lipschitz continuous, p_n is bounded, and hence, up to the extraction of a subsequence, p_n converges to some $p = (p_t, p_x) \in \mathbb{R} \times \mathbb{R}^d$. Then $p \in D^\star\varphi(t_0, x_0)$ and thus $p_x \in \Pi_x(D^\star\varphi(t_0, x_0)) \subset \Pi_x(D^+\varphi(t_0, x_0)) \subset \nabla^+\varphi(t_0, x_0) = \{\nabla\varphi(t_0, x_0)\}$. Hence $|\nabla\varphi(t_0, x_0)| = |p_x| = \lim_{n\to\infty}|\nabla\varphi(t_n, x_n)| \geq c$, as required. □

Another consequence of the semi-concavity of φ is the following.

Proposition 3.28. *Let γ be an optimal trajectory for x_0, at time t_0, and u be the associated optimal control. If φ is differentiable at (t_0, x_0), then φ is differentiable at $(t, \gamma(t))$, for all $t \in [t_0, t_0 + \tau_0)$, where $\tau_0 = \tau^{t_0, x_0, u}$.*

Proof. Fix $t \in [t_0, t_0+\tau_0)$. If φ is differentiable at (t_0, x_0), then the subdifferential $\nabla^-\varphi(t_0, x_0)$ is a singleton, say $\nabla^-\varphi(t_0, x_0) = \{p_0\}$. Now, let p be a solution of (3.11) with initial condition $p(t_0) = p_0$. By Proposition 3.19, $p(t) \in \nabla^-\varphi(t, \gamma(t))$, which implies, in particular, that $\nabla^-\varphi(t, \gamma(t)) \neq \emptyset$. On the other hand, as φ is semi-concave, then $D^+\varphi(t, \gamma(t)) \neq \emptyset$ and so, $\nabla^+\varphi(t, \gamma(t)) \neq \emptyset$. Hence, φ is differentiable with respect to x at $(t, \gamma(t))$. Now, take $(p_t, p_x) \in D^+\varphi(t, \gamma(t))$. Then, $p_x = \nabla\varphi(t, \gamma(t))$. Yet, by Proposition 3.6, we have

$$-p_t + k(t, \gamma(t))|p_x| = 1, \tag{3.58}$$

which implies that p_t is uniquely determined by $\nabla\varphi(t, \gamma(t))$. Consequently, $D^+\varphi(t, \gamma(t))$ is a singleton and so, φ is differentiable at $(t, \gamma(t))$ (thanks again to the semi-concavity of the value function φ). ☐

Remark 3.29. The proof of Proposition 3.28 cannot be extended to include the final time $t_0 + \tau_0$ as (3.58) does not hold a priori at the endpoint of an optimal trajectory.

Proposition 3.30. *Let t_0, x_0, γ, u, τ_0, and p be as in Proposition 3.19. Fix $t_1 \in (t_0, t_0 + \tau_0)$ and set $x_1 := \gamma(t_1)$. Suppose that $p(t_1) \in \Pi_x(D^\star\varphi(t_1, x_1))$, then $p(t) \in \Pi_x(D^\star\varphi(t, \gamma(t)))$, for all $t \in [t_1, t_0 + \tau_0]$.*

Proof. If $p(t_1) \in \Pi_x(D^\star\varphi(t_1, x_1))$, then there is a sequence $(t_{1,n}, x_{1,n}) \in \mathbb{R}^+ \times \Omega$ such that $t_{1,n} \to t_1$, $x_{1,n} \to x_1$ and φ is differentiable at $(t_{1,n}, x_{1,n})$ with $\nabla\varphi(t_{1,n}, x_{1,n}) \to p(t_1)$. As φ is differentiable at $(t_{1,n}, x_{1,n})$, then, by Proposition 3.28, φ is differentiable at $(t, \gamma_n(t))$, for all $t \in [t_{1,n}, t_{1,n} + \tau_{1,n})$, where γ_n is an optimal trajectory for $x_{1,n}$, at time $t_{1,n}$, and $\tau_{1,n} = \tau^{t_{1,n}, x_{1,n}, u_n}$, u_n being the optimal control associated with γ_n. Let p_n be the solution of

$$\begin{cases} \quad p_n'(t) = -\nabla k(t, \gamma_n(t)) u_n(t) \cdot p_n(t), \qquad t \in [t_{1,n}, t_{1,n} + \tau_{1,n}], \\ p_n(t_{1,n}) = \nabla\varphi(t_{1,n}, x_{1,n}). \end{cases} \tag{3.59}$$

By Proposition 3.19, we have $p_n(t) = \nabla\varphi(t, \gamma_n(t))$ for all $t \in [t_{1,n}, t_{1,n} + \tau_{1,n})$. Yet, it is clear, from Lemma 3.16 & Proposition 3.17, that $u_n \to u$ and $\gamma_n \to \gamma$ uniformly, where u is the unique optimal control for x_1, at time t_1, and γ is its associated optimal trajectory. So, we also have $p_n \to p$ uniformly. Now, fix $t \in [t_1, t_0 + \tau_0]$ and let $(t_n)_n$ be any sequence such that $t_n \in (t_{1,n}, t_{1,n} + \tau_{1,n})$, for all n, and $t_n \to t$. As $p_n(t_n) = \nabla\varphi(t_n, \gamma_n(t_n))$, we get that $p(t) = \lim_n \nabla\varphi(t_n, \gamma_n(t_n))$, which means that $p(t) \in \Pi_x(D^\star\varphi(t, \gamma(t)))$. ☐

We are now ready to prove the main result of this subsection.

Theorem 3.31. *Given $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$, let $\gamma : [t_0, t_0 + \tau_0] \to \Omega$ be an optimal trajectory for x_0, at time t_0, where $\tau_0 = \tau^{t_0, x_0, u}$; u being the associated optimal control. Then, φ is differentiable at all points $(t, \gamma(t))$, with $t \in (t_0, t_0 + \tau_0)$.*

Proof. Let us argue by contradiction and suppose that $D^+\varphi(t, \gamma(t))$ is not a singleton for some $t \in (t_0, t_0 + \tau_0)$. Then, thanks to Proposition 2.2, $D^\star\varphi(t, \gamma(t))$ contains at least two elements, say $(p_{t,0}, p_{x,0})$ and $(p_{t,1}, p_{x,1})$. Yet, from Proposition 3.6, we see that different elements of $D^\star\varphi(t, \gamma(t))$ have different space components, i.e. $p_{x,0} \neq p_{x,1}$. For any $\theta \in (0, 1)$, we have $(1-\theta)(p_{t,0}, p_{x,0}) + \theta(p_{t,1}, p_{x,1}) \in D^+\varphi(t, \gamma(t))$ and so, recalling again Proposition 3.6, one has

$$\begin{aligned} -p_{t,0} + k(t, \gamma(t))|p_{x,0}| - 1 = 0, \\ -p_{t,1} + k(t, \gamma(t))|p_{x,1}| - 1 = 0, \end{aligned}$$

and

$$-(1-\theta)p_{t,0} - \theta p_{t,1} + k(t, \gamma(t))|(1-\theta)p_{x,0} + \theta p_{x,1}| - 1 = 0.$$

Hence, we have

$$|(1-\theta)p_{x,0}+\theta p_{x,1}| = (1-\theta)|p_{x,0}|+\theta|p_{x,1}|, \tag{3.60}$$

which implies that $p_{x,1} = \alpha p_{x,0}$, for some $\alpha > 0$, $\alpha \neq 1$. Now, let p_0 and p_1 be the solutions of (3.11), associated with the optimal (γ, u), with initial condition $p_0(t) = p_{x,0}$ and $p_1(t) = p_{x,1}$, respectively. Then $p_1 = \alpha p_0$. In particular, we have $p_1(t_0+\tau_0) = \alpha p_0(t_0+\tau_0)$. Yet, by Proposition 3.30, we know that both $p_0(t_0+\tau_0)$ and $p_1(t_0+\tau_0)$ belong to $\Pi_x(D^\star\varphi(t_0+\tau_0, \gamma(t_0+\tau_0)))$. As $\varphi(t,x) = g(x)$ at every $(t,x) \in \mathbb{R}^+ \times \partial\Omega$, then φ is differentiable with respect to t on $\mathbb{R}^+ \times \partial\Omega$ and, $\partial_t\varphi = 0$. This implies that $\Pi_t(D^\star\varphi(t_0+\tau_0, \gamma(t_0+\tau_0))) = \{0\}$. Hence, we obtain, using Proposition 3.6, that if $q_0, q_1 \in \Pi_x(D^\star\varphi(t_0+\tau_0, \gamma(t_0+\tau_0)))$, then $|q_0| = |q_1|$. This implies that $|p_0(t_0+\tau_0)| = |p_1(t_0+\tau_0)| = \alpha|p_0(t_0+\tau_0)|$, which is a contradiction as $\alpha \neq 1$. Hence, φ is differentiable at $(t, \gamma(t))$, for all $t \in (t_0, t_0+\tau_0)$. □

Remark 3.32. Contrarily to other classical results on the differentiability of the value function along optimal trajectories such as [17, Theorem 8.4.6], we cannot conclude the proof of Theorem 3.31 using only local information on the superdifferential at the point $(t, \gamma(t))$. The main conclusion we obtain from local information is (3.60), which allows us to deduce that $p_{x,0}$ and $p_{x,1}$ are colinear and point to the same direction. In order to obtain the desired contradiction, we need to propagate this information to the boundary, using Proposition 3.30, and exploit the additional information that $\partial_t\varphi$ vanishes on $\partial\Omega$ in order to conclude.

As a consequence of Propositions 3.13 & 3.18 and Theorem 3.31, one can characterize an optimal control u in terms of the normalized gradient, with respect to x, of the value function φ.

Corollary 3.33. *Let $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$ and $\gamma = \gamma^{t_0,x_0,u}$ be an optimal trajectory for x_0, at time t_0, where u is the associated optimal control. Then, for all $t \in (t_0, t_0+\tau_0)$, where $\tau_0 := \tau^{t_0,x_0,u}$, one has*

$$\gamma'(t) = -k(t,\gamma(t))\frac{\nabla\varphi(t,\gamma(t))}{|\nabla\varphi(t,\gamma(t))|}. \tag{3.61}$$

Our final result of this section provides a converse to Corollary 3.33, proving that any solution of (3.61) is an optimal trajectory, and also that such solutions are unique for almost every initial condition.

Proposition 3.34. *Fix $(t_0, x_0) \in \mathbb{R}^+ \times \Omega$ and let $\gamma : [t_0, +\infty) \to \mathbb{R}^d$ be an absolutely continuous function satisfying, for almost every $t \in [t_0, +\infty)$,*

$$\begin{aligned} &\gamma'(t) = \begin{cases} -k(t,\gamma(t))\dfrac{\nabla\varphi(t,\gamma(t))}{|\nabla\varphi(t,\gamma(t))|}, & \text{if } \gamma(t) \in \mathring{\Omega}, \\ 0, & \text{otherwise}, \end{cases} \\ &\gamma(t_0) = x_0. \end{aligned} \tag{3.62}$$

Then γ is an optimal trajectory for x_0 at time t_0. Moreover, for every $t_0 \in \mathbb{R}^+$ and for a.e. $x_0 \in \Omega$, (3.62) admits a unique solution γ.

Proof. Let $t \geq t_0$ be such that (3.62) holds at t. So, this implies, in particular, that φ is differentiable with respect to x at $(t, \gamma(t))$. Thanks to Proposition 3.26, we infer that φ is also differentiable with respect to t. This follows from the fact that if $(p_t, p_x) \in D^\star\varphi(t, \gamma(t))$ then, using Proposition 3.5, p_t is uniquely determined by $p_x = \nabla\varphi(t, \gamma(t))$. Then $D^\star\varphi(t, \gamma(t))$ is a singleton, which implies, by Proposition 2.2, that $D^+\varphi(t, \gamma(t))$ is also a singleton.

Hence, we have

$$-\frac{d}{dt}\varphi(t,\gamma(t)) = -\partial_t\varphi(t,\gamma(t)) - \nabla\varphi(t,\gamma(t))\cdot\gamma'(t) = -\partial_t\varphi(t,\gamma(t)) + k(t,\gamma(t))|\nabla\varphi(t,\gamma(t))| = 1.$$

Integrating the above inequality over $[t_0, t_0+\tau^{t_0,x_0,u}]$ we finally obtain, since $\varphi(t_0+\tau^{t_0,x_0,u}, \gamma_\tau^{t_0,x_0,u}) = g(\gamma_\tau^{t_0,x_0,u})$ where u is the control associated with γ,

$$\varphi(t_0,x_0) = \tau^{t_0,x_0,u} + g(\gamma_\tau^{t_0,x_0,u}).$$

Therefore u is optimal.

For a given $t_0 \in \mathbb{R}^+$, $x \mapsto \varphi(t_0,x)$ is Lipschitz continuous by Proposition 3.8, and then $\nabla\varphi(t_0,x_0)$ exists for almost every $x_0 \in \Omega$. The last statement of the proposition is then a direct consequence of Proposition 3.20. □

4. Optimal-exit mean field games

After the preliminary study of the corresponding optimal control problem in Section 3, we are ready to consider in this section the mean field game model treated in this paper, which we briefly recall. Let $\Omega \subset \mathbb{R}^d$ be compact and $k : \mathcal{P}(\Omega)\times\Omega \to \mathbb{R}^+$ and $g : \partial\Omega \to \mathbb{R}^+$ be continuous (recall that $\mathcal{P}(\Omega)$ is endowed with the topology of weak convergence of measures). We consider the mean field game in which agents evolve in Ω, their distribution at time t being given by a probability measure $\rho_t \in \mathcal{P}(\Omega)$. We assume the initial distribution ρ_0 to be known. The goal of each agent is to leave Ω through its boundary $\partial\Omega$ minimizing the sum of their exit time with the cost $g(z)$ at their exit position $z \in \partial\Omega$. The speed of an agent at the position x at time t is assumed to be bounded by $k(\rho_t, x)$, which means that, for a given agent, their trajectory γ satisfies $|\gamma'(t)| \leq k(\rho_t,\gamma(t))$, and thus depends on the distribution of all agents ρ_t. On the other hand, the distribution of the agents ρ_t itself depends on how agents choose their trajectories γ. We are interested here in *equilibrium* situations, i.e., in situations where, starting from a time evolution of the density of agents $\rho : \mathbb{R}^+ \to \mathcal{P}(\Omega)$, the trajectories γ chosen by agents induce an evolution of the initial distribution of agents ρ_0 that is precisely given by ρ.

In this section, we first provide a precise definition of equilibrium and prove existence of equilibria, obtaining as well a system of PDEs, called the *MFG system*, satisfied by the time-dependent measure ρ_t and the value function of the corresponding optimal control problem. We then prove that, if ρ_0 is absolutely continuous with L^p density, the same holds for ρ_t for $t \geq 0$, with a control on its L^p norm. Finally, thanks to these L^p estimates, we extend the result of existence of equilibria and the corresponding MFG system to a case where k is less regular.

4.1. **Existence of equilibria and the MFG system.** In order to provide the definition of equilibrium used in this paper, let us introduce some notation. Let $\Gamma = C(\mathbb{R}^+,\Omega)$. For a given $\gamma \in \Gamma$, we define its arrival time at $\partial\Omega$ by

$$\tau_\gamma := \inf\{s \geq 0 \,:\, \gamma(s) \in \partial\Omega\},$$

and, if $\tau_\gamma < +\infty$, we write

$$\gamma_\tau := \gamma(\tau_\gamma) \in \partial\Omega.$$

Given $\rho : \mathbb{R}^+ \to \mathcal{P}(\Omega)$ and $x \in \Omega$, we define the set $\Gamma[\rho,x]$ of *admissible trajectories* from x by

$$\Gamma[\rho,x] := \big\{\gamma \in \Gamma \,: \gamma(0) = x,\ |\gamma'(s)| \leq k(\rho_s,\gamma(s)) \text{ for a.e. } s \in (0,\tau_\gamma), \\ \text{and } \gamma'(s) = 0 \text{ for every } s > \tau_\gamma\big\}.$$

With these definitions, one can write the optimal control problem solved by each agent of the mean field game as

$$\inf\{J(\gamma) \,:\, \gamma \in \Gamma[\rho, x]\}, \tag{4.1}$$

where

$$J(\gamma) = \begin{cases} \tau_\gamma + g(\gamma_\tau) & \text{if } \tau_\gamma < +\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

Remark 4.1. If $\gamma \in \Gamma[\rho, x]$, then there is a measurable control $u : \mathbb{R}^+ \to \bar{B}(0,1)$ such that

$$\begin{cases} \gamma'(t) = k(\rho_t, \gamma(t))u(t), & \text{for a.e. } t, \\ \gamma(0) = x. \end{cases} \tag{4.2}$$

System (4.2) can be seen as a control system under the form (3.1) where the dynamic is given by $\widetilde{k}(t, x) = k(\rho_t, x)$ for every $(t, x) \in \mathbb{R}^+ \times \Omega$. This point of view allows one to formulate (4.1) as an optimal control problem when ρ is fixed.

Remark 4.2. Due to the interaction between agents stemming from k, which may be non-local, the behavior of players who have not yet arrived at $\partial\Omega$ may be influenced by the players who already arrived. However, after arriving at $\partial\Omega$, players are no longer submitted to the minimization criterion (4.1), and thus their trajectory γ might in principle be arbitrary after their arrival time τ_γ. The condition that $\gamma'(s) = 0$ for every $s > \tau_\gamma$ is imposed on admissible trajectories γ in order to avoid ambiguity.

The above choice leads to a concentration of agents on the boundary, which is quite artificial from a modeling point of view. For this reason, one may consider, for modeling purposes, that, for k given by (1.1), the function ψ is a cut-off function, equal to 1 everywhere on Ω except on a neighborhood of $\partial\Omega$ and vanishing at $\partial\Omega$ together with all its derivatives. In this way, the interaction term does not take into account agents who already left Ω. Notice, however, that such assumptions on k are not necessary for the results proved in this paper.

We use in this paper a relaxed notion of MFG equilibrium based on a Lagrangian formulation, following the ideas in [9, 15, 20, 24, 58], for which we give existence result. Such a formulation consists of replacing curves of probability measures on Ω with measures on arcs in Ω. For any $t \in \mathbb{R}^+$, we denote by $e_t : \Gamma \to \Omega$ the evaluation map defined by

$$e_t(\gamma) = \gamma(t), \quad \text{for all } \gamma \in \Gamma.$$

For any $\eta \in \mathcal{P}(\Gamma)$, we define the curve ρ^η of probability measures on Ω as

$$\rho^\eta(t) = (e_t)_{\#}\eta, \quad \text{for all } t \in \mathbb{R}^+.$$

Since $e_t : \Gamma \to \Omega$ is continuous, we observe that, if $\eta_n, \eta \in \mathcal{P}(\Gamma)$, $n \geq 1$, are such that $\eta_n \rightharpoonup \eta$, then $\rho^{\eta_n}(t) \rightharpoonup \rho^\eta(t)$ for all $t \in \mathbb{R}^+$. For any fixed $\rho_0 \in \mathcal{P}(\Omega)$, we denote by $\mathcal{P}_{\rho_0}(\Gamma)$ the set of all Borel probability measures η on Γ such that $(e_0)_{\#}\eta = \rho_0$. Notice that $\mathcal{P}_{\rho_0}(\Gamma)$ is nonempty, since it contains $j_{\#}\rho_0$, where $j : \Omega \to \Gamma$ is the continuous map defined by $j(x)(t) = x$ for all $t \in \mathbb{R}^+$. For all $x \in \Omega$ and $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$, we define the set $\Gamma'[\rho^\eta, x]$ of *optimal trajectories* from x by

$$\Gamma'[\rho^\eta, x] := \left\{\gamma \in \Gamma[\rho^\eta, x] \,:\, J(\gamma) = \min_{\Gamma[\rho^\eta, x]} J\right\}.$$

We also find it useful to introduce the set $\Gamma_{k_{\max}}$ of $k_{\max}$-Lipschitz trajectories $\gamma \in \Gamma$, i.e.,

$$\Gamma_{k_{\max}} = \{\gamma \in \Gamma \,:\, |\gamma'(t)| \leq k_{\max} \text{ for a.e. } t \in \mathbb{R}^+\}.$$

Recall that $\Gamma_{k_{\max}}$ is a compact subset of Γ, and, for every $\eta \in \mathcal{P}(\Gamma)$ and $x \in \Omega$, one has $\Gamma'[\rho^\eta, x] \subset \Gamma[\rho^\eta, x] \subset \Gamma_{k_{\max}}$.

The definition of equilibrium used in this paper is the following.

Definition 4.3. Let $\rho_0 \in \mathcal{P}(\Omega)$. We say that $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$ is a *MFG equilibrium* for ρ_0 if

$$\operatorname{spt}(\eta) \subset \bigcup_{x\in\Omega} \Gamma'[\rho^\eta, x].$$

Let us state the assumptions used to guarantee existence of equilibria. The function $k : \mathcal{P}(\Omega) \times \Omega \to \mathbb{R}^+$ is assumed to be continuous. It is reasonable to suppose that k is bounded from above, since it is not natural to assume that an agent's speed might approach $+\infty$. For simplicity, and in order to affirm that there is at least one admissible trajectory γ starting from a point x that reaches the boundary in finite time, we also suppose that k is bounded from below. Hence, as in Section 3, we assume that k satisfies (H1). We also suppose that the counterpart of (H2) holds, namely,

(H11)

$$\exists L_x > 0 \quad \text{such that} \quad |k(\mu, x_0) - k(\mu, x_1)| \leq L_x |x_0 - x_1| \quad \text{for all } x_0, x_1 \in \Omega,\ \mu \in \mathcal{P}(\Omega).$$

Notice that (H1) and (H11) are satisfied for (1.1) if $V : \mathbb{R}^+ \to (0, +\infty)$ and $\chi : \mathbb{R}^d \to \mathbb{R}^+$ are Lipschitz continuous and $\psi : \mathbb{R}^d \to \mathbb{R}^+$ is continuous. Moreover, we suppose, as in Section 3, that $g : \partial\Omega \to \mathbb{R}^+$ satisfies (H3). In particular, from Proposition 3.2, we infer that (4.1) reaches a minimum.

We can now state our result on the existence of equilibria.

Theorem 4.4. *Let $\rho_0 \in \mathcal{P}(\Omega)$, $k : \mathcal{P}(\Omega) \times \Omega \to \mathbb{R}^+$ be continuous, $g : \partial\Omega \to \mathbb{R}^+$, and assume that (H1), (H3), and (H11) hold. Then there exists a MFG equilibrium for ρ_0.*

The proof of Theorem 4.4 is based on the same fixed-point strategy used in [15, 58]. Notice that the above theorem is slightly stronger than [58, Theorem 5.1] since existence of equilibria is obtained under weaker assumptions. For this reason, and also for the sake of completeness, we provide a detailed proof of Theorem 4.4. The first step is the following property of the map $(\eta, x) \mapsto \Gamma'[\rho^\eta, x]$.

Lemma 4.5. *Let ρ_0, k, and g be as the statement of Theorem 4.4. Let $(\eta_n)_n$ be a sequence in $\mathcal{P}_{\rho_0}(\Gamma)$, $(x_n)_n$ a sequence in Ω, and $(\gamma_n)_n$ a sequence in Γ such that $\gamma_n \in \Gamma'[\rho^{\eta_n}, x_n]$ for every $n \in \mathbb{N}$ and $\eta_n \rightharpoonup \eta$, $x_n \to x$, and $\gamma_n \to \bar\gamma$ for some $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$, $x \in \Omega$, and $\bar\gamma$ in Γ. Then $\bar\gamma \in \Gamma'[\rho^\eta, x]$. Consequently, $(\eta, x) \mapsto \Gamma'[\rho^\eta, x]$ has a closed graph.*

Proof. We set, for simplicity, $\tau_n := \tau_{\gamma_n}$ and $z_n := \gamma_n(\tau_n)$. Using Proposition 3.7, $(\tau_n)_n$ is bounded and, up to extracting a subsequence, τ_n converges to some $\bar\tau$. On the other hand, we see easily that $\bar\gamma$ is $k_{\max}$-Lipschitz continuous. In addition, for a.e. $t \in (0, \tau_n)$, we have $|\gamma_n'(t)| \leq k(\rho^{\eta_n}(t), \gamma_n(t))$. Letting $n \to +\infty$, we get that $|\bar\gamma'(t)| \leq k(\rho^\eta(t), \bar\gamma(t))$ for a.e. $t \in (0, \bar\tau)$. In the same way, one can prove that $\bar\gamma'(t) = 0$ for all $t > \bar\tau$. Moreover, we have $z_n \to \bar\gamma(\bar\tau)$, which implies that $\bar\gamma(\bar\tau) \in \partial\Omega$ and $\tau := \tau_{\bar\gamma} \leq \bar\tau$. Notice that $\bar\gamma \in \Gamma[\rho^\eta, x]$ if and only if $\tau = \bar\tau$.

Define the trajectory $\gamma \in \Gamma[\rho^\eta, x]$ by

$$\gamma(t) = \begin{cases} \bar\gamma(t), & \text{if } t \leq \tau, \\ \bar\gamma(\tau), & \text{if } t > \tau. \end{cases}$$

Notice that, by Lemma 3.1, $J(\gamma) \leq \bar\tau + g(\bar\gamma(\bar\tau))$, with a strict inequality if and only if $\tau < \bar\tau$. Suppose, to obtain a contradiction, that $\bar\gamma \notin \Gamma'[\rho^\eta, x]$. Then there exists a trajectory $\widehat\gamma \in$

$\Gamma'[\rho^\eta, x]$ such that $J(\widehat{\gamma}) < \bar{\tau} + g(\bar{\gamma}(\bar{\tau}))$; indeed, this follows by the definition of $\Gamma'[\rho^\eta, x]$ if $\tau = \bar{\tau}$ or by the fact that $\gamma \in \Gamma[\rho^\eta, x]$ and $J(\gamma) < \bar{\tau} + g(\bar{\gamma}(\bar{\tau}))$ if $\tau < \bar{\tau}$.

For each $n \in \mathbb{N}$, let $\widetilde{\gamma}_n : [0, |x_n - x|] \to \mathbb{R}^d$ be the segment from x_n to x with $|\widetilde{\gamma}_n'(t)| = 1$ for every $t \in [0, |x_n - x|]$. Let $\phi_n : [k_{\min}^{-1}|x_n - x|, +\infty) \to \mathbb{R}^+$ be a function satisfying

$$\begin{cases} \phi_n'(t) = \dfrac{k(\rho^{\eta_n}(t), \widehat{\gamma}(\phi_n(t)))}{k(\rho^\eta(\phi_n(t)), \widehat{\gamma}(\phi_n(t)))}, \\ \phi_n(k_{\min}^{-1}|x_n - x|) = 0. \end{cases} \tag{4.3}$$

Define $\widehat{\gamma}_n : \mathbb{R}^+ \to \mathbb{R}^d$ by

$$\widehat{\gamma}_n(t) = \begin{cases} \widetilde{\gamma}_n(k_{\min} t) & \text{if } t \in [0, k_{\min}^{-1}|x_n - x|], \\ \widehat{\gamma}(\phi_n(t)) & \text{otherwise.} \end{cases}$$

One has $\widehat{\gamma}_n(\phi_n^{-1}(\tau_{\widehat{\gamma}})) = \widehat{\gamma}(\tau_{\widehat{\gamma}})$, and thus $\tau_{\widehat{\gamma}_n} \leq \phi_n^{-1}(\tau_{\widehat{\gamma}})$. Hence, by Lemma 3.1,

$$\tau_{\widehat{\gamma}_n} + g(\widehat{\gamma}_n(\tau_{\widehat{\gamma}_n})) \leq \phi_n^{-1}(\tau_{\widehat{\gamma}}) + g\Big(\widehat{\gamma}_n\big(\phi_n^{-1}(\tau_{\widehat{\gamma}})\big)\Big) = \phi_n^{-1}(\tau_{\widehat{\gamma}}) + g(\widehat{\gamma}_\tau). \tag{4.4}$$

We modify $\widehat{\gamma}_n$ on the interval $(\tau_{\widehat{\gamma}_n}, +\infty)$ by setting $\widehat{\gamma}_n(t) = \widehat{\gamma}_n(\tau_{\widehat{\gamma}_n})$ for $t > \tau_{\widehat{\gamma}_n}$. This modification does not change $\tau_{\widehat{\gamma}_n}$ and one has now $\widehat{\gamma}_n \in \Gamma[\rho^{\eta_n}, x_n]$. In particular, (4.4) reads

$$J(\widehat{\gamma}_n) \leq \phi_n^{-1}(\tau_{\widehat{\gamma}}) + g(\widehat{\gamma}_\tau). \tag{4.5}$$

Since $(\phi_n)_n$ and $(\phi_n^{-1})_n$ are equi-Lipschitz sequences, it follows from Arzelà–Ascoli Theorem that, up to extracting subsequences, there exists a bi-Lipschitz function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\phi_n \to \phi$ and $\phi_n^{-1} \to \phi^{-1}$ uniformly on compact sets of $\mathbb{R}^+$. In addition, it is easy to check by integrating (4.3) that, for all $t \in [k_{\min}^{-1}|x_n - x|, +\infty)$,

$$\int_0^{\phi_n(t)} k(\rho^\eta(s), \widehat{\gamma}(s))\,\mathrm{d}s = \int_{k_{\min}^{-1}|x_n - x|}^t k(\rho^{\eta_n}(s), \widehat{\gamma}(\phi_n(s)))\,\mathrm{d}s.$$

So, letting $n \to +\infty$, we get, for all $t \in \mathbb{R}^+$,

$$\int_0^{\phi(t)} k(\rho^\eta(s), \widehat{\gamma}(s))\,\mathrm{d}s = \int_0^t k(\rho^\eta(s), \widehat{\gamma}(\phi(s)))\,\mathrm{d}s.$$

Set

$$G(\theta) = \int_0^\theta k(\rho^\eta(s), \widehat{\gamma}(s))\,\mathrm{d}s, \quad \forall \theta \in \mathbb{R}^+.$$

Then $G : \mathbb{R}^+ \to \mathbb{R}^+$ is a bi-Lipschitz bijection and, for $t \in \mathbb{R}^+$,

$$\begin{aligned} |\phi(t) - t| &= \left| G^{-1}\left(\int_0^t k(\rho^\eta(s), \widehat{\gamma}(\phi(s)))\,\mathrm{d}s\right) - G^{-1}\left(\int_0^t k(\rho^\eta(s), \widehat{\gamma}(s))\,\mathrm{d}s\right)\right| \\ &\leq C \int_0^t |k(\rho^\eta(s), \widehat{\gamma}(\phi(s))) - k(\rho^\eta(s), \widehat{\gamma}(s))|\,\mathrm{d}s \\ &\leq C \int_0^t |\phi(s) - s|\,\mathrm{d}s. \end{aligned}$$

By Grönwall's lemma, we get that $\phi(t) = t$ for all $t \in \mathbb{R}^+$. Passing to the limit in (4.5), we get

$$\limsup_n J(\widehat{\gamma}_n) \leq \tau_{\widehat{\gamma}} + g(\widehat{\gamma}_\tau) = J(\widehat{\gamma}) < \bar{\tau} + g(\bar{\gamma}(\bar{\tau})). \tag{4.6}$$

Yet,

$$\lim_n J(\gamma_n) = \lim_n \tau_n + g(z_n) = \bar\tau + g(\bar\gamma(\bar\tau)).$$

Using (4.6), we infer that, for n large enough,

$$J(\widehat\gamma_n) < J(\gamma_n),$$

which is a contradiction, as $\widehat\gamma_n \in \Gamma[\rho^{\eta_n}, x_n]$ and $\gamma_n \in \Gamma'[\rho^{\eta_n}, x_n]$. Then $\bar\gamma \in \Gamma'[\rho^\eta, x]$. □

Remark 4.6. As a consequence of Lemma 4.5, for a given $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$, the graph G of the map $x \mapsto \Gamma'[\rho^\eta, x]$ is closed in $\Omega\times\Gamma$. Since $\Gamma'[\rho^\eta, x] \subset \Gamma_{k_{\max}}$, G is compact, since it is a closed subset of the compact set $\Omega\times\Gamma_{k_{\max}}$. Hence, the set $\bigcup_{x\in\Omega}\Gamma'[\rho^\eta, x]$, which is the projection of G onto Γ, is also compact, and, in particular, a measure $\widetilde\eta \in \mathcal{P}(\Gamma)$ satisfies $\operatorname{spt}(\widetilde\eta) \subset \bigcup_{x\in\Omega}\Gamma'[\rho^\eta, x]$ if and only if $\widetilde\eta\big[\bigcup_{x\in\Omega}\Gamma'[\rho^\eta, x]\big] = 1$.

In particular, one can reformulate Definition 4.3 in an equivalent way by saying that $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$ is a MFG equilibrium for ρ_0 if

$$\eta\left[\bigcup_{x\in\Omega}\Gamma'[\rho^\eta, x]\right] = 1,$$

i.e., if for η-a.e. $\bar\gamma \in \Gamma$, we have

$$J(\bar\gamma) \le J(\gamma), \quad \text{for all } \gamma \in \Gamma[\rho^\eta, \bar\gamma(0)].$$

We now reformulate the notion of equilibrium as a fixed point problem, in order to prove Theorem 4.4 using a fixed-point argument. We introduce the set-valued map $E : \mathcal{P}_{\rho_0}(\Gamma) \rightrightarrows \mathcal{P}_{\rho_0}(\Gamma)$ given, for $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$, by

$$E(\eta) = \left\{\widetilde\eta \in \mathcal{P}_{\rho_0}(\Gamma) \,:\, \operatorname{spt}(\widetilde\eta) \subset \bigcup_{x\in\Omega}\Gamma'[\rho^\eta, x]\right\}.$$

It follows immediately that $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$ is a MFG equilibrium for ρ_0 if and only if $\eta \in E(\eta)$, which is precisely the definition of fixed point for a set-valued map. We will therefore prove Theorem 4.4 by showing that E admits a fixed point using Kakutani's Theorem (see, e.g., [40, §7, Theorem 8.6], [51]), whose assumptions we verify in the next lemma.

Lemma 4.7. *Let ρ_0, k, and g be as the statement of Theorem 4.4. Then*
(a) *for any $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$, $E(\eta)$ is a nonempty convex set; and*
(b) *$E : \mathcal{P}_{\rho_0}(\Gamma) \rightrightarrows \mathcal{P}_{\rho_0}(\Gamma)$ has a closed graph.*

Proof. To prove (a), fix $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$. Using Remark 4.6, one immediately verifies that $E(\eta)$ is convex. To see that it is nonempty, notice that, by Lemma 4.5 and [8, Theorem 8.1.3], the map $x \mapsto \Gamma'[\rho^\eta, x]$ has a Borel measurable selection $\boldsymbol{\gamma}^\eta : x \mapsto \gamma^\eta_x \in \Gamma'[\rho^\eta, x]$, and one immediately verifies that $\boldsymbol{\gamma}^\eta_{\#}\rho_0 \in E(\eta)$.

Now, to prove (b), let $(\eta_n)_n$ and $(\widehat\eta_n)_n$ be sequences in $\mathcal{P}_{\rho_0}(\Gamma)$ and $\eta, \widehat\eta \in \mathcal{P}_{\rho_0}(\Gamma)$ such that $\widehat\eta_n \in E(\eta_n)$ for every $n \in \mathbb{N}$, $\eta_n \rightharpoonup \eta$, and $\widehat\eta_n \rightharpoonup \widehat\eta$. For $k \in \mathbb{N}^\star$, let $V_k := \{\gamma \in \Gamma \,:\, d(\gamma, \bigcup_x \Gamma'[\rho^\eta, x]) \le \frac{1}{k}\}$. Notice that the graph G of the set-valued map $\widetilde\eta \mapsto \bigcup_x \Gamma'[\rho^{\widetilde\eta}, x]$ is the projection onto $\mathcal{P}_{\rho_0}(\Gamma)\times\Gamma$ of the graph of the set-valued map from Lemma 4.5, and thus, since Ω is compact, G is closed. Then, using [8, Proposition 1.4.8], it follows that there exists a neighborhood W of η such that $\bigcup_x \Gamma'[\rho^{\widetilde\eta}, x] \subset V_k$ for every $\widetilde\eta \in W$. Then, for n large enough, $\bigcup_x \Gamma'[\rho^{\eta_n}, x] \subset V_k$. Since $\widehat\eta_n(\bigcup_x \Gamma'[\rho^{\eta_n}, x]) = 1$, one obtains that $\widehat\eta_n(V_k) = 1$, for large n. Yet, $\widehat\eta_n \rightharpoonup \widehat\eta$ and V_k is closed, hence it follows that $\widehat\eta(V_k) \ge \limsup_n \widehat\eta_n(V_k) = 1$ and thus, $\widehat\eta(V_k) = 1$. As this holds for every $k \in \mathbb{N}^\star$, one concludes that $\widehat\eta(\bigcup_x \Gamma'[\rho^\eta, x]) = 1$. Hence $\widehat\eta \in E(\eta)$, which proves that the graph of E is closed. □

Remark 4.8. The set $\mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$ is a compact convex subset of $\mathcal{P}_{\rho_0}(\Gamma)$. Indeed, the convexity of $\mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$ follows immediately. As for compactness, if $(\eta_k)_k$ is a sequence in $\mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$, then, since $\Gamma_{k_{\max}}$ is compact, $(\eta_k)_k$ is tight, and so, by Prokhorov's Theorem, one finds a subsequence which converges weakly to some probability measure $\eta \in \mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$.

Notice that, by the definition of E, we have

$$E(\eta) \subset \mathcal{P}_{\rho_0}(\Gamma_{k_{\max}}), \quad \text{for all } \eta \in \mathcal{P}_{\rho_0}(\Gamma).$$

In particular, any fixed point of E belongs to $\mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$. We will thus restrict our domain of interest to $\mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$ with no loss of generality, denoting hereafter by E the restriction $E|_{\mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})}$. Notice that Lemma 4.7 still holds for this restriction. One can now complete the proof of Theorem 4.4.

Proof of Theorem 4.4. Lemma 4.7 guarantees that the set-valued map E has a closed graph and $E(\eta)$ is a nonempty convex set for any $\eta \in \mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$. Since, by Remark 4.8, $\mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$ is a nonempty compact convex set, all assumptions of Kakutani's Theorem are satisfied and thus there exists $\eta \in \mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$ such that $\eta \in E(\eta)$, i.e., η is a MFG equilibrium for ρ_0. □

Now that existence of a MFG equilibrium $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$ has been established, we wish to prove that, similarly to most mean field game models, the corresponding time-dependent measure $\rho_t = \rho^\eta(t)$ satisfies, together with the value function φ of the corresponding optimal control problem, a system of PDEs, known as MFG system, composed of a continuity equation under the form $\partial_t \rho + \nabla \cdot (\rho v) = 0$ for some velocity field v and a Hamilton–Jacobi equation on φ. The Hamilton–Jacobi equation on φ is the one from Proposition 3.5, and one can easily obtain that ρ satisfies some continuity equation (for instance, by proving that $t \mapsto \rho_t$ is Lipschitz continuous with respect to the Wasserstein distance W_p for $p > 1$, as in [58, Proposition 5.2(a)], and then applying [7, Theorem 8.3.1]). The main point here is to identify the velocity field of the continuity equation. To do so, we shall use the results from Section 3.4, which in particular require assumption (H10). We then introduce the following notion.

Definition 4.9. Let $k : \mathcal{P}(\Omega) \times \Omega \to \mathbb{R}^+$ be continuous, $g : \partial\Omega \to \mathbb{R}^+$, $\rho_0 \in \mathcal{P}(\Omega)$, and assume that (H1), (H3), and (H11) hold. We say that k is $C^{1,1}$ *on MFG equilibria for* ρ_0 if, for every MFG equilibrium $\eta \in \mathcal{P}_{\rho_0}(\Gamma)$ for ρ_0, the function $(t, x) \mapsto k(\rho^\eta(t), x)$ is $C^{1,1}$ on $\mathbb{R}^+ \times \Omega$.

To motivate this definition, we prove that the function k given by (1.1) is $C^{1,1}$ on MFG equilibria under suitable regularity assumptions on V, χ, ψ, $\partial\Omega$, and g.

Proposition 4.10. *Let $V \in C^{1,1}(\mathbb{R}^+, (0, +\infty))$ be Lipschitz continuous, $\chi \in C^{1,1}(\mathbb{R}^d, \mathbb{R}^+)$, $\psi \in C^{1,1}(\mathbb{R}^d, \mathbb{R}^+)$, $k : \mathcal{P}(\Omega) \times \mathbb{R}^d \to \mathbb{R}^+$ be given by (1.1), $k_{\max} = \sup_{\mathbb{R}^+ \times \Omega} k$, and $\rho_0 \in \mathcal{P}(\Omega)$. Suppose that $\psi(x) = 0$ and $\nabla\psi(x) = 0$ for every $x \in \partial\Omega$, (H4) holds, and $g : \partial\Omega \to \mathbb{R}^+$ satisfies (H3) and (H6). Then k is $C^{1,1}$ on MFG equilibria for ρ_0.*

Proof. Notice first that $k : \mathcal{P}(\Omega) \times \mathbb{R}^d \to \mathbb{R}^+$ is continuous and satisfies (H1) and (H11). Let $\rho_0 \in \mathcal{P}(\Omega)$, $\eta \in \mathcal{P}_{\rho_0}(\Gamma_{k_{\max}})$ be a MFG equilibrium for ρ_0, and $\rho_t = \rho^\eta(t)$ for $t \geq 0$. Let $\kappa : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^+$ be given by

$$\kappa(t, x) = \int_\Omega \chi(x - y)\psi(y) \, \mathrm{d}\rho_t(y).$$

Since $V \in C^{1,1}(\mathbb{R}^+, (0, +\infty))$, it suffices to prove that $\kappa \in C^{1,1}(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R}^+)$.

Set $\Gamma' = \bigcup_{x \in \Omega} \Gamma'[\rho^\eta, x] \subset \Gamma_{k_{\max}}$. Since η is a MFG equilibrium, one has $\eta(\Gamma') = 1$. Notice that

$$\kappa(t, x) = \int_{\Gamma'} \chi(x - \gamma(t))\psi(\gamma(t)) \, \mathrm{d}\eta(\gamma),$$

and, since every $\gamma \in \Gamma'$ is $k_{\max}$-Lipschitz, one obtains that $(t,x) \mapsto k(\rho_t, x)$ is Lipschitz continuous.

For $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$,

$$\nabla\kappa(t,x) = \int_\Omega \nabla\chi(x-y)\psi(y)\,\mathrm{d}\rho_t(y) = \int_{\Gamma'} \nabla\chi(x-\gamma(t))\psi(\gamma(t))\,\mathrm{d}\eta(\gamma),$$

and this function can be easily seen to be Lipschitz continuous on $\mathbb{R}^+ \times \mathbb{R}^d$. In particular, the function $(t,x) \mapsto k(\rho_t, x)$ satisfies (H5) and (H7). Hence, the results of Section 3.2 apply to the optimal control problem (4.1), and, in particular, by Proposition 3.14, one obtains that $\gamma \in C^{1,1}([0,\tau_\gamma],\Omega)$ for every $\gamma \in \Gamma'$.

For every $\gamma \in \Gamma'$, the function $t \mapsto \chi(x-\gamma(t))\psi(\gamma(t))$ is differentiable everywhere on $\mathbb{R}^+$, except possibly at $t = \tau_\gamma$, with

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[\chi(x-\gamma(t))\psi(\gamma(t))\big] = -\nabla\chi(x-\gamma(t))\cdot\gamma'(t)\psi(\gamma(t)) + \chi(x-\gamma(t))\nabla\psi(\gamma(t))\cdot\gamma'(t).$$

Since $\psi(x) = 0$ and $\nabla\psi(x) = 0$ for $x \in \partial\Omega$ and $\gamma(t) \in \partial\Omega$ for $t = \tau_\gamma$, one can also prove that the above function is differentiable and its derivative is zero at $t = \tau_\gamma$. Moreover, its derivative is Lipschitz continuous and upper bounded, and thus $\partial_t\kappa(t,x)$ exists, with

$$\partial_t\kappa(t,x) = \int_{\Gamma'}\Big[-\nabla\chi(x-\gamma(t))\cdot\gamma'(t)\psi(\gamma(t)) + \chi(x-\gamma(t))\nabla\psi(\gamma(t))\cdot\gamma'(t)\Big]\,\mathrm{d}\eta(\gamma),$$

and one immediately verifies using the previous assumptions that $\partial_t\kappa$ is Lipschitz continuous in $\mathbb{R}^+ \times \mathbb{R}^d$. Together with the corresponding property for $\nabla\kappa$, we obtain that $\kappa \in C^{1,1}(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R}^+)$. □

We now show that, for every MFG equilibrium η, ρ^η and the corresponding value function satisfy a MFG system.

Theorem 4.11. *Let $k : \mathcal{P}(\Omega)\times\Omega \to \mathbb{R}^+$ be continuous, $g : \partial\Omega \to \mathbb{R}^+$, $\rho_0 \in \mathcal{P}(\Omega)$, and assume that (H1), (H3), (H4), (H9), and (H11) hold. Suppose that k is $C^{1,1}$ on MFG equilibria for ρ_0. Let $\eta \in \mathcal{P}_{\rho_0}(\Omega)$ be a MFG equilibrium for ρ_0, $\rho = \rho^\eta$, and φ be the value function of the optimal control problem (4.1) with dynamic $(t,x) \mapsto k(\rho_t, x)$. Then (ρ,φ) solve the* MFG system

$$\begin{cases} \partial_t\rho(t,x) - \nabla\cdot\left(\rho(t,x)k(\rho_t,x)\dfrac{\nabla\varphi(t,x)}{|\nabla\varphi(t,x)|}\right) = 0, & (t,x) \in (0,\infty)\times\mathring{\Omega}, \\ -\partial_t\varphi(t,x) + k(\rho_t,x)|\nabla\varphi(t,x)| - 1 = 0, & (t,x) \in \mathbb{R}^+\times\Omega, \\ \varphi(t,x) = g(x), & (t,x) \in \mathbb{R}^+\times\partial\Omega, \\ \rho(0,x) = \rho_0(x), & x \in \Omega, \end{cases} \tag{4.7}$$

where the first and second equations are satisfied, respectively, in the sense of distributions and in the viscosity sense.

Proof. The second equation in (4.7) and the corresponding boundary condition have already been established in Proposition 3.5. We are left to prove that ρ satisfies the continuity equation in (4.7).

Let $\phi \in C_c^\infty((0,\infty)\times\mathring{\Omega})$ and set $\Gamma' = \bigcup_{x\in\Omega}\Gamma'[\rho^\eta, x]$. Then, recalling Theorem 3.31 and Corollary 3.33, we have

$$-\int_0^{+\infty}\int_\Omega \partial_t\phi(t,x)\,\mathrm{d}\rho_t(x)\,\mathrm{d}t + \int_0^{+\infty}\int_\Omega k(\rho_t,x)\nabla\phi(t,x)\cdot\frac{\nabla\varphi(t,x)}{|\nabla\varphi(t,x)|}\,\mathrm{d}\rho_t(x)\,\mathrm{d}t$$

$$= -\int_0^{+\infty}\int_{\Gamma'} \partial_t\phi(t,\gamma(t))\,\mathrm{d}\eta(\gamma)\,\mathrm{d}t + \int_0^{+\infty}\int_{\Gamma'} k(\rho_t,\gamma(t))\nabla\phi(t,\gamma(t))\cdot\frac{\nabla\varphi(t,\gamma(t))}{|\nabla\varphi(t,\gamma(t))|}\,\mathrm{d}\eta(\gamma)\,\mathrm{d}t$$
$$= -\int_0^{+\infty}\int_{\Gamma'} \partial_t\phi(t,\gamma(t))\,\mathrm{d}\eta(\gamma)\,\mathrm{d}t - \int_0^{+\infty}\int_{\Gamma'} \nabla\phi(t,\gamma(t))\cdot\gamma'(t)\,\mathrm{d}\eta(\gamma)\,\mathrm{d}t$$
$$= -\int_{\Gamma'}\int_0^{+\infty} \frac{d}{dt}\big[\phi(t,\gamma(t))\big]\,\mathrm{d}t\,\mathrm{d}\eta(\gamma) = 0. \qquad \square$$

4.2. L^p **estimates.** Recall that our motivation for the mean field game model in this paper comes from crowd motion, where a reasonable expression for k is (1.1). In order to apply the existence result from Theorem 4.4 to this setting, one should require the function ψ in (1.1) to be at least continuous. On the other hand, as stated in Remark 4.2, agents concentrate on the boundary. A reasonable feature of our model would be to assume that agents do not take into account in their congestion term other agents that have already left the domain, which can be done by assuming that $\psi(x) = 0$ for $x \in \partial\Omega$. However, due to the continuity of ψ, this implies that agents that are too close to the boundary, but have not yet left, will also be somehow discounted.

From a modeling point of view, an interesting choice would be to take $\psi = \mathbb{1}_{\mathring{\Omega}}$, but this yields a function k that is discontinuous on measures μ such that $\mu(\partial\Omega) > 0$, and the arguments used in the proof of Theorem 4.4 do not apply. On the other hand, one may still expect to have existence of equilibria, at least when ρ_0 is absolutely continuous with respect to the Lebesgue measure. The goal of this section and the following is to establish a result on the existence of equilibria in this setting. We first prove that, as soon as ρ_0 is absolutely continuous and with an L^p density, $\rho_t|_{\mathring{\Omega}}$ is also absolutely continuous and with an L^p density, with a control on the L^p norm that is, in some sense, independent of ψ. This will be a key result for the proof of existence of an equilibrium with $\psi = \mathbb{1}_{\mathring{\Omega}}$ in Section 4.3, which is based on a limit argument on a sequence ψ_ε converging to $\mathbb{1}_{\mathring{\Omega}}$ as $\varepsilon \to 0$.

The control of the L^p norm we prove in this section depends essentially on the semi-concavity constant of the value function φ at equilibrium. On the other hand, for k given by (1.1), it follows from Theorem 3.23 that, for uniformly bounded functions ψ, the semi-concavity constant of φ may depend on ψ only through a lower bound on $\partial_t k$. We then start by proving that, for reasonable choices of ψ, one can obtain a lower bound on $\partial_t k$ independent of ψ. We shall consider as reasonable choices of ψ those belonging to the class Ψ_δ defined for $\delta > 0$ by

$$\begin{aligned}\Psi_\delta = \{\psi : \mathbb{R}^d \to [0,1] \mid \exists \alpha \in C^{1,1}(\mathbb{R},[0,1]) &\text{ such that } \alpha \text{ is non-increasing},\\ &\alpha(x) = 0 \text{ for } x \geq 0,\ \alpha'(0) = 0,\ \alpha(x) = 1 \text{ for } x \leq -\delta,\\ &\text{and } \psi(x) = \alpha(d^{\pm}(x))\}.\end{aligned}$$

Proposition 4.12. *Let $V \in C^{1,1}(\mathbb{R}^+,(0,+\infty))$ be Lipschitz continuous and non-increasing, $\chi \in C^{1,1}(\mathbb{R}^d,\mathbb{R}^+)$ be Lipschitz continuous, and $g : \partial\Omega \to \mathbb{R}^+$ satisfy (H3) and (H6). Suppose also that (H4) holds. Then there exist $C, \delta > 0$ such that, for every $\psi \in \Psi_\delta$, if k is given by (1.1) and η is a MFG equilibrium, defining $\tilde{k}$ by $\tilde{k}(t,x) = k(\rho^\eta(t),x)$, one has*

$$\partial_t\tilde{k}(t,x) \geq -C, \qquad \forall (t,x) \in \mathbb{R}^+ \times \Omega.$$

Proof. Notice first that, for every $\delta > 0$ small enough, $d^{\pm}$ is $C^{1,1}$ in a closed δ-neighborhood of $\partial\Omega$, and thus one has $\psi \in C^{1,1}(\mathbb{R}^d,\mathbb{R}^+)$ for every $\psi \in \Psi_\delta$. Then, by Proposition 4.10, k is $C^{1,1}$ on MFG equilibria.

Let $M > 0$ be such that $\sup_{x,y\in\Omega} \chi(x-y) \leq M$ and $\sup_{x,y\in\Omega}|\nabla\chi(x-y)| \leq M$. Let $\underline{V}' = -\inf_{x\in[0,M]} V'(x) \geq 0$ and $\overline{V} = \sup_{x\in[0,M]} V(x) > 0$. Let $c > 0$ and $\delta > 0$ be as in the statement of Proposition 3.21. Notice that, for every $\delta > 0$, $\psi \in \Psi_\delta$, $x \in \Omega$, and $\mu \in \mathcal{P}(\Omega)$, one has $\int_\Omega \chi(x-y)\psi(y)\,\mathrm{d}\mu(y) \leq M$, and then $k(\mu, x) \leq \overline{V}$.

Let $\psi \in \Psi_\delta$, k be given by (1.1), η be a MFG equilibrium, and $\tilde{k}$ be defined from k as in the statement. Let $\kappa : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^+$ be given by

$$\kappa(t,x) = \int_\Omega \chi(x-y)\psi(y)\,\mathrm{d}\rho_t(y).$$

Notice that $\tilde{k}(t,x) = V(\kappa(t,x))$ and $\kappa(t,x) \in [0,M]$ for every $(t,x) \in \mathbb{R}^+ \times \Omega$. Since V is non-increasing and $V'(\kappa(t,x)) \geq -\underline{V}'$ for every $(t,x) \in \mathbb{R}^+ \times \Omega$, the proposition is proved if one obtains an upper bound on $\partial_t \kappa(t,x)$.

Let $\alpha \in C^{1,1}(\mathbb{R}, [0,1])$ be a non-increasing function with $\alpha(x) = 0$ for $x \geq 0$, $\alpha(x) = 1$ for $x \leq -\delta$, $\alpha'(0) = 0$, and $\psi(x) = \alpha(d^\pm(x))$ for $x \in \mathbb{R}^d$. As in the proof of Proposition 4.10, κ is $C^{1,1}$ and

$$\partial_t \kappa(t,x) = \int_{\Gamma'} \Big[-\nabla\chi(x-\gamma(t))\cdot\gamma'(t)\psi(\gamma(t)) + \chi(x-\gamma(t))\nabla\psi(\gamma(t))\cdot\gamma'(t)\Big]\,\mathrm{d}\eta(\gamma), \tag{4.8}$$

where $\Gamma' = \bigcup_{x\in\Omega} \Gamma'[\rho^\eta, x] \subset \Gamma_{k_{\max}}$. For every $\gamma \in \Gamma'$, one has $|\gamma'(t)| \leq \overline{V}$. On the other hand, denoting by u the optimal control associated with γ, one has

$$\nabla\psi(\gamma(t))\cdot\gamma'(t) = k(t,\gamma(t))\nabla\psi(\gamma(t))\cdot u(t) = k(t,\gamma(t))\alpha'(d^\pm(\gamma(t)))\nabla d^\pm(\gamma(t))\cdot u(t). \tag{4.9}$$

If $d(\gamma(t), \partial\Omega) > \delta$, then $\alpha'(d^\pm(\gamma(t))) = 0$ and thus $\nabla\psi(\gamma(t))\cdot\gamma'(t) = 0$. Otherwise, by Proposition 3.21, one has $\nabla d^\pm(\gamma(t))\cdot u(t) \geq c$, and, since $\alpha'(x) \leq 0$ for every $x \in \mathbb{R}$, one has $\nabla\psi(\gamma(t))\cdot\gamma'(t) \leq 0$. It then follows from (4.8) and (4.9) that

$$\partial_t \kappa(t,x) \leq M\overline{V},$$

providing the required upper bound. □

Our main result of this section is the following.

Theorem 4.13. *Let $p \in (1,+\infty]$, $k : \mathcal{P}(\Omega) \times \Omega \to \mathbb{R}^+$ be continuous, $g : \partial\Omega \to \mathbb{R}^+$, and assume that (H1), (H3), (H4), (H9), and (H11) hold. Suppose that k is $C^{1,1}$ on MFG equilibria. Let $\rho_0 \in \mathcal{P}(\Omega)$, $\eta \in \mathcal{P}_{\rho_0}(\Omega)$ be a MFG equilibrium for ρ_0, $\rho = \rho^\eta$, and φ be the value function of the optimal control problem (4.1) with dynamic $(t,x) \mapsto k(\rho_t, x)$. There exists $C > 0$ such that, if ρ_0 is absolutely continuous and $\rho_0 \in L^p(\mathring{\Omega})$, then, for every $t \geq 0$, $\rho_t|_{\mathring{\Omega}}$ is absolutely continuous, $\rho_t \in L^p(\mathring{\Omega})$, and*

$$\|\rho_t\|_{L^p(\mathring{\Omega})} \leq C\|\rho_0\|_{L^p(\mathring{\Omega})}. \tag{4.10}$$

Moreover, C depends only on λ, $k_{\min}$, $k_{\max}$, $\operatorname{diam}(\Omega)$, a bound κ on the curvature of $\partial\Omega$, L_x, L_{xx}, ℓ_t, the semi-concavity constant of g, and the semi-concavity constant w.r.t. x of the value function φ.

Before proving Theorem 4.13, we need the following auxiliary results.

Lemma 4.14. *Let $O \subset \mathbb{R}^d$ be a bounded open set, $\alpha : O \to \mathbb{R}$ be a semi-concave function with semi-concavity constant $C \geq 0$, and $\beta : \mathbb{R}^d \to \mathbb{R}$ be a C^2 convex function with $\nabla\beta$ Lipschitz continuous and bounded by some constant $C' \geq 0$. Then $\nabla\beta \circ \nabla\alpha$ is a function of locally bounded variation and $\nabla\cdot(\nabla\beta\circ\nabla\alpha) \leq CC'$ in the sense of distributions.*

Proof. Since α is semi-concave with semi-concavity constant C, $\nabla\alpha : O \to \mathbb{R}^d$ is a function of locally bounded variation and $\nabla^2\alpha \leq C$ in the sense of measures (see, e.g., [17, Proposition 1.1.3 and Theorem 2.3.1]). Then, by [6, Theorem 3.96], $\nabla\beta \circ \nabla\alpha$ is a function of locally bounded variation, with

$$\nabla(\nabla\beta \circ \nabla\alpha) = \xi \nabla^2\alpha$$

and $\xi : O \to \mathcal{M}_d(\mathbb{R})$ given by

$$\xi(x) = \int_0^1 \nabla\beta(t\nabla\alpha^+(x) + (1-t)\nabla\alpha^-(x))\,\mathrm{d}t,$$

where $\nabla\alpha^+$ and $\nabla\alpha^-$ have their usual definitions at jump points (see, e.g., [6, Section 3.6]) and are defined at points $x \in O$ where $\nabla\alpha$ is approximately continuous by setting $\nabla\alpha^+(x) = \nabla\alpha^-(x) = \nabla\alpha(x)$. In particular, since β is convex, $\nabla\beta(y)$ is a positive semidefinite matrix for every $y \in \mathbb{R}^d$, and then $\xi(x)$ is also positive semidefinite for every $x \in O$ and bounded by C'. Then $\xi(x)$ admits a positive semidefinite square root $\sqrt{\xi(x)}$, bounded by $\sqrt{C'}$, and one has, in the sense of distributions,

$$\nabla \cdot (\nabla\beta \circ \nabla\alpha) = \mathrm{Tr}(\xi\nabla^2\alpha) = \mathrm{Tr}(\sqrt{\xi}\nabla^2\alpha\sqrt{\xi}) \leq CC',$$

as required. □

Lemma 4.15. *Let $\beta \in C^\infty(\mathbb{R}^d, \mathbb{R}^+)$ be such that $\mathrm{spt}(\beta) \subset B(0,1)$, $\inf_{x \in B(0,1/2)} \beta(x) > 0$, and $\int_{\mathbb{R}^d} \beta(x)\,\mathrm{d}x = 1$. For $\varepsilon > 0$, let $\beta_\varepsilon \in C^\infty(\mathbb{R}^d, \mathbb{R}^+)$ be defined by $\beta_\varepsilon(x) = \varepsilon^{-d}\beta(\frac{x}{\varepsilon})$. Then*

$$\inf_{\substack{x \in \Omega \\ \varepsilon \in (0,1]}} \int_{B(x,\varepsilon)\cap\Omega} \beta_\varepsilon(x-y)\,\mathrm{d}y > 0.$$

Proof. Let $\underline{\beta} = \inf_{x \in B(0,1/2)} \beta(x) > 0$. Notice that, for every $\varepsilon > 0$, if $x \in \mathbb{R}^d$ is such that $|x| < \varepsilon/2$, then $\beta_\varepsilon(x) \geq \varepsilon^{-d}\underline{\beta}$. Hence, for every $x \in \Omega$ and $\varepsilon > 0$, one has

$$\int_{B(x,\varepsilon)\cap\Omega} \beta_\varepsilon(x-y)\,\mathrm{d}y \geq \int_{B(x,\frac{\varepsilon}{2})\cap\Omega} \beta_\varepsilon(x-y)\,\mathrm{d}y \geq \underline{\beta}\varepsilon^{-d}\,\mathrm{Leb}\Big(B\Big(x,\frac{\varepsilon}{2}\Big)\cap\Omega\Big),$$

where Leb denotes the Lebesgue measure in $\mathbb{R}^d$. Thus, it suffices to prove that

$$\inf_{\substack{x \in \Omega \\ \varepsilon \in (0,1]}} \varepsilon^{-d}\,\mathrm{Leb}\Big(B\Big(x,\frac{\varepsilon}{2}\Big)\cap\Omega\Big) > 0. \tag{4.11}$$

Let $(x_n)_{n\in\mathbb{N}}$ and $(\varepsilon_n)_{n\in\mathbb{N}}$ be sequences in Ω and $(0,1]$, respectively, such that

$$\lim_{n\to\infty} \varepsilon_n^{-d}\,\mathrm{Leb}\Big(B\Big(x_n,\frac{\varepsilon_n}{2}\Big)\cap\Omega\Big) = \inf_{\substack{x \in \Omega \\ \varepsilon \in (0,1]}} \varepsilon^{-d}\,\mathrm{Leb}\Big(B\Big(x,\frac{\varepsilon}{2}\Big)\cap\Omega\Big).$$

Up to extracting subsequences, there exist $x_* \in \Omega$ and $\varepsilon_* \in [0,1]$ such that $x_n \to x_*$ and $\varepsilon_n \to \varepsilon_*$ as $n \to \infty$.

Let us consider first the case $x_* \in \mathring{\Omega}$. If $\varepsilon_* > 0$, then

$$\lim_{n\to\infty} \varepsilon_n^{-d}\,\mathrm{Leb}\Big(B\Big(x_n,\frac{\varepsilon_n}{2}\Big)\cap\Omega\Big) = \varepsilon_*^{-d}\,\mathrm{Leb}\Big(B\Big(x_*,\frac{\varepsilon_*}{2}\Big)\cap\Omega\Big),$$

and there exists $\varepsilon' \in (0,\varepsilon_*]$ such that $B(x_*,\varepsilon'/2) \subset \Omega$, in which case

$$\varepsilon_*^{-d}\,\mathrm{Leb}\Big(B\Big(x_*,\frac{\varepsilon_*}{2}\Big)\cap\Omega\Big) \geq \varepsilon_*^{-d}\,\mathrm{Leb}\Big(B\Big(x_*,\frac{\varepsilon'}{2}\Big)\Big) > 0,$$

and (4.11) holds. If $\varepsilon_* = 0$, then, for n large enough, $B(x_n, \varepsilon_n/2) \subset \Omega$, and thus

$$\lim_{n\to\infty} \varepsilon_n^{-d} \operatorname{Leb}\Bigl(B\Bigl(x_n, \frac{\varepsilon_n}{2}\Bigr) \cap \Omega\Bigr) = \lim_{n\to\infty} \varepsilon_n^{-d} \operatorname{Leb}\Bigl(B\Bigl(x_n, \frac{\varepsilon_n}{2}\Bigr)\Bigr) = 2^{-d} v_d > 0,$$

where v_d is the volume of the d-dimensional unit ball. Hence (4.11) holds.

Consider now the case $x_* \in \partial\Omega$. For $x \in \mathbb{R}^d$, $r > 0$, and $\omega \in \mathbb{S}^{d-1}$, let

$$P(x, r, \omega) = \Bigl\{y \in B(x, r) \,:\, (y - x)\cdot\omega > \frac{1}{2}|y - x|\Bigr\}.$$

Notice that $\operatorname{Leb}(P(x, r, \omega)) = r^d \operatorname{Leb}(P(0, 1, \omega'))$ for every $x \in \mathbb{R}^d$, $r > 0$, and $\omega, \omega' \in \mathbb{S}^{d-1}$. Let $v_d' = \operatorname{Leb}(P(0, 1, \omega))$, which is positive and independent of $\omega \in \mathbb{S}^{d-1}$. If $\varepsilon_* = 0$, we claim that, for n large enough,

$$P\Bigl(x_n, \frac{\varepsilon_n}{2}, -\nabla d^{\pm}(x_n)\Bigr) \subset \Omega. \tag{4.12}$$

Indeed, let U be a neighborhood of $\partial\Omega$ such that $d^{\pm}$ is $C^{1,1}$ on U. For n large enough, one has $B(x_n, \varepsilon_n/2) \subset U$, and in particular $\nabla d^{\pm}(y)$ exists (and is a unit vector) for every $y \in B(x_n, \varepsilon_n/2)$. Let L_d be a Lipschitz constant for $\nabla d^{\pm}$. For $y \in P(x_n, \frac{\varepsilon_n}{2}, -\nabla d^{\pm}(x_n))$, one has

$$\begin{aligned} d^{\pm}(y) &\leq d^{\pm}(x_n) + \nabla d^{\pm}(x_n)\cdot(y - x_n) + L_d|y - x_n|^2 \\ &\leq -\frac{1}{2}|y - x_n| + L_d|y - x_n|^2 \\ &\leq -\frac{1}{2}(1 - L_d\varepsilon_n)|y - x_n|. \end{aligned}$$

For n large enough, one has $1 - L_d\varepsilon_n > 0$, and thus $d^{\pm}(y) \leq 0$, i.e., $y \in \Omega$. Hence (4.12) holds. Then, since $P(x_n, \varepsilon_n/2, -\nabla d^{\pm}(x_n)) \subset B(x_n, \varepsilon_n/2)$, one has

$$\lim_{n\to\infty} \varepsilon_n^{-d} \operatorname{Leb}\Bigl(B\Bigl(x_n, \frac{\varepsilon_n}{2}\Bigr) \cap \Omega\Bigr) \geq \limsup_{n\to\infty} \varepsilon_n^{-d} \operatorname{Leb}\Bigl(P\Bigl(x_n, \frac{\varepsilon_n}{2}, -\nabla d^{\pm}(x_n)\Bigr)\Bigr) = 2^{-d} v_d' > 0,$$

yielding (4.11). Finally, if $\varepsilon_* > 0$, one can prove similarly that $P(x_*, \varepsilon'/2, -\nabla d^{\pm}(x_*)) \subset \Omega$ for some $\varepsilon' \in (0, \varepsilon_*]$, yielding that the infimum in (4.11) is lower bounded by $(\varepsilon'/\varepsilon_*)^d v_d'$, and thus (4.11) holds. ☐

Proof of Theorem 4.13. Let $T = \frac{1+\lambda k_{\max}}{1-\lambda k_{\max}} k_{\min}^{-1} \sup_{x\in\Omega} d(x, \partial\Omega)$. It follows from Proposition 3.7 that $\rho_t|_{\mathring{\Omega}} = 0$ for $t \geq T$, and thus it suffices to prove (4.10) for $t \in [0, T]$.

For $t \in [0, T]$, define the vector field $v_t : \mathbb{R}^d \to \mathbb{R}^d$ by

$$v_t(x) = \begin{cases} -k(\rho_t, x)\dfrac{\nabla\varphi(t, x)}{|\nabla\varphi(t, x)|}, & \text{if } x \in \mathring{\Omega}, \\ 0, & \text{otherwise.} \end{cases}$$

Notice that v_t is well-defined almost everywhere since $x \mapsto \varphi(t, x)$ is Lipschitz continuous and, by Proposition 3.27, $\nabla\varphi(t, x) \neq 0$ wherever it exists. Let $c > 0$ be the constant from Corollary 3.11 and let $F : \mathbb{R}^d \to \mathbb{R}$ be a convex C^2 function such that $F(x) = |x|$ for every $x \in \mathbb{R}^d$ with $|x| \geq c$. Notice that F can be chosen in such a way that ∇F is bounded by some constant c' depending only on c. It follows from Proposition 3.27 that, for almost every $x \in \mathring{\Omega}$, one has $v_t(x) = -k(\rho_t, x)\nabla F(\nabla\varphi(t, x))$. Since $x \mapsto k(\rho_t, x)$ is Lipschitz continuous, it follows from [6, Proposition 3.2(b)] that v_t is of locally bounded variation and that its divergence satisfies, in the sense of distributions,

$$\nabla\cdot v_t = -\nabla k\cdot(\nabla F \circ \nabla\varphi) - k\nabla\cdot(\nabla F \circ \nabla\varphi).$$

It then follows from (H1), (H11), and Lemma 4.14 that there exists $C > 0$ depending on c, the constant $k_{\max}$ from (H1), the constant L_x from (H11), and the semi-concavity constant of φ such that

$$\nabla \cdot v_t \geq -C \tag{4.13}$$

in the sense of distributions.

For $\varepsilon > 0$, let $\beta, \beta_\varepsilon \in C^\infty(\mathbb{R}^d, \mathbb{R}^+)$ be defined as in the statement of Lemma 4.15, so that $\operatorname{spt}(\beta_\varepsilon) \subset B(0, \varepsilon)$ and $\int_{\mathbb{R}^d} \beta_\varepsilon(x) \, \mathrm{d}x = 1$. Let $\Omega_\varepsilon = \{x \in \Omega : d(x, \partial\Omega) > \varepsilon\}$. For $t \in [0, T]$, define $v_t^\varepsilon : \mathbb{R}^d \to \mathbb{R}^d$ by $v_t^\varepsilon = v_t * \beta_\varepsilon$. It then follows from (4.13) and [6, Proposition 3.2(c)] that

$$\nabla \cdot v_t^\varepsilon(x) \geq -C \qquad \forall (t, x) \in [0, T] \times \Omega_\varepsilon. \tag{4.14}$$

Notice also that, for every $q_t, q_x \in [1, +\infty)$, one has $v^\varepsilon \to v$ in $L^{q_t}([0, T], L^{q_x}(\mathbb{R}^d))$ as $\varepsilon \to 0$. Let $d^\pm : \mathbb{R}^d \to \mathbb{R}$ be the signed distance to $\partial\Omega$ defined in (3.12).

Claim 1. *There exists $\bar{c} > 0$ and $\bar{\varepsilon} > 0$ such that, for every $\varepsilon \in (0, \bar{\varepsilon}]$, $t \in [0, T]$, and $x \in \Omega$ with $d(x, \partial\Omega) \leq \varepsilon$, one has*

$$\nabla d^\pm(x) \cdot v_t^\varepsilon(x) \geq \bar{c}. \tag{4.15}$$

Proof. Let $c > 0$ and $\delta > 0$ be as in the statement of Proposition 3.21. Up to reducing $\delta > 0$, $d^\pm$ is $C^{1,1}$ on the set of all points at a distance at most $\delta > 0$ from $\partial\Omega$. Let L_d be a Lipschitz constant for $\nabla d^\pm$ on this set and define

$$c' = \inf_{\substack{x \in \Omega \\ \varepsilon \in (0,1]}} \int_{B(x,\varepsilon) \cap \Omega} \beta_\varepsilon(x - y) \, \mathrm{d}y,$$

which is positive by Lemma 4.15. By (3.14), one deduces that, for every $t \in [0, T]$, one has $\nabla d^\pm(x) \cdot v_t(x) \geq c k_{\min}$ for almost every $x \in \Omega$ with $d(x, \partial\Omega) \leq \delta$. Let $\bar{\varepsilon} = \min\left\{\delta/2, 1, \frac{cc' k_{\min}}{2 L_d k_{\max}}\right\}$ and fix $\varepsilon \in (0, \bar{\varepsilon}]$, $t \in [0, T]$, and $x \in \Omega$ with $d(x, \partial\Omega) \leq \varepsilon$. Then

$$\begin{aligned}
\nabla d^\pm(x) \cdot v_t^\varepsilon(x) &= \nabla d^\pm(x) \cdot \int_{B(x,\varepsilon) \cap \Omega} v_t(y) \beta_\varepsilon(x - y) \, \mathrm{d}y \\
&= \int_{B(x,\varepsilon) \cap \Omega} \nabla d^\pm(y) \cdot v_t(y) \beta_\varepsilon(x - y) \, \mathrm{d}y \\
&\quad + \int_{B(x,\varepsilon) \cap \Omega} \left[\nabla d^\pm(x) - \nabla d^\pm(y)\right] \cdot v_t(y) \beta_\varepsilon(x - y) \, \mathrm{d}y \\
&\geq cc' k_{\min} - L_d k_{\max} \varepsilon \geq \frac{1}{2} cc' k_{\min}.
\end{aligned}$$

Hence (4.15) holds with $\bar{c} = \frac{1}{2} cc' k_{\min}$. $\square$

Let $X_\varepsilon : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ satisfy

$$\begin{cases} \partial_t X_\varepsilon(t, x) = v_t^\varepsilon(X_\varepsilon(t, x)), & (t, x) \in [0, T] \times \mathbb{R}^d, \\ X_\varepsilon(0, x) = x, & x \in \mathbb{R}^d, \end{cases} \tag{4.16}$$

i.e., X_ε is the flow of the differential equation $\gamma' = v_t^\varepsilon(\gamma)$ restricted to the fixed initial time 0. By standard properties of flows, for every $t \in [0, T]$, the map $X_\varepsilon(t, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is invertible, and, with a slight abuse of notation, we denote its inverse by $X_\varepsilon^{-1}(t,\cdot)$. Since $v_t^\varepsilon \in C^\infty(\mathbb{R}^d,\mathbb{R}^d)$, one has $X_\varepsilon(t,\cdot) \in C^\infty(\mathbb{R}^d,\mathbb{R}^d)$ and, in particular,

$$\begin{cases} \partial_t \nabla X_\varepsilon(t,x) = \nabla v_t^\varepsilon(X_\varepsilon(t,x))\nabla X_\varepsilon(t,x), & (t,x) \in [0,T]\times\mathbb{R}^d, \\ \nabla X_\varepsilon(0,x) = I, & x \in \mathbb{R}^d. \end{cases}$$

Let $J_\varepsilon : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ be given by $J_\varepsilon(t,x) = \det(\nabla X_\varepsilon(t,x))$. Then J_ε satisfies

$$\begin{cases} \partial_t J_\varepsilon(t,x) = \nabla\cdot v_t^\varepsilon(X_\varepsilon(t,x))J_\varepsilon(t,x), & (t,x) \in [0,T]\times\mathbb{R}^d, \\ J_\varepsilon(0,x) = 1, & x \in \mathbb{R}^d, \end{cases}$$

which yields

$$J_\varepsilon(t,x) = \exp\biggl(\int_0^t \nabla\cdot v_s^\varepsilon(X_\varepsilon(s,x))\,\mathrm{d}s\biggr).$$

Let $\rho_t^\varepsilon = X_\varepsilon(t,\cdot)_\# \rho_0$. Then ρ^ε satisfies, in the sense of distributions in $[0,T)\times\mathbb{R}^d$,

$$\begin{cases} \partial_t \rho^\varepsilon + \nabla\cdot(\rho^\varepsilon v^\varepsilon) = 0, \\ \rho_0^\varepsilon = \rho_0. \end{cases} \tag{4.17}$$

Moreover, since ρ_0 is absolutely continuous with respect to the Lebesgue measure, so is ρ_t^ε, and their densities (also denoted by ρ_0 and ρ_t^ε for simplicity) satisfy

$$\rho_t^\varepsilon(x) = \frac{\rho_0(X_\varepsilon^{-1}(t,x))}{J_\varepsilon(t,X_\varepsilon^{-1}(t,x))}.$$

Let $K \subset \mathring{\Omega}$ be compact. For $t \in [0,T]$ and $\varepsilon > 0$, set

$$K_t^\varepsilon = \Bigl\{X_\varepsilon(s,x) \,:\, s \in [0,t],\ x \in \mathbb{R}^d \text{ and } X_\varepsilon(t,x) \in K\Bigr\},$$

i.e., K_t^ε is the set of all points which belong to some trajectory of (4.16) passing through K at time t.

Claim 2. *There exists $\varepsilon_0 > 0$ such that, for every $t \in [0,T]$ and $\varepsilon \in (0,\varepsilon_0)$, one has $K_t^\varepsilon \subset \Omega_\varepsilon$.*

Proof. Assume, to obtain a contradiction, that there exist sequences $(t_n)_{n\in\mathbb{N}}$ and $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\varepsilon_n \to 0$ as $n \to \infty$ such that $t_n \in [0,T]$ and $K_{t_n}^{\varepsilon_n} \not\subset \Omega_{\varepsilon_n}$ for every $n \in \mathbb{N}$. Then, for every $n \in \mathbb{N}$, there exists $s_n \in [0,t_n]$ and $x_n \in \mathbb{R}^d$ such that, setting $z_n = X_{\varepsilon_n}(s_n,x_n)$, one has $z_n \notin \Omega_{\varepsilon_n}$. Notice that $d(x_n,\Omega) \leq \varepsilon_n$, for otherwise one would have $v_s^{\varepsilon_n}(x_n) = 0$ for every $s \geq 0$ and then $X_{\varepsilon_n}(s,x_n) = x_n$ for every $s \geq 0$, contradicting the fact that $X_{\varepsilon_n}(t_n,x_n) \in K \subset \mathring{\Omega}$. For the same reason, one must have $d(z_n,\Omega) \leq \varepsilon_n$, and, since $z_n \notin \Omega_{\varepsilon_n}$, this implies that $d(z_n,\partial\Omega) \leq \varepsilon_n$.

Let $\bar\varepsilon > 0$ and $\bar c > 0$ be such that (4.15) holds for every $\varepsilon \in (0,\bar\varepsilon]$, $t \in [0,T]$, and $x \in \Omega$ with $d(x,\partial\Omega) \leq \varepsilon$. Up to reducing $\bar\varepsilon$, one may assume that $d(K,\partial\Omega) > \bar\varepsilon$.

Fix $n \in \mathbb{N}$ such that $\varepsilon_n \leq \bar\varepsilon$. Let $\alpha : [0,T] \to \mathbb{R}$ be defined for $s \in [0,T]$ by $\alpha(s) = d^\pm(X_{\varepsilon_n}(s,x_n))$. Then $\alpha'(s) = \nabla d^\pm(X_{\varepsilon_n}(s,x_n)) \cdot v_s^{\varepsilon_n}(X_{\varepsilon_n}(s,x_n))$. In particular, by (4.15), $\alpha'(s) \geq \bar c > 0$ whenever $\alpha(s) \in [-\varepsilon_n, 0]$ (i.e., whenever $X_{\varepsilon_n}(s,x_n) \in \Omega$ and $d(X_{\varepsilon_n}(s,x_n),\partial\Omega) \leq \varepsilon_n$). Since $d(z_n,\partial\Omega) \leq \varepsilon_n$, one has $\alpha(s_n) = d^\pm(z_n) \in [-\varepsilon_n,\varepsilon_n]$, and thus $\alpha(s) \geq -\varepsilon_n$ for every $s \in [s_n,T]$. This is a contradiction, since $\alpha(t_n) = d^\pm(X_{\varepsilon_n}(t_n,x_n)) < -\varepsilon_n$ due to the fact that $X_{\varepsilon_n}(t_n,x_n) \in K \subset \mathring{\Omega}$ and $d(K,\partial\Omega) > \bar\varepsilon \geq \varepsilon_n$. ☐

Let $\varepsilon_0 > 0$ be as in the statement of Claim 2. We consider here only the case $p \in (1, \infty)$, the remaining case $p = \infty$ following from the fact that our constants do not depend on p. For $t \in [0, T]$ and $\varepsilon \in (0, \varepsilon_0)$, one has

$$\begin{aligned}\|\rho_t^\varepsilon\|_{L^p(K)}^p &= \int_K \frac{\rho_0(X_\varepsilon^{-1}(t,x))^p}{J_\varepsilon(t, X_\varepsilon^{-1}(t,x))^p} \,\mathrm{d}x = \int_{K_0} \frac{\rho_0(x)^p}{J_\varepsilon(t,x)^{p-1}} \,\mathrm{d}x \\ &= \int_{K_0} \rho_0(x)^p \left[\exp\left(\int_0^t \nabla \cdot v_s^\varepsilon(X_\varepsilon(s,x)) \,\mathrm{d}s\right)\right]^{1-p} \mathrm{d}x,\end{aligned}$$

where $K_0 = \{x \in \mathbb{R}^d \,:\, X_\varepsilon(t,x) \in K\}$. For every $x \in K_0$ and $s \in [0,t]$, one has $X_\varepsilon(s,x) \in K_t^\varepsilon \subset \Omega_\varepsilon$, and thus one obtains from the above expression and (4.14) that

$$\|\rho_t^\varepsilon\|_{L^p(K)}^p \leq e^{C(p-1)T} \|\rho_0\|_{L^p(\mathring{\Omega})}^p. \tag{4.18}$$

Let $(K_n)_{n\in\mathbb{N}}$ be an increasing sequence of compact subsets of $\mathring{\Omega}$ such that $\mathring{\Omega} = \bigcup_{n\in\mathbb{N}} K_n$. For $i \in \mathbb{N}$, we construct by induction on i a sequence $(\varepsilon_n^i)_{n\in\mathbb{N}}$ such that $\varepsilon_n^i \to 0$ as $n \to \infty$. Let $K = K_0$ and take $\varepsilon_0 > 0$ as in the statement of Claim 2. Since, by (4.18), $(\rho^\varepsilon)_{\varepsilon\in(0,\varepsilon_0]}$ is bounded in $L^\infty([0,T], L^p(K_0))$, there exists a sequence $(\varepsilon_n^0)_{n\in\mathbb{N}}$ in $(0, \varepsilon_0]$ with $\varepsilon_n^0 \to 0$ as $n \to \infty$ such that $(\rho^{\varepsilon_n^0})_{n\in\mathbb{N}}$ converges weakly-$*$ in $L^\infty([0,T], L^p(K_0))$. Now, assume that $i \in \mathbb{N}$ is such that $(\varepsilon_n^i)_{n\in\mathbb{N}}$ is constructed and $\varepsilon_n^i \to 0$ as $n \to \infty$. Since, by (4.18), $(\rho^{\varepsilon_n^i})_{n\in\mathbb{N}}$ is bounded in $L^\infty([0,T], L^p(K_{i+1}))$, there exists a subsequence $(\varepsilon_n^{i+1})_{n\in\mathbb{N}}$ of $(\varepsilon_n^i)_{n\in\mathbb{N}}$ such that $(\rho^{\varepsilon_n^{i+1}})_{n\in\mathbb{N}}$ converges weakly-$*$ in $L^\infty([0,T], L^p(K_{i+1}))$.

For $n \in \mathbb{N}$, let $\varepsilon_n = \varepsilon_n^n$. Then $(\rho^{\varepsilon_n})_{n\in\mathbb{N}}$ converges weakly-$*$ in $L^\infty([0,T], L^p(K_i))$ for every $i \in \mathbb{N}$. Let $\bar\rho \in L^\infty([0,T], L^p_{\mathrm{loc}}(\mathring{\Omega}))$ denote the weak-$*$ limit of $(\rho^{\varepsilon_n})_{n\in\mathbb{N}}$. One deduces from the weak convergence of $(\rho^{\varepsilon_n})_{n\in\mathbb{N}}$ and (4.18) that, for every $i \in \mathbb{N}$ and almost every $t \in [0,T]$,

$$\|\bar\rho_t\|_{L^p(K_i)}^p \leq e^{C(p-1)T} \|\rho_0\|_{L^p(\mathring{\Omega})}^p,$$

and thus

$$\|\bar\rho_t\|_{L^p(\mathring{\Omega})} = \lim_{i\to\infty} \|\bar\rho_t\|_{L^p(K_i)} \leq e^{C\left(1-\frac{1}{p}\right)T} \|\rho_0\|_{L^p(\mathring{\Omega})}$$

for almost every $t \in [0,T]$. In particular, one obtains that $\bar\rho \in L^\infty([0,T], L^p(\mathring{\Omega}))$.

Since $v^\varepsilon \to v$ in $L^1([0,T], L^{p'}(\mathbb{R}^d))$ as $\varepsilon \to 0$, one obtains from (4.17) that $\bar\rho$ satisfies, in the sense of distributions in $[0,T) \times \mathring{\Omega}$,

$$\begin{cases} \partial_t \bar\rho + \nabla \cdot (\bar\rho v) = 0, \\ \bar\rho_0 = \rho_0. \end{cases}$$

On the other hand, the measure $\rho = \rho^\eta$ obtained from the MFG equilibrium η also satisfies the continuity equation $\partial_t \rho + \nabla \cdot (\rho v) = 0$ with initial condition ρ_0. It follows from Proposition 3.34 and [5, Theorem 3.1] that solutions to this equation are unique, and thus $\bar\rho = \rho$. In particular,

$$\|\rho_t\|_{L^p(\mathring{\Omega})}^p \leq e^{C(p-1)T} \|\rho_0\|_{L^p(\mathring{\Omega})}^p \tag{4.19}$$

for almost every $t \in [0,T]$. To conclude that (4.19) holds for every $t \in [0,T]$, let $\bar t \in [0,T]$ and $(t_n)_{n\in\mathbb{N}}$ be a sequence in $[0,T]$ such that $t_n \to t$ as $n \to \infty$ and (4.19) holds at t_n for every $n \in \mathbb{N}$. The sequence $(\rho_{t_n})_{n\in\mathbb{N}}$ is bounded in $L^p(\mathring{\Omega})$, and thus, up to the extraction of a subsequence, it admits a weak limit $\widetilde\rho$. On the other hand, $t \mapsto \rho_t = (e_t)_\# \eta$ is continuous with respect to the weak convergence of measures, and thus $\widetilde\rho = \rho_{\bar t}$. One concludes that (4.19) holds for $\bar{t}$ by the L^p-weak convergence of $(\rho_{t_n})_{n\in\mathbb{N}}$ to $\rho_{\bar{t}}$ and the weak lower semi-continuity of the L^p norm. $\square$

4.3. **Equilibria in a less regular model.** In this section, we use the L^p estimates on ρ_t from Theorem 4.13 to study equilibria of the MFG model with k given by (1.1) and $\psi = \mathbb{1}_{\mathring{\Omega}}$, i.e.,

$$k(\mu, x) = V\left(\int_\Omega \chi(x-y)\mathbb{1}_{\mathring{\Omega}}(y)\,\mathrm{d}\mu(y)\right), \quad \text{for all } (\mu, x) \in \mathcal{P}(\Omega)\times\Omega. \tag{4.20}$$

Notice that the lack of continuity of the dynamic k with respect to μ prevents us from using the result of Section 4.1. So, the idea is to consider a sequence of *cut-off* functions $(\psi^\varepsilon)_{\varepsilon>0}$ taken in Ψ_δ, for δ as in Proposition 4.12, and converging as $\varepsilon \to 0$ to $\mathbb{1}_{\mathring{\Omega}}$ in $L^q(\mathbb{R}^d)$ for all $q \in [1, +\infty)$, and to replace the dynamic k with k_ε defined from ψ^ε as in (1.1), i.e.,

$$k_\varepsilon(\mu, x) = V\left(\int_\Omega \chi(x-y)\psi^\varepsilon(y)\,\mathrm{d}\mu(y)\right), \quad \text{for all } (\mu, x) \in \mathcal{P}(\Omega)\times\Omega. \tag{4.21}$$

Our first result of this section shows that, under some suitable convergence assumptions on k_ε as $\varepsilon \to 0$, one has uniform convergence of the value functions of the corresponding optimal control problems and that the limit of MFG equilibria is a MFG equilibrium for the limiting model.

Proposition 4.16. *Let $\rho_0 \in \mathcal{P}(\Omega)$. For $n \in \mathbb{N}$, let $k_n, k : \mathcal{P}(\Omega)\times\Omega \to \mathbb{R}^+$ be such that k_n is continuous on $\mathcal{P}(\Omega)\times\Omega$ and Lipschitz continuous with respect to the second variable. Let η_n be a MFG equilibrium for ρ_0 associated with the control problem with dynamic k_n. In addition, assume the following:*

- *As $n \to \infty$, $(\eta_n)_{n\in\mathbb{N}}$ converges weakly in $\mathcal{P}(\Gamma)$ to some measure η.*
- *There exist two constants $k_{\min}$ and $k_{\max}$ such that $0 < k_{\min} \le k_n \le k_{\max} < +\infty$.*
- *There exists a constant M independent of n such that $|\nabla k_n| \le M$.*
- *For every $(t, x) \in \mathbb{R}^+\times\Omega$, we have $k_n((e_t)_\#\eta_n, x) \to k((e_t)_\#\eta, x)$ as $n \to \infty$.*
- *For every $x \in \Omega$, $t \mapsto k((e_t)_\#\eta, x)$ is continuous on $\mathbb{R}^+$.*

For $n \in \mathbb{N}$, let φ_n (resp. φ) be the value function associated with the control problem with dynamic k_n (resp. k). Then

(a) *$\varphi_n \to \varphi$ as $n \to \infty$ uniformly in $\mathbb{R}^+\times\Omega$, and*
(b) *η is a MFG equilibrium for ρ_0 with dynamic k.*

Proof. Let us first prove (a). Notice that, up to extracting a subsequence, $(\varphi_n)_{n\in\mathbb{N}}$ converges uniformly to some function $\widetilde{\varphi}$ on $\mathbb{R}^+\times\Omega$. Indeed, from Proposition 3.7, the sequence $(\varphi_n)_n$ is equibounded. Moreover, by Proposition 3.8, the value function φ_n is Lipschitz in $\mathbb{R}^+\times\Omega$ with a Lipschitz constant depending only on the Lipschitz constant of the dynamic k_n with respect to x, which is independent of n. Then, by Arzelà–Ascoli Theorem, $(\varphi_n)_{n\in\mathbb{N}}$ admits a uniform limit $\widetilde{\varphi}$ up to the extraction of a subsequence.

We now prove that the limit of $(\varphi_n)_{n\in\mathbb{N}}$ is φ. Fix $(t, x) \in \mathbb{R}^+\times\Omega$. For every $n \in \mathbb{N}$, let γ_n be an optimal trajectory for x, at time t, in the control problem with dynamic k_n. It is easy to observe that, up to extracting a subsequence, $\gamma_n \to \gamma$ uniformly for some $\gamma \in \Gamma_{k_{\max}}$. Yet, this γ is, in fact, an admissible trajectory for x, at time t, in the control problem with dynamic k. Indeed, for a.e. $s \in (t, \infty)$, we have $|\gamma_n'(s)| \le k_n((e_s)_\#\eta_n, \gamma_n(s))$. So, letting $n \to \infty$, we get

$$|\gamma'(s)| \le k((e_s)_\#\eta, \gamma(s)), \quad \text{for a.e. } s \in (t, \infty).$$

Let u_n be the optimal control associated with γ_n and u the control associated with γ. Set

$$\tau_n := \tau^{t,x,u_n}, \qquad \tau_\gamma := \tau^{t,x,u}, \qquad z_n := \gamma_n(t+\tau_n) \in \partial\Omega.$$

It is clear that there exist $\bar\tau \geq 0$ and $z \in \partial\Omega$ such that, up to extracting subsequences, $\tau_n \to \bar\tau$ and $z_n \to z$ as $n \to \infty$. In particular, we have $z = \gamma(t+\bar\tau)$ and then $\tau_\gamma \leq \bar\tau$. Consequently, by Lemma 3.1, $\varphi(t,x) \leq \bar\tau + g(z) = \lim_{n\to\infty} \varphi_n(t,x) = \widetilde\varphi(t,x)$.

To prove the converse inequality, let γ be an optimal trajectory for x, at time t, in the control problem with dynamic k, and u be the associated optimal control with γ. For $n \in \mathbb{N}$, let ϕ_n be a solution of

$$\begin{cases} \phi_n'(s) = \dfrac{k_n((e_s)_\# \eta_n, \gamma(\phi_n(s)))}{k((e_{\phi_n(s)})_\# \eta, \gamma(\phi_n(s)))}, \\ \phi_n(t) = t. \end{cases} \tag{4.22}$$

Set

$$\gamma_n(s) = \gamma(\phi_n(s)), \quad \text{for all } s \in [t,\infty).$$

It is clear that γ_n is admissible for x, at time t, in the control problem with dynamic k_n, and its corresponding control u_n is given by $u_n(s) = u(\phi_n(s))$ for $s \geq t$. Let $\tau_n = \tau^{t,x,u_n}$. Hence, we have

$$\varphi_n(t,x) \leq \tau_n + g(\gamma_n(t+\tau_n)). \tag{4.23}$$

Yet, we observe easily that $\tau_n = \phi_n^{-1}(t+\tau) - t$, where $\tau := \tau^{t,x,u}$. From (4.22), we have

$$\int_t^{\phi_n(s)} k((e_r)_\#\eta, \gamma(r))\,\mathrm{d}r = \int_t^s k_n((e_r)_\#\eta_n, \gamma(\phi_n(r)))\,\mathrm{d}r.$$

Set

$$\Psi(\theta) = \int_t^\theta k((e_r)_\#\eta, \gamma(r))\,\mathrm{d}r, \quad \text{for all } \theta \in [t,\infty).$$

Then Ψ is a bijective map from $[t,+\infty)$ to $[0,+\infty)$, whose inverse is $k_{\min}^{-1}$-Lipschitz continuous. We have

$$\begin{aligned} |\phi_n(s) - s| &= \left|\Psi^{-1}\left(\int_t^s k_n((e_r)_\#\eta_n, \gamma(\phi_n(r)))\,\mathrm{d}r\right) - \Psi^{-1}\left(\int_t^s k((e_r)_\#\eta, \gamma(r))\,\mathrm{d}r\right)\right| \\ &\leq k_{\min}^{-1} \int_t^s |k_n((e_r)_\#\eta_n, \gamma(\phi_n(r))) - k((e_r)_\#\eta, \gamma(r))|\,\mathrm{d}r \\ &\leq k_{\min}^{-1} \int_t^s \Big(|k_n((e_r)_\#\eta_n, \gamma(\phi_n(r))) - k((e_r)_\#\eta, \gamma(\phi_n(r)))| \\ &\qquad\qquad + |k((e_r)_\#\eta, \gamma(\phi_n(r))) - k((e_r)_\#\eta, \gamma(r))|\Big)\,\mathrm{d}r \\ &\leq k_{\min}^{-1}\Big(\int_t^s |k_n((e_r)_\#\eta_n, \gamma(\phi_n(r))) - k((e_r)_\#\eta, \gamma(\phi_n(r)))|\,\mathrm{d}r \\ &\qquad + M k_{\max} \int_t^s |\phi_n(r) - r|\,\mathrm{d}r\Big). \end{aligned}$$

Yet,

$$\begin{aligned}
&|k_n((e_r)_\#\eta_n, \gamma(\phi_n(r))) - k((e_r)_\#\eta, \gamma(\phi_n(r)))| \\
\leq &\left|\Big(k_n((e_r)_\#\eta_n, \gamma(\phi_n(r))) - k_n((e_r)_\#\eta_n, \gamma(r))\Big) - \Big(k((e_r)_\#\eta, \gamma(\phi_n(r))) - k((e_r)_\#\eta, \gamma(r))\Big)\right| \\
&+ \left|k_n((e_r)_\#\eta_n, \gamma(r)) - k((e_r)_\#\eta, \gamma(r))\right| \\
\leq &\, 2M k_{\max}|\phi_n(r) - r| + \left|k_n((e_r)_\#\eta_n, \gamma(r)) - k((e_r)_\#\eta, \gamma(r))\right|.
\end{aligned}$$

Hence, one has

$$|\phi_n(s) - s| \leq C\left(\int_t^s \left|k_n((e_r)_\#\eta_n, \gamma(r)) - k((e_r)_\#\eta, \gamma(r))\right| \mathrm{d}r + \int_t^s |\phi_n(r) - r| \,\mathrm{d}r\right),$$

where $C > 0$ depends only on M, $k_{\max}$, and $k_{\min}$. Using Gronwall's inequality, we get

$$|\phi_n(s) - s| \leq Ce^{C(s-t)} \int_t^s \left|k_n((e_r)_\#\eta_n, \gamma(r)) - k((e_r)_\#\eta, \gamma(r))\right| \mathrm{d}r.$$

Consequently, for every $s \geq t$, $\phi_n(s) \to s$ as $n \to \infty$. In particular, we have $\tau_n = \phi_n^{-1}(t+\tau) - t \to \tau$. So, passing to the limit in (4.23), we get

$$\widetilde{\varphi}(t, x) \leq \tau + g(\gamma(t + \tau)) = \varphi(t, x).$$

This concludes the proof of (a).

To prove (b), we define, for $k \in \mathbb{N}^*$, the set $V_k := \{\gamma \in \Gamma \,:\, d(\gamma, \bigcup_x \Gamma'[\rho^\eta, x]) \leq \frac{1}{k}\}$. We claim that, for every $k \in \mathbb{N}^*$, there is some $N_0 \in \mathbb{N}$ such that

$$\bigcup_{x\in\Omega} \Gamma'[\rho^{\eta_n}, x] \subset V_k \qquad \text{for every } n \geq N_0. \tag{4.24}$$

Indeed, if this is not the case, then there exists $k \in \mathbb{N}^*$ and sequences $(n_j)_{j\in\mathbb{N}}$, $(x_j)_{j\in\mathbb{N}}$, and $(\gamma_j)_{j\in\mathbb{N}}$ with $n_j \to \infty$ as $j \to \infty$ and, for every $j \in \mathbb{N}$, $x_j \in \Omega$ and $\gamma_j \in \Gamma'[\rho^{\eta_{n_j}}, x_j] \setminus V_k$. Up to extracting subsequences, there exist $x \in \Omega$ and $\gamma \in \Gamma_{k_{\max}}$ such that, as $j \to \infty$, $x_j \to x$ and $\gamma_j \to \gamma$ on Γ. For $j \in \mathbb{N}$, set $\tau_j = \tau^{0,x_j,u_j}$, where u_j is the control corresponding to γ_j. Then, using Proposition 3.7, we infer that, up to extracting subsequences, there exists $\bar\tau \geq 0$ such that $\tau_j \to \bar\tau$ and $\gamma_j(\tau_j) \to \gamma(\bar\tau) \in \partial\Omega$, which implies that $\tau_\gamma \leq \bar\tau$. Moreover, it is easy to check that γ is admissible in the control problem with dynamic k. Yet, we have

$$\varphi_{n_j}(0, x_j) = \tau_j + g(\gamma_j(\tau_j)).$$

Then, passing to the limit when $j \to \infty$, we obtain from (a) and Lemma 3.1 that

$$\varphi(0, x) = \bar\tau + g(\gamma(\bar\tau)) \geq \tau_\gamma + g(\gamma(\tau_\gamma)).$$

This implies that $\gamma \in \Gamma'[\rho^\eta, x]$, which is a contradiction.

As a consequence of (4.24) and the fact that V_k is a closed subset of Γ, we have, for every $k \in \mathbb{N}^*$,

$$\eta(V_k) \geq \limsup_{n\to\infty} \eta_n(V_k) \geq \limsup_{n\to\infty} \eta_n\left(\bigcup_x \Gamma'[\rho^{\eta_n}, x]\right) = 1.$$

Hence, $\eta(V_k) = 1$ and, since k is arbitrary and $\bigcap_{k\in\mathbb{N}^*} V_k = \bigcup_x \Gamma'[\rho^\eta, x]$, we infer that $\eta\Big(\bigcup_x \Gamma'[\rho^\eta, x]\Big) = 1$, concluding the proof of (b). $\square$

One can now use Proposition 4.16 to obtain the existence of an equilibrium to the less regular dynamic k defined in (4.20).

Theorem 4.17. *Let k be given by (4.20) and $\rho_0 \in L^p(\mathring{\Omega})$ for some $p \in (1, +\infty]$. Then there exists a MFG equilibrium η for ρ_0. Moreover, letting $\rho = \rho^\eta$ and φ be the value function of the optimal control problem (4.1), then (ρ, φ) solves the MFG system (4.7).*

Proof. For $\varepsilon > 0$, let k_ε be given by (4.21) and η_ε be a MFG equilibrium for ρ_0 associated with the control problem with dynamic k_ε. Then there exists $\eta \in \mathcal{P}(\Gamma)$ and a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\varepsilon_n \to 0$ as $n \to \infty$ such that $\eta_{\varepsilon_n} \rightharpoonup \eta$ as $n \to \infty$. To prove that η is a MFG equilibrium for ρ_0, it suffices to show that the hypotheses of Proposition 4.16 are verified for the sequences $(k_{\varepsilon_n})_{n\in\mathbb{N}}$ and $(\eta_{\varepsilon_n})_{n\in\mathbb{N}}$. For $t \geq 0$, let $\rho_t^\varepsilon = (e_t)_{\#}\eta_\varepsilon$.

One easily obtains from (4.21) and Proposition 4.12 that there exist $k_{\min}, k_{\max}, M, C > 0$ such that, for every $\varepsilon > 0$, $0 < k_{\min} \leq k_\varepsilon \leq k_{\max} < +\infty$, $\lvert\nabla k_\varepsilon\rvert \leq M$, and $\partial_t k_\varepsilon \geq -C$. As a consequence of that, by Theorem 3.23, the value function φ_ε, associated with the control problem with the dynamic k_ε, is semi-concave with respect to x, and its semi-concavity constant is independent of ε. Then, Theorem 4.13 implies that

$$\lVert\rho_t^\varepsilon\rVert_{L^p(\mathring{\Omega})} \leq C\lVert\rho_0\rVert_{L^p(\mathring{\Omega})}, \quad \text{for all } t \in \mathbb{R}^+,\ \varepsilon > 0,$$

where the constant $C > 0$ is independent of t and ε.

Since $\eta_{\varepsilon_n} \rightharpoonup \eta$ in $\mathcal{P}(\Gamma)$ as $n \to \infty$, one deduces from the above uniform L^p estimate that $\rho_t^{\varepsilon_n} \rightharpoonup \rho_t$ in L^p. In addition, $\psi^\varepsilon \to \mathbb{1}_{\mathring{\Omega}}$ in L^q as $\varepsilon \to 0$, for all $q \in [1, +\infty)$. Using these facts, we get, for every $(t, x) \in \mathbb{R}^+ \times \Omega$, that, as $n \to \infty$,

$$k_{\varepsilon_n}((e_t)_{\#}\eta_{\varepsilon_n}, x) = V\left(\int_\Omega \chi(x-y)\psi^{\varepsilon_n}(y)\rho_t^{\varepsilon_n}(y)\,\mathrm{d}y\right) \to V\left(\int_{\mathring{\Omega}} \chi(x-y)\rho_t(y)\,\mathrm{d}y\right) = k((e_t)_{\#}\eta, x).$$

Moreover, for any $x \in \Omega$, the function $t \mapsto k((e_t)_{\#}\eta, x)$ is continuous on $\mathbb{R}^+$. Indeed, if $(t_n)_{n\in\mathbb{N}}$ is a sequence with $t_n \to t$, then $\rho_{t_n} \rightharpoonup \rho_t$ in L^p and so we have

$$k((e_{t_n})_{\#}\eta, x) = V\left(\int_{\mathring{\Omega}} \chi(x-y)\rho_{t_n}(y)\,\mathrm{d}y\right) \to V\left(\int_{\mathring{\Omega}} \chi(x-y)\rho_t(y)\,\mathrm{d}y\right) = k((e_t)_{\#}\eta, x).$$

Hence the hypotheses of Proposition 4.16 are satisfied, and then η is a MFG equilibrium for ρ_0.

To obtain the MFG system (4.7) for this equilibrium, notice first that $(t, x) \mapsto k(\rho_t, x)$ is continuous and satisfies (H1) and (H2), and thus it follows from Proposition 3.5 that φ satisfies the Hamilton–Jacobi equation in (4.7) in the viscosity sense.

By Theorem 4.11, for every $\varepsilon > 0$, ρ^ε satisfies the continuity equation in (4.7) with dynamic k_ε and the corresponding value function φ_ε. This means that, for every $\phi \in C_c^\infty((0, +\infty) \times \mathring{\Omega})$, one has

$$-\int_0^\infty \int_{\mathring{\Omega}} \partial_t\phi(t,x)\rho_t^\varepsilon(x)\,\mathrm{d}x\,\mathrm{d}t + \int_0^\infty \int_{\mathring{\Omega}} \rho_t^\varepsilon(x) k_\varepsilon(\rho_t^\varepsilon, x)\frac{\nabla\varphi_\varepsilon(t,x)}{\lvert\nabla\varphi_\varepsilon(t,x)\rvert}\cdot\nabla\phi(t,x)\,\mathrm{d}x\,\mathrm{d}t = 0.$$

Recall that, by Proposition 3.7, one has $\rho_t^\varepsilon|_{\mathring{\Omega}} = 0$ for every $\varepsilon > 0$ and $t \geq T$, where $T = \frac{1+\lambda k_{\max}}{1-\lambda k_{\max}} k_{\min}^{-1} \sup_{x\in\Omega} d(x, \partial\Omega)$. Hence,

$$-\int_0^T \int_{\mathring{\Omega}} \partial_t\phi(t,x)\rho_t^\varepsilon(x)\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_{\mathring{\Omega}} \rho_t^\varepsilon(x) k_\varepsilon(\rho_t^\varepsilon, x)\frac{\nabla\varphi_\varepsilon(t,x)}{\lvert\nabla\varphi_\varepsilon(t,x)\rvert}\cdot\nabla\phi(t,x)\,\mathrm{d}x\,\mathrm{d}t = 0 \tag{4.25}$$

for every $\phi \in C_c^\infty((0,T)\times \mathring{\Omega})$. Recall that, for every $t \in \mathbb{R}^+$, one has $\rho_t^{\varepsilon_n} \rightharpoonup \rho_t$ in L^p, and thus

$$\lim_{n\to\infty} \int_{\mathring{\Omega}} \partial_t \phi(t,x)\rho_t^{\varepsilon_n}(x)\,\mathrm{d}x = \int_{\mathring{\Omega}} \partial_t\phi(t,x)\rho_t(x)\,\mathrm{d}x. \tag{4.26}$$

Moreover, $k_{\varepsilon_n}(\rho_t^{\varepsilon_n}, x) \to k(\rho_t, x)$ for every $(t,x) \in \mathbb{R}^+ \times \Omega$. On the other hand, for every $t \in \mathbb{R}^+$ and $\varepsilon > 0$, $x \mapsto \varphi_\varepsilon(t,x)$ is semi-concave and its semi-concavity constant C is independent of ε. Then, by Proposition 4.16(a), $x \mapsto \varphi(t,x)$ is also semi-concave with the same semi-concavity constant. For every $t \in \mathbb{R}^+$, $n \in \mathbb{N}$, and almost every $x \in \mathring{\Omega}$, $\nabla\varphi_{\varepsilon_n}(t,x)$ and $\nabla\varphi(t,x)$ exist. Then, for every $h > 0$ small enough, one has

$$\varphi_{\varepsilon_n}(t,x+h) - \varphi_{\varepsilon_n}(t,x) - \nabla\varphi_{\varepsilon_n}(t,x)\cdot h \le C|h|^2.$$

Letting $n \to \infty$, up to extracting a subsequence, $\nabla\varphi_{\varepsilon_n}(t,x)$ converges to some $p \in \mathbb{R}^d$, and then

$$\varphi(t,x+h) - \varphi(t,x) - p\cdot h \le C|h|^2.$$

This means that $p \in \nabla^+\varphi(t,x)$ and, since $\nabla\varphi(t,x)$ exists, one concludes that $\nabla\varphi_{\varepsilon_n}(t,x) \to \nabla\varphi(t,x)$ as $n \to \infty$. Moreover, by Proposition 3.27, there exists $c > 0$ such that $|\nabla\varphi_{\varepsilon_n}(t,x)| \ge c$, implying that $\frac{\nabla\varphi_{\varepsilon_n}(t,x)}{|\nabla\varphi_{\varepsilon_n}(t,x)|} \to \frac{\nabla\varphi(t,x)}{|\nabla\varphi(t,x)|}$ as $n \to \infty$. One then concludes that, for every $t \ge 0$,

$$\lim_{n\to\infty}\int_{\mathring{\Omega}} \rho_t^{\varepsilon_n}(x) k_{\varepsilon_n}(\rho_t^{\varepsilon_n}, x)\frac{\nabla\varphi_{\varepsilon_n}(t,x)}{|\nabla\varphi_{\varepsilon_n}(t,x)|}\cdot\nabla\phi(t,x)\,\mathrm{d}x = \int_{\mathring{\Omega}} \rho_t(x)k(\rho_t,x)\frac{\nabla\varphi(t,x)}{|\nabla\varphi(t,x)|}\cdot\nabla\phi(t,x)\,\mathrm{d}x. \tag{4.27}$$

Combining (4.26) and (4.27), one obtains from (4.25) that

$$-\int_0^T\int_{\mathring{\Omega}} \partial_t\phi(t,x)\rho_t(x)\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_{\mathring{\Omega}}\rho_t(x)k(\rho_t,x)\frac{\nabla\varphi(t,x)}{|\nabla\varphi(t,x)|}\cdot\nabla\phi(t,x)\,\mathrm{d}x\,\mathrm{d}t = 0,$$

yielding the conclusion. □

Acknowledgements

The authors would like to thank Filippo Santambrogio for the fruitful discussions that lead to this paper.

References

[1] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM J. Numer. Anal.*, 48(3):1136–1162, 2010.

[2] Y. Achdou and A. Porretta. Convergence of a finite difference scheme to weak solutions of the system of partial differential equations arising in mean field games. *SIAM J. Numer. Anal.*, 54(1):161–186, 2016.

[3] Y. Achdou and A. Porretta. Mean field games with congestion. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 35(2):443–480, 2018.

[4] G. Albi, M. Bongini, E. Cristiani, and D. Kalise. Invisible control of self-organizing agents leaving unknown environments. *SIAM J. Appl. Math.*, 76(4):1683–1710, 2016.

[5] L. Ambrosio. Transport equation and Cauchy problem for non-smooth vector fields. In *Calculus of variations and nonlinear partial differential equations*, volume 1927 of *Lecture Notes in Math.*, pages 1–41. Springer, Berlin, 2008.

[6] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[7] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[8] J.-P. Aubin and H. Frankowska. *Set-valued analysis*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2009. Reprint of the 1990 edition.

[9] 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., pages 141–171. Birkhäuser/Springer, Cham, 2017.

[10] A. Bensoussan, J. Frehse, and S. C. P. Yam. On the interpretation of the Master Equation. *Stochastic Process. Appl.*, 127(7):2093–2137, 2017.

[11] N. Bourbaki. *Topologie Générale. Chapitres 5 à 10.* Éléments de Mathématique. Springer, 2007.

[12] M. Burger, M. D. 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, dec 2013.

[13] S. Cacace, F. Camilli, and C. Marchi. A numerical method for mean field games on networks. *ESAIM Math. Model. Numer. Anal.*, 51(1):63–88, 2017.

[14] F. Camilli, E. Carlini, and C. Marchi. A model problem for mean field games on networks. *Discrete Contin. Dyn. Syst.*, 35(9):4173–4192, 2015.

[15] P. Cannarsa and R. Capuani. Existence and uniqueness for mean field games with state constraints. Preprint arXiv:1711.01063.

[16] P. Cannarsa and H. Frankowska. Local regularity of the value function in optimal control. *Systems Control Lett.*, 62(9):791–794, 2013.

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

[18] P. Cardaliaguet. Long time average of first order mean field games and weak KAM theory. *Dyn. Games Appl.*, 3(4):473–488, 2013.

[19] P. Cardaliaguet. Notes on mean field games (from P.-L. Lions' lectures at Collège de France). Available at https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf, 2013.

[20] P. Cardaliaguet. Weak solutions for first order mean field games with local coupling. In *Analysis and geometry in control theory and its applications*, volume 11 of *Springer INdAM Ser.*, pages 111–158. Springer, Cham, 2015.

[21] P. Cardaliaguet. The convergence problem in mean field games with local coupling. *Applied Mathematics & Optimization*, 76(1):177–215, jun 2017.

[22] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. The master equation and the convergence problem in mean field games. Preprint arXiv:1509.02505.

[23] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta. Long time average of mean field games with a nonlocal coupling. *SIAM J. Control Optim.*, 51(5):3558–3591, 2013.

[24] P. Cardaliaguet, A. R. Mészáros, and F. Santambrogio. First order mean field games with density constraints: pressure equals price. *SIAM J. Control Optim.*, 54(5):2672–2709, 2016.

[25] G. Carlier, C. Jimenez, and F. Santambrogio. Optimal transportation with traffic congestion and Wardrop equilibria. *SIAM J. Control Optim.*, 47(3):1330–1350, 2008.

[26] G. Carlier and F. Santambrogio. A continuous theory of traffic congestion and Wardrop equilibria. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 390(Teoriya Predstavleniĭ, Dinamicheskie Sistemy, Kombinatornye Metody. XX):69–91, 307–308, 2011.

[27] E. Carlini and F. J. Silva. A fully discrete semi-Lagrangian scheme for a first order mean field game problem. *SIAM J. Numer. Anal.*, 52(1):45–67, 2014.

[28] R. Carmona and F. Delarue. The master equation for large population equilibriums. In *Stochastic analysis and applications 2014*, volume 100 of *Springer Proc. Math. Stat.*, pages 77–128. Springer, Cham, 2014.

[29] R. M. Colombo and M. D. Rosini. Pedestrian flows and non-classical shocks. *Math. Methods Appl. Sci.*, 28(13):1553–1567, 2005.

[30] E. Cristiani and F. S. Priuli. A destination-preserving model for simulating Wardrop equilibria in traffic flow on networks. *Netw. Heterog. Media*, 10(4):857–876, 2015.

[31] E. Cristiani, F. S. Priuli, and A. Tosin. Modeling rationality to control self-organization of crowds: an environmental approach. *SIAM J. Appl. Math.*, 75(2):605–629, 2015.

[32] M. C. Delfour and J.-P. Zolésio. Shape analysis via oriented distance functions. *J. Funct. Anal.*, 123(1):129–201, 1994.

[33] S. Dweik. Optimal transportation with boundary costs and summability estimates on the transport density. *J. Convex Anal.*, 25(1):135–160, 2018.

[34] D. Evangelista, R. Ferreira, D. A. Gomes, L. Nurbekyan, and V. Voskanyan. First-order, stationary mean-field games with congestion. *Nonlinear Anal.*, 173:37–74, 2018.

[35] S. Faure and B. Maury. Crowd motion from the granular standpoint. *Math. Models Methods Appl. Sci.*, 25(3):463–493, 2015.
[36] D. A. Gomes and H. Mitake. Existence for stationary mean-field games with congestion and quadratic Hamiltonians. *NoDEA Nonlinear Differential Equations Appl.*, 22(6):1897–1910, 2015.
[37] D. A. Gomes, J. Mohr, and R. R. Souza. Continuous time finite state mean field games. *Appl. Math. Optim.*, 68(1):99–143, 2013.
[38] D. A. Gomes and J. Saúde. Mean field games models—a brief survey. *Dyn. Games Appl.*, 4(2):110–154, 2014.
[39] D. A. Gomes and V. K. Voskanyan. Short-time existence of solutions for mean-field games with congestion. *J. Lond. Math. Soc. (2)*, 92(3):778–799, 2015.
[40] A. Granas and J. Dugundji. *Fixed point theory.* Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.
[41] O. Guéant. New numerical methods for mean field games with quadratic costs. *Netw. Heterog. Media*, 7(2):315–336, 2012.
[42] O. Guéant. Existence and uniqueness result for mean field games with congestion effect on graphs. *Appl. Math. Optim.*, 72(2):291–303, 2015.
[43] O. Guéant, J.-M. Lasry, and P.-L. Lions. Mean field games and applications. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 205–266. Springer, Berlin, 2011.
[44] D. Helbing, I. Farkas, and T. Vicsek. Simulating dynamical features of escape panic. *Nature*, 407(6803):487, 2000.
[45] L. F. Henderson. The statistics of crowd fluids. *Nature*, 229(5284):381–383, feb 1971.
[46] 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*, volume 1, pages 98–103. IEEE, 2003.
[47] 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(9):1560–1571, 2007.
[48] 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(3):221–251, 2006.
[49] R. L. Hughes. A continuum theory for the flow of pedestrians. *Transportation Research Part B: Methodological*, 36(6):507–535, jul 2002.
[50] R. L. Hughes. The flow of human crowds. *Annu. Rev. Fluid Mech.*, 35:169–182, 2003.
[51] S. Kakutani. A generalization of Brouwer's fixed point theorem. *Duke Math. J.*, 8:457–459, 1941.
[52] V. N. Kolokoltsov, M. Troeva, and W. Yang. On the rate of convergence for the mean-field approximation of controlled diffusions with large number of players. *Dyn. Games Appl.*, 4(2):208–230, 2014.
[53] A. Lachapelle and M.-T. Wolfram. On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation Research Part B: Methodological*, 45(10):1572–1589, dec 2011.
[54] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.
[55] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.
[56] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
[57] B. Maury, A. Roudneff-Chupin, and F. Santambrogio. A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.*, 20(10):1787–1821, 2010.
[58] G. Mazanti and F. Santambrogio. Minimal-time mean field games. Preprint arXiv: 1804.03246.
[59] A. R. Mészáros and F. J. Silva. A variational approach to second order mean field games with density constraints: the stationary case. *J. Math. Pures Appl. (9)*, 104(6):1135–1159, 2015.
[60] B. Piccoli and A. Tosin. Time-evolving measures and macroscopic modeling of pedestrian flow. *Arch. Ration. Mech. Anal.*, 199(3):707–738, 2011.
[61] A. Prosinski and F. Santambrogio. Global-in-time regularity via duality for congestion-penalized mean field games. *Stochastics*, 89(6-7):923–942, 2017.
[62] F. Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.

[63] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[64] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proc. Inst. Civ. Eng.*, 1(3):325–362, 1952.

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France.

E-mail address: samer.dweik@math.u-psud.fr

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France.

E-mail address: guilherme.mazanti@math.u-psud.fr