HAL
open science

On the Properties of the Value Function Associated to a Mean-Field Optimal Control Problem of Bolza Type *

Benoît Bonnet, Hélène Frankowska

To cite this version:

Benoît Bonnet, Hélène Frankowska. On the Properties of the Value Function Associated to a Mean-Field Optimal Control Problem of Bolza Type *. 2021. hal-03306905

HAL Id: hal-03306905

https://hal.science/hal-03306905

Preprint submitted on 29 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the Properties of the Value Function Associated to a Mean-Field Optimal Control Problem of Bolza Type*

Benoît Bonnet† and Hélène Frankowska‡

July 29, 2021

Abstract

In this paper, we obtain several structural results for the value function associated to a mean-field optimal control problem of Bolza type in the space of measures. After establishing the sensitivity relations bridging between the costates of the maximum principle and metric superdifferentials of the value function, we investigate semiconcavity properties of this latter with respect to both variables. We then characterise optimal trajectories using set-valued feedback mappings defined in terms of suitable directional derivatives of the value function.

1 Introduction

During the past decade, variational problems formulated on mean-field approximations of collective dynamics have gained a tremendous amount of steam. Indeed, the mathematical field of *multi-agent systems* analysis has been for several years the stage on which modern theories such as mean-field control [4, 15] and mean-field games [21, 22] are being developed.

The main rationale behind studying the former of these two fields of research can be heuristically summarised as follows. From a practical standpoint, it is usually easier to model collective dynamics by means of large families of *microscopic* entities – called the agents –, which evolution is prescribed e.g. by ODEs [18] or time-dependent graphs [23]. However, such formulations often lead to numerically intractable problems, and rarely allow to capture global properties of the system in a meaningful way. For these reasons, an ever-expanding literature has been focusing on the development of mathematical tools for the investigation of adequate *macroscopic approximations* of such systems, see e.g. [9, 13, 14, 25] and references therein. In this context, an intensive research effort has been directed towards the establishment of a theory of *mean-field optimal control* [5, 7, 8, 17, 16, 20]. This denomination usually refers to optimal control problems formulated in the so-called *Wasserstein spaces* (see Definition 2 below) of optimal transport [2], which constitute a convenient and natural framework for the generalisation of the classical geometric results of control theory to this setting.

In this paper, we transpose to this class of problems a series of result that were originally derived in [10], dealing with fine properties of the value function associated to a general family of smooth optimal control problems. With this aim, we first investigate the so-called *sensitivity relations* – originally established in [19] – which provide a link between the costate variables stemming from the Pontryagin maximum principle [26] ("PMP" for short) and the superdifferentials of the value function. Among their many possible applications, the sensitivity relations are very useful for expressing sufficient conditions to ensure the optimality of Pontryagin's extremals. We proceed by providing a list of conditions under which the value function is *semiconcave* [11] along suitably chosen interpolating curves. This notion of regularity is indeed very rich, and appears in a wide number of problems

* *This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-18-1-0254.*

†CNRS, IMJ-PRG, UMR 7586, Sorbonne Université, 4 place Jussieu, 75252 Paris, France benoit.bonnet@imj-prg.fr

‡CNRS, IMJ-PRG, UMR 7586, Sorbonne Université, 4 place Jussieu, 75252 Paris, France. helene.frankowska@imj-prg.fr

originating both in control theory and in the calculus of variations, see for instance [12, 27]. Moreover in the context of Wasserstein spaces, it has been well established in the monograph [2] that semiconcavity properties along interpolating curves are in fact the natural path leading to meaningful results on metric superdifferentials in the space of measures. By leveraging some of these results, it is then possible to characterise optimal feedbacks for mean-field optimal control problems, and to provide useful regularity properties for the latter.

The structure of the article is the following. In Section 2, we recollect notions of measure theory and optimal transport, and proceed by recalling several results which are more specific to mean-field optimal control problems in Section 3. In Section 4, we study the sensitivity relations and semiconcavity properties of the value function for a general Bolza problem in Wasserstein spaces. We then make use of these results to investigate optimal feedbacks and their regularity in Section 5.

2 Preliminaries

In this section, we recall preliminary notions of measure theory and optimal transport, for which we refer to [1] and [2] respectively.

2.1 Optimal transport and calculus in Wasserstein spaces

In the sequel, $\mathscr{P}(\mathbb{R}^d)$ denotes the space of Borel probability measures over $\mathbb{R}^d$, endowed with the standard *narrow topology* (see [2, Chapter 5]). Given $p \in [1, +\infty)$, let $\mathscr{P}_p(\mathbb{R}^d)$ be the subset of probability measure which p-momentum $\mathcal{M}_p(\mu) := \left(\int_{\mathbb{R}^d} |x|^p \mathrm{d}\mu(x)\right)^{1/p}$ is finite, and $\mathscr{P}_c(\mathbb{R}^d)$ be that of measures with compact support. For $\mu \in \mathscr{P}(\mathbb{R}^d)$, we denote by $L^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ the Banach space of maps from $\mathbb{R}^d$ into itself that are p-summable with respect to μ. When $p = 2$, the latter is a Hilbert space endowed with the scalar product

$$\langle \zeta, \xi \rangle_{L^2(\mu)} := \int_{\mathbb{R}^d} \langle \zeta(x), \xi(x) \rangle \mathrm{d}\mu(x)$$

defined here for any $\zeta, \xi \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$. We will also denote by $\mathscr{L}^1$ the standard Lebesgue measure over $\mathbb{R}$, and consider the associated spaces $L^p(I, \mathbb{R}_+)$ of p-summable maps from an interval $I \subset \mathbb{R}$ into $\mathbb{R}_+$. Given a Borel map $f : \mathbb{R}^d \to \mathbb{R}^d$, we define the *pushforward* $f_\# \mu \in \mathscr{P}(\mathbb{R}^d)$ of μ through f as the measure $f_\# \mu(B) := \mu(f^{-1}(B))$, for every Borel set $B \subset \mathbb{R}^d$.

Definition 1 (Transport plans)**.** *The set of* transport plans *between two measures $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$ – denoted by $\Gamma(\mu, \nu)$ – is the collection of elements $\gamma \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d)$ such that $\pi^1_\# \gamma = \mu$ and $\pi^2_\# \gamma = \nu$, where $\pi^1, \pi^2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ represent the projection operator onto the first and second factor respectively.*

In the sequel, we will also need the following simplified version of [2, Theorem 5.3.1].

Theorem 1 (Barycentric projection)**.** *Let $p \in [1, +\infty)$, $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$ and $\gamma \in \Gamma(\mu, \nu)$. Then, there exists a unique map $\bar{\gamma} \in L^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ – called the* barycentric projection *of γ on $\pi^1_\# \gamma = \mu$ –, such that*

$$\int_{\mathbb{R}^{2d}} \langle y, \mathcal{F}(x) \rangle \mathrm{d}\gamma(x, y) = \langle \bar{\gamma}, \mathcal{F} \rangle_{L^2(\mu)},$$

for every $\mathcal{F} \in L^\infty(\mathbb{R}^d, \mathbb{R}^d; \mu)$.

Definition 2 (Wasserstein distances)**.** *The quantity*

$$W_p(\mu, \nu) := \min_{\gamma \in \Gamma(\mu, \nu)} \left(\int_{\mathbb{R}^{2d}} |x - y|^p \mathrm{d}\gamma(x, y) \right)^{1/p},$$

defines a distance between any two measures $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$, and we denote by $\Gamma_o(\mu, \nu)$ the set of transport plans for which the minimum is attained.

The Wasserstein space of order $p \in [1, +\infty)$ is defined as the metric space $(\mathscr{P}_p(\mathbb{R}^d), W_p)$. In the sequel, we will often consider the (non complete) metric space $(\mathscr{P}_c(\mathbb{R}^d), W_p)$ of compactly supported measures equipped with the W_p-metric. When $p = 2$, it has been known since [2, 24] that $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ could be endowed – at least formally – with a differentiable manifold-like structure. We present below a revisited definition of the corresponding calculus, tailored for $(\mathscr{P}_c(\mathbb{R}^d), W_2)$, which is borrowed from [7]. Given $R > 0$, we will use the notation $B_R(\mu) := \cup_{x \in \mathrm{supp}(\mu)} B(x, R)$.

Definition 3 (Local differentials)**.** *A functional $\phi : \mathscr{P}_2(\mathbb{R}^d) \mapsto \mathbb{R} \cup \{\pm\infty\}$ is* locally differentiable *at $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ if and only if there exists $\nabla\phi(\mu) \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$ such that for all $R > 0$ and $\nu \in \mathscr{P}(B_R(\mu))$, it holds*

$$\phi(\nu) = \phi(\mu) + \int_{\mathbb{R}^{2d}} \langle \nabla\phi(\mu), y - x\rangle \mathrm{d}\gamma(x, y) + o_R(W_2(\mu, \nu)),$$

for every $\gamma \in \Gamma_o(\mu, \nu)$. We also say that $\phi(\cdot)$ is locally continuously differentiable *if the application $(\mu, x) \in \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \nabla\phi(\mu)(x) \in \mathbb{R}^d$ is continuous.*

In [7], we proved the following chain rule for locally differentiable functionals.

Proposition 1 (Chain rule)**.** *Consider elements $\mu \in \mathscr{P}_c(\mathbb{R}^d)$, $\mathcal{F} \in L^\infty(\mathbb{R}^d, \mathbb{R}^d; \mu)$ and let $\phi : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{\pm\infty\}$ be locally differentiable at μ. Then, it holds*

$$\phi((\mathrm{Id} + \epsilon\mathcal{F})_\# \mu) = \phi(\mu) + \epsilon\langle \nabla\phi(\mu), \mathcal{F}\rangle_{L^2(\mu)} + o(\epsilon),$$

for every $\epsilon \in \mathbb{R}$.

2.2 Non-local continuity equations

Consider the *non-local continuity equation* in a Wasserstein space

$$\partial_t \mu(t) + \mathrm{div}_x(v(t, \mu(t))\mu(t)) = 0, \quad \mu(0) = \mu^0, \tag{1}$$

where the *non-local velocity-field* $v : [0, T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies the following assumptions.

Hypotheses (CE). *Suppose that for any compact set $K \subset \mathbb{R}^d$, the following holds.*

(i) The mapping $t \mapsto v(t, \mu, x) \in \mathbb{R}^d$ is $\mathscr{L}^1$-measurable with respect to $t \in [0, T]$ for all $(\mu, x) \in \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$, and there exists $m(\cdot) \in L^1([0, T], \mathbb{R}_+)$ such that

$$|v(t, \mu, x)| \leq m(t)\Big(1 + |x| + \mathcal{M}_1(\mu)\Big).$$

(ii) There exists a map $\ell_K(\cdot) \in L^1([0, T], \mathbb{R}_+)$ such that

$$|v(t, \nu, y) - v(t, \mu, x)| \leq \ell_K(t)\big(W_1(\mu, \nu) + |x - y|\big)$$

for $\mathscr{L}^1$-almost every $t \in [0, T]$, any $\mu, \nu \in \mathscr{P}(K)$ and all $x, y \in K$.

Under these standard Cauchy-Lipschitz assumptions, we have the following well-posedness result for (1) (see e.g. [6]).

Theorem 2 (Well-posedness of (1))**.** *Let $r > 0$, $\mu^0 \in \mathscr{P}(B(0, r))$ and suppose that hypotheses* **(CE)** *hold. Then, there exists a unique distributional solution $\mu(\cdot)$ of (1), which also satisfies*

$$\mathrm{supp}(\mu(t)) \subset B(0, R_r) \qquad \text{and} \qquad W_1(\mu(t), \mu(\tau)) \leq \int_\tau^t m_r(s)\mathrm{d}s,$$

for all $0 \leq \tau \leq t \leq T$, with $R_r > 0$ and $m_r(\cdot) \in L^1([0, T], \mathbb{R}_+)$ depending on $r > 0$ and $m(\cdot)$. Moreover, the curve $\mu(\cdot)$ can be represented explicitly as

$$\mu(t) = \Phi_{(\tau,t)}[\mu(\tau)](\cdot)_\# \mu(\tau),$$

for every $\tau, t \in [0, T]$, where $(\Phi_{(\tau,t)}[\mu](\cdot))_{\tau,t \in [0,T]}$ denotes the non-local flow *starting from $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ at time $\tau \in [0, T]$, defined as the unique solution of*

$$\begin{cases} \partial_t \Phi_{(\tau,t)}[\mu](x) = v\Big(t, \Phi_{(\tau,t)}[\mu](\cdot)_\# \mu, \Phi_{(\tau,t)}[\mu](x)\Big), \\ \Phi_{(\tau,\tau)}[\mu](x) = x, \end{cases} \tag{2}$$

for all $x \in \mathbb{R}^d$.

3 Mean-field optimal control

In this section, we focus more in details on the following Bolza mean-field optimal control problem

$$(\mathcal{P})\quad \begin{cases} \min\limits_{u(\cdot)\in\mathcal{U}} \Big[\displaystyle\int_0^T L(\mu(t),u(t))\mathrm{d}t + \varphi(\mu(T)) \Big], \\ \text{s.t.} \begin{cases} \partial_t\mu(t) + \mathrm{div}_x(v(t,\mu(t),u(t))\mu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \end{cases}$$

Here $T > 0$, $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times U \times \mathbb{R}^d \to \mathbb{R}^d$ is a controlled velocity-field, the set of admissible controls is $\mathcal{U} := L^2([0,T],U)$ where $U \subset \mathbb{R}^m$ is a closed set, and $\mu^0 \in \mathscr{P}_c(\mathbb{R}^d)$ is a fixed initial datum. In the present paper, we make the following assumptions on $(\mathcal{P})$.

Hypotheses (OCP). *Suppose that for any compact set $K \subset \mathbb{R}^d$, the following holds.*

(i) The map $u \in U \mapsto v(t,\mu,u,x) \in \mathbb{R}^d$ is continuous, and there exist constants $m, \ell_K > 0$ such that

$$\begin{cases} |v(t,\mu,u,x)| \le m\Big(1 + |x| + \mathcal{M}_1(\mu)\Big)(1+|u|^2), \\ |v(\tau,\nu,u,y) - v(t,\mu,u,x)| \le \ell_K\big(|t-\tau| + W_1(\mu,\nu) + |x-y|\big)(1+|u|^2), \end{cases}$$

for all $t,\tau \in [0,T]$, any $\mu,\nu \in \mathscr{P}(K)$, every $u \in U$ and each $x,y \in K$.

(ii) The map $u \in U \mapsto L(\mu,u) \in \mathbb{R}_+$ is lower semicontinuous, and there exist constants $l^1, \mathcal{C}^1, \mathcal{L}^1_K > 0$ such that $L(\mu,u) \ge \mathcal{C}^1|u|^2$ and

$$\begin{cases} L(\mu,u) \le l^1(1+\mathcal{M}_1(\mu))(1+|u|^2), \\ |L(\nu,u) - L(\mu,u)| \le \mathcal{L}^1_K W_1(\mu,\nu)(1+|u|^2), \end{cases}$$

for any $\mu,\nu \in \mathscr{P}(K)$ and every $u \in U$.

(iii) There exist $\mathcal{C}^2 \in \mathbb{R}$ and $\mathcal{L}^2_K > 0$ such that

$$\varphi(\mu) \ge \mathcal{C}^2 \quad \text{and} \quad |\varphi(\nu) - \varphi(\mu)| \le \mathcal{L}^2_K W_1(\mu,\nu),$$

for any $\mu,\nu \in \mathscr{P}(K)$.

(iv) The maps $(x,\mu) \mapsto v(t,\mu,u,x), L(u,\mu), \varphi(\mu)$ are locally continuously differentiable in the sense of Definition 3.

While the assumptions listed above are not minimal for the developments of this section (see e.g. more general conditions in [7, Section 4.2]), they will be necessary to establish several results in Section 4 and Section 5.

Theorem 3 (Existence)**.** *Suppose that* **(OCP)** *hold and that for any non-empty compact $K \subset \mathbb{R}^d$ and every $(t,\mu) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d)$, the restricted sets*

$$\Big\{ (v(t,\mu,u)_{|K}, L(\mu,u) + r) \ \text{s.t.}\ u \in U,\ r \ge 0 \Big\} \subset C^0(K,\mathbb{R}^d) \times \mathbb{R},$$

are closed and convex. Then, there exists an optimal trajectory-control pair $(\mu^*(\cdot), u^*(\cdot))$ for $(\mathcal{P})$.

Proof. Under hypothesis **(OCP)**-(ii), there exists for any minimising sequence $(\mu^*_N(\cdot), u^*_N(\cdot))$ a constant $\mathcal{C}_U > 0$ such that

$$\sup_{N\ge1} \int_0^T |u^*_N(t)|^2 \mathrm{d}t \le \mathcal{C}_U.$$

The result then follows up to minor modifications in the proofs of [6, Theorem 6 and Theorem 7]. □

In the sequel, we will assume that $(\mathcal{P})$ admits such an optimal pair $(\mu^*(\cdot), u^*(\cdot))$. By a simple adaptation of the arguments of [7], it can be shown that the latter satisfies the following Pontryagin-type optimality conditions.

Theorem 4 (Pontryagin Maximum Principle)**.** *Let $(\mu^*(\cdot), u^*(\cdot))$ be an optimal pair for $(\mathcal{P})$. Then, there exists a unique state-costate curve $\nu^* : [0,T] \mapsto \mathscr{P}_c(\mathbb{R}^{2d})$ which satisfies the following.*

(i) It solves the Hamiltonian continuity equation

$$\begin{cases} \partial_t \nu^*(t) + \mathrm{div}_{x,r}(\nabla_\nu \mathbb{H}(t, \nu^*(t), u^*(t))\nu^*(t)) = 0, \\ \pi^1_\# \nu^*(t) = \mu^*(t) \qquad\qquad \text{for all times } t \in [0,T], \\ \nu^*(T) = (\mathrm{Id}, -\nabla\varphi(\mu^*(T)))_\# \mu^*(T), \end{cases}$$

where the Hamiltonian $\mathbb{H}$ is defined by

$$\mathbb{H}(t, \nu, u) := \int_{\mathbb{R}^{2d}} \langle r, v(t, \mu, u, x)\rangle \mathrm{d}\nu(x, r) - L(\mu, u), \tag{3}$$

with $\mu := \pi^1_\# \nu$ and $\nabla_\nu \mathbb{H}(t, \nu^*(t), u^*(t))$ being its gradient given in the sense of Definition 3.

(ii) The maximisation condition

$$\mathbb{H}(t, \nu^*(t), u^*(t)) = \max_{u \in U} \mathbb{H}(t, \nu^*(t), u), \tag{4}$$

holds for $\mathscr{L}^1$-almost every $t \in [0,T]$.

4 Properties of the value function

In this section, we use some of the tools introduced above to study fine properties of the *value function*

$$\mathscr{V}(\tau, \mu) := \begin{cases} \min\limits_{u(\cdot) \in \mathcal{U}} \Big[\displaystyle\int_\tau^T L(\mu(t), u(t))\mathrm{d}t + \varphi(\mu(T)) \Big], \\ \quad \text{s.t.} \begin{cases} \partial_t \mu(t) + \mathrm{div}_x(v(t, \mu(t), u(t))\mu(t)) = 0, \\ \mu(\tau) = \mu, \end{cases} \end{cases}$$

associated to $(\mathcal{P})$, defined for $(\tau, \mu) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d)$. A first key observation is that under our working assumptions, the mapping $\mathscr{V}(\cdot, \cdot)$ is absolutely continuous in its first argument and Lipschitz in the second one.

Proposition 2 (Regularity)**.** *Suppose that hypotheses* **(OCP)** *hold, and let $r > 0$. Then, there exist $\mathcal{M}_r(\cdot) \in L^1([0,T], \mathbb{R}_+)$ and $\mathcal{L}_r > 0$ such that*

$$|\mathscr{V}(\tau_2, \mu_2) - \mathscr{V}(\tau_1, \mu_1)| \leq \int_{\tau_1}^{\tau_2} \mathcal{M}_r(t)\mathrm{d}t + \mathcal{L}_r W_1(\mu_1, \mu_2),$$

for every $\tau_1, \tau_2 \in [0,T]$ and all $\mu_1, \mu_2 \in \mathscr{P}(B(0,r))$.

Proof. This result follows from the Lipschitz dependence of solutions of non-local continuity equations on their initial data (see [6]) together with the sublinearity and Lipschitz regularity of the cost functionals. □

In what follows, we will first investigate the so-called sensitivity relations involving the costates of the PMP of Theorem 4, and then semiconcavity properties of $\mathscr{V}(\cdot, \cdot)$. Given $u(\cdot) \in \mathcal{U}$, we will use the notation $(\Phi^u_{(\tau,t)}[\cdot](\cdot))_{\tau,t \in [0,T]}$ for the non-local flows generated by $(t, \mu, x) \mapsto v(t, \mu, u(t), x) \in \mathbb{R}^d$ defined as in (2).

4.1 Sensitivity analysis and Pontryagin costates

In our context, the sensitivity relations link the state-costate curves $\nu^*(\cdot)$ of the maximum principle and the *Hadamard superdifferential* of the value function.

Definition 4 (Hadamard superdifferentials)**.** *Let $(\tau,\mu)\in[0,T]\times\mathscr{P}_c(\mathbb{R}^d)$. Then, we say that a pair $(\delta,\xi)\in\mathbb{R}\times L^2(\mathbb{R}^d,\mathbb{R}^d;\mu)$ belongs to the Hadamard superdifferential $\eth^+\mathscr{V}(\tau,\mu)$ of $\mathscr{V}(\cdot,\cdot)$ at (τ,μ) provided that for all $(h,\mathcal{F})\in\mathbb{R}\times L^\infty(\mathbb{R}^d,\mathbb{R}^d;\mu)$*

$$\limsup_{\substack{\epsilon\to 0^+\\ \tau+\epsilon h\in[0,T]}}\Big[\frac{\mathscr{V}(\tau+\epsilon h,(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu)-\mathscr{V}(\tau,\mu)}{\epsilon}\Big]\le\delta h+\langle\xi,\mathcal{F}\rangle_{L^2(\mu)}.$$

On vector spaces, Hadamard superdifferentials are usually defined in terms of contingent directional derivatives. Yet, we showed in Proposition 2 above that under **(OCP)**, the map $\mathscr{V}(t,\cdot)$ is locally Lipschitz over $\mathscr{P}_c(\mathbb{R}^d)$ with local Lipschitz constants independent of $t\in[0,T]$. This allows us to use the simpler upper Dini directional derivatives to define Hadamard superdifferentials.

Theorem 5 (Hadamard-type sensitivity)**.** *Suppose that hypotheses* **(OCP)** *hold. Let $(\mu^*(\cdot),u^*(\cdot))$ be a solution of $(\mathcal{P})$ and $\nu^*(\cdot)$ be the corresponding state-costate curve given by Theorem 4. Then*

$$\big(\mathbb{H}(t,\nu^*(t),u^*(t)),-\bar\nu^*(t)\big)\in\eth^+\mathscr{V}(t,\mu^*(t)),$$

for $\mathscr{L}^1$-almost every $t\in[0,T]$, where the map $\bar\nu^*(t)\in L^2(\mathbb{R}^d,\mathbb{R}^d;\mu^*(t))$ denotes the barycentric projection *of $\nu^*(t)$ onto its first marginal $\pi^1_\#\nu^*(t)=\mu^*(t)$.*

Proof. The proof of this result is fairly long and technical, and will thus be published in greater details elsewhere. For the convenience of the reader, we sketch the main steps of the underlying arguments. Given $(\tau,\epsilon)\in[0,T]\times\mathbb{R}_+$ and $(h,\mathcal{F})\in\mathbb{R}\times L^\infty(\mathbb{R}^d,\mathbb{R}^d;\mu^*(\tau))$, the perturbed curve $\tilde\mu_\epsilon(\cdot)$ defined by

$$\tilde\mu_\epsilon(t):=\Big(\Phi^{u^*}_{(\tau+\epsilon h,t)}[(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu^*(\tau)]\circ(\mathrm{Id}+\epsilon\mathcal{F})\Big)_\#\mu^*(\tau),$$

for all times $t\in[\tau,T]$ satisfies

$$\mathscr{V}(\tau+\epsilon h,(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu^*(\tau))\le\int_{\tau+\epsilon h}^T L(\tilde\mu_\epsilon(t),u^*(t))\mathrm{d}t+\varphi(\tilde\mu_\epsilon(T)),$$

by definition of $\mathscr{V}(\cdot,\cdot)$. Following technical linearisation arguments in the spirit of [5, 8], one can show

$$\Phi^{u^*}_{(\tau+\epsilon h,t)}[(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu^*(\tau)](x+\epsilon\mathcal{F}(x))=\Phi^*_{(\tau,t)}(x)+\epsilon w_{h,\mathcal{F}}(t,\Phi^*_{(t,\tau)}(x))+o_{\tau,t,x}(\epsilon),$$

with $\sup_{(\tau,t,x)}|o_{\tau,t,x}(\epsilon)|=o(\epsilon)$ and where we denoted $\Phi^*_{(\tau,t)}(x):=\Phi^{u^*}_{(\tau,t)}[\mu^*(\tau)](x)$. Therein, $t\in[\tau,T]\mapsto w_{h,\mathcal{F}}(t,y)$ solves a suitable linearised Cauchy problem.

Combining the chain rule of Proposition 1 with the optimality of $(\mu^*(\cdot),u^*(\cdot))$, the following inequality

$$\begin{aligned}\tfrac{1}{\epsilon}\Big(\mathscr{V}(\tau+\epsilon h,(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu^*(\tau))-\mathscr{V}(\tau,\mu^*(\tau))\Big)\le&\int_\tau^T\langle\nabla_\mu L(\mu^*(t),u^*(t)),w_{h,\mathcal{F}}(t,\Phi^*_{(t,\tau)}(\cdot))\rangle_{L^2(\mu^*(\tau))}\mathrm{d}t\\&+\langle\nabla\varphi(\mu^*(T)),w_{h,\mathcal{F}}(T,\Phi^*_{(T,\tau)}(\cdot))\rangle_{L^2(\mu^*(T))}\\&-hL(\mu^*(\tau),u^*(\tau))+o_\tau(1),\end{aligned}$$

holds for every $\epsilon>0$, provided that $\tau\in[0,T]$ belongs to a suitable subset of full $\mathscr{L}^1$-measure. Upon performing lengthy computations involving the dynamics of the state-costate curve $\nu^*(\cdot)$ introduced in Theorem 4, this inequality can be further reformulated as

$$\begin{aligned}&\tfrac{1}{\epsilon}\Big(\mathscr{V}(\tau+\epsilon h,(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu^*(\tau))-\mathscr{V}(\tau,\mu^*(\tau))\Big)\\&\qquad\le\mathbb{H}(\tau,\nu^*(\tau),u^*(\tau))\,h-\int_{\mathbb{R}^{2d}}\langle r,\mathcal{F}(x)\rangle\mathrm{d}\nu^*(\tau)(x,r)+o_\tau(1)\\&\qquad\le\mathbb{H}(\tau,\nu^*(\tau),u^*(\tau))\,h-\langle\bar\nu^*(\tau),\mathcal{F}\rangle_{L^2(\mu^*(\tau))}+o_\tau(1),\end{aligned}\tag{5}$$

for any $\epsilon>0$ small enough, and where we used the barycentric projection of $\nu^*(\tau)$ defined as in Theorem 1. We end the proof by taking the limsup as $\epsilon\to0^+$ in (5) for $\mathscr{L}^1$-almost every $t\in[0,T]$. ☐

The main application of the sensitivity relations is to provide *sufficient* optimality conditions for the PMP.

Theorem 6 (Sufficient optimality conditions)**.** *Let $(\mu^*(\cdot), u^*(\cdot))$ be an admissible pair for $(\mathcal{P})$ and $\nu^*(\cdot)$ be an associated state-costate curve satisfying the PMP of Theorem 4 and the sensitivity relation of Theorem 5. Then, the pair $(\mu^*(\cdot), u^*(\cdot))$ is optimal for $(\mathcal{P})$.*

Proof. Observe that a pair $(\mu^*(\cdot), u^*(\cdot))$ is optimal for $(\mathcal{P})$ if and only if

$$\mathscr{V}(\tau_2, \mu^*(\tau_2)) = \mathscr{V}(\tau_1, \mu^*(\tau_1)) + \int_{\tau_1}^{\tau_2} L(\mu^*(t), u^*(t))\mathrm{d}t,$$

for all $0 \leq \tau_1 \leq \tau_2 \leq T$. By Proposition 2, the map $t \in [0,T] \mapsto \mathscr{V}(t, \mu^*(t))$ is absolutely continuous, and by definition of the value function it necessary holds

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathscr{V}(t, \mu^*(t)) \geq L(\mu^*(t), u^*(t)), \tag{6}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$. Following [7, Section 3.2], it can be shown that

$$\Phi^*_{(t,t+\epsilon)}(x) = x + \epsilon v(t, \mu^*(t), u^*(t), x) + o_{t,x}(\epsilon),$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and for any $\epsilon \in \mathbb{R}$, with $\int_0^T \sup_x |o_{t,x}(\epsilon)|\mathrm{d}t = o(\epsilon)$. This combined with (2) and Proposition 2 then yields

$$\mathscr{V}(t+\epsilon, \mu^*(t+\epsilon)) = \mathscr{V}(t+\epsilon, (\mathrm{Id} + \epsilon v(t, \mu^*(t), u^*(t)))_{\#}\mu^*(t)) + o_t(\epsilon) \tag{7}$$

with $\int_0^T |o_t(\epsilon)|\mathrm{d}t = o(\epsilon)$. By the sensitivity relation of Theorem 5, it further holds

$$\begin{aligned}\limsup_{\epsilon\to 0^+} \Big[\tfrac{1}{\epsilon}\Big(\mathscr{V}(t+\epsilon, (\mathrm{Id} + \epsilon v(t, \mu^*(t), u^*(t)))_{\#}\mu^*(t)) - \mathscr{V}(t, \mu^*(t))\Big)\Big] \\ \leq -\langle \bar{\nu}^*(t), v(t, \mu^*(t), u^*(t))\rangle_{L^2(\mu^*(t))} + \mathbb{H}(t, \nu^*(t), u^*(t)) \\ = L(\mu^*(t), u^*(t)),\end{aligned} \tag{8}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$, where we used the definition (3) of the Hamiltonian. Upon combining (7) and (8), we can thus recover

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathscr{V}(t, \mu^*(t)) \leq L(\mu^*(t), u^*(t)),$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$, which together with (6) concludes the proof of our claim up to an integration over any time interval of the form $[\tau_1, \tau_2]$. □

4.2 Regularity of the value function

Our aim now is to study semiconcavity estimates for $\mathscr{V}(\cdot,\cdot)$, taking inspiration from [2, Chapter 9]. We start our developments by recalling in which sense this notion is understood in Wasserstein spaces.

Definition 5 (Interpolation semiconcavity)**.** *A map $\phi : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{\pm\infty\}$ is* locally semiconcave *if it is finite on $\mathscr{P}_c(\mathbb{R}^d)$, and for every $r > 0$ there exists $\mathcal{C}_r > 0$ such that for any $\lambda \in [0,1]$,*

$$(1-\lambda)\phi(\mu_1) + \lambda\phi(\mu_2) - \phi(\boldsymbol{\mu}^{1\to 2}_\lambda) \leq \mathcal{C}_r \lambda(1-\lambda) W^2_{2,\boldsymbol{\mu}}(\mu_1, \mu_2),$$

for any $\mu_1, \mu_2 \in \mathscr{P}(B(0,r))$ and each $\boldsymbol{\mu} \in \Gamma(\mu,\nu)$, where $\boldsymbol{\mu}^{1\to 2}_\lambda := ((1-\lambda)\pi^1 + \lambda\pi^2)_{\#}\boldsymbol{\mu}$, and

$$W_{2,\boldsymbol{\mu}}(\mu,\nu) := \Big(\int_{\mathbb{R}^{2d}} |x-y|^2 \mathrm{d}\boldsymbol{\mu}(x,y)\Big)^{1/2}.$$

Below, we make the following additional assumptions on the data of $(\mathcal{P})$.

Hypotheses (R). *Suppose that the following holds.*

(*i*) *For every $r > 0$, there exists a constant $\mathcal{C}_r^v > 0$ such that for all $\lambda \in [0,1]$, it holds*

$$\begin{aligned}\big|(1-\lambda)v(t,\mu_1,u,x) + \lambda v(t,\mu_2,u,y) - v(t,\boldsymbol{\mu}_\lambda^{1\to 2},u,(1-\lambda)x+\lambda y)\big| \\ \leq \mathcal{C}_r^v \lambda(1-\lambda)\Big(W_{2,\boldsymbol{\mu}}^2(\mu_1,\mu_2) + |x-y|^2\Big),\end{aligned}$$

for all $(t,u) \in [0,T] \times U$, any $\mu_1, \mu_2 \in \mathscr{P}(B(0,r))$ and each $x, y \in B(0,r)$.

(*ii*) *The maps $\mu \in \mathscr{P}_c(\mathbb{R}^d) \mapsto L(\mu,u), \varphi(\mu)$ are locally semiconcave, uniformly with respect to $u \in U$.*

Theorem 7 (Semiconcavity)**.** *Suppose that hypotheses* **(OCP)** *and* **(R)** *hold. Then for every $r > 0$, there exists $\mathscr{C}_r > 0$ such that for any $\lambda \in [0,1]$, it holds*

$$(1-\lambda)\mathscr{V}(\tau_1,\mu_1) + \lambda\mathscr{V}(\tau_2,\mu_2) - \mathscr{V}\big((1-\lambda)\tau_1 + \lambda\tau_2, \boldsymbol{\mu}_\lambda^{1\to 2}\big) \leq \mathscr{C}_r\lambda(1-\lambda)\Big(|\tau_1-\tau_2|^2 + W_{2,\boldsymbol{\mu}}^2(\mu_1,\mu_2)\Big),$$

for every $\tau_1, \tau_2 \in [0,T]$, any $\mu_1, \mu_2 \in \mathscr{P}(B(0,r))$ and each $\boldsymbol{\mu} \in \Gamma(\mu_1,\mu_2)$.

Proof. The proof of this result is again fairly long and technical, and will be exposed in greater details in a subsequent publication. The core idea is to use the refined superposition results from [6, Lemma 1] to prove that a norm-semiconcavity estimate similar to **(R)**-(*i*) holds for the non-local flows $(\Phi^u_{(\tau,t)}[\cdot](\cdot))_{\tau,t\in[0,T]}$. It is then possible to recover the local semiconcavity of $\mathscr{V}(\tau,\cdot)$ for any fixed $\tau \in [0,T]$ by combining the latter estimates with hypotheses **(OCP)**-(*ii*),(*iii*) and **(R)**-(*ii*).

To recover the semiconcavity with respect to both variables, one must perform a suitable rescaling of the time variable in the spirit of [10, Section 5], combined with the semiconcavity estimates already established with respect to the measure variable. □

5 Optimal feedbacks

By leveraging several results of Section 4, it is possible to derive a general characterisation of optimal feedbacks for mean-field optimal control problems. Throughout this section, we assume that hypotheses **(OCP)** hold. We start by recalling the definition of an *upper-semicontinuous* set-valued map ("u.s.c." for short).

Definition 6 (Regular set-valued maps)**.** *A multifunction $\mathcal{G} : \mathscr{S} \rightrightarrows X$ from a metric space $(\mathscr{S}, d_{\mathscr{S}})$ into a Banach space $(X, \|\cdot\|_X)$ is upper-semicontinuous if for any $s \in \mathscr{S}$ and every neighbourhood Ω of $\mathcal{G}(s)$, there exists $\eta > 0$ such that $\mathcal{G}(s') \subset \Omega$ for every $s' \in B(s,\eta)$. In particular, if $\mathcal{G}(\cdot)$ is u.s.c. and single-valued, it is then continuous in the usual sense.*

In order to characterise optimal feedbacks for $(\mathcal{P})$, we need to introduce two notions of inferior derivatives for functionals defined over the space of measures.

Definition 7 (Two notions of lower-derivatives)**.** *Let $\phi : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{\pm\infty\}$ and $\mu \in \mathscr{P}_c(\mathbb{R}^d)$, $\phi(\mu) \neq \{\pm\infty\}$. Given $\mathcal{F} \in C^0(\mathbb{R}^d,\mathbb{R}^d)$, we define the* lower Dini derivative *of $\phi(\cdot)$ at μ in the direction $\mathcal{F}$ as*

$$\partial_{\mathrm{D}}^- \phi(\mu)(\mathcal{F}) := \liminf_{\epsilon\to 0^+}\Big[\frac{\phi((\mathrm{Id}+\epsilon\mathcal{F})_\#\mu) - \phi(\mu)}{\epsilon}\Big].$$

Given $R > 0$, we also define the regularised R-lower derivative *of $\phi(\cdot)$ at μ in the direction $\mathcal{F}$ as*

$$\partial_{o,R}^- \phi(\mu)(\mathcal{F}) := \liminf_{\substack{\nu\to\mu,\,\epsilon\to 0^+\\ \nu\in\mathscr{P}(B_R(\mu))}}\Big[\frac{\phi((\mathrm{Id}+\epsilon\mathcal{F})_\#\nu) - \phi(\nu)}{\epsilon}\Big].$$

Proposition 3 (The case of semiconcave functions)**.** *Suppose that $\phi : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{\pm\infty\}$ is locally semiconcave and let $\mu \in \mathscr{P}_c(\mathbb{R}^d)$. Then, it holds that*

$$\partial_{\mathrm{D}}^- \phi(\mu)(\mathcal{F}) = \partial_{o,R}^- \phi(\mu)(\mathcal{F}),$$

for every $R > 0$ and all $\mathcal{F} \in C^0(\mathbb{R}^d,\mathbb{R}^d)$.

Proof. Adapt the arguments of [10, Section 3]. □

In the following theorem, we combine these notions of non-smooth analysis to provide a general characterisation of optimal feedbacks for $(\mathcal{P})$.

Theorem 8 (Optimal feedbacks)**.** *An admissible pair $(\mu^*(\cdot), u^*(\cdot))$ is optimal for $(\mathcal{P})$ if and only if*

$$v(t, \mu^*(t), u^*(t)) \in \mathcal{G}(t, \mu^*(t)), \tag{9}$$

for $\mathscr{L}^1$-almost every $t \in [0, T]$, where we defined the sets

$$\mathcal{G}(t, \mu) := \Big\{ \boldsymbol{v} \in v(t, \mu, U) \ \text{s.t.} \ \partial_{\mathrm{D}}^{-} \mathscr{V}(t, \mu)(1, \boldsymbol{v}) \leq 0 \Big\}, \tag{10}$$

for all $(t, \mu) \in [0, T] \times \mathscr{P}_c(\mathbb{R}^d)$. If in addition hypotheses **(R)** *hold, the restricted set-valued maps $\mathcal{G}_{|K} : [0, T] \times \mathscr{P}(K) \rightrightarrows C^0(K, \mathbb{R}^d)$ are upper-semicontinuous for every compact set $K \subset \mathbb{R}^d$.*

Proof. The proof of the first part of this result – namely that $(\mu^*(\cdot), u^*(\cdot))$ is optimal if and only if (9) holds $\mathscr{L}^1$-almost everywhere in $[0, T]$ – is quite long, and can be obtained by performing computations similar to those in the proof of Theorem 5 above.

The regularity of the set-valued map $\mathcal{G}(\cdot, \cdot)$ under hypotheses **(R)** can be established via the following arguments. By Theorem 7, the value function $\mathscr{V}(\cdot, \cdot)$ is locally semiconcave. Thus by Proposition 3, the lower-derivative $\partial_{\mathrm{D}}^{-} \mathscr{V}(t, \mu^*(t))(1, v(t, \mu^*(t), u^*(t)))$ coincides with the regularised R-lower derivative $\partial_{o,R}^{-} \mathscr{V}(t, \mu^*(t))(1, v(t, \mu^*(t), u^*(t)))$ for any $R > 0$. By definition of the latter together with the regularity **(OCP)**-(i) of the controlled velocity field $v : [0, T] \times \mathscr{P}_c(\mathbb{R}^d) \times U \times \mathbb{R}^d \to \mathbb{R}^d$, it can then be shown that for every $K \subset \mathbb{R}^d$, the graph of the restricted set-valued map $\mathcal{G}_{|K}(\cdot, \cdot)$ is closed. Moreover under hypotheses **(OCP)**, it can be noted that $\mathcal{G}_{|K}(\cdot, \cdot)$ takes its values in a compact subset of $C^0(K, \mathbb{R}^d)$. By classical results in set-valued analysis (see e.g. [3, Section 1.4]), every such multifunction with closed graph is upper-semicontinuous, which ends the proof of our claim. □

In conclusion, the results of Theorem 8 inform us that every optimal curve $\mu^*(\cdot)$ to a problem of the form $(\mathcal{P})$ can be written as the solution of a *continuity inclusion* defined in the sense of [6].

Corollary 1 (Continuity inclusion)**.** *A curve of measures $\mu^*(\cdot)$ is optimal for $(\mathcal{P})$ if and only if*

$$\partial_t \mu^*(t) \in -\mathrm{div}_x(\mathcal{G}(t, \mu^*(t)\mu^*(t)), \tag{11}$$

with $\mathcal{G}(\cdot, \cdot)$ being defined as in (10)*, namely if and only if there exists an $\mathscr{L}^1$-measurable selection $t \in [0, T] \mapsto \boldsymbol{v}^*(t) \in \mathcal{G}(t, \mu^*(t))$ such that the pair $(\mu^*(\cdot), \boldsymbol{v}^*(\cdot))$ satisfies*

$$\partial_t \mu^*(t) + \mathrm{div}_x(\boldsymbol{v}^*(t)\mu^*(t)) = 0$$

in the sense of distributions.

Even though the theory developed in [6] only accounts at the present moment for set-valued velocities that are Lipschitz with respect to the measure variable, the existence of solutions to (11) is guaranteed whenever there exists at least one optimal solution to $(\mathcal{P})$. In this context if we further assume that hypotheses **(R)** hold, then for every compact set $K \subset \mathbb{R}^d$ the restricted set-valued maps $\mathcal{G}_{|K}(\cdot, \cdot)$ are upper-semicontinuous. The importance of this regularity result comes from the fact that set-valued maps with closed graphs are more convenient when Euler constructions of solutions are involved. Also as previously mentioned, the applications $\mathcal{G}_{|K}(\cdot, \cdot)$ are also continuous as soon as they are single-valued, e.g. when the value function is differentiable.

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variations and Free Discontinuity Problems*. Oxford Mathematical Monographs, 2000.

[2] 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, 2008.

[3] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser Basel, 1990.

[4] A. Bensoussan, J. Frehse, and P. Yam. *Mean Field Games and Mean Field Type Control Theory*. Springer Briefs in Mathematics. Springer, 2013.

[5] B. Bonnet. A Pontryagin Maximum Principle in Wasserstein Spaces for Constrained Optimal Control Problems. *ESAIM COCV*, 25(52), 2019.

[6] B. Bonnet and H. Frankowska. Differential Inclusions in Wasserstein Spaces: The Cauchy-Lipschitz Framework. *Journal of Differential Equations*, 271:594–637, 2021.

[7] B. Bonnet and H. Frankowska. Necessary Optimality Conditions for Optimal Control Problems in Wasserstein Spaces. *Applied Mathematics and Optimization*, Published Online, 2021.

[8] B. Bonnet and F. Rossi. The Pontryagin Maximum Principle in the Wasserstein Space. *Calculus of Variations and Partial Differential Equations*, 58:11, 2019.

[9] B. Bonnet and F. Rossi. Intrinsic Lipschitz Regularity of Mean-Field Optimal Controls. *SIAM Journal on Control and Optimization*, 59(3):2011–2046, 2021.

[10] P. Cannarsa and H. Frankowska. Some Characterizations of Optimal Trajectories in Control Theory. *SIAM Journal on Control and Optimization*, 29(6):1322–1347, 1991.

[11] P. Cannarsa and C. Sinestrari. *Semiconcave Functions, Hamilton-Jacobi Equations and Optimal Control*. Progresses in Nonlinear Differential Equations and Their Applications. Birkhaäuser Basel, 2004.

[12] I. Capuzzo Dolcetta and H. Ishii. Approximate Solutions of the Bellman Equation of Deterministic Control. *Applied Mathematics and Optimization*, (11):161–181, 1984.

[13] P. Cardaliaguet. Notes on Mean-Field Games. 2012. https://www.ceremade.dauphine.fr/ cardaliaguet/MFG20130420.pdf.

[14] P Cardaliaguet, F. Delarue, J-M. Lasry, and P.-L. Lions. *The Master Equation and the Convergence Problem in Mean Field Games*. Annals of Mathematics Studies. Princeton Univesrity Press, 2019.

[15] R. Carmona and F. Delarue. *Probabilistic Theory of Mean-Field Games with Applications*. Springer, 2018.

[16] G. Cavagnari, S. Lisini, C. Orrieri, and G. Savaré. Lagrangian, Eulerian and Kantorovich Formulations of Multi-Agent Optimal Control Problems: Equivalence and Gamma-Convergence. *arxiv preprint arXiv:2011.07117*, 2021.

[17] G. Cavagnari, A. Marigonda, and B. Piccoli. Generalized Dynamic Programming Principle and Sparse Mean-Field Control Problems. *Journal of Mathematical Analysis and Applications*, 481(1):123437, 2020.

[18] Y.-P. Choi, J.A. Carrillo, and M. Hauray. The Derivation of Swarming Models: Mean-Field Limit and Wasserstein Distances. In *Collective Dynamics from Bacteria to Crowds, CISM International Centre for Mechanical Sciences*, volume 553, pages 1–46. Springer, 2014.

[19] W.H. Flemming and R.W. Rischel. *Deterministic and Stochastic Optimal Control*, volume 1 of *Stochastic Modelling and Applied Probability*. Springer, 1975.

[20] M. Fornasier, S. Lisini, C. Orrieri, and G. Savaré. Mean-Field Optimal Control as Gamma-Limit of Finite Agent Controls. *European Journal of Applied Mathematics*, 30(6):1153–1186, 2019.

[21] M.Y. Huang, R. Malhamé, and P.E. Caines. Large Population Stochastic Dynamic Games : Closed-Loop McKean-Vlasov Systems and the Nash Certainty Equivalence Principle. *Communications in Information and Systems*, 6(3):221–252, 2006.

[22] J-M. Lasry and P.-L. Lions. Mean Field Games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[23] M. Mesbahi and M. Egerstedt. *Graph Theoretic Multi-Agent Systems*. 2010.

[24] F. Otto. The Geometry of Dissipative Equations : The Porous Medium Equation. *Communications in Partial Differential Equations*, 26:101–174, 2001.

[25] B. Piccoli, F. Rossi, and E. Trélat. Control to Flocking of the Kinetic Cucker-Smale model. *SIAM Journal on Mathematical Analysis*, 47(6):4685–4719, 2015.

[26] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mischenko. *The Mathematical Theory of Optimal Processes*, volume 4. CRC Press, 1987.

[27] L. Rifford. Semiconcave Control-Lyapunov Functions and Stabilizing Feedbacks. *SIAM Journal on Control and Optimization*, 41(3):659–681, 2002.