Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 19–36

ON A MEAN FIELD OPTIMAL CONTROL PROBLEM

JOSÉ A. CARRILLO, EDGARD A. PIMENTEL AND VARDAN K. VOSKANYAN

ABSTRACT: In this paper we consider a mean field optimal control problem with an aggregation-diffusion constraint, where agents interact through a potential, in the presence of a Gaussian noise term. Our analysis focuses on a PDE system coupling a Hamilton-Jacobi and a Fokker-Planck equation, describing the optimal control aspect of the problem and the evolution of the population of agents, respectively. The main contribution of the paper is a result on the existence of solutions for the aforementioned system. We notice this model is in close connection with the theory of mean-field games systems. However, a distinctive feature concerns the nonlocal character of the interaction; it affects the drift term in the Fokker-Planck equation as well as the Hamiltonian of the system, leading to new difficulties to be addressed.

KEYWORDS: Mean field games; Nonlocal Fokker-Planck equations; Nonlocal Hamiltonians; Existence of solutions.
MATH. SUBJECT CLASSIFICATION (2010): 35K10, 35B45, 35A01.

1. Introduction

The multi-agent framework has been used to model swarming or collective behavior in a number of applications. Those include animal herding or flocking [36, 16, 28, 53, 64], cell movement [54, 65, 66], cell adhesion [44, 10, 9], alloy clustering [55], opinion formation [37], robotics [70], among others.

Controlling these collective-behavior models, both at the microscopic and macroscopic levels, has recently become a very popular research direction [17, 41, 18, 14, 40, 67, 7]. Of particular interest is the mean-field limit of a system of N-agents interacting through a potential W in presence of Gaussian

Received November 26, 2019.

JAC was partially supported by the EPSRC grant number EP/P031587/1.

EAP was partially supported by FAPERJ (# E26/200.002/2018), CNPq-Brazil (#433623/2018-7 and #307500/2017-9) and Instituto Serrapilheira.

VKV was partially supported by FCT – Fundação para a Ciência e a Tecnologia, I.P., through projects PTDC/MAT-PUR/28686/2017 and by CMUC – UID/MAT/00324/2019, funded by the Portuguese government through FCT and co-funded by the European Regional Development Fund through Partnership Agreement PT2020.

We would like to acknowledge the Institute Mittag-Leffler, Imperial College London and King Abdullah University of Science and Technology for hosting us and providing with constant help and vivid research environment.

noise. It leads to the control problem of a density ρ solving an aggregation-diffusion equation of the form:

$$\begin{cases} \frac{\partial \rho}{\partial t} = \operatorname{div}\left((\nabla W \star \rho)\rho + F\rho\right) + \Delta\rho & \text{in} \quad \mathbb{T}^d \times (0,\infty) \\ \rho(x,0) = \rho_0(x) & \text{in} \quad \mathbb{T}^d \end{cases} \tag{1}$$

where the control F is chosen to minimize the functional:

$$\inf_{\rho,F} \left\{ E(F) + \int_{\mathbb{T}^d} \phi_T(x)\rho(x,T)dx \right\} \tag{2}$$

with

$$E(F) = \int_0^T\!\!\int_{\mathbb{T}^d} \left[L(x,\rho) + \frac{|F|^2}{2} \right] \rho dx dt. \tag{3}$$

A distinctive feature in (1) regards the drift term. In addition to the optimal control component F, it depends on ρ itself through the interactions driven by W. Perhaps more involving is the fact that such a dependence is nonlocal; this character of the drift term poses important mathematical difficulties to the program developed in this paper.

Typical potentials used in applications are radially symmetric $W(x) = w(|x|)$ in the whole space chosen to be repulsive in the short range and attractive in the long range. Some examples include the Morse or power-law potentials

$$w(r) = -C_A e^{-\frac{r}{l_A}} + C_R e^{-\frac{r}{l_R}}$$

or

$$w(r) = \frac{r^a}{a} - \frac{r^b}{b},$$

for well-prepared parameters $Cl^d < 1$, with $C := C_A/C_R$ and $l := l_A/l_R$, and $a > b > -d$ respectively. See [36, 72, 29] and the references therein. Many other biological applications use finite range potentials, i.e. compactly supported, but with local behaviors near the origin, similar to the Morse or power-laws potentials above. We refer the reader to [33, 10, 9, 27], to name just a few. Notice that these potentials can be used in the periodic setting $x \in \mathbb{T}^d$, see [33, 10, 9].

The optimal control problem in (2)-(3) can be written as

$$\inf_{\rho,F\in(1)} \sup_{\phi} \left\{ E(F) + \int_{\mathbb{T}^d} \phi_T(x)\rho(x,T)dx - \int_0^T\!\!\int_{\mathbb{T}^d} \phi \left[\frac{\partial \rho}{\partial t} - \operatorname{div}\left((\nabla W \star \rho)\rho + F\rho\right) - \Delta\rho \right] \right\},$$

reminiscent of optimal transport problems [15]. By changing the order of the infimum and the supremum, we obtain the dual problem

$$\sup_{\phi} \inf_{(\rho,F)\in(1)} \left\{ \int_0^T \int_{\mathbb{T}^d} \left[L(x,\rho) + \frac{|F|^2}{2} + \frac{\partial \phi}{\partial t} - (\nabla W \star \rho)\cdot \nabla \phi - F \cdot \nabla \phi + \Delta \phi \right] \rho(x,t) dx dt \right\}.$$

At least heuristically, this optimization problem leads to a system of PDEs of the form

$$\begin{cases} -\phi_t + \dfrac{|\nabla \phi|^2}{2} + (\nabla W \star \rho)\cdot \nabla \phi + \nabla W \star (\rho \nabla \phi) - U(x,\rho) = \Delta \phi & \text{in } \mathbb{T}^d \times (0,T) \\ \rho_t = \operatorname{div}\left((\nabla W \star \rho)\rho + \rho \nabla \phi\right) + \Delta \rho & \text{in } \mathbb{T}^d \times (0,T) \\ \phi(x,T) = \phi_T(x),\ \rho(x,0) = \rho_0(x) & \text{in } \mathbb{T}^d, \end{cases} \tag{4}$$

where $U(x,\rho) = L(x,\rho) + \frac{\delta L}{\delta \rho}(x,\rho)\rho$, $F = -\nabla \phi$, and

$$\nabla W \star (\rho \nabla \phi) = \sum_{i=1}^d \frac{\partial W}{\partial x_i} \star \left(\rho \frac{\partial \phi}{\partial x_i} \right).$$

It is worthy noticing that the system in (4) can be regarded as a first order condition associated with (2)-(3). These first-order optimality conditions were obtained both formally and rigorously in [7]. We pose the problem in the d-dimensional torus $\mathbb{T}^d$ to focus on the main difficulties related to the regularity of the Hamilton-Jacobi equation in the system (4).

Remark 1. *Throughout the paper we use the convention to identify functions on the torus $\mathbb{T}^d$ with 1-periodic functions on the cube $Q^d = [-\frac{1}{2}, \frac{1}{2}]^d$. Hence, since the function $|x|$ has periodic boundary data on that cube it can be considered as function on the torus, and therefore radial potentials can be considered in the torus as well. Furthermore, the convolution $\nabla W \star \rho$ is defined by:*

$$\nabla W \star \rho(x) := \int_{Q^d} \nabla W(x-y)\rho(y) dy = \int_{Q^d} \nabla W(y) \rho(x-y) dy,$$

where the values $\nabla W(x-y)$, $\rho(x-y)$ are defined by periodicity.

The contribution of the present paper concerns the existence and regularity of the solutions to (4). Our main result is the following:

Theorem 2 (Existence of solutions)**.** *Let the following assumptions hold:*

(A1) *Regularity of the potential: $W \in W^{2,\infty}(\mathbb{T}^d) = C^{1,1}(\mathbb{T}^d)$.*

(A2) *Regularity of the coupling:* $U : \mathbb{T}^d \times L^2(\mathbb{T}^d) \to \mathbb{R}$ *is uniformly bounded in* x *in* $\mathcal{C}^2$ *norm and continuous in* m *in* L^2 *norm, i.e., there exists a constant* $C > 0$ *such that*

$$\|U(\cdot, m)\|_{\mathcal{C}^2},\ \|\phi_T\|_{\mathcal{C}^2} \le C, \forall m \in L^2(\mathbb{T}^d),$$

for every $m \in L^2(\mathbb{T}^d)$*, and*

$$|U(x, m_1) - U(x, m_2)| \le C\|m_1 - m_2\|_{L^2(\mathbb{T}^d)},$$

for every $x \in \mathbb{T}^d \times (0, T)$ *and for all* m_1*,* $m_2 \in L^2(\mathbb{T}^d)$*.*

(A3) *Initial-terminal boundary conditions:* $\rho_0 \in L^2(\mathbb{T}^d)$ *with* $\int \rho_0 = 1$*,* $\rho_0 \geq 0$*, and* $\phi_T \in \mathcal{C}^1(\mathbb{T}^d)$*.*

Then, there exists a solution (ϕ, ρ) *to the optimal control system* (4) *with* $\rho \in L^\infty(0, T; L^2(\mathbb{T}^d))$ *and* $\phi \in \mathcal{C}^{1,2}(\mathbb{T}^d \times [0, T])$*.*

Remark 3. *Typical potentials with power law behavior satisfying our assumptions are*

$$W(x) \simeq \frac{|x|^a}{a} - \frac{|x|^b}{b}, \text{ with } a, b \geq 2,$$

to be understood in the sense of Remark 1.

The system in (4) relates to the mean-field games (MFG, for short) introduced by Jean-Michel Lasry and Pierre-Louis Lions [59, 60, 61, 62], and, independently, by Minyi Huang, Roland P. Malhamé, and Peter E. Caines [56, 57]. Indeed, it couples a Hamilton-Jacobi equation describing an optimization problem, with the Fokker-Planck equation accounting for the evolution of the population. An interesting feature of (4) regards the lack of an adjoint structure, which yields further difficulties from the regularity viewpoint.

In the recent years, the MFG theory developed in a number of directions. The existence and uniqueness of solutions is the topic of [22, 20, 23, 24, 21, 48, 49, 47, 34, 35, 8, 30], whereas the study of numerical methods is the object of [1, 6, 4, 5], to name just a few. In [25, 12, 13, 26, 31, 43] the authors examine the master equation. Applications of the MFG framework to social sciences can be found in [3, 58, 42, 2]. We also refer the reader to the monographs [19, 11, 50] and the lists of references therein.

An important toy-model in the context of time-dependent mean-field games has the form

$$\begin{cases} -u_t - \nu\Delta u + H(D_x u, x, \theta) = 0 & \text{in} \quad \mathbb{T}^d \times (0,T) \\ \theta_t - \nu\Delta\theta - \operatorname{div}(D_p H(D_x u, x, \theta)\theta) = 0 & \text{in} \quad \mathbb{T}^d \times (0,T) \\ u(x,T) = \psi(x, \theta(T)), \quad \theta(x,0) = \theta_0 & \text{in} \quad \mathbb{T}^d. \end{cases}$$

This class of systems consists of a Hamilton-Jacobi equation coupled with a Fokker-Plank equation; the latter is the formal adjoint, in the L^2-sense, of the linearization of the former one. This (adjoint) structure was exploited in proving the existence of classical solutions [46, 51, 52, 68] through the so-called non-linear adjoint methods, introduced by L.C. Evans [38]. The system in (4) does not present this adjoint structure because of the additional term $\nabla W \star (\rho \nabla \phi)$, which substantially complicates the arguments. However, the non-linear adjoint method is still useful in proving the Lipschitz continuity of a solution ϕ to the Hamilton-Jacobi equation in (4).

The nonlocal interaction among agents affects the PDE system accounting for the optimality conditions of the model. Here, the drift term in the Fokker-Planck equation becomes nonlocal. A similar phenomenon takes place in the Hamilton-Jacobi counterpart of the system: the Hamiltonian becomes nonlocal *with respect to the gradient of the solutions.*

In this context, at least two genuine difficulties appear. First, the existence and uniqueness of the solutions to the Fokker-Planck equation becomes a nontrivial matter. Also the (uniform) compactness of the solutions to the Hamilton-Jacobi equation does not follow from the usual techniques. At first, one could resort to semiconvexity/semiconcavity properties to bypass this issue - however, we notice the equation falls short in preserving those conditions.

To circumvent the first question, we resort to an argument in [69], combined with a fixed-point strategy performed in an appropriate space of measures. As for the compactness for the solutions to the Hamilton-Jacobi, we reason through a nonlocal L^p-regularity theory, together with (compact) embedding results. We believe our techniques are flexible enough to produce information on a larger class of problems.

The remainder of this article is structured as follows: in Section 2 we present a brief outline of the proof of Theorem 2. Section 3 establishes the well-posedness for the Fokker-Planck equation in (4), whereas in Section 4 we produce a number of a priori estimates for the solutions of the Hamilton-Jacobi. Finally, a section reporting the proof of Theorem 2 closes the paper.

2. Set up and outline of the proof

In what follows, let us introduce the general lines along which we establish Theorem 2. Let us start with the definition of (weak) solution used in the paper.

Definition 4 (Weak solution)**.** *A pair* (ϕ, ρ) *is a solution to* (4) *if*

(1) $\phi \in \mathcal{C}(\mathbb{T}^d \times (0,T))$ *satisfies the first equation in* (4) *in the viscosity sense;*
(2) $\rho \in L^\infty(0,T; L^2(\mathbb{T}^d))$ *satisfies the second equation in* (4) *in the sense of distributions.*

Throughout the paper, we denote by $\mathcal{C}^{1,2}(\mathbb{T}^d \times [0,T])$ the space of functions defined over $\mathbb{T}^d \times [0,T]$ that are of class $\mathcal{C}^1$ with respect to time and class $\mathcal{C}^2$ with respect to space. Similarly, for any $\alpha \in (0,1)$, $\mathcal{C}^{\alpha,1+\alpha}(\mathbb{T}^d \times [0,T])$ stands for the functions that are $\alpha-$Hőlder continuous in time, $\mathcal{C}^1$ with respect to space with $\alpha-$Hőlder continuous derivative.

We establish Theorem 2 by using a fixed-point argument. The argument starts by taking an element ϕ in $\mathcal{C}^{0,1}(\mathbb{T}^d \times [0,T])$. To such a function $\phi \in \mathcal{C}^{0,1}(\mathbb{T}^d \times [0,T])$, we can assign the unique weak solution $\rho \in L^\infty(0,T; L^2(\mathbb{T}^d))$ to the Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left((\nabla W \star \rho)\rho + \nabla \phi\, \rho\right) + \Delta \rho, \tag{5}$$

equipped with initial condition $\rho(x,0) = \rho_0(x)$ in the d-dimensional torus.

Then, standard results in the literature ensure the existence of a unique solution, $\Phi \in \mathcal{C}^{1,2}(\mathbb{T}^d \times [0,T])$ to

$$\begin{cases} -\Phi_t + \frac{|\nabla \Phi|^2}{2} + (\nabla W \star \rho) \cdot \nabla \Phi + \nabla W \star (\rho \nabla \phi) = \Delta \Phi + U(x,\rho) & \text{in } \mathbb{T}^d \times (0,T) \\ \Phi(x,T) = \phi_T(x) & \text{in } \mathbb{T}^d \times (0,T); \end{cases} \tag{6}$$

see, for instance, [19, Section 3.2]. This procedure induces a mapping $\phi \to \rho \to \Phi$, which we denote

$$\mathcal{F} \,:\, \cap_{\alpha\in(0,1)} \mathcal{C}^{\alpha,1+\alpha}(\mathbb{T}^d \times [0,T]) \longrightarrow \cap_{\alpha\in(0,1)} \mathcal{C}^{\alpha,1+\alpha}(\mathbb{T}^d \times [0,T]),$$

and define as

$$\mathcal{F}(\phi) := \Phi.$$

To prove the existence of a solution (ϕ, ρ) to problem (4) is tantamount to verify that the mapping $\mathcal{F}$ has a fixed point. To that end, we start with the definition of an appropriate subset $\mathcal{K}_\alpha \subset \mathcal{C}^{\alpha,1+\alpha}(\mathbb{T}^d \times [0,T])$. For some constants A, B, and $C > 0$ to be determined further, we define

$$\mathcal{K}_\alpha = \left\{ \phi \in \mathcal{C}^{\alpha,1+\alpha}(\mathbb{T}^d \times [0,T]) \mid \|\Phi\|_{\mathcal{C}^{\alpha,1+\alpha}} \leq C \text{ and } \|\Phi(\cdot,t)\|_{Lip} \leq Ae^{B(T-t)} \right\}.$$

Remark 5. *We remark here that the set $\mathcal{K}_\alpha$ is nonempty, since for any smooth function ϕ appropriately rescaled $\phi(\lambda x, \lambda t)$ is in $\mathcal{K}_\alpha$. Furthermore, the set of smooth function $\mathcal{C}^\infty(\mathbb{T}^d \times [0,T])$ is dense in $\mathcal{K}_\alpha$, this can be proved by a simple mollification argument.*

The first step towards the proof of Theorem 2 is a result on the well-posedness for the Fokker-Planck equation in (4).

Proposition 6 (Well-posedness for the Fokker-Planck equation)**.** *Suppose the assumptions* **(A1)-(A3)** *hold and $\phi \in \mathcal{K}_\alpha$ is fixed. Then, there exists a unique solution to* (8)*.*

Once the existence and uniqueness for the Fokker-Planck equation (8) is assured, we turn our attention to the compactness of the solutions to the Hamilton-Jacobi equation (6). We proceed with the following proposition.

Proposition 7. *Let the assumptions* **(A1)-(A3)** *hold. Then, there exists a choice of constants A, B, C depending only on T, ϕ_T, ρ_0, $\|W\|_{W^{2,\infty}(\mathbb{T}^d)}$ and U, such that for any $\phi \in \mathcal{C}(\mathbb{T}^d \times [0,T])$, satisfying*

$$|\phi(x,t) - \phi(y,t)| \leq Ae^{B(T-t)}|x-y|,$$

we have

$$\|\Phi(\cdot,t)\|_{Lip} \leq Ae^{B(T-t)}, \quad \|\Phi\|_\infty \leq C,$$

for every solution Φ to (6)*.*

From now on, the set $\mathcal{K}_\alpha$ is defined with the constants A and B given by Proposition 7. We now turn to fix C. The L^p-regularity theory for the Hamilton-Jacobi equation is pivotal in proving uniform bounds for the solutions in appropriate Hölder spaces. This is the content of the next proposition.

Proposition 8. *Let Φ be a solution to* (6) *corresponding to $\phi \in \mathcal{K}_\alpha$. Suppose the assumptions* **(A1)-(A3)** *hold true. Then, $\Phi \in \mathcal{C}^{\alpha,1+\alpha}(\mathbb{T}^d \times [0,T])$ for* *some $\alpha \in (0,1)$. Moreover, there exists $C>0$ depending solely on T, ϕ_T, ρ_0, $\|W\|_{\mathcal{C}^2(\mathbb{T}^d)}$ and U, such that*

$$\|\Phi\|_{\mathcal{C}^{\alpha,1+\alpha}(\mathbb{T}^d\times[0,T])} \leq C.$$

From now on, the set $\mathcal{K}_\alpha$ is fixed with the definition of C given in the previous proposition. We detail next the proof of Proposition 6.

3. Well-posedness for the Fokker-Planck equation

In this section, we prove Proposition 6 on the existence and uniqueness of solutions to

$$\begin{cases} \rho_t \,=\, \operatorname{div}\left[\left(\left(\nabla W \star \rho\right) \,+\, \nabla\phi\right)\rho\right] \,+\, \Delta\rho & \text{in} \quad \mathbb{T}^d \times [0,T] \\ \rho(x,0) \,=\, \rho_0(x) & \text{in} \quad \mathbb{T}^d, \end{cases} \tag{7}$$

provided W and ϕ are well-prepared. We reason through a fixed-point argument; the next proposition details the functional space we work. Before proceeding let us recall the 1-Wasserstein metric on the space of probability measures $\mathcal{P}(\mathbb{T}^d)$:

Definition 9 (Wasserstein metric)**.** *The* 1*-Wasserstein metric between probability measures $\mu,\nu \in \mathcal{P}(\mathbb{T}^d)$ is given by*

$$d_1(\mu,\,\nu) := \inf_{\pi\in\Gamma(\mu,\,\nu)} \int_{\mathbb{T}^d\times\mathbb{T}^d} |x-y| d\pi(x,y),$$

where $\Gamma(\mu,\nu)$ is the collection of all measures on $\mathbb{T}^d\times\mathbb{T}^d$ with marginals μ and ν.

Recall that the d_1 distance in bounded sets is the same as the bounded Lipschitz distance defined via duality with respect to Lipschitz functions, see [71].

Proposition 10. *Define the function*

$$N(t) := \int_{\mathbb{T}^d} \left|\left(\nabla W \star \mu\right) + \nabla\phi\right|^2 d\mu(x,t).$$

Take $C>0$ to be determined later. Set

$$\mathcal{M} := \left\{\mu \in \mathcal{C}([0,T],\mathcal{P}(\mathbb{T}^d)) : d_1(\mu(s),\mu(t)) \leq C|t-s|^{\frac{1}{2}}, \|N\|_{L^\infty([0,T])} \leq C\right\}.$$

Then, $\mathcal{M}$ is a convex and compact subset of $\mathcal{C}([0,T],\mathcal{P}(\mathbb{T}^d))$.

The proof of the Proposition 10 follows along the same lines as in [19, Lemma 5.7], except for minor modifications, and is omitted here.

In what follows, we define a mapping

$$\mathcal{T} : \mathcal{C}([0,T], \mathcal{P}(\mathbb{T}^d)) \to \mathcal{C}([0,T], \mathcal{P}(\mathbb{T}^d)).$$

Let $\phi \in \mathcal{K}_\alpha$ be fixed and take $\rho^0 \in \mathcal{M}$. Solve

$$\rho_t = \operatorname{div}\left[\left(\left(\nabla W \star \rho^0\right) + \nabla\phi\right)\rho\right] + \Delta\rho \quad \text{in} \quad \mathbb{T}^d \times [0,T] \tag{8}$$

equipped with initial condition $\rho(x,0) = \rho_0$. By [19, Lemma 3.3], we know that there exists a solution ρ^1 to (8). Notice, that

$$\left|\left(\nabla W \star \rho^0\right) + \nabla\phi\right|^2 \rho^1 \in L^1(\mathbb{T}^d \times (0,T)).$$

In fact,

$$\int_0^T \int_{\mathbb{T}^d} \left|\left(\nabla W \star \rho^0\right) + \nabla\phi\right|^2 d\rho^1(x,t) \le C(W,\phi,T).$$

Therefore, it follows from [69, Theorem 1.1] that (8) has at most one solution, which is precisely ρ^1. As a consequence, we define

$$\mathcal{T}(\rho^0) := \rho^1. \tag{9}$$

Next we examine properties of the mapping $\mathcal{T}$ which are related to the fixed-point arguments presented further in this section.

Proposition 11. *Let $\mathcal{T} : \mathcal{C}([0,T], \mathcal{P}(\mathbb{T}^d)) \to \mathcal{C}([0,T], \mathcal{P}(\mathbb{T}^d))$ be defined as in* (9)*. Then, $\mathcal{T}$ maps $\mathcal{M}$ into itself. In addition, $\mathcal{T}$ is a continuous mapping when restricted to the set $\mathcal{M}$.*

Proof: First, we verify the inclusion $\mathcal{T}(\mathcal{M}) \subset \mathcal{M}$. Let $\mu \in \mathcal{M}$ and $\sigma := \mathcal{T}(\mu)$. We have

$$d_1(\sigma(t), \sigma(s)) \le \mathbb{E}(X_t, X_s),$$

where

$$\begin{cases} dX_t = [(\nabla W \star \mu) + \nabla\phi]\, dt + \sqrt{2} I dW_t \\ X(0) = X_0, \end{cases} \tag{10}$$

and W_t is a d-dimensional Brownian motion. Because the vector field $(\nabla W \star \mu) + \nabla\phi$ is Hölder continuous with respect to space, there exists a unique solution to (10); see, for example, [32]. Hence, using the Lipschitz continuity of W and ϕ, we obtain

$$d_1(\sigma(t), \sigma(s)) \le \mathbb{E}\left[\int_s^t |(\nabla W \star \mu) + \nabla\phi|\, d\tau + \sqrt{2}\, |W_t - W_s|\right]$$

$$\leq C_1 |t - s|^{\frac{1}{2}},$$

with $C_1 = C_1(W, \phi, T)$. Moreover, once more from the Lipschitz continuity of W and ϕ, we infer

$$\int_0^T \int_{\mathbb{T}^d} |\nabla W \star \sigma + \nabla\phi|^2 \, d\sigma(x,t) \leq C \int_0^T \int_{\mathbb{T}^d} \sigma(x,t) dx dt \leq C_2,$$

where $C_2 = C_2(W, \phi, T)$. Finally, by choosing the constant $C > 0$ in the definition of $\mathcal{M}$ as $C := \max\{C_1, C_2\}$, we have $\mathcal{T}(\mathcal{M}) \subset \mathcal{M}$.

It remains to prove that $\mathcal{T}$ is continuous. However, it follows from standard results in stability theory for the Fokker-Planck equation in the presence of Hölder-continuous drift terms [69]. ■

Remark 12. *We notice that the constant $C > 0$ in the definition of $\mathcal{M}$ depends solely on W and ϕ.*

We close this section with the proof of Proposition 6.

Proof of Proposition 6: We start by noticing that $\mathcal{T}$ has a fixed point. In fact, Proposition 10 ensures that $\mathcal{M}$ is a convex and compact subset of $\mathcal{C}([0,T], \mathcal{P}_1)$. From Proposition 11 we infer that $\mathcal{T}(\mathcal{M}) \subset \mathcal{M}$ and that this mapping is continuous. Therefore, the Schauder's Fixed Point Theorem yields the existence of a fixed point for $\mathcal{T}$. The uniqueness is ensured by a straightforward application of [69, Theorem 1.1] and the proposition is established. ■

Proposition 13 (Propagation of L^2 regularity)**.** *Let the assumptions of Proposition 6 hold, then $\rho \in L^\infty(0,T; L^2(\mathbb{T}^d))$.*

Proof: We take a standard radially symmetric mollifier $\varphi_\varepsilon \in C^\infty(\mathbb{T}^d)$ and a smooth function $\psi \in C_c^\infty([0,T) \times \mathbb{T}^d)$, and consider $\varphi_\varepsilon \star \psi(t, \cdot)$ as test function for the weak solution ρ:

$$\int_0^T \int_{\mathbb{T}^d} (\varphi_\varepsilon \star \psi_t - \Delta\varphi_\varepsilon \star \psi + bD\varphi_\varepsilon \star \psi)\rho \, dx dt = \int_{\mathbb{T}^d} \rho_0 \varphi_\varepsilon \star \psi(0) dx,$$

where $b(x,t) := \nabla W \star \rho + \nabla\phi$. Using the identity

$$\int_{\mathbb{T}^d} (\varphi_\varepsilon \star g)(x) \, f(x) \, dx = \int_{\mathbb{T}^d} g(x) \, (\varphi_\varepsilon \star f)(x) \, dx,$$

we get

$$\int_0^T \int_{\mathbb{T}^d} (\psi_t - \Delta\psi)(\varphi_\varepsilon \star \rho) + D\psi(\varphi_\varepsilon \star (b\rho))\, dxdt = \int_{\mathbb{T}^d} \varphi_\varepsilon \star \psi\, \rho_0 dx\,,$$

implying $\varphi_\varepsilon \star \rho$ is weak solution to

$$\begin{cases} (\varphi_\varepsilon \star \rho)_t = \operatorname{div}[\varphi_\varepsilon \star (b\rho))] + \Delta(\varphi_\varepsilon \star \rho) & \text{in} \quad \mathbb{T}^d \times [0,T] \\ \rho(x,0) = \varphi_\varepsilon \star \rho_0(x) & \text{in} \quad \mathbb{T}^d. \end{cases}$$

Since $t \mapsto \rho\, dx$ is continuous with respect to d_1 and ∇W is in $C^1(\mathbb{T}^d)$, then $\nabla W * \rho$ is continuous in time and is in $C^1(\mathbb{T}^d)$. Since $\nabla\phi \in C^\alpha(\mathbb{T}^d)$ the vector field b is continuous in time and is in $C^\alpha(\mathbb{T}^d)$. This implies that the first term of the right hand side of the above equation, $\operatorname{div}[\varphi_\varepsilon \star (b\rho)]$, is a bounded function, hence $\rho_\varepsilon := \varphi_\varepsilon \star \rho$ is a classical solution due to classical regularity of the heat equation.

Multiplying it by ρ_ε and integrating in space we get by integration by parts

$$\begin{aligned} \frac{d}{dt}\|\rho_\varepsilon(\cdot,t)\|^2_{L^2(\mathbb{T}^d)} =& 2\int_{\mathbb{T}^d} \left[-|\nabla\rho_\varepsilon|^2 - \varphi_\varepsilon \star (b\rho)\cdot\nabla\rho_\varepsilon\right] dx \\ \le& \frac{1}{2}\int_{\mathbb{T}^d} (\varphi_\varepsilon \star (b\rho))^2 dx \le \frac{1}{2}\|b\|_\infty \int_{\mathbb{T}^d} (\varphi_\varepsilon \star \rho)^2 dx. \end{aligned}$$

Since $\|b\|_\infty \le \|\nabla W\|_\infty + \|\nabla\phi\|_\infty < +\infty$, Gronwall's inequality implies

$$\|\rho_\varepsilon(\cdot,t)\|^2_{L^2(\mathbb{T}^d)} \le e^{\frac{1}{2}Ct}\|\rho_\varepsilon(\cdot,0)\|^2_{L^2(\mathbb{T}^d)}.$$

By taking the limit $\varepsilon \to 0$ we finally obtain $\|\rho\|_{L^\infty([0,T],L^2(\mathbb{T}^d))} < +\infty$. ∎

4. Lipschitz continuity and Hölder regularity

In this section we detail the proofs of Propositions 7 and 8. Under the assumption $W \in W^{2,\infty}$ the coefficients of the equation (6) are C^1 in the x variable. As a consequence, we have that Φ is C^2 in the x variable. This fact can be verified through a Hopf-Cole transformation; see e.g. [19, Section 3.2].

We resort to the nonlinear adjoint method, introduced by L. C. Evans in [39]. For a detailed discussion on the application of this method to the theory of mean field games, we refer the reader to [46, 50].

Start by fixing $(x_0, t_0) \in \mathbb{T}^d \times [0, T]$ and consider two adjoint variables η, ζ given by the equations

$$\textit{Optimal Flow:} \quad \begin{cases} \eta_t = \operatorname{div}\left((\nabla W \star \rho)\eta + \nabla\Phi\eta\right) + \Delta\eta & \text{in } \mathbb{T}^d \times (0,T) \\ \eta(\cdot, t_0) = \delta_{x_0} & \text{in } \mathbb{T}^d \end{cases}$$

and

$$\textit{Zero velocity Flow:} \quad \begin{cases} \zeta_t = \operatorname{div}\left((\nabla W \star \rho)\zeta\right) + \Delta\zeta & \text{in } \quad \mathbb{T}^d \times (0,T) \\ \zeta(\cdot, t_0) = \delta_{x_0} & \text{in } \quad \mathbb{T}^d. \end{cases} \tag{11}$$

The first one represents the evolution of the distribution of the optimal trajectories in the stochastic optimal control problem associated with the first equation in (4). The second one accounts for the evolution of the distribution of the trajectory corresponding to the zero velocity. Note that

$$\int_{\mathbb{T}^d} \eta(x,t)dx = \int_{\mathbb{T}^d} \zeta(x,t)dx = 1$$

and η, $\zeta \geq 0$.

Note that in general η and ζ are measure valued. However since the vector fields $(\nabla W \star \rho) + \nabla\Phi$ and $(\nabla W \star \rho)$ are C^1 in space, for times $t > t_0$, η and ζ can be viewed as functions. Alternatively, to make the rest of the arguments rigorous one can approximate the delta distributions δ_{x_0} by smooth approximations of unity and pass to the limit at the end. Below we will avoid these complications for simplicity.

Proof of Proposition 7: We split the proof of the proposition in several steps.

Step 1. ***An estimates on*** $\int_{t_0}^T \int_{\mathbb{T}^d} \frac{|\nabla\Phi|^2}{2}\eta dxdt$***:***

We integrate (6) against the measure ζ and use the equation (11) in the distributional sense to obtain

$$\Phi(x_0,t_0) - \int_{\mathbb{T}^d} \phi_T\zeta_T dx + \int_{t_0}^T\int_{\mathbb{T}^d} \frac{|\nabla\Phi|^2}{2}\zeta dxdt + \int_{t_0}^T\int_{\mathbb{T}^d} \nabla W \star (\rho\nabla\phi)\zeta dxdt = \int_{t_0}^T\int_{\mathbb{T}^d} U\zeta dxdt.$$

Subsequently, we infer that

$$\Phi(x_0,t_0) \leq \|\nabla W\|_\infty \int_{t_0}^T \|\nabla\phi(\cdot,t)\|_\infty dt + (T - t_0)\|U\|_\infty + \|\phi_T\|_\infty. \tag{12}$$

Similarly, multiplying (6) by η and using (4) in the distributional sense, we have

$$\Phi(x_0,t_0) - \int_{\mathbb{T}^d} \phi_T \eta_T dx - \int_{t_0}^T\int_{\mathbb{T}^d} \frac{|\nabla\Phi|^2}{2}\eta dxdt + \int_{t_0}^T\int_{\mathbb{T}^d} \nabla W \star (\rho\nabla\phi)\eta dxdt = \int_{t_0}^T\int_{\mathbb{T}^d} U\eta dxdt.$$

Thus, we get

$$\int_{t_0}^T\int_{\mathbb{T}^d} \frac{|\nabla\Phi|^2}{2}\eta dxdt \leq \Phi(x_0,t_0) + \|\phi_T\|_\infty + \|\nabla W\|_\infty \int_{t_0}^T \|\nabla\phi(\cdot,t)\|_\infty dt + (T-t_0)\|U\|_\infty. \quad (13)$$

From (12) and (13) we deduce

$$\int_{t_0}^T\int_{\mathbb{T}^d} \frac{|\nabla\Phi|^2}{2}\eta dxdt \leq 2\|\phi_T\|_\infty + 2\|\nabla W\|_\infty \int_{t_0}^T \|\nabla\phi(\cdot,t)\|_\infty dt + 2(T-t_0)\|U\|_\infty. \quad (14)$$

Step 2. ***Lipschitz continuity in space:***

Fix $\xi \in \mathbb{R}^d$ and let $u = \nabla\Phi\cdot\xi := \Phi_\xi$, then differentiating the equation for Φ, we get

$$-u_t + \nabla\Phi\cdot\nabla u + (\nabla W_\xi \star \rho)\cdot\nabla\Phi + (\nabla W\star\rho)\cdot\nabla u + \nabla W_\xi\star(\rho\nabla\phi) = \Delta u + U_\xi.$$

As above, integrating the equation for Φ against ζ and using integration by parts, we obtain

$$u(x_0,t_0) - \int_{\mathbb{T}^d} \nabla\phi_T\eta_T dx + \int_{t_0}^T\int_{\mathbb{T}^d} (\nabla W_\xi\star\rho)\cdot\nabla\Phi\eta dxdt + \int_{t_0}^T\int_{\mathbb{T}^d} \nabla W_\xi\star(\rho\nabla\phi)\eta dxdt = \int_{t_0}^T\int_{\mathbb{T}^d} U_\xi \eta dxdt.$$

Hence

$$|u(x_0,t_0)| \leq \|\nabla\phi_T\|_\infty + \|\nabla^2 W\|_\infty \int_{t_0}^T \left(\int_{\mathbb{T}^d} |\nabla\Phi|\eta dx + \|\nabla\phi(\cdot,t)\|_\infty\right) dt + (T-t_0)\|\nabla U\|_\infty + C.$$

Since

$$\int_{t_0}^T\int_{\mathbb{T}^d} |\nabla\Phi|\eta dxdt \leq \frac{1}{2}\int_{t_0}^T\int_{\mathbb{T}^d} [1+|\nabla\Phi|^2]\eta dxdt,$$

we use (14) to produce

$$|u(x_0,t_0)| \leq A + B\int_{t_0}^T \|\nabla\phi(\cdot,s)\|_\infty ds$$

with

$$A = \|\nabla\phi_T\|_\infty + \|\nabla^2 W\|_\infty\left(\frac{1}{2}T + 2\|\phi_T\|_\infty + 2T\|U\|_\infty\right) + T\|\nabla U\|_\infty + C$$

and $B = \|\nabla^2 W\|_\infty (2\|\nabla W\|_\infty + 1)$. Since x_0 and t_0 were arbitrary, we have proved

$$\|\nabla\Phi(\cdot,t)\|_\infty \le A + B\int_t^T \|\nabla\phi(\cdot,s)\|_\infty ds, \quad \forall t \in [0,T].$$

Now it is easy to check that

$$\|\nabla\phi(\cdot,t)\|_\infty \le Ae^{B(T-t)} \implies \|\nabla\Phi(\cdot,t)\|_\infty \le Ae^{B(T-t)}.$$

This together with (12) and (13) yields $\|\Phi\|_\infty \le C_1$, where C_1 depends on $\|W\|_{\mathcal{C}^2}$,$\|U\|_{\mathcal{C}^2}$,$\|\phi_T\|_{\mathcal{C}^1}$, A and B. ■

Next we present the proof of Proposition 8.

Proof of Proposition 8: We observe that (6) can be written as

$$-\Phi_t - \Delta\Phi = f(x,t) \in L^\infty(\mathbb{T}^d \times [0,T]),$$

with a bound on $\|f\|_{L^\infty}$ depending only on $\|W\|_{\mathcal{C}^2}$,$\|U\|_{\mathcal{C}^2}$,$\|\phi_T\|_{\mathcal{C}^1}$, A and B. Therefore, by standard elliptic regularity theory, we have

$$\Phi_t,\ \nabla^2\Phi \in L^p(\mathbb{T}^d \times [0,T]),$$

for every $p > 1$, see for example [63, Theorem D.1. and Remark D.1.4]. Morrey's Embedding Theorem implies the result. ■

5. Proof of Theorem 2

In this section we present the proof of Theorem 2. In fact, we show that $\mathcal{F}$ restricted to $\mathcal{K}_\alpha$ is a continuous mapping. This fact builds upon Propositions 6-8 to cover the existence of a fixed point of $\mathcal{F}$ under the scope of the Schauder's Fixed-Point Theorem [45, Theorem 11.1].

Theorem (Schauder's Fixed-Point Theorem)**.** *Let K be a compact convex set in a Banach space $\mathbb{B}$ and let T be a continuous mapping of K into itself. Then T has a fixed point, that is, $Tx = x$ for some $x \in \mathbb{B}$.*

Proof of Theorem 2: Take $0 < \beta < \alpha < 1$, it is obvious that $\mathcal{K}_\alpha$ is a closed and convex set which is compact in $\mathcal{C}^{\beta,1+\beta}(\mathbb{T}^d \times [0,T])$. Its compactness follows from Propositions 7 and 8, together with the Arzelà-Ascoli Theorem. The fact that $\mathcal{F}(\mathcal{K}_\alpha) \subset \mathcal{K}_\alpha$ follows from Proposition 7.

It remains to prove that the map $\mathcal{F}\big|_{\mathcal{K}_\alpha}$ is continuous in $\mathcal{C}^{\beta,1+\beta}(\mathbb{T}^d \times [0,T])$. For that we need to show that for any sequence $(\phi_n)_{n\in\mathbb{N}} \subset \mathcal{K}_\alpha$ with $\phi_n \to$

ϕ in $\mathcal{C}^{\beta,1+\beta}(\mathbb{T}^d \times [0,T])$, we have $\Phi_n := \mathcal{F}(\phi_n) \to \mathcal{F}(\phi)$, in $\mathcal{C}^{\beta,1+\beta}(\mathbb{T}^d \times [0,T])$. We split the proof in two steps.

Step 1 *Stability property of the Fokker-Planck equation:* Let $(\rho_n)_{n\in\mathbb{N}}$ be such that

$$\frac{\partial \rho_n}{\partial t} = \operatorname{div}\left((\nabla W \star \rho_n)\rho_n + \nabla\phi_n\rho_n\right) + \Delta\rho_n \quad \text{in} \quad \mathbb{T}^d \times (0,T),$$

under the initial condition $\rho_n(x,0) = \rho_0(x)$. We will prove that $\rho_n \to \rho$ in $L^\infty(0,T;L^2(\mathbb{T}^d))$. We have

$$\|\phi_n - \phi\|_{L^\infty(\mathbb{T}^d\times[0,T])} \to 0, \ \|\nabla\phi_n\|_{L^\infty(\mathbb{T}^d\times[0,T])} \leq Ae^{BT}, \\ \text{and } \|\nabla\phi_n(\cdot,t)\|_{\mathcal{C}^\alpha(\mathbb{T}^d)} \leq C;$$

hence, we can conclude that $\nabla\phi_n(\cdot,t) \to \nabla\phi(\cdot,t)$, uniformly in $\mathbb{T}^d$. In particular, by the Lebesgue's Dominated Convergence Theorem

$$\int_0^T \int_{\mathbb{T}^d} \rho^2(\nabla\phi_n - \nabla\phi)^2 dxdt \to 0.$$

Now, denote $m^n = \rho_n - \rho$; then

$$m^n_t - \operatorname{div}((\nabla W \star \rho)m^n) - \operatorname{div}((\nabla W \star m^n)\rho) \\ - \operatorname{div}(\nabla\phi_n m^n) - \operatorname{div}\left(\rho(\nabla\phi_n - \nabla\phi)\right) = \Delta m^n,$$

under the homogeneous initial condition $m_n(x,0) = 0$. Proceeding analogously to the previous section, we get

$$\frac{d}{dt}\|m^n(\cdot,t)\|^2_{L^2(\mathbb{T}^d)} = 2\int_{\mathbb{T}^d}\left[-|\nabla m^n|^2 - m^n\nabla m^n\cdot\nabla\phi_n - \rho\nabla m^n\cdot(\nabla\phi_n - \nabla\phi)\right]dx \\ + 2\int_{\mathbb{T}^d}\left[(\nabla W\star\rho)m^n\nabla m^n + (\nabla W \star m^n)\nabla m^n\rho\right]dx\,.$$

Therefore, using the inequality $a\,b \leq \varepsilon a^2 + \frac{b^2}{4\varepsilon}$, with ε small enough and $a = \nabla m^n$, on each term above starting from the second one , we obtain

$$\frac{d}{dt}\|m^n(\cdot,t)\|^2_{L^2(\mathbb{T}^d)} \leq C(\|\nabla\phi_n\|^2_\infty + \|\nabla W\|^2_{L^2}\|\rho\|^2_{L^2})\|m^n(\cdot,t)\|^2_{L^2(\mathbb{T}^d)} \\ + C\|\rho(\nabla\phi_n - \nabla\phi)\|^2_{L^2(\mathbb{T}^d)}.$$

Thus, by Gronwall's inequality, we conclude

$$\|m^n(\cdot,t)\|^2_{L^2(\mathbb{T}^d)} \leq \|\rho(\nabla\phi_n - \nabla\phi)\|^2_{L^2([0,t]\times\mathbb{T}^d)} e^{C(T-t)} \to 0.$$

As a result,

$$\rho_n \to \rho \quad \text{in} \quad L^\infty([0,T], L^2(\mathbb{T}^d)),$$

as desired.

Step 2 *Stability and Uniqueness of viscosity solutions for the Hamilton-Jacobi equation:* We now examine the convergence $\Phi_n \to \Phi$. We first observe that

$$\|(\nabla W \star \rho_n) - (\nabla W \star \rho)\|_\infty \leq \|\nabla W\|_{L^2} \|\rho_n - \rho\|_{L^2} \to 0$$

and

$$\|U(\cdot, \rho_n) - U(\cdot, \rho)\|_\infty \leq C\|\rho_n - \rho\|_{L^2} \to 0.$$

Since $\nabla \phi_n(\cdot, t) \to \nabla \phi(\cdot, t)$ uniformly due to our functional setting, then

$$\nabla W \star \rho \nabla \phi_n \longrightarrow \nabla W \star \rho \nabla \phi$$

uniformly on $[0, T] \times \mathbb{T}^d$. Hence, the nonlocal terms

$$\begin{aligned} |\nabla W \star (\rho_n \nabla \phi_n) - \nabla W \star (\rho \nabla \phi)| \leq & \|\nabla \phi_n\|_\infty \|\nabla W\|_{L^2} \|\rho_n - \rho\|_{L^2} \\ & + |\nabla W \star \rho(\nabla \phi_n - \nabla \phi)| \end{aligned}$$

converge to zero uniformly. By the compactness of $\mathcal{K}_\alpha$ there is a converging subsequence $\Phi_{n_k} \to \tilde{\Phi}$, in $\mathcal{C}^{\beta,1+\beta}(\mathbb{T}^d \times [0, T])$, in particular $\nabla \Phi_{n_k} \to \nabla \tilde{\Phi}$ uniformly. Then $\tilde{\Phi}$ is a weak solution to

$$-\tilde{\Phi}_t + \frac{|\nabla \tilde{\Phi}|^2}{2} + (\nabla W \star \rho) \cdot \nabla \tilde{\Phi} + \nabla W \star (\rho \nabla \phi) = \Delta \tilde{\Phi} + U(x, \rho), \ \tilde{\Phi}(x, T) = \phi_T(x).$$

By parabolic regularity $\tilde{\Phi}$ is also a classical solution. Thus Φ and $\tilde{\Phi}$ solve the same equation, to prove that $\Phi = \tilde{\Phi}$, we introduce two adjoint variables $\eta, \tilde{\eta}$ which solve the following equations respectively

$$\begin{cases} \frac{\partial \eta}{\partial t} = \mathrm{div}\left((\nabla W \star \rho)\eta + \nabla \Phi \eta\right) + \Delta \eta, \\ \eta(\cdot, 0) = \delta_{x_0}, \end{cases}$$

and

$$\begin{cases} \frac{\partial \tilde{\eta}}{\partial t} = \mathrm{div}\left((\nabla W \star \rho)\tilde{\eta} + \nabla \tilde{\Phi} \tilde{\eta}\right) + \Delta \tilde{\eta}, \\ \tilde{\eta}(\cdot, 0) = \delta_{x_0}. \end{cases}$$

By subtracting the equation for Φ from the one for $\tilde{\Phi}$, we deduce

$$-\left(\tilde{\Phi} - \Phi\right)_t + \frac{|\nabla \tilde{\Phi}|^2}{2} - \frac{|\nabla \tilde{\Phi}|^2}{2} + (\nabla W \star \rho) \cdot \nabla \left(\tilde{\Phi} - \Phi\right) = \Delta \left(\tilde{\Phi} - \Phi\right),$$

and similarly,

$$(\tilde{\eta} - \eta)_t = \mathrm{div}\left((\nabla W \star \rho)(\tilde{\eta} - \eta) + \nabla \tilde{\Phi} \tilde{\eta} - \nabla \Phi \eta\right) + \Delta(\tilde{\eta} - \eta).$$

Multiplying the first equation above by $\tilde{\eta} - \eta$ and subtracting the second equation multiplied by $\tilde{\Phi} - \Phi$ and then integrating by parts on $[0,T] \times \mathbb{T}^d$, we obtain

$$\int_{\mathbb{T}^d} (\tilde{\Phi} - \Phi)(\tilde{\eta} - \eta) dx \Big|_0^T = \int_0^T \int_{\mathbb{T}^d} \left[\left(\frac{|\nabla \tilde{\Phi}|^2}{2} - \frac{|\nabla \tilde{\Phi}|^2}{2} \right) (\tilde{\eta} - \eta) - \left(\nabla \tilde{\Phi} \tilde{\eta} - \nabla \Phi \eta \right) \cdot \nabla (\tilde{\Phi} - \Phi) \right] dx dt.$$

Recalling the terminal conditions on Φ, $\tilde{\Phi}$ and initial conditions on $\eta, \tilde{\eta}$, we have that the left hand side of the equality is zero, and thus

$$\int_0^T \int_{\mathbb{T}^d} \frac{|\nabla \tilde{\Phi} - \nabla \Phi|^2}{2} (\tilde{\eta} + \eta) dx dt = 0.$$

Hence $\nabla \tilde{\Phi} = \nabla \Phi$ and then $\Phi_t = \tilde{\Phi}_t$. Since $\Phi(T,x) = \tilde{\Phi}(T,x) = \phi_T(x)$, we obtain $\Phi = \tilde{\Phi}$. The above arguments show that any subsequence of the sequence $\{\Phi_n\}_{n \geq 1}$ has a further subsequence converging to Φ in $\mathcal{C}^{\beta,1+\beta}(\mathbb{T}^d \times [0,T])$. This proves that $\Phi_n \to \Phi$ in $\mathcal{C}^{\beta,1+\beta}(\mathbb{T}^d \times [0,T])$. Thus we have proved that the conditions of Schauder's Fixed-Point Theorem are satisfied, we infer that $\mathcal{F}$ has a fixed point in $\mathcal{K}_\alpha$ and the proof is complete. ■

References

[1] Y. Achdou, *Finite difference methods for mean field games*, in Hamilton-Jacobi Equations: Approximations, Numerical Analysis and Applications, Springer, 2013, pp. 1–47.

[2] Y. Achdou, F. Buera, J.-M. Lasry, P.-L. Lions, and B. Moll, *Partial differential equation models in macroeconomics*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 372 (2014), pp. 20130397, 19.

[3] Y. Achdou, F. J. Buera, J.-M. Lasry, P.-L. Lions, and B. Moll, *Partial differential equation models in macroeconomics*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 372 (2014), pp. 20130397, 19.

[4] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta, *Mean field games: numerical methods for the planning problem*, SIAM J. Control Optim., 50 (2012), pp. 77–109.

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

[6] Y. Achdou and V. Perez, *Iterative strategies for solving linearized discrete mean field games systems*, Netw. Heterog. Media, 7 (2012), pp. 197–217.

[7] G. Albi, Y.-P. Choi, M. Fornasier, and D. Kalise, *Mean Field Control Hierarchy*, Appl. Math. Optim., 76 (2017), pp. 93–135.

[8] M. Bardi and M. Cirant, *Uniqueness of solutions in mean field games with several populations and Neumann conditions*, in PDE models for multi-agent phenomena, vol. 28 of Springer INdAM Ser., Springer, Cham, 2018, pp. 1–20.

[9] J. Barré, J. A. Carrillo, P. Degond, D. Peurichard, and E. Zatorska, *Particle interactions mediated by dynamical networks: assessment of macroscopic descriptions*, arXiv:1701.01435 , to appear in J. Nonlinear Sci., (2017).

[10] J. Barré, P. Degond, and E. Zatorska, *Kinetic theory of particle interactions mediated by dynamical networks*, arXiv:1607.01975, (2016).

[11] A. Bensoussan, J. Frehse, and P. Yam, *Mean field games and mean field type control theory*, Springer Briefs in Mathematics, Springer, New York, 2013.
[12] ——, *The master equation in mean field theory*, J. Math. Pures Appl. (9), 103 (2015), pp. 1441–1474.
[13] ——, *On the interpretation of the Master Equation*, Stochastic Process. Appl., 127 (2017), pp. 2093–2137.
[14] M. Bongini, M. Fornasier, and D. Kalise, *(Un)conditional consensus emergence under perturbed and decentralized feedback controls*, Discrete Contin. Dyn. Syst., 35 (2015), pp. 4071–4094.
[15] Y. Brenier, *Extended Monge-Kantorovich theory*, in Optimal transportation and applications (Martina Franca, 2001), vol. 1813 of Lecture Notes in Math., Springer, Berlin, 2003, pp. 91–121.
[16] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau, *Self-organization in biological systems*, Princeton Studies in Complexity, Princeton University Press, Princeton, NJ, 2003. Reprint of the 2001 original.
[17] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat, *Sparse stabilization and optimal control of the Cucker-Smale model*, Math. Control Relat. Fields, 3 (2013), pp. 447–466.
[18] ——, *Sparse stabilization and control of alignment models*, Math. Models Methods Appl. Sci., 25 (2015), pp. 521–564.
[19] P. Cardaliaguet, *Notes on mean-field games*, (2013).
[20] ——, *Weak solutions for first order mean field games with local coupling*, in Analysis and geometry in control theory and its applications, vol. 11 of Springer INdAM Ser., Springer, Cham, 2015, pp. 111–158.
[21] P. Cardaliaguet and P. Graber, *Mean field games systems of first order*, ESAIM Control Optim. Calc. Var., 21 (2015), pp. 690–722.
[22] P. Cardaliaguet, P. Graber, A. Porretta, and D. Tonon, *Second order mean field games with degenerate diffusion and local coupling*, NoDEA Nonlinear Differential Equations Appl., 22 (2015), pp. 1287–1317.
[23] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta, *Long time average of mean field games*, Netw. Heterog. Media, 7 (2012), pp. 279–301.
[24] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta, *Long time average of mean field games with a nonlocal coupling*, SIAM Journal on Control and Optimization, 51 (2013), pp. 3558–3591.
[25] P. Cardaliaguet and A. Porretta, *Long time behavior of the master equation in mean field game theory*, Anal. PDE, 12 (2019), pp. 1397–1453.
[26] R. Carmona and F. Delarue, *The master equation for large population equilibriums*, in Stochastic analysis and applications 2014, vol. 100 of Springer Proc. Math. Stat., Springer, Cham, 2014, pp. 77–128.
[27] J. A. Carrillo, A. Colombi, and M. Scianna, *Adhesion and volume constraints via nonlocal interactions lead to cell sorting*, J. Theo. Biology, 445 (2018), pp. 75–91.
[28] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil, *Particle, kinetic, and hydrodynamic models of swarming*, in Mathematical modeling of collective behavior in socio-economic and life sciences, Model. Simul. Sci. Eng. Technol., Birkhäuser Boston, Inc., Boston, MA, 2010, pp. 297–336.
[29] J. A. Carrillo, Y. Huang, and S. Martin, *Explicit flock solutions for Quasi-Morse potentials*, European J. Appl. Math., 25 (2014), pp. 553–578.

[30] A. Cesaroni, M. Cirant, S. Dipierro, M. Novaga, and E. Valdinoci, *On stationary fractional mean field games*, J. Math. Pures Appl. (9), 122 (2019), pp. 1–22.

[31] J.-F. Chassagneux, D. Crisan, and F. Delarue, *Classical solutions to the master equation for large population equilibria*, 2014.

[32] P. E. Chaudru de Raynal, *Strong existence and uniqueness for degenerate SDE with Hölder drift*, Ann. Inst. Henri Poincaré Probab. Stat., 53 (2017), pp. 259–286.

[33] L. Chayes and V. Panferov, *The McKean-Vlasov equation in finite volume*, J. Stat. Phys., 138 (2010), pp. 351–380.

[34] M. Cirant, *Multi-population Mean Field Games systems with Neumann boundary conditions*, J. Math. Pures Appl. (9), 103 (2015), pp. 1294–1315.

[35] ——, *Stationary focusing mean-field games*, Comm. Partial Differential Equations, 41 (2016), pp. 1324–1346.

[36] M. R. D'Orsogna, Y. L. Chuang, A. L. Bertozzi, and L. S. Chayes, *Self-propelled particles with soft-core interactions: Patterns, stability, and collapse*, Phys. Rev. Lett., 96 (2006), p. 104302.

[37] B. Düring, P. Markowich, J.-F. Pietschmann, and M.-T. Wolfram, *Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 465 (2009), pp. 3687–3708.

[38] L. C. Evans, *Adjoint and compensated compactness methods for Hamilton-Jacobi PDE*, Arch. Ration. Mech. Anal., 197 (2010), pp. 1053–1088.

[39] L. C. Evans, *Adjoint and compensated compactness methods for Hamilton-Jacobi PDE*, Arch. Ration. Mech. Anal., 197 (2010), pp. 1053–1088.

[40] M. Fornasier, B. Piccoli, and F. Rossi, *Mean-field sparse optimal control*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 372 (2014), pp. 20130400, 21.

[41] M. Fornasier and F. Solombrino, *Mean-field optimal control*, ESAIM Control Optim. Calc. Var., 20 (2014), pp. 1123–1152.

[42] X. Gabaix, J.-M. Lasry, P.-L. Lions, and B. Moll, *The dynamics of inequality*, Econometrica, 84 (2016), pp. 2071–2111. Appendix G with Zhaonan Qu, Appendix G appearing in the supplementary materials not found within the manuscript.

[43] W. Gangbo and A. Świech ech, *Existence of a solution to an equation arising from the theory of mean field games*, J. Differential Equations, 259 (2015), pp. 6573–6643.

[44] A. Gerisch and M. A. J. Chaplain, *Mathematical modelling of cancer cell invasion of tissue: local and non-local models and the effect of adhesion*, J. Theoret. Biol., 250 (2008), pp. 684–704.

[45] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[46] D. Gomes, S. Patrizi, and V. Voskanyan, *On the existence of classical solutions for stationary extended mean field games*, Nonlinear Anal., 99 (2014), pp. 49–79.

[47] D. Gomes and E. Pimentel, *Time-dependent mean-field games with logarithmic nonlinearities*, SIAM J. Math. Anal., 47 (2015), pp. 3798–3812.

[48] D. Gomes, E. Pimentel, and H. Sánchez-Morgado, *Time-dependent mean-field games in the subquadratic case*, Comm. Partial Differential Equations, 40 (2015), pp. 40–76.

[49] ——, *Time-dependent mean-field games in the superquadratic case*, ESAIM Control Optim. Calc. Var., 22 (2016), pp. 562–580.

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

[51] D. Gomes, G. E. Pires, and H. Sánchez-Morgado, *A-priori estimates for stationary mean-field games*, Netw. Heterog. Media, 7 (2012), pp. 303–314.

[52] D. Gomes and H. Sánchez Morgado, *A stochastic Evans-Aronsson problem*, Trans. Amer. Math. Soc., 366 (2014), pp. 903–929.

[53] H. Hildenbrandt, C. Carere, and C. K. Hemelrijk, *Self-organised complex aerial displays of thousands of starlings: a model*, Behavioral Ecology, 107 (2010), pp. 1349–1359.

[54] T. Hillen and K. J. Painter, *A user's guide to PDE models for chemotaxis*, J. Math. Biol., 58 (2009), pp. 183–217.

[55] D. D. Holm and V. Putkaradze, *Aggregation of finite-size particles with variable mobility*, Phys. Rev. Lett., 95:226106 (2005).

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

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

[58] O. Itskhoki and B. Moll, *Optimal development policies with financial frictions*, Econometrica, 87 (2019), pp. 139–173.

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

[60] ——, *Jeux à champ moyen. II. Horizon fini et contrôle optimal*, C. R. Math. Acad. Sci. Paris, 343 (2006), pp. 679–684.

[61] ——, *Mean field games*, Jpn. J. Math., 2 (2007), pp. 229–260.

[62] ——, *Mean field games*, Cahiers de la Chaire Finance et Développement Durable, (2007).

[63] P.-L. Lions, *Generalized solutions of Hamilton-Jacobi equations*, vol. 69 of Research Notes in Mathematics, Pitman (Advanced Publishing Program), Boston, Mass.-London, 1982.

[64] R. Lukeman, Y. Li, and L. Edelstein-Keshet, *Inferring individual rules from collective behavior*, Proc. Natl. Acad. Sci. U.S.A., 107 (2010), pp. 12576–12580.

[65] K. J. Painter, J. M. Bloomfield, J. A. Sherratt, and A. Gerisch, *A nonlocal model for contact attraction and repulsion in heterogeneous cell populations*, Bull. Math. Biol., 77 (2015), pp. 1132–1165.

[66] B. t. Perthame, *Transport equations in biology*, Frontiers in Mathematics, Birkhäuser Verlag, Basel, 2007.

[67] B. Piccoli, F. Rossi, and E. Trélat, *Control to flocking of the kinetic Cucker-Smale model*, SIAM J. Math. Anal., 47 (2015), pp. 4685–4719.

[68] E. Pimentel and V. Voskanyan, *Regularity for second-order stationary mean-field games*, Indiana Univ. Math. J., 66 (2017), pp. 1–22.

[69] A. Porretta, *Weak solutions to Fokker-Planck equations and mean field games*, Arch. Ration. Mech. Anal., 216 (2015), pp. 1–62.

[70] J. P. Taylor-King, B. Franz, C. A. Yates, and R. Erban, *Mathematical modelling of turning delays in swarm robotics*, IMA J. Appl. Math., 80 (2015), pp. 1454–1474.

[71] C. Villani, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.

[72] J. H. von Brecht, D. Uminsky, T. Kolokolnikov, and A. L. Bertozzi, *Predicting pattern formation in particle interactions*, Math. Models Methods Appl. Sci., 22 (2012), pp. 1140002, 31.

José A. Carrillo
Department of Mathematics, Imperial College London, London SW7 2AZ, UK.
E-mail address: carrillo@imperial.ac.uk

Edgard A. Pimentel
Department of Mathematics, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, 22495-900, Brazil.

E-mail address: pimentel@puc-rio.br

Vardan K. Voskanyan
CMUC, Department of Mathematics, University of Coimbra, Apartado 3008, EC Santa Cruz 3001-501, Coimbra, Portugal

E-mail address: vardan.voskanyan@mat.uc.pt