LEARNING IN MEAN FIELD GAMES: THE FICTITIOUS PLAY

Pierre Cardaliaguet[1] and Saeed Hadikhanloo[2]

Abstract. Mean Field Game systems describe equilibrium configurations in differential games with infinitely many infinitesimal interacting agents. We introduce a learning procedure (similar to the Fictitious Play) for these games and show its convergence when the Mean Field Game is potential.

1. Introduction

Mean Field Game is a class of differential games in which each agent is infinitesimal and interacts with a huge population of other agents. These games have been introduced simultaneously by Lasry and Lions [20–23] and Huang *et al.* [18], (actually a discrete in time version of these games were previously known under the terminology of heterogenous models in economics. See for instance [2]). The classical notion of equilibrium solution in mean field game (abbreviated MFG) is given by a pair of maps (u, m), where $u = u(t, x)$ is the value function of a typical small player while $m = m(t, x)$ denotes the density at time t and at position x of the population. The value function u satisfies a Hamilton–Jacobi equation – in which m enters as a parameter and describes the influence of the population on the cost of each agent –, while the density m evolves in time according to a Fokker–Planck equation in which u enters as a drift. More precisely the pair (u, m) is a solution of the *MFG system*, which reads

$$
\begin{cases}
\text{(i)} & -\partial_t u - \sigma \Delta u + H(x, \nabla u(t, x)) = f(x, m(t)) \\
\text{(ii)} & \partial_t m - \sigma \Delta m - \operatorname{div}(m D_p H(x, \nabla u)) = 0 \\
& m(0, x) = m_0(x), \ u(T, x) = g(x, m(T)).
\end{cases}
\tag{1.1}
$$

In the above system, $T > 0$ is the horizon of the game, σ is a nonnegative parameter describing the intensity of the (individual) noise each agent is submitted to (for simplicity we assume that either $\sigma = 0$ (no noise) or $\sigma = 1$, some individual noise). The map H is the Hamiltonian of the control problem (thus typically convex in the gradient variable). The running cost f and the terminal cost g depend on the one hand on the position of the agent and, on the other hand, on the population density. Note that, in order to solve the (backward)

[1] Université Paris-Dauphine, PSL Research University, Ceremade, Place du Maréchal de Lattre de Tassigny 75775 Paris cedex 16, France. cardaliaguet@ceremade.dauphine.fr

[2] Université Paris-Dauphine, PSL Research University, Lamsade, Place du Maréchal de Lattre de Tassigny 75775 Paris cedex 16, France.

Hamilton–Jacobi equation (*i.e.*, the optimal control problem of each agent) one has to know the evolution of the population density, while the Fokker–Planck equation depends on the optimal strategies of the agents (through the drift term $-\text{div}(mD_pH(x,\nabla u)))$. The MFG system formalizes therefore *an equilibrium configuration*.

Under suitable assumptions recalled below, the MFG system (1.1) has at least one solution. This solution is even unique under a monotonicity condition on f and g. Under this condition, one can also show that it is the limit of symmetric Nash equilibria for a finite number of players as the number of players tends to infinity [10]; moreover, the optimal strategy given by the solution of the MFG system can be implemented in the game with finitely many players to give an approximate Nash equilibrium [12, 18]. MFG systems have been widely used in several areas ranging from engineering to economics, either under the terminology of heterogeneous agent model [2], or under the name of MFG [1, 15, 17].

In the present paper we raise the question of the actual formation of the MFG equilibrium. Indeed, the game being quite involved, it is unrealistic to assume that the agents can actually *compute* the equilibrium configuration. This seems to indicate that, if the equilibrium configuration arises, it is because the agents have *learned* how to play the game. For instance, people driving every day from home to work are dealing with such a learning issue. Every day they try to forecast the traffic and choose their optimal path accordingly, minimizing the journey and/or the consumed fuel for instance. If their belief on the traffic turns out not to be correct, they update their estimation, and so on. . . The question is wether such a procedure leads to stability or not.

The question of *learning* is a very classical one in game theory (see, for instance, the monograph [16]). There is by now a very large number of learning procedures for one-shot games in the literature. In the present paper we focus on a very classical and simple one: the Fictitious Play. The Fictitious Play was first introduced by Brown [5]. In this learning procedure, every player plays at each step the best response action with respect to the average of the previous actions of the other players. Fictitious Play does not necessarily converge, as shows the counter-example by Shapley [28], but it is known to converge for several classes of one shot games: for instance for zero-sum games (Robinson [27]), for 2×2 games (Miyasawa [24]), for potential games (Monderer and Shapley [26]). . .

Note that, in our setting, the question of learning makes all the more sense that the game is particularly intricate. Our aim is to define a Fictitious Play for the MFG system and to prove the convergence of this procedure under suitable assumption on the coupling f and g. The Fictitious Play for the MFG system runs as follows: the players start with a smooth initial belief $(m^0(t))_{t\in[0,T]}$. At the beginning of stage $n+1$, the players having observed the same past, share the same belief $(\overline{m}^n(t))_{t\in[0,T]}$ on the evolving density of the population. They compute their corresponding optimal control problem with value function u^{n+1} accordingly. When all players actually implement their optimal control the population density evolves in time and the players observe the resulting evolution $(m^{n+1}(t))_{t\in[0,T]}$. At the end of stage $n+1$ the players update their belief according to the rule (the same for all the players), which consists in computing the *average* of their observation up to time $n+1$. This yields to define by induction the sequences $u^n, m^n, \bar{m}^n$ by:

$$\begin{cases} \text{(i)} & -\partial_t u^{n+1} - \sigma \Delta u^{n+1} + H(x, \nabla u^{n+1}(t,x)) = f(x, \bar{m}^n(t)), \\ \text{(ii)} & \partial_t m^{n+1} - \sigma \Delta m^{n+1} - \text{div}(m^{n+1}D_pH(x, \nabla u^{n+1})) = 0, \\ & m^{n+1}(0) = m_0, \ u^{n+1}(x,T) = g(x, \bar{m}^n(T)) \end{cases} \tag{1.2}$$

where $\bar{m}^n = \frac{1}{n}\sum_{k=1}^n m^k$. Indeed, u^{n+1} is the value function at stage $n+1$ if the belief of players on the evolving density is $\bar{m}^n$, and thus solves (1.2)-(i). The actual density then evolves according to the Fokker–Planck equation (1.2)-(ii).

Our main result is that, under suitable assumption, this learning procedure converges, *i.e.*, any cluster point of the pre-compact sequence (u^n, m^n) is a solution of the MFG system (1.1) (by compact, we mean compact for the uniform convergence). Of course, if in addition the solution of the MFG system (1.1) is unique, then the full sequence converges. Let us recall (see [22]) that this uniqueness holds for instance if f and g are monotone:

$$\int (f(x,m) - f(x,m')) \, \mathrm{d}(m-m')(x) \geq 0, \qquad \int (g(x,m) - g(x,m') \, \mathrm{d}(m-m')(x) \geq 0$$

for any probability measures m, m'. This condition is generally interpreted as an aversion for congestion for the agents. Our key assumptions for the convergence result is that *f and g derive from potentials*. By this we mean that there exists $F = F(m)$ and $G = G(m)$ such that

$$f(x, m) = \frac{\delta F}{\delta m}(x, m) \qquad \text{and} \qquad g(x, m) = \frac{\delta G}{\delta m}(x, m).$$

The above derivative – in the space of measure – is introduced in Section 1.2, the definition being borrowed from [10]. Our assumption actually ensures that our MFG system is also "a potential game" (in the flavor of Monderer and Shapley [25]) so that the MFG system falls into a framework closely related to that of Monderer and Shapley [26]. Compared to [26], however, we face two issues. First we have an infinite population of players and the state space and the actions are also infinite. Second the game has a much more involved structure than in [26]. In particular, the potential for our game is far from being straightforward. We consider two different frameworks. In the first one, the so-called second order MFG systems where $\sigma = 1$ – which corresponds to the case where the players have a dynamic perturbed by independent noise – the potential is defined as a map of the evolving population density. This is reminiscent of the variational structure for the MFG system as introduced in [22] and exploited in [7, 11] for instance. The proof of the convergence then strongly relies on the regularity properties of the value function and of the population density (*i.e.*, of the u^n and m^n). The second framework is for first order MFG systems, where $\sigma = 0$. In contrast with the previous case, the lack of regularity of the value function and of the population density prevent to define the same Fictitious Play and the same potential. To overcome the difficulty, we lift the problem to the space of curves, which is the natural space of strategies. We define the Fictitious Play and a potential in this setting, and then prove the convergence, first for the infinite population and then for a large, but finite, one.

As far as we are aware of, our paper is the first one to consider a learning procedure in the framework of mean field games. Let us nevertheless point out that, for a particular class of MFG systems (quadratic Hamiltonians, local coupling), Guéant introduces in [14] an algorithm which is closely related to a replicator dynamics: namely it is exactly (1.2) in which one replaces $\bar{m}^n$ by m^n in (1.2)-(i)). The convergence is proved by using a kind of monotonicity of the sequence. This monotonicity does not hold in the more intricate framework considered here.

For simplicity we work in the periodic setting: we assume that the maps H, f and g are periodic in the space variable (and thus actually defined on the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$). This simplifies the estimates and the notation. However we do not think that the result changes in a substantial way if the state space is $\mathbb{R}^d$ or a subdomain of $\mathbb{R}^d$, with suitable boundary conditions.

The paper is organized as follows: we complete the introduction by fixing the main notation and stating the basic assumptions on the data. Then we define the notion of potential MFG and characterize the conditions of deriving from a potential. Section 2 is devoted to the Fictitious Play for second order MFG systems while Section 3 deals with the first order ones.

1.1. Preliminaries and assumptions

If X is a metric space, we denote by $\mathcal{P}(X)$ the set of Borel probability measures on X. When $X = \mathbb{T}^d$ ($\mathbb{T}^d$ being the torus $\mathbb{R}^d/\mathbb{Z}^d$), we endow $\mathcal{P}(\mathbb{T}^d)$ with the distance

$$\mathbf{d}_1(\mu, \nu) = \sup_h \left\{ \int_{\mathbb{T}^d} h(x) \, \mathrm{d}(\mu - \nu)(x) \right\} \qquad \mu, \nu \in \mathcal{P}(\mathbb{T}^d), \tag{1.3}$$

where the supremum is taken over all the maps $h : \mathbb{T}^d \to \mathbb{R}$ which are 1-Lipschitz continuous. Then $\mathbf{d}_1$ metricizes the weak-* convergence of measures on $\mathbb{T}^d$.

The maps H, f and g are periodic in the space arguments: $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ while $f, g : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$. In the same way, the initial condition $m_0 \in \mathcal{P}(\mathbb{T}^d)$ is periodic in space and is assumed to be absolutely continuous with a smooth density.

We now state our key assumptions on the data: these conditions are valid throughout the paper. On the initial measure m_0, we assume that

$$m_0 \text{ has a smooth density (again denoted } m_0). \tag{1.4}$$

Concerning the Hamiltonian, we suppose that H is of class $\mathcal{C}^2$ on $\mathbb{T}^d \times \mathbb{R}^d$ and quadratic-like in the second variable: there is $\bar{C} > 0$ such that

$$H \in \mathcal{C}^2(\mathbb{T}^d \times \mathbb{R}^d) \text{ and } \frac{1}{C}I_d \leq D_{pp}^2 H(x,p) \leq \bar{C}I_d \qquad \forall(x,p) \in \mathbb{T}^d \times \mathbb{R}^d. \tag{1.5}$$

Moreover, we suppose that $D_x H$ satisfies the lower bound:

$$\langle D_x H(x,p), p \rangle \geq -\bar{C}\left(|p|^2 + 1\right). \tag{1.6}$$

The maps f and g are supposed to be globally Lipschitz continuous (in both variables) and regularizing:

$$\text{The map } m \to f(\cdot, m) \text{ is Lipschitz continuous from } \mathcal{P}(\mathbb{T}^d) \text{ to } \mathcal{C}^2(\mathbb{T}^d)$$
$$\text{while the map } m \to g(\cdot, m) \text{ is Lipschitz continuous from } \mathcal{P}(\mathbb{T}^d) \text{ to } \mathcal{C}^3(\mathbb{T}^d). \tag{1.7}$$

In particular,

$$\sup_{m \in P(\mathbb{T}^d)} \|f(\cdot, m)\|_{\mathcal{C}^2} + \|g(\cdot, m)\|_{\mathcal{C}^3} < \infty. \tag{1.8}$$

Assumptions (1.4), (1.5), (1.6), (1.7), (1.9) are in force throughout the paper. As explained below, they ensure the MFG system to have at least one solution.

To ensure the uniqueness of the solution, we sometime require f and g to be monotone: for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d} (f(x,m) - f(x,m'))d(m-m')(x) \geq 0, \quad \int_{\mathbb{T}^d} (g(x,m) - g(x,m'))d(m-m')(x) \geq 0. \tag{1.9}$$

1.2. Potential mean field games

In this section we introduce the main structure condition on the data f and g of the game: we assume that f and g are the derivative, with respect to the measure, of potential maps F and G. In this case we say that f and g derive from a potential.

Let us first explain what we mean by a derivative with respect to a measure. Let $F : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ be a continuous map. We say that the continuous map $\frac{\delta F}{\delta m} : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ is the derivative of F if, for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$,

$$\lim_{s \to 0} \frac{F((1-s)m + sm') - F(m)}{s} = \int_{\mathbb{T}^d} \frac{\delta F}{\delta m}(m,x)d(m'-m)(x). \tag{1.10}$$

As $\frac{\delta F}{\delta m}$ is continuous, this equality can be equivalently written as

$$F(m') - F(m) = \int_0^1 \int_{\mathbb{T}^d} \frac{\delta F}{\delta m}((1-s)m + sm'), x)d(m'-m)(x)ds,$$

for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$. Note that $\frac{\delta F}{\delta m}$ is defined only up to an additive constant. To fix the ideas we assume therefore that

$$\int_{\mathbb{T}^d} \frac{\delta F}{\delta m}(m,x)dm(x) = 0 \qquad \forall m \in \mathcal{P}(\mathbb{T}^d).$$

We also use the notation $\frac{\delta F}{\delta m}(m)(m'-m) := \int_{\mathbb{T}^d} \frac{\delta F}{\delta m}(m,x)d(m'-m)(x)$ and often see the map $\frac{\delta F}{\delta m}$ as a continuous function from $\mathcal{P}(\mathbb{T}^d)$ to $\mathcal{C}(\mathbb{T}^d, \mathbb{R})$.

Definition 1.1. A Mean Field Game is called a Potential Mean Field Game if the instantaneous and final cost functions $f, g : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ derive from potentials, *i.e.*, there exists continuously differentiable maps $F, G : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that

$$\frac{\delta F}{\delta m} = f, \qquad \frac{\delta G}{\delta m} = g.$$

In the rest of the section we characterize the maps f which derive from a potential. Although this is not used in the rest of the paper, this characterization is natural and we believe that it has its own interest.

To proceed we assume for the rest of the section that, for any $x \in \mathbb{T}^d$, $f(x, \cdot)$ has a derivative and that this derivative $\frac{\delta f}{\delta m} : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \to \mathbb{R}$ is continuous.

Proposition 1.2. *The map $f : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ derives from a potential, if and only if,*

$$\frac{\delta f}{\delta m}(y, m, x) = \frac{\delta f}{\delta m}(x, m, y) \qquad \forall x, y \in \mathbb{T}^d, \ \forall m \in \mathcal{P}(\mathbb{T}^d).$$

Proof. First assume that f derives from a potential $F : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$. Deriving in m the relation $\frac{\delta F}{\delta m} = f$ we obtain

$$\frac{\delta^2 F}{\delta m^2}(m, x, y) = \frac{\delta f}{\delta m}(x, m, y) \qquad \forall x, y \in \mathbb{T}^d, \ m \in \mathcal{P}(\mathbb{T}^d).$$

As $\frac{\delta^2 F}{\delta m^2}(m, x, y)$ is symmetric in (x, y) (see [10]), so is $\frac{\delta f}{\delta m}(x, m, y)$.

Let us now assume that $\frac{\delta f}{\delta m}(x, m, y)$ is symmetric in (x, y). Let us fix $m_0 \in \mathcal{P}(\mathbb{T}^d)$ and set, for any $m \in \mathcal{P}(\mathbb{T}^d)$,

$$F(m) = \int_0^1 \int_{\mathbb{T}^d} f(x, (1-t)m_0 + tm)\mathrm{d}(m - m_0)(x)\mathrm{d}t.$$

We claim that F is a potential for f. Indeed, as f has a continuous derivative, so has F, with

$$\frac{\delta F}{\delta m}(m, y) = \int_0^1 t \int_{\mathbb{T}^d} \frac{\delta f}{\delta m}(x, (1-t)m_0 + tm, y) \ \mathrm{d}(m - m_0)(x)\mathrm{d}t$$
$$+ \int_0^1 f(y, (1-t)m_0 + tm)\mathrm{d}t. \tag{1.11}$$

As, by symmetry assumption,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(y, (1-t)m_0 + tm) = \int_{\mathbb{T}^d} \frac{\delta f}{\delta m}(y, (1-t)m_0 + tm, x)\mathrm{d}(m - m_0)(x)$$
$$= \int_{\mathbb{T}^d} \frac{\delta f}{\delta m}(x, (1-t)m_0 + tm, y)\mathrm{d}(m - m_0)(x),$$

we have therefore after integration by parts in (1.11),

$$\frac{\delta F}{\delta m}(m, y) = \left[t \ f(x, (1-t)m_0 + tm) \right]_0^1 = f(x, m). \qquad \square$$

2. The fictitious play for second order MFG systems

In this section, we study a learning procedure for the second order MFG system:

$$\begin{cases} \text{(i)} & -\partial_t u - \Delta u + H(x, \nabla u(t, x)) = f(x, m(t)), \quad (t, x) \in [0, T] \times \mathbb{T}^d \\ \text{(ii)} & \partial_t m - \Delta m - \mathrm{div}(m D_p H(x, \nabla u)) = 0, \quad (t, x) \in [0, T] \times \mathbb{T}^d \\ & m(0) = m_0, \ u(x, T) = g(x, m(T)), \quad x \in \mathbb{T}^d. \end{cases} \tag{2.1}$$

Let us recall (see [22]) that, under our assumptions (1.4)–(1.7), there exists at least one classical solution to (2.1) (*i.e.*, for which all the involved derivative exists and are continuous). If furthermore (1.9) holds, then the solution is unique.

2.1. The learning rule and the convergence result

The Fictitious Play can be written as follows: given a smooth initial guess $m^0 \in C^0([0,T], \mathcal{P}(\mathbb{T}^d))$, we define by induction sequences $u^n, m^n : [0,T] \times \mathbb{T}^d \to \mathbb{R}$ by:

$$
\begin{cases}
\text{(i)} & -\partial_t u^{n+1} - \Delta u^{n+1} + H(x, \nabla u^{n+1}(t,x)) = f(x, \bar{m}^n(t)), \quad (t,x) \in [0,T] \times \mathbb{T}^d \\
\text{(ii)} & \partial_t m^{n+1} - \Delta m^{n+1} - \mathrm{div}(m^{n+1} D_p H(x, \nabla u^{n+1})) = 0, \quad (t,x) \in [0,T] \times \mathbb{T}^d \\
& m^{n+1}(0) = m_0, \ u^{n+1}(x,T) = g(x, \bar{m}^n(T)), \quad x \in \mathbb{T}^d
\end{cases} \tag{2.2}
$$

where $\bar{m}^n(t,x) = \frac{1}{n} \sum_{k=1}^n m^k(t,x)$. The interpretation is that, at the beginning of stage $n+1$, the players have the same belief of the future density of the population $(\overline{m}^n(t))_{t \in [0,T]}$ and compute their corresponding optimal control problem with value function u^{n+1}. Their optimal (closed-loop) control is then $(t,x) \to -D_p H(x, \nabla u^{n+1}(t,x))$. When all players actually implement this control the population density evolves in time according to (2.2)-(ii). We assume that the players observe the resulting evolution of the population density $(m^{n+1}(t))_{t \in [0,T]}$. At the end of stage $n+1$ the players update their guess by computing the *average* of their observation up to time $n+1$.

In order to show the convergence of the Fictitious Play, we assume that the MFG is potential, *i.e.* there are potential functions $F, G : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that

$$
f(x,m) = \frac{\delta F}{\delta m}(m,x) \qquad \text{and} \qquad g(x,m) = \frac{\delta G}{\delta m}(m,x). \tag{2.3}
$$

We also assume, besides the smoothness assumption (1.4), that m_0 is smooth and positive.

Theorem 2.1. *Under the assumptions* (1.4)–(1.7) *and* (2.3), *the family* $\{(u^n, m^n)\}_{n \in \mathbb{N}}$ *is uniformly continuous and any cluster point is a solution to the second order MFG* (2.1).

If, in addiction, the monotonicity condition (1.9) *holds, then the whole sequence* $\{(u^n, m^n)\}_{n \in \mathbb{N}}$ *converges to the unique solution of* (2.1).

The key remark to prove Theorem 2.1 is that the game itself has a *potential*. Given $m \in C^0([0,T] \times \mathbb{T}^d)$ and $w \in C^0([0,T] \times \mathbb{T}^d)$ such that, in the sense of distribution,

$$
\partial_t m - \Delta m + \mathrm{div}(w) = 0 \text{ in } (0,T) \times \mathbb{T}^d \qquad m(0) = m_0, \tag{2.4}
$$

let

$$
\Phi(m,w) = \int_0^T \int_{\mathbb{T}^d} m(t,x) H^*(x, -w(t,x)/m(t,x)) \mathrm{d}x \mathrm{d}t + \int_0^T F(m(t)) \mathrm{d}t + G(m(T)),
$$

where H^* is the convex conjugate of H:

$$
H^*(x,q) = \sup_{p \in \mathbb{R}^d} \langle p, q \rangle - H(x,p).
$$

In the definition of Φ, we set by convention, when $m = 0$,

$$
mH^*(x, -w/m) = \begin{cases} 0 & \text{if } w = 0 \\ +\infty & \text{otherwise.} \end{cases}
$$

For sake of simplicity, we often drop the integration and the variable (t,x) to write the potential in a shorter form:

$$
\Phi(m,w) = \int_0^T \int_{\mathbb{T}^d} mH^*(x, -w/m) + \int_0^T F(m(t)) \mathrm{d}t + G(m(T)).
$$

It is explained in [22] Section 2.6 that (u,m) is a solution to (2.1) if and only if (m,w) is a minimizer of Φ and $w = -m D_p H(\cdot, \nabla u)$ and also constrained to (2.4). We show here that the same map can be used as a potential in the Fictitious Play: Φ (almost) decreases at each step of the Fictitious Play and the derivative of Φ does not vary too much at each step. Then the proof of [26] applies.

2.2. Proof of the convergence

Before starting the proof of Theorem 2.1, let us fix some notations. First we set

$$w^n(t,x) = -m^n(t,x)D_pH(x, \nabla u^n(t,x)) \text{ and } \bar{w}^n(t,x) = \frac{1}{n}\sum_{k=1}^{n} w^k(t,x). \tag{2.5}$$

Since the Fokker–Planck equation is linear we have:

$$\partial_t \bar{m}^{n+1} - \Delta\bar{m}^{n+1} + \text{div}(\bar{w}^{n+1}) = 0, \quad t \in [0,T], \qquad \bar{m}^{n+1}(0) = m_0. \tag{2.6}$$

Recall that H^* is the convex conjugate of H:

$$H^*(x,q) = \sup_{p\in\mathbb{R}^d} \langle p,q\rangle - H(x,p).$$

We define $\hat{p}(x,q)$ as the minimum in the above right-hand side:

$$H^*(x,p) = \langle \hat{p}(x,q), q\rangle - H(x,\hat{p}(x,q)). \tag{2.7}$$

Note that $\hat{p}$ is characterized by $q = D_pH(x, \hat{p}(x,q))$. The uniqueness comes from the fact that H satisfies $D_{pp}H \geq \frac{1}{C}I_d$, which yields that $D_pH(x,\cdot)$ is one-to-one. We note for later use that

$$mH^*(x, -\frac{q}{m}) = \sup_{p\in\mathbb{R}^d} -\langle p,q\rangle - mH(x,p).$$

Next we state a standard result on uniformly convex functions, the proof of which is postponed:

Lemma 2.2. *Under assumption* (1.5), *we have for any* $x \in \mathbb{T}^d$, $p,q \in \mathbb{R}^d$:

$$H(x,p) + H^*(x,q) - \langle p,q\rangle \geq \frac{1}{2C}|q - D_pH(x,p)|^2.$$

The following Lemma explains that Φ is "almost decreasing" along the sequence $(\bar{m}^n, \bar{w}^n)$.

Lemma 2.3. *There exists a constant* $C > 0$ *such that, for any* $n \in \mathbb{N}^*$,

$$\Phi(\bar{m}^{n+1}, \bar{w}^{n+1}) - \Phi(\bar{m}^n, \bar{w}^n) \leq -\frac{1}{C}\frac{a_n}{n} + \frac{C}{n^2}, \tag{2.8}$$

where $a_n = \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1}|\bar{w}^{n+1}/\bar{m}^{n+1} - w^{n+1}/m^{n+1}|^2$.

Throughout the proofs, C denotes a constant which depends on the data of the problem only (*i.e.*, on H, f, g and m_0) and might change from line to line. We systematically use the fact that, as f and g admit F and G as a potential and are globally Lipschitz continuous, there exists a constant $C > 0$ such that, for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$ and $s \in [0,1]$,

$$\left|F(m + s(m'-m)) - F(m) - s\int_{\mathbb{T}^d} f(x,m)\mathrm{d}(m'-m)(x)\right| < C|s|^2,$$

$$\left|G(m + s(m'-m)) - G(m) - s\int_{\mathbb{T}^d} g(x,m)\mathrm{d}(m'-m)(x)\right| < C|s|^2.$$

Proof of Lemma 2.3. We have

$$\Phi(\bar{m}^{n+1}, \bar{w}^{n+1}) = \Phi(\bar{m}^n, \bar{w}^n) + A + B,$$

where

$$A = \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} H^*(-\bar{w}^{n+1}/\bar{m}^{n+1}) - \bar{m}^n H^*(-\bar{w}^n/\bar{m}^n) \tag{2.9}$$

$$B = \int_0^T \left(F(\bar{m}^{n+1}(t)) - F(\bar{m}^n(t)) \right) \mathrm{d}t + \left(G(\bar{m}^{n+1}(T)) - G(\bar{m}^n(T)) \right). \tag{2.10}$$

Since F is C^1 with respect to m with derivative f, we have

$$B \leq \int_0^T \int_{\mathbb{T}^d} f(x, \bar{m}^n(t))(\bar{m}^{n+1} - \bar{m}^n) + \int_{\mathbb{T}^d} g(x, \bar{m}^n(T))(\bar{m}^{n+1} - \bar{m}^n) + \frac{C}{n^2}.$$

As $\bar{m}^{n+1} - \bar{m}^n = \frac{1}{n}(m^{n+1} - \bar{m}^{n+1})$, we find after rearranging:

$$B \leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} f(x, \bar{m}^n(t))(m^{n+1} - \bar{m}^{n+1}) + \frac{1}{n} \int_{\mathbb{T}^d} g(x, \bar{m}^n(T))(m^{n+1}(T) - \bar{m}^{n+1}(T)) + \frac{C}{n^2}.$$

Using now the equation satisfied by u^{n+1} we get

$$B \leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} \left(-\partial_t u^{n+1} - \Delta u^{n+1} + H(x, \nabla u^{n+1}) \right)(m^{n+1} - \bar{m}^{n+1})$$

$$+ \frac{1}{n} \int_{\mathbb{T}^d} g(x, \bar{m}^n(T))(m^{n+1}(T) - \bar{m}^{n+1}(T)) + \frac{C}{n^2}$$

$$\leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} \left(\partial_t(m^{n+1} - \bar{m}^{n+1}) - \Delta(m^{n+1} - \bar{m}^{n+1}) \right) u^{n+1}$$

$$+ \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} H(x, \nabla u^{n+1})(m^{n+1} - \bar{m}^{n+1}) + \frac{C}{n^2},$$

where we have integrated by parts in the second inequality. Using now the equation satisfied by $m^{n+1} - \bar{m}^{n+1}$ derived from (2.5) and integrating again by parts, we obtain

$$B \leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} \langle w^{n+1} - \bar{w}^{n+1}, \nabla u^{n+1} \rangle + H(x, \nabla u^{n+1})(m^{n+1} - \bar{m}^{n+1}) + \frac{C}{n^2}.$$

Note that by Lemma 2.2,

$$-\langle \bar{w}^{n+1}, \nabla u^{n+1} \rangle - H(x, \nabla u^{n+1})\bar{m}^{n+1} \leq \bar{m}^{n+1} H^*(-\bar{w}^{n+1}/\bar{m}^{n+1})$$

$$- \frac{1}{2\bar{C}} \bar{m}^{n+1} \left| \bar{w}^{n+1}/\bar{m}^{n+1} - w^{n+1}/m^{n+1} \right|^2$$

while, by the definition of w^{n+1},

$$\langle w^{n+1}, \nabla u^{n+1} \rangle + H(x, \nabla u^{n+1})m^{n+1} = -m^{n+1} H^*(-w^{n+1}/m^{n+1}).$$

Therefore

$$B \leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} H^*(-\bar{w}^{n+1}/\bar{m}^{n+1}) - m^{n+1} H^*(-w^{n+1}/m^{n+1})$$

$$- \frac{1}{2\bar{C}n} \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} \left| \bar{w}^{n+1}/\bar{m}^{n+1} - w^{n+1}/m^{n+1} \right|^2 + \frac{C}{n^2}. \tag{2.11}$$

On the other hand, recalling the definition of $\hat{p}$ in (2.7) and setting $\bar{p}^{n+1} = \hat{p}(\cdot, -\bar{w}^{n+1}/\bar{m}^{n+1})$, we can estimate A as follows:

$$
\begin{aligned}
A &\leq \int_0^T \int_{\mathbb{T}^d} -\langle \bar{p}^{n+1}, \bar{w}^{n+1} \rangle - \bar{m}^{n+1} H(x, \bar{p}^{n+1}) + \langle \bar{p}^{n+1}, \bar{w}^n \rangle + \bar{m}^n H(x, \bar{p}^{n+1}) \\
&= \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} \langle \bar{p}^{n+1}, \bar{w}^{n+1} \rangle + \bar{m}^{n+1} H(x, \bar{p}^{n+1}) - \langle \bar{p}^{n+1}, w^{n+1} \rangle - m^{n+1} H(x, \bar{p}^{n+1}) \\
&\leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} m^{n+1} H^*(-w^{n+1}/m^{n+1}) - \bar{m}^{n+1} H^*(-\bar{w}^{n+1}/\bar{m}^{n+1}).
\end{aligned} \tag{2.12}
$$

Putting together (2.11) and (2.12) we find:

$$
\Phi(\bar{m}^{n+1}, \bar{w}^{n+1}) - \Phi(\bar{m}^n, \bar{w}^n) \leq -\frac{1}{2\bar{C}} \frac{a_n}{n} + \frac{C}{n^2}
$$

where $a_n = \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} |\bar{w}^{n+1}/\bar{m}^{n+1} - w^{n+1}/m^{n+1}|^2$. $\qquad \square$

In order to proceed, let us recall some basic estimates on the system (2.2), the proof of which is postponed:

Lemma 2.4. *For any $\alpha \in (0, 1/2)$ there exist a constant $C > 0$ such that for any $n \in \mathbb{N}^*$*

$$
\|u^n\|_{C^{1+\alpha/2, 2+\alpha}} + \|m^n\|_{C^{1+\alpha/2, 2+\alpha}} \leq C, \qquad m^n \geq 1/C,
$$

where $C^{1+\alpha/2, 2+\alpha}$ is the usual Hölder space on $[0, T] \times \mathbb{T}^d$.

As a consequence, the u^n, the m^n and the w^n do not vary too much between two consecutive steps:

Lemma 2.5. *There exists a constant $C > 0$ such that*

$$
\|u^{n+1} - u^n\|_\infty + \|\nabla u^{n+1} - \nabla u^n\|_\infty + \|m^{n+1} - m^n\|_\infty + \|w^{n+1} - w^n\|_\infty \leq \frac{C}{n}.
$$

Proof. As $\bar{m}^n - \bar{m}^{n-1} = ((n-1)\bar{m}^{n-1} + m^n)/n$, where the m^n (and thus the $\bar{m}^n$) are uniformly bounded thanks to Lemma 2.4, we have by Lipschitz continuity of f and g that

$$
\sup_{t \in [0,T]} \|f(\cdot, \bar{m}^{n+1}(t)) - f(\cdot, \bar{m}^n(t))\|_\infty + \|g(\cdot, \bar{m}^{n+1}(T)) - g(\cdot, \bar{m}^n(T))\|_\infty \leq \frac{C}{n}. \tag{2.13}
$$

Thus, by comparison for the solution of the Hamilton–Jacobi equation, we get

$$
\|u^{n+1} - u^n\|_\infty \leq \frac{C}{n}. \tag{2.14}
$$

Let us set $z := u^{n+1} - u^n$. Then z satisfies

$$
-\partial_t z - \Delta z + H(x, \nabla u^n + \nabla z) - H(x, \nabla u^n) = f(x, \bar{m}^n(t)) - f(x, \bar{m}^{n-1}(t)).
$$

Multiplying by z and integrating over $[0, T] \times \mathbb{T}^d$ we find by (2.13) and (2.14):

$$
-\left[\int_{\mathbb{T}^d} \frac{z^2}{2}\right]_0^T + \int_0^T \int_{\mathbb{T}^d} |\nabla z|^2 + z(H(x, \nabla u^n + \nabla z) - H(x, \nabla u^n)) \leq \frac{C}{n^2}.
$$

Then we use the uniform bound on the ∇u^n given by Lemma 2.4 as well as (2.14) to get

$$
\int_0^T \int_{\mathbb{T}^d} \left(|\nabla z|^2 - \frac{C}{n}|\nabla z|\right) \leq \frac{C}{n^2}.
$$

Thus

$$\int_0^T \int_{\mathbb{T}^d} |\nabla z|^2 \leq \frac{C}{n^2},$$

which implies that $\|\nabla z\|_\infty \leq C/n$ since $\|\nabla^2 z\|_\infty + \|\partial_t \nabla z\|_\infty \leq C$ by Lemma 2.4.

We argue in a similar way for $\mu := m^{n+1} - m^n$: μ satisfies

$$\partial_t \mu - \Delta \mu - \operatorname{div}(\mu D_p H(x, Du^{n+1})) - \operatorname{div}(R) = 0,$$

where we have set $R = m^n \left(D_p H(x, \nabla u^{n+1}) - D_p H(x, \nabla u^n) \right)$. As $\|R\|_\infty \leq C/n$ by the previous step, we get the bound on $\|m^{n+1} - m^n\|_\infty \leq C/n$ by standard parabolic estimates. This implies the bound on $\|w^{n+1} - w^n\|_\infty$ by the definition of the w^n.   □

Combining Lemma 2.4 with Lemma 2.5 we immediately obtain that the sequence (a_n) defined in Lemma 2.3 is slowly varying in time.

Corollary 2.6. *There exists a constant $C > 0$ such that, for any $n \in \mathbb{N}^*$,*

$$|a_{n+1} - a_n| \leq \frac{C}{n}.$$

Proof of Theorem 2.1. From Lemma 2.3, we have for any $n \in \mathbb{N}^*$,

$$\Phi(\bar{m}^{n+1}, \bar{w}^{n+1}) - \Phi(\bar{m}^n, \bar{w}^n) \leq -\frac{1}{C}\frac{a_n}{n} + \frac{C}{n^2}$$

where $a_n = \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} |\bar{w}^{n+1}/\bar{m}^{n+1} - w^{n+1}/m^{n+1}|^2.$

Since the potential Φ is bounded from below the above inequality implies that

$$\sum_{n \geq 1} a_n/n < +\infty.$$

From Corollary 2.6, we also have, for any $n \in \mathbb{N}^*$,

$$|a_{n+1} - a_n| \leq \frac{C}{n}.$$

Then Lemma 2.7 below implies that $\lim_{n \to \infty} a_n = 0$.

In particular we have, by Lemma 2.4:

$$\lim_{n \to \infty} \int_0^T \int_{\mathbb{T}^d} |\bar{w}^n/\bar{m}^n - w^n/m^n|^2 \leq C \lim_{n \to \infty} \int_0^T \int_{\mathbb{T}^d} \bar{m}^n |\bar{w}^n/\bar{m}^n - w^n/m^n|^2 = 0.$$

This implies that the sequence $\{\bar{w}^n/\bar{m}^n - w^n/m^n\}_{n \in \mathbb{N}}$ – which is uniformly continuous from Lemma 2.4 – uniformly converges to 0 on $[0, T] \times \mathbb{T}^d$.

Recall that, by Lemma 2.4, the sequence $\{(u^{n+1}, m^n, \bar{m}^n, \bar{w}^n)\}_{n \in \mathbb{N}}$ is pre-compact for the uniform convergence. Let $(u, m, \bar{m}, \bar{w})$ be a cluster point of the sequence $\{(u^{n+1}, m^n, \bar{m}^n, \bar{w}^n)\}_{n \in \mathbb{N}}$. Our aim is to show that (u, m) is a solution to the MFG system (2.1), that $\bar{m} = m$ and that $\bar{w} = -m D_p H(\cdot, \nabla u)$.

Let $n_i \in \mathbb{N}, i \in \mathbb{N}$ be a subsequence such that $(u^{n_i+1}, m^{n_i}, \bar{m}^{n_i}, w^{n_i})$ uniformly converges to $(u, m, \bar{m}, \bar{w})$. By the estimates in Lemma 2.4, we have $D_p H(x, \nabla u^{n_j})$ converges uniformly to $D_p H(x, \nabla u)$, so that by (2.5) and the fact that the sequence $\{\bar{w}^n/\bar{m}^n - w^n/m^n\}_{n \in \mathbb{N}}$ converges to 0,

$$-D_p H(x, \nabla u) = \frac{w}{m} = \frac{\bar{w}}{\bar{m}}. \tag{2.15}$$

We now pass to the limit in (2.2) (in the viscosity sense for the Hamilton–Jacobi equation and in the sense of distribution for the Fokker–Planck equation) to get

$$
\begin{aligned}
&\text{(i)} \quad && -\partial_t u - \Delta u + H(x, \nabla u(t,x)) = f(x, \bar{m}(t)), \quad (t,x) \in [0,T] \times \mathbb{T}^d \\
&\text{(ii)} \quad && \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, \nabla u)) = 0, \quad (t,x) \in [0,T] \times \mathbb{T}^d \\
& && m(0) = m_0, \ u(x,T) = g(x, \bar{m}(T)), \quad x \in \mathbb{T}^d.
\end{aligned}
\tag{2.16}
$$

Letting $n \to +\infty$ in (2.6) we also have

$$
\partial_t \bar{m} - \Delta \bar{m} + \operatorname{div}(\bar{w}) = 0, \quad t \in [0,T], \qquad \bar{m}(0) = m_0.
$$

By (2.15), this means that m and $\bar{m}$ are both solutions to the same Fokker–Planck equation. Thus they are equal and (u,m) is a solution to the MFG system.

If (1.9) holds, then the MFG system has a unique solution (u,m), so that the compact sequence $\{(u^n, m^n)\}$ has a unique accumulation point (u,m) and thus converges to (u,m). □

In the proof of Theorem 2.1, we have used the following Lemma, which can be found in [26].

Lemma 2.7. *Consider a sequence of positive real numbers $\{a_n\}_{n \in \mathbb{N}}$ such that $\sum_{n=1}^{\infty} a_n/n < +\infty$. Then we have*

$$
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} a_n = 0.
$$

In addition, if there is a constant $C > 0$ such that $|a_n - a_{n+1}| < \frac{C}{n}$ then $\lim_{n \to \infty} a_n = 0$.

Proof. We reproduce the proof of [26] for the sake of completeness. For every $k \in \mathbb{N}$ define $b_k = \sum_{n=k}^{\infty} a_n/n$. Since $\sum_{n=1}^{\infty} a_n/n < +\infty$ we have $\lim_{k \to \infty} b_k = 0$. So we have:

$$
\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} b_k = 0,
$$

which yields the first result since:

$$
\sum_{n=1}^{N} a_n \le \sum_{k=1}^{N} b_k.
$$

For the second result, consider $\epsilon > 0$. We know that for every $\lambda > 0$ we have:

$$
\lim_{N \to \infty} \frac{1}{N} + \frac{1}{N+1} + \cdots + \frac{1}{[(1+\lambda)N]} = \log(1+\lambda),
$$

where $[a]$ denotes the integer part of the real number a. So if $\lambda_\epsilon > 0$ is so small that $\log(1+\lambda_\epsilon) < \frac{\epsilon}{2C}$, then there exist $N_\epsilon \in \mathbb{N}$ so large that for $N \ge N_\epsilon$ we have

$$
\frac{1}{N} + \frac{1}{N+1} + \cdots + \frac{1}{[(1+\lambda_\epsilon)N]} < \frac{\epsilon}{2C}.
\tag{2.17}
$$

Let $N \ge N_\epsilon$. Assume for a while that $a_N > \varepsilon$. As $|a_{k+1} - a_k| \le C/k$, (2.17) implies that $a_k > \frac{\epsilon}{2}$ for $N \le k \le [N(1+\lambda_\epsilon)]$. Thus

$$
\frac{1}{[N(1+\lambda_\epsilon)]} \sum_{k=1}^{[N(1+\lambda_\epsilon)]} a_k \ge \frac{\lambda_\epsilon}{1+\lambda_\epsilon} \frac{\epsilon}{2}.
$$

Since the average $N^{-1} \sum_{k=1}^{N} a_k$ converges to zero, the above inequality cannot hold for N large enough. This implies that $a_N \le \varepsilon$ for N sufficiently large, so that (a_k) converges to 0. □

Proof of Lemma 2.2. For simplicity of notation, we omit the x dependence in the various quantities. As by assumption (1.5) we have $\frac{1}{C}I_d \leq D_{pp}^2 H \leq \bar{C}I_d$, H^* is differentiable with respect to q and the following inequality holds: for any $q_1, q_2 \in \mathbb{R}^d$,

$$\langle D_q H^*(q_1) - D_q H^*(q_2), q_1 - q_2 \rangle \geq \frac{1}{\bar{C}}|q_1 - q_2|^2.$$

Let us fix $p, q \in \mathbb{R}^d$ and let $\hat{q} \in \mathbb{R}^d$ be the maximum in

$$\max_{q' \in \mathbb{R}^d} \langle q', p \rangle - H^*(q') = H(p).$$

Recall that $p = D_q H^*(\hat{q})$ and thus $\hat{q} = D_p H(p)$. Then

$$\begin{aligned}
H(p) + H^*(q) - \langle p, q \rangle &= H^*(q) - H^*(\hat{q}) - \langle q - \hat{q}, p \rangle \\
&= \int_0^1 \langle D_q H^*((1-t)\hat{q} + tq) - D_q H^*(\hat{q}), q - \hat{q} \rangle \mathrm{d}t \\
&= \int_0^1 \frac{1}{t} \langle D_q H^*((1-t)\hat{q} + tq) - D_q H^*(\hat{q}), ((1-t)\hat{q} + tq) - \hat{q} \rangle \mathrm{d}t \\
&\geq \int_0^1 t\, \frac{1}{\bar{C}}|\hat{q} - q|^2 = \frac{1}{2\bar{C}}|D_p H(p) - q|^2. \qquad \square
\end{aligned}$$

Proof of Lemma 2.4. Given $\bar{m}^n \in C^0([0,T], \mathcal{P}(\mathbb{T}^d))$, the solution u^{n+1} is uniformly Lipschitz continuous. Hence any weak solution to the Fokker–Planck equation is uniformly Hölder continuous in $C^0([0,T], \mathcal{P}(\mathbb{T}^d))$. This shows that the right-hand side of the Hamilton–Jacobi equation is uniformly Hölder continuous; then the Schauder estimate provide the bound in $C^{1+\alpha/2, 2+\alpha}$ for $\alpha \in (0, 1/2)$, by combining ([19], Thm. 2.2), to get a uniform Holder estimates for Du^{n+1}, with ([19], Thm. 12.1), to obtain the full bound by considering $-H(x, \nabla u^{n+1}(t,x)) + f(x, \bar{m}^n(t))$ as a Holder continuous right-hand side for the heat equation. Plugging this estimate into the Fokker–Planck equation and using again the Schauder estimates gives the bounds in $C^{1+\alpha/2, 2+\alpha}$ on the the m^n. The bound from below for the m^n comes from the strong maximum principle. $\qquad \square$

3. The fictitious play for first order MFG systems

We now consider the first order order MFG system:

$$\begin{cases}
\text{(i)} & -\partial_t u + H(x, \nabla u(t,x)) = f(x, m(t)), \quad (t,x) \in [0,T] \times \mathbb{T}^d \\
\text{(ii)} & \partial_t m + \mathrm{div}(-m D_p H(x, \nabla u(t,x))) = 0, \quad (t,x) \in [0,T] \times \mathbb{T}^d \\
& m(0) = m_0, \ u(x,T) = g(x, m(T)), \quad x \in \mathbb{T}^d.
\end{cases} \tag{3.1}$$

In contrast with second order MFG systems, we cannot expect existence of classical solutions: namely both the Hamilton–Jacobi equation and the Fokker–Planck equation have to be understood in a generalized sense. In particular, the solutions of the Fictitious Play are not smooth enough to justify the various computations of Section 2. For this reason we introduce another method – based on another potential –, which also has the interest that it can be adapted to a finite population of players.

Let us start by recalling the notion of solution for (3.1). Following [22], we say that the pair (u, m) is a solution to the MFG system (3.1) if u is a Lipschitz continuous viscosity solution to (3.1)-(i) while $m \in L^\infty((0,T) \times \mathbb{T}^d)$ is a solution of (3.1)-(ii) in the sense of distribution.

Under our standing assumptions (1.4)–(1.7), there exists at least one solution (u, m) to the mean field game system (3.1). If furthermore (1.9) holds, then the solution is unique (see [22] and Thm. 5.1 in [8]).

3.1. The learning rule and the potential

The learning rule is basically the same as for second order MFG systems: given a smooth initial guess $m^0 : [0,T] \times \mathbb{T}^d \to \mathbb{R}$, we define by induction sequences $u^n, m^n : [0,T] \times \mathbb{T}^d \to \mathbb{R}$ *heuristically* given by:

$$
\begin{aligned}
&\text{(i)} & -\partial_t u^{n+1} + H(x, \nabla u^{n+1}(t,x)) &= f(x, \bar{m}^n(t)), & (t,x) &\in [0,T] \times \mathbb{T}^d \\
&\text{(ii)} & \partial_t m^{n+1} + \operatorname{div}(-m^{n+1} D_p H(x, \nabla u^{n+1})) &= 0, & (t,x) &\in [0,T] \times \mathbb{T}^d \\
& & m^{n+1}(0) = m_0, \ u^{n+1}(x,T) &= g(x, \bar{m}^n(T)), & x &\in \mathbb{T}^d
\end{aligned}
\tag{3.2}
$$

where $\bar{m}^n(t,x) = \frac{1}{n} \sum_{k=1}^n m^k(t,x)$. If equation (3.2)-(i) is easy to interpret, the meaning of (3.2)-(ii) would be more challenging and, actually, would make little sense for a finite population. For this reason we are going to rewrite the problem in a completely different way, as a problem on the space of curves.

Let us fix the notation. Let $\Gamma = C^0([0,T], \mathbb{T}^d)$ be the set of curves. It is endowed with usual topology of the uniform convergence and we denote by $\mathcal{B}(\Gamma)$ the associated σ-field. We define $\mathcal{P}(\Gamma)$ as the set of Borel probability measures on $\mathcal{B}(\Gamma)$. We view Γ and $\mathcal{P}(\Gamma)$ as the set of pure and mixed strategies for the players. For any $t \in [0,T]$ the evaluation map $e_t : \Gamma \to \mathbb{T}^d$, defined by:

$$
e_t(\gamma) = \gamma(t), \quad \forall \gamma \in \Gamma
$$

is continuous and thus measurable. For any $\eta \in \mathcal{P}(\Gamma)$ we define $m^\eta(t) = e_t \sharp \eta$ as the push forward of the measure η to $\mathbb{T}^d$ *i.e.*

$$
m^\eta(t)(A) = \eta(\{\gamma \in \Gamma \mid \gamma(t) \in A\})
$$

for any measurable set $A \subset \mathbb{T}^d$. We denote by $\mathcal{P}_0(\Gamma)$ the set of probability measures on Γ such that $e_0 \sharp \eta = m_0$. Note that $\mathcal{P}_0(\Gamma)$ is the set of strategies compatible with the initial density m_0.

Given an initial time $t \in [0,T]$ and an initial position x, it is convenient to define the cost of a path $\gamma \in C^0([t,T], \mathbb{T}^d)$ payed by a small player starting from that position when the repartition of strategies of the other players is η. It is given by

$$
J(t,x,\gamma,\eta) := \begin{cases} \displaystyle\int_t^T L(\gamma(s), \dot{\gamma}(s)) + f(\gamma(s), m^\eta(s)) \mathrm{d}s + g(\gamma(T), m^\eta(T)) & \text{if } \gamma \in H^1([t,T], \mathbb{T}^d) \\ +\infty & \text{otherwise.} \end{cases}
$$

where $L(x,v) := H^*(x,-v)$ and H^* is the Fenchel conjugate of H with respect to the last variable. If $t = 0$, we simply abbreviate $J(x,\gamma,\eta) := J(0,x,\gamma,\eta)$. We note for later use that $J(t,x,\cdot,\eta)$ is lower semi-continuous on Γ.

We now define the Fictitious Play. We start with an initial configuration $\eta^0 \in \mathcal{P}_0(\Gamma)$ (the belief before the first step of a typical player on the actions of the other players). We now build by induction the sequences (θ^n) and (η^n) of $\mathcal{P}(\Gamma)$, η^n being interpreted as the belief at the end of stage n of a typical player on the actions of the other agents and θ^{n+1} the repartition of strategies of the players when they play optimally in the game against η^n. More precisely, for any $x \in \mathbb{T}^d$, let $\bar{\gamma}_x^{n+1} \in H^1([0,T], \mathbb{T}^d)$ be an optimal solution to

$$
\inf_{\gamma \in H^1, \ \gamma(0)=x} J(x,\gamma,\eta^n).
$$

In view of our coercivity assumptions on H and the definition of L, the optimum is known to exist. Moreover, by the measurable selection theorem we can (and will) assume that the map $x \to \bar{\gamma}_x^{n+1}$ is Borel measurable. We then consider the measure $\theta^{n+1} \in P_0(\Gamma)$ defined by

$$
\theta^{n+1} := \bar{\gamma}_\cdot^{n+1} \sharp m_0 \qquad \forall t \in [0,T]
$$

and set

$$
\eta^{n+1} := \frac{1}{n+1} \sum_{k=1}^{n+1} \theta^k = \eta^n + \frac{1}{n+1}(\theta^{n+1} - \eta^n).
\tag{3.3}
$$

As in Section 2, we assume that our MFG is potential, *i.e.*, that there exists of potential functions $F, G :$ $\mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that:

$$f(x, m) = \frac{\delta F}{\delta m}(x, m), \quad g(x, m) = \frac{\delta G}{\delta m}(x, m). \tag{3.4}$$

Here is our main convergence result.

Theorem 3.1. *Assume that* (1.4)–(1.7) *and* (3.4) *hold. Then the sequences* (η^n, θ^n) *is pre-compact in* $\mathcal{P}(\Gamma) \times \mathcal{P}(\Gamma)$ *and any cluster point* $(\bar{\eta}, \bar{\theta})$ *satisfies the following:* $\bar{\theta} = \bar{\eta}$ *and, if we set*

$$\bar{m}(t) := e_t \sharp \bar{\eta}, \quad \bar{u}(t, x) = \inf_{\gamma \in H^1, \ \gamma(t) = x} J(t, x, \gamma, \bar{\eta}), \tag{3.5}$$

then the pair $(\bar{u}, \bar{m})$ *is a solution to the MFG system* (3.1).

If furthermore (1.9) *holds, then the entire sequence* (η^n, θ^n) *converges.*

The proof of Theorem 3.1 is postponed to the next subsection. As for the second order problem, the key idea is that our MFG system has a potential. However, in contrast with the second order case, the potential is now written on the space of probability on curves and reads, for $\eta \in \mathcal{P}(\Gamma)$,

$$\Phi(\eta) := \int_\Gamma \int_0^T L(\gamma(t), \dot{\gamma}(t)) \ \mathrm{d}t \mathrm{d}\eta(\gamma) + \int_0^T F(e_t \sharp \eta) \ \mathrm{d}t + G(e_T \sharp \eta). \tag{3.6}$$

Note that $\Phi(\eta)$ is well-defined and belongs to $(-\infty, +\infty]$. The potential defined above is reminiscent of [9] or [7]. For instance, in [7] – but for MFG system with a local dependence and under the monotonicity condition (1.9) – it is proved that the MFG equilibrium can be found as a global minimum of Φ. We will show in the proof of Theorem 3.1 that the limit measure $\bar{\eta}$ is characterized by the optimality condition

$$\frac{\delta \Phi}{\delta m}(\bar{\eta})(\bar{\eta}) \leq \frac{\delta \Phi}{\delta m}(\bar{\eta})(\theta) \qquad \forall \theta \in \mathcal{P}(\Gamma).$$

Before proving that Φ is a potential for the game, let us start with preliminary remarks. The first one explains that the optimal curves are uniformly Lipschitz continuous.

Lemma 3.2. *There exists a constant* $C > 0$ *such that, for any* $x \in \mathbb{T}^d$ *and any* $n \geq 0$,

$$\|\dot{\bar{\gamma}}_x^{n+1}\|_\infty \leq C. \tag{3.7}$$

In particular, the sequences (η^n) *and* (θ^n) *are tight and*

$$\mathbf{d}_1(e_t \sharp \eta^{n+1}, e_{t'} \sharp \eta^{n+1}) \leq C|t - t'| \qquad \forall t, t' \in [0, T].$$

Proof. Under our assumption on H, f and g, it is known that the (u^n) are uniformly Lipschitz continuous (see, for instance, the appendix of [8]). As a byproduct the optimal solutions are also uniformly Lipschitz continuous thanks to the classical link between the derivative of the value function and the optimal trajectories (Thm. 6.4.8 of [6]): this is (3.7). The rest of the proof is a straightforward consequence of (3.7). □

Next we compute the derivative of Φ with respect to the measure η. Let us point out that, since Φ is not continuous and can take the value $+\infty$, the derivative, although defined by the formula (1.10), has to be taken only at points and direction along which Φ is finite. This is in particular the case for the η^n and the θ^n.

Lemma 3.3. *For any* $\eta, \eta' \in \mathcal{P}(\Gamma)$ *such that* $\Phi(\eta), \Phi(\eta') < +\infty$, *we have*

$$\frac{\delta \Phi}{\delta \eta}(\eta)(\eta' - \eta) \ = \int_\Gamma J(\gamma(0), \gamma, \eta) \ d(\eta' - \eta)(\gamma).$$

Proof. This is a straightforward application of the definition of Φ in (3.6) and of the continuous derivability of F and G. □

By abuse of notation, we also define $\frac{\delta\Phi}{\delta\eta}(\eta)(\theta)$ for a positive Borel measure θ on Γ by setting

$$\frac{\delta\Phi}{\delta\eta}(\eta)(\theta) = \int_\Gamma J(\gamma(0), \gamma, \eta) \mathrm{d}\theta(\gamma).$$

Note that, as J is bounded below, the quantity $\frac{\delta\Phi}{\delta\eta}(\eta)(\theta)$ is well-defined and belongs to $(-\infty, +\infty]$.

Next we translate the optimality property of $\bar\gamma_x^n$ to an optimality property of η^n.

Lemma 3.4. *For any $n \in \mathbb{N}^*$,*

$$\frac{\delta\Phi}{\delta\eta}(\eta^n)(\theta^{n+1}) = \int_{\mathbb{T}^d} J(x, \bar\gamma_x^{n+1}, \eta^n) m_0(x) \mathrm{d}x = \min_{\theta \in \mathcal{P}_0(\Gamma)} \frac{\delta\Phi}{\delta\eta}(\eta^n)(\theta).$$

Proof. The first equality is just the definition of θ^{n+1}. It remains to check that, for any $\theta \in \mathcal{P}_0(\Gamma)$,

$$\int_{\mathbb{T}^d} J(x, \bar\gamma_x^{n+1}, \eta^n) m_0(x) \mathrm{d}x \leq \int_\Gamma J(\gamma(0), \gamma, \eta^n) \mathrm{d}\theta(\gamma).$$

As $m_0 = e_0 \sharp \theta$, we can disintegrate θ into $\theta = \int_{\mathbb{T}^d} \theta_x \mathrm{d}m_0(x)$, where $\theta_x \in \mathcal{P}(\Gamma)$ with $\gamma(0) = x$ for θ_x-a.e. γ. By optimality of $\bar\gamma_x^{n+1}$ we have, for m_0-a.e. $x \in \mathbb{T}^d$,

$$J(x, \bar\gamma_x^{n+1}, \eta^n) \leq \int_\Gamma J(x, \gamma, \eta^n) \, \mathrm{d}\theta_x(\gamma)$$

and therefore, integrating with respect to m_0:

$$\int_{\mathbb{T}^d} J(x, \bar\gamma_x^{n+1}, \eta^n) m_0(x) \mathrm{d}x \leq \int_{\mathbb{T}^d} \int_\Gamma J(x, \gamma, \eta^n) \, \mathrm{d}\theta_x(\gamma) m_0(x) \mathrm{d}x = \int_\Gamma J(\gamma(0), \gamma, \eta^n) \mathrm{d}\theta(\gamma). \qquad □$$

The next proposition states that the potential Φ is indeed almost decreasing along the sequence (η^n).

Proposition 3.5. *There is a constant $C > 0$ such that, for any $n \in \mathbb{N}^*$, we have*

$$\Phi(\eta^{n+1}) \leq \Phi(\eta^n) + \frac{1}{n+1} \frac{\delta\Phi}{\delta\eta}(\eta^n)(\theta^{n+1} - \eta^n) + \frac{C}{(n+1)^2} \tag{3.8}$$

where

$$\frac{\delta\Phi}{\delta\eta}(\eta^n)(\theta^{n+1} - \eta^n) = \int_\Gamma J(\gamma(0), \gamma, \eta^n) \, \mathrm{d}(\theta^{n+1} - \eta^n)(\gamma) \leq 0. \tag{3.9}$$

Proof. Recalling (3.3), we have

$$\Phi(\eta^{n+1}) - \Phi(\eta^n) = \int_0^1 \frac{\delta\Phi}{\delta\eta}((1-s)\eta^n + s\eta^{n+1})(\eta^{n+1} - \eta^n) \mathrm{d}s$$

$$= \frac{1}{(n+1)} \int_0^1 \frac{\delta\Phi}{\delta\eta}((1-s)\eta^n + s\eta^{n+1})(\theta^{n+1} - \eta^n) \mathrm{d}s. \tag{3.10}$$

Let us estimate the right-hand side of the inequality. For any $s \in [0, 1]$, Lemma 3.3 states that

$$\frac{\delta\Phi}{\delta\eta}((1-s)\eta^n + s\eta^{n+1})(\theta^{n+1} - \eta^n) = \int_\Gamma J(\gamma(0), \gamma, (1-s)\eta^n + s\eta^{n+1})) \mathrm{d}(\theta^{n+1} - \eta^n)(\gamma)$$

$$= \int_\Gamma J(\gamma(0), \gamma, \eta^n) \mathrm{d}(\theta^{n+1} - \eta^n)(\gamma) + R(s) \tag{3.11}$$

where, by the definition of J and Lipschitz continuity of f and g,

$$
\begin{aligned}
R(s) = &\int_\Gamma \int_0^T \left(f(\gamma(t), e_t \sharp ((1-s)\eta^n + s\eta^{n+1})) - f(\gamma(t), e_t \sharp \eta^n) \right) \mathrm{d}t \mathrm{d}(\theta^{n+1} - \eta^n)(\gamma) \\
&+ \int_\Gamma \left(g(\gamma(T), e_T \sharp ((1-s)\eta^n + s\eta^{n+1})) - g(\gamma(T), e_T \sharp \eta^n) \right) \mathrm{d}(\theta^{n+1} - \eta^n)(\gamma) \\
\leq\; &C \sup_{t \in [0,T]} \mathbf{d}_1 \left(e_t \sharp ((1-s)\eta^{n+1} + s\eta^n)), e_t \sharp \eta^n \right).
\end{aligned} \tag{3.12}
$$

Note that, by the definition of $\mathbf{d}_1$, we have for any $t \in [0, T]$,

$$
\begin{aligned}
\mathbf{d}_1 (&e_t \sharp ((1-s)\eta^{n+1} + s\eta^n)), e_t \sharp \eta^n) \\
&\leq \sup_\xi \int_{\mathbb{T}^d} \xi(x) \mathrm{d}(e_t \sharp ((1-s)\eta^{n+1} + s\eta^n)(x) - \int_{\mathbb{T}^d} \xi(x) \; d(e_t \sharp \eta^n)(x) \\
&\leq (1-s) \sup_\xi \int_{\mathbb{T}^d} \xi(x) \; d(e_t \sharp \eta^{n+1})(x) - \int_{\mathbb{T}^d} \xi(x) \; d(e_t \sharp \eta^n)(x) \\
&\leq \frac{(1-s)}{n+1} \sup_\xi \int_{\mathbb{T}^d} \xi(x) \; d(e_t \sharp (\theta^{n+1} - \eta^n))(x) \\
&\leq \frac{(1-s)}{n+1} \sup_\xi \int_{\mathbb{T}^d} (\xi(x) - \xi(0)) \; d(e_t \sharp \theta^{n+1} - e_t \sharp \eta^n)(x) \leq \frac{C}{n+1},
\end{aligned}
$$

where the supremum is taken over the set of Lipschitz maps $\xi : \mathbb{T}^d \to \mathbb{R}$ with Lipschitz constant not larger than 1. Therefore

$$
\Phi(\eta^{n+1}) - \Phi(\eta^n) \leq \frac{1}{(n+1)} \int_\Gamma J(\gamma(0), \gamma, \eta^n) \; d(\theta^{n+1} - \eta^n)(\gamma) + \frac{C}{(n+1)^2},
$$

where the first term in the right-hand side is nonpositive thanks to Lemma 3.4.   □

3.2. Convergence of the fictitious play

In this subsection, we prove Theorem 3.1. Recall that Lemma 3.2 states that the sequence (η^n) is tight. We next Lemma characterizes the cluster distribution:

Lemma 3.6. *Any cluster point $\bar{\eta}$ of the sequence (η^n) satisfies*

$$
\frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\bar{\eta}) \leq \frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\theta) \qquad \forall \theta \in P_0(\Gamma), \tag{3.13}
$$

which means that $\bar{\eta}$-a.e. γ is optimal for the map $\tilde{\gamma} \to J(\gamma(0), \tilde{\gamma}, \bar{\eta})$ under the constraint $\tilde{\gamma}(0) = \gamma(0)$.

Proof. Let us define:

$$
\begin{aligned}
a^{n+1} :&= -\frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta^{n+1} - \eta^n) = -\int_\Gamma J(\gamma(0), \gamma, \eta^n) \mathrm{d}(\theta^{n+1} - \eta^n) \\
&= \int_\Gamma J(\gamma(0), \gamma, \eta^n) \mathrm{d}\eta^n(\gamma) - \min_{\theta \in \mathcal{P}_0(\mathbb{T}^d)} \int_\Gamma J(\gamma(0), \gamma, \eta^n) \mathrm{d}\theta(\gamma),
\end{aligned}
$$

where the last equality come from Lemma 3.4. Then according to Proposition 3.5 the sequence (a^n) is non-negative and, by (3.8), the quantity $\sum_k a^k / k$ is finite (because Φ is bounded below). Therefore by Lemma 2.7 we have:

$$
\lim_{N \to +\infty} \frac{1}{N} \sum_{k=1}^N a^k = 0. \tag{3.14}
$$

Let us now check that $|a^{n+1} - a^n| \leq C/n$ for some constant C. By arguments similar to the ones in the proof of Proposition 3.5, we have, for any $\theta \in P_0(\Gamma)$,

$$\left| \frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta) - \frac{\delta \Phi}{\delta \eta}(\eta^{n+1})(\theta) \right| \leq \frac{C}{n}. \tag{3.15}$$

On the other hand, by optimality of θ^{n+1} and θ^{n+2} in Lemma 3.4 and (3.15), we have

$$\frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta^{n+1}) = \min_{\theta \in \mathcal{P}_0(\mathbb{T}^d)} \int_\Gamma J(\gamma(0), \gamma, \eta^n) \mathrm{d}\theta(\gamma) \leq \int_\Gamma J(\gamma(0), \gamma, \eta^n) \mathrm{d}\theta^{n+2}(\gamma)$$

$$\leq \int_\Gamma J(\gamma(0), \gamma, \eta^{n+1}) \mathrm{d}\theta^{n+2}(\gamma) + C/n = \frac{\delta \Phi}{\delta \eta}(\eta^{n+1})(\theta^{n+2}) + C/n$$

$$= \min_{\theta \in \mathcal{P}_0(\mathbb{T}^d)} \int_\Gamma J(\gamma(0), \gamma, \eta^{n+1}) \mathrm{d}\theta(\gamma) + C/n$$

$$\leq \int_\Gamma J(\gamma(0), \gamma, \eta^{n+1}) \mathrm{d}\theta^{n+1}(\gamma) + C/n = \frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta^{n+1}) + C/n,$$

which proves that

$$\left| \frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta^{n+1}) - \frac{\delta \Phi}{\delta \eta}(\eta^{n+1})(\theta^{n+2}) \right| \leq C/n.$$

So we have:

$$|a^{n+1} - a^n| = \left| \frac{\delta \Phi}{\delta \eta}(\eta^n)(\eta^n - \theta^{n+1}) - \frac{\delta \Phi}{\delta \eta}(\eta^{n+1})(\eta^{n+1} - \theta^{n+2}) \right|$$

$$\leq \left| \frac{\delta \Phi}{\delta \eta}(\eta^n)(\eta^n) - \frac{\delta \Phi}{\delta \eta}(\eta^{n+1})(\eta^{n+1}) \right| + \left| \frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta^{n+1}) - \frac{\delta \Phi}{\delta \eta}(\eta^{n+1})(\theta^{n+2}) \right|$$

$$\leq \left| \frac{\delta \Phi}{\delta \eta}(\eta^n)(\eta^n - \eta^{n+1}) \right| + C/n = \frac{1}{n+1} \left| \frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta^{n+1} - \eta^n) \right| + C/n \leq C/n.$$

By (3.14) and the above estimate, we conclude that $a_n \to 0$ thanks to Lemma 2.7.

Let now $\bar{\eta}$ be any cluster point of the sequence (η^n). Let us check that (3.13) holds. Let $\theta \in \mathcal{P}_0(\mathbb{T}^d)$. Then, from Lemma 3.4, for every $n \in \mathbb{N}$ we have:

$$\frac{\delta \Phi}{\delta \eta}(\eta^n)(\eta^n) - a_n = \frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta^{n+1}) \leq \frac{\delta \Phi}{\delta \eta}(\eta^n)(\theta).$$

If $(\eta^{n_i})_{i \in \mathbb{N}}$ is such that $\eta^{n_i} \to \bar{\eta}$, then:

$$\forall \gamma \in \Gamma : \quad |J(\gamma, \gamma(0), \bar{\eta}) - J(\gamma, \gamma(0), \eta^{n_i})| \leq K \sup_{t \in [0,T]} \mathbf{d}_1(e_t \sharp \eta^{n_i}, e_t \sharp \bar{\eta}),$$

where the last term tends to 0 because the maps $t \to e_t \sharp \eta^{n_i}$ are uniformly continuous (from Lem. 3.2) and converges pointwisely (and thus uniformly) to $t \to e_t \sharp \bar{\eta}$. This yields that $(\frac{\delta \Phi}{\delta \eta}(\eta^{n_i})(\theta))$ converges to $\frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\theta)$. On the other hand, by lower semicontinuity of the map $\gamma \to J(\gamma, \gamma(0), \bar{\eta})$ on Γ, we have

$$\frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\bar{\eta}) \leq \liminf \frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\eta^{n_i}) = \liminf \frac{\delta \Phi}{\delta \eta}(\eta^{n_i})(\eta^{n_i}),$$

which proves (3.13).

Let us check that $\bar{\eta}$-a.e. γ is optimal for the map $\tilde{\gamma} \to J(\gamma(0), \tilde{\gamma}, \bar{\eta})$ under the constraint $\tilde{\gamma}(0) = \gamma(0)$. Let $\theta = \int_{\mathbb{T}^d} \delta_{\bar{\gamma}_x} m_0(x) \mathrm{d}x$ where $\bar{\gamma}_x$ is (a measurable selection of) an optimal solution for $\tilde{\gamma} \to J(x, \tilde{\gamma}, \bar{\eta})$ under the constraint $\tilde{\gamma}(0) = x$. If we disintegrate $\bar{\eta}$ into $\bar{\eta} = \int_{\mathbb{T}^d} \bar{\eta}_x m_0(x) \mathrm{d}x$, then, for m_0-a.e. x and $\bar{\eta}_x$-a.e. γ we have

$$J(x, \bar{\gamma}_x, \bar{\eta}) \leq J(x, \gamma, \bar{\eta}). \tag{3.16}$$

Integrating over $\bar{\eta}_x$ and then against m_0 then implies that

$$\frac{\delta \Phi}{\delta m}(\bar{\eta})(\theta) = \int_{\mathbb{T}^d} J(x, \bar{\gamma}_x, \bar{\eta}) m_0(x) \mathrm{d}x \leq \int_{\Gamma} J(\gamma(0), \gamma, \bar{\eta}) \mathrm{d}\bar{\eta}(\gamma) = \frac{\delta \Phi}{\delta m}(\bar{\eta})(\bar{\eta}).$$

As the reverse inequality always holds, this proves that there must be an equality in (3.16) a.e., which proves the claim. $\qquad \square$

Proof of Theorem 3.1. Let $(\bar{\eta}, \bar{\theta})$ be the limit of a converging subsequence $(\eta^{n_i}, \theta^{n_i})$. We set

$$\bar{u}(t, x) := \inf_{\gamma \in \Gamma, \ \gamma(t) = x} J(t, x, \gamma, \bar{\eta}) \qquad \text{and} \qquad \bar{m}(t) := e_t \sharp \bar{\eta}.$$

By standard argument in optimal control, we know that $\bar{u}$ is a viscosity solution to (3.1)-(i) with terminal condition $\bar{u}(T, x) = g(x, \bar{m}(T))$. Moreover, $\bar{u}$ is Lipschitz continuous and semiconcave (*cf.* for instance Lem. 5.2 in [8]).

It remains to check that $\bar{m}$ satisfies (3.1)-(ii). By Lemma 3.6, we know that

$$\frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\bar{\eta}) \leq \frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\theta) \qquad \forall \theta \in P_0(\Gamma),$$

which means that $\bar{\eta}$-a.e. γ is optimal for the map $\tilde{\gamma} \to J(\gamma(0), \tilde{\gamma}, \bar{\eta})$ under the constraint $\tilde{\gamma}(0) = \gamma(0)$. Following Theorem 6.4.9 in [6], the optimal solution for $J(x, \cdot, \bar{\eta})$ is unique at any point of differentiability of $\bar{u}(0, \cdot)$ (let us call it $\bar{\gamma}_x$). Disintegrating $\bar{\eta}$ into $\bar{\eta} = \int_{\mathbb{T}^d} \bar{\eta}_x \mathrm{d}m_0(x)$, we have therefore, since m_0 is absolutely continuous,

$$\bar{\eta}_x = \delta_{\bar{\gamma}_x} \qquad \text{for } m_0\text{-a.e. } x \in \mathbb{T}^d,$$

so that

$$\bar{\eta} = \int_{\mathbb{T}^d} \delta_{\bar{\gamma}_x} m_0(x) \mathrm{d}x \qquad \text{and} \qquad \bar{m}(t) = \bar{\gamma}.(t) \sharp m_0 \qquad \forall t \in [0, T]. \tag{3.17}$$

Let us also recall that the derivative of $\bar{u}(t, \cdot)$ exists along the optimal solution $\bar{\gamma}_x$ and that

$$\dot{\bar{\gamma}}_x(t) = -D_p H(\bar{\gamma}_x(t), \nabla \bar{u}(t, \bar{\gamma}_x(t)) \qquad \forall t \in (0, T]$$

(see Thms. 6.4.7 and 6.4.8 of [6]). This proves that $\bar{m}$ is a solution in the sense of distribution of (3.1)-(ii) (where we denote by $\nabla \bar{u}$ any fixed Borel measurable selection of the map $(t, x) \to D^* u(t, x)$, the set of reachable gradients of u at (t, x), see [6]). Proposition A.1 in appendix states that (3.1)-(ii) has a unique solution and that this solution has a density in L^∞: thus $\bar{m}$ is in L^∞, which shows that the pair $(\bar{u}, \bar{m})$ is a solution of the MFG system (3.1).

In order to identify the cluster point $\bar{\theta}$, let us recall that θ^n is defined by

$$\theta^n = \bar{\gamma}.^n \sharp m_0,$$

where, for any $x \in \mathbb{T}^d$, $\bar{\gamma}_x^n$ is a minimum of $J(x, \cdot, \eta^n)$ under the constraint $\gamma(0) = x$. As the criterion $J(x, \cdot, \eta^{n_i})$ Γ-converges to $J(x, \cdot, \bar{\eta})$ and since at any point of differentiability of $\bar{u}(0, \cdot)$ the optimal solution $\bar{\gamma}_x$ is unique, standard compactness arguments show that $(\bar{\gamma}_x^{n_i})$ converges to $\bar{\gamma}_x$ for a.e. $x \in \mathbb{T}^d$. Therefore (θ^{n_i}) converges to $\bar{\gamma}.\sharp m_0$, which is nothing but $\bar{\eta}$ by (3.17). So we conclude that $\bar{\theta} = \bar{\eta}$.

Finally, if (1.9) holds, then we claim that $\bar{\eta}$ is independent of the chosen subsequence. Indeed, since from its very definition the dependence with respect to $\bar{\eta}$ of $J(x, \gamma, \bar{\eta})$ is only through the family of measures $(\bar{m}(t) = e_t \sharp \bar{\eta})$ and since, by (1.9), there exists a unique solution to the MFG system and thus $\bar{m}$ is uniquely defined, $J(x, \gamma, \bar{\eta})$ is independent of the choice of the subsequence. Then $\bar{\gamma}_x$ defined above is also independent of the subsequence, which characterizes $\bar{\eta}$ in a unique way thanks to (3.17). Therefore the entire sequence (η^n, θ^n) converges to $(\bar{\eta}, \bar{\eta})$. $\hfill\square$

Remark 3.7. The proof shows that a measure $\bar{\eta} \in \mathcal{P}_0(\Gamma)$ which satisfies (3.13) can be understood as the representation of a MFG equilibrium. Indeed, if we define $(\bar{u}, \bar{m})$ as in (3.5), then $(\bar{u}, \bar{m})$ is a solution to the MFG system (3.1). Conversely, if $(\bar{u}, \bar{m})$ is a solution to the MFG system (3.1), then the relation (3.17) identifies uniquely a measure $\bar{\eta} \in \mathcal{P}_0(\Gamma)$. For this reason, we call such a measure an equilibrium measure.

3.3. The learning procedure in N-players first order games

In this part we show that the Fictitious Play in the Mean Field Game with large (but finite) number of players $N \in \mathbb{N}$ converges in some sense to the equilibrium of our Mean Field Game with infinite number of players. For every $N \in \mathbb{N}$, fix a sequence of initial states $x_1^N, x_2^N, \ldots, x_N^N \in \mathbb{T}^d$ such that:

$$\lim_{N \to \infty} \mathbf{d}_1(m_0^N, m_0) = 0$$

where $m_0^N = \frac{1}{N} \sum_{i+1}^N \delta_{x_i^N}$ is the empirical measure associated with the $\{x_i^N\}_{i=1,\ldots,N}$. As in the case of an infinite population, let us define the sequences $\eta^{n,N}, \theta^{n,N} \in \mathcal{P}(\Gamma)$, for $n \in \mathbb{N}^*$ in the following way:

$$\eta^{n+1,N} = \frac{1}{n+1} \left(\theta^{1,N} + \theta^{2,N} + \cdots + \theta^{n+1,N} \right)$$

$$\theta^{n+1,N} = \frac{1}{N} \left(\delta_{\gamma_{x_1^N}^{n+1,N}} + \delta_{\gamma_{x_2^N}^{n+1,N}} + \cdots + \delta_{\gamma_{x_N^N}^{n+1,N}} \right) \tag{3.18}$$

where $\gamma_{x_i^N}^{n+1,N}$ is an optimal path which minimizes $J(x_i^N, \cdot, \eta^{n,N})$. As before one can show that if

$$a^{n+1,N} := -\frac{\delta \Phi}{\delta \eta}(\eta^{n,N})(\theta^{n+1,N} - \eta^{n,N}) = -\int_\Gamma J(\gamma(0), \gamma, \eta^{n,N}) \mathrm{d}(\theta^{n+1,N} - \eta^{n,N})(\gamma)$$

$$= \int_\Gamma J(\gamma(0), \gamma, \eta^{n,N}) \mathrm{d}\eta^{n,N}(\gamma) - \min_{\theta \in \mathcal{P}(\Gamma), e_0 \sharp \theta = m_0^N} \int_\Gamma J(\gamma(0), \gamma, \eta^{n,N}) \mathrm{d}\theta(\gamma),$$

then we have $\lim_{n \to \infty} a^{n,N} = 0$. This proves that any accumulation distribution $\bar{\eta}^N$ of the sequence $\{\eta^{n,N}\}_{n \in \mathbb{N}^*}$ satisfies:

$$\int_\Gamma J(\gamma(0), \gamma, \bar{\eta}^N) \mathrm{d}\bar{\eta}^N(\gamma) = \min_{\theta \in \mathcal{P}(\Gamma), e_0 \sharp \theta = m_0^N} \int_\Gamma J(\gamma(0), \gamma, \bar{\eta}^N) \mathrm{d}\theta(\gamma). \tag{3.19}$$

So if $\bar{\eta}^N = \frac{1}{N}(\bar{\eta}_{x_1}^N + \bar{\eta}_{x_2}^N + \cdots + \bar{\eta}_{x_N}^N)$ then

$$\mathrm{supp}(\bar{\eta}_{x_i}^N) \subseteq \mathrm{argmin}_{\gamma(0) = x_i} J(x_i, \gamma, \bar{\eta}^N).$$

Note that, in contrast with the case of an infinite population, this is not an equilibrium condition, since the deviation of a player changes the measure $\bar{\eta}^N$ as well.

In the following Theorem we prove that any accumulation point $\bar{\eta}$ of $\{\bar{\eta}_N\}$ satisfies:

$$\int_\Gamma J(\gamma(0), \gamma, \bar{\eta}) \mathrm{d}\bar{\eta}(\gamma) = \min_{\theta \in \mathcal{P}_0(\Gamma)} \int_\Gamma J(\gamma(0), \gamma, \bar{\eta}) \mathrm{d}\theta(\gamma), \tag{3.20}$$

where $\mathcal{P}_0(\Gamma)$ is the set of measure $\theta \in \mathcal{P}(\Gamma)$ such that $e_0 \sharp \theta = m_0$. We have seen in Remark 3.7 that this condition characterizes an MFG equilibrium.

Theorem 3.8. *Assume that* (1.4)–(1.7) *and* (3.4) *hold. Consider the Fictitious Play for the N-player game as described in* (3.18) *and let* $\bar{\eta}^N$ *by an accumulation distribution of* $(\eta^{n,N})_{n\in\mathbb{N}}$. *Then every accumulation point of pre-compact set of* $\{\bar{\eta}^N\}_{N\in\mathbb{N}}$ *is an MFG equilibrium.*

If furthermore the monotonicity condition (1.9) *holds, then* $(\bar{\eta}^N)$ *has a limit which is the MFG equilibrium.*

Proof. Consider $\bar{\eta}$ as an accumulation point of the set $\{\bar{\eta}^N\}_{N\in\mathbb{N}}$. It is sufficient to show that for every $\theta \in \mathcal{P}(\Gamma)$ such that $e_0\sharp\theta = m_0$, we have

$$\int_\Gamma J(\gamma(0),\gamma,\bar{\eta})\mathrm{d}\bar{\eta}(\gamma) \leq \int_\Gamma J(\gamma(0),\gamma,\bar{\eta})\mathrm{d}\theta(\gamma). \tag{3.21}$$

Since m_0 is absolutely continuous with respect to the Lebesgue measure, there exists an optimal transport map $\tau_N : \mathbb{T}^d \to \mathbb{T}^d$ such that:

$$\tau_N\sharp m_0 = m_0^N, \quad \mathbf{d}_1(m_0,m_0^N) = \int_{\mathbb{T}^d} |x - \tau_N(x)|\mathrm{d}m_0(x)$$

(see [4]). We define the functions $\xi_N : \Gamma \to \Gamma$ as follows:

$$\xi_N(\gamma) = \gamma - \gamma(0) + \tau_N(\gamma(0))$$

and set $\theta^N = \xi_N\sharp\theta$. Then we have

$$e_0\sharp\theta^N = e_0\sharp(\xi_N\sharp\theta) = (e_0\circ\xi_N)\sharp\theta = (\tau_N\circ e_0)\sharp\theta = \tau_N\sharp(e_0\sharp\theta) = \tau_N\sharp m_0 = m_0^N.$$

Then the characterization (3.19) of $\bar{\eta}^N$ yields:

$$\int_\Gamma J(\gamma(0),\gamma,\bar{\eta}^N)\mathrm{d}\bar{\eta}^N(\gamma) \leq \int_\Gamma J(\gamma(0),\gamma,\bar{\eta}^N)\mathrm{d}\theta^N(\gamma). \tag{3.22}$$

By lower semicontinuity of J we have

$$\int_\Gamma J(\gamma(0),\gamma,\bar{\eta})\mathrm{d}\bar{\eta}(\gamma) \leq \liminf_N \int_\Gamma J(\gamma(0),\gamma,\bar{\eta}^N)\mathrm{d}\bar{\eta}^N(\gamma).$$

On the other hand, by the definition of ξ^N and θ^N and the decomposition $\theta = \int_{\mathbb{T}^d}\theta_x m_0(x)\mathrm{d}x$, we have

$$\int_\Gamma J(\gamma(0),\gamma,\bar{\eta}^N)\mathrm{d}\theta^N(\gamma) = \int_{\mathbb{T}^d}\int_\Gamma (\int_0^T L(\gamma(t)-\gamma(0)+\tau_N(\gamma(0)),\dot{\gamma}(t)) + f(\gamma(t)-\gamma(0)+\tau_N(\gamma(0)),e_t\sharp\bar{\eta}^N)\,\mathrm{d}t$$
$$+ g(\gamma(t)-\gamma(0)+\tau_N(\gamma(0)),e_T\sharp\bar{\eta}^N))m_0(x)\mathrm{d}\theta_x(\gamma)\mathrm{d}x,$$

where, by dominate convergence, the right-hand side converges to the right-hand side of (3.21). So letting $N \to \infty$ in (3.22) gives exactly (3.21).

Under (1.9), the MFG equilibrium is unique. Hence, for any $\epsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that for any $N > N_\varepsilon$ and any accumulation point $\bar{\eta}^N$ we have $\mathbf{d}_1(\bar{\eta},\bar{\eta}^N) < \epsilon$. $\qquad\square$

Corollary 3.9. *Assume* (1.4)–(1.7) *and* (3.4) *and* (1.9). *Then, for any* $\epsilon > 0$ *there is* $N_\varepsilon \in \mathbb{N}$ *such that for any* $N > N_\varepsilon$,

$$\exists n(N,\epsilon) \in \mathbb{N}: \quad \forall n > n(N,\epsilon): \quad \mathbf{d}_1(\eta^{n,N},\bar{\eta}) < \epsilon,$$

where $\bar{\eta}$ *is the MFG equilibrium. In other words, for every* $\epsilon > 0$, *one can reach to the* ϵ-*neighborhood of the equilibrium point if the number of players* N *is large enough.*

APPENDIX A. WELL-POSEDNESS OF A CONTINUITY EQUATION

We consider the continuity equation

$$\begin{cases} \partial_t m - \operatorname{div}(m D_p H(x, \nabla \bar{u})) = 0 & \text{in } (0, T) \times \mathbb{T}^d \\ m(0, x) = m_0(x). \end{cases} \tag{A.1}$$

where $\bar{u}$ is the viscosity solution to

$$\begin{cases} -\partial_t u + H(x, \nabla u(t, x)) = f(x, \bar{m}(t)), & (t, x) \in [0, T] \times \mathbb{T}^d \\ u(T, x) = g(x, m(T)), & x \in \mathbb{T}^d. \end{cases}$$

Let us recall that $\bar{u}$ is semi-concave. In (A.1) we denote by $\nabla \bar{u}$ any fixed Borel measurable selection of the map $(t, x) \to D^* u(t, x)$ (the set of reachable gradients of u at (t, x), see [6]). The section is devoted to the proof of the following statement.

Proposition A.1. *Given a fixed map $\bar{u}$, there exists a unique solution $\bar{m}$ of (A.1) in the sense of distribution. Moreover $\bar{m}$ is absolutely continuous and satisfies*

$$\sup_{t \in [0, T]} \|\bar{m}(t, \cdot)\|_\infty \leq C.$$

The difficulty for the proof comes from the fact that the vector field $-D_p H(t, x, \nabla u)$ is not smooth: it is even discontinuous in general. The analysis of transport equations with non smooth vector fields has attracted a lot of attention since the DiPerna–Lions seminal paper [13]. We face here a simple situation where the vector field generates almost everywhere a unique solution. Nevertheless uniqueness of solution of the associated continuity equation requires the combination of several arguments. We rely here on Ambrosio's approach [3], in particular for the "superposition principle" (see Thm. A.3 below).

Let us start with the existence of a bounded solution to (A.1): this is the easy part.

Lemma A.2. *There exists a solution to (A.1) which belongs to L^∞.*

Proof. We follow (at least partially) the perturbation argument given in the proof of Theorem 5.1 of [8]. For $\varepsilon > 0$, let $(u^\varepsilon, m^\varepsilon)$ be the unique classical solution to

$$\begin{cases} -\partial_t u^\varepsilon - \varepsilon \Delta u^\varepsilon + H(x, \nabla u^\varepsilon) = f(x, \bar{m}(t)) & \text{in } (0, T) \times \mathbb{T}^d \\ \partial_t m^\varepsilon - \varepsilon \Delta m^\varepsilon - \operatorname{div}(m^\varepsilon D_p H(x, \nabla u^\varepsilon)) = 0 & \text{in } (0, T) \times \mathbb{T}^d \\ m^\varepsilon(0, x) = m_0(x), \ u^\varepsilon(T, x) = g(x, \bar{m}(t)) & \text{in } \mathbb{T}^d. \end{cases}$$

Following the same argument as in [8], we know that the (m^ε) are uniformly bounded in L^∞: there exists $C > 0$ such that

$$\|m^\varepsilon\|_\infty \leq C \qquad \forall \varepsilon > 0.$$

Moreover (by semi-concavity) the (∇u^ε) are uniformly bounded and converge a.e. to $\nabla \bar{u}$ as ε tends to 0. Letting $\varepsilon \to 0$, we can extract a subsequence such that m^ε converges in L^∞-weak* to a solution m of (A.1). □

The difficult part of the proof of Proposition A.1 is to check that the solution to (A.1) is unique. Let us first point out some basic properties of the solution $\bar{u}$: we already explained that $\bar{u}$ is Lipschitz continuous and semiconcave in space for any t, with a modulus bounded independently of t. We will repetitively use the fact that $\bar{u}$ can be represented as the value function of a problem of calculus of variation:

$$\bar{u}(t, x) = \inf_{\gamma, \ \gamma(t) = x} \int_t^T \tilde{L}(s, \gamma(s), \dot{\gamma}(s), \bar{m}(s)) \mathrm{d}s + \tilde{g}(\gamma(T)) \tag{A.2}$$

where we have set, for simplicity of notation,

$$\tilde{L}(s, x, v) = L(x, v) + f(x, \bar{m}(s)), \qquad \tilde{g}(x) = g(x, \bar{m}(T)).$$

For $(t, x) \in [0, T] \times \mathbb{T}^d$ we denote by $\mathcal{A}(t, x)$ the set of optimal trajectories for the control problem (A.2).

We need to analyze precisely the connexion between the differentiability of $\bar{u}$ with respect to the x variable and the uniqueness of the minimizer in (A.2) (see [6], Thms. 6.4.7 and 6.4.9 and Cor. 6.4.10). Let $(t, x) \in [0, T] \times \mathbb{T}^d$ and $\gamma \in \Gamma$. Then

1. (Uniqueness of the optimal control along optimal trajectories) Assume that $\gamma \in \mathcal{A}(t, x)$. Then, for any $s \in (t, T]$, $\bar{u}(s, \cdot)$ is differentiable at $\gamma(s)$ for $s \in (t, T)$ and one has $\dot{\gamma}(s) = -D_p H(\gamma(s), \nabla \bar{u}(s, \gamma(s)))$.
2. (Uniqueness of the optimal trajectories) $\nabla u(t, x)$ exists if and only if $\mathcal{A}(t, x)$ is a reduced to singleton. In this case, $\dot{\gamma}(t) = -D_p H(x, \nabla \bar{u}(t, x))$ where $\mathcal{A}(t, x) = \{\gamma\}$.
3. (Optimal synthesis) conversely, if $\gamma(\cdot)$ is an absolutely continuous solution of the differential equation

$$\begin{cases} \dot{\gamma}(s) = -D_p H(s, \gamma(s), \nabla \bar{u}(s, \gamma(s))) & \text{a.e. in } [t, T] \\ \gamma(t) = x, \end{cases} \tag{A.3}$$

then the trajectory γ is optimal for $\bar{u}(t, x)$. In particular, if $\bar{u}(t, \cdot)$ is differentiable at x, then equation (A.3) has a unique solution, corresponding to the optimal trajectory.

The next ingredient is Ambrosio's superposition principle, which says that any weak solution to the transport equation (A.1) can be represented by a measure on the space of trajectories of the ODE

$$\dot{\gamma}(s) = -D_p H(\gamma(s), \nabla \bar{u}(s, \gamma(s))). \tag{A.4}$$

Theorem A.3 (Ambrosio superposition principle). *Let μ be a solution to (A.1). Then there exists a Borel probability measure η on $C^0([0, T], \mathbb{T}^d)$ such that $\mu(t) = e_t \sharp \eta$ for any t and, for η-a.e. $\gamma \in C^0([0, T], \mathbb{T}^d)$, γ is a solution to the ODE (A.4).*

See, for instance, Theorem 8.2.1. from [4].

We are now ready to prove the uniqueness part of the result:

Proof of Proposition A.1. Let μ be a solution of the transport equation (A.1). From Ambrosio superposition principle, there exists a Borel probability measure η on $C^0([0, T], \mathbb{T}^d)$ such that $\mu(t) = e_t \sharp \eta$ for any t and, for η-a.e. $\gamma \in C^0([0, T], \mathbb{T}^d)$, γ is a solution to the ODE $\dot{\gamma} = -D_p H(t, \gamma(t), \nabla u(t, \gamma(t)))$. As $m_0 = e_0 \sharp \eta$, we can disintegrate the measure η into $\eta = \int_{\mathbb{T}^d} \eta_x dm_0(x)$, where $\gamma(0) = x$ for η_x-a.e. γ and m_0-a.e. $x \in \mathbb{T}^d$. Since m_0 is absolutely continuous, for m_0-a.e. $x \in \mathbb{T}^d$, η_x-a.e. map γ is a solution to the ODE starting from x. By the optimal synthesis explained above, such a solution γ is optimal for the calculus of variation problem (A.2). As, moreover, for a.e. $x \in \mathbb{T}^d$ the solution of this problem is reduced to a singleton $\{\bar{\gamma}_x\}$, we can conclude that $d\eta_x(\gamma) = \delta_{\bar{\gamma}_x}$ for m_0-a.e. $x \in \mathbb{T}^d$. Hence, for any continuous map $\phi : \mathbb{T}^d \to \mathbb{R}$, one has

$$\int_{\mathbb{T}^d} \phi(x) \mu(t, dx) = \int_{\mathbb{T}^d} \phi(\bar{\gamma}_x(t)) m_0(x) dx$$

which defines μ in a unique way. $\qquad\square$

REFERENCES

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

[2] S.R. Aiyagari, Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics* **109** (1994) 659–84.

[3] L. Ambrosio, Transport equation and Cauchy problem for BV vector fields. *Inv. Math.* **158** (2004) 227–260.

[4] L. Ambrosio, N. Gigli and G. Savarè, Gradient flows in metric spaces and in the space of probability measures. *Lectures in Mathematics ETH Zürich*. Birkhäuser-Verlag, Basel (2008).

[5] G.W. Brown, Iterative solution of games by Fictitious Play. *Activity Anal. Prod. Alloc.* **13** (1951) 374–376.

[6] P. Cannarsa and C. Sinestrari, *Semiconcave functions, Hamilton–Jacobi equations and optimal control*. Birkhäuser, Boston (2004).

[7] P. Cardaliaguet, *Weak solutions for first order mean field games with local coupling.* Preprint hal-00827957.

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

[9] P. Cardaliaguet, G. Carlier and B. Nazaret, Geodesics for a class of distances in the space of probability measures. *Calc. Var. Partial Differ. Eq.* **48** (2013) 395–420.

[10] P. Cardaliaguet, F. Delarue, J.-M. Lasry and P.-L. Lions, The master equation and the convergence problem in mean field games. Preprint arXiv:1509.0205 (2015).

[11] P. Cardaliaguet, J. Graber, A. Porretta and D. Tonon, Second order mean field games with degenerate diffusion and local coupling. *Nonlin. Differ. Equ. Appl.* **22** (2015) 1287–1317.

[12] R. Carmona and F. Delarue, Probabilist analysis of Mean-Field Games. *SIAM J. Control Optim.* **51** (2013) 2705–2734.

[13] R.-J. DiPerna and P.-L. Lions, Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98** (1989) 511–547.

[14] O. Guéant, Mean field games equations with quadratic hamiltonian: a specific approach. *Math. Models Methods Appl. Sci.* **22** (2012) 1250022.

[15] O. Guéant, P.-L. Lions and J.-M. Lasry, Mean Field Games and Applications. *Paris-Princeton Lectures on Mathematical Finance 2010*, edited by P. Tankov, P.-L. Lions, J.-P. Laurent, J.-M. Lasry, M. Jeanblanc, D. Hobson, O. Guéant, S. Crépey, A. Cousin. Springer, Berlin (2011) 205–266.

[16] D. Fudenberg and D.K. Levine, *The theory of learning in games*. MIT Press, Cambridge, MA (1998).

[17] M. Huang, P.E. Caines and R.P. Malhamé, Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. *Proc. of 42nd IEEE Conf. Decision Contr., Maui, Hawaii* (2003) 98–103.

[18] 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. Inform. Syst.* **6** (2006) 221–252.

[19] O.A. Ladyženskaja, V.A. Solonnikov and N.N. Ural'ceva, Linear and quasilinear equations of parabolic type. In vol. 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, R.I. (1967).

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

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

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

[23] P.L. Lions, Cours au Collège de France. Available at www.college-de-france.fr.

[24] K. Miyasawa, *On the convergence of the learning process in a 2×2 non-zero-sum two-person game*. Princeton University, NJ (1961).

[25] D. Monderer and L.S. Shapley, Potential games. *Games Econ. Behav.* **14** (1996) 124–143.

[26] D. Monderer and L.S. Shapley, Fictitious play property for games with identical interests. *J. Econ. Theory* **68** (1996) 258–265.

[27] J. Robinson, An iterative method of solving a game. *Ann. Math.* (1951) 296–301.

[28] L.S. Shapley, Some topics in two-person games. *Ann. Math. Stud.* **5** (1964) 1–28.