



Learning in Mean Field Games

Saeed Hadikhanloo

► To cite this version:

Saeed Hadikhanloo. Learning in Mean Field Games. Other [cs.OH]. Université Paris sciences et lettres, 2018. English. NNT : 2018PSLED001 . tel-01747579

HAL Id: tel-01747579

<https://theses.hal.science/tel-01747579>

Submitted on 29 Mar 2018

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

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

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Université de Paris-Dauphine

Apprentissage dans les jeux à champ moyen

École doctorale de Dauphine - ED 543

UNIVERSITÉ PARIS-DAUPHINE

Spécialité MATHÉMATIQUES

Soutenue par **Saeed Hadikhanloo**
le 29 January 2018

Dirigée par
Pierre Cardaliaguet
Rida Laraki



COMPOSITION DU JURY :

M Fabio Camilli
Sapienza Università di Roma,
Rapporteur

M Sylvain Sorin
Université Pierre-et-Marie-Curie,
Rapporteur

M Yves Achdou
Université Paris-Diderot,
Président

M Pierre Cardaliaguet
CNRS & Université Paris-Dauphine,
Directeur

M Rida Laraki
CNRS & Université Paris-Dauphine,
Directeur

Mme Catherine Rainer
Université de Brest, Examineur

Mme Daniela Tonon
Université Paris-Dauphine, Examineur

Acknowledgement

This thesis is the fruit of more than twenty years of study which started from primary school and ended in this very thesis; and undoubtedly, it would not have been possible without the precious supports of many people during these years.

First of all, I appreciate very much the commendable tutoring of my supervisors, Pierre Cardaliaguet and Rida Laraki. I should thank them for their great supervision and availability, and moreover for their patience and tolerance that made this path easy to pass.

I remember the atmosphere of the mathematics Olympiad competitions in Iran, which initiated my interests in mathematics. I would like to thank the people who made that community alive and dynamic, whose names are Omid Naghshineh Arjomand, Mohammad Reza Razvan, Arash Rastegar, Seyyed Reza Moghaddasi, Kasra Alishahi and Erfan Salavati. Likewise, a great community of friends and mentors was formed at that time, helping each other not just in studies, but also in life. They are Mahdi Enshai, Pedram Pad, Seyyed Mohammad Mansouri, Shayan Dashmiz, Pourya Kabir, Mohammad Hossein Shafinia, Seyyed Hessam Firouzi, Ali Kashi, Mohammad Khodabakhsh, Hamid Firouz and Mehran Ebrahimian. I acknowledge their favours and brotherhood.

I enjoyed the guidance of my professors, Ali Reza Bahreini and Omid Amini who paved my way to suitable direction and destinations for studying maths in France.

Unfortunately, we no longer have Gerard Tronel among us. He would have been pleased to see this thesis too. He used to be my tutor during my stay in France. Even though he had been suffering from some health problems, he helped me a lot to get acquainted to the new environment. I will never forget his attitude and compassionate behaviour.

I also enjoyed a great collaboration with Francisco J. Silva Álvarez. I am grateful for his preciseness, availability and sympathetic approach that were very valuable to me. He, apart from the scientific ideas, helped me a lot in writing the manuscript to achieve a good shape for the results.

Some ideas for the online mirror descent part were also created during a great discussion with Panayotis Mertikopoulos. I acknowledge his availability and helpful suggestions.

I am also obliged to the members of Parisian weekly game theory seminars, Sylvain Sorin in particular for his great mentoring the group. I should similarly thank Thomas Rivera for organizing the PhD seminars that gave me the chance of presenting my works.

I have passed recent years among the Lamsade Phd students; and I should remember them because they provided me with a vibrant and enjoyable environment.

Again, this thesis is a result of many people's contributions, direct or indirect; it is however difficult to name them all and I appreciate all those helps. Specifically, I would like to say many thanks to my family members because of their support, especially my father, mother, wife and my sisters. They are the ones who have devoted their lives to me. I dedicate this thesis to them.

Abstract

Mean Field Game (MFG) systems describe equilibrium configurations in differential games with infinitely many infinitesimal interacting agents. The main question in this thesis is to find learning procedures in these games and to investigate if they converge to an equilibrium. This thesis is delivered in 4 chapters.

We present in chapter 1, a unified review of different concepts we use throughout the thesis. First, we give a brief explanation of two subjects, *learning in game theory* and the model of *mean field games*. Second, we present our main contributions to the question of the thesis, that are explained more in details in its subsequent chapters.

In chapter 2, we introduce a learning procedure, similar to the Fictitious Play [26], for MFGs and show its convergence when the MFG is potential. Potential MFG were introduced by Lasry and Lions [74] and are such that the equilibrium can be obtained by minimizing a functional, called the potential. We can formally show that this class of game is a very natural extension of the potential game model of Monderer and Shapley [80], defined for finitely many players and finite dimensional strategy spaces.

In chapter 3, we introduce a model of non-atomic anonymous games with the player dependent action sets; typical examples of this model are first-order mean field games. We propose several learning procedures based on the fictitious play and the online mirror descent and prove their convergence to equilibrium under the classical monotonicity condition.

In chapter 4, we consider finite MFGs, i.e. with finite time and finite states. We adopt the framework introduced by Gomes et al. [55] and study two seemingly unexplored subjects. In the first one, we analyze the convergence of the fictitious play learning procedure, inspired by the results in continuous MFGs, in chapters 2 and 3. In the second one, we consider the relation of some finite MFGs and continuous first order MFGs. Namely, given a continuous first order MFG problem and a sequence of refined space/time grids, we construct a sequence finite MFGs whose solutions admit limits points and every such limit point solves the continuous first order MFG problem.

Résumé

Les systèmes de jeux à champ moyen (MFG) décrivent des configurations d'équilibre dans des jeux différentiels avec un nombre infini d'agents infinitésimaux. L'objectif principal de cette thèse est de trouver des procédures d'apprentissage dans ces jeux et d'étudier leur convergence vers un équilibre. Cette thèse se structure autour de 4 chapitres.

Le chapitre 1 présente une revue unifiée des différents concepts que nous utilisons tout au long de la thèse. Nous donnons en premier lieu une brève explication de *l'apprentissage en théorie des jeux* et modèle de *jeux à champ moyen*. Nous présentons nos principales contributions, qui sont expliquées plus en détails dans les chapitres suivants.

Le chapitre 2 introduit une procédure d'apprentissage, similaire au 'fictitious play' [26] pour les MFGs et montrons sa convergence lorsque le MFG est potentiel. Les MFG potentiels ont été introduits par Lasry et Lions [74] et sont tels que l'équilibre peut être obtenu en minimisant une fonction, appelée 'potentiel'. Nous pouvons montrer formellement que cette classe de jeu est une extension naturelle du modèle de jeu potentiel de Monderer et Shapley [80], défini pour un nombre fini de joueurs et d'espaces de stratégie de dimension finie.

Le chapitre 3 introduit un modèle de jeux anonymes non atomiques avec des jeux d'action dépendants du joueur. Des exemples typiques de ce modèle sont les jeux de champ à moyen de premier ordre. Nous proposons plusieurs procédures d'apprentissage basées sur le 'fictitious play' et le 'online mirror descent' et prouvons leur convergence vers un équilibre sous la condition de monotonie classique.

Enfin, le chapitre 4, considère les MFG finis, c'est-à-dire des MFGs en temps et avec des états finis. Nous utilisons le cadre introduit par Gomes et al. [55] en étudiant deux sujets jusqu'à présent inexplorés. Dans le premier, nous analysons la convergence de la procédure d'apprentissage de 'fictitious play', inspirée par les résultats des MFG continus, obtenus dans les chapitres 2 et 3. Dans le second, nous considérons la relation entre certains MFG finis et les MFG continus de premier ordre. Notamment, étant donné un problème MFG continu de premier ordre et une séquence de grilles spatiales/temporelles raffinées, nous construisons une séquence MFG finie dont les solutions admettent des points limites où chacun de ces points limite résout le problème MFG continu de premier ordre.

Contents

Contents	vii
1 Introduction	1
1.1 Game theory	1
1.1.1 Finite games	2
1.1.2 Static non atomic games	3
1.1.3 Monotone games	4
1.2 Learning schemes in games	5
1.3 Mean field games	10
1.3.1 Assumptions on data	12
1.3.2 Second order	13
1.3.3 First order	13
1.4 Our contributions	18
1.4.1 Fictitious play in potential MFG	19
1.4.2 Non atomic anonymous games	22
1.4.3 Finite MFG: fictitious play and convergence to classical MFG	24
2 Learning in potential MFG with fictitious play	29
2.1 Introduction	29
2.1.1 Preliminaries and assumptions	31
2.1.2 Potential mean field games	32
2.2 The fictitious play for second order MFG systems	34
2.2.1 The learning rule and the convergence result	34
2.2.2 Proof of the convergence	35
2.3 The fictitious play for first order MFG systems	41
2.3.1 The learning rule and the potential	41
2.3.2 Convergence of the fictitious play	45
2.3.3 Fictitious play scheme in N -players first order games	47
2.4 Well-posedness of a continuity equation	49
3 Learning in anonymous non atomic games	53
3.1 Introduction	53
3.2 Non atomic anonymous games	55
3.2.1 Model	55
3.2.2 Nash equilibria	56
3.2.3 Anonymous games with monotone cost	58
3.3 Fictitious play in anonymous games	59

3.4	Online mirror descent	63
3.4.1	Preliminaries	63
3.4.2	OMD algorithm and convergence result	64
3.5	Application to first order MFG	66
3.5.1	Model	66
3.5.2	Fictitious play in monotone first order MFG	68
3.5.3	Online mirror descent in monotone first order MFG	70
3.6	Appendix	72
4	Finite MFG: fictitious play and convergence to classical MFG	75
4.1	Introduction	75
4.2	The finite state and discrete time Mean Field Game problem	76
4.3	Fictitious play for the finite MFG system	78
4.3.1	Generalized fictitious play	78
4.3.2	Convergence of the fictitious play for finite MFG	81
4.4	First order MFG as limits of finite MFG	84
4.4.1	Convergence analysis	86
	References	94

Chapter 1

Introduction

In this chapter we briefly introduce the different concepts, notions and models we encounter throughout the thesis. We first review preliminary game theory frameworks and related notions e.g. finite games, potential games, non atomic games, monotone games, Nash equilibria, etc. Afterwards, we recall the question of learning and adaptive schemes in games and review the two most classical procedures, fictitious play and online convex optimisation schemes. We then give an introduction to Mean Field Games (MFGs) by defining the second and first order ones and then recalling the different regularity notions for solutions in cases with noise (stochastic and second order) and without noise (deterministic and first order). At the end of the chapter, a summary of our contributions to the question of learning in mean field games will be given; those are separately explained in more details in upcoming chapters. In order to concisely cover all the material, we usually refer the reader to the corresponding references for detailed explanations and proofs.

1.1 Game theory

Game theory studies situations where there is a conflict of interest on a result which is produced by the decisions made by many decision makers. The decision-makers (or players) make decisions which yield a final situation whose desirability is different for every individual. For each player involved in the game, their actions do not just affect their own utility, they can also affect other players' pay-offs. Since the concept is quite general and covers many different problems, many different types of games appear in the literature. There are differences reflecting various aspects of a game; for example finite versus infinite number of agents, cooperative (coalition-wise) versus non-cooperative, finite versus infinite set of actions, complete information and partial information. Some of the surveys and books covering the main topics in game theory are [93][54][85][84][71].

Let us fix our notation for a few common concepts. We usually use I as the set of players. For every player $i \in I$ the set X_i denotes the set of decisions available to player i . The set of all profiles of decisions is $X = \prod_{i \in I} X_i$. Typical element in X_i and X are denoted respectively by x_i and $\mathbf{x} = (x_i)_{i \in I}$. For every profile $\mathbf{x} \in X$ and every player $i \in I$ we write $\mathbf{x} = (x_i, \mathbf{x}_{-i})$, where $x_i \in X_i$, $\mathbf{x}_{-i} \in X_{-i} = \prod_{j \in I, j \neq i} X_j$ are respectively the decisions of player i and of its adversaries in the profile \mathbf{x} . For every individual $i \in I$, the desirability of different profiles for player i is modelled by a preference relation on X , that is a complete order \succeq_i on X . A numerical version of this preference is captured by a cost function, $c_i : X \rightarrow \mathbb{R}$, where the more desirable a profile is for i , the smaller the cost, i.e.

$$\forall \mathbf{x}, \mathbf{x}' \in X : \quad c_i(\mathbf{x}) \leq c_i(\mathbf{x}') \quad \text{if and only if} \quad \mathbf{x} \succeq_i \mathbf{x}'.$$

1.1.1 Finite games

We now review the most classical non-cooperative model in game theory, i.e. finite games. Let the set of players I be finite. For every player $i \in I$, the set of its decisions is of the form $X_i = \Delta(A_i)$ that is the set of all probability measures over a finite set A_i , usually called the set of actions. An arbitrary element $x_i \in X_i$ is called a *mixed strategy* of player i ; more specifically, the elements of the action set A_i which can be regarded as the singular measures in X_i , usually are called *pure strategies*. The profiles of mixed and pure strategies are denoted respectively by $X = \prod_{i \in I} X_i$ and $A = \prod_{i \in I} A_i$. Every cost function $c_i : A \rightarrow \mathbb{R}$ which is defined on the profile of pure actions A , can be extended multi-linearly to the set of all profiles of mixed strategies X :

$$\forall \mathbf{x} = (x_i)_{i \in I} \in X : \quad c_i(\mathbf{x}) = \sum_{\mathbf{a} = (a_i)_{i \in I} \in A} \mathbf{x}(\mathbf{a}) c_i(\mathbf{a}),$$

where $\mathbf{x}(\mathbf{a}) = \prod_{i \in I} x_i(a_i)$, reflecting the fact that players choose their mixed strategies independently.

Definition 1.1.1. A profile of mixed strategies $\tilde{\mathbf{x}} = (\tilde{x}_i)_{i \in I} \in X$ is a mixed Nash equilibrium iff

$$\forall i \in I : \quad c_i(\tilde{x}_i, \tilde{\mathbf{x}}_{-i}) = \min_{x \in X_i} c_i(x, \tilde{\mathbf{x}}_{-i}).$$

The mixed Nash equilibrium $\tilde{\mathbf{x}} = (\tilde{x}_i)_{i \in I}$ is called a pure Nash equilibrium if for all $i \in I, x_i \in A_i$.

The notion of Nash equilibrium can be regarded as a stable profile where no player can be better-off by individually deviating from their original decision. The existence of mixed equilibrium in finite games is a classical theorem proved by Nash (1951) by using the Brouwer's fixed point theorem:

Theorem 1.1.1 (Nash 1951 [81]). *There is at least one mixed Nash equilibrium in every finite game.*

The existence of an equilibrium profile can be seen as an existence of a fixed point for a map related to the game. For every player $i \in I$ the best reply of player i to a profile $\mathbf{x} \in X$ denoting by $BR(i, \mathbf{x})$, is defined as

$$BR(i, \mathbf{x}) = \operatorname{argmin}_{x \in X_i} c_i(x, \mathbf{x}_{-i}).$$

The *best reply* correspondence $BR : X \mapsto X$ is obtained as the product of individual best response sets $BR(\mathbf{x}) = \prod_{i \in I} BR(i, \mathbf{x})$. By definition $\tilde{\mathbf{x}}$ is a Nash equilibrium if and only if it is a fixed point of the best reply correspondence, i.e. if $\tilde{\mathbf{x}} \in BR(\tilde{\mathbf{x}})$.

Finite potential games. Finite potential games were defined by Monderer and Shapley [80] as those games whose the players face to an identical cost quantity called the *potential*. In these games, the existence of a pure equilibrium can be obtained by finding a minimiser of the potential.

Definition 1.1.2. A finite game G is a potential game if there is a function $\phi : A \rightarrow \mathbb{R}$ such that

$$\forall i \in I, a_i, a'_i \in A_i, \mathbf{a}_{-i} \in A_{-i} : \quad c_i(a_i, \mathbf{a}_{-i}) - c_i(a'_i, \mathbf{a}_{-i}) = \phi(a_i, \mathbf{a}_{-i}) - \phi(a'_i, \mathbf{a}_{-i}).$$

The function ϕ is called a *potential* of the game.

A game is potential if and only if the cost functions are in the form

$$\forall i \in I, \forall \mathbf{a} = (a_i, \mathbf{a}_{-i}) \in A : \quad c_i(a_i, \mathbf{a}_{-i}) = G(a_i, \mathbf{a}_{-i}) + F_i(\mathbf{a}_{-i}) + f_i(a_i).$$

This game is a potential game with $\phi(\mathbf{a}) = G(\mathbf{a}) + \sum_{i \in I} f_i(a_i)$.

1.1.2 Static non atomic games

Non atomic games represent the strategic interactions with an infinite number of small deciders. These games are called non atomic when every individual decision is negligible in the overall result and only the aggregative behaviour of non-zero measure sets of players can change the pay-offs. The applications are numerous, from traffic, internet routing, voting, etc. This branch of the literature was started by the seminal works of Aumann [11][12], Schmeidler [89] and Mas-Colell [76]. We try to summarize the most important and fundamental models and results in this topic.

Finite action set. We start by the model of Schmeidler 1972 ([89]) which is a starting point in the literature. This is a game model with an infinite number of players where decisions are mixed strategies over a finite set of actions. The set of players is the closed interval $I = [0, 1]$ endowed with the Lebesgue measure λ . For every $i \in I$ the decision set of player i is $X_i = \Delta(A)$ where A is a finite set, usually called the set of actions. The set of all profiles is $\Delta(A)^I$ or equivalently all functions $\Psi : I \rightarrow \Delta(A)$. We shall work with a subset of profile of decisions, consisting of all measurable maps $\Psi : I \rightarrow \Delta(A)$, where the measurability is with respect to the Borel σ -fields over I and $\Delta(A)$. We denote this *admissible* set of profiles as \mathcal{A} . For every $\Psi \in \mathcal{A}$ there are measurable functions $\Psi_a : I \rightarrow [0, 1]$ for all $a \in A$, such that

$$\Psi(i) = (\Psi_a(i))_{a \in A} \in \Delta(A) \quad \text{and} \quad \sum_{a \in A} \Psi_a(i) = 1, \quad \text{for all } i \in I.$$

The cost of player $i \in I$ facing the profile Ψ is constructed by auxiliary cost functions $C_i : A \times \mathcal{A} \rightarrow \mathbb{R}, i \in I$ as follows: $c_i(\Psi) = \sum_{a \in A} \Psi_a(i) C_i(a, \Psi)$.

Definition 1.1.3. A profile $\Psi^* \in \mathcal{A}$ is called a Nash equilibrium if

$$\text{for } \lambda\text{-almost every } i \in I : \quad \text{supp}(\Psi^*(i)) \subseteq \text{argmin}_{a \in A} C_i(a, \Psi^*),$$

where $\text{supp}(m)$ represents the support of measure m .

Theorem 1.1.2 (Schmeidler [89]). Let \mathcal{A} be endowed with the L^1 -weak topology. If the auxiliary cost functions C_i ($i \in I$) are continuous in Ψ , then there is at least one mixed Nash equilibrium. In addition, if C_i 's depend on Ψ only through $(\int_I \Psi_a(i) d\lambda(i))_{a \in A}$ then there is an equilibrium with pure actions for each player.

Infinite action set. In contrast to the approach of Schmeidler, the Mas-Colell model [76] allows to have players with infinite action spaces, e.g. positions or trajectories in \mathbb{R}^d .

Before we start the model, let us fix a few notations. We call a tuple (S, \mathcal{F}_S) a measure space if \mathcal{F}_S be a σ -field over set S . The probability measures over S respect to the σ -field \mathcal{F}_S , are denoted by $\mathcal{P}_{\mathcal{F}_S}(S)$ or for simplicity $\mathcal{P}(S)$ when \mathcal{F}_S is known; we use also the notation $\Delta(S)$ as the probability measures over S , when S is finite and $\mathcal{F}_S = 2^S$. For every measurable map $\rho : (S, \mathcal{F}_S) \rightarrow (W, \mathcal{F}_W)$, we can *push-forward* the measures over S to measures over W . That is for every $\mu \in \mathcal{P}(S)$, the push-forward of μ by ρ is an element in $\mathcal{P}(W)$ denoted by $\rho\# \mu$, and is defined by:

$$\forall B \in \mathcal{F}_W : \quad \rho\# \mu(B) = \mu(\rho^{-1}(B)).$$

Let us define the model of the game with continuum of players and infinite set of actions. Let $(\mathcal{X}, \mathcal{F}_{\mathcal{X}}), (\mathcal{Y}, \mathcal{F}_{\mathcal{Y}})$ be two measure spaces representing the set of *types* and *actions* of players. There is a fixed given distribution $\mu \in \mathcal{P}(\mathcal{X})$ capturing the distribution of types of players. Each player with type $x \in \mathcal{X}$ choosing $y \in \mathcal{Y}$ has to pay a cost equal to $\phi(x, y, \nu)$ where $\nu \in \mathcal{P}(\mathcal{Y})$ represents the induced measure of actions chosen by all players. The definition of the Nash equilibria is as follows.

Definition 1.1.4. A measure $\rho \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ is called a Nash equilibrium if $\pi_{\mathcal{X}}\# \rho = \mu$ and

$$\rho(\{ (x, y) \in \mathcal{X} \times \mathcal{Y} \mid y \in \operatorname{argmin}_{z \in \mathcal{Y}} \phi(x, z, \nu) \}) = 1,$$

where $\nu = \pi_{\mathcal{Y}}\# \rho$.

The existence of equilibria for compact metric \mathcal{X}, \mathcal{Y} and continuous ϕ is a direct application of Kakutani's fixed point theorem (see for example [24], section 2). The continuity condition on ϕ can be relaxed in an extent, which covers the cost functions depending on the density of ν ; on a series of papers, Blanchet, Carlier [23][24][25] provided an approach inspired by optimal transport theory, implying a full characterisation of such equilibria, and convergent numerical computation schemes.

1.1.3 Monotone games

Inspired by the notion of maximal monotone operators, the so-called monotonicity condition in terms of the cost functions was first defined by Rosen [88] under the terminology of the diagonal strict concavity condition. This notion, usually yielding to uniqueness of a Nash equilibrium, has been used afterwards in many games with different terminologies; Lasry, Lions [72][73] in mean field games dealt with monotone couplings; Hofbauer, Sandholm [65] considered population games with monotone costs called stable games; Blanchet, Carlier [24] worked in the framework of games with continuum of players and actions sets, etc.

Let us illustrate the idea in a game with finitely many players. Let the set of players I be finite and $(c_i)_{i \in I}$ be the cost functions. For every player i , denote a convex compact set $X_i \subseteq \mathbb{R}^d$ as the set of decisions of i . Let for every $i \in I$, the cost function $c_i(x_i; x_{-i})$ be \mathcal{C}^1 with respect to the i -th variable x_i , and denote

$$\forall \mathbf{x} = (x_i; x_{-i}) \in X : \quad \nabla_{x_i} c_i(\mathbf{x}) = v_i(\mathbf{x}).$$

We denote $v(\mathbf{x}) = (v_i(\mathbf{x}))_{i \in I}$ and for all $\mathbf{z} = (z_i; z_{-i}) \in X$ we define:

$$\langle v(\mathbf{x}), \mathbf{z} \rangle = \sum_{i \in I} \langle v_i(\mathbf{x}), z_i \rangle.$$

With this formulation, the Nash equilibria have a variational representation.

Proposition 1.1.1. If the profile $\mathbf{x}^* \in X$ is an equilibrium then:

$$\forall \mathbf{x} \in X : \quad \langle v(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle \geq 0.$$

Proof. By definition, it is sufficient to prove that for every $i \in I$, we have

$$\forall x_i \in X_i : \quad \langle v_i(\mathbf{x}^*), x_i - x_i^* \rangle \geq 0.$$

Set $\mathbf{x}^\lambda = (\lambda x_i + (1 - \lambda)x_i^*, \mathbf{x}_{-i}^*)$ for $\lambda \in [0, 1]$. Since \mathbf{x}^* is an equilibrium, we have $c_i(\mathbf{x}^*) \leq c_i(\mathbf{x}^\lambda)$. So

$$0 \leq \lim_{\lambda \rightarrow 0} \frac{c_i(\mathbf{x}^\lambda) - c_i(\mathbf{x}^*)}{\lambda} = \langle v_i(\mathbf{x}^*), x_i - x_i^* \rangle.$$

□

Definition 1.1.5. A finite game with differentiable cost function $(c_i)_{i \in I}$ is called monotone if

$$\forall \mathbf{x}, \mathbf{x}' \in X : \quad \langle v(\mathbf{x}) - v(\mathbf{x}'), \mathbf{x} - \mathbf{x}' \rangle \geq 0,$$

and it is called strictly monotone if the above inequality holds strictly for $\mathbf{x} \neq \mathbf{x}'$.

The monotonicity condition gives another characterization of the equilibria.

Theorem 1.1.3. *For a monotone game:*

1. *a profile $\mathbf{x}^* \in X$ is an equilibrium if and only if:*

$$\forall \mathbf{x} \in X : \quad \langle v(\mathbf{x}), \mathbf{x} - \mathbf{x}^* \rangle \geq 0, \quad (1.1)$$

2. *if the game is strictly monotone, then the equilibrium $\mathbf{x}^* \in X$ is unique.*

Proof. If \mathbf{x}^* is an equilibrium, by Proposition 1.1.1 and monotonicity definition 1.1.5, we have

$$\forall \mathbf{x} \in X : \quad \langle v(\mathbf{x}), \mathbf{x} - \mathbf{x}^* \rangle \geq \langle v(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle \geq 0.$$

Conversely, suppose (1.1) hold for \mathbf{x}^* . Set $\mathbf{x} = (x_i, \mathbf{x}_{-i}^*)$ for an arbitrary $x_i \in X_i$. By the mean value theorem we have for some $\lambda \in (0, 1)$ and $z_i = \lambda x_i + (1 - \lambda)x_i^*$,

$$c_i(x_i, \mathbf{x}_{-i}^*) - c_i(x_i^*, \mathbf{x}_{-i}^*) = \langle v_i(z_i, \mathbf{x}_{-i}^*), x_i - x_i^* \rangle.$$

Using the fact that $x_i - x_i^* = \frac{1}{\lambda}(z_i - x_i^*)$ and (1.1) we can conclude $c_i(x_i, \mathbf{x}_{-i}^*) \geq c_i(x_i^*, \mathbf{x}_{-i}^*)$; this means \mathbf{x}^* is an equilibrium since $i \in I$ and $x_i \in X_i$ were arbitrary.

For the second statement, suppose the game is strictly monotone and there are two equilibria $\mathbf{x}^*, \tilde{\mathbf{x}}^*$. We have

$$\langle v(\mathbf{x}^*) - v(\tilde{\mathbf{x}}^*), \mathbf{x}^* - \tilde{\mathbf{x}}^* \rangle = \langle v(\mathbf{x}^*), \mathbf{x}^* - \tilde{\mathbf{x}}^* \rangle + \langle v(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^* - \mathbf{x}^* \rangle \leq 0,$$

by Proposition 1.1.1. Hence $\mathbf{x}^* = \tilde{\mathbf{x}}^*$ by strict monotonicity condition. \square

We refer to [65][77] for more properties of monotone games. The inequality (1.1) informally says, for every profile $\mathbf{x} \in X$ if we move slightly in the direction of $v(\mathbf{x})$ we get closer to the equilibrium, even without knowing where it is located. Hofbauer, Sandholm [65] proved that several dynamics in strict monotone population games converge to the set of Nash equilibria. Mertikopoulos [77] applied the mirror descent dynamics to monotone games and proved convergence to Nash equilibria.

1.2 Learning schemes in games

After the definition of various game frameworks and corresponding equilibria, the question of formation of an equilibrium arises naturally. Actually, it is unreasonable to assume that all the players coordinate their strategies to an equilibrium. The situation gets worse as the game becomes more complex, with a large set of players and a large set of actions. We refer to [53][44][94] for an overview of different learning procedures in games.

Fictitious play. Here we review a classical learning procedure in games called *Fictitious Play (FP)*. It was introduced by Brown [26] in the context of 2-players zero-sum games. Convergence towards Nash equilibria has been proven in the case of 2×2 games [78], zero-sum games [87], potential games [79], etc. The relation between the discrete procedure of fictitious play and best response dynamics has been investigated using stochastic approximation techniques, see Benaïm, Hofbauer and Sorin [19][20].

Suppose we have a finite game with set of players I . Let for every player $i \in I$, A_i be the finite set of pure actions and $X_i = \Delta(A_i)$ the set of decisions (or mixed strategies). The cost functions c_i are defined on profiles $\prod_{i \in I} A_i$ and extended multi-linearly to the set of mixed strategies $\prod_{i \in I} X_i$.

Suppose the game is played repeatedly, and at every round, the action chosen by a player is a best response with respect to the empirical average of actions of adversaries at previous rounds. More formally, let the action played by the player $i \in I$ at round n be denoted by $a_i^n \in A_i$, and $\bar{a}_i^n \in X_i$ be the empirical average of actions up to round n . Then the fictitious play scheme reads as follows: for $n = 1, 2, \dots$

$$\begin{aligned} (i) \quad & a_i^{n+1} \in BR(i, \bar{a}_{-i}^n) = \operatorname{argmin}_{a \in A_i} c_i(a, \bar{a}_{-i}^n), \quad \text{for every } i \in I, \\ (ii) \quad & \bar{a}_i^{n+1} = \frac{1}{n+1} \sum_{k=1}^{n+1} a_i^k, \quad \text{for every } i \in I, \\ (iii) \quad & \bar{a}^{n+1} = \prod_{i \in I} \bar{a}_i^{n+1}, \end{aligned} \tag{1.2}$$

where $\prod_{i \in I}$ denotes the Cartesian product. The main question in fictitious play is whether the empirical average \bar{a}^n (or realized actions a^n) converges to the set of Nash equilibria or not. This question is answered in the case of finite potential games by Monderer and Shapley [79]. Let us give a sketch of their proof; their approach will give an idea for some of our arguments in the case of mean field games.

Theorem 1.2.1 (Monderer, Shapley [79]). *Let $\{\bar{a}^n\}_{n \in \mathbb{N}}$ be constructed by a fictitious play scheme proposed in (1.2). If the game is potential, then $\lim_{n \rightarrow \infty} d(\bar{a}^n, NE) = 0$, where NE is the set of Nash equilibria.*

Sketch of the proof. Let ϕ be a potential of the game. By definition of average profiles

$$\bar{a}^{n+1} = (\bar{a}_i^{n+1})_{i \in I} = \left(\frac{1}{n+1} (a_i^{n+1} - \bar{a}_i^n) + \bar{a}_i^n \right)_{i \in I}$$

and by multi-linearity of ϕ , we have

$$\phi(\bar{a}^{n+1}) - \phi(\bar{a}^n) = \frac{1}{n+1} \left(\sum_{i \in I} c_i(a_i^{n+1}, \bar{a}_{-i}^n) - c_i(\bar{a}_i^n, \bar{a}_{-i}^n) \right) + \frac{K_n}{(n+1)^2}.$$

The quantities $\{K_n\}_{n \in \mathbb{N}}$ are uniformly bounded, i.e. there is $K > 0$ such that $|K_n| < K$ for all $n \in \mathbb{N}$. If we denote

$$b_n = \sum_{i \in I} c_i(\bar{a}_i^n, \bar{a}_{-i}^n) - c_i(a_i^{n+1}, \bar{a}_{-i}^n), \quad n \in \mathbb{N},$$

then $b_n \geq 0$, since $a_i^{n+1} \in \operatorname{argmin}_{a \in A_i} c_i(a, \bar{a}_{-i}^n)$ for all $i \in I$. Writing

$$\phi(\bar{a}^{n+1}) - \phi(\bar{a}^n) = -b_n/(n+1) + K_n/(n+1)^2,$$

and summing up over all $n \in \mathbb{N}$ gives,

$$\sum_{n \in \mathbb{N}} b_n/(n+1) = \sum_{n \in \mathbb{N}} \phi(\bar{a}^n) - \phi(\bar{a}^{n+1}) + \sum_{n \in \mathbb{N}} K_n/(n+1)^2 < +\infty.$$

The boundedness of the first sum comes from the boundedness of the potential function and the fact that the telescopic terms cancel each other consecutively; the second sum is finite since K_n 's are uniformly bounded. We can deduce from $\sum_{n \in \mathbb{N}} b_n/(n+1) < +\infty$ and positiveness of b_n 's that

$$\lim_{k \rightarrow \infty} \frac{\sum_{n=1}^k b_n}{k} = 0. \tag{1.3}$$

We next show that for every $\epsilon > 0$, for all enough large n , the average profile \bar{a}^n is an ϵ -equilibrium; it yields $\lim_{n \rightarrow \infty} d(\bar{a}^n, NE) = 0$. Using the fact that $\|\bar{a}^n - \bar{a}^{n+1}\| = O(1/n)$, we can prove that there is $C > 0$ with $|b_n - b_{n+1}| \leq C/n$, for all $n \in \mathbb{N}$. This property with equation (1.3) gives $\lim_{n \rightarrow \infty} b_n = 0$. On the other hand, if \bar{a}^n is not an ϵ -equilibrium then by definition $b_n \geq \epsilon$. Thus for all $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that for all $n > N_\epsilon$ the average profile \bar{a}^n is an ϵ -equilibrium. \square

This summary would be incomplete without recalling the application of stochastic approximation in convergence of fictitious play schemes. Due to techniques developed by Benaïm, Hofbauer and Sorin [19][20] one can assert that the convergence of best response dynamics implies the convergence of FP. Their approach for the case of fictitious play is as follows: we can rewrite (1.2) as

$$(n+1)\bar{a}_i^{n+1} - n\bar{a}_i^n \in BR(i, \bar{a}_{-i}^n) \quad \text{for all } i \in I,$$

since $a_i^{n+1} = (n+1)\bar{a}_i^{n+1} - n\bar{a}_i^n$. Setting $\bar{a}^{n+1} = (\bar{a}_i^{n+1})_{i \in I}$ it gives us

$$\bar{a}^{n+1} - \bar{a}^n \in \frac{1}{n+1} (BR(\bar{a}^n) - \bar{a}^n),$$

where $BR(\bar{a}^n) = \prod_{i \in I} BR(i, \bar{a}_{-i}^n)$. The stochastic approximation method relates the asymptotic behaviour of fictitious play scheme to the continuous dynamic $\dot{a}(t) \in BR(a(t)) - a(t)$, called the best response dynamics. For a survey on stochastic approximation, we refer to [18].

Online learning in convex optimization. Classical optimization problems deal with minimizing some given function on some given region. In *online optimisation* problems, one has to optimize over a flow of functions which are unknown at the beginning and become revealed after each step. The examples are very frequent; from routing problems to applications in machine learning. For a few surveys on this topic, we refer to [44][90].

Let us describe this framework more precisely. Let \mathcal{X} be the set of *choices* and $\mathcal{S} \subseteq \mathcal{X}^{\mathbb{R}}$ be the set of *cost functions*. Consider a decision maker (DM) who chooses elements in \mathcal{X} and pays costs according to the following scheme: at every step $n \in \mathbb{N}$,

- the DM chooses an element $x_n \in \mathcal{X}$,
- a cost function $f_n \in \mathcal{S}$ is revealed,
- then the DM has to pay $f_n(x_n)$.

The goal of the decision maker is to choose "optimally" the choices x_n . The revealed history up to step n , is defined by $H_n = (\mathcal{X} \times \mathcal{S})^n$, and a typical element h_n in H_n is in the form

$$h_n = (x_1, f_1, \dots, x_n, f_n)$$

where x_m, f_m represent the choice of the decision maker and revealed cost function at step $1 \leq m \leq n$. We should notice that in a complete information scheme, the decision maker at step $n+1$ is completely aware of past history $h_n = (x_1, f_1, \dots, x_n, f_n)$. Accordingly, a *strategy* is a map $\sigma : \cup_{n \in \mathbb{N}} H_n \rightarrow \mathcal{X}$ that gives a rule to the decision maker to choose an element in \mathcal{X} at step $n+1$ as a function of known history up to step n , for all $n \in \mathbb{N}$.

One of the criteria of optimality in these classes of problems is defined by Hannan [60] with the notion of regret: the *regret* is a map $R : \cup_{n \in \mathbb{N}} H_n \rightarrow \mathbb{R} \cup \{+\infty\}$ such that for all $n \in \mathbb{N}$:

$$\forall h_n = (x_1, f_1, \dots, x_n, f_n) \in H_n : \quad R(h_n) = \sum_{m=1}^n f_m(x_m) - \min_{x \in \mathcal{X}} \sum_{m=1}^n f_m(x).$$

The value $R(h_n)$ captures the regret of the decision maker of not having chosen a fixed choice for the steps up to n . This value can be considered as a tool to give a sense to optimal strategies in online optimization problems.

Definition 1.2.1. A strategy $\sigma : \cup_{n \in \mathbb{N}} H_n \rightarrow \mathcal{X}$ is called a no-regret strategy if for all revelation of cost functions $f_1, f_2, \dots \in \mathcal{S}$, we have

$$\limsup_{n \rightarrow \infty} \frac{R(h_n)}{n} \leq 0,$$

where $h_n = (x_1, f_1, \dots, x_n, f_n) \in H_n$ is constructed under the strategy σ on the revealed cost functions, i.e. $x_{m+1} = \sigma(h_m)$ for all $m \in \mathbb{N}$.

Follow the regularized leader. A series of examples of strategies with relatively low regrets, are the ones constructed by the Follow the Regularized Leader (FtRL) strategies introduced by Kalai, Vempala [68].

Here we borrowed from [90]. We suppose that $\mathcal{X} \subseteq \mathbb{R}^d$ is compact convex and there is $L > 0$ such that all cost functions $f \in \mathcal{S}$ are L -Lipschitz convex functions from \mathcal{X} to \mathbb{R} . We call a map $T : \mathcal{X} \rightarrow \mathbb{R}$ a strongly convex function, if there is $K > 0$ such that for all $x, y \in \mathcal{X}, \lambda \in [0, 1]$:

$$T(\lambda x + (1 - \lambda)y) \leq \lambda T(x) + (1 - \lambda)T(y) - K\lambda(1 - \lambda)\|x - y\|^2.$$

Definition 1.2.2. Let $T : \mathcal{X} \rightarrow \mathbb{R}$ be strongly convex and $\{\epsilon_n\}_{n \in \mathbb{N}}$ be a sequence of positive numbers. The strategy $\sigma : \cup_{n \in \mathbb{N}} H_n \rightarrow \mathcal{X}$ defined with

$$\sigma(h_n) \in \operatorname{argmin}_{x \in \mathcal{X}} \sum_{m=1}^n f_m(x) + \frac{1}{\epsilon_n} T(x), \quad \text{for } n \in \mathbb{N}, \quad (1.4)$$

is called a Follow the Regularised Leader (FtRL) procedure.

The reason for this name is as follows: suppose each element $x \in \mathcal{X}$ represents an expert and $f(x)$ the cost of obeying the expert x while the cost function is f . At each round the decision maker has to choose which expert to follow. So the equation (1.4) describes that the decision maker has followed the best expert in performance up to the current round, regularized with a function T . The following lemma gives an estimation of the regret imposed by FtRL procedure.

Lemma 1.2.1. Let $\sigma : \cup_{n \in \mathbb{N}} H_n \rightarrow \mathcal{X}$ be a FtRL strategy defined in (1.4) with $\epsilon_n = \epsilon$ for all $n \in \mathbb{N}$. For every sequence of histories $\{h_n\}_{n \in \mathbb{N}}, h_n \in H_n$ with

$$h_n = (x_1, f_1, \dots, x_n, f_n), \quad x_{n+1} = \sigma(h_n), \quad \text{for all } n \in \mathbb{N},$$

we have

$$R(h_n) \leq \frac{1}{\epsilon} T(x_1) + \sum_{m=2}^n f_m(x_m) - f_m(x_{m+1}). \quad (1.5)$$

Example 1.2.1. Let $\mathcal{X} \subseteq \mathbb{R}^d$ be bounded, convex, closed and the cost functions be of the form

$$f_n(x) = \langle z_n, x \rangle \quad \text{for some } z_n \in \mathbb{R}^d.$$

Fix $N \in \mathbb{N}$. If $T(x) = \frac{1}{2}\|x\|^2, \epsilon_n = \epsilon$ for $n \leq N$, then the FtRL procedure takes the following form:

$$x_{n+1} = \pi_{\mathcal{X}}(-\epsilon(z_1 + z_2 + \dots + z_n)), \quad \text{or} \quad x_{n+1} = \pi_{\mathcal{X}}(x_n - \epsilon z_n), \quad \text{for } n \in \mathbb{N}, n \leq N, \quad (1.6)$$

where $\pi_{\mathcal{X}} : \mathbb{R}^d \rightarrow \mathcal{X}$ is the projection on set \mathcal{X} . The regret also can be bounded as follows:

$$R(h_n) \leq \frac{1}{\epsilon} T(x_1) + \epsilon \sum_{m=1}^n \|z_m\|^2.$$

Hence, if the z_n 's are uniformly bounded, one can chooses $\epsilon = \frac{1}{\sqrt{N}}$ that makes $R(h_n) \leq C\sqrt{N}$, for all $n \leq N$ and for a quantity $C > 0$ independent of n, N . We can even set $\{\epsilon_n\}_{n \in \mathbb{N}}$ such that $R(h_n) \leq C\sqrt{n}$,

for all $n \in \mathbb{N}$ (see [44], section 2.3). The idea called double trick, is as follows. Set $\epsilon_m = \frac{1}{\sqrt{2^k}}$ for $2^k \leq m < 2^{k+1}$. For all n with $2^k \leq n < 2^{k+1}$ we denote:

$$h_n = \hat{h}_1 \oplus \hat{h}_2 \oplus \dots \oplus \hat{h}_{2^{k-1}} \oplus \hat{h}_{n-2^k+1},$$

where

$$\forall l \in \mathbb{N}^* : \quad \hat{h}_{2^l} = (x_{2^l}, f_{2^l}, \dots, x_{2^{l+1}-1}, f_{2^{l+1}-1}), \quad \hat{h}_{n-2^k+1} = (x_{2^k}, f_{2^k}, \dots, x_n, f_n),$$

and the operator \oplus is the concatenation of the vectors. Then using the previous result we can write

$$\begin{aligned} R(h_n) &\leq \sum_{l=0}^{k-1} R(\hat{h}_{2^l}) + R(\hat{h}_{n-2^k+1}) \leq C \sum_{l=0}^k \sqrt{2^l} \\ &\leq \frac{C}{\sqrt{2}-1} \sqrt{2^{k+1}} \leq \frac{C\sqrt{2}}{\sqrt{2}-1} \sqrt{n}. \end{aligned} \tag{1.7}$$

Thus in this way, for $C' = \frac{C\sqrt{2}}{\sqrt{2}-1}$ we have $R(h_n) \leq C' \sqrt{n}$ for all $n \in \mathbb{N}$.

As shown in [43][10] the best asymptotic one can propose for regret at step n is in the order of \sqrt{n} .

Online mirror descent. The mirror descent methodologies started with the work of Nemirovski and Yudin [82]. This procedure is a modification of FtRL when the cost functions are convex. Before we give an exact definition, let us recall a property concerning convex maps.

Remark 1.2.1. Let $\mathcal{X} \subseteq \mathbb{R}^d$ be an convex set. For every convex Lipschitz function $f : \mathcal{X} \rightarrow \mathbb{R}$ and every $x \in \mathcal{X}$ there is a non-empty set $\partial f(x) \subseteq \mathbb{R}^d$, called sub-gradient set of f at point x , such that

$$\forall z \in \partial f(x), y \in \mathcal{X} : \quad f(y) - f(x) \geq \langle z, y - x \rangle.$$

Recalling that the cost functions $f \in \mathcal{S}$ are convex, we can rewrite the relation (1.5) as follows:

$$R(h_n) = \sum_{m=1}^n f_n(x_n) - \min_{x \in \mathcal{X}} \sum_{m=1}^n f_n(x) \leq \tilde{R}_N = \sum_{m=1}^n \langle z_n, x_n \rangle - \min_{x \in \mathcal{X}} \sum_{m=1}^n \langle z_n, x \rangle.$$

with $z_n \in \partial f_n(x_n)$. We should note that \tilde{R}_N is the regret with respect to the functions $l_n(x) = \langle z_n, x \rangle$, so if we can bound the regret \tilde{R}_N for linear functions $\{l_n\}_{n \in \mathbb{N}}$ we can do so for $R(h_n)$ and convex functions $\{f_n\}_{n \in \mathbb{N}}$. The FtRL procedure for convex cost functions takes the form:

$$x_{k+1} \in \operatorname{argmin}_{x \in \mathcal{X}} \left\langle \sum_{n=1}^k z_n, x \right\rangle + \frac{1}{\epsilon} T(x), \quad z_{k+1} \in \partial f_{k+1}(x_{k+1}) \tag{1.8}$$

For $k \in \mathbb{N}$. If we set $y_k = -\epsilon(z_1 + z_2 + \dots + z_k)$, we obtain the definition of online mirror descent:

Definition 1.2.3 (Online mirror descent). Suppose $T : \mathcal{X} \rightarrow \mathbb{R}$ is such that the mirror correspondence

$$Q_{\mathcal{X}} : \mathbb{R}^d \mapsto \mathcal{X} : \quad Q_{\mathcal{X}}(y_k) := \operatorname{argmin}_{x \in \mathcal{X}} \{T(x) - \langle y_k, x \rangle\},$$

is well-defined. For an arbitrary $(x_1, y_1) \in \mathcal{X} \times \mathbb{R}^d$, we say the sequence $\{(x_k, y_k)\}_{k \in \mathbb{N}}$ is constructed by an online mirror descent (OMD) scheme if

$$x_{k+1} \in Q_{\mathcal{X}}(y_k), \quad y_{k+1} = y_k - \epsilon z_{k+1} \quad \text{with } z_{k+1} \in \partial f_{k+1}(x_{k+1}). \tag{1.9}$$

Online learning in games. The online optimisation framework can be used as a learning procedure in games. Suppose a game being played repeatedly. So a given player i , at round $n \in \mathbb{N}$ faces to the cost function $f_n = c_i(\cdot, \mathbf{x}_{-i}^n)$ where \mathbf{x}_{-i}^n is the action of other players at the current round. Since \mathbf{x}_{-i}^n is priorly unknown before player i plays at stage n , this player is dealing with an optimal optimization framework with its set of decisions X_i as the set of choices \mathcal{X} . Consider for example the framework of games with finitely many players. Let I be the set of players, $X_i \subseteq \mathbb{R}^d$ compact convex be the set of decisions of $i \in I$ and $X = \prod_{i \in I} X_i$. If all players apply the online mirror descent (1.9), then $\{\mathbf{x}_k\}_{k \in \mathbb{N}} \subseteq X$ takes the form

$$\mathbf{x}_{k+1} \in Q_X(\mathbf{y}_k), \quad \mathbf{y}_{k+1} = \mathbf{y}_k - \epsilon_k v(\mathbf{x}_{k+1}) \quad (1.10)$$

where $v(\mathbf{x}) = (v_i(\mathbf{x}))_{i \in I}$ with $v_i(\mathbf{x}) = \nabla_{x_i} c_i(\mathbf{x}), i \in I$.

For the seminal applications of online algorithms for games, we refer to the series of works by Hart and Mas-Colell applying the no-regret algorithms [61][62][63][64]. Foster and Vohra [52] proved the convergence of a class of online algorithms to the set of correlated equilibria. Mertikopoulos [77] applied the online mirror descent to monotone type games and prove convergence to Nash equilibria:

Theorem 1.2.2 (Mertikopoulos [77]). *Consider a game with finitely many players and strictly monotone cost functions defined in 1.1.5 with X_i compact convex. Then the sequence $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ constructed by an OMD scheme (1.10) with $\epsilon_k = 1/k, k \in \mathbb{N}$ and T strictly convex, converge to the unique equilibrium.*

The proof relies on the definition of *Fenchel coupling* corresponding to map T . The Fenchel coupling is a function $F : X \times \mathbb{R}^{Nd} \rightarrow \mathbb{R}$ defined as

$$F(y, x) = T(x) + T^*(y) - \langle y, x \rangle, \quad \text{where} \quad T^*(y) = \sup_{z \in X} \langle y, z \rangle - T(z).$$

By definition $F(y, x) \geq 0$ and equality occurs if and only if $x \in Q_X(y)$. Setting \mathbf{x}^* as the unique Nash equilibrium in the game, Mertikopoulos [77] showed that $\lim_{k \rightarrow \infty} F(\mathbf{y}_k, \mathbf{x}^*) = 0$; because T is strictly convex this is equivalent to say $\mathbf{x}_k \rightarrow \mathbf{x}^*$.

1.3 Mean field games

In this thesis, we specifically study learning procedures for Mean Field Games. In this section, we briefly present the main model of Mean Field Games, technical details and the related literature. Mean Field Games (MFGs) were introduced by parallel works of Lasry, Lions [72][73][74][75] and Caines, Huang, Malhamé [67]. The MFGs are symmetric differential games with a continuum of players. As shown by Cardaliaguet et al. [37], the mean field games are the limit of the symmetric differential games with finite number of players N , as N tends to infinity. It is called *mean field* since the players take into account the role of other players using a mean field measure term, created by the states of the players. The MFG equilibria satisfy a system of partial differential equations (PDE) of the Hamilton-Jacobi-Bellman (HJB) type coupled with a Fokker-Planck or continuity equation. For references which studies these two equations separately, we refer to [46][16][32][51][1] for HJB equations and its numerical analysis; and [49] for Fokker-Planck equation.

The literature on MFG has been growing fast since its creation. The existence of solutions under different growth conditions on data, and local/non-local couplings are studied in [58][34][35][38][86]. Probabilistic approach on MFG, dealing with backward SDE is proposed by Carmona and Delarue [42]. Numerical approaches have been developed in [5][3][30][4][1][40][41][21][27]. A discrete analogous of MFG was proposed by Gomes et al. [55]. For a survey on the MFG, we refer to [33][59].

Let us introduce the MFGs framework precisely. Let $T > 0$ be finite, as the time horizon of the game. Set a fixed filtration $(\mathcal{F}_t)_{t \in [0, T]}$ and a constant $\sigma \geq 0$. A typical player chooses a (\mathcal{F}_t) -adapted

random process $(\alpha_t)_{t \in [0, T]}$ with values in \mathbb{R}^d , called *control*. Then its *state* $x_t \in \mathbb{R}^d, t \in [0, T]$ evolves by the dynamic $dx_t = \alpha_t dt + \sqrt{2\sigma} dB_t$, where B_t is an adapted d -dimensional Brownian motion. The agent aims to minimise the total cost function

$$\mathbb{E} \left(\int_0^T \left(L(x_t, \alpha_t) + f(x_t, m(t)) \right) dt + g(x_T, m(T)) \right), \quad (1.11)$$

over all adapted controls $(\alpha_t)_{t \in [0, T]}$. In the cost function, the map $m : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$ describes the evolving distribution of states of all players. We suppose that the agent is infinitesimal, this means that the change of its states, does not affect the measures m_t and it can assume the map m as given. For solving the optimal control problem in (1.11), one introduces an auxiliary map called the value function $u : \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ as:

$$u(x, s) = \inf_{(\alpha_t)_{t \in [s, T]}} \mathbb{E} \left(\int_s^T \left(L(x_t, \alpha_t) + f(x_t, m(t)) \right) dt + g(x_T, m(T)) \right),$$

where $dx_t = \alpha_t dt + \sqrt{2\sigma} dB_t, x_s = x$ and infimum is taken over all adapted controls $(\alpha_t)_{t \in [s, T]}$. The value function satisfies the dynamic programming relation i.e. for all $s \in [0, T]$ and $\epsilon \in [0, T - s]$ we have:

$$u(x, s) = \inf_{(\alpha_t)_{t \in [s, s+\epsilon]}} \mathbb{E} \left(\int_s^{s+\epsilon} \left(L(x_t, \alpha_t) + f(x_t, m(t)) \right) dt + u(x(s+\epsilon), s+\epsilon) \right).$$

We can deduce from the dynamic programming relation that the value function satisfies the Hamilton-Jacobi-Bellman (HJB) equation:

$$-\partial_t u - \sigma \Delta u + H(x, \nabla u(x, t)) = f(x, m(t)) \quad \text{with boundary condition} \quad u(T, x) = g(x, m(T)), \quad (1.12)$$

where the Hamiltonian $H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is defined by

$$H(x, p) = L^*(x, -p), \quad \text{with} \quad L^*(x, p) = \max_{v \in \mathbb{R}^d} \langle v, p \rangle - L(x, v).$$

Under suitable assumptions, the value function u corresponding to the variational problem (1.11), is the only solution satisfying the HJB equation (1.12). We can obtain the optimal control of problem (1.11) as a function of the value function by $\alpha(x, t) = -D_p H(x, \nabla u(x, t))$, where by ∇ we usually mean the derivative with respect to the input x .

Up to now, the evolving repartition of players $m : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$ appears in the cost function just as an input. In the case of equilibrium, the distributions $(m(t))_{t \in [0, T]}$ are sustainable if they are produced by the optimal behaviours of the players. For a given $m_0 \in \mathcal{P}(\mathbb{R}^d)$ as an initial distribution of initial states, let $(X_t)_{t \in [0, T]}$ be the process solution of the following stochastic differential equation

$$dX_t = -D_p H(X_t, \nabla u(X_t, t)) dt + \sqrt{2\sigma} dB_t, \quad \mathcal{L}(X_0) = m_0.$$

where for every random variable $Y, \mathcal{L}(Y)$ represents the law of Y on \mathbb{R} . If $m(t) = \mathcal{L}(X_t)$, then $m(t)$ has a density (denoting by $m(x, t)$) that solves the Fokker-Planck equation:

$$\partial_t m - \sigma \Delta m - \text{div}(m(x, t) D_p H(x, \nabla u(x, t))) = 0.$$

Hence the equilibrium in MFG is captured by all couples (u, m) (usually called the MFG solution) satisfying the coupled HJB and Fokker-Planck with suitable boundary conditions:

$$\begin{cases} (i) & -\partial_t u - \sigma \Delta u + H(x, \nabla u(x, t)) = f(x, m(t)) \\ (ii) & \partial_t m - \sigma \Delta m - \text{div}(m D_p H(x, \nabla u)) = 0 \\ & m(0) = m_0, u(x, T) = g(x, m(T)). \end{cases} \quad (1.13)$$

We will make more precise later in what different senses (weak or strong) the solutions satisfy the MFG equations (1.13). In contrast to the existence of classical solutions in the stochastic case ($\sigma \neq 0$ and called second order), the deterministic case ($\sigma = 0$ and first order) requires defining a more general notion of solution. We will explain these two different cases in later subchapters.

The reasonings for existence of solution satisfying MFG system (1.13) are through a fixed point argument. Informally, the idea is as follows. For an evolving distribution $\mu : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$ let u_μ be a solution of HJB equation

$$-\partial_t u - \sigma \Delta u + H(x, \nabla u(x, t)) = f(x, \mu(t)), \quad u(T, x) = g(x, \mu(T)).$$

Then, set $\Lambda(\mu) = m_\mu$ where $m_\mu : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$ is a solution of the continuity equation

$$\partial_t m - \sigma \Delta m - \operatorname{div}(m(x, t) D_p H(x, \nabla u_\mu)) = 0, \quad m(0) = m_0.$$

So the existence of MFG solution is equivalent to find a fixed point μ for the map Λ , i.e. $\Lambda(\mu) = \mu$. For the fixed point arguments, we first need to propose a compact set $\mathcal{Z} \subseteq \mathcal{P}(\mathbb{R}^d)^{[0, T]}$ and ensuring that $\Lambda(\mathcal{Z}) \subseteq \mathcal{Z}$ and $\Lambda : \mathcal{Z} \rightarrow \mathcal{Z}$ is continuous. This is where the technicalities concerning HJB and continuity solutions come into the argument.

1.3.1 Assumptions on data

The existence of solutions are ensured under assumptions on couplings f, g and initial measure m_0 . Let $\mathcal{C}^{2,1}(\mathbb{R}^d \times [0, T], \mathbb{R})$ denotes the set of functions $h(x, t)$ twice derivable in x and once derivable in t , $W^{1,\infty}(\mathbb{R}^d \times [0, T], \mathbb{R})$ denotes the Sobolev space of functions with bounded weak derivatives, $L^\infty(\mathbb{R}^d \times [0, T], \mathbb{R})$ the set of measurable functions with bounded essential supremum. Let $\mathcal{P}_1(\mathbb{R}^d)$ be the set of all probability distributions with bounded first order moment, i.e.

$$\mathcal{P}_1(\mathbb{R}^d) = \left\{ m \in \mathcal{P}(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} \|x\| \, dm(x) < +\infty \right\}.$$

The set $\mathcal{P}_1(\mathbb{R}^d)$ is equipped with Kantorovitch-Rubinstein metric d_1 defined as

$$d_1(m_1, m_2) = \sup_{f: \mathbb{R}^d \rightarrow \mathbb{R}, \text{ 1-Lipschitz}} \int_{\mathbb{R}^d} f(x) \, d(m_1 - m_2)(x).$$

Let $f : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathcal{C}^2(\mathbb{R}^d)$, $g : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathcal{C}^3(\mathbb{R}^d)$ be Lipschitz continuous and

$$\sup_{m \in \mathcal{P}_1(\mathbb{R}^d)} \|f(\cdot, m)\|_{\mathcal{C}^2} + \|g(\cdot, m)\|_{\mathcal{C}^3} < +\infty. \quad (1.14)$$

Here we mean by $\|h\|_{\mathcal{C}^k}$, for all functions $h \in \mathcal{C}^k(\mathbb{R}^d, \mathbb{R})$, the quantity $\|h\|_{\mathcal{C}^k} = \sup_x \sum_{l=0}^k \|D^l h(x)\|$. For the Hamiltonian $H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ (with $H(x, p) = L^*(x, -p)$), we assume

$$H(x, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R} \text{ is twice differentiable for all } x \in \mathbb{R}^d, \quad (1.15)$$

and there exists $C > 0$ such that

$$C^{-1} I_d \leq D_{pp} H(x, p) \leq C I_d, \quad \langle D_x H(x, p), p \rangle \geq -C(1 + \|p\|^2). \quad (1.16)$$

for all $(x, p) \in \mathbb{R}^d \times \mathbb{R}^d$. Let us recall the notation $D_v L = L_v$ and in the same way for L_x, H_x, H_p, f_x, g_x .

The initial distribution $m_0 \in \mathcal{P}(\mathbb{R}^d)$ has a smooth density with compact support, which still is denoted with m_0 . The couplings f, g are called monotone if for all $m, m' \in \mathcal{P}(\mathbb{R}^d)$:

$$\int_{\mathbb{T}^d} (f(x, m) - f(x, m')) d(m - m')(x) \geq 0, \quad \int_{\mathbb{T}^d} (g(x, m) - g(x, m')) d(m - m')(x) \geq 0. \quad (1.17)$$

1.3.2 Second order

The main theorem concerning second order MFGs is the following:

Theorem 1.3.1. *Under assumptions (1.14)(1.15)(1.16) and $\sigma > 0$, the system (1.13) has a classical solution $u, m \in \mathcal{C}^{2,1}(\mathbb{R}^d \times [0, T], \mathbb{R})$ such that*

$$\forall (x, t) \in \mathbb{R}^d \times [0, T] : \quad m(x, t) \geq 0, \quad \int_{\mathbb{R}^d} m(z, t) \, dz = 1.$$

Moreover, this solution is unique if the couplings f, g are monotone (1.17).

It can be proved that there exists a $C > 0$ depending on the constants of the data, such that for

$$\mathcal{Z} = \left\{ \mu : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d) \mid \forall t, s \in [0, T] : \int_{\mathbb{R}^d} \|x\|^2 d\mu(x, t) \leq C, \, d_1(\mu(t), \mu(s)) \leq C\sqrt{|t-s|} \right\}$$

we know $\Lambda : \mathcal{Z} \rightarrow \mathcal{Z}$ is continuous. The continuity of map Λ relies on estimation bounds for solutions of parabolic equations of Hamilton-Jacobi and Fokker-Planck type. Then, the rest of the argument is to use Schauder fixed point theorem implying a fixed point for map Λ . For more details, we refer to [33], section 3.

1.3.3 First order

Here we review the notions around the first order mean field game solutions, from the viscosity solutions of HJB equation to the weak solutions of continuity equation. The main theorem is the following:

Theorem 1.3.2 (Lasry, Lions [74]). *Under assumptions (1.14)(1.15)(1.16), there exist $u \in W^{1,\infty}(\mathbb{R}^d \times [0, T], \mathbb{R})$ and $m \in L^\infty(\mathbb{R}^d \times [0, T], \mathbb{R})$ such that (u, m) satisfies the first order MFG system*

$$\left\{ \begin{array}{ll} (i) & -\partial_t u + H(x, \nabla u(x, t)) - f(x, m(t)) = 0 \\ (ii) & \partial_t m - \operatorname{div}(m D_p H(x, \nabla u)) = 0 \\ & m(0) = m_0, \, u(x, T) = g(x, m(T)) \\ & \text{for a.e. } (x, t) \in \mathbb{R}^d \times [0, T] : \, m(x, t) \geq 0, \text{ and } \int_{\mathbb{R}^d} m(z, t) \, dz = 1. \end{array} \right. \quad (1.18)$$

in weak sense, i.e. u solves (i) in viscosity sense and m solves (ii) in distribution sense.

We will give the definition of viscosity solution later in definition 1.3.2. We say m satisfies (ii)(1.18) in distribution sense, if for all $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^d \times [0, T], \mathbb{R})$ we have

$$\int_{\mathbb{R}^d} \phi(x, 0) m_0(x) \, dx + \int_0^T \int_{\mathbb{R}^d} (\partial_t \phi(x, t) - \langle \nabla_x \phi(x, t), D_p H(x, \nabla u(x, t)) \rangle) m(x, t) \, dx = 0. \quad (1.19)$$

Theorem 1.3.2 uses the Schauder fixed point theorem by using convergence results for semi-concave functions. In the proof the crucial point is that if $(u_k)_{k \in \mathbb{N}}, u \in \mathcal{C}(\mathbb{R}^d \times [0, T], \mathbb{R})$ are uniformly semi-concave (the definition will be given later), if $u_k \rightarrow u$ locally uniformly, then $Du_k(x, t) \rightarrow Du(x, t)$ for almost every $(x, t) \in \mathbb{R}^d \times [0, T]$. The semi-concave viscosity solution u usually lacks the \mathcal{C}^1 regularity and is differentiable only almost everywhere and one of the main challenge for system (ii)(1.18), is to show that the weak solution m , that satisfies (1.19), is indeed unique.

Semi-concave functions. This property is crucial for the value functions coming from the deterministic optimal control problems (that we will see in (1.20)). The locally uniform convergence of semi-concave functions yields the convergence of derivatives almost everywhere. Look at [32] chapter 2 for the properties of semi-concave functions.

Definition 1.3.1. Let $S \subseteq \mathbb{R}^d$ be convex. We say that a function $u : S \rightarrow \mathbb{R}$ is semi-concave with linear modulus if there is $C > 0$ such that for all $x, y \in S$:

$$\lambda u(x) + (1 - \lambda)u(y) - u(\lambda x + (1 - \lambda)y) \leq C\lambda(1 - \lambda)\|x - y\|^2.$$

The function u is semi-convex if $-u$ is semi-concave.

We call the maps $(u_k)_{k \in \mathbb{N}}$ uniformly semi-concave, if the coefficient $C > 0$ in above definition, is common for all of these maps.

Theorem 1.3.3 ([32] Theorem 2.1.7, Theorem 3.3.3). Suppose $(u_k)_{k \in \mathbb{N}}, u \in \mathcal{C}(\mathbb{R}^d \times [0, T], \mathbb{R})$ are uniformly semi-concave, then

1. $\{u_k\}_{k \in \mathbb{N}}, u \in \mathcal{C}(\mathbb{R}^d \times [0, T], \mathbb{R})$ are locally Lipschitz and hence almost every where differentiable,
2. if $u_k \rightarrow u$ locally uniformly, then $Du_k(x, t) \rightarrow Du(x, t)$ for almost every $(x, t) \in \mathbb{R}^d \times [0, T]$.

Viscosity solutions. The weak solutions for linear PDEs are defined by the help of integrals and they are the solutions that satisfy the equation in distribution sense. In non-linear PDEs instead, it is no longer possible to extend the notion of solution from strong to weak by passing the derivatives to the test functions. The idea of viscosity solution is to define a weak version of solutions for non-linear PDEs. These weak solutions are compatible with classical ones if they are derivable enough.

Viscosity solutions were introduced by Crandall and Lions [47][45] for the HJB type equations. The idea stemmed from an approach called *vanishing viscosity*. The HJB equation

$$-\partial_t u + H(t, x, \nabla u(x, t)) = 0,$$

lacks the existence of global classical solution even for convex smooth Hamiltonians H (see for example [50], Chapter 3.2). Crandall, Lions [47] instead proved that the limit of solutions $\{u^\epsilon\}_{\epsilon > 0}$ of the perturbed equation

$$-\partial_t u^\epsilon + H(t, x, \nabla u^\epsilon(x, t)) = \epsilon \Delta u^\epsilon,$$

as $\epsilon \rightarrow 0$, should satisfy a series of conditions characterizing a viscosity solution. A suitable comparison principle was proposed as well, implying the uniqueness of such solution.

Definition 1.3.2. We call $u \in \mathcal{C}(\mathbb{R}^d \times [0, T], \mathbb{R})$ a viscosity solution of Hamilton-Jacobi equation $-\partial_t u + H(t, x, \nabla u(x, t)) = 0$,

- if u is a **subsolution**, that is for every test function $\phi \in \mathcal{C}^\infty(\mathbb{R}^d \times [0, T])$ such that $u - \phi$ has a local strict maximum at (t^*, x^*) we have

$$-\partial_t \phi(x^*, t^*) + H(t^*, x^*, \nabla \phi(x^*, t^*)) \leq 0.$$

- if u is a **supersolution**, that is for every test function $\phi \in \mathcal{C}^\infty(\mathbb{R}^d \times [0, T])$ such that $u - \phi$ has a local strict minimum at (t^*, x^*) we have

$$-\partial_t \phi(x^*, t^*) + H(t^*, x^*, \nabla \phi(x^*, t^*)) \geq 0.$$

The comparison principle for viscosity solutions plays a key role for the uniqueness. There are a variety of comparison principles for viscosity solutions which differs on the regularity assumptions on H and solution u .

Theorem 1.3.4 ([51] section 5, [32] section 5). *Let $H : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ be continuous and satisfies*

$$\forall t, x, p, q : |H(t, x, p) - H(t, x, q)| \leq K(\|x\| + 1)\|p - q\|,$$

for some $K > 0$. Suppose also that for all $R > 0$, there exists $m_R : [0, \infty[\rightarrow [0, \infty[$ continuous, non-decreasing, with $m_R(0) = 0$ such that

$$\forall x, y \in B(0, R), p \in \mathbb{R}^d, t \in [0, T] : |H(t, x, p) - H(t, y, p)| \leq m_R(\|x - y\|) + m_R(\|x - y\|\|p\|).$$

Let $u_1, u_2 \in \mathcal{C}([0, T] \times \mathbb{R}^d)$ be, respectively, viscosity sub-solution and super-solution of the equation

$$-u_t + H(t, x, \nabla u(x, t)) = 0, \quad (x, t) \in (0, T) \times \mathbb{R}^d$$

Then if $u_1(T, x) \leq u_2(T, x)$ for all $x \in \mathbb{R}^d$ then

$$\forall (x, t) \in [0, T] \times \mathbb{R} : u_1(x, t) \leq u_2(x, t).$$

The comparison principle gives the uniqueness of HJB solution with a boundary condition: if there are two viscosity solutions u_1, u_2 which are equal on the boundary, by the comparison principle we have $u_1 = u_2$. The previous theorem can be used only for linear-like Hamiltonians and cannot be applied for our very first example $H(t, x, p) = \frac{1}{2}\|p\|^2$. Here is another comparison principle covering the quadratic Hamiltonian but with assuming the periodicity for solutions u .

Theorem 1.3.5. *Suppose there are two continuous functions $u_1, u_2 : \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ such that they are periodic in space input with same periodicity, and $u_1(x, T) = u_2(x, T)$ for all $x \in \mathbb{R}^d$. Assume a continuous Hamiltonian $H : \mathbb{R}^d \times \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ with*

$$|H(x, s, p) - H(x, t, p)| \leq C|s - t|, \quad \langle D_x H(x, t, p), p \rangle \geq -C(1 + \|p\|^2).$$

If u_1 (resp. u_2) is the sub-solution (resp. super-solution) of HJB equation $-\partial_t u(x, t) + H(x, t, \nabla u(x, t)) = 0$, then we have $u_1(x, t) \leq u_2(x, t)$ for all $(x, t) \in \mathbb{R}^d \times [0, T]$.

Proof. Without loss of generality, suppose u_1, u_2 are 1-periodic in space input x ; we work then with d -dimensional torus \mathbb{T}^d as the set for input x . Suppose there is $(x, t) \in \mathbb{T}^d \times [0, T]$ such that $u_1(x, t) - u_2(x, t) > \sigma$ for some $\sigma > 0$. Let

$$\Phi_\epsilon(x, t, y, s) = u_1(x, t) - u_2(y, s) - \lambda(2T - t - s) - \frac{1}{2\epsilon}(\|x - y\|^2 + (t - s)^2),$$

and

$$(x_\epsilon, t_\epsilon, y_\epsilon, s_\epsilon) \in \operatorname{argmax}_{(x, t, y, s) \in (\mathbb{T}^d \times [0, T])^2} \Phi_\epsilon(x, t, y, s).$$

We can choose $\lambda > 0$ small enough such that $|t_\epsilon - T|, |s_\epsilon - T| > \delta > 0$ with δ independent of ϵ . If we set

$$\phi(x, t) = u_2(y_\epsilon, s_\epsilon) + \lambda(2T - t - s_\epsilon) + \frac{1}{2\epsilon}(\|x - y_\epsilon\|^2 + (t - s_\epsilon)^2),$$

then $(x_\epsilon, t_\epsilon) \in \operatorname{argmax} u_1(x, t) - \phi(x, t)$. Since u_1 is sub-solution we can conclude:

$$-\partial_t \phi(x_\epsilon, t_\epsilon) + H(x_\epsilon, t_\epsilon, \nabla_x \phi(x_\epsilon, t_\epsilon)) \leq 0, \quad \text{which gives} \quad \lambda - \frac{1}{\epsilon}(t_\epsilon - s_\epsilon) + H(x_\epsilon, t_\epsilon, \frac{1}{\epsilon}(x_\epsilon - y_\epsilon)) \leq 0.$$

Similarly if

$$\phi(y, s) = u_1(x_\epsilon, t_\epsilon) - \lambda(2T - t_\epsilon - s) - \frac{1}{2\epsilon}(\|x_\epsilon - y\|^2 + (t_\epsilon - s)^2),$$

then $(y_\epsilon, s_\epsilon) \in \operatorname{argmin} u_2(y, s) - \phi(y, s)$ and u_2 is super-solution. Hence

$$-\partial_s \phi(y_\epsilon, s_\epsilon) + H(y_\epsilon, s_\epsilon, \nabla_y \phi(y_\epsilon, s_\epsilon)) \geq 0, \quad \text{which gives} \quad -\lambda - \frac{1}{\epsilon}(t_\epsilon - s_\epsilon) + H(y_\epsilon, s_\epsilon, \frac{1}{\epsilon}(x_\epsilon - y_\epsilon)) \geq 0$$

So then, comparing two last inequalities gives:

$$H(y_\epsilon, s_\epsilon, \frac{1}{\epsilon}(x_\epsilon - y_\epsilon)) - H(x_\epsilon, t_\epsilon, \frac{1}{\epsilon}(x_\epsilon - y_\epsilon)) \geq 2\lambda.$$

Choosing (x^*, t^*, y^*, s^*) an accumulation point of $\{(x_\epsilon, t_\epsilon, y_\epsilon, s_\epsilon)\}_{\epsilon \in \mathbb{R}^+}$ as $\epsilon \rightarrow 0$, then by

$$\|x_\epsilon - y_\epsilon\|, |t_\epsilon - s_\epsilon| = o(\sqrt{\epsilon}),$$

we have $x^* = y^*, t^* = s^*$. On the other hand if we set $p_\epsilon = \frac{1}{\epsilon}(x_\epsilon - y_\epsilon)$ then $x_\epsilon = y_\epsilon + \epsilon p_\epsilon$ and $\epsilon \|p_\epsilon\|^2 \rightarrow 0$. We have

$$\begin{aligned} 0 < 2\lambda &\leq H(y_\epsilon, s_\epsilon, \frac{1}{\epsilon}(x_\epsilon - y_\epsilon)) - H(x_\epsilon, t_\epsilon, \frac{1}{\epsilon}(x_\epsilon - y_\epsilon)) = H(y_\epsilon, s_\epsilon, p_\epsilon) - H(y_\epsilon + \epsilon p_\epsilon, t_\epsilon, p_\epsilon) \\ &\leq C|t_\epsilon - s_\epsilon| - \epsilon \langle D_x H(z_\epsilon, t_\epsilon, p_\epsilon), p_\epsilon \rangle \leq C|t_\epsilon - s_\epsilon| + \epsilon C(1 + \|p_\epsilon\|^2), \end{aligned}$$

for some $z_\epsilon \in [x_\epsilon, y_\epsilon]$. The last expression tends to 0 when $\epsilon \rightarrow 0$, it is a contradiction with $\lambda > 0$. \square

Optimal control problem. The optimal control problem

$$\inf_{\alpha: [t, T] \rightarrow \mathbb{R}^d} \int_t^T \left(L(X^{x,t}[\alpha](s), \alpha(s)) + f(t, X^{x,t}[\alpha](s)) \right) ds + g(X^{x,t}[\alpha](T)), \quad (1.20)$$

with $X^{x,t}[\alpha](s) = x + \int_t^s \alpha_r dr$, has a close relation with the Hamilton-Jacobi equation

$$-\partial_t u + H(x, \nabla u(x, t)) - f(x, t) = 0 \quad (1.21)$$

with $H(x, p) = L^*(x, -p)$. Indeed, if $u(x, t)$ is the value of the minimisation problem (1.20), then u is a viscosity solution of (1.21) satisfying $u(x, T) = g(x)$.

Theorem 1.3.6 ([32], Theorem 7.2.4). *Suppose the following conditions hold:*

1. *the conditions (1.15)(1.16) hold for H with $H(x, p) = L^*(x, -p)$,*
2. *the function $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is Lipschitz continuous in time and there is $C > 0$ such that*

$$\max_{t \in [0, T]} \|f(\cdot, t)\|_{C^2}, \|g(\cdot)\|_{C^2} \leq C, \quad (1.22)$$

Then the value function u of optimal control problem (1.20) is the only bounded uniformly continuous viscosity solution of the equation (1.21) satisfying $u(x, T) = g(x)$. Moreover, u is semi-concave, locally Lipschitz and almost everywhere derivable.

For all $(x, t) \in \mathbb{R}^d \times [0, T]$ let $\mathcal{A}(x, t)$ be the set of optimal control $\alpha : [t, T] \rightarrow \mathbb{R}^d$ minimizing the variational problem (1.20). The Euler-Lagrange optimality condition, characterize the elements in $\mathcal{A}(x, t)$. Let us recall the notation $D_v L = L_v$ and the same for L_x, f_x, g_x .

Theorem 1.3.7. *Suppose (1.15)(1.16)(1.22) hold. If $\alpha \in \mathcal{A}(x, t)$, then α is of class C^1 on (t, T) with*

$$\begin{aligned} \frac{d}{ds} L_v(X^{x,t}[\alpha](s), \alpha(s)) &= L_x(X^{x,t}[\alpha](s), \alpha(s)) + f_x(X^{x,t}[\alpha](s), s), \quad s \in (t, T), \\ L_v(X^{x,t}[\alpha](T), \alpha(T)) &= -g_x(X^{x,t}[\alpha](T)), \end{aligned} \quad (1.23)$$

where $X^{x,t}[\alpha](s) = x + \int_t^s \alpha(r) dr$. In particular, there is a constant $C > 0$ such that, for $(x, t) \in [0, T] \times \mathbb{R}^d$ and any $\alpha \in \mathcal{A}(x, t)$ we have $\|\alpha\|_\infty \leq C$.

Proof. The characterization in (1.23) is classical, see for example ([32] Theorem 6.2.4). For the boundedness problem, set $X, P : (t, T] \rightarrow \mathbb{R}^d$ as $X(s) = X^{x,t}[\alpha](s)$, $P(s) = -L_v(X(s), \alpha(s))$ for $s \in (t, T]$. By (1.23) we have

$$\dot{P}(s) = -L_x(X(s), \alpha(s)) - f_x(X(s), s), \quad s \in (t, T].$$

By relation $H(x, p) = L^*(x, -p)$, we have $\alpha(s) = -H_p(X(s), P(s))$, $L_x(X(s), \alpha(s)) = -H_x(X(s), P(s))$. So rewriting the last ODE gives

$$\dot{P}(s) = H_x(X(s), P(s)) - f_x(X(s), s), \quad s \in (t, T].$$

Multiplying both side by $P(s)$ and using (1.16)(1.22) implies

$$\exists C' > 0 : \quad \frac{d}{ds} \left(\frac{1}{2} \|P(s)\|^2 \right) \geq -C'(1 + \|P(s)\|^2), \quad \|P(T)\| \leq C',$$

By a direct application of Gronwall theorem we can assert the existence of a constant $C'' > 0$ independent of t , such that $\|P(s)\| \leq C''$ for all $s \in (t, T]$. Afterwards, by strong convexity condition (1.16) we can demonstrate as well the uniform boundedness of $\alpha(s)$ for all $s \in (t, T]$. \square

The next theorem asserts that the points (x, t) where the value function u is derivable coincide with the points (x, t) where the optimal control set $\mathcal{A}(x, t)$ is singleton.

Theorem 1.3.8 ([33], Lemma 4.9). *Let $(x, t) \in [0, T] \times \mathbb{R}^d$, $\alpha \in \mathcal{A}(x, t)$ and let us set $x(s) = x + \int_t^s \alpha_r \, dr$. Then*

- *(Uniqueness of the optimal control along optimal trajectories) for any $s \in (t, T]$, the restriction of α to $[s, T]$ is the unique element of $\mathcal{A}(s, x(s))$.*
- *(Uniqueness of the optimal trajectories) $\nabla u(x, t)$ exists if and only if $\mathcal{A}(x, t)$ is reduced to a singleton. In this case, $-D_p H(x, \nabla u(x, t)) = \alpha(t)$ where $\mathcal{A}(x, t) = \{\alpha\}$.*

In general, the value function u is not necessarily derivable but since it is semi-concave, it is almost everywhere derivable; hence $\mathcal{A}(x, t)$ is singleton for almost every (x, t) .

Continuity equation. Since the value function u is Lipschitz continuous, then it is almost everywhere derivable and by Theorem 1.3.8, for almost every (x, t) the optimal control $\mathcal{A}(x, t)$ is singleton. Let us consider a measurable selection β of the correspondence \mathcal{A} . Let Φ be the associated flow, that is

$$\Phi(x, t, s) = x + \int_t^s \beta(x, t)(\tau) \, d\tau, \quad x \in \mathbb{R}^d, \quad t, s \in [0, T], \quad t \leq s.$$

Let $\mu(t) \in \mathcal{P}(\mathbb{R}^d)$ be the transportation of initial distribution m_0 by the flow Φ , i.e.

$$\mu(t) = \Phi(\cdot, 0, t) \# m_0, \quad t \in [0, T].$$

Despite the lack of existence of derivative of u everywhere, the following theorem asserts that $\mu(t)$ actually satisfies in distribution sense the continuity equation corresponding to the vector field $-D_p H(t, \nabla u(x, t))$ and it is indeed absolutely continuous.

Theorem 1.3.9. *The transported distribution $\mu(t) = \Phi(\cdot, 0, t) \# m_0$, $t \in [0, T]$, is absolutely continuous and satisfies the continuity equation*

$$\partial_t \mu - \operatorname{div}(\mu D_p H(x, \nabla u)) = 0$$

in a weak sense. That is for all $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^d \times [0, T], \mathbb{R})$ we have

$$\int_{\mathbb{R}^d} \phi(0, x) m_0(x) \, dx + \int_0^T \int_{\mathbb{R}^d} (\partial_t \phi(x, t) - \langle \nabla_x \phi(x, t), D_p H(x, \nabla u(x, t)) \rangle) \mu(x, t) \, dx = 0.$$

1.4 Our contributions

The main question in this thesis is to find learning procedures in mean field games and to investigate if they converge to an equilibrium. Those are games involving a non atomic set of players each of them is choosing an action in an infinite dimensional space (a trajectory in the euclidean space starting from some player dependent initial position). The situation is therefore much more complex than usual finite games. However, our thesis takes inspiration from the learning schemes in static finite games to design adaptive procedures that converge to equilibria in several classes of MFGs.

More precisely we extend the fictitious play and the online mirror descent procedures to MFGs and prove their convergence when the game is potential or monotone and provide approximations theorems when the game is discretized in time and space. For example, the fictitious play algorithm extends to MFGs as follows. Suppose the differential game is played in many rounds, each round containing the whole time interval $[0, T]$. At every round n , an estimation of the evolving measure $(m_t)_{t \in [0, T]}$ is computed as the time average of the observed distributions in precedent rounds $0, \dots, n-1$. The agents then behave optimally regarding to this estimation and then a new estimate is compute similarly in the next step.

Chapter 2 concentrates on convergence of fictitious play in potential mean field games. Those have already been defined by Lasry and Lions [74] and are such that the equilibrium can be obtained by minimizing a functional, called the Potential, over a suitable complex space of functions. We can formally show that this class of game is a very natural extension of the potential game model of Monderer and Shapley [80] (defined for finitely many players and finite dimensional strategy spaces). Hence, it is quite reasonable to start our thesis with this framework and to expect that the fictitious play will slightly decrease the potential. This would imply that the time average behaviour converges to the minimiser of the potential and so to a MFG equilibrium. This is what happens. Our approach for attacking the problem is however more complex than in Monderer and Shapley and uses different spaces for second and first order MFG. The space we work with in the case of second order are the classical solutions of PDEs of HJB and Fokker-Planck type, while one needs to work with the space of trajectories in the first order type. The potential for the case of the second order MFG was already defined by Lasry and Lions however, for the case of first order MFGs, we provide a new and convenient representation of the potential as a function of measures over trajectories. We finish this chapter by proving an approximation theorem showing that the fictitious play procedure applied to a differential game with a finite number of players, converges to the MFG equilibrium as the number of players goes to infinity.

In chapter 3, the goal is to prove convergence of fictitious play, and more generally of online mirror descent schemes, in monotone mean field games (the second natural class of game in which one should expect convergence). We start by observing that a MFG model deals with a continuum of players choosing each from a player dependent infinite dimensional space (i.e. the space of trajectories with an initial condition depending on the player's position) and that the cost of a player does not depend on the identities of the players but only on their distribution. This leads us to generalize this model by working in an abstract model we call anonymous game. This is a normal form game with a non atomic set of players, a player dependent action set and an anonymous payoff function. Keeping in mind that the first order MFGs is our principal application, we provide conditions on the game (such as the unique minimizer condition) under which fictitious play and online mirror descent procedures converge to equilibria in all monotone anonymous games (and so also in all monotone first order MFGs, as desired). The question of convergence in potential anonymous games (expected to hold), is postponed to a future work, because we already know the answer in MFGs from the previous chapter.

In chapter 4 we look at a discrete (in time and space) version of MFGs introduced by Gomes et al. [55], and investigate the convergence of fictitious play in this model. We prove that this framework

is a particular instance of our general anonymous game model studied in the previous chapter 3 and conclude convergence to the equilibrium in all monotone discrete MFGs. Finally, we provide a general approximation theorem (which has nothing to do with learning per se) by proving convergence of the discrete MFGs to the first order continuous MFGs when the mesh of the discretization (in time and space) goes to zero (in a well chosen way). This confirms that the discrete model is close to the continuous system (as it was expected in the field).

Before going to a more detailed exposition of the results, we want to remark that the dynamic nature of MFGs implies that one can design more sophisticated learning algorithms better adapted to the dynamic nature of the game (for example at half of the game a player can observe that the assumptions he made was false and adapts his behaviour accordingly). Investigating this more complex behaviour is delegated to a future work. Nevertheless, we believe that this thesis is a good starting point that builds a bridge between learning in mean field games and learning in classical game theory.

1.4.1 Fictitious play in potential MFG

Chapter 2 is devoted to our first contribution, that is an application of fictitious play in Potential MFGs. Assuming the data are periodic in state variable x , we work specially with the solutions of MFG system which are periodic in x . Due to this, we consider the d -dimensional torus \mathbb{T}^d as the set of states.

We deal with the definition and corresponding results of derivative of functions with respect to the measure arguments; they are borrowed from [37]. For a function $K : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$, a derivative of K with respect to the measure argument is a map denoted by $\frac{\delta K}{\delta m}$ such that $\frac{\delta K}{\delta m} : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ is continuous and for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$

$$\lim_{s \rightarrow 0} \frac{K((1-s)m + sm') - K(m)}{s} = \int_{\mathbb{T}^d} \frac{\delta K}{\delta m}(x, m) d(m' - m)(x). \quad (1.24)$$

We denote the last expression with $\frac{\delta K}{\delta m}(m)(m' - m)$ as well, that is the derivative of K at point m and direction $m' - m$. We call a MFG (second or first order) a Potential mean field game if its couplings f, g possess potentials. That means there exist continuously differentiable maps $F, G : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ such that

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

where $\frac{\delta F}{\delta m}, \frac{\delta G}{\delta m}$ are the derivative with respect to the measure argument. The following Proposition characterize all derivable functions which can be written as a derivative of a potential.

Proposition 1.4.1. *The map $f : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ derives from a potential, if and only if,*

$$\frac{\delta f}{\delta m}(x, m, y) + \phi(x, m) = \frac{\delta f}{\delta m}(y, m, x) + \phi(y, m) \quad \forall x, y \in \mathbb{T}^d, \forall m \in \mathcal{P}(\mathbb{T}^d), \quad (1.25)$$

for some $\phi : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ where $\frac{\delta f}{\delta m}(x, m, y) = \frac{\delta f(x, \cdot)}{\delta m}(y, m)$ for all $x, y \in \mathbb{T}^d, m \in \mathcal{P}(\mathbb{T}^d)$.

If the couplings f, g possess potentials, then the second order MFG equilibrium can be obtained by solving a variational problem.

Theorem 1.4.1 ([74], section 2.6). *Let \mathcal{K} be consisting of all $(m, w) \in \mathcal{C}^0([0, T] \times \mathbb{T}^d; \mathbb{R}) \times \mathcal{C}^0((0, T) \times \mathbb{T}^d; \mathbb{R}^d)$ satisfying*

$$\partial_t m - \sigma \Delta m + \operatorname{div}(w) = 0 \quad \text{in } (0, T) \times \mathbb{T}^d, \quad m(0) = m_0, \quad (1.26)$$

in sense of distribution. Define $\Phi : \mathcal{K} \rightarrow \mathbb{R}$ by

$$\Phi(m, w) = \int_0^T \int_{\mathbb{T}^d} m(x, t) H^*(x, -w(x, t)/m(x, t)) dx dt + \int_0^T F(m(t)) dt + G(m(T)), \quad (1.27)$$

for $(m, w) \in \mathcal{K}$. Then $(\tilde{m}, \tilde{w}) \in \mathcal{K}$ is a solution of $\min_{(m, w) \in \mathcal{K}} \Phi(m, w)$ if and only if

$$\tilde{w}(x, t) = -\tilde{m}(x, t)DH_p(x, \nabla \tilde{u}(x, t)), \quad \text{for all } (x, t) \in (0, T) \times \mathbb{T}^d,$$

for a solution (\tilde{u}, \tilde{m}) of MFG system (1.13).

The fictitious play for the case of second order MFGs, where the solutions satisfy the MFG system in the classical sense, takes the form :

$$\begin{cases} (i) & -\partial_t u^{n+1} - \sigma \Delta u^{n+1} + H(x, \nabla u^{n+1}(x, t)) = f(x, \bar{m}^n(t)), \\ (ii) & \partial_t m^{n+1} - \sigma \Delta m^{n+1} - \operatorname{div}(m^{n+1} D_p H(x, \nabla u^{n+1})) = 0, \\ & m^{n+1}(0) = m_0, \quad u^{n+1}(x, T) = g(x, \bar{m}^n(T)) \end{cases} \quad (1.28)$$

with

$$\bar{m}^{n+1}(t) = \frac{1}{n+1} \sum_{k=1}^{n+1} m^k(t), \quad \text{for all } t \in [0, T]. \quad (1.29)$$

The equation (1.28)(i) refers to an optimal control problem with $m = \bar{m}^n$; this means the players set their belief at round $n+1$ equal to the average of the measures in previous steps. After the players fixed their optimal control derived by $m = \bar{m}^n$, the equation (1.28)(ii) describes how the realized distribution is computed as a result of the optimal control $\alpha(x, t) = -D_p H(x, \nabla u^{n+1}(x, t))$. At the end of the $(n+1)$ -th round the players refine their estimation with (1.29).

Theorem 1.4.2. *Under suitable assumptions (see section 2.1.1), 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 (1.13). If, in addition, the monotonicity condition (1.17) holds, then the whole sequence $\{(u^n, m^n)\}_{n \in \mathbb{N}}$ converges to the unique solution of (1.13).*

The proof is based on the definition of the potential Φ as in (1.27). If $w^n(x, t) = -m^n(x, t)D_p H(x, \nabla u^n(x, t))$ then $(m^n, w^n) \in \mathcal{K}$. Moreover, the values $\phi_n = \Phi(m^n, w^n)$ are almost decreasing, that means there are $C > 0$ and $a_n > 0$ for $n \in \mathbb{N}$, such that

$$\forall n \in \mathbb{N}: \quad \phi_{n+1} - \phi_n \leq -\frac{a_n}{n} + \frac{C}{n^2}. \quad (1.30)$$

Writing the exact expression for a_n and using the above inequality, it implies that $a_n \rightarrow 0$ which yields our desired result.

For the case of first order, due to the lack of regularity for solutions, we instead work with the space of continuous trajectories $\Gamma = \mathcal{C}([0, T], \mathbb{T}^d)$ and measures over Γ . Set the potential for first order case as $\Phi : \mathcal{P}(\Gamma) \rightarrow \mathbb{R}$ with

$$\Phi(\eta) := \int_{\Gamma} \int_0^T L(\gamma(t), \dot{\gamma}(t)) \, dt \, d\eta(\gamma) + \int_0^T F(e_t \# \eta) \, dt + G(e_T \# \eta), \quad (1.31)$$

where for every $t \in [0, T]$, the evaluation at instant t is the map $e_t : \Gamma \rightarrow \mathbb{R}^d$, $e_t(\gamma) = \gamma(t)$; hence the push-forward measure $e_t \# \eta$ captures the margin of distribution η at time t .

Remark 1.4.1. *Informally speaking, the definition of potential in second order (1.27) and first order case (1.31) are identical. First, the measures $m(t)$ and $e_t \# \eta$ are both capturing the distribution of agents' states at instant $t \in [0, T]$. Second, $w(x, t)/m(x, t)$ is equal to the drift $\alpha(x, t)$ that makes the measures $(m(t))_{t \in [0, T]}$ evolve; these drifts can be considered as the derivative of the trajectories in distribution η as well. Hence we can write*

$$\int_{\Gamma} L(\gamma(t), \dot{\gamma}(t)) \, d\eta(\gamma) = \int_{\Gamma} L(\gamma(t), \alpha(\gamma(t), t)) \, d\eta(\gamma) = \int_{\mathbb{T}^d} L(x, \alpha(x, t)) \, d(e_t \# \eta)(x)$$

$$= \int_{\mathbb{T}^d} m(x, t) L(x, \alpha(x, t)) \, dx = \int_{\mathbb{T}^d} m(x, t) H^*(x, -w(x, t)/m(x, t)) \, dx \, dt,$$

since $L(x, v) = H^*(x, -v)$.

Remark 1.4.2. The crucial point about the potential Φ defined in (1.31), is that for all $\eta, \eta' \in \mathcal{P}(\Gamma)$ we have

$$\frac{\delta \Phi}{\delta \eta}(\eta)(\eta' - \eta) = \int_{\Gamma} \left(\int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), e_t \# \eta)) \, dt + g(\gamma(T), e_T \# \eta) \right) d(\eta' - \eta)(\gamma).$$

Roughly speaking, it implies that the minimiser η of Φ are concentrated on optimal curves with respect to the $(e_t \# \eta)_{t \in [0, T]}$, i.e. they are equilibria.

For a probability measure over set of trajectories $\eta \in \mathcal{P}(\Gamma)$, let $\gamma^\eta : \mathbb{T}^d \rightarrow \mathcal{AC}([0, T], \mathbb{T}^d)$ be a measurable function such that for any $x \in \mathbb{T}^d$ the trajectory $\gamma_x^\eta \in \mathcal{AC}([0, T], \mathbb{T}^d)$ be an optimal solution to

$$\inf_{\gamma \in \mathcal{AC}([0, T], \mathbb{T}^d), \gamma(0)=x} \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), e_t \# \eta)) \, dt + g(\gamma(T), m(T)).$$

The fictitious play in the case of first order MFG takes the following form:

$$\begin{aligned} (i) \quad \eta^{n+1} &:= \gamma^{\bar{\eta}^n} \# m_0, \\ (ii) \quad \bar{\eta}^{n+1} &:= \frac{1}{n+1} \sum_{k=1}^{n+1} \eta^k. \end{aligned} \tag{1.32}$$

The equation (i)–(1.32) captures a distribution of curves η^{n+1} with support on the optimal curves with respect to the average distribution $\bar{\eta}^n$; the equation (ii)–(1.32) uses η^{n+1} to revise the average.

Theorem 1.4.3. Under suitable assumptions (see section 2.1.1) the sequences $(\bar{\eta}^n, \eta^n)$ is pre-compact in $\mathcal{P}(\Gamma) \times \mathcal{P}(\Gamma)$ and any cluster point $(\bar{\eta}, \eta)$ satisfies the following: $\bar{\eta} = \eta$ and, if we set

$$\begin{aligned} m(t) &:= e_t \# \eta, \\ u(t, x) &:= \inf_{\gamma \in H^1, \gamma(t)=x} \int_0^T (L(\gamma(s), \dot{\gamma}(s)) + f(\gamma(s), e_s \# \eta)) \, ds + g(\gamma(T), m(T)), \end{aligned} \tag{1.33}$$

then the pair (u, m) is a solution to the first order MFG system (1.18). If furthermore (1.17) holds, then the entire sequence $(\bar{\eta}^n, \eta^n)$ converges.

The proof idea is similar to the case of second order by using the potential (1.31). We can prove an inequality similar to (1.30) with $\phi_n = \Phi(\bar{\eta}^n)$ and obtaining our desired result. We proved that the accumulation points of distributions $\bar{\eta}^n$ are the equilibrium distributions $\tilde{\eta}$, i.e.

$$e_0 \# \tilde{\eta} = m_0, \quad \text{supp}(\tilde{\eta}) \subseteq \operatorname{argmin}_{\gamma \in \mathcal{AC}([0, T], \mathbb{T}^d)} J(\gamma, \tilde{\eta}). \tag{1.34}$$

where

$$J(\gamma, \eta) = \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), e_t \# \eta)) \, dt + g(\gamma(T), e_T \# \eta).$$

We show in section 2.4 that we can construct the first order MFG solution (u, m) from an equilibrium distribution $\tilde{\eta}$ as in (1.33).

Our last contribution in chapter 2 concerns with the question of convergence of fictitious play in a symmetric differential game with finite number of players; as one expects, these fictitious play should converge to an equilibrium of first order MFG, as N tends to infinity. For every $N \in \mathbb{N}$, fix a sequence of initial states $x_1^N, x_2^N, \dots, x_N^N \in \mathbb{T}^d$ such that:

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

where $m_0^N = \frac{1}{N} \sum_{i=1}^N \delta_{x_i^N}$. Define the sequences $\bar{\eta}^{n,N}, \eta^{n,N} \in \mathcal{P}(\Gamma)$, for $n \in \mathbb{N}$ in the following way:

$$\begin{aligned}\eta^{n+1,N} &= \frac{1}{N} (\delta_{\gamma_{x_1^N}^{n+1,N}} + \delta_{\gamma_{x_2^N}^{n+1,N}} + \dots + \delta_{\gamma_{x_N^N}^{n+1,N}}) \\ \bar{\eta}^{n+1,N} &= \frac{1}{n+1} (\eta^{1,N} + \eta^{2,N} + \dots + \eta^{n+1,N})\end{aligned}\tag{1.35}$$

where $\gamma_{x_i^N}^{n+1,N}$ is an optimal path which is a solution of $\inf_{\gamma \in H^1(\gamma(0)=x_i^N)} J(\gamma, \bar{\eta}^{n,N})$.

Theorem 1.4.4. *Consider the fictitious play for the N -player game as described in (1.35) and let $\bar{\eta}^N$ be an accumulation distribution of $(\bar{\eta}^{n,N})_{n \in \mathbb{N}}$. Then every accumulation point of the pre-compact set of $\{\bar{\eta}^N\}_{N \in \mathbb{N}}$ is an MFG equilibrium. If furthermore the monotonicity condition (1.17) holds, then $(\bar{\eta}^N)$ has a limit which is the MFG equilibrium.*

1.4.2 Non atomic anonymous games

Inspired specifically by the games with continuum of players appeared in [89][76][24][25], we try to propose in chapter 3 a general framework, with first order mean field games as a special case; this is the model of non atomic anonymous games.

Let I be the set of players and $\lambda \in \mathcal{P}(I)$ a prior non-atomic probability measure on I modelling the repartition of players on I . An individual $i \in I$ chooses an action a from a player dependent set $A_i \subset V$ and pays a cost of the form $J(a, \eta)$ where η is the distribution of actions chosen by other players. The set of admissible profiles are measurable functions $\Psi : I \rightarrow V$ such that $\Psi(i) \in A_i$ for all $i \in I$. For every admissible profile $\Psi : I \rightarrow V$, the measure $\Psi\#\lambda \in \mathcal{P}(V)$ is the push-forward of λ by map Ψ and captures the distribution of actions chosen by players in profile Ψ . An admissible profile $\tilde{\Psi}$ is called a Nash equilibrium if

$$\tilde{\Psi}(i) \in \arg \min_{a \in A_i} J(a, \tilde{\Psi}\#\lambda) \quad \text{for } \lambda\text{-almost every } i \in I,$$

and the corresponding distribution $\tilde{\eta} = \tilde{\Psi}\#\lambda$ over set of actions V , is called a Nash (or equilibrium) distribution. We especially work with the sub-class of anonymous games with a monotone cost. The cost function $J : V \times \mathcal{P}(V) \rightarrow \mathbb{R}$ is called monotone if for every $\eta, \eta' \in \mathcal{P}(V)$ the following inequality holds:

$$\int_V (J(a, \eta) - J(a, \eta')) \, d(\eta - \eta')(a) \geq 0, \tag{1.36}$$

Theorem 1.4.5. *Let $G = (I, \lambda, V, (A_i)_{i \in I}, J)$ be a non atomic anonymous game. Under suitable assumptions the game G will admit at least a Nash equilibrium. Moreover, the equilibrium is unique under monotonicity condition (1.36).*

Our next step is to propose a learning procedure similar to fictitious play and online mirror descent for anonymous games and prove their convergence when the monotonicity condition (1.36) holds. The fictitious play in anonymous games reads as follows:

$$\begin{aligned}(i) \quad \Psi_{n+1}(i) &= \arg \min_{a \in A_i} J(a, \bar{\eta}_n), \quad \text{for } \lambda\text{-almost every } i \in I, \\ (ii) \quad \eta_{n+1} &= \Psi_{n+1}\#\lambda, \\ (iii) \quad \bar{\eta}_{n+1} &= \frac{n}{n+1} \bar{\eta}_n + \frac{1}{n+1} \eta_{n+1}.\end{aligned}\tag{1.37}$$

Theorem 1.4.6. *Consider a non atomic anonymous game with a monotone cost. Under suitable conditions (appeared in Theorem 3.3.1), for the sequence $\eta^n, \bar{\eta}^n$ constructed in (1.37) we have $\eta_n, \bar{\eta}_n \xrightarrow{d_1} \tilde{\eta}$ where $\tilde{\eta} \in \mathcal{P}_G(V)$ is the unique Nash equilibrium distribution.*

The proof is done by defining the quantities $\phi_n = \int_V J(a, \bar{\eta}_n) d(\bar{\eta}_n - \eta_{n+1})(a)$ for $n \in \mathbb{N}$, and showing the inequalities:

$$\forall n \in \mathbb{N} : \quad \phi_{n+1} - \phi_n \leq -\frac{1}{n+1} \phi_n + \frac{\epsilon_n}{n}. \quad (1.38)$$

hold for some values $\{\epsilon_n\}_{n \in \mathbb{N}}$ with $\lim_{n \rightarrow \infty} \epsilon_n = 0$. The inequalities (1.38) with $\phi_n \geq 0$ give $\phi_n \rightarrow 0$, which implies the convergence of $\bar{\eta}_n$ towards the equilibrium distribution.

The second learning procedure we study for anonymous games is an analogue of the online mirror descent. It reads as follows:

$$\begin{aligned} (i) \quad \Phi_{n+1}(i) &= \Phi_n(i) - \beta_n \nabla_a J(\Psi_n(i), \eta_n), \quad \text{for every } i \in I \\ (ii) \quad \Psi_{n+1}(i) &= Q_{A_i}(\Phi_{n+1}(i)), \quad \text{for every } i \in I \\ (iii) \quad \eta_{n+1} &= \Psi_{n+1} \# \lambda. \end{aligned} \quad (1.39)$$

where $Q_A(y) = \arg \max_{a \in A} \langle y, a \rangle - h(a)$ for a strongly convex map h . The first two (i, ii)–(1.39) expressions are just an OMD procedure that the player i follows. The third expression (iii)–(1.39) entangles the parallel OMD procedures that are being done simultaneously by all players. In the case of monotone cost function J , this collective learning yields the convergence of η_n to equilibrium.

Theorem 1.4.7. *Suppose J is monotone and convex with respect to the first input. Let one applies the OMD algorithm proposed in (3.9) for $\beta_n = \frac{1}{n}$. Under suitable conditions (appeared in Theorem 3.4.1) $\eta_n = \Psi_n \# \lambda$ converges to $\tilde{\eta} = \tilde{\Psi} \# \lambda$ where $\tilde{\eta} \in \mathcal{P}(V)$ is the unique Nash equilibrium distribution.*

The proof is inspired by the method in [77], and it proceeds by using the quantity:

$$\forall n \in \mathbb{N} : \quad \phi_n = \int_I (h(\tilde{\Psi}(i)) + h_{A_i}^*(\Phi_n(i)) - \langle \Phi_n(i), \tilde{\Psi}(i) \rangle) d\lambda(i). \quad (1.40)$$

where $\tilde{\Psi}$ is the profile of actions in equilibrium. Calculating the differences $\phi_{n+1} - \phi_n$ for all $n \in \mathbb{N}$, implies that the quantities $\psi_n = \int_V J(a, \tilde{\eta}) d(\eta_n - \tilde{\eta})(a)$ converge to 0, where $\tilde{\eta}$ denotes the equilibrium distribution.

We complete chapter 3 by showing that we can write the first order MFG as a non atomic anonymous game. Let the set of players be $I = \mathbb{R}^d$ and $m_0 \in \mathcal{P}(I)$ a given non atomic Borel probability measure on I . Let $\mathcal{AC}([0, T], \mathbb{R}^d)$ denotes the set of all absolutely continuous paths γ from $[0, T]$ to \mathbb{R}^d . Set $V = \mathcal{C}^0([0, T], \mathbb{R}^d)$ and for each player $i \in \mathbb{R}^d$ its set of actions be

$$A_i := \{\gamma \in \mathcal{AC}([0, T], \mathbb{R}^d) \mid \gamma(0) = i, \int_0^T \|\dot{\gamma}(t)\|^2 dt \leq M\}. \quad (1.41)$$

for suitable $M > 0$. The MFG cost function $J : V \times \mathcal{P}(V) \rightarrow \mathbb{R}$ is defined as follows:

$$J(\gamma, \eta) = \begin{cases} \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), e_t \# \eta)) dt + g(\gamma(T), e_T \# \eta), & \text{if } \gamma \in \mathcal{AC}([0, T], \mathbb{R}^d) \\ +\infty & \text{otherwise,} \end{cases}$$

where for every $t \in [0, T]$ the evaluation function $e_t : \mathcal{C}^0([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^d$ is defined by $e_t(\gamma) = \gamma(t)$. We call this non atomic anonymous game a *first-order mean field game*.

Theorem 1.4.8. *If $f, g : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$ are monotone in sense of (1.17), then the MFG cost function J will be so in sense of (1.36).*

Theorem 1.4.9. *The first-order MFGs defined as above has at least a Nash Equilibrium $\tilde{\Psi} \in \mathcal{A}$ under suitable conditions (appeared in assumptions 3.5.1). This equilibrium is unique under monotonicity assumption (1.36).*

If we set $\tilde{\eta} = \tilde{\Psi} \# m_0$ for equilibrium profile $\tilde{\Psi}$, and $e_t \# \tilde{\eta} = \tilde{m}_t$ for all $t \in [0, T]$, then for m_0 -almost every $i \in \mathbb{R}^d$:

$$\tilde{\Psi}(i) = \operatorname{argmin}_{\gamma \in \mathcal{AC}([0, T], \mathbb{R}^d), \gamma(0)=i} \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), \tilde{m}_t)) \, dt + g(\gamma(T), \tilde{m}_T).$$

This is exactly as in (1.34); hence as in section 2.4 proved, we can construct the first order MFG solution (u, m) from an equilibrium distribution $\tilde{\eta}$ as in (1.33).

Completely compatible with (1.32), the fictitious play in this framework for MFG reads as follows:

$$\begin{aligned} (i) \quad \Psi_{n+1}(i) &= \operatorname{argmin}_{\gamma \in \mathcal{AC}([0, T], \mathbb{R}^d), \gamma(0)=i} \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), e_t \# \tilde{\eta}_n)) \, dt + g(\gamma(T), e_T \# \tilde{\eta}_n), \\ (ii) \quad \eta_{n+1} &= \Psi_{n+1} \# \lambda, \\ (iii) \quad \tilde{\eta}_{n+1} &= \frac{1}{n+1} \sum_{i=1}^{n+1} \eta_i. \end{aligned} \tag{1.42}$$

Theorem 1.4.10. *Let $\{\eta_n\}_{n \in \mathbb{N}}$ be constructed as in (1.42). Then under suitable conditions (appeared in assumptions 3.5.1) with monotonicity of cost function (1.36), the sequence η_n converges to the unique Nash equilibrium distribution $\tilde{\eta}$.*

The online mirror descent in the case of first order MFG is more complicated due to the calculation of derivative of J :

$$\begin{aligned} \langle D_\gamma J(\gamma, \eta), z \rangle &= \lim_{\epsilon \rightarrow 0} \frac{J(\gamma + \epsilon z, \eta) - J(\gamma, \eta)}{\epsilon} \\ &= \int_0^T (L_x(\gamma_t, \dot{\gamma}(t)) \cdot z_t + L_v(\gamma_t, \dot{\gamma}(t)) \cdot \dot{z}_t + f_x(\gamma(t), e_t \# \eta) \cdot z_t) \, dt + g_x(\gamma(T), e_T \# \eta) \cdot z_T \end{aligned}$$

Using this formulation for the gradient $\nabla_\gamma J(\cdot, \eta)$, one can conclude the convergence of OMD to the equilibrium under good assumptions for L, f, g .

Theorem 1.4.11. *If the cost function J is monotone and convex with respect to the first argument, then under suitable conditions (appeared in assumptions 3.5.1), the online mirror descent algorithm proposed in (1.39) for $\beta_n = \frac{1}{n}$ ($n \in \mathbb{N}$), converges to the unique first-order mean field game equilibrium.*

1.4.3 Finite MFG: fictitious play and convergence to classical MFG

In chapter 4 we address the finite MFG model introduced by Gomes et al. [55]. This is a discrete version of mean field games where the time interval and set of states are both finite sets. The model is as follows. Let \mathcal{S} and $\mathcal{T} = \{t_0, t_1, \dots, t_m\}$ (with $0 = t_0 < t_1 < \dots < t_m = T$), be finite sets representing the states set and time set. We call a tuple (U, M) with $U : \mathcal{T} \times \mathcal{S} \rightarrow \mathbb{R}$, $M : \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$, an equilibrium solution to the finite MFG, if there exists $\hat{P} : \mathcal{S} \times \mathcal{S} \times \mathcal{T} \setminus \{T\} \rightarrow [0, 1]$ such that

$$\begin{aligned} (i) \quad U(x, t_k) &= \inf_{p \in \Delta(\mathcal{S})} \sum_{y \in \mathcal{S}} p_y (c_{xy}(p, M(t_k)) + U(y, t_{k+1})), \quad x \in \mathcal{S}, 0 \leq k < m, \\ (ii) \quad \hat{P}(x, \cdot, t_k) &\in \operatorname{argmin}_{p \in \Delta(\mathcal{S})} \sum_{y \in \mathcal{S}} p_y (c_{xy}(p, M(t_k)) + U(y, t_{k+1})), \quad x \in \mathcal{S}, 0 \leq k < m, \\ (iii) \quad M(x, t_{k+1}) &= \sum_{y \in \mathcal{S}} M(y, t_k) \hat{P}(y, x, t_k), \quad x \in \mathcal{S}, 0 \leq k < m, \\ (iv) \quad U(x, T) &= g(x, M(T)), \quad M(0) = M_0 \end{aligned} \tag{1.43}$$

The relation (i)–(1.43) is the dynamic programming computing the value function from inter-temporal cost c_{xy} and the value function at its next time. For all $x, y \in \mathcal{S}$ the inter-temporal cost $c_{xy}(p, M)$ capturing the cost of moving from x to y taking into account the dispersion $p \in \Delta(\mathcal{S})$ and distribution

over states $M \in \mathcal{P}(\mathcal{S})$. The relation (ii, iii)–(1.43) describing the equilibrium configuration, where the optimal solution to (i)–(1.43) constructs the evolving measure M .

The existence of an equilibrium is proved in [55] under some assumptions on data. In our approach, we give a non atomic anonymous game representation to this game and uses our previous result in chapter 3. Set $I = [0, 1]$ as the set of players and λ the Lebesgue measure over I . For every player $i \in I$ define its action set A_i and V as

$$\mathcal{K}_{\mathcal{S}, \mathcal{T}} = \left\{ P : \mathcal{S} \times \mathcal{S} \times \mathcal{T} \setminus \{T\} \rightarrow [0, 1] \mid \sum_{y \in \mathcal{S}} P(x, y, t) = 1, \text{ for all } x \in \mathcal{S}, t \in \mathcal{T} \setminus \{T\} \right\}. \quad (1.44)$$

We recall that for a typical element $P \in V$ that is a function $P : \mathcal{S} \times \mathcal{S} \times \mathcal{T} \setminus \{T\} \rightarrow [0, 1]$, the quantity $P(x, y, t)$ captures the probability of passing from state x to state y at time t . The set of profile of actions is $\mathcal{A} = \{\Psi : [0, 1] \rightarrow V \mid \Psi \text{ measurable}\}$. For a typical $\Psi \in \mathcal{A}$ the aggregated distribution at time t produced by the profile Ψ is defined by

$$M_\Psi(t) := \int_I M_{\Psi(i)}^{M_0}(t) d\lambda(i) = \int_V M_P^{M_0}(t) d\eta(P), \quad \text{for } t \in \mathcal{T} \setminus \{T\}, \quad (1.45)$$

where for every $P \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$ the measure $M_P^{M_0}(t)$ is the induced measure at time t from initial measure M_0 and P a Markovian transition. We will abuse the notation and use $M_\eta(t)$ (with $\eta = \Psi\#\lambda$ for $\Psi \in \mathcal{A}$) instead of $M_\Psi(t)$, to insist on the dependency of aggregated distribution $M_\Psi(t)$ through the induced measure η . We suppose the following form of the cost function $c_i(\Psi) = J(\Psi(i), \Psi\#\lambda)$, where

$$J(P, \eta) = \sum_{k=0}^{m-1} \sum_{x, y \in \mathcal{S}} M_P^{M_0}(x, t_k) P(x, y, t_k) c_{xy}(P(x, t_k), M_\eta(t_k)) + \sum_{x \in \mathcal{S}} M_P^{M_0}(x, T) g(x, M_\eta(T)). \quad (1.46)$$

Theorem 1.4.12. *The finite MFG possesses at least one equilibrium under suitable assumptions (see assumptions 4.2.1).*

The fictitious play scheme in finite MFG reads as follows: let $M_1 = \bar{M}_1 : \mathcal{T} \rightarrow \Delta(\mathcal{S})$ is arbitrary, for every iteration $n = 1, \dots$ let:

1. for $\bar{M}_n : \mathcal{T} \rightarrow \Delta(\mathcal{S})$ known, construct (U_n, P_n) as follows:

$$\begin{aligned} (i) \quad & U_n(x, T) = g(x, \bar{M}_n(T)), \quad x \in \mathcal{S}, \\ (ii) \quad & U_n(x, t_k) = \min_{p \in \Delta(\mathcal{S})} \sum_{y \in \mathcal{S}} p_y c_{xy}(p, \bar{M}_n(t_k)) + p_y U_n(y, t_{k+1}), \quad x \in \mathcal{S}, 0 \leq k < m, \\ (iii) \quad & P_n(x, \cdot, t_k) = \operatorname{argmin}_{p \in \Delta(\mathcal{S})} \sum_{y \in \mathcal{S}} p_y c_{xy}(p, \bar{M}_n(t_k)) + p_y U_n(y, t_{k+1}), \quad x \in \mathcal{S}, 0 \leq k < m, \end{aligned} \quad (1.47)$$

2. construct $M_{n+1} : \mathcal{T} \rightarrow \Delta(\mathcal{S})$ with

$$M_{n+1}(x, 0) = M_0(x), \quad M_{n+1}(x, t_{k+1}) = \sum_{y \in \mathcal{S}} M_{n+1}(y, t_k) P_n(y, x, t_k), \quad x \in \mathcal{S}, 0 \leq k < m,$$

3. define $\bar{M}_{n+1} : \mathcal{T} \rightarrow \Delta(\mathcal{S})$ with

$$\bar{M}_{n+1} = \frac{n}{n+1} \bar{M}_n + \frac{1}{n+1} M_{n+1}.$$

Theorem 1.4.13. *Under suitable conditions (see Theorem 4.3.2) the sequence $\{(\bar{M}^n, M^{n+1})\}_{n \in \mathbb{N}}$ converges to (M^*, M^*) , where M^* is the equilibrium.*

Our second question in chapter 4 refers to the convergence of finite scheme to continuous scheme when the discretization becomes finer. Our main framework is as follows. Let (N_n^s) and (N_n^t) be two sequences of natural numbers such that $\lim_{n \rightarrow \infty} N_n^s = \lim_{n \rightarrow \infty} N_n^t = +\infty$ and let (ϵ_n) be a sequence of positive real numbers such that $\lim_{n \rightarrow \infty} \epsilon_n = 0$. Define $\Delta x_n := 1/N_n^s$ and $\Delta t_n := T/N_n^t$. For a fixed $n \in \mathbb{N}$, consider the discrete state set \mathcal{S}_n and the discrete time set \mathcal{T}_n defined as

$$\begin{aligned}\mathcal{S}_n &:= \{x_q := q\Delta x_n \mid q \in \mathbb{Z}^d, \|q\|_\infty \leq (N_n^s)^2\} \subseteq \mathbb{R}^d, \\ \mathcal{T}_n &:= \{t_k := k\Delta t_n \mid k = 0, \dots, N_n^t\} \subseteq [0, T].\end{aligned}\tag{1.48}$$

For every $x \in \mathcal{S}_n$ let $E_x := \{x' \in \mathbb{R}^d \mid \|x' - x\|_\infty \leq (\Delta x_n)/2\}$ and define $M_{n,0} \in \Delta(\mathcal{S}_n)$ as $(M_{n,0})_x = m_0(E_x)$ for all $x \in \mathcal{S}_n$. We consider a finite MFG with inter-temporal cost function

$$c_{xy}(p, M) := \Delta t_n \left(\frac{1}{q} \left| \frac{y-x}{\Delta t_n} \right|^q + f(x, M) \right) + \epsilon_n \log(p_y).$$

for a $q > 1$ and set $1/q + 1/q' = 1$. The finite MFG system will be

$$\begin{aligned}U_n(x, t_k) &= \min_{p \in \Delta(\mathcal{S}_n)} \left\{ \sum_{y \in \mathcal{S}_n} p_y \left(\frac{\Delta t_n}{q} \left| \frac{y-x}{\Delta t_n} \right|^q + U_n(y, t_{k+1}) \right) + \epsilon_n \mathcal{E}_n(p) \right\} + \Delta t_n f(x, M_n(t_k)) \\ &\quad \forall x \in \mathcal{S}_n, \quad 0 \leq k \leq N_n^t - 1, \\ U_n(x, T) &= g(x, M_n(T)) \quad \forall x \in \mathcal{S}_n, \\ M_n(x, t_{k+1}) &= \sum_{y \in \mathcal{S}_n} \hat{P}_n(y, x, t_k) M_n(y, t_k) \quad \forall x \in \mathcal{S}_n, \quad 0 \leq k \leq N_n^t - 1, \\ M_n(x, 0) &= (M_{n,0})_x \quad \forall x \in \mathcal{S}_n,\end{aligned}\tag{1.49}$$

where for all $x \in \mathcal{S}_n, 0 \leq k \leq N_n^t - 1$

$$\left(\hat{P}_n(x, y, t_k) \right)_{y \in \mathcal{S}_n} = \operatorname{argmin}_{p \in \Delta(\mathcal{S})} \left\{ \sum_{y \in \mathcal{S}_n} p_y \left(\frac{\Delta t_n}{q} \left| \frac{y-x}{\Delta t_n} \right|^q + U_n(y, t_{k+1}) \right) + \epsilon_n \mathcal{E}_n(p) \right\},\tag{1.50}$$

and $\mathcal{E}_n : \Delta(\mathcal{S}_n) \rightarrow \mathbb{R}$ is the (non positive) entropy function defined by $\mathcal{E}_n(p) = \sum_{x \in \mathcal{S}_n} p_x \log(p_x)$ for all $p \in \Delta(\mathcal{S}_n)$. The main question here is the convergence of (U_n, M_n) solving (1.49), to the solution (u, m) of first order MFG system (1.18). Our theorem is the following

Theorem 1.4.14. *Suppose that, as $n \rightarrow \infty$, $N_n^t/N_n^s \rightarrow 0$, $\epsilon_n = o\left(\frac{1}{N_n^t \log(N_n^s)}\right)$ and consider the corresponding sequence (U_n, M_n) of solutions to the finite MFGs (1.49). Then, there exists a solution (u, m) to (1.18) such that, up to some subsequence, $U_n \rightarrow u$ uniformly on compact subsets of $\mathbb{R}^d \times [0, T]$ and $M_n \rightarrow m$ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$.*

The proof includes several steps; here we give a sketch of it. The set of evolving measures $\{M_n\}_{n \in \mathbb{N}}$, that are extended to entire $[0, T]$ in affine manner, are compact in a suitable function space. For every converging subsequence of $\{M_n\}_{n \in \mathbb{N}}$ (that is still denoted by M_n 's), define

$$U^*(x, t) := \limsup_{\substack{n \rightarrow \infty \\ \mathcal{S}_n \ni y \rightarrow x \\ \mathcal{T}_n \ni s \rightarrow t}} U_n(y, s), \quad U_*(x, t) := \liminf_{\substack{n \rightarrow \infty \\ \mathcal{S}_n \ni y \rightarrow x \\ \mathcal{T}_n \ni s \rightarrow t}} U_n(y, s) \quad \forall x \in \mathbb{R}^d, \quad t \in [0, T].\tag{1.51}$$

We can prove $U^*(x, T) = U_*(x, T) = g(x, m(T))$ for all $x \in \mathbb{R}^d$ where m is the limit function of M_n 's. Using the assumptions for N_n^t, N_n^s, ϵ_n and a suitable comparison principle, we can deduce that $U^* = U_* = u$, where u is the unique continuous viscosity solution to

$$\begin{aligned}-\partial_t u + \frac{1}{q'} |\nabla u(x, t)|^{q'} - f(x, m(t)) &= 0 & x \in \mathbb{R}^d, \quad t \in (0, T), \\ u(x, T) &= g(x, m(T)) & x \in \mathbb{R}^d.\end{aligned}\tag{1.52}$$

or equivalently, u will be the value function of

$$u(x, t) = \inf_{\alpha \in L^2([0, T]; \mathbb{R}^d)} \int_t^T \left[\frac{1}{q} |\alpha(s)|^q + f(X^{x, t}[\alpha](s), m(s)) \right] ds + g(X^{x, t}[\alpha](T), m(T)), \quad (1.53)$$

where $X^{x, t}[\alpha](s) := x + \int_t^s \alpha(s') ds'$ for all $s \in [t, T]$. The rest of the proof relies on the compactness of measures \mathbb{P}_n , that are defined over set of trajectories by the aide of transitions \hat{P}_n . We prove that every accumulation point of measures $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ is an equilibrium distribution $\tilde{\eta}$ in sense of (1.34). The rest of the argument is again by following section 2.4; that we can construct the first order MFG system solution (u, m) from an equilibrium distribution $\tilde{\eta}$ as in (1.33).

Chapter 2

Learning in potential MFG with fictitious play

*Joint work with Pierre Cardaliaguet,
published in ESAIM: Control, Optimisation and Calculus of Variations*

2.1 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, Lions [72, 73, 74] and Huang, Malhamé and Caines [67], (actually a discrete in time version of these games were previously known under the terminology of heterogenous models in economics. See for instance [6]). The classical notion of equilibrium solution in Mean Field Game (abbreviated MFG) is given by a pair of maps (u, m) , where $u = u(x, t)$ is the value function of a typical small player while $m = m(x, t)$ 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

$$\begin{cases} (i) & -\partial_t u - \sigma \Delta u + H(x, \nabla u(x, t)) = f(x, m(t)) \\ (ii) & \partial_t m - \sigma \Delta m - \operatorname{div}(m D_p H(x, \nabla u)) = 0 \\ & m(0, x) = m_0(x), \quad u(T, x) = g(x, m(T)). \end{cases} \quad (2.1)$$

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) 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 $-\operatorname{div}(m D_p H(x, \nabla u))$). The MFG system formalizes therefore *an equilibrium configuration*.

Under suitable assumptions recalled below, the MFG system (2.1) 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 [37]; 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 [67, 42]. MFG systems have been widely used in several areas ranging from engineering to economics, either under the terminology of heterogeneous agent model [6], or under the name of MFG [2, 59, 66].

In the present paper we raise the question of the actual formation of an 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 whether 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 [53]). 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: fictitious play. The fictitious play was first introduced by Brown [26]. In this learning procedure, every player plays at each step a 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 [91], but it is known to converge for several classes of one shot games: for instance for zero-sum games (Robinson [87]), for 2×2 games (Miyasawa [78]), for potential games (Monderer and Shapley [79])...

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 $(\bar{m}^n(t))_{t \in [0, T]}$ 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:

$$\begin{cases} (i) & -\partial_t u^{n+1} - \sigma \Delta u^{n+1} + H(x, \nabla u^{n+1}(x, t)) = f(x, \bar{m}^n(t)), \\ (ii) & \partial_t m^{n+1} - \sigma \Delta m^{n+1} - \operatorname{div}(m^{n+1} D_p H(x, \nabla u^{n+1})) = 0, \\ & m^{n+1}(0) = m_0, u^{n+1}(x, T) = g(x, \bar{m}^n(T)) \end{cases} \quad (2.2)$$

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 (2.2)-(i). The actual density then evolves according to the Fokker-Planck equation (2.2)-(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 (2.1) (by compact, we mean compact for the uniform convergence). Of course, if in addition the solution of the MFG system (2.1) is unique, then the full sequence converges. Let us recall (see [74]) that this uniqueness holds for instance if f and g are monotone:

$$\int (f(x, m) - f(x, m')) \, d(m - m')(x) \geq 0, \quad \int (g(x, m) - g(x, m')) \, d(m - m')(x) \geq 0$$

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

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

The above derivative—in the space of measure—is introduced in subsection 2.1.2, the definition being borrowed from [37]. Our assumption actually ensures that our MFG system is also “a potential game” (in the flavor of Monderer and Shapley [80]) so that the MFG system falls into a framework closely related to that of Monderer and Shapley [79]. Compared to [79], 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 [79]. 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 [74] and exploited in [35, 38] 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 actions. 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, this chapter is the first work that considers 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 [58] an algorithm which is closely related to a replicator dynamics: namely it is exactly (2.2) in which one replaces \bar{m}^n by m^n in (2.2)-(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.

This chapter 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 2.2 is devoted to the fictitious play for second order MFG systems while section 2.3 deals with the first order ones.

Acknowledgement. I wish to thank the support of ANR (Agence Nationale de la Recherche) MFG (ANR-16-CE40-0015-01).

2.1.1 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

$$\mathbf{d}_1(\mu, \nu) = \sup_h \left\{ \int_{\mathbb{T}^d} h(x) \, \mathrm{d}(\mu - \nu)(x) \right\} \quad \mu, \nu \in \mathcal{P}(\mathbb{T}^d), \quad (2.3)$$

where the supremum is taken over all the maps $h : \mathbb{T}^d \rightarrow \mathbb{R}$ 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: $H : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ while $f, g : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$. 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

$$m_0 \text{ has a smooth density (again denoted } m_0). \quad (2.4)$$

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

$$H \in \mathcal{C}^2(\mathbb{T}^d \times \mathbb{R}^d) \text{ and } \frac{1}{\bar{C}} I_d \leq D_{pp}^2 H(x, p) \leq \bar{C} I_d \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d. \quad (2.5)$$

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

$$\langle D_x H(x, p), p \rangle \geq -\bar{C} (|p|^2 + 1). \quad (2.6)$$

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

$$\begin{aligned} &\text{The map } m \rightarrow f(\cdot, m) \text{ is Lipschitz continuous from } \mathcal{P}(\mathbb{T}^d) \text{ to } \mathcal{C}^2(\mathbb{T}^d) \\ &\text{while the map } m \rightarrow g(\cdot, m) \text{ is Lipschitz continuous from } \mathcal{P}(\mathbb{T}^d) \text{ to } \mathcal{C}^3(\mathbb{T}^d). \end{aligned} \quad (2.7)$$

In particular,

$$\sup_{m \in \mathcal{P}(\mathbb{T}^d)} \|f(\cdot, m)\|_{\mathcal{C}^2} + \|g(\cdot, m)\|_{\mathcal{C}^3} < \infty. \quad (2.8)$$

Assumptions (2.4), (2.5), (2.6), (2.7), (2.9) 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 $m, m' \in \mathcal{P}(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d} (f(x, m) - f(x, m')) d(m - m')(x) \geq 0, \quad \int_{\mathbb{T}^d} (g(x, m) - g(x, m')) d(m - m')(x) \geq 0. \quad (2.9)$$

2.1.2 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 $F : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ be a continuous map. We say that the continuous map $\frac{\delta F}{\delta m} : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}$ is a derivative of F if, for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$,

$$\lim_{s \rightarrow 0} \frac{F((1-s)m + sm') - F(m)}{s} = \int_{\mathbb{T}^d} \frac{\delta F}{\delta m}(m, x) d(m' - m)(x). \quad (2.10)$$

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 $m, m' \in \mathcal{P}(\mathbb{T}^d)$. 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 $\mathcal{C}(\mathbb{T}^d, \mathbb{R})$.

Note that $\frac{\delta F}{\delta m}$ is defined only up to an additive constant. To fix the ideas we assume the derivative $\widetilde{\frac{\delta F}{\delta m}}$ as the one that satisfies

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

Definition 2.1.1. A Mean Field Game is called a Potential Mean Field Game if the instantaneous and final cost functions $f, g : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ derive from potentials, i.e., there exists continuously differentiable maps $F, G : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ such that

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

In the following 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 $\frac{\delta f}{\delta m} : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}$ is continuous. Then, for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$,

$$f(x, (1-s)m + sm') = f(x, m) + s \int_{\mathbb{T}^d} \frac{\delta f}{\delta m}(x, m, y) d(m' - m)(y) + o(s),$$

where $\lim_{s \rightarrow 0} \frac{o(s)}{s} = 0$.

Proposition 2.1.1. The map $f : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ derives from a potential, if and only if,

$$\frac{\delta f}{\delta m}(x, m, y) + \phi(x, m) = \frac{\delta f}{\delta m}(y, m, x) + \phi(y, m) \quad \forall x, y \in \mathbb{T}^d, \forall m \in \mathcal{P}(\mathbb{T}^d), \quad (2.11)$$

for some $\phi : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$.

Proof. First assume that f derives from a potential $F : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$. Taking derivate the relation $\frac{\delta F}{\delta m} = f$ respect to m gives

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

As shown in [37] (Section 2.2) there are some $\tilde{\phi} : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ such that

$$\widetilde{\frac{\delta^2 F}{\delta m^2}}(m, x, y) + \tilde{\phi}(x, m) = \widetilde{\frac{\delta^2 F}{\delta m^2}}(m, y, x) + \tilde{\phi}(y, m) \quad \forall x, y \in \mathbb{T}^d, \forall m \in \mathcal{P}(\mathbb{T}^d),$$

so the same is true for $\frac{\delta f}{\delta m}(x, m, y)$. It yields the relation (2.11) since for any derivative $\frac{\delta f}{\delta m} = \frac{\delta f}{\delta m}(x, m, y)$ there is some $\hat{\phi} : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ such that

$$\widetilde{\frac{\delta f}{\delta m}}(x, m, y) - \frac{\delta f}{\delta m}(x, m, y) = \hat{\phi}(x, m) \quad \forall x, y \in \mathbb{T}^d, \forall m \in \mathcal{P}(\mathbb{T}^d).$$

Let us now assume the relation (2.11) for some $\phi : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$. Let us fix $m_0 \in \mathcal{P}(\mathbb{T}^d)$ and set, for any $m \in \mathcal{P}(\mathbb{T}^d)$,

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

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

$$\frac{\delta F}{\delta m}(m, y) = \int_0^1 t \int_{\mathbb{T}^d} \frac{\delta f}{\delta m}(x, (1-t)m_0 + tm, y) d(m - m_0)(x) dt + \int_0^1 f(y, (1-t)m_0 + tm) dt \quad (2.12)$$

By assumption 2.11 we can write,

$$\begin{aligned}
\frac{d}{dt}f(y, (1-t)m_0 + tm) &= \int_{\mathbb{T}^d} \frac{\delta f}{\delta m}(y, (1-t)m_0 + tm, x) d(m - m_0)(x) \\
&= \int_{\mathbb{T}^d} \left(\frac{\delta f}{\delta m}(y, (1-t)m_0 + tm, x) + \phi(y, (1-t)m_0 + tm) \right) d(m - m_0)(x) \\
&= \int_{\mathbb{T}^d} \left(\frac{\delta f}{\delta m}(x, (1-t)m_0 + tm, y) + \phi(x, (1-t)m_0 + tm) \right) d(m - m_0)(x),
\end{aligned} \tag{2.13}$$

So

$$\begin{aligned}
\frac{\delta F}{\delta m}(m, y) &= \int_0^t \frac{d}{dt}f(y, (1-t)m_0 + tm) dt + \int_0^1 f(y, (1-t)m_0 + tm) dt \\
&\quad - \int_0^1 t \int_{\mathbb{T}^d} \phi(x, (1-t)m_0 + tm) d(m - m_0)(x) dt.
\end{aligned} \tag{2.14}$$

We have therefore after integration by parts in (2.14),

$$\begin{aligned}
\frac{\delta F}{\delta m}(m, y) &= \left[t f(y, (1-t)m_0 + tm) \right]_0^1 - \int_0^1 t \int_{\mathbb{T}^d} \phi(x, (1-t)m_0 + tm) d(m - m_0)(x) dt \\
&= f(y, m) - \int_0^1 t \int_{\mathbb{T}^d} \phi(x, (1-t)m_0 + tm) d(m - m_0)(x) dt.
\end{aligned} \tag{2.15}$$

Since $\int_0^1 t \int_{\mathbb{T}^d} \phi(x, (1-t)m_0 + tm) d(m - m_0)(x) dt$ is independent of y , we can write

$$F(m') - F(m) = \int_0^1 \int_{\mathbb{T}^d} f(y, (1-t)m + tm') d(m' - m)(y) dt,$$

for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$. □

2.2 The fictitious play for second order MFG systems

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

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

Let us recall (see [74]) that, under our assumptions (2.4), (2.5), (2.6), (2.7), there exists at least one classical solution to (2.16) (i.e., for which all the involved derivative exists and are continuous). If furthermore (2.9) holds, then the solution is unique.

2.2.1 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 $u^n, m^n : \mathbb{T}^d \times [0, T] \rightarrow \mathbb{R}$ by:

$$\begin{cases} (i) & -\partial_t u^{n+1} - \Delta u^{n+1} + H(x, \nabla u^{n+1}(x, t)) = f(x, \bar{m}^n(t)), \quad (x, t) \in \mathbb{T}^d \times [0, T] \\ (ii) & \partial_t m^{n+1} - \Delta m^{n+1} - \operatorname{div}(m^{n+1} D_p H(x, \nabla u^{n+1})) = 0, \quad (x, t) \in \mathbb{T}^d \times [0, T] \\ & m^{n+1}(0) = m_0, u^{n+1}(x, T) = g(x, \bar{m}^n(T)), \quad x \in \mathbb{T}^d \end{cases} \tag{2.17}$$

where $\bar{m}^n(x, t) = \frac{1}{n} \sum_{k=1}^n m^k(x, t)$. The interpretation is that, at the beginning of stage $n + 1$, the players have the same belief of the future density of the population $(\bar{m}^n(t))_{t \in [0, T]}$ and compute their corresponding optimal control problem with value function u^{n+1} . Their optimal (closed-loop) control is then $(x, t) \rightarrow -D_p H(x, \nabla u^{n+1}(x, t))$. When all players actually implement this control the population density evolves in time according to (2.17)-(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 $F, G : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ such that

$$f(x, m) = \frac{\delta F}{\delta m}(m, x) \quad \text{and} \quad g(x, m) = \frac{\delta G}{\delta m}(m, x). \quad (2.18)$$

We also assume, besides the smoothness assumption (2.4), that m_0 is smooth and positive.

Theorem 2.2.1. *Under the assumptions (2.4), (2.5), (2.6), (2.7) and (2.18), 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 (2.16). If, in addition, the monotonicity condition (2.9) holds, then the whole sequence $\{(u^n, m^n)\}_{n \in \mathbb{N}}$ converges to the unique solution of (2.16).*

The key remark to prove Theorem 2.2.1 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,

$$\partial_t m - \Delta m + \operatorname{div}(w) = 0 \text{ in } (0, T) \times \mathbb{T}^d \quad m(0) = m_0, \quad (2.19)$$

let

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

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 (x, t) to write the potential in a shorter form:

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

It is explained in [74] section 2.6 that (u, m) is a solution to (2.16) if and only if (m, w) is a minimizer of Φ and $w = -mD_p H(\cdot, \nabla u)$ and also constrained to (2.19). 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 [79] applies.

2.2.2 Proof of the convergence

Before starting the proof of Theorem 2.2.1, let us fix some notations. First we set

$$w^n(x, t) = -m^n(x, t)D_p H(x, \nabla u^n(x, t)) \text{ and } \bar{w}^n(x, t) = \frac{1}{n} \sum_{k=1}^n w^k(x, t). \quad (2.20)$$

Since the Fokker-Planck equation is linear we have :

$$\partial_t \bar{m}^{n+1} - \Delta \bar{m}^{n+1} + \operatorname{div}(\bar{w}^{n+1}) = 0, \quad t \in [0, T], \quad \bar{m}^{n+1}(0) = m_0. \quad (2.21)$$

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 minimizer in the above right-hand side:

$$H^*(x, q) = \langle \hat{p}(x, q), q \rangle - H(x, \hat{p}(x, q)). \quad (2.22)$$

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

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

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

Lemma 2.2.1. *Under assumption (2.5), we have for any $x \in \mathbb{T}^d$, $p, q \in \mathbb{R}^d$:*

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

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

Lemma 2.2.2. *There exists a constant $C > 0$ such that, for any $n \in \mathbb{N}^*$,*

$$\Phi(\bar{m}^{n+1}, \bar{w}^{n+1}) - \Phi(\bar{m}^n, \bar{w}^n) \leq -\frac{1}{C} \frac{a_n}{n} + \frac{C}{n^2}, \quad (2.23)$$

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 $s \in [0, 1]$,

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

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

Proof of Lemma 2.2.2. We have

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

where

$$A = \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} H^*(-\bar{w}^{n+1}/\bar{m}^{n+1}) - \bar{m}^n H^*(-\bar{w}^n/\bar{m}^n) \quad (2.24)$$

$$B = \int_0^T (F(\bar{m}^{n+1}(t)) - F(\bar{m}^n(t))) dt + (G(\bar{m}^{n+1}(T)) - G(\bar{m}^n(T))). \quad (2.25)$$

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

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

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

$$\begin{aligned} B &\leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} (-\partial_t u^{n+1} - \Delta u^{n+1} + H(x, \nabla u^{n+1}))(m^{n+1} - \bar{m}^{n+1}) \\ &\quad + \frac{1}{n} \int_{\mathbb{T}^d} g(x, \bar{m}^n(T))(m^{n+1}(T) - \bar{m}^{n+1}(T)) + \frac{C}{n^2} \\ &\leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} (\partial_t(m^{n+1} - \bar{m}^{n+1}) - \Delta(m^{n+1} - \bar{m}^{n+1}))u^{n+1} \\ &\quad + \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} H(x, \nabla u^{n+1})(m^{n+1} - \bar{m}^{n+1}) + \frac{C}{n^2}, \end{aligned}$$

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 (2.20) 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 2.2.1,

$$\begin{aligned} -\langle \bar{w}^{n+1}, \nabla u^{n+1} \rangle - H(x, \nabla u^{n+1})\bar{m}^{n+1} &\leq \bar{m}^{n+1} H^*(x, -\bar{w}^{n+1}/\bar{m}^{n+1}) \\ &\quad - \frac{1}{2\bar{C}} \bar{m}^{n+1} |\bar{w}^{n+1}/\bar{m}^{n+1} - w^{n+1}/m^{n+1}|^2 \end{aligned}$$

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

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

Therefore

$$\begin{aligned} B &\leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} H^*(x, -\bar{w}^{n+1}/\bar{m}^{n+1}) - m^{n+1} H^*(x, -w^{n+1}/m^{n+1}) \\ &\quad - \frac{1}{2\bar{C}n} \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} |\bar{w}^{n+1}/\bar{m}^{n+1} - w^{n+1}/m^{n+1}|^2 + \frac{C}{n^2}. \end{aligned} \tag{2.26}$$

On the other hand, recalling the definition of \hat{p} in (2.22) and setting $\bar{p}^{n+1} = \hat{p}(\cdot, -\bar{w}^{n+1}/\bar{m}^{n+1})$, we can estimate A as follows:

$$\begin{aligned} A &\leq \int_0^T \int_{\mathbb{T}^d} -\langle \bar{p}^{n+1}, \bar{w}^{n+1} \rangle - \bar{m}^{n+1} H(x, \bar{p}^{n+1}) + \langle \bar{p}^{n+1}, \bar{w}^n \rangle + \bar{m}^n H(x, \bar{p}^{n+1}) \\ &= \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} \langle \bar{p}^{n+1}, \bar{w}^{n+1} \rangle + \bar{m}^{n+1} H(x, \bar{p}^{n+1}) - \langle \bar{p}^{n+1}, w^{n+1} \rangle - m^{n+1} H(x, \bar{p}^{n+1}) \\ &\leq \frac{1}{n} \int_0^T \int_{\mathbb{T}^d} m^{n+1} H^*(x, -w^{n+1}/m^{n+1}) - \bar{m}^{n+1} H^*(x, -\bar{w}^{n+1}/\bar{m}^{n+1}). \end{aligned} \tag{2.27}$$

Putting together (2.26) and (2.27) 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 $a_n = \int_0^T \int_{\mathbb{T}^d} \bar{m}^{n+1} |\bar{w}^{n+1}/\bar{m}^{n+1} - w^{n+1}/m^{n+1}|^2$. □

In order to proceed, let us recall some basic estimates on the system (2.17), the proof of which is postponed:

Lemma 2.2.3. *For any $\alpha \in (0, 1/2)$ there exist a constant $C > 0$ such that for any $n \in \mathbb{N}^*$*

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

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 2.2.4. *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 $\bar{m}^n - \bar{m}^{n-1} = (m^n - \bar{m}^{n-1})/n$, where the m^n (and thus the \bar{m}^n) are uniformly bounded thanks to Lemma 2.2.3, we have by Lipschitz continuity of f and g that

$$\sup_{t \in [0, T]} \|f(\cdot, \bar{m}^{n+1}(t)) - f(\cdot, \bar{m}^n(t))\|_\infty + \|g(\cdot, \bar{m}^{n+1}(T)) - g(\cdot, \bar{m}^n(T))\|_\infty \leq \frac{C}{n}. \quad (2.28)$$

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

$$\|u^{n+1} - u^n\|_\infty \leq \frac{C}{n}. \quad (2.29)$$

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 $\mathbb{T}^d \times [0, T]$ we find by (2.28) and (2.29):

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

Then we use the uniform bound on the ∇u^n given by Lemma 2.2.3 as well as (2.29) to get

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

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 2.2.3.

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 $R = m^n (D_p H(x, \nabla u^{n+1}) - D_p H(x, \nabla u^n))$. 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 . \square

Combining Lemma 2.2.3 with Lemma 2.2.4 we immediately obtain that the sequence (a_n) defined in Lemma 2.2.2 is slowly varying in time:

Corollary 2.2.1. *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 2.2.1. From Lemma 2.2.2, 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 2.2.1, we also have, for any $n \in \mathbb{N}^*$,

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

Then Lemma 2.2.5 below implies that $\lim_{n \rightarrow \infty} a_n = 0$.

In particular we have, by Lemma 2.2.3:

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

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 2.2.3—uniformly converges to 0 on $[0, T] \times \mathbb{T}^d$.

Recall that, by Lemma 2.2.3, 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 (2.16), 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 2.2.3, we have $D_p H(x, \nabla u^{n_i})$ converges uniformly to $D_p H(x, \nabla u)$, so that by (2.20) and the fact that the sequence $\{\bar{w}^n/\bar{m}^n - w^n/m^n\}_{n \in \mathbb{N}}$ converges to 0,

$$-D_p H(x, \nabla u) = \frac{w}{m} = \frac{\bar{w}}{\bar{m}}. \quad (2.30)$$

We now pass to the limit in (2.17) (in the viscosity sense for the Hamilton-Jacobi equation and in the sense of distribution for the Fokker-Planck equation) to get

$$\begin{aligned} (i) \quad & -\partial_t u - \Delta u + H(x, \nabla u(x, t)) = f(x, \bar{m}(t)), \quad (x, t) \in \mathbb{T}^d \times [0, T] \\ (ii) \quad & \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, \nabla u)) = 0, \quad (x, t) \in \mathbb{T}^d \times [0, T] \\ & m(0) = m_0, \quad u(x, T) = g(x, \bar{m}(T)), \quad x \in \mathbb{T}^d. \end{aligned} \quad (2.31)$$

Letting $n \rightarrow +\infty$ in (2.21) we also have

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

By (2.30), 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 (2.9) 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) . \square

In the proof of Theorem 2.2.1, we have used the following Lemma, which can be found in [79].

Lemma 2.2.5. *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*

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

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

Proof. We reproduce the proof of [79] 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 $\lim_{k \rightarrow \infty} b_k = 0$. So we have:

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

which yields the first result since:

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

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

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

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 $N_\epsilon \in \mathbb{N}$ large enough that for $N \geq N_\epsilon$ we have

$$\frac{1}{N} + \frac{1}{N+1} + \cdots + \frac{1}{[(1+\lambda_\epsilon)N]} < \frac{\epsilon}{2C}. \quad (2.32)$$

Let $N \geq N_\epsilon$. Assume for a while that $a_N > \epsilon$. As $|a_{k+1} - a_k| \leq C/k$, (2.32) implies that $a_k > \frac{\epsilon}{2}$ for $N \leq k \leq [N(1+\lambda_\epsilon)]$. Thus

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

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 $a_N \leq \epsilon$ for N sufficiently large, so that (a_k) converges to 0. \square

Proof of Lemma 2.2.1. For simplicity of notation, we omit the x dependence in the various quantities. As by assumption (2.5) 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

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

\square

Proof of Lemma 2.2.3. 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 [70, Theorem 2.2], to get a uniform Holder estimates for Du^{n+1} , with [70, Theorem 12.1], to obtain the full bound by considering $-H(x, \nabla u^{n+1}(x, t)) + f(x, \bar{m}^n(t))$ 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 m^n . The bound from below for the m^n comes from the strong maximum principle. \square

2.3 The fictitious play for first order MFG systems

We now consider the first order order MFG system:

$$\begin{cases} (i) & -\partial_t u + H(x, \nabla u(x, t)) = f(x, m(t)), \quad (x, t) \in \mathbb{T}^d \times [0, T] \\ (ii) & \partial_t m + \operatorname{div}(-m D_p H(x, \nabla u(x, t))) = 0, \quad (x, t) \in \mathbb{T}^d \times [0, T] \\ & m(0) = m_0, \quad u(x, T) = g(x, m(T)), \quad x \in \mathbb{T}^d \end{cases} \quad (2.33)$$

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 2.2. For this reason we introduce another method—based on another potential—, which also has the interest that it can be adapted to a finite number of players.

Let us start by recalling the notion of solution for (2.33). Following [74], we say that the pair (u, m) is a solution to the MFG system (2.33) if u is a Lipschitz continuous viscosity solution to (2.33)-(i) while $m \in L^\infty((0, T) \times \mathbb{T}^d)$ is a solution of (2.33)-(ii) in the sense of distribution.

Under our standing assumptions (2.4), (2.5), (2.6), (2.7), there exists at least one solution (u, m) to the mean field game system (2.33). If furthermore (2.9) holds, then the solution is unique (see [74] and Theorem 5.1 in [34]).

2.3.1 The learning rule and the potential

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

$$\begin{aligned} (i) & \quad -\partial_t u^{n+1} + H(x, \nabla u^{n+1}(x, t)) = f(x, \bar{m}^n(t)), \quad (x, t) \in \mathbb{T}^d \times [0, T] \\ (ii) & \quad \partial_t m^{n+1} + \operatorname{div}(-m^{n+1} D_p H(x, \nabla u^{n+1})) = 0, \quad (x, t) \in \mathbb{T}^d \times [0, T] \\ & \quad m^{n+1}(0) = m_0, \quad u^{n+1}(x, T) = g(x, \bar{m}^n(T)), \quad x \in \mathbb{T}^d \end{aligned} \quad (2.34)$$

where $\bar{m}^n(x, t) = \frac{1}{n} \sum_{k=1}^n m^k(x, t)$. If equation (2.34)-(i) is easy to interpret, the meaning of (2.34)-(ii) would be more challenging and, actually, would make little sense for a finite number of players. 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. We denote

$$H^1([0, T], \mathbb{T}^d) = \left\{ \gamma \in \mathcal{AC}([0, T], \mathbb{T}^d) \mid \int_0^T \|\dot{\gamma}(t)\|^2 dt < +\infty \right\}.$$

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 $e_t : \Gamma \rightarrow \mathbb{T}^d$, 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 $m^\eta(t) = e_{t\#}\eta$ 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 $e_{0\#}\eta = m_0$. 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

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

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 $J(\gamma, \eta) := J(0, \gamma, \eta)$. We note for later use that $J(t, \cdot, \eta)$ is lower semi-continuous on Γ .

We now define the fictitious play process. We start with an initial configuration $\bar{\eta}^0 = \eta^0 \in \mathcal{P}_0(\Gamma)$ (the belief before the first step of a typical player on the actions of the other players). We now build by induction the sequences (η^n) and $(\bar{\eta}^n)$ of $\mathcal{P}(\Gamma)$, $\bar{\eta}^n$ being interpreted as the belief at the end of stage n of a typical player on the actions of the other agents and η^{n+1} the repartition of strategies of the players when they play optimally in the game against $\bar{\eta}^n$. 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

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

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 $x \rightarrow \bar{\gamma}_x^{n+1}$ is Borel measurable. We then consider the measure $\eta^{n+1} \in \mathcal{P}_0(\Gamma)$ defined by

$$\eta^{n+1} := \bar{\gamma}_\cdot^{n+1\#} m_0 \quad \forall t \in [0, T]$$

and set

$$\bar{\eta}^{n+1} := \frac{1}{n+1} \sum_{k=1}^{n+1} \eta^k = \bar{\eta}^n + \frac{1}{n+1} (\eta^{n+1} - \bar{\eta}^n). \quad (2.35)$$

As in section 2.2, we assume that our MFG is potential, i.e., that there exists of potential functions $F, G : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ such that:

$$f(x, m) = \frac{\delta F}{\delta m}(x, m), \quad g(x, m) = \frac{\delta G}{\delta m}(x, m). \quad (2.36)$$

Here is our main convergence result.

Theorem 2.3.1. *Assume that (2.4), (2.5), (2.6), (2.7) and (2.36) hold. Then the sequences $(\bar{\eta}^n, \eta^n)$ is pre-compact in $\mathcal{P}(\Gamma) \times \mathcal{P}(\Gamma)$ and any cluster point $(\bar{\eta}, \eta)$ satisfies the following: $\eta = \bar{\eta}$ and, if we set*

$$\bar{m}(t) := e_{t\#}\bar{\eta}, \quad \bar{u}(x, t) = \inf_{\gamma \in H^1, \gamma(t)=x} J(t, \gamma, \bar{\eta}), \quad (2.37)$$

then the pair (\bar{u}, \bar{m}) is a solution to the MFG system (2.33). If furthermore (2.9) holds, then the entire sequence $(\bar{\eta}^n, \eta^n)$ converges.

The proof of Theorem 2.3.1 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 $\eta \in \mathcal{P}(\Gamma)$,

$$\Phi(\eta) := \int_{\Gamma} \int_0^T L(\gamma(t), \dot{\gamma}(t)) \, dt \, d\eta(\gamma) + \int_0^T F(e_t \# \eta) \, dt + G(e_T \# \eta). \quad (2.38)$$

Note that $\Phi(\eta)$ is well-defined and belongs to $(-\infty, +\infty]$. The potential defined above is reminiscent of [36] or [35]. For instance, in [35]—but for MFG system with a local dependence and under the monotonicity condition (2.9)—it is proved that the MFG equilibrium can be found as a global minimum of Φ . We will show in the proof of Theorem 2.3.1 that the limit measure $\bar{\eta}$ is characterized by the optimality condition

$$\frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\bar{\eta}) \leq \frac{\delta \Phi}{\delta \eta}(\bar{\eta})(\theta) \quad \forall \theta \in \mathcal{P}_0(\Gamma).$$

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

$$\|\dot{\bar{\gamma}}_x^{n+1}\|_{\infty} \leq C. \quad (2.39)$$

In particular, the sequences $(\bar{\eta}^n)$ and (η^n) are tight and

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

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 [34]). 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 (Theorem 6.4.8 of [32]): this is (2.39). The rest of the proof is a straightforward consequence of (2.39). \square

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 (2.10), 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 2.3.2. *For any $\eta, \eta' \in \mathcal{P}(\Gamma)$ such that $\Phi(\eta), \Phi(\eta') < +\infty$, we have*

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

Proof. This is a straightforward application of the definition of Φ in (2.38) and of the continuous derivability of F and G . \square

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

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

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 $\bar{\eta}^n$.

Lemma 2.3.3. *For any $n \in \mathbb{N}^*$,*

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

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

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

As $m_0 = e_0 \# \theta$, we can disintegrate θ into $\theta = \int_{\mathbb{T}^d} \theta_x dm_0(x)$, 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. $x \in \mathbb{T}^d$,

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

and therefore, integrating with respect to m_0 :

$$\int_{\mathbb{T}^d} J(\bar{\gamma}_x^{n+1}, \bar{\eta}^n) m_0(x) dx \leq \int_{\mathbb{T}^d} \int_{\Gamma} J(\gamma, \bar{\eta}^n) d\theta_x(\gamma) m_0(x) dx = \int_{\Gamma} J(\gamma, \bar{\eta}^n) d\theta(\gamma).$$

□

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

Proposition 2.3.1. *There is a constant $C > 0$ such that, for any $n \in \mathbb{N}^*$, we have*

$$\Phi(\bar{\eta}^{n+1}) \leq \Phi(\bar{\eta}^n) + \frac{1}{n+1} \frac{\delta \Phi}{\delta \eta}(\bar{\eta}^n)(\eta^{n+1} - \bar{\eta}^n) + \frac{C}{(n+1)^2} \quad (2.40)$$

where

$$\frac{\delta \Phi}{\delta \eta}(\bar{\eta}^n)(\eta^{n+1} - \bar{\eta}^n) = \int_{\Gamma} J(\gamma, \bar{\eta}^n) d(\eta^{n+1} - \bar{\eta}^n)(\gamma) \leq 0. \quad (2.41)$$

Proof. Recalling (2.35), we have

$$\begin{aligned} \Phi(\bar{\eta}^{n+1}) - \Phi(\bar{\eta}^n) &= \int_0^1 \frac{\delta \Phi}{\delta \eta}((1-s)\bar{\eta}^n + s\bar{\eta}^{n+1})(\bar{\eta}^{n+1} - \bar{\eta}^n) ds \\ &= \frac{1}{(n+1)} \int_0^1 \frac{\delta \Phi}{\delta \eta}((1-s)\bar{\eta}^n + s\bar{\eta}^{n+1})(\eta^{n+1} - \bar{\eta}^n) ds. \end{aligned} \quad (2.42)$$

Let us estimate the right-hand side of the inequality. For any $s \in [0, 1]$, Lemma 2.3.2 states that

$$\begin{aligned} \frac{\delta \Phi}{\delta \eta}((1-s)\bar{\eta}^n + s\bar{\eta}^{n+1})(\eta^{n+1} - \bar{\eta}^n) &= \int_{\Gamma} J(\gamma, (1-s)\bar{\eta}^n + s\bar{\eta}^{n+1}) d(\eta^{n+1} - \bar{\eta}^n)(\gamma) \\ &= \int_{\Gamma} J(\gamma, \bar{\eta}^n) d(\eta^{n+1} - \bar{\eta}^n)(\gamma) + R(s) \end{aligned} \quad (2.43)$$

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

$$\begin{aligned} R(s) &= \int_{\Gamma} \int_0^T (f(\gamma(t), e_t \# ((1-s)\bar{\eta}^n + s\bar{\eta}^{n+1})) - f(\gamma(t), e_t \# \bar{\eta}^n)) dt d(\eta^{n+1} - \bar{\eta}^n)(\gamma) \\ &\quad + \int_{\Gamma} (g(\gamma(T), e_T \# ((1-s)\bar{\eta}^n + s\bar{\eta}^{n+1})) - g(\gamma(T), e_T \# \bar{\eta}^n)) d(\eta^{n+1} - \bar{\eta}^n)(\gamma) \\ &\leq C \sup_{t \in [0, T]} \mathbf{d}_1(e_t \# ((1-s)\bar{\eta}^{n+1} + s\bar{\eta}^n), e_t \# \bar{\eta}^n). \end{aligned} \quad (2.44)$$

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

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

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

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

where the first term in the right-hand side is non-positive thanks to Lemma 2.3.3. \square

2.3.2 Convergence of the fictitious play

In this subsection, we prove Theorem 2.3.1. Recall that Lemma 2.3.1 states that the sequence $(\bar{\eta}^n)$ is tight. We next Lemma characterizes the cluster distribution :

Lemma 2.3.4. *Any cluster point $\bar{\eta}$ of the sequence $(\bar{\eta}^n)$ satisfies*

$$\frac{\delta\Phi}{\delta\eta}(\bar{\eta})(\bar{\eta}) \leq \frac{\delta\Phi}{\delta\eta}(\bar{\eta})(\theta) \quad \forall \theta \in P_0(\Gamma), \quad (2.45)$$

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

Proof. Let us define:

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

where the last equality come from Lemma 2.3.3. Then according to Proposition 2.3.1 the sequence (a^n) is non-negative and, by (2.40), the quantity $\sum_k a^k/k$ is finite (because Φ is bounded below). Therefore by Lemma 2.2.5 we have:

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{k=1}^N a^k = 0. \quad (2.46)$$

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 2.3.1, we have, for any $\theta \in P_0(\Gamma)$,

$$\left| \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^n)(\theta) - \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^{n+1})(\theta) \right| \leq \frac{C}{n}. \quad (2.47)$$

On the other hand, by optimality of η^{n+1} and η^{n+2} in Lemma 2.3.3 and (2.47), we have

$$\begin{aligned} \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^n)(\eta^{n+1}) &= \min_{\theta \in \mathcal{P}_0(\mathbb{T}^d)} \int_{\Gamma} J(\gamma, \bar{\eta}^n) d\theta(\gamma) \leq \int_{\Gamma} J(\gamma, \bar{\eta}^n) d\eta^{n+2}(\gamma) \\ &\leq \int_{\Gamma} J(\gamma, \bar{\eta}^{n+1}) d\eta^{n+2}(\gamma) + C/n = \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^{n+1})(\eta^{n+2}) + C/n \\ &= \min_{\theta \in \mathcal{P}_0(\mathbb{T}^d)} \int_{\Gamma} J(\gamma, \bar{\eta}^{n+1}) d\theta(\gamma) + C/n \\ &\leq \int_{\Gamma} J(\gamma, \bar{\eta}^{n+1}) d\eta^{n+1}(\gamma) + C/n = \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^n)(\eta^{n+1}) + C/n, \end{aligned}$$

which proves that

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

So we have:

$$\begin{aligned}
|a^{n+1} - a^n| &= \left| \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^n)(\bar{\eta}^n - \eta^{n+1}) - \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^{n+1})(\bar{\eta}^{n+1} - \eta^{n+2}) \right| \\
&\leq \left| \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^n)(\bar{\eta}^n) - \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^{n+1})(\bar{\eta}^{n+1}) \right| + \left| \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^n)(\eta^{n+1}) - \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^{n+1})(\eta^{n+2}) \right| \\
&\leq \left| \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^n)(\bar{\eta}^n - \bar{\eta}^{n+1}) \right| + C/n = \frac{1}{n+1} \left| \frac{\delta\Phi}{\delta\eta}(\bar{\eta}^n)(\eta^{n+1} - \bar{\eta}^n) \right| + C/n \leq C/n.
\end{aligned}$$

By (2.46) and the above estimate, we conclude that $a_n \rightarrow 0$ thanks to Lemma 2.2.5.

Let now $\bar{\eta}$ be any cluster point of the sequence $(\bar{\eta}^n)$. Let us check that (2.45) holds. Let $\theta \in \mathcal{P}_0(\mathbb{T}^d)$. Then, from Lemma 2.3.3, for every $n \in \mathbb{N}$ we have:

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

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

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

where the last term tends to 0 because the maps $t \rightarrow e_t \# \bar{\eta}^{n_i}$ are uniformly continuous (from Lemma 2.3.1) and converges pointwisely (and thus uniformly) to $t \rightarrow e_t \# \bar{\eta}$. This yields that $(\frac{\delta\Phi}{\delta\eta}(\bar{\eta}^{n_i})(\theta))$ converges to $\frac{\delta\Phi}{\delta\eta}(\bar{\eta})(\theta)$. On the other hand, by lower semicontinuity of the map $\gamma \rightarrow J(\gamma, \bar{\eta})$ on Γ , we have

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

which proves (2.45).

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

$$J(\tilde{\gamma}_x, \bar{\eta}) \leq J(\gamma, \bar{\eta}). \quad (2.48)$$

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

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

As the reverse inequality always holds, this proves that there must be an equality in (2.48) a.e., which proves the claim. \square

Proof of Theorem 2.3.1. Let $(\bar{\eta}, \eta)$ be the limit of a converging subsequence $(\bar{\eta}^{n_i}, \eta^{n_i})$. We set

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

By standard argument in optimal control, we know that \bar{u} is a viscosity solution to (2.34)-(i) with terminal condition $\bar{u}(T, x) = g(x, \bar{m}(T))$. Moreover, \bar{u} is Lipschitz continuous and semiconcave (cf. for instance Lemma 5.2 in [34]).

It remains to check that \bar{m} satisfies (2.34)-(ii). By Lemma 2.3.4, we know that

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

which means that $\bar{\eta}$ -a.e. γ is optimal for the map $\tilde{\gamma} \rightarrow J(\tilde{\gamma}, \bar{\eta})$ under the constraint $\tilde{\gamma}(0) = \gamma(0)$. Following Theorem 6.4.9 in [32], the optimal solution for $J(\cdot, \bar{\eta})$ is unique at any point of differentiability of $\bar{u}(0, \cdot)$ (let us call it $\bar{\gamma}_x$). Disintegrating $\bar{\eta}$ into $\bar{\eta} = \int_{\mathbb{T}^d} \bar{\eta}_x dm_0(x)$, we have therefore, since m_0 is absolutely continuous,

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

so that

$$\bar{\eta} = \int_{\mathbb{T}^d} \delta_{\bar{\gamma}_x} m_0(x) dx \quad \text{and} \quad \bar{m}(t) = \bar{\gamma}(\cdot, t) \# m_0 \quad \forall t \in [0, T]. \quad (2.49)$$

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

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

(see Theorems 6.4.7 and 6.4.8 of [32]). This proves that \bar{m} is a solution in the sense of distribution of (2.34)-(ii) (where we denote by $\nabla \bar{u}$ any fixed Borel measurable selection of the map $(x, t) \rightarrow D^*u(x, t)$, the set of reachable gradients of u at (x, t) , see [32]). Proposition 2.4.1 in appendix states that (2.34)-(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 (2.34).

In order to identify the cluster point η , let us recall that η^n is defined by

$$\eta^n = \bar{\gamma}^n \# m_0,$$

where, for any $x \in \mathbb{T}^d$, $\bar{\gamma}_x^n$ is a minimum of $J(\cdot, \bar{\eta}^n)$ under the constraint $\gamma(0) = x$. As the criterion $J(\cdot, \bar{\eta}^{n_i})$ Γ -converges to $J(\cdot, \bar{\eta})$ 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 (η^{n_i}) converges to $\bar{\gamma} \# m_0$, which is nothing but $\bar{\eta}$ by (2.49). So we conclude that $\eta = \bar{\eta}$.

Finally, if (2.9) 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 $J(\gamma, \bar{\eta})$ is only through the family of measures $(\bar{m}(t) = e_t \# \bar{\eta})$ and since, by (2.9), there exists a unique solution to the MFG system and thus \bar{m} is uniquely defined, $J(\gamma, \bar{\eta})$ 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 (2.49). Therefore the entire sequence $(\bar{\eta}^n, \eta^n)$ converges to $(\bar{\eta}, \bar{\eta})$. \square

Remark 2.3.1. *The proof shows that a measure $\bar{\eta} \in \mathcal{P}_0(\Gamma)$ which satisfies (2.45) can be understood as the representation of a MFG equilibrium. Indeed, if we define (\bar{u}, \bar{m}) as in (2.37), then (\bar{u}, \bar{m}) is a solution to the MFG system (2.34). Conversely, if (\bar{u}, \bar{m}) is a solution to the MFG system (2.34), then the relation (2.49) identifies uniquely a measure $\bar{\eta} \in \mathcal{P}_0(\Gamma)$. For this reason, we call such a measure an equilibrium measure.*

2.3.3 Fictitious play scheme 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 $x_1^N, x_2^N, \dots, x_N^N \in \mathbb{T}^d$ such that:

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

where $m_0^N = \frac{1}{N} \sum_{i=1}^N \delta_{x_i^N}$ is the empirical measure associated with the $\{x_i^N\}_{i=1,\dots,N}$. 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:

$$\begin{aligned}\eta^{n+1,N} &= \frac{1}{n+1}(\theta^{1,N} + \theta^{2,N} + \dots + \theta^{n+1,N}) \\ \theta^{n+1,N} &= \frac{1}{N}(\delta_{\gamma_{x_1^N}^{n+1,N}} + \delta_{\gamma_{x_2^N}^{n+1,N}} + \dots + \delta_{\gamma_{x_N^N}^{n+1,N}})\end{aligned}\tag{2.50}$$

where $\gamma_{x_i^N}^{n+1,N}$ is an optimal path which minimizes $J(\cdot, \eta^{n,N})$ under constraint $\gamma(0) = x_i^N$. As before one can show that if

$$\begin{aligned}a^{n+1,N} &:= -\frac{\delta\Phi}{\delta\eta}(\eta^{n,N})(\theta^{n+1,N} - \eta^{n,N}) = -\int_{\Gamma} J(\gamma, \eta^{n,N}) d(\theta^{n+1,N} - \eta^{n,N})(\gamma) \\ &= \int_{\Gamma} J(\gamma, \eta^{n,N}) d\eta^{n,N}(\gamma) - \min_{\theta \in \mathcal{P}(\Gamma), e_0\#\theta = m_0^N} \int_{\Gamma} J(\gamma, \eta^{n,N}) d\theta(\gamma),\end{aligned}$$

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

$$\int_{\Gamma} J(\gamma, \bar{\eta}^N) d\bar{\eta}^N(\gamma) = \min_{\theta \in \mathcal{P}(\Gamma), e_0\#\theta = m_0^N} \int_{\Gamma} J(\gamma, \bar{\eta}^N) d\theta(\gamma).\tag{2.51}$$

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

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

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:

$$\int_{\Gamma} J(\gamma, \bar{\eta}) d\bar{\eta}(\gamma) = \min_{\theta \in \mathcal{P}_0(\Gamma)} \int_{\Gamma} J(\gamma, \bar{\eta}) d\theta(\gamma),\tag{2.52}$$

where $\mathcal{P}_0(\Gamma)$ is the set of measure $\theta \in \mathcal{P}(\Gamma)$ such that $e_0\#\theta = m_0$. We have seen in Remark 2.3.1 that this condition characterizes an MFG equilibrium.

Theorem 2.3.2. *Assume that (2.4), (2.5), (2.6), (2.7) and (2.36) hold. Consider the fictitious play for the N -player game as described in (2.50) 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 (2.9) 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 $e_0\#\theta = m_0$, we have

$$\int_{\Gamma} J(\gamma, \bar{\eta}) d\bar{\eta}(\gamma) \leq \int_{\Gamma} J(\gamma, \bar{\eta}) d\theta(\gamma).\tag{2.53}$$

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

$$\tau_N\#m_0 = m_0^N, \quad \mathbf{d}_1(m_0, m_0^N) = \int_{\mathbb{T}^d} |x - \tau_N(x)| dm_0(x)$$

(see [8]). We define the functions $\xi_N : \Gamma \rightarrow \Gamma$ as follows:

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

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

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

Then the characterization (2.51) of $\bar{\eta}^N$ yields:

$$\int_{\Gamma} J(\gamma, \bar{\eta}^N) d\bar{\eta}^N(\gamma) \leq \int_{\Gamma} J(\gamma, \bar{\eta}^N) d\theta^N(\gamma). \quad (2.54)$$

By lower semicontinuity of J we have

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

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

$$\begin{aligned} & \int_{\Gamma} J(\gamma, \bar{\eta}^N) d\theta^N(\gamma) \\ &= \int_{\mathbb{T}^d} \int_{\Gamma} \left(\int_0^T L(\gamma(t) - \gamma(0) + \tau_N(\gamma(0)), \dot{\gamma}(t)) + f(\gamma(t) - \gamma(0) + \tau_N(\gamma(0)), e_t \# \bar{\eta}^N) dt \right. \\ & \quad \left. + g(\gamma(t) - \gamma(0) + \tau_N(\gamma(0)), e_T \# \bar{\eta}^N) \right) m_0(x) d\theta_x(\gamma) dx, \end{aligned}$$

where, by dominate convergence, the right-hand side converges to the right-hand side of (2.53). So letting $N \rightarrow \infty$ in (2.54) gives exactly (2.53).

Under (2.9), the MFG equilibrium is unique. Hence, for any $\epsilon > 0$ there exists $N_{\epsilon} \in \mathbb{N}$ such that for any $N > N_{\epsilon}$ and any accumulation point $\bar{\eta}^N$ we have $\mathbf{d}_1(\bar{\eta}, \bar{\eta}^N) < \epsilon$. \square

Corollary 2.3.1. *Assume (2.4), (2.5), (2.6), (2.7) and (2.36) and (2.9). Then, for any $\epsilon > 0$ there is $N_{\epsilon} \in \mathbb{N}$ such that for any $N > N_{\epsilon}$,*

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

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.

2.4 Well-posedness of a continuity equation

We consider the continuity equation

$$\begin{cases} \partial_t m - \operatorname{div}(m D_p H(x, \nabla \bar{u})) = 0 & \text{in } (0, T) \times \mathbb{T}^d \\ m(0, x) = m_0(x). \end{cases} \quad (2.55)$$

where \bar{u} is the viscosity solution to

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

Let us recall that \bar{u} is semi-concave. In (2.55) we denote by $\nabla \bar{u}$ any fixed Borel measurable selection of the map $(x, t) \rightarrow D^*u(x, t)$ (the set of reachable gradients of u at (x, t) , see [32]). The section is devoted to the proof of the following statement.

Proposition 2.4.1. *Given a fixed map \bar{u} , there exists a unique solution \bar{m} of (2.55) 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 [49]. 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 [7], in particular for the “superposition principle” (see Theorem 2.4.1 below).

Let us start with the existence of a bounded solution to (2.55): this is the easy part.

Lemma 2.4.1. *There exists a solution to (2.55) which belongs to L^∞ .*

Proof. We follow (at least partially) the perturbation argument given in the proof of Theorem 5.1 of [34]. For $\varepsilon > 0$, let $(u^\varepsilon, m^\varepsilon)$ be the unique classical solution to

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

Following the same argument as in [34], we know that the (m^ε) are uniformly bounded in L^∞ : there exists $C > 0$ such that

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

Moreover (by semi-concavity) the (∇u^ε) are uniformly bounded and converge a.e. to $\nabla \bar{u}$ as ε tends to 0. Letting $\varepsilon \rightarrow 0$, we can extract a subsequence such that m^ε converges in L^∞ -weak* to a solution m of (2.55). \square

The difficult part of the proof of Proposition 2.4.1 is to check that the solution to (2.55) 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:

$$\bar{u}(x, t) = \inf_{\gamma, \gamma(t)=x} \int_t^T \tilde{L}(s, \gamma(s), \dot{\gamma}(s), \bar{m}(s)) ds + \tilde{g}(\gamma(T)) \quad (2.56)$$

where we have set, for simplicity of notation,

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

For $(x, t) \in [0, T] \times \mathbb{T}^d$ we denote by $\mathcal{B}(x, t)$ the set of optimal trajectories for the control problem (2.56).

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 (2.56) (see [32], Theorems 6.4.7 and 6.4.9 and Corollary 6.4.10). Let $(x, t) \in \mathbb{T}^d \times [0, T]$ and $\gamma \in \Gamma$. Then

1. (Uniqueness of the optimal control along optimal trajectories) Assume that $\gamma \in \mathcal{B}(x, t)$. 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) $\nabla u(x, t)$ exists if and only if $\mathcal{B}(x, t)$ is reduced to singleton. In this case, $\dot{\gamma}(t) = -D_p H(x, \nabla \bar{u}(x, t))$ where $\mathcal{B}(x, t) = \{\gamma\}$.
3. (Optimal synthesis) conversely, if $\gamma(\cdot)$ is an absolutely continuous solution of the differential equation

$$\begin{cases} \dot{\gamma}(s) = -D_p H(s, \gamma(s), \nabla \bar{u}(s, \gamma(s))) & \text{a.e. in } [t, T] \\ \gamma(t) = x, \end{cases} \quad (2.57)$$

then the trajectory γ is optimal for $\bar{u}(x, t)$. In particular, if $\bar{u}(t, \cdot)$ is differentiable at x , then equation (2.57) 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 (2.55) can be represented by a measure on the space of trajectories of the ODE

$$\dot{\gamma}(s) = -D_p H(\gamma(s), \nabla u(s, \gamma(s))). \quad (2.58)$$

Theorem 2.4.1 (Ambrosio superposition principle). *Let μ be a solution to (2.55). Then there exists a Borel probability measure η on $C^0([0, T], \mathbb{T}^d)$ such that $\mu(t) = e_t \# \eta$ for any t and, for η -a.e. $\gamma \in C^0([0, T], \mathbb{T}^d)$, γ is a solution to the ODE (2.58).*

See, for instance, Theorem 8.2.1. from [8].

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

Proof of Proposition 2.4.1. Let μ be a solution of the transport equation (2.55). From Ambrosio superposition principle, there exists a Borel probability measure η on $C^0([0, T], \mathbb{T}^d)$ such that $\mu(t) = e_t \# \eta$ 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 $m_0 = e_0 \# \eta$, 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 (2.56). 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 $\phi : \mathbb{T}^d \rightarrow \mathbb{R}$, 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. □

Chapter 3

Learning in anonymous non atomic games

Submitted in INFORMS: Mathematics of Operations Research

3.1 Introduction

Mean field games (MFGs) are symmetric differential games with an infinite number of non-atomic players. The model was first introduced simultaneously by Lasry and Lions [73][74][70] and Huang, Caines and Malhamé [66][67]. In this game, each player chooses a control and accordingly, incurs a cost that depends on the distribution of all the other players' states. More formally, a typical player chooses a path $\gamma : [0, T] \rightarrow \mathbb{R}^d$, $\gamma(0) = x$ via a control $\alpha : [0, T] \rightarrow \mathbb{R}^d$, with the dynamic $d\gamma(t) = \alpha_t dt$, and incurs the cost:

$$J(\gamma, (m_t)_{t \in [0, T]}) = \int_0^T \left(L(\gamma(t), \alpha_t) + f(\gamma(t), m_t) \right) dt + g(\gamma(T), m_T), \quad (3.1)$$

where $(m_t)_{t \in [0, T]} \subseteq \mathcal{P}(\mathbb{R}^d)$ is the evolving distribution of other players. The Lagrangian $L : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ captures the running cost depending on the tuple $(\gamma(t), \alpha_t)$ and $f, g : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$ are the maps describing the cost of interaction of this player with other players. Under classical assumptions, the optimal control of this player, that minimizes the cost (3.1), can be obtained by solving the Hamilton-Jacobi equation:

$$-\partial_t u + H(x, \nabla u(x, t)) = f(x, m_t), \quad u(x, T) = g(x, m_T)$$

with $H(x, p) = -\inf_{v \in \mathbb{R}^d} \langle p, v \rangle + L(x, v)$. The desired optimal control will be computed as

$$\tilde{\alpha}(x, t) = -D_p H(x, \nabla u(x, t)).$$

for almost every $(x, t) \in \mathbb{R}^d \times [0, T]$. If every player chooses its optimal control, the evolving distribution of players is given by the Fokker-Planck equation:

$$\partial_t m - \operatorname{div}(m D_p H(x, \nabla u)) = 0, \quad m(0, x) = m_0(x).$$

Hence the notion of Nash Equilibrium (or stability) is captured by the system of coupled Hamilton-Jacobi (backward) and Fokker-Planck (forward) equations written above.

The equilibrium configuration in MFGs is quite complicated and its occurrence requires a huge amount of information and a large degree of cooperation between players. The question of formation of equilibrium arises naturally. Thus, one would conclude that the formation of MFG equilibrium is justifiable

because there is a reasonable way of adapting (or learning) of players via observation and revision of the beliefs about the other players' behaviour.

In the current chapter, our main purpose is to prove the convergence of some learning procedures to the Nash equilibrium in first-order MFGs with monotone costs; however, since the approach can be used for a larger class of games, we work under a more general framework, that is the model of non atomic anonymous games.

Non atomic anonymous games model the strategic situations where there is a huge set of negligible agents (reflecting the non atomic nature), and cost functions depend to the distribution of actions (reflecting the anonymity characteristic). The non atomic games are known in the literature; look at Schmeidler [89], Mas-Colell [76], as the seminal works in this area. Contrasting to the approach by Mas-Colell [76], we work with a non atomic game with player dependent action sets and an identical cost function for all of players. This is the case in first-order MFGs; the players choose the paths with fixed (player dependent) initial positions as their actions, and the cost function as in (3.1), is identical for all players.

We provided sufficient conditions proving the existence of an equilibrium. Moreover, we proved the uniqueness of the equilibrium under an adapted monotonicity notion. The monotonicity condition in game frameworks, introduced by Rosen [88]. The strict monotonicity yields the uniqueness of the Nash equilibrium in several games (Lasry and Lions [73],[74], Hofbauer and Sandholm [65], Blanchet and Carlier [24]). In non atomic anonymous games with (not necessarily strict) monotone costs, equilibrium uniqueness is a direct consequence of monotonicity and an additional assumption, called the unique minimiser condition.

There are several learning procedures in static games with finitely many players and/or a finite number of actions per player (see for example the monograph [53]). Here we extend two of the most known of them to non atomic anonymous games: fictitious play and online mirror descent.

Fictitious play introduced by Brown[26], describes a learning procedure in which a fixed game is played over and over in repeated discrete rounds. At every round, each player sets their belief as the empirical frequency of play of the player's opponents, and then chooses its best action with respect to this belief. Convergence to a Nash equilibrium has been proved for different classes of finite games, for example potential games (Monderer, Shapley [79]), zero sum games (Robinson [87]) and 2×2 games (Miyasawa [78]). Cardaliaguet, Hadikhanloo [39] proved the convergence of fictitious play in first and second order potential MFGs. Our approach here covers a different class of first-order MFGs, i.e. the ones with monotone costs.

The second procedure we consider is the online mirror descent (OMD). The method was first introduced by Nemirovski, Yudin [82], as a generalization of standard gradient descent. The form of the algorithm is closely related to the notion of no-regret procedures in online optimization. A good explanatory introduction can be found in Shalev Shwartz[90]. Roughly speaking, the procedure deals with two variables, a primal one and a dual one. They are revised at every round; the dual is revised by using the sub-gradient of the objective function and the primal is obtained by a *quasi* projection via a strongly convex penalty function on the convex domain. Mertikopolous [77] proved the convergence of OMD to equilibria in the class of games with convex action sets and concave costs. Here we examine the convergence properties of OMD in monotone anonymous games.

In the proof of convergence of both procedures to the Nash equilibrium, we define a value $\phi_n \in \mathbb{R}, n \in \mathbb{N}$ measuring how much the actual behaviour at step n is far from being an equilibrium; in fictitious play the quantity ϕ_n is calculated by using the best response function and in OMD by the Fenchel coupling. We then prove that indeed $\lim_{n \rightarrow \infty} \phi_n = 0$; this gives our desired convergence toward the equilibrium.

Here is how the paper is organized: in section 3.2 a general model of anonymous game is proposed.

The notion of Nash equilibrium is reviewed and the existence is proved under general continuity conditions. Then we define monotonicity in terms of the cost function, and its consequence on the uniqueness of the Nash equilibrium. Section 3.3 is devoted to the definition of fictitious play and its convergence under Lipschitz regularity conditions. Section 3.4 deals with the online mirror descent algorithm and its convergence. Section 3.5 shows that the first order MFG can be considered as an example of anonymous games and shows that the previous results can be applied under suitable conditions. For sake of completeness, we provide in the Appendix some disintegration theorems which are used in the proofs.

Acknowledgement. The extension of online mirror descent to the case of non atomic anonymous games was inspired from the explanations of Panayotis Mertikopoulos; I would like to sincerely thank him for his permanent supports. I wish to thank as well the support of ANR (Agence Nationale de la Recherche) MFG (ANR-16-CE40-0015-01).

3.2 Non atomic anonymous games

3.2.1 Model

Let us introduce our general model of anonymous game G . For a measure space X let $\mathcal{P}(X)$ denotes the set of probability measures on X . Let I be the set of players and $\lambda \in \mathcal{P}(I)$ a prior non-atomic probability measure on I modelling the repartition of players on I . Let V be a measure space. For every player $i \in I$, let $A_i \subset V$ be the action set of i . Define the set of admissible profiles of actions

$$\mathcal{A} = \{\Psi : I \rightarrow V \text{ measurable} \mid \Psi(i) \in A_i \text{ for } \lambda\text{-almost every } i \in I\}.$$

We identify the action profiles up to λ -zero measure subsets of I , i.e. $\Psi_1 = \Psi_2$ iff $\Psi_1(i) = \Psi_2(i)$ for λ -almost every $i \in I$. The induced measure of a typical profile $\Psi \in \mathcal{A}$ on the set of actions, that captures the portion of players who have chosen a given subset of actions, is denoted by $\Psi\#\lambda \in \mathcal{P}(V)$. More precisely, $\Psi\#\lambda$ is the push-forward of the measure of λ by the application Ψ , that is for every measurable set $B \subseteq V$ we have $\Psi\#\lambda(B) = \lambda(\Psi^{-1}(B))$. Since the set consisting of measures $\Psi\#\lambda$ for all admissible profiles Ψ , may be different from $\mathcal{P}(V)$, it is sufficient to work with:

$$\mathcal{P}_G(V) = \{\eta \in \mathcal{P}(V) \mid \exists \Psi \in \mathcal{A} : \eta = \Psi\#\lambda\}.$$

For every $i \in I$ let $c_i : \mathcal{A} \rightarrow \mathbb{R}$ be the cost paid by player i . We call the game anonymous, if for every player $i \in I$, there exists $J_i : A_i \times \mathcal{P}_G(V) \rightarrow \mathbb{R}$ such that $c_i(\Psi) = J_i(\Psi(i), \Psi\#\lambda)$. In other words, $J_i(a, \eta)$ captures the cost endured by a typical player $i \in I$, whose action is $a \in A_i$ while facing the distribution of actions $\eta \in \mathcal{P}(V)$ chosen by other players. We consider here anonymous games where the players have identical cost function, i.e. there is $J : V \times \mathcal{P}_G(V) \rightarrow \mathbb{R}$ such that for every $i \in I$ we have $J_i = J$. We use the following notation for referring to such game:

$$G = (I, \lambda, V, (A_i)_{i \in I}, J).$$

Example 3.2.1 (Population Game [65]). Set $I = [0, 1]$ be the set of players and λ the Lebesgue measure as the distribution of players on I . Let $N \in \mathbb{N}$ represents the number of populations in the game i.e. there is a partition of players $I_1, I_2, \dots, I_N \subseteq I$ where for every $1 \leq p \leq N, I_p \subseteq I$ represents the set of players belonging to population p . For every player $i \in I$ suppose the set of actions A_i is finite and depends only on the population where the player i comes from, i.e. for every population p there is S_p such that for all $i \in I_p$ we have $A_i = S_p$. Set $V = \cup_p S_p$. For every population p the cost function has the form $J_p : S_p \times \Delta(V) \rightarrow \mathbb{R}$ where $J_p(a, (m_j)_{1 \leq j \leq |V|})$ is the cost payed by a typical player in population

p whose action is $a \in S_p$ while facing $(m_j)_{1 \leq j \leq |V|}$ where for every $1 \leq j \leq |V|$, $m_j \geq 0$ is the portion of players who have chosen action $j \in V$. The form of the cost function illustrates the fact that the population games are anonymous.

Example 3.2.2. In section 5, we show that the First order MFG is an anonymous game with suitable actions sets and cost function.

3.2.2 Nash equilibria

Inspired from the notion of Nash equilibrium in non-atomic games (see Schmeidler [89], Mas-Colell [76]), we omit the effect of λ -zero measure subsets of players in the definition of equilibria:

Definition 3.2.1. A profile $\tilde{\Psi} \in \mathcal{A}$ is called a Nash equilibrium if

$$\tilde{\Psi}(i) \in \arg \min_{a \in A_i} J(a, \tilde{\Psi} \# \lambda) \quad \text{for } \lambda\text{-almost every } i \in I.$$

The corresponding distribution $\tilde{\eta} = \tilde{\Psi} \# \lambda$ is called a Nash (or equilibrium) distribution.

One can note that the definition of Nash equilibrium highly depends on the prior distribution of players λ . The following theorem gives a sufficient condition under which the game possesses at least one equilibrium. Let I be a topological and V be a metric space (with $\mathcal{B}(I), \mathcal{B}(V)$ as their σ -fields). Suppose the A_i 's are uniformly bounded for λ -almost every $i \in I$, i.e. there exist $M > 0, v \in V$ such that:

$$\text{for } \lambda\text{-almost every } i \in I \text{ and every } a \in A_i : \quad d_V(v, a) < M. \quad (3.2)$$

This condition gives us $\mathcal{P}_G(V) \subseteq \mathcal{P}_1(V)$ where:

$$\mathcal{P}_1(V) = \{ \eta \in \mathcal{P}(V) \mid \exists v \in V : \int_V d_V(v, a) \, d\eta(a) < +\infty \}$$

endowed with the metric:

$$d_1(\eta_1, \eta_2) = \sup_{h: V \rightarrow \mathbb{R}, \text{ 1-Lipschitz}} \int_V h(a) \, d(\eta_1 - \eta_2)(a).$$

For technical reasons we work with closure convex hull of $\mathcal{P}_G(V)$ i.e. $\overline{\text{cov}(\mathcal{P}_G(V))}$.

Definition 3.2.2. We say $G = (I, \lambda, V, (A_i)_{i \in I}, J)$ satisfies the unique minimiser condition, if for every $\eta \in \overline{\text{cov}(\mathcal{P}_G(V))}$, there exists $I_\eta \subseteq I$ with $\lambda(I \setminus I_\eta) = 0$, such that for all $i \in I_\eta$ there is exactly one $a \in A_i$ minimizing $J(\cdot, \eta)$ in A_i .

Informally, the definition says facing to every distribution of actions, (almost) every player has a unique best response.

Definition 3.2.3. A correspondence $A : I \rightarrow V, A(i) = A_i$ is called continuous if:

- it is upper semi continuous i.e. the graph $\{(i, a) \in I \times V \mid a \in A_i\}$ is closed in $I \times V$,
- it is lower semi continuous i.e. for every open set $U \subseteq V$ the set $\{i \in I \mid A_i \cap U \neq \emptyset\}$ is open in I .

For more detailed theorems about set valued maps, see [9].

Assumption 3.2.1. Here are the assumptions we consider for the non atomic anonymous games:

1. the correspondence $A : I \rightarrow V, A(i) = A_i$ is continuous and compact valued,

2. there is an extension $J : V \times \overline{\text{cov}(\mathcal{P}_G(V))} \rightarrow \mathbb{R}$ which is lower semi-continuous,
3. the function $\text{Min} : I \times \mathcal{P}_G(V) \rightarrow \mathbb{R}$, $\text{Min}(i, \eta) := \min_{a \in A_i} J(a, \eta)$ is continuous,
4. $\overline{\text{cov}(\mathcal{P}_G(V))}$ is compact,
5. G satisfies the unique minimiser condition.

Theorem 3.2.1. *Let $G = (I, \lambda, V, (A_i)_{i \in I}, J)$ be an anonymous game. Suppose the assumptions (3.2.1) hold. Then G will admit at least a Nash equilibrium.*

Assumptions (3.2.1)(1-4) provide enough continuity and compactness conditions we need for the fixed point theorem. The assumption (3.2.1)(5) allows us to prove the existence of pure Nash equilibrium; it is crucial as well for the uniqueness of equilibrium and convergence results in learning procedures that we will propose. So we add it here as an assumption for being coherent in the entire chapter. Before we start the proof let us provide some lemmas which will be used here and in the rest of paper:

Lemma 3.2.1. *Define the best response correspondence as follows*

$$BR : I \times \overline{\text{cov}(\mathcal{P}_G(V))} \rightarrow V, \quad BR(i, \eta) = \arg \min_{a \in A_i} J(a, \eta).$$

If the assumptions (3.2.1) hold, then for every $\eta \in \mathcal{P}_G(V)$ the correspondence $BR(\cdot, \eta) : I \rightarrow V$, that is almost everywhere singleton, is almost everywhere continuous and hence measurable.

Proof. Fix $\eta \in \overline{\text{cov}(\mathcal{P}_G(V))}$. According to the unique minimiser condition there exists $I_\eta \subseteq I$ with $\lambda(I \setminus I_\eta) = 0$, such that $BR(i, \eta)$ is singleton for every $i \in I_\eta$. We will show the continuity of the restricted best response function $BR(\cdot, \eta) : I_\eta \rightarrow V$ which completes our proof. Consider $i, i_n \in I_\eta$ such that $i_n \rightarrow i$. Set $a_n = BR(i_n, \eta)$. The set $\{a_n\}_{n \in \mathbb{N}}$ is pre-compact since $A : I \rightarrow V$ is a compact valued correspondence and hence $A(\{i_n\}_{n \in \mathbb{N}} \cup \{i\}) = \cup_n A_{i_n} \cup A_i$ is compact. Suppose $\tilde{a} \in V$ is an accumulation point of $\{a_n\}_{n \in \mathbb{N}}$. So there is a sub-sequence $\{a_{n_k}\}_{k \in \mathbb{N}}$ such that $\lim_{k \rightarrow \infty} a_{n_k} = \tilde{a}$. We have $\tilde{a} \in A_i$ since the correspondence $A : I \rightarrow V$ is upper semi continuous and $a_n \in A_{i_n}$. By definition $J(a_n, \eta) = \text{Min}(i_n, \eta)$ which gives:

$$J(\tilde{a}, \eta) \leq \liminf_{n_k} J(a_{n_k}, \eta) = \liminf_{n_k} \text{Min}(i_{n_k}, \eta) = \text{Min}(i, \eta),$$

since the Min function is continuous. It yields $\tilde{a} = BR(i, \eta)$. So every accumulation point of $\{a_n\}_{n \in \mathbb{N}}$ should be $BR(i, \eta)$ which shows $a_n \rightarrow BR(i, \eta)$. \square

Lemma 3.2.2. *Define the best response distribution function $\Theta : \overline{\text{cov}(\mathcal{P}_G(V))} \rightarrow \mathcal{P}_G(V)$ as follows:*

$$\Theta(\eta) = BR(\cdot, \eta) \# \lambda, \quad \text{for every } \eta \in \overline{\text{cov}(\mathcal{P}_G(V))}.$$

If the assumptions (3.2.1) hold then Θ is continuous.

Proof. Let $\eta_n \rightarrow \eta$. If $J = I_\eta \cap_{n \in \mathbb{N}} I_{\eta_n}$ then we have $\lambda(I \setminus J) = 0$. One can show as for Lemma 3.2.1 that for every $i \in J$:

$$BR(i, \eta_n) \rightarrow BR(i, \eta).$$

Since the A_i 's are uniformly bounded for λ -almost every $i \in J$, the dominated Lebesgue convergence theorem implies $\int_I d_V(BR(i, \eta_n), BR(i, \eta)) d\lambda(i) \rightarrow 0$. Thus $\Theta(\eta_n) \xrightarrow{d_1} \Theta(\eta)$ since:

$$\begin{aligned} d_1(\Theta(\eta_n), \Theta(\eta)) &= \sup_{f: V \rightarrow \mathbb{R}, 1\text{-Lipschitz}} \int_V f(v) d(\Theta(\eta_n) - \Theta(\eta))(v) = \\ &= \sup_{f: V \rightarrow \mathbb{R}, 1\text{-Lipschitz}} \int_I (f(BR(i, \eta_n)) - f(BR(i, \eta))) d\lambda(i) \leq \int_I d_V(BR(i, \eta_n), BR(i, \eta)) d\lambda(i) \rightarrow 0. \end{aligned}$$

\square

Proof of Theorem 3.2.1. Consider the best response distribution function Θ defined in Lemma 3.2.2. We have by definition

$$\Theta(\overline{\text{cov}(\mathcal{P}_G(V))}) \subset \mathcal{P}_G(V) \subset \overline{\text{cov}(\mathcal{P}_G(V))},$$

which implies that the image of Θ is pre-compact. Since Θ is continuous (Lemma 3.2.2) and $\overline{\text{cov}(\mathcal{P}_G(V))}$ is convex, by the Schauder's fixed point theorem, there is $\tilde{\eta} \in \overline{\text{cov}(\mathcal{P}_G(V))}$ such that $\Theta(\tilde{\eta}) = \tilde{\eta}$. Since $\Theta(\tilde{\eta}) = BR(\cdot, \tilde{\eta})\sharp\lambda \in \mathcal{P}_G(V)$ so if we set $\tilde{\Psi}(\cdot) = BR(\cdot, \tilde{\eta}) \in \mathcal{A}$ then

$$\tilde{\Psi}\sharp\lambda = \tilde{\eta}, \quad \tilde{\Psi}(i) \in \arg \min_{a \in A_i} J(a, \tilde{\eta}) \quad \text{for } \lambda\text{-almost every } i \in I.$$

This means $\tilde{\Psi}$ is the desired Nash equilibrium. \square

3.2.3 Anonymous games with monotone cost

Here we give a definition of monotonicity and its additional consequences on the structure of the game and its equilibria.

Definition 3.2.4. *The anonymous game $G = (I, \lambda, V, (A_i)_{i \in I}, J)$ has a monotone cost J if for any $\eta, \eta' \in \overline{\text{cov}(\mathcal{P}_G(V))}$:*

$$\int_V |J(a, \eta)| \, d\eta'(a) < +\infty,$$

and

$$\int_V (J(a, \eta) - J(a, \eta')) \, d(\eta - \eta')(a) \geq 0.$$

We call J a strict monotone cost function if the later inequality holds strictly for $\eta \neq \eta'$.

This condition is usually interpreted as the aversion of players for choosing actions that are chosen by many of players i.e. congestion avoiding effect.

Remark 3.2.1. *If J is monotone and if $\tilde{\Psi} \in \mathcal{A}$ is a Nash equilibrium, then for every $\Psi \in \mathcal{A}$ we have:*

$$\text{if } \tilde{\eta} = \tilde{\Psi}\sharp\lambda, \quad \eta = \Psi\sharp\lambda : \quad \int_V J(a, \eta) \, d(\eta - \tilde{\eta})(a) \geq \int_V J(a, \tilde{\eta}) \, d(\eta - \tilde{\eta})(a) \geq 0.$$

Proof. Since J is monotone we have $\int_V (J(a, \eta) - J(a, \tilde{\eta})) \, d(\eta - \tilde{\eta})(a) \geq 0$ and so:

$$\int_V J(a, \eta) \, d(\eta - \tilde{\eta})(a) \geq \int_V J(a, \tilde{\eta}) \, d(\eta - \tilde{\eta})(a).$$

On the other hand

$$\int_V J(a, \tilde{\eta}) \, d(\eta - \tilde{\eta})(a) = \int_I (J(\Psi(i), \tilde{\eta}) - J(\tilde{\Psi}(i), \tilde{\eta})) \, d\lambda(i)$$

by the definition of push-forward measures. Since $\tilde{\Psi}$ is an equilibrium, for λ -almost every $i \in I$, we have $J(\Psi(i), \tilde{\eta}) - J(\tilde{\Psi}(i), \tilde{\eta}) \geq 0$, which gives our result. \square

The strict monotonicity yields the uniqueness of the Nash equilibrium in different frameworks, e.g. Haufbauer, Sandholm [65], Blanchet, Carlier [24], Lasry, Lions [73]. In the following we show that in non atomic anonymous games, the monotonicity and unique minimiser conditions are sufficient for the uniqueness of the equilibrium.

Theorem 3.2.2. *Consider a game $G = (I, \lambda, V, (A_i)_{i \in I}, J)$. Then the game G admits at most one Nash equilibrium if J is monotone and G satisfies the unique minimiser condition.*

Proof. Let $\Psi_1, \Psi_2 \in \mathcal{A}$ be two Nash equilibria. We will show that $\Psi_1(i) = \Psi_2(i)$ for λ -almost every $i \in I$. Set $\eta_i = \Psi_i \# \lambda$ for $i = 1, 2$. Since Ψ_1 is an equilibrium, we have:

$$\int_I (J(\Psi_1(i), \eta_1) - J(\Psi_2(i), \eta_1)) \, d\lambda(i) \leq 0,$$

since $J(\Psi_1(i), \eta_1) \leq J(\Psi_2(i), \eta_1)$ for λ -almost every $i \in I$. On the other hand:

$$\int_I (J(\Psi_1(i), \eta_1) - J(\Psi_2(i), \eta_1)) \, d\lambda(i) = \int_V J(a, \eta_1) \, d(\eta_1 - \eta_2)(a),$$

from the definition since $\Psi_i \# \lambda = \eta_i$ for $i = 1, 2$. So

$$\int_V J(a, \eta_1) \, d(\eta_1 - \eta_2)(a) \leq 0 \quad \text{and (similarly)} \quad \int_V J(a, \eta_2) \, d(\eta_2 - \eta_1)(a) \leq 0.$$

Summing up the last inequalities gives us:

$$\int_V (J(a, \eta_1) - J(a, \eta_2)) \, d(\eta_1 - \eta_2)(a) \leq 0.$$

Hence by monotonicity of J we should have the equality in the previous inequalities. So for λ -almost every $i \in I$, one has $J(\Psi_1(i), \eta_1) = J(\Psi_2(i), \eta_1)$. Since $\Psi_1(i) \in A_i$ is the unique minimiser of $J(\cdot, \eta_1)$ on A_i so $\Psi_1(i) = \Psi_2(i)$ for λ -almost every $i \in I$. \square

Remark 3.2.2. *One can similarly show that if J is strictly monotone and not necessarily satisfies the unique minimizer condition, then there exists at most one Nash equilibrium distribution.*

3.3 Fictitious play in anonymous games

Here we introduce a learning procedure similar to the fictitious play and prove its convergence to the unique Nash equilibrium under monotonicity condition.

Let $G = (I, \lambda, V, (A_i)_{i \in I}, J)$. For technical reasons, we suppose that assumptions (3.2.1) hold throughout this section. Suppose G is being played repeatedly on discrete rounds $n = 1, 2, \dots$. At every round, the players set their belief equals to the average of the action distribution observed in the previous rounds and then react their best to such belief. At the end of the round players revise their beliefs by a new observation. More formally, consider $\Psi_1 \in \mathcal{A}$, $\bar{\eta}_1 = \eta_1 = \Psi_1 \# \lambda \in \mathcal{P}_G(V)$ an arbitrary initial belief. Construct recursively $(\Psi_n, \eta_n, \bar{\eta}_n) \in \mathcal{A} \times \mathcal{P}_G(V) \times \text{cov}(\mathcal{P}_G(V))$ for $n = 1, 2, \dots$ as follows:

$$\begin{aligned} (i) \quad \Psi_{n+1}(i) &= BR(i, \bar{\eta}_n), \quad \text{for } \lambda\text{-almost every } i \in I, \\ (ii) \quad \eta_{n+1} &= \Psi_{n+1} \# \lambda, \\ (iii) \quad \bar{\eta}_{n+1} &= \frac{1}{n+1} \sum_{k=1}^{n+1} \eta_k = \frac{n}{n+1} \bar{\eta}_n + \frac{1}{n+1} \eta_{n+1}. \end{aligned} \tag{3.3}$$

One should notice that by assumption (3.2.1)(5) and Lemma 3.2.1 the expressions in (3.3)(i, ii) are well defined. We will show now that this procedure converges to the Nash Equilibrium when G is monotone.

Theorem 3.3.1. *Consider a non atomic anonymous game $G = (I, \lambda, V, (A_i)_{i \in I}, J)$ satisfying assumptions 3.2.1. Suppose the cost function J is monotone and there exists $C > 0$ such that for all $a, b \in V$, $\eta, \eta' \in \overline{\text{cov}(\mathcal{P}_G(V))}$:*

$$\begin{aligned} |J(a, \eta) - J(a, \eta') - J(b, \eta) + J(b, \eta')| &\leq C \, d_V(a, b) \, d_1(\eta, \eta'), \\ |J(a, \eta) - J(a, \eta')| &\leq C \, d_1(\eta, \eta'). \end{aligned} \tag{3.4}$$

Construct $(\Psi_n, \eta_n, \bar{\eta}_n) \in \mathcal{A} \times \mathcal{P}_G(V) \times \text{cov}(\mathcal{P}_G(V))$ for $n \in \mathbb{N}$ by applying the fictitious play procedure proposed in (3.3). Then:

$$\eta_n, \bar{\eta}_n \xrightarrow{d_1} \tilde{\eta}$$

where $\tilde{\eta} \in \mathcal{P}_G(V)$ is the unique Nash equilibrium distribution.

Inspired from [65], the proof requires several steps. The key idea is to use the quantities $\phi_n \in \mathbb{R}$ defined by

$$\phi_n = \int_V J(a, \bar{\eta}_n) d(\bar{\eta}_n - \eta_{n+1})(a), \quad \text{for every } n \in \mathbb{N}.$$

Since the best response distribution of $\bar{\eta}_n$ is η_{n+1} , the quantity ϕ_n describes how much $\bar{\eta}_n$ is far from being an equilibrium. By using monotonicity and the regularity conditions, one gets

$$\forall n \in \mathbb{N}: \quad \phi_{n+1} - \phi_n \leq -\frac{1}{n+1}\phi_n + \frac{\epsilon_n}{n},$$

for suitable $\{\epsilon_n\}_{n \in \mathbb{N}}$ such that $\lim_{n \rightarrow \infty} \epsilon_n = 0$. We show the later inequality is sufficient to prove $\lim_{n \rightarrow \infty} \phi_n = 0$ and then we conclude that the accumulation points of $\bar{\eta}_n, \eta_n$ is the equilibrium distribution $\bar{\eta}$. As one will see, the unique minimiser assumption plays a key role in Lemma 3.3.2 and hence in our main result.

Lemma 3.3.1. *Consider a sequence of real numbers $\{\phi_n\}_{n \in \mathbb{N}}$ such that $\liminf_n \phi_n \geq 0$. If there exists a real sequence $\{\epsilon_n\}_{n \in \mathbb{N}}$ such that $\lim_{n \rightarrow \infty} \epsilon_n = 0$ and :*

$$\forall n \in \mathbb{N}: \quad \phi_{n+1} - \phi_n \leq -\frac{1}{n+1}\phi_n + \frac{\epsilon_n}{n},$$

then $\lim_{n \rightarrow \infty} \phi_n = 0$.

Proof. Let $b_n = n\phi_n$ for every $n \in \mathbb{N}$. We have:

$$\forall n \in \mathbb{N}: \quad \frac{b_{n+1}}{n+1} - \frac{b_n}{n} \leq -\frac{b_n}{n(n+1)} + \frac{\epsilon_n}{n},$$

which implies $b_{n+1} \leq b_n + (n+1)\epsilon_n/n \leq b_n + 2|\epsilon_n|$. Then we get $b_n \leq b_1 + 2 \sum_{i=1}^{n-1} |\epsilon_i|$ for $n \in \mathbb{N}$ and so:

$$0 \leq \liminf_n \phi_n \leq \limsup_n \phi_n \leq \limsup_n \frac{b_1 + 2 \sum_{i=1}^{n-1} |\epsilon_i|}{n} = 0.$$

which proves $\lim_{n \rightarrow \infty} \phi_n = 0$. □

Lemma 3.3.2. *Let $(\eta_n)_{n \in \mathbb{N}}$ be defined by (3.3). Then*

$$d_1(\bar{\eta}_n, \bar{\eta}_{n+1}) = O(1/n), \quad \lim_{n \rightarrow \infty} d_1(\eta_n, \eta_{n+1}) = 0.$$

Proof. Let $M > 0, v \in V$ be chosen from (3.2). For every 1-Lipschitz continuous map $h : V \rightarrow \mathbb{R}$ we have:

$$\begin{aligned} \left| \int_V h(a) d(\bar{\eta}_{n+1} - \bar{\eta}_n) \right| &= \frac{1}{n+1} \left| \int_V h(a) d(\eta_{n+1} - \bar{\eta}_n)(a) \right| \\ &= \frac{1}{n+1} \left| \int_V (h(a) - h(v)) d(\eta_{n+1} - \bar{\eta}_n)(a) \right| \\ &\leq \frac{1}{n+1} \left(\int_V d_V(a, v) d\eta_{n+1}(a) + \frac{1}{n} \sum_{k=1}^n \int_V d_V(a, v) d\eta_k(a) \right). \end{aligned} \tag{3.5}$$

By the definition we have:

$$\int_V d_V(a, v) d\eta_k(a) = \int_V d_V(\Psi_k(i), v) d\lambda(i) \leq M, \quad \text{for every } k \in \mathbb{N}.$$

So we can write

$$\left| \int_V h(a) d(\bar{\eta}_{n+1} - \bar{\eta}_n) \right| \leq \frac{2M}{n+1},$$

or $d_1(\bar{\eta}_n, \bar{\eta}_{n+1}) \leq \frac{2M}{n+1}$ since h is an arbitrary 1-Lipschitz continuous function.

For the second part of the lemma, let us consider the best response distribution function Θ defined in Lemma 3.2.2. Since Θ is continuous (Lemma 3.2.2) and $\overline{\text{cov}(\mathcal{P}_G(V))}$ is compact, there exists a non decreasing continuity modulus

$$\omega : \mathbb{R}^+ \rightarrow \mathbb{R}^+, \quad \lim_{x \rightarrow 0^+} \omega(x) = 0$$

such that:

$$\forall \eta_1, \eta_2 \in \overline{\text{cov}(\mathcal{P}_G(V))} : d_1(\Theta(\eta_1), \Theta(\eta_2)) \leq \omega(d_1(\eta_1, \eta_2)).$$

Since for all $n \in \mathbb{N}$ we have $\bar{\eta}_n \in \overline{\text{cov}(\mathcal{P}_G(V))}$ and $\Theta(\bar{\eta}_n) = \eta_{n+1}$ we have

$$0 \leq d_1(\eta_{n+1}, \eta_{n+2}) = d_1(\Theta(\bar{\eta}_n), \Theta(\bar{\eta}_{n+1})) \leq \omega(d_1(\bar{\eta}_n, \bar{\eta}_{n+1})).$$

It gives our desired result since $d_1(\bar{\eta}_n, \bar{\eta}_{n+1}) = O(1/n)$. \square

The proof of previous lemma relies heavily on the unique minimizer assumption. Instead without it, one cannot conclude that η_n, η_{n+1} are close even if $\bar{\eta}_n, \bar{\eta}_{n+1}$ are so. Even for $\bar{\eta}_n = \bar{\eta}_{n+1}$, one might have very different best responses η_n and η_{n+1} .

Proof of Theorem 3.3.1. Let $\{\phi_n\}_{n \in \mathbb{N}}$ be defined by:

$$\phi_n = \int_V J(a, \bar{\eta}_n) d(\bar{\eta}_n - \eta_{n+1})(a), \quad \text{for every } n \in \mathbb{N}.$$

We have $\phi_n \geq 0$ for all $n \in \mathbb{N}$. Indeed, rewriting the definition of ϕ_n , we have:

$$\phi_n = \int_I \frac{1}{n} \sum_{k=1}^n (J(\Psi_k(i), \bar{\eta}_n) - J(BR(i, \bar{\eta}_n), \bar{\eta}_n)) d\lambda(i),$$

and the positiveness comes from the definition of the best response. We now prove that exists $C > 0$ such that:

$$\phi_{n+1} - \phi_n \leq -\frac{1}{n+1} \phi_n + C \frac{d_1(\eta_n, \eta_{n+1}) + 1/n}{n}, \quad \text{for every } n \in \mathbb{N}. \quad (3.6)$$

Let us rewrite $\phi_{n+1} - \phi_n = A + B$, where:

$$\begin{aligned} A &= \int_V J(a, \bar{\eta}_{n+1}) d\bar{\eta}_{n+1}(a) - \int_V J(a, \bar{\eta}_n) d\bar{\eta}_n(a), \\ B &= \int_V J(a, \bar{\eta}_n) d\eta_{n+1}(a) - \int_V J(a, \bar{\eta}_{n+1}) d\eta_{n+2}(a). \end{aligned}$$

We have:

$$\begin{aligned} B &\leq \int_V J(a, \bar{\eta}_n) d\eta_{n+2}(a) - \int_V J(a, \bar{\eta}_{n+1}) d\eta_{n+2}(a) \\ &= \int_V (J(a, \bar{\eta}_n) - J(a, \bar{\eta}_{n+1})) d\eta_{n+2}(a) \\ &\leq \int_V (J(a, \bar{\eta}_n) - J(a, \bar{\eta}_{n+1})) d\eta_{n+1}(a) + \frac{C}{n} d_1(\eta_{n+1}, \eta_{n+2}), \end{aligned}$$

since by (3.4) and Lemma 3.3.2 there exists C such that the function $J(\cdot, \bar{\eta}_n) - J(\cdot, \bar{\eta}_{n+1}) : V \rightarrow \mathbb{R}$ is a C/n -Lipschitz continuous function. Let us rewrite the expression A as follows:

$$\begin{aligned} A &= \int_V J(a, \bar{\eta}_{n+1}) d(\bar{\eta}_n + \frac{1}{n+1}(\eta_{n+1} - \bar{\eta}_n))(a) - \int_V J(a, \bar{\eta}_n) d\bar{\eta}_n(a) \\ &= \int_V (J(a, \bar{\eta}_{n+1}) - J(a, \bar{\eta}_n)) d\bar{\eta}_n(a) + \frac{1}{n+1} \int_V J(a, \bar{\eta}_{n+1}) d(\eta_{n+1} - \bar{\eta}_n)(a) \\ &\leq \int_V (J(a, \bar{\eta}_{n+1}) - J(a, \bar{\eta}_n)) d\bar{\eta}_n(a) + \frac{1}{n+1} \int_V J(a, \bar{\eta}_n) d(\eta_{n+1} - \bar{\eta}_n)(a) + \frac{C}{n^2} \end{aligned}$$

since by (3.4) and Lemma 3.3.2 we have $|J(a, \bar{\eta}_n) - J(a, \bar{\eta}_{n+1})| \leq C d_1(\bar{\eta}_{n+1}, \bar{\eta}_n) = O(1/n)$. So

$$A \leq \int_V (J(a, \bar{\eta}_{n+1}) - J(a, \bar{\eta}_n)) d\bar{\eta}_n(a) - \frac{\phi_n}{n+1} + \frac{C}{n^2}.$$

Then if we set $\epsilon_n = C(d_1(\eta_{n+1}, \eta_{n+2}) + 1/n)$, by using the above inequalities for A, B , we have :

$$\begin{aligned} A + B &\leq \int_V (J(a, \bar{\eta}_{n+1}) - J(a, \bar{\eta}_n)) d(\bar{\eta}_n - \eta_{n+1})(a) - \frac{\phi_n}{n+1} + \frac{\epsilon_n}{n} \\ &= -(n+1) \int_V (J(a, \bar{\eta}_{n+1}) - J(a, \bar{\eta}_n)) d(\bar{\eta}_{n+1} - \bar{\eta}_n)(a) - \frac{\phi_n}{n+1} + \frac{\epsilon_n}{n} \\ &\leq -\frac{\phi_n}{n+1} + \frac{\epsilon_n}{n}, \end{aligned} \quad (3.7)$$

and the last inequality comes from the monotonicity assumption. By Lemmas 4.3.1 and 3.3.2, the inequality (3.6) implies $\phi_n \rightarrow 0$. Let $(\eta, \bar{\eta}) \in \overline{\mathcal{P}_G(V) \times \text{cov}(\mathcal{P}_G(V))}$ be an accumulation point of the set $\{(\eta_{n+1}, \bar{\eta}_n)\}_{n \in \mathbb{N}}$. We have $\eta = \Theta(\bar{\eta})$ due to the continuity of best response distribution function Θ (Lemma 3.2.2) and the fact that $\eta_{n+1} = \Theta(\bar{\eta}_n)$.

Take an arbitrary $\theta \in \mathcal{P}_G(V)$. Since J is lower semi-continuous we have (see [8] section 5.1.1):

$$\begin{aligned} \int_V J(a, \bar{\eta}) d(\bar{\eta} - \theta)(a) &\leq \liminf \int_V J(a, \bar{\eta}) d(\bar{\eta}_n - \theta)(a) = \liminf \int_V J(a, \bar{\eta}_n) d(\bar{\eta}_n - \theta)(a) \\ &= \liminf \int_V J(a, \bar{\eta}_n) d(\eta_{n+1} - \theta)(a) + \phi_n \leq \liminf \phi_n = 0 \end{aligned}$$

since $\eta_{n+1} = \Theta(\bar{\eta}_n)$ and $\int_V J(a, \bar{\eta}_n) d(\eta_{n+1} - \theta)(a) \leq 0$ for every $\theta \in \mathcal{P}_G(V)$. So:

$$\forall \theta \in \mathcal{P}_G(V) : \int_V J(a, \bar{\eta}) d(\bar{\eta} - \theta)(a) \leq 0. \quad (3.8)$$

We rewrite the above inequality as follows: since $\bar{\eta} \in \overline{\text{cov}(\mathcal{P}_G(V))}$ by Corollary 3.6.1 we can disintegrate it with respect to $(A_i)_{i \in I}$ i.e. there are $\{\bar{\eta}^i\}_{i \in I} \subseteq \mathcal{P}(V)$ such that for λ -almost every $i \in I$ we have $\text{supp}(\bar{\eta}^i) \subset A_i$ and for every integrable function $h : V \rightarrow \mathbb{R}$:

$$\int_V h(a) d\bar{\eta}(a) = \int_I \int_{A_i} h(a) d(\bar{\eta}^i)(a) d\lambda(i).$$

Specially for $h = J(\cdot, \eta)$ we have:

$$\int_V J(a, \bar{\eta}) d\bar{\eta}(a) = \int_I \int_{A_i} J(a, \bar{\eta}) d(\bar{\eta}^i)(a) d\lambda(i),$$

and for all $\Psi \in \mathcal{A}$:

$$\int_V J(a, \bar{\eta}) d(\Psi \# \lambda)(a) = \int_I J(\Psi(i), \bar{\eta}) d\lambda(i) = \int_I \int_{A_i} J(\Psi(i), \bar{\eta}) d(\bar{\eta}^i)(a) d\lambda(i).$$

Combining the previous equalities with (3.8), gives us:

$$\forall \Psi \in \mathcal{A} : \int_I \int_{A_i} (J(a, \bar{\eta}) - J(\Psi(i), \bar{\eta})) d(\bar{\eta}^i)(a) d\lambda(i) = \int_V J(a, \bar{\eta}) d(\bar{\eta} - \Psi \# \lambda)(a) \leq 0.$$

In particular if $\Psi = BR(\cdot, \bar{\eta})$ we have:

$$\int_I \int_{A_i} (J(a, \bar{\eta}) - J(BR(i, \bar{\eta}), \bar{\eta})) d(\bar{\eta}^i)(a) d\lambda(i) \leq 0,$$

which gives the equality by definition of best response action. So by unique minimizer we have $\bar{\eta}^i = \delta_{BR(i, \bar{\eta})}$ for λ -almost every $i \in I$. It means $\bar{\eta} = BR(\cdot, \bar{\eta}) \# \lambda$ or $\bar{\eta} = \Theta(\bar{\eta})$. Hence $\bar{\eta} = \eta$ and they are both equal to $\bar{\eta} \in \mathcal{P}_G(V)$, the unique fixed point of Θ , or equivalently, the unique equilibrium distribution. \square

3.4 Online mirror descent

Here we investigate the convergence to a Nash equilibrium by applying Online Mirror Descent (OMD) in anonymous games. The form of OMD algorithm is closely related to the online optimization and no regret algorithms. The reader can find a good explanatory note in [90]. The goal of the algorithm is to act optimally in online manner by "minimizing" a function that itself changes at each step. In the game frameworks, the cost function changes due to change of the actions chosen by adversaries in each round. As one can notice in the following, we need the structure of vector space for the action sets.

3.4.1 Preliminaries

Before we propose the main OMD, let us review some definitions and lemmas.

Definition 3.4.1. Let $(W, \|\cdot\|_W)$ be a normed vector space. For $K > 0$ we say that $h : W \rightarrow \mathbb{R}$ is a K -strongly convex function if

$$\forall a_1, a_2 \in W, \forall \lambda \in [0, 1] : \quad h(\lambda a_1 + (1 - \lambda)a_2) \leq \lambda h(a_1) + (1 - \lambda)h(a_2) - K\lambda(1 - \lambda)\|a_1 - a_2\|_W^2.$$

Definition 3.4.2. The Fenchel conjugate of a function $h : W \rightarrow \mathbb{R}$ on a set $A \subseteq W$ is defined by:

$$h_A^* : W^* \rightarrow \mathbb{R} \cup \{+\infty\} : \quad h_A^*(y) = \sup_{a \in A} \langle y, a \rangle - h(a), \quad \text{for all } y \in W^*$$

and the related maximiser correspondence by:

$$Q_A : W^* \rightarrow A : \quad Q_A(y) = \arg \max_{a \in A} \langle y, a \rangle - h(a), \quad \text{for all } y \in W^*.$$

Remark 3.4.1. The corresponding Q_A is not empty if A is weakly closed and h is weakly lower semi-continuous and coercive, i.e. $\lim_{a \rightarrow \infty} h(a)/\|a\|_W = +\infty$.

If W be a Hilbert space (so $W^* = W$) and $h(a) = \frac{1}{2}\|a\|_W^2$ then the correspondence Q_A will be the classical projection on A :

$$Q_A(y) = \arg \max_{a \in A} \langle y, a \rangle_W - \frac{1}{2}\|a\|_W^2 = \arg \max_{a \in A} -\|y - a\|_W^2 = \pi_A(y).$$

Lemma 3.4.1. Let $h : W \rightarrow \mathbb{R}$ be a K -strongly convex function and A a convex subset of W . For any $y_1, y_2 \in W^*$ let $a_i \in Q_A(y_i)$, $i = 1, 2$. Then we have:

$$2K\|a_1 - a_2\|_W^2 \leq \langle y_1 - y_2, a_1 - a_2 \rangle.$$

It implies $\|a_1 - a_2\|_W \leq \frac{1}{2K}\|y_1 - y_2\|_{W^*}$. In particular if $y_1 = y_2$ then $a_1 = a_2$ i.e. the correspondence $Q_A(y)$ is either empty or single valued for every $y \in W^*$.

Proof. Since A is convex, for every $\epsilon \in (0, 1]$ we have $(1 - \epsilon)a_1 + \epsilon a_2 \in A$. By definition:

$$\langle y_1, a_1 \rangle - h(a_1) \geq \langle y_1, (1 - \epsilon)a_1 + \epsilon a_2 \rangle - h((1 - \epsilon)a_1 + \epsilon a_2),$$

and K -strongly convex condition for h gives:

$$h((1 - \epsilon)a_1 + \epsilon a_2) \leq (1 - \epsilon)h(a_1) + \epsilon h(a_2) - K\epsilon(1 - \epsilon)\|a_1 - a_2\|^2.$$

So by combining the above inequalities:

$$\langle y_1, a_1 \rangle - h(a_1) \geq \langle y_1, (1 - \epsilon)a_1 + \epsilon a_2 \rangle - (1 - \epsilon)h(a_1) - \epsilon h(a_2) + K\epsilon(1 - \epsilon)\|a_1 - a_2\|^2,$$

which gives:

$$\epsilon \langle y_1, a_1 - a_2 \rangle \geq \epsilon h(a_1) - \epsilon h(a_2) + K\epsilon(1 - \epsilon)\|a_1 - a_2\|^2.$$

After dividing the both sides by ϵ and then tending $\epsilon \rightarrow 0^+$ we will get:

$$\langle y_1, a_1 - a_2 \rangle \geq h(a_1) - h(a_2) + K\|a_1 - a_2\|^2.$$

By exchanging the role of (a_1, y_1) and (a_2, y_2) we have:

$$\langle y_2, a_2 - a_1 \rangle \geq h(a_2) - h(a_1) + K\|a_2 - a_1\|^2.$$

It yields the desired result if we sum up the two last inequalities. \square

Definition 3.4.3. Let $F : W \rightarrow \mathbb{R}$ be a convex function. We say that $v \in W^*$ is a sub-gradient of F at $a \in W$ if:

$$\forall b \in W : F(b) - F(a) \geq \langle v, b - a \rangle,$$

and set $\partial F(a) \subseteq W^*$ the set of all sub-gradients at a .

One can notice that if $F : W \rightarrow \mathbb{R}$ is differentiable (in sense of Fréchet) at $a \in W$, then $\partial F(a) = \{DF(a)\}$.

3.4.2 OMD algorithm and convergence result

Consider an anonymous game $G = (I, \lambda, V, (A_i)_{i \in I}, J)$. Suppose that the following conditions hold:

- there is a normed vector space $(W, \|\cdot\|_W)$ such that

$$\bigcup_{i \in I} A_i \subseteq W \subseteq V,$$

and let $h : W \rightarrow \mathbb{R}$ be a K -strongly convex function for a real $K > 0$.

- for every $i \in I$ the action sets A_i are weakly closed in W and h is weakly lower semi-continuous and coercive (and hence Q_{A_i} is single valued by Remark 3.4.1),
- for every $(a, \eta) \in W \times \mathcal{P}_G(V)$ the function $J(\cdot, \eta) : W \rightarrow \mathbb{R}$ is convex and exists a subgradient $y(a, \eta) \in \partial_a J(\cdot, \eta) \subseteq W^*$,

Let $\{\beta_n\}_{n \in \mathbb{N}}$ be a sequence of real positive numbers. Set an arbitrary initial measurable functions $\Psi_0 \in \mathcal{A}$, $\eta_0 = \Psi_0 \# \lambda$, $\Phi_0 : I \rightarrow W^*$. The following procedure (3.9) is called the *Online Mirror Descent (OMD)* on anonymous game G :

$$\begin{aligned} (i) \quad \Phi_{n+1}(i) &= \Phi_n(i) - \beta_n y(\Psi_n(i), \eta_n), & \text{for every } i \in I \\ (ii) \quad \Psi_{n+1}(i) &= Q_{A_i}(\Phi_{n+1}(i)), & \text{for every } i \in I \\ (iii) \quad \eta_{n+1} &= \Psi_{n+1} \# \lambda. \end{aligned} \tag{3.9}$$

Theorem 3.4.1. Suppose one applies the OMD algorithm proposed in (3.9) for $\beta_n = \frac{1}{n}$. Suppose the following conditions hold:

1. the game G satisfies assumptions (3.2.1),
2. for every $i \in I$ the action sets A_i are convex and exists $M > 0$ such that for λ -almost every $i \in I$ we have $\|a\|_W \leq M$ for all $a \in A_i$ and we have $R(M) := \sup_{\|a\| \leq M} |h(a)| < +\infty$,
3. the map $\Phi_0 : I \rightarrow W^*$ is bounded,

4. the cost function J is monotone,

5. there exists $\delta > 0$ such that for λ -almost every $i \in I$ and all $a \in A_i, \eta \in \mathcal{P}_G(V)$,

$$\|y(a, \eta)\|_{W^*} \leq \delta. \quad (3.10)$$

Then $\eta_n = \Psi_n \# \lambda$ converges to $\tilde{\eta} = \tilde{\Psi} \# \lambda$ where $\tilde{\eta} \in \mathcal{P}_G(V)$ is the unique Nash equilibrium distribution.

Remark 3.4.2. For every $y, z \in W^*$ and any $A \subseteq W$ we have :

$$\forall a \in Q_A(y) : \quad h_A^*(y) - h_A^*(z) \leq \langle y - z, a \rangle.$$

This is obvious since $h_A^*(y) - \langle y, a \rangle + h(a) = 0 \leq h_A^*(z) - \langle z, a \rangle + h(a)$.

Proof of Theorem 3.4.1. Let $\tilde{\Psi} \in \mathcal{A}$ be a Nash equilibrium profile. Define the real sequence $\{\phi_n\}_{n \in \mathbb{N}}$ as follows:

$$\forall n \in \mathbb{N} : \quad \phi_n = \int_I (h(\tilde{\Psi}(i)) + h_{A_i}^*(\Phi_n(i)) - \langle \Phi_n(i), \tilde{\Psi}(i) \rangle) \, d\lambda(i).$$

By definition of Fenchel conjugate we know $\phi_n \geq 0$. For making the rest of argument well-defined, we first show that ϕ_n is indeed finite. We have

$$\int_I (h(\tilde{\Psi}(i)) + h_{A_i}^*(\Phi_n(i)) - \langle \Phi_n(i), \tilde{\Psi}(i) \rangle) \, d\lambda(i) = \int_I (h(\tilde{\Psi}(i)) - h(\Psi_n(i)) - \langle \Phi_n(i), \tilde{\Psi}(i) - \Psi_n(i) \rangle) \, d\lambda(i)$$

since $\Psi_n(i) = Q_{A_i}(\Phi_n(i))$ for λ -almost every $i \in I$. Moreover,

$$|h(\tilde{\Psi}(i)) - h(\Psi_n(i)) - \langle \Phi_n(i), \tilde{\Psi}(i) - \Psi_n(i) \rangle| \leq 2R(M) + 2\|\Phi_n(i)\|_{W^*}M,$$

since $\|\tilde{\Psi}(i)\|_W, \|\Psi_n(i)\|_W \leq M$ for λ -almost every $i \in I$. By (3.9)(i) we have:

$$\forall n \in \mathbb{N} : \quad \|\Phi_n\|_\infty \leq \delta(1 + \frac{1}{2} + \dots + \frac{1}{n-1}) + \|\Phi_0\|_\infty$$

which yields $|\phi_n| < \infty$. Let us compute the difference $\phi_{n+1} - \phi_n$:

$$\phi_{n+1} - \phi_n = \int_I (h_{A_i}^*(\Phi_{n+1}(i)) - h_{A_i}^*(\Phi_n(i)) - \langle \Phi_{n+1}(i) - \Phi_n(i), \tilde{\Psi}(i) \rangle) \, d\lambda(i)$$

So from Remark 3.4.2:

$$\begin{aligned} \phi_{n+1} - \phi_n &\leq \int_I \langle \Phi_{n+1}(i) - \Phi_n(i), \Psi_{n+1}(i) - \tilde{\Psi}(i) \rangle \, d\lambda(i) \\ &= -\beta_n \int_I \langle y(\Psi_n(i), \eta_n), \Psi_{n+1}(i) - \tilde{\Psi}(i) \rangle \, d\lambda(i) \\ &= -\beta_n \int_I (\langle y(\Psi_n(i), \eta_n), \Psi_n(i) - \tilde{\Psi}(i) \rangle + \langle y(\Psi_n(i), \eta_n), \Psi_{n+1}(i) - \Psi_n(i) \rangle) \, d\lambda(i) \\ &\leq -\beta_n \alpha_n + C\beta_n^2 \end{aligned}$$

where $\alpha_n = \int_I \langle y(\Psi_n(i), \eta_n), \Psi_n(i) - \tilde{\Psi}(i) \rangle \, d\lambda(i)$ and since by condition (3.10) we have:

$$\begin{aligned} |\langle y(\Psi_n(i), \eta_n), \Psi_{n+1}(i) - \Psi_n(i) \rangle| &\leq \delta \|\Psi_{n+1}(i) - \Psi_n(i)\|_W \\ &\leq \frac{\delta}{2K} \|\Phi_{n+1}(i) - \Phi_n(i)\|_{W^*} = \beta_n \frac{\delta}{2K} \|y(\Psi_n(i), \eta_n)\|_{W^*} \leq \beta_n \frac{\delta^2}{2K}. \end{aligned}$$

By definition of the sub-gradient we have:

$$\forall b \in W : \quad \langle y(a, \eta_n), a - b \rangle \geq J(a, \eta_n) - J(b, \eta_n).$$

So:

$$\begin{aligned}\alpha_n &= \int_I \langle y(\Psi_n(i), \eta_n), \Psi_n(i) - \tilde{\Psi}(i) \rangle d\lambda(i) \geq \int_I (J(\Psi_n(i), \eta_n) - J(\tilde{\Psi}(i), \eta_n)) d\lambda(i) = \\ &= \int_X J(a, \eta_n) d(\eta_n - \tilde{\eta})(a) \geq \int_X J(a, \tilde{\eta}) d(\eta_n - \tilde{\eta})(a) \geq 0,\end{aligned}$$

by Remark 3.2.1. If we set $\psi_n = \int_X J(a, \tilde{\eta}) d(\eta_n - \tilde{\eta})(a)$, then since $\beta_n = \frac{1}{n}$ we have:

$$\sum_{n=1}^N \frac{\psi_n}{n} \leq \sum_{n=1}^N \frac{\alpha_n}{n} = \sum_{n=1}^N \beta_n \alpha_n \leq \sum_{n=1}^N \left(\phi_n - \phi_{n+1} + \frac{C}{n^2} \right) = \phi_1 - \phi_{N+1} + \sum_{n=1}^N \frac{C}{n^2} < +\infty \quad (3.11)$$

so $\sum_{n=1}^{\infty} \frac{\psi_n}{n} < +\infty$. We show then that $|\psi_{n+1} - \psi_n| = O(1/n)$. We remind from Lemma 2.2.5 that this yields $\lim_{n \rightarrow \infty} \psi_n = 0$. We have

$$\psi_{n+1} - \psi_n = \int_X J(a, \tilde{\eta}) d(\eta_{n+1} - \eta_n)(a) = \int_X (J(\Psi_{n+1}(i), \tilde{\eta}) - J(\Psi_n(i), \tilde{\eta})) d\lambda(i),$$

and from the definition of sub-gradient:

$$\langle y(\Psi_n(i), \tilde{\eta}), \Psi_{n+1}(i) - \Psi_n(i) \rangle \leq J(\Psi_{n+1}(i), \tilde{\eta}) - J(\Psi_n(i), \tilde{\eta}) \leq \langle y(\Psi_{n+1}(i), \tilde{\eta}), \Psi_{n+1}(i) - \Psi_n(i) \rangle$$

so $|J(\Psi_{n+1}(i), \tilde{\eta}) - J(\Psi_n(i), \tilde{\eta})| = O(1/n)$ which gives $|\psi_{n+1} - \psi_n| = O(1/n)$.

Since $\mathcal{P}_G(V)$ is pre-compact, there exist a sequence $\{n_i\}_{i \in \mathbb{N}} \subseteq \mathbb{N}$ and $\eta' \in \overline{\mathcal{P}_G(V)}$ such that $\lim_{i \rightarrow \infty} \eta_{n_i} = \eta'$. Since $J(\cdot, \tilde{\eta}) : V \rightarrow \mathbb{R}$ is lower semi-continuous, we have:

$$\int_V J(a, \tilde{\eta}) d(\eta' - \tilde{\eta})(a) \leq \liminf_i \int_V J(a, \tilde{\eta}) d(\eta_{n_i} - \tilde{\eta}) = \liminf_i \psi_{n_i} = 0,$$

which yields $\eta' = \tilde{\eta}$ due to the Corollary 3.6.1 and the definition of Nash equilibrium distribution. So every accumulation point of set $\{\eta_n\}_{n \in \mathbb{N}} \subseteq \mathcal{P}_G(V)$ is $\tilde{\eta}$ which gives $\lim_{n \rightarrow \infty} \eta_n = \tilde{\eta}$ since $\mathcal{P}_G(V)$ is pre-compact. \square

3.5 Application to first order MFG

3.5.1 Model

Let us show the first-order mean field games are special case of non atomic anonymous games proposed in section 3.2. Set $I = \mathbb{R}^d$ with the usual topology, as the set of players and $m_0 \in \mathcal{P}(I)$ a given non atomic Borel probability measure on \mathbb{R}^d . Let $V = \mathcal{C}^0([0, T], \mathbb{R}^d)$ endowed with the supremum norm $\|\gamma\|_\infty = \sup_{t \in [0, T]} \|\gamma(t)\|$. For each player $i \in \mathbb{R}^d$ let $A_i = S_{i, M} \subseteq \mathcal{C}^0([0, T], \mathbb{R}^d)$ where:

$$\forall x \in \mathbb{R}^d, M > 0 : \quad S_{x, M} := \{\gamma \in \mathcal{AC}([0, T], \mathbb{R}^d) \mid \gamma(0) = x, \int_0^T \|\dot{\gamma}(t)\|^2 dt \leq M\}, \quad (3.12)$$

where $\mathcal{AC}([0, T], \mathbb{R}^d)$ denotes the set absolutely continuous function from $[0, T]$ to \mathbb{R}^d . We will explain later how to choose $M > 0$ properly.

Let $H^1([0, T], \mathbb{R}^d)$ defined as

$$H^1([0, T], \mathbb{R}^d) = \left\{ \gamma \in \mathcal{AC}([0, T], \mathbb{R}^d) \mid \int_0^T \|\dot{\gamma}(t)\|^2 dt < +\infty \right\}.$$

We denote $\mathcal{P}_1(\mathcal{C}^0([0, T], \mathbb{R}^d))$ be the set of Borel probability measures with finite first moment on $\mathcal{C}^0([0, T], \mathbb{R}^d)$. Set for every $t \in [0, T]$ the evaluation function $e_t : \mathcal{C}^0([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^d$ as $e_t(\gamma) = \gamma(t)$. The MFG cost function $J : \mathcal{C}^0([0, T], \mathbb{R}^d) \times \mathcal{P}_1(\mathcal{C}^0([0, T], \mathbb{R}^d)) \rightarrow \mathbb{R}$ is defined as follows:

$$J(\gamma, \eta) = \begin{cases} \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), e_t \# \eta)) dt + g(\gamma(T), e_T \# \eta), & \text{if } \gamma \in H^1([0, T], \mathbb{R}^d) \\ +\infty & \text{otherwise,} \end{cases}$$

We call the anonymous game $G = (\mathbb{R}^d, m_0, \mathcal{C}^0([0, T], \mathbb{R}^d), (S_{i,M})_{i \in \mathbb{R}^d}, J)$ a *first-order mean field game*.

Remark 3.5.1. For every admissible profile of actions $\Psi : I \rightarrow \mathcal{C}^0([0, T], \mathbb{R}^d)$ that $\Psi(i) \in S_{i,M}$, for $\eta = \Psi \# m_0$ we have:

$$d_1(e_t \# \eta, e_s \# \eta) \leq \int_{\Gamma} \|\gamma(t) - \gamma(s)\| d\eta(\gamma) \leq \sqrt{|t-s|} \int_{\Gamma} \sqrt{\int_t^s \|\dot{\gamma}(r)\|^2 dr} d\eta(\gamma) \leq \sqrt{M|t-s|},$$

due to definition of M in (3.12). That means for every $\eta \in \mathcal{P}_G(V)$ the map $t \rightarrow e_t \# \eta$ is $\frac{1}{2}$ -Holder continuous.

Suppose that the following conditions hold for the data:

Assumption 3.5.1. Let

1. m_0 has a compact support,
2. for every $x \in \mathbb{R}^d$ the map $L(x, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is twice differentiable and there exists $C > 0$ such that for all $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$ we have:

$$\frac{1}{C} I_d \leq D_{vv} L(x, \cdot) \leq C I_d, \quad \|L_x(x, v)\| \leq C,$$

3. the functions $f, g : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$ are continuous and for every $m \in \mathcal{P}_1(\mathbb{R}^d)$ the maps $f(\cdot, m), g(\cdot, m) : \mathbb{R}^d \rightarrow \mathbb{R}$ are $\mathcal{C}^1(\mathbb{R}^d; \mathbb{R})$,
4. suppose that there exist $C > 0$ such that:

$$\forall x \in \mathbb{R}^d, m \in \mathcal{P}_1(\mathbb{R}^d) : \quad \|f_x(x, m)\|, \|g_x(x, m)\| \leq C.$$

Remark 3.5.2 ([32], Theorem 7.2.4). If conditions 3.5.1(2,3,4) hold, then there is at least one minimizer of variational problem

$$\min_{\gamma \in \mathcal{AC}([0, T], \mathbb{R}^d), \gamma(0)=x} \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), e_t \# \eta)) dt + g(\gamma(T), e_T \# \eta). \quad (3.13)$$

The minimizer $\gamma : [0, T] \rightarrow \mathbb{R}^d$ belongs to $\mathcal{C}^1([0, T], \mathbb{R}^d)$, $L_v(\gamma(t), \dot{\gamma}(t))$ is absolutely continuous and

$$\frac{d}{dt} L_v(\gamma(t), \dot{\gamma}(t)) = L_x(\gamma(t), \dot{\gamma}(t)) + f_x(\gamma(t), e_t \# \eta), \quad \text{for almost every } t \in [0, T], \quad (3.14)$$

with $\dot{\gamma}(T) = -g_x(\gamma(T), e_T \# \eta)$. In addition there is $M > 0$ such that $\|\dot{\gamma}\|_{\infty} \leq \sqrt{M/T}$ for every solution of (3.14). This is the way we set M in (3.12) as a function of constants of data in 3.5.1(2,3,4).

The following remark asserts that the definition of action sets in (3.12) and conditions in 3.5.1(2,3) imply the assumptions (3.2.1) for first order mean field game.

Remark 3.5.3. If $K \subseteq \mathbb{R}^d$ be compact such that $\text{supp}(m_0) \subseteq K$, then

1. for $x \in K$ we have $\|\gamma\|_{\infty} \leq \max_{y \in K} \|y\| + MT$, for $\gamma \in S_{x,M}$, which gives the condition (3.2),
2. the correspondence $A : I \rightarrow V$, $A(i) = S_{i,M}$ is continuous and by Arzela-Ascoli $S_{i,M}$ is compact for all $i \in I$,
3. the convexity of $L(x, \cdot)$ implies that for every $\eta \in \mathcal{P}(V)$ the function $J(\cdot, \eta) : \mathcal{C}^0([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$ is lower semi-continuous,

4. the function $\text{Min} : I \times \mathcal{P}_G(V) \rightarrow \mathbb{R}$, $\text{Min}(i, \eta) := \min_{a \in A_i} J(a, \eta)$ is continuous,

5. by Arzela-Ascoli the following set:

$$S = S_{K,M} = \{\gamma \in C^0([0, T], \mathbb{R}^d) \mid \gamma(0) \in K, \int_0^T \|\dot{\gamma}(t)\|^2 dt \leq M\}$$

is precompact in V . For every $\eta \in \text{cov}(\mathcal{P}_G(V))$ we have $\text{supp}(\eta) \subseteq S$, so $\text{cov}(\mathcal{P}_G(V))$ is tight and hence it is pre-compact in $(\mathcal{P}_1(V), d_1)$,

6. the minimiser of problem (3.13) is unique as is explained in section 2.4. Hence the unique minimiser condition holds.

Corollary 3.5.1. *The first-order MFGs defined above, satisfies the assumptions (3.2.1) and hence by Theorem 3.2.1 has at least a Nash equilibrium $\tilde{\Psi} \in \mathcal{A}$. If we set $\tilde{\eta} = \tilde{\Psi} \# m_0$ and $e_t \# \tilde{\eta} = \tilde{m}_t$ for all $t \in [0, T]$, then for m_0 -almost every $i \in \mathbb{R}^d$:*

$$\tilde{\Psi}(i) = \text{argmin}_{\gamma \in H^1([0, T], \mathbb{R}^d), \gamma(0)=i} \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), \tilde{m}_t)) dt + g(\gamma(T), \tilde{m}_T).$$

The measure $\tilde{\eta}$ is an equilibrium distribution in sense of (1.34). Under stronger assumptions, by following section 2.4 we can construct the first order MFG system solution (u, m) from the equilibrium distribution $\tilde{\eta}$ as in (2.37).

We prove that the uniqueness of equilibrium is a consequence of the monotonicity of f, g and the unique minimizer condition. This is the counterpart for the uniqueness result in [73].

Lemma 3.5.1. *If $f, g : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$ are monotone, then the MFG cost function will be so.*

Proof. Let $\eta_1, \eta_2 \in \mathcal{P}(V)$. If we define $m_{i,t} = e_t \# \eta_i$ for $i = 1, 2$ and $t \in [0, T]$, we then have:

$$\begin{aligned} & \int_V (J(\gamma, \eta_1) - J(\gamma, \eta_2)) d(\eta_1 - \eta_2)(\gamma) = \\ & \int_V \left(\int_0^T (f(\gamma(t), m_{1,t}) - f(\gamma(t), m_{2,t})) dt + g(\gamma(T), m_{1,T}) - g(\gamma(T), m_{2,T}) \right) d(\eta_1 - \eta_2)(\gamma) = A + B \end{aligned}$$

where

$$\begin{aligned} A &= \int_0^T \left(\int_{\mathbb{R}^d} (f(x, m_{1,t}) - f(x, m_{2,t})) d(m_{1,t} - m_{2,t})(x) \right) dt \geq 0 \\ B &= \int_{\mathbb{R}^d} (g(x, m_{1,T}) - g(x, m_{2,T})) d(m_{1,T} - m_{2,T})(x) \geq 0, \end{aligned}$$

since the couplings f, g are monotone. □

Corollary 3.5.2. *The monotone first order MFG satisfying assumptions 3.5.1 possesses a unique equilibrium.*

3.5.2 Fictitious play in monotone first order MFG

The fictitious play in first-order MFG takes such form: for initial profile of actions

$$\Psi_1 \in \mathcal{A}, \bar{\eta}_1 = \eta_1 = \Psi_1 \# \lambda \in \mathcal{P}(V)$$

the players play as follows at round $n = 1, 2, \dots$:

$$\begin{aligned} (i) \quad \Psi_{n+1}(i) &= \arg \min_{\gamma \in H^1, \gamma(0)=i} \int_0^T (L(\gamma(t), \dot{\gamma}(t)) + f(\gamma(t), e_t \# \bar{\eta}_n)) \, dt + g(\gamma(T), e_T \# \bar{\eta}_n), \\ (ii) \quad \eta_{n+1} &= \Psi_{n+1} \# \lambda, \\ (iii) \quad \bar{\eta}_{n+1} &= \frac{1}{n+1} \sum_{i=1}^{n+1} \eta_i. \end{aligned} \tag{3.15}$$

where (i) holds for m_0 -almost every $i \in \mathbb{R}^d$. Here we apply the convergence result in fictitious play (Section 3) for monotone first-order MFG. We suppose the assumptions 3.5.1 (and hence (3.2.1)) conditions hold.

Lemma 3.5.2. *If $f, g : m \rightarrow f(\cdot, m), g(\cdot, m)$ are Lipschitz from $\mathcal{P}(\mathbb{R}^d)$ to $\mathcal{C}^1(\mathbb{R}^d)$ then there is a constant $C > 0$ such that:*

$$|J(\gamma, \eta) - J(\gamma, \eta') - J(\gamma', \eta) + J(\gamma', \eta')| \leq C \|\gamma - \gamma'\|_\infty d_1(\eta, \eta')$$

$$|J(\gamma, \eta) - J(\gamma, \eta')| \leq C d_1(\eta, \eta')$$

for every $\gamma, \gamma' \in H^1([0, T], \mathbb{R}^d)$ and $\eta, \eta' \in \mathcal{P}(V)$.

Proof. Since $f : m \rightarrow f(\cdot, m)$ is Lipschitz from $\mathcal{P}(\mathbb{R}^d)$ to $\mathcal{C}^1(\mathbb{R}^d)$ there is $C > 0$ such that:

$$\|f(\cdot, m) - f(\cdot, m')\|_{\mathcal{C}^1} \leq C d_1(m, m'), \quad \|g(\cdot, m) - g(\cdot, m')\|_{\mathcal{C}^1} \leq C d_1(m, m')$$

which means that for every $x, x' \in \mathbb{R}^d$ we have

$$|f(x, m) - f(x, m') - f(x', m) + f(x', m')| \leq C \|x - x'\| d_1(m, m'),$$

$$|f(x, m) - f(x, m')| \leq C d_1(m, m').$$

Similar inequalities hold with respect to g . We have:

$$\begin{aligned} & |J(\gamma, \eta) - J(\gamma, \eta') - J(\gamma', \eta) + J(\gamma', \eta')| \\ & \leq \int_0^T |f(\gamma(t), e_t \# \eta) - f(\gamma(t), e_t \# \eta') + f(\gamma'(t), e_t \# \eta) - f(\gamma'(t), e_t \# \eta')| \, dt \\ & \quad + |g(\gamma(T), e_T \# \eta) - g(\gamma(T), e_T \# \eta') - g(\gamma'(T), e_T \# \eta) + g(\gamma'(T), e_T \# \eta')| \\ & \leq C \int_0^T \|\gamma(t) - \gamma'(t)\| d_1(e_t \# \eta, e_t \# \eta') \, dt + \|\gamma(T) - \gamma'(T)\| d_1(e_T \# \eta, e_T \# \eta') \\ & \leq C \int_0^T \|\gamma - \gamma'\|_\infty d_1(\eta, \eta') \, dt + \|\gamma - \gamma'\|_\infty d_1(\eta, \eta') = (CT + 1) \|\gamma - \gamma'\|_\infty d_1(\eta, \eta'), \end{aligned}$$

and

$$\begin{aligned} |J(\gamma, \eta) - J(\gamma, \eta')| & \leq \int_0^T |f(\gamma(t), e_t \# \eta) - f(\gamma(t), e_t \# \eta')| \, dt + |g(\gamma(T), e_T \# \eta) - g(\gamma(T), e_T \# \eta')| \\ & \leq C \int_0^T d_1(e_t \# \eta, e_t \# \eta') \, dt + d_1(e_T \# \eta, e_T \# \eta') \leq (CT + 1) d_1(\eta, \eta'). \end{aligned}$$

□

Corollary 3.5.3. *If $f, g : m \rightarrow f(\cdot, m), g(\cdot, m)$ are Lipschitz, then by Lemma 3.5.2, the convergence result of fictitious play (Theorem 3.3.1) holds for the first-order monotone MFG.*

3.5.3 Online mirror descent in monotone first order MFG

Here we use the convergence result proved in section 4 for the first-order MFG with a monotone convex cost function J . Let us suppose that the couplings f, g are monotone and $L(\cdot, \cdot), f(\cdot, m), g(\cdot, m)$ are convex for every $m \in \mathcal{P}(\mathbb{R}^d)$. It easily yields that J is monotone (by Lemma 3.5.1) and for every $\eta \in \mathcal{P}(V)$, the function $J(\cdot, \eta) : H^1([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$ is convex.

Remark 3.5.4. *We propose an example of data L, f, g such that they are convex in x, v inputs. Before we start the precise definition, let us point out that we can relax the condition 3.5.1(4) and replace it with the following assumption. Suppose that there exists $M > 0$ in (3.12) such that for all solution γ of Euler-Lagrange equation (3.14) with $\eta \in \mathcal{P}_G(V)$, we have $\|\dot{\gamma}\|_\infty \leq \sqrt{M/T}$. This assumption with the conditions (3.5.1)(1,2,3) give the existence of equilibrium as in Corollary 3.5.1.*

For the example, set

$$L(x, v) = \frac{1}{2}\|v\|^2, \quad f(x, m) = \alpha\langle x, \mathbb{E}_m z \rangle, \quad g(x, m) = \beta\langle x, \mathbb{E}_m z \rangle,$$

for some $\alpha, \beta > 0$ where $\mathbb{E}_m z = \int_{\mathbb{R}^d} z \, dm(z)$. Set $\alpha, \beta, R > 0$ with

$$\alpha T^2 R + \beta T R + \sup_{x \in \text{supp}(m_0)} \|x\| \leq R.$$

and the constant $M > 0$ in (3.12) with $M = T(\alpha T R + \beta R)^2$. For every $\eta \in \mathcal{P}_G(V)$ we have

$$\text{for } \eta\text{-almost every } \gamma: \quad \|\gamma(t)\| \leq \sup_{x \in \text{supp}(m_0)} \|x\| + \sqrt{MT} \leq \sup_{x \in \text{supp}(m_0)} \|x\| + \alpha T^2 R + \beta T R \leq R.$$

Hence for every $\eta \in \mathcal{P}_G(V)$:

$$\sup_{t \in [0, T], x \in \text{supp}(e_t \# \eta)} \|x\| \leq R. \quad (3.16)$$

The Euler Lagrange equation (3.14) in this example read as

$$\frac{d}{dt} \dot{\gamma}(t) = \alpha \mathbb{E}_{e_t \# \eta} z, \quad \text{for almost every } t \in [0, T], \quad (3.17)$$

and $\dot{\gamma}(T) = \beta \mathbb{E}_{e_T \# \eta} z$, $\gamma(0) \in \text{supp}(m_0)$. That yields

$$\sup_{t \in [0, T]} \|\dot{\gamma}(t)\| \leq \alpha T R + \beta R, \quad \sup_{t \in [0, T]} \|\gamma(t)\| \leq R, \quad (3.18)$$

for all $\eta \in \mathcal{P}_G(V)$ since (3.16) holds. That means for every $\eta \in \mathcal{P}_G(V)$ the optimal trajectories γ satisfies (3.18) and hence $\|\dot{\gamma}\|_\infty \leq \sqrt{M/T}$.

Let us set $W = H^1([0, T], \mathbb{R}^d)$ endowed with inner product:

$$\forall \gamma_1, \gamma_2 \in W : \quad \langle \gamma_1, \gamma_2 \rangle_W = \langle \gamma_1(0), \gamma_2(0) \rangle_{\mathbb{R}^d} + \int_0^T \langle \dot{\gamma}_1(t), \dot{\gamma}_2(t) \rangle_{\mathbb{R}^d} dt.$$

We clearly have

$$\bigcup_{i \in I} A_i \subseteq W \subseteq V,$$

and A_i are uniformly bounded in W for m_0 -almost every $i \in I$. For integrable functions $F, D \in L^2([0, T], \mathbb{R})$ and $G \in \mathbb{R}$ we define $y = [[F, D, G]] \in W^*$ by:

$$\langle y, \gamma \rangle = \int_0^T (F(t) \cdot \gamma(t) + D(t) \cdot \dot{\gamma}(t)) dt + G \cdot \gamma(T), \quad \text{for every } \gamma \in W$$

After a few computation we have:

$$\langle y, \gamma \rangle = \int_0^T \left(\int_t^T F(s) \, ds + D(t) + G \right) \cdot \dot{\gamma}(t) \, dt + \left(\int_0^T F(s) \, ds + G \right) \cdot \gamma_0.$$

We can find $\gamma_y \in W$ as a representation of $y \in W^*$ i.e. for all $\gamma \in H^1([0, T], \mathbb{R}^d)$ we have $\langle y, \gamma \rangle = \langle \gamma_y, \gamma \rangle_W$. The representation γ_y corresponding to y should solve

$$\gamma_y(0) = \int_0^T F(s) \, ds + G, \quad \frac{d}{dt}(\gamma_y)(t) = \int_t^T F(s) \, ds + D(t) + G. \quad (3.19)$$

or

$$\gamma_y(t) = \int_0^T F(s) \min(t, s) \, ds + \int_0^t D(s) \, ds + (t+1)G + \int_0^T F(s) \, ds. \quad (3.20)$$

By assumptions 3.5.1(2,3) and using dominated Lebesgue convergence theorem, we can conclude that the function $J(\cdot, \eta) : W \rightarrow \mathbb{R}$ is differentiable for every $\eta \in \mathcal{P}(V)$. So the sub-differential set is singleton $\partial J(\cdot, \eta)(\gamma) = \{D_\gamma J(\gamma, \eta)\} \subseteq W^*$ and the derivative is calculated by:

$$\begin{aligned} \forall z \in W : \quad \langle D_\gamma J(\gamma, \eta), z \rangle &= \lim_{\epsilon \rightarrow 0} \frac{J(\gamma + \epsilon z, \eta) - J(\gamma, \eta)}{\epsilon} \\ &= \int_0^T (L_x(\gamma_t, \dot{\gamma}(t)) \cdot z_t + L_v(\gamma_t, \dot{\gamma}(t)) \cdot \dot{z}_t + f_x(\gamma(t), e_t \# \eta) \cdot z_t) \, dt + g_x(\gamma(T), e_T \# \eta) \cdot z_T \end{aligned}$$

or according to our representation:

$$D_\gamma J(\gamma, \eta) = [[L_x(\gamma(\cdot), \dot{\gamma}(\cdot)) + f_x(\gamma(\cdot), e_{(\cdot)} \# \eta), L_v(\gamma(\cdot), \dot{\gamma}(\cdot)), g_x(\gamma(T), e_T \# \eta)]].$$

So by the computation in (3.20) the gradient $\nabla_\gamma J(\gamma, \eta) \in W$ is obtained as follows:

$$\begin{aligned} \nabla_\gamma J(\gamma, \eta)(t) &= \int_0^T (L_x(\gamma_s, \dot{\gamma}_s) + f_x(\gamma_s, e_s \# \eta)) \min(t, s) \, ds + \int_0^t L_v(\gamma_s, \dot{\gamma}_s) \, ds \\ &\quad + (t+1)g_x(\gamma(T), e_T \# \eta) + \int_0^T (L_x(\gamma_s, \dot{\gamma}_s) + f_x(\gamma_s, e_s \# \eta)) \, ds. \end{aligned} \quad (3.21)$$

Theorem 3.5.1. *Suppose a first-order MFG satisfies the assumptions 3.5.1. If the cost function J is monotone and convex w.r.t. first argument, then the online mirror descent algorithm proposed in (3.9) for $h : W \rightarrow \mathbb{R}$, $h(\gamma) = \frac{1}{2} \|\gamma\|_{H^1}^2$ and $\beta_n = \frac{1}{n}$ ($n \in \mathbb{N}$), converges to the unique first-order mean field game equilibrium.*

Proof. The function $h : W \rightarrow \mathbb{R}$, $h(\gamma) = \frac{1}{2} \|\gamma\|_{H^1}^2$ is $\frac{1}{2}$ -strongly convex function and lower semi-continuous for the weak topology, so the mirror projection Q_{A_i} will have singleton values.

The game satisfies the assumptions (3.2.1). Since the assumptions 3.5.1 hold, there is $C' > 0$ such that:

$$\forall \gamma \in H^1, \eta \in \mathcal{P}_G(V) : \quad \|D_\gamma J(\gamma, \eta)\|_{W^*} \leq C'(\|\dot{\gamma}\|_{L^2} + 1).$$

So all of the conditions in Theorem 3.4.1 are satisfied and the desired convergence result holds. \square

Remark 3.5.5. *Since the space $H^1([0, T], \mathbb{R}^d)$ is Hilbert, we identify it by its dual space. Hence by choice $h(\gamma) = \frac{1}{2} \|\gamma\|_{H^1}^2$ we have:*

$$Q_{A_i}(\gamma) = \pi_{A_i}(\gamma) = \frac{\min(\|\dot{\gamma}\|_{L^2}, \sqrt{M})}{\|\dot{\gamma}\|_{L^2}}(\gamma - \gamma_0) + i.$$

by the choice of A_i . Then, the OMD algorithm have such form

$$\begin{aligned}
(i) \quad \Phi_{n+1}(i) &= \Phi_n(i) - \frac{1}{n} \nabla J(\Psi_n(i), \eta_n), & \text{for every } i \in I \\
(ii) \quad \Psi_{n+1}(i) &= \frac{\min(\|\Phi_{n+1}(i)\|_{L^2}, \sqrt{M})}{\|\Phi_{n+1}(i)\|_{L^2}} (\Phi_{n+1}(i) - \Phi_{n+1}(i)_0) + i, & \text{for every } i \in I \\
(iii) \quad \eta_{n+1} &= \Psi_{n+1} \# \lambda.
\end{aligned} \tag{3.22}$$

or in explicit way it takes the following form: let $\hat{\gamma}_{0,x} = 0$ for every $x \in \mathbb{R}^d$ and:

$$\begin{aligned}
\hat{\gamma}_{n+1,x}(t) &= \hat{\gamma}_{n,x}(t) - \frac{1}{n} \int_0^T (L_x(\gamma_{n,x}(s), \dot{\gamma}_{n,x}(s)) + f_x(\gamma_{n,x}(s), e_s \# \eta_n)) \min(t, s) \, ds \\
&\quad - \frac{1}{n} \int_0^t L_v(\gamma_{n,x}(s), \dot{\gamma}_{n,x}(s)) \, ds - \frac{t}{n} g_x(\gamma_{n,x}(T), e_T \# \eta_n), \\
\gamma_{n+1,x} &= c_{n+1} \hat{\gamma}_{n+1,x} + x, \quad c_{n+1} = \frac{\min(\|\hat{\gamma}_{n+1,x}\|_{L^2}, \sqrt{M})}{\|\hat{\gamma}_{n+1,x}\|_{L^2}}, \\
\eta_{n+1} &= \gamma_{n+1, \cdot} \# \lambda.
\end{aligned} \tag{3.23}$$

3.6 Appendix

Here we extend the disintegration Theorem 5.3.1 in [8], and demonstrate its modification that is used in the precedent proofs. Suppose I a Polish space and V a metric space. Let $A : I \rightarrow V$ be a correspondence with $A(i) = A_i$. For a Borel probability measure $\lambda \in \mathcal{P}(I)$ we say $\eta \in \mathcal{P}(V)$ *disintegrates* with respect to $(A_i)_{i \in I}$ if there are $\{\eta^i\}_{i \in I} \subset \mathcal{P}(V)$ such that for λ -almost every $i \in I$ we have $\text{supp}(\eta^i) \subseteq A_i$, and for every bounded measurable $f : V \rightarrow \mathbb{R}$:

$$\int_V f(a) \, d\eta(a) = \int_I \int_V f(a) \, d\eta^i(a) \, d\lambda(i).$$

Theorem 3.6.1. *Suppose $A : I \rightarrow V$ be upper semi continuous. Let $\{\eta_n\}_{n \in \mathbb{N}} \subseteq \mathcal{P}_1(V)$ with $\eta_n \rightarrow \eta$ in weak sense. If for every $n \in \mathbb{N}$, η_n disintegrates with respect to $(A_i)_{i \in I}$ then the same holds true for η .*

Proof. For every $n \in \mathbb{N}$, define $m_n \in \mathcal{P}(I \times V)$ as follows: for every bounded measurable $f : I \times V \rightarrow \mathbb{R}$ let:

$$\int_{I \times V} f(i, a) \, dm_n(i, a) = \int_I \int_V f(i, a) \, d\eta_n^i(a) \, d\lambda(i).$$

Obviously $\pi_I \# m_n = \lambda, \pi_V \# m_n = \eta_n$ where π_I, π_V are respectively projections of $I \times V$ on I, V . Since $\{\eta_n\}$ are tight and I is a Polish space, for every $\epsilon > 0$, there is a compact set $I_\epsilon \subseteq I, K_\epsilon \subseteq V$ such that $\lambda(I \setminus I_\epsilon), \eta_n(V \setminus K_\epsilon) < \epsilon$ for all $n \in \mathbb{N}$. In addition

$$\begin{aligned}
m_n(I_\epsilon \times K_\epsilon) &\geq 1 - m_n(I \times V \setminus K_\epsilon) - m_n(I \setminus I_\epsilon \times V) \\
&= 1 - \eta_n(V \setminus K_\epsilon) - \lambda(I \setminus I_\epsilon) \geq 1 - 2\epsilon,
\end{aligned}$$

which means the set $\{m_n\}_{n \in \mathbb{N}}$ is tight too. Hence there exists $m \in \mathcal{P}(I \times V)$ and a subsequence $\{m_{n_k}\}_{k \in \mathbb{N}}$ such that $m_{n_k} \rightarrow m$. We directly have $\eta_{n_k} = \pi_V \# m_{n_k} \rightarrow \pi_V \# m$ which means $\pi_V \# m = \eta$. On the other hand, due to the disintegration theorem (see [8] Theorem 5.3.1) there are $m^i \in \mathcal{P}(V)$ for every $i \in I$, such that for every bounded measurable $f : I \times V \rightarrow \mathbb{R}$:

$$\int_{I \times V} f(i, a) \, dm(i, a) = \int_I \int_V f(i, a) \, dm^i(a) \, d\lambda(i).$$

So since the second marginal of m is η , we can write: for every bounded measurable $f : I \times V \rightarrow \mathbb{R}$:

$$\int_V f(a) \, d\eta(a) = \int_I \int_V f(a) \, dm^i(a) \, d\lambda(i).$$

So what is left is to show that for λ -almost every $i \in I$ we have $\text{supp}(m^i) \subseteq A_i$. Set $f : I \times V \rightarrow \mathbb{R}$ as $f(i, a) = 1_{a \in A_i}$. We know the function f is upper semi continuous since the correspondence $A : I \rightarrow V$, $A(i) = A_i$ is upper semi continuous. For every $n \in \mathbb{N}$ we have:

$$\int_{I \times V} f(i, a) \, dm_n(i, a) = \int_I \int_V f(i, a) \, d\eta_n^i(a) \, d\lambda(i) = 1.$$

Hence

$$1 = \limsup_k \int_{I \times V} f(i, a) \, dm_{n_k}(i, a) \leq \int_{I \times V} f(i, a) \, dm(i, a) \leq 1,$$

so $\int_{I \times V} f(i, a) \, dm(i, a) = 1$ which is equivalent to say for λ -almost every $i \in I$ we have $\text{supp}(m^i) \subseteq A_i$. \square

Corollary 3.6.1. *Every element $\eta \in \overline{\text{cov}(\mathcal{P}_G(V))}$ disintegrates with respect to $(A_i)_{i \in I}, \lambda \in \mathcal{P}(I)$.*

Proof. Let $\mathcal{S} \subset \mathcal{P}(V)$ be the set of all measures which disintegrates with respect to $(A_i)_{i \in I}$. Clearly \mathcal{S} is convex and due to Theorem 3.6.1 it is closed. Also, we have $\mathcal{P}_G(V) \subseteq \mathcal{S}$ since for all $\Psi \in \mathcal{A}$ we have for every bounded measurable $f : I \times V \rightarrow \mathbb{R}$:

$$\int_V f(a) \, d(\Psi \# \lambda)(a) = \int_I \int_V f(a) \, d\delta_{\Psi(i)}(a) \, d\lambda(i),$$

hence it gives $\overline{\text{cov}(\mathcal{P}_G(V))} \subseteq \mathcal{S}$. \square

Chapter 4

Finite MFG: fictitious play and convergence to classical MFG

Joint work with Francisco José Silva

4.1 Introduction

Mean Field Games (MFGs) were introduced by Lasry and Lions in [72, 73, 74] and, independently, by Huang, Caines and Malhamé in [67]. One of the main purposes of the theory is to develop a notion of Nash equilibria for dynamic games, which can be deterministic or stochastic, with an infinite number of players. More precisely, if we consider a N -player game and we assume that the players are indistinguishable and small, in the sense that a change of strategy of player j has a small impact on the cost for player i , then, under some assumptions, it is possible to show that as $N \rightarrow \infty$ the sequence of equilibria admits limit points (see [37]). The latter correspond to probability measures on the set of actions and define the notion of equilibria with a continuum of agents. An interesting feature of the theory is that it allows to obtain important qualitative information on the equilibria and the resulting problem is amenable to numerical computation. We refer the reader to the lessons by P.-L. Lions [75] and to [33, 59, 57, 56] for surveys on the theory and its applications.

Most of the literature about MFGs deals with games in continuous time and where the agents are distributed on a continuum of states (see [33]). In this article we consider a MFG problem where the number of states and times are finite. For the sake of simplicity, we will call *finite MFGs* the games of this type. This framework has been introduced by Gomes, Mohr and Souza in [55], where the authors prove results related to the existence and uniqueness of equilibria, as well as the convergence to a stationary equilibrium as time goes to infinity.

Our contribution to these type of games is twofold. First, we consider the *fictitious play* procedure, which is a learning process introduced by Brown in [26]. We refer the reader to [53, Chapter 2] and the references therein for a survey on this subject. Loosely speaking, the procedure is that, at each iteration, a typical player implements a best response strategy to his *belief* on the action of the remaining players. The belief at iteration $n \in \mathbb{N}$ is given, by definition, by the average of outputs of decisions of the remaining players in the previous iterations $1, \dots, n-1$. In the context of continuous MFGs, the study of the convergence of such procedure to an equilibrium has been first addressed in [39], for a particular class of MFGs called *potential MFGs*. This analysis has then been extended in chapter 3, by assuming that the MFG is monotone, which means that agents have aversion to imitate the strategies of other

players. Under an analogous monotonicity assumption, we prove in Theorem 4.4 that the fictitious play procedure converges also in the case of finite MFGs. As pointed out in [39], the convergence of a learning procedure in MFGs theory is interesting, because, in practice, it is related with the formation of the equilibria.

Our second contribution concerns the relation between continuous and finite MFGs. We consider here a first order continuous MFG and we associate to it a family of finite MFGs defined on finite space/time grids. By applying the results in [55], we know that for any fixed space/time grid the associated finite MFG admits at least one solution. Moreover, any such solution induces a probability measure on the space of strategies. Letting the grid length tend to zero, we prove that the aforementioned sequence of probability measures is precompact and, hence, has at least one limit point. The main result of this article is given in Theorem 4.4.1 and states that any such limit point is an equilibrium of the continuous MFG problem. To the best of our knowledge, this is the first result relating the equilibria for continuous MFGs, introduced in [74], with the equilibria for finite MFGs, introduced in [55]. Let us point out that, contrary to [41], where the authors propose a discretization of a first order continuous MFG, the approximation result in Theorem 4.4.1 has no practical application. Indeed, for fine space/time grids the numerical computation of finite MFG equilibrium is very costly because of the very large number of unknowns involved in the problem. Thus, we insist that our main result in Theorem 4.4.1 has, for the time being, only a theoretical importance since it relate two interesting MFG models.

The article is organized as follows. In Section 4.2 we recall the finite MFG introduced in [55] and we state our first assumption that ensures the existence of at least one equilibrium. In Section 4.3 we describe the fictitious play procedure for the finite MFG and prove its convergence under a monotonicity assumption on the data. In Section 4.4 we introduce the first order continuous MFG under study, as well as the corresponding space/time discretization and the associated finite MFGs. As the length of the space/time grid tends to zero, we prove several asymptotic properties of the finite MFGs equilibria and we also prove our main result showing their convergence to a solution of the continuous MFG problem.

4.2 The finite state and discrete time Mean Field Game problem

We begin this section by presenting the MFG problem introduced in [55] with finite state and discrete time. Let \mathcal{S} be a finite set, and, given $T > 0$, let $\mathcal{T} = \{0, \dots, m\}$. We denote by $|\mathcal{S}|$ the number of elements in \mathcal{S} , and by

$$\mathcal{P}(\mathcal{S}) := \left\{ m : \mathcal{S} \rightarrow [0, 1] \mid \sum_{x \in \mathcal{S}} m(x) = 1 \right\},$$

the simplex in $\mathbb{R}^{|\mathcal{S}|}$, which is identified with the set of probability measures over \mathcal{S} . We define now the notion of transition kernel associated to \mathcal{S} and \mathcal{T} .

Definition 4.2.1. *We denote by $\mathcal{K}_{\mathcal{S}, \mathcal{T}}$ the set of all maps $P : \mathcal{S} \times \mathcal{S} \times (\mathcal{T} \setminus \{m\}) \rightarrow [0, 1]$, called the transition kernels, such that $P(x, \cdot, k) \in \mathcal{P}(\mathcal{S})$ for all $x \in \mathcal{S}$ and $k \in \mathcal{T} \setminus \{m\}$.*

Note that $\mathcal{K}_{\mathcal{S}, \mathcal{T}}$ can be seen as a compact subset of $\mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}| \times m}$. Given an initial distribution $M_0 \in \mathcal{P}(\mathcal{S})$ and $P \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$, the pair (M_0, P) induces a probability distribution over \mathcal{S}^{m+1} , with marginal distributions given by

$$\begin{aligned} M_P^{M_0}(x_k, 0) &:= M_0(x_0), \quad \forall x_0 \in \mathcal{S}, \\ M_P^{M_0}(x_k, k) &:= \sum_{(x_0, x_1, \dots, x_{k-1}) \in \mathcal{S}^k} M_0(x_0) \prod_{k'=0}^{k-1} P(x_{k'}, x_{k'+1}, t_{k'}) \quad \forall k = 1, \dots, m, \quad x_k \in \mathcal{S}, \end{aligned} \tag{4.1}$$

or equivalently, written in a recursively form,

$$\begin{aligned} M_P^{M_0}(x_k, 0) &:= M_0(x_0), \quad \forall x_0 \in \mathcal{S}, \\ M_P^{M_0}(x_k, k) &:= \sum_{x_{k-1} \in \mathcal{S}} M_P^{M_0}(x_{k-1}, k-1) P(x_{k-1}, x_k, k-1) \quad \forall k = 1, \dots, m, \quad x_k \in \mathcal{S}. \end{aligned} \quad (4.2)$$

Now, let $c : \mathcal{S} \times \mathcal{S} \times \mathcal{P}(\mathcal{S}) \times \mathcal{P}(\mathcal{S}) \rightarrow \mathbb{R}$, $g : \mathcal{S} \times \mathcal{P}(\mathcal{S}) \rightarrow \mathbb{R}$, $M : \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$ and define $J_M : \mathcal{K}_{\mathcal{S}, \mathcal{T}} \rightarrow \mathbb{R}$ as

$$J_M(P) := \sum_{k=0}^{m-1} \sum_{x, y \in \mathcal{S}} M_P^{M_0}(x, k) P(x, y, k) c_{xy}(P(x, k), M(k)) + \sum_{x \in \mathcal{S}} M_P^{M_0}(x, m) g(x, M(m)),$$

where, for notational convenience, we have set $c_{xy}(\cdot, \cdot) := c(x, y, \cdot, \cdot)$ and $P(x, k) := P(x, \cdot, k) \in \mathcal{P}(\mathcal{S})$. We consider the following MFG problem: find $\hat{P} \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$ such that

$$\hat{P} \in \operatorname{argmin}_{P \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}} J_M(P) \quad \text{with } M = M_{\hat{P}}^{M_0}. \quad (\text{MFG}_d)$$

In order to rewrite (MFG_d) in a recursive form (as in [55]), given $k = 0, \dots, m-1$, $x \in \mathcal{S}$ and $P \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$, we define a probability distribution in \mathcal{S}^{m-k+1} whose marginals are given by

$$\begin{aligned} M_P^{x,k}(x_k, k) &:= \delta_{x, x_k}, \quad \forall x_k \in \mathcal{S}, \\ M_P^{x,k}(x_{k'}, k') &:= \sum_{x_{k'-1} \in \mathcal{S}} M_P^{x,k}(x_{k'-1}, k'-1) P(x_{k'-1}, x_{k'}, k'-1) \quad \forall k' = k+1, \dots, m, \quad x_{k'} \in \mathcal{S}, \end{aligned}$$

where $\delta_{x, x_k} := 1$ if $x = x_k$ and $\delta_{x, x_k} := 0$, otherwise. Given $M : \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$, we also set

$$J_M^{x,k}(P) := \sum_{k'=k}^{m-1} \sum_{x, y \in \mathcal{S}} M_P^{x,k}(x_{k'}, k') P(x, y, k') c_{xy}(P(x, k'), M(k')) + \sum_{x \in \mathcal{S}} M_P^{x,k}(x, m) g(x, M(m)).$$

Since for every $M : \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$ the function

$$U_M(x, k) := \inf_{P \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}} J_M^{x,k}(P) \quad \forall k = 0, \dots, m-1, \quad x \in \mathcal{S}, \quad U_M(x, m) := g(x, M(m)), \quad \forall x \in \mathcal{S},$$

satisfies the Dynamic Programming Principle (DPP),

$$U_M(x, k) = \inf_{p \in \mathcal{P}(\mathcal{S})} \sum_{y \in \mathcal{S}} p(y) [c_{xy}(p, M(k)) + U_M(y, k+1)], \quad \forall k = 0, \dots, m-1, \quad x \in \mathcal{S}, \quad (4.3)$$

problem (MFG_d) is equivalent to find $U : \mathcal{S} \times \mathcal{T} \rightarrow \mathbb{R}$ and $M : \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$ such that

$$\begin{aligned} \text{(i)} \quad U(x, k) &= \sum_{y \in \mathcal{S}} \hat{P}(x, y, k) [c_{xy}(\hat{P}(x, k), M(k)) + U(y, k+1)], \quad \forall k = 0, \dots, m-1, \quad x \in \mathcal{S}, \\ \text{(ii)} \quad M(x, k) &= \sum_{y \in \mathcal{S}} M(y, k-1) \hat{P}(y, x, k-1), \quad \forall k = 1, \dots, m, \quad x \in \mathcal{S}, \\ \text{(iii)} \quad U(x, m) &= g(x, m), \quad M(x, 0) = M_0(x) \quad \forall x \in \mathcal{S}, \end{aligned} \quad (4.4)$$

where $\hat{P} \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$ satisfies

$$\hat{P}(x, \cdot, k) \in \operatorname{argmin}_{p \in \mathcal{P}(\mathcal{S})} \sum_{y \in \mathcal{S}} p(y) [c_{xy}(p, M(k)) + U(y, k+1)], \quad \forall k = 0, \dots, m-1, \quad x \in \mathcal{S}, \quad (4.5)$$

and in the argument of c_{xy} in (i) we have written $\hat{P}(x, k)$ for $\hat{P}(x, \cdot, k)$. As in [55], we will assume that

(H1) The following properties hold true:

Assumption 4.2.1. (i) For every $x \in \mathcal{S}$ the functions $g(x, \cdot)$ and $\mathcal{P}(\mathcal{S}) \times \mathcal{P}(\mathcal{S}) \ni (p, M) \mapsto \sum_{y \in \mathcal{S}} p(y) c_{xy}(p, M)$ are continuous.

(ii) For every $U : \mathcal{S} \rightarrow \mathbb{R}$, $M \in \mathcal{P}(\mathcal{S})$ and $x \in \mathcal{S}$, the optimization problem

$$\inf_{p \in \mathcal{P}(\mathcal{S})} \sum_{y \in \mathcal{S}} p(y) [c_{xy}(p, M) + U(y)], \quad (4.6)$$

admits a unique solution $\hat{p}(x, \cdot) \in \mathcal{P}(\mathcal{S})$.

Remark 4.2.1. (i) By using Brower's fixed point theorem, it is proved in [55, Theorem 5] that under **(H1)**, problem **(MFG_d)** admits at least one solution.

(ii) As a consequence of the DPP, we have that **(H1)(ii)** implies that for every $M : \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$, problem

$$\inf_{P \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}} J_M(P)$$

admits a unique solution.

(iii) An example running cost c_{xy} satisfying that $\mathcal{P}(\mathcal{S}) \times \mathcal{P}(\mathcal{S}) \ni (p, M) \mapsto \sum_{y \in \mathcal{S}} p(y) c_{xy}(p, M)$ is continuous and **(H1)(ii)** is given by

$$c_{xy}(p, M) := K(x, y, M) + \epsilon \log(p(y)) \quad (4.7)$$

where $\epsilon > 0$, $K(x, y, \cdot)$ is continuous for all $x, y \in \mathcal{S}$, with the convention that $0 \log 0 = 0$. This type of cost has been already considered in [55], and, given $x \in \mathcal{S}$, the unique solution of (4.6) is given by

$$\hat{p}(x, y) = \frac{\exp(-[K(x, y, M) + U(y)]/\epsilon)}{\sum_{y' \in \mathcal{S}} \exp(-[K(x, y', M) + U(y')]/\epsilon)}.$$

In Section 4.4 we will consider this type of cost in order to approximate continuous MFGs by finite ones.

4.3 Fictitious play for the finite MFG system

Inspired by the fictitious play procedure introduced for continuous MFGs in chapter 3, we consider in this section the convergence problem for the sequence of functions transition kernels $P_n \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$ and marginal distributions $M_n : \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$ constructed as follows: given $M_1 : \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$ arbitrary, set $\bar{M}_1 = M_1$ and, for $n \geq 1$, define

$$\begin{aligned} P_n &:= \operatorname{argmin}_{P \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}} J_{\bar{M}_n}(P), \\ M_{n+1}(\cdot, k) &:= M_{P_n}^{M_0}(\cdot, k), \quad \forall k = 0, \dots, m, \\ \bar{M}_{n+1}(\cdot, k) &:= \frac{n}{n+1} \bar{M}_n(\cdot, k) + \frac{1}{n+1} M_{n+1}(\cdot, k), \quad \forall k = 0, \dots, m, \end{aligned} \quad (4.8)$$

where we recall that M_0 is given and for $P \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$, the function $M_P^{M_0} : \mathcal{S} \times \mathcal{T} \rightarrow [0, 1]$ is defined by (4.1) (or recursively by (4.2)). Note that by Remark 4.2.1(ii), the sequences (P_n) and (M_n) are well defined under **(H1)**.

The main object of this section is to show that, under suitable conditions, the sequence (P_n) converges to a solution \hat{P} to **(MFG_d)** and (M_n) converges to $M_{\hat{P}}^{M_0}$, i.e. the marginal distributions at the equilibrium. In practice, in order to compute M_{n+1} from \bar{M}_n , we find first P_n backwards in time by using the DPP expression for $U_{\bar{M}_n}$ in (4.3) and then we compute M_{n+1} forward in time by using (4.2). Notice that both computations are explicit in time.

4.3.1 Generalized fictitious play

For the sake of simplicity, we present here an abstract framework that will allow us to prove the convergence of the sequence constructed in (4.8). We begin by introducing some notations that will be also

used in Section 4.4. Let \mathcal{X} and \mathcal{Y} be two Polish spaces and $\Psi : \mathcal{X} \rightarrow \mathcal{Y}$ be a Borel measurable function. Given a Borel probability measure μ on \mathcal{X} , we denote by $\Psi\#\mu$ the probability measure on \mathcal{Y} defined by $\Psi\#\mu(A) := \mu(\Psi^{-1}(A))$ for all $A \in \mathcal{B}(\mathcal{Y})$. Denoting by $\mathcal{P}(\mathcal{X})$ the set of Borel probability measures on \mathcal{X} and by d the metric on \mathcal{X} , we set $\mathcal{P}_p(\mathcal{X})$ for the subset of $\mathcal{P}(\mathcal{X})$ consisting on measures μ such that $\int_{\mathcal{X}} d_{\mathcal{X}}(x, x_0)^p d\mu(x) < +\infty$ for some $x_0 \in \mathcal{X}$. For $\mu_1, \mu_2 \in \mathcal{P}_p(\mathcal{X})$ define

$$\Pi(\mu_1, \mu_2) := \{ \gamma \in \mathcal{P}(\mathcal{X} \times \mathcal{X}) \mid \rho\#\pi_1 = \mu_1 \text{ and } \rho\#\pi_2 = \mu_2 \},$$

where $\pi_1, \pi_2 : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$, are defined by $\pi_i(x_1, x_2) := x_i$ for $i = 1, 2$. Endowed with the Monge-Kantorovic metric

$$d_p(\mu_1, \mu_2) = \inf_{\gamma \in \Pi(\mu_1, \mu_2)} \left(\int_{\mathcal{X} \times \mathcal{X}} d(x, y)^p d\gamma(x, y) \right)^{1/p},$$

the set $\mathcal{P}_p(\mathcal{X})$ is shown to be a Polish space (see e.g. [8, Proposition 7.1.5]). Let us recall that d_1 corresponds to the Kantorovic-Rubinstein metric, i.e.

$$d_1(\mu_1, \mu_2) = \sup \left\{ \int_{\mathcal{X}} f(x) d(\mu_1 - \mu_2)(x) ; f \in \text{Lip}_1(\mathbb{R}^d) \right\}, \quad (4.9)$$

where $\text{Lip}_1(\mathcal{X})$ denotes the set of Lipschitz functions defined in \mathcal{X} with Lipschitz constant less or equal than 1 (see e.g. [92]).

Let $\mathcal{C} \subseteq \mathcal{X}$ be a compact set. Then, by definition, $\mathcal{P}(\mathcal{C}) = \mathcal{P}_p(\mathcal{C})$ for all $p \geq 1$, and d_p metricizes the weak convergence of probability measures on \mathcal{C} (see e.g. [8, Proposition 7.1.5]). Moreover, the set $\mathcal{P}(\mathcal{C})$ is compact.

Now, let $F : \mathcal{C} \times \mathcal{P}(\mathcal{C}) \rightarrow \mathbb{R}$ be a given continuous function. Given $x_1 \in \mathcal{C}$ set $\bar{\eta}_1 := \delta_{x_1}$, the Dirac mass at x_1 , and for $n \geq 1$ define:

$$x_{n+1} \in \text{argmin}_{x \in \mathcal{C}} F(x, \bar{\eta}_n), \quad \bar{\eta}_{n+1} = \frac{1}{n+1} \sum_{k=1}^{n+1} \delta_{x_k} = \frac{n}{n+1} \bar{\eta}_n + \frac{1}{n+1} \delta_{x_{n+1}}. \quad (4.10)$$

We consider now the convergence problem of the sequence $(\bar{\eta}_n)$ to some $\tilde{\eta} \in \mathcal{P}(\mathcal{C})$ satisfying that

$$\text{supp}(\tilde{\eta}) \subseteq \text{argmin}_{x \in \mathcal{C}} F(x, \tilde{\eta}), \quad (4.11)$$

where $\text{supp}(\tilde{\eta})$ denotes the support of the measure $\tilde{\eta}$. We call such $\tilde{\eta}$ an equilibrium and its existence can be easily proved by using Fan's fixed point theorem.

We will prove the convergence of $(\bar{\eta}_n)$ under a monotonicity and unique minimizer condition for F .

Definition 4.3.1 (Monotonicity). *The function F is called monotone, if*

$$\int_{\mathcal{C}} (F(x, \mu_1) - F(x, \mu_2)) d(\mu_1 - \mu_2)(x) \geq 0, \quad \forall \mu_1, \mu_2 \in \mathcal{P}(\mathcal{C}), \quad \mu_1 \neq \mu_2. \quad (4.12)$$

Moreover, F is called strictly monotone if the inequality in (4.12) is strict.

Definition 4.3.2 (Unique minimizer condition). *The function F satisfies the unique minimizer condition if for every $\eta \in \mathcal{P}(\mathcal{C})$ the optimization problem $\inf_{x \in \mathcal{C}} F(x, \eta)$ admits a unique solution.*

The following remark states some elementary consequence of the previous definitions.

Remark 4.3.1. (i) *If the unique minimizer condition holds then any equilibrium must be a Dirac mass. Moreover, the application $\mathcal{P}(\mathcal{C}) \ni \eta \mapsto x_{\eta} := \text{argmin}_{x \in \mathcal{C}} F(x, \eta) \in \mathcal{C}$ is well defined and uniformly continuous.*

(ii) *If F is monotone and the unique minimizer condition holds then the equilibrium must be unique.*

Indeed, suppose that there are two different equilibria $\tilde{\eta} = \delta_{\tilde{x}}$ and $\tilde{\eta}' = \delta_{\tilde{x}'}$. Then, by the unique minimizer condition,

$$F(\tilde{x}, \delta_{\tilde{x}}) < F(\tilde{x}', \delta_{\tilde{x}}), \quad \text{and} \quad F(\tilde{x}', \delta_{\tilde{x}'}) < F(\tilde{x}, \delta_{\tilde{x}'}).$$

This gives $\int_{\mathcal{C}} (F(x, \delta_{\tilde{x}}) - F(x, \delta_{\tilde{x}'})) \, d(\delta_{\tilde{x}} - \delta_{\tilde{x}'})(x) < 0$, which contradicts the monotonicity assumption.

Arguing as in [33, Proposition 2.9]), it is easy to see that uniqueness of the equilibrium also holds if F is strictly monotone but does not necessarily satisfy the unique minimizer condition.

Theorem 4.3.1. Assume that

- (i) F is monotone and satisfies the unique minimizer condition.
- (ii) F is Lipschitz, when $\mathcal{P}(\mathcal{C})$ is endowed with the distance d_1 , and there exists $C > 0$ such that

$$|F(x_1, \eta_1) - F(x_1, \eta_2) - F(x_2, \eta_1) + F(x_2, \eta_2)| \leq C d_{\mathcal{X}}(x_1, x_2) d_1(\eta_1, \eta_2), \quad (4.13)$$

for all $x_1, x_2 \in \mathcal{C}$, and $\mu_1, \mu_2 \in \mathcal{P}(\mathcal{C})$

Then, there exists $\tilde{x} \in \mathcal{C}$ such that $\tilde{\eta} = \delta_{\tilde{x}}$ is the unique equilibrium and the sequence $(x_n, \bar{\eta}_n)$ defined by (4.10) converges to $(\tilde{x}, \delta_{\tilde{x}})$.

Before we prove the theorem, let us recall a preliminary result (see chapter 3).

Lemma 4.3.1. Consider a sequence of real numbers (ϕ_n) such that $\liminf_{n \rightarrow \infty} \phi_n \geq 0$. If there exists a real sequence (ϵ_n) such that $\lim_{n \rightarrow \infty} \epsilon_n = 0$ and

$$\phi_{n+1} - \phi_n \leq -\frac{1}{n+1} \phi_n + \frac{\epsilon_n}{n}, \quad \forall n \in \mathbb{N},$$

then $\lim_{n \rightarrow \infty} \phi_n = 0$.

Proof. Let $b_n = n\phi_n$ for every $n \in \mathbb{N}$. We have

$$\frac{b_{n+1}}{n+1} - \frac{b_n}{n} \leq -\frac{b_n}{n(n+1)} + \frac{\epsilon_n}{n}, \quad \forall n \in \mathbb{N},$$

which implies that $b_{n+1} \leq b_n + (n+1)\epsilon_n/n \leq b_n + 2|\epsilon_n|$. Then, we get $b_n \leq b_1 + 2 \sum_{i=1}^{n-1} |\epsilon_i|$ and, hence,

$$0 \leq \liminf_{n \rightarrow \infty} \phi_n \leq \limsup_{n \rightarrow \infty} \phi_n \leq \lim_{n \rightarrow \infty} \frac{b_1 + 2 \sum_{i=1}^{n-1} |\epsilon_i|}{n} = 0,$$

from which the result follows. \square

Proof of Theorem 4.3.1. Let us define the real sequence (ϕ_n) as

$$\phi_n := \int_{\mathcal{C}} F(x, \bar{\eta}_n) d\bar{\eta}_n(x) - F(x_{n+1}, \bar{\eta}_n).$$

We claim that $\phi_n \rightarrow 0$. Assuming that the claim is true, then any limit point $(\tilde{x}, \tilde{\eta})$ of $(x_n, \bar{\eta}_n)$ satisfies that

$$F(\tilde{x}, \tilde{\eta}) \leq F(x, \tilde{\eta}) \quad \forall x \in \mathcal{C}, \quad \text{and} \quad F(\tilde{x}, \tilde{\eta}) = \int_{\mathcal{C}} F(x, \tilde{\eta}) d\tilde{\eta}(x),$$

which implies that $\tilde{\eta}$ satisfies (4.11), i.e. $\tilde{\eta}$ is an equilibrium. Using that F is monotone and Remark 4.3.1(ii), the assertions on the theorem follows.

Thus, it remains to show that $\phi_n \rightarrow 0$, which will be proved with the help of Lemma 4.3.1. By definition of x_{n+1} we have that $\phi_n \geq 0$. Let us write $\phi_{n+1} - \phi_n = A + B$, where

$$A = \int_{\mathcal{C}} F(x, \bar{\eta}_{n+1}) \, d\bar{\eta}_{n+1}(x) - \int_{\mathcal{C}} F(x, \bar{\eta}_n) \, d\bar{\eta}_n(x), \quad B = F(x_{n+1}, \bar{\eta}_n) - F(x_{n+2}, \bar{\eta}_{n+1}).$$

We have

$$\begin{aligned}
B &\leq F(x_{n+2}, \bar{\eta}_n) - F(x_{n+2}, \bar{\eta}_{n+1}) \\
&\leq F(x_{n+1}, \bar{\eta}_n) - F(x_{n+1}, \bar{\eta}_{n+1}) + C \, d_{\mathcal{X}}(x_{n+2}, x_{n+1}) d_1(\bar{\eta}_n, \bar{\eta}_{n+1}) \\
&\leq F(x_{n+1}, \bar{\eta}_n) - F(x_{n+1}, \bar{\eta}_{n+1}) + \frac{C}{n+1} \, d_{\mathcal{X}}(x_{n+2}, x_{n+1}) d_1(\delta_{x_{n+1}}, \bar{\eta}_n),
\end{aligned} \tag{4.14}$$

where we have used (4.13) to pass from the first to the second inequality and (4.9) from the second to the third inequality. Similarly, using (4.10) and that F is Lipschitz,

$$\begin{aligned}
A &= \int_{\mathcal{C}} (F(x, \bar{\eta}_{n+1}) - F(x, \bar{\eta}_n)) \, d\bar{\eta}_n(x) + \frac{1}{n+1} \left[F(x_{n+1}, \bar{\eta}_{n+1}) - \int_{\mathcal{C}} F(x, \bar{\eta}_{n+1}) \, d\bar{\eta}_n(x) \right] \\
&\leq \int_{\mathcal{C}} (F(x, \bar{\eta}_{n+1}) - F(x, \bar{\eta}_n)) \, d\bar{\eta}_n(x) + \frac{1}{n+1} \left[F(x_{n+1}, \bar{\eta}_n) - \int_{\mathcal{C}} F(x, \bar{\eta}_n) \, d\bar{\eta}_n(x) \right] + \frac{C}{n+1} d_1(\bar{\eta}_n, \bar{\eta}_{n+1}) \\
&\leq \int_{\mathcal{C}} (F(x, \bar{\eta}_{n+1}) - F(x, \bar{\eta}_n)) \, d\bar{\eta}_n(x) - \frac{1}{n+1} \phi_n + \frac{C}{(n+1)^2} d_1(\bar{\eta}_n, \delta_{x_{n+1}}).
\end{aligned} \tag{4.15}$$

On the other hand, the second relation in (4.10) yields $-(n+1)(\bar{\eta}_{n+1} - \bar{\eta}_n) = \bar{\eta}_n - \delta_{x_{n+1}}$. Therefore,

$$\begin{aligned}
&F(x_{n+1}, \bar{\eta}_n) - F(x_{n+1}, \bar{\eta}_{n+1}) + \int_{\mathcal{C}} (F(x, \bar{\eta}_{n+1}) - F(x, \bar{\eta}_n)) \, d\bar{\eta}_n(x) = \\
&\quad -(n+1) \int_{\mathcal{C}} (F(x, \bar{\eta}_{n+1}) - F(x, \bar{\eta}_n)) \, d(\bar{\eta}_{n+1} - \bar{\eta}_n)(x) \leq 0,
\end{aligned} \tag{4.16}$$

by the monotonicity condition of F . From estimates (4.14)-(4.15) and inequality (4.16) we deduce that

$$\phi_{n+1} - \phi_n \leq -\frac{1}{n+1} \phi_n + \frac{C}{n+1} d_1(\delta_{x_{n+1}}, \bar{\eta}_n) \left(\frac{1}{n+1} + d_{\mathcal{X}}(x_{n+2}, x_{n+1}) \right). \tag{4.17}$$

Using that $\mathcal{P}(\mathcal{C})$ is compact (and so bounded in d_1), we get that

$$\phi_{n+1} - \phi_n \leq -\frac{1}{n+1} \phi_n + \frac{\epsilon_n}{n},$$

where $\epsilon_n := C'(\frac{1}{n+1} + d_{\mathcal{X}}(x_{n+2}, x_{n+1}))$, with $C' > 0$ and independent of n . Remark 4.3.1 implies that $d_{\mathcal{X}}(x_{n+2}, x_{n+1}) \rightarrow 0$ as $n \rightarrow \infty$ (because $d_1(\bar{\eta}_n, \bar{\eta}_{n+1}) = d_1(\bar{\eta}_n, \delta_{x_{n+1}})/(n+1) \rightarrow 0$). Thus, $\epsilon_n \rightarrow 0$ and the result follows from Lemma 4.3.1. \square

4.3.2 Convergence of the fictitious play for finite MFG

In this section, we apply the abstract result in Theorem 4.3.1 to the finite MFG problem (MFG_d). Under the notations of Section 4.2, in what follows, will assume that $c_{xy}(\cdot, \cdot)$ has a separable form. Namely,

$$c_{xy}(p, M) = K(x, y, p) + f(x, M), \quad \forall x, y \in \mathcal{S}, \quad p, M \in \mathcal{P}(\mathcal{S}), \tag{4.18}$$

where $K : \mathcal{S} \times \mathcal{S} \times \mathcal{P}(\mathcal{S}) \rightarrow \mathbb{R}$ and $f : \mathcal{S} \times \mathcal{P}(\mathcal{S}) \rightarrow \mathbb{R}$ are given. In order to write (MFG_d) as a particular instance of (4.11), given $\eta \in \mathcal{P}(\mathcal{K}_{\mathcal{S}, \mathcal{T}})$ we define $M_\eta := \mathcal{T} \rightarrow \mathcal{P}(\mathcal{S})$ and $F : \mathcal{K}_{\mathcal{S}, \mathcal{T}} \times \mathcal{P}(\mathcal{K}_{\mathcal{S}, \mathcal{T}}) \rightarrow \mathbb{R}$ as

$$M_\eta(k) := \int_{\mathcal{K}_{\mathcal{S}, \mathcal{T}}} M_P^{M_0}(k) \, d\eta(P), \quad \forall k = 0, \dots, m, \quad \text{and} \quad F(P, \eta) := J_{M_\eta}(P). \tag{4.19}$$

Under assumption (H1), we have that F is continuous and satisfies the unique minimizer condition in Definition 4.3.2. Therefore, by Remark 4.3.1(i), associated to any equilibrium $\eta \in \mathcal{P}(\mathcal{K}_{\mathcal{S}, \mathcal{T}})$ for F , i.e. η satisfies (4.11) with $\mathcal{C} = \mathcal{K}_{\mathcal{S}, \mathcal{T}}$, there exists $P_\eta \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$ such that $\eta = \delta_{P_\eta}$, from which we get that

P_η solves (MFG_d). Conversely, for any solution P to (MFG_d) we can associate the measure $\eta_P := \delta_P$, which solves (4.10). An analogous argument shows that the fictitious play procedures (4.8) and (4.10) are equivalent.

We consider now some assumptions on the data of the finite MFG problem that will ensure the validity of assumptions (i)-(ii) for F in Theorem 4.3.1.

(H2) We assume that

(i) f and g are monotone, in the sense that setting $h = f, g$, we have

$$\sum_{x \in \mathcal{S}} (h(x, p_1) - h(x, p_2)) (p_1(x) - p_2(x)) \geq 0 \quad \forall p_1, p_2 \in \mathcal{P}(\mathcal{S}).$$

(ii) f and g are Lipschitz with respect to their second argument.

The following result is a straightforward consequence of the definitions.

Lemma 4.3.2. *If f and g are monotone, then F is monotone in sense of Definition 4.3.1.*

Proof. For any two distributions $\eta, \eta' \in \mathcal{P}(\mathcal{C})$ we want to show $\int_{\mathcal{C}} (F(P, \eta) - F(P, \eta')) d(\eta - \eta')(P) \geq 0$. By using the exact form of the cost function F by equation (4.19) and taking into account the separable form of the running cost (4.18), we have:

$$\begin{aligned} F(P, \eta) - F(P, \eta') &= \sum_{k=0}^{m-1} \sum_{x \in \mathcal{S}} M_P^{M_0}(x, k) [f(x, M_\eta(k)) - f(x, M_{\eta'}(k))] \\ &\quad + \sum_{x \in \mathcal{S}} M_P^{M_0}(x, m) [g(x, M_\eta(m)) - g(x, M_{\eta'}(m))]. \end{aligned}$$

Thus,

$$\begin{aligned} \int_{\mathcal{K}_{\mathcal{S}, \tau}} (F(P, \eta) - F(P, \eta')) d(\eta - \eta')(P) &= \sum_{k=0}^{m-1} \sum_{x \in \mathcal{S}} [f(x, M_\eta(k)) - f(x, M_{\eta'}(k))] \int_{\mathcal{K}_{\mathcal{S}, \tau}} M_P^{M_0}(x, k) d(\eta - \eta')(P) \\ &\quad + \sum_{x \in \mathcal{S}} [g(x, M_\eta(m)) - g(x, M_{\eta'}(m))] \int_{\mathcal{K}_{\mathcal{S}, \tau}} M_P^{M_0}(x, m) d(\eta - \eta')(P) \\ &= \sum_{k=0}^{m-1} \sum_{x \in \mathcal{S}} [f(x, M_\eta(k)) - f(x, M_{\eta'}(k))] (M_\eta(x, k) - M_{\eta'}(x, k)) \\ &\quad + \sum_{x \in \mathcal{S}} [g(x, M_\eta(m)) - g(x, M_{\eta'}(m))] (M_\eta(x, m) - M_{\eta'}(x, m)) \geq 0, \end{aligned}$$

where the positiveness follows from the monotonicity of f and g . \square

By Remark 4.3.1 we directly deduce the following result.

Proposition 4.3.1. *If (H1) and (H2)(ii) hold, then the finite MFG (MFG_d) has a unique equilibrium.*

Remark 4.3.2. *The previous result slightly improves [55, Theorem 6], where the uniqueness of the equilibrium is proved under a stronger strict monotonicity assumption on f and g .*

In order to check assumption (ii) in Theorem 4.3.1, we need first a preliminary result.

Lemma 4.3.3. *There exists a constant $C > 0$ such that*

$$|M_P^{M_0}(k) - M_{P'}^{M_0}(k)| \leq C|P - P'|_\infty \quad \forall P, P' \in \mathcal{K}_{\mathcal{S}, \tau}, \quad k = 0, \dots, m. \quad (4.20)$$

In particular,

$$|M_\eta^{M_0}(k) - M_{\eta'}^{M_0}(k)| \leq C d_1(\eta, \eta') \quad \forall \eta, \eta' \in \mathcal{P}(\mathcal{K}_{\mathcal{S}, \tau}), \quad k = 0, \dots, m. \quad (4.21)$$

Proof. For any $k = 0, \dots, m-1$ and $x \in \mathcal{S}$ we have

$$\begin{aligned}
M_P^{M_0}(x, k+1) - M_{P'}^{M_0}(x, k+1) &= \sum_{y \in \mathcal{S}} M_P^{M_0}(y, k) P(y, x, k) - \sum_{y \in \mathcal{S}} M_{P'}^{M_0}(y, k) P'(y, x, k) \\
&\leq \sum_{y \in \mathcal{S}} M_P^{M_0}(y, k) (P(y, x, k) - P'(y, x, k)) \\
&\quad + |M_P^{M_0}(k) - M_{P'}^{M_0}(k)|_\infty \sum_{y \in \mathcal{S}} P'(y, x, k) \\
&\leq |P - P'|_\infty + |\mathcal{S}| |M_P^{M_0}(k) - M_{P'}^{M_0}(k)|_\infty,
\end{aligned} \tag{4.22}$$

where we have used that $\sum_{y \in \mathcal{S}} M_P^{M_0}(y, k) = 1$. Using that $M_P^{M_0}(0) = M_{P'}^{M_0}(0) = M_0$, inequality (4.20) follows by applying (4.22) recursively. Now, given $\gamma \in \Pi(\eta, \eta')$, i.e. $\gamma \in \mathcal{P}(\mathcal{K}_{\mathcal{S}, \mathcal{T}} \times \mathcal{K}_{\mathcal{S}, \mathcal{T}})$ with marginals given by η and η' , we have

$$\begin{aligned}
|M_\eta^{M_0}(k) - M_{\eta'}^{M_0}(k)| &= \left| \int_{\mathcal{K}_{\mathcal{S}, \mathcal{T}}} M_P^{M_0}(k) d\eta(P) - \int_{\mathcal{K}_{\mathcal{S}, \mathcal{T}}} M_{P'}^{M_0}(k) d\eta'(P') \right| \\
&= \left| \int_{\mathcal{K}_{\mathcal{S}, \mathcal{T}}} (M_P^{M_0}(k) - M_{P'}^{M_0}(k)) d\gamma(P, P') \right| \\
&\leq C \int_{\mathcal{K}_{\mathcal{S}, \mathcal{T}}} |P - P'|_\infty d\gamma(P, P').
\end{aligned}$$

Inequality (4.21) follows by taking the infimum over $\gamma \in \Pi(\eta, \eta')$. \square

Lemma 4.3.4. Assume that (H2)(ii) holds. Then, there exists $C > 0$ such that

$$\begin{aligned}
|F(P, \eta) - F(P, \eta') - F(P', \eta) + F(P', \eta')| &\leq C |P - P'|_\infty d_1(\eta, \eta'), \\
|F(P, \eta) - F(P, \eta')| &\leq C d_1(\eta, \eta'),
\end{aligned} \tag{4.23}$$

for all $P, P' \in \mathcal{K}_{\mathcal{S}, \mathcal{T}}$ and $\eta, \eta' \in \mathcal{P}(\mathcal{K}_{\mathcal{S}, \mathcal{T}})$.

Proof. Let us first prove the second relation in (4.23). Denoting by $c > 0$, the maximum between the Lipschitz constants of f and g , we can write $|F(P, \eta) - F(P, \eta')| \leq A + B$ with

$$A := \sum_{k=0}^{m-1} \sum_{x \in \mathcal{S}} M_P^{M_0}(x, k) |f(x, M_\eta(k)) - f(x, M_{\eta'}(k))| \leq \sum_{k=0}^{m-1} \sum_{x \in \mathcal{S}} M_P^{M_0}(x, k) c m d_1(\eta, \eta') = c m^2 d_1(\eta, \eta'),$$

and

$$B := \sum_{x \in \mathcal{S}} M_P^{M_0}(x, m) |g(x, M_\eta(m)) - g(x, M_{\eta'}(m))| \leq \sum_{x \in \mathcal{S}} M_P^{M_0}(x, m) c d_1(\eta, \eta') = c m d_1(\eta, \eta'),$$

where the inequalities follow from (4.21). Thus, the second estimate in (4.23) follows. In order to prove the first relation in (4.23), let us write $|F(P, \eta) - F(P, \eta') - F(P', \eta) + F(P', \eta')| \leq A' + B'$ with

$$A' := \sum_{k=0}^{m-1} \sum_{x \in \mathcal{S}} |M_P(x, k) - M_{P'}(x, k)| |f(x, M_\eta(k)) - f(x, M_{\eta'}(k))| \leq C m |\mathcal{S}| |P - P'|_\infty d_1(\eta, \eta'),$$

$$B' := \sum_{x \in \mathcal{S}} |M_P(x, m) - M_{P'}(x, m)| |g(x, M_\eta(m)) - g(x, M_{\eta'}(m))| \leq C |\mathcal{S}| |P - P'|_\infty d_1(\eta, \eta').$$

The result follows. \square

By combining Lemma 4.3.2, Lemma 4.3.4 and Theorem 4.3.1, we get the following convergence result.

Theorem 4.3.2. Assume (H1) and (H2) and let (P_n, M_n, \bar{M}_n) be the sequence generated in the fictitious play procedure (4.8). Then, $(P_n, M_n, \bar{M}_n) \rightarrow (\hat{P}, M_{\hat{P}}^{M_0}, \bar{M}_{\hat{P}}^{M_0})$, where \hat{P} is the unique solution to (MFG_d).

4.4 First order MFG as limits of finite MFG

In this section we consider a relaxed first order MFG problem in continuous time and with a continuum of states. We define a natural finite MFG associated to a discretization of the space and time variables. We address our second main question in this work, which is the convergence of the solutions of finite MFGs to solutions of continuous MFGs when the discretization parameters tend to zero.

In order to introduce the MFG problem, we need first to introduce some definitions. Let us define $\Gamma = C([0, T]; \mathbb{R}^d)$ and given $m_0 \in \mathcal{P}(\mathbb{R}^d)$, called the initial distribution, let

$$\mathcal{P}_{m_0}(\Gamma) = \{\eta \in \mathcal{P}(\Gamma) ; e_0 \# \eta = m_0\},$$

where, for each $t \in [0, T]$, the function $e_t : \Gamma \rightarrow \mathbb{R}^d$ is defined by $e_t(\gamma) = \gamma(t)$. Let $q \in (1, +\infty)$, with conjugate exponent $q' := q/(q-1)$, and $f, g : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$. Given $m \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ consider the following family of optimal control problems, parametrized by the initial condition,

$$\inf \left\{ \int_0^T \left[\frac{1}{q} |\dot{z}(t)|^q + f(z(t), m(t)) \right] dt + g(z(T), m(T)) \mid z \in W^{1,q}([0, T]; \mathbb{R}^d), z(0) = x \right\}, \quad x \in \mathbb{R}^d. \quad (4.24)$$

Definition 4.4.1. We call $\xi^* \in \mathcal{P}_{m_0}(\Gamma)$ a MFG equilibrium for (4.24) if $[0, T] \ni t \mapsto e_t \# \xi^*$ belongs to $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ and ξ^* -almost every γ solves the optimal control problem in (4.24) with $x = \gamma(0)$ and $m(t) = e_t \# \xi^*$ for all $t \in [0, T]$.

Assuming that the cost functional of the optimal control problem in (4.24) is meaningful, which is ensured by the conditions on f and g in assumption **(H3)** below, the interpretation of a MFG equilibrium is as follows: the measure ξ^* is an equilibrium if it only charges trajectories in \mathbb{R}^d , distributed as m_0 at the initial time, minimizing a cost depending on the collection of time marginals of ξ^* in $[0, T]$.

Remark 4.4.1. Usually, see e.g. [74] and [33], a first order MFG equilibrium is presented in the form of a system of PDEs consisting in a HJB equation, modelling the fact that a typical agent solves an optimal control problem, which depends on the marginal distributions of the agents at each time $t \in [0, T]$, coupled with a continuity equation, describing the evolution of the aforementioned marginal distributions if the agents follow the optimal dynamics. The definition of equilibrium that we adopted in this work corresponds to a relaxation of the PDE notion of equilibrium, and has been used, for instance, in [39], [22, Section 3] and, recently, in [31].

Throughout this section, we will suppose that the following assumption holds.

(H3)(i) For $h = f, g$ we have that h is continuous and there exists $C > 0$ such that

$$\sup_{m \in \mathcal{P}_1(\mathbb{R}^d)} \{ \|h(\cdot, m)\|_\infty + \|D_x h(\cdot, m)\|_\infty \} \leq C. \quad (4.25)$$

(ii) The initial distribution $m_0 \in \mathcal{P}(\mathbb{R}^d)$ has a compact support.

Now we will focus on a particular class of finite MFGs and relate their solutions, asymptotically, with the MFG equilibria for (4.24). Let (N_n^s) and (N_n^t) be two sequences of natural numbers such that $\lim_{n \rightarrow \infty} N_n^s = \lim_{n \rightarrow \infty} N_n^t = +\infty$ and let (ϵ_n) be a sequence of positive real numbers such that $\lim_{n \rightarrow \infty} \epsilon_n = 0$. Define $\Delta x_n := 1/N_n^s$ and $\Delta t_n := T/N_n^t$. For a fixed $n \in \mathbb{N}$, consider the discrete state set \mathcal{S}_n and the discrete time set \mathcal{T}_n defined as

$$\begin{aligned} \mathcal{S}_n &:= \{x_q := q\Delta x_n \mid q \in \mathbb{Z}^d, |q|_\infty \leq (N_n^s)^2\} \subseteq \mathbb{R}^d, \\ \mathcal{T}_n &:= \{t_k := k\Delta t_n \mid k = 0, \dots, N_n^t\} \subseteq [0, T]. \end{aligned} \quad (4.26)$$

Let us also define the (non positive) entropy function $\mathcal{E}_n : \mathcal{P}(\mathcal{S}_n) \rightarrow \mathbb{R}$ by

$$\mathcal{E}_n(p) = \sum_{x \in \mathcal{S}_n} p_x \log(p_x) \quad \forall p \in \mathcal{P}(\mathcal{S}_n),$$

with the convention that $0 \log 0 = 0$. For every $x \in \mathcal{S}_n$ set $E_x := \{x' \in \mathbb{R}^d \mid |x' - x|_\infty \leq \Delta x_n/2\}$. Since we will be interested in the asymptotic as $n \rightarrow \infty$, we can assume, without loss of generality, that $m_0(\partial E_x) = 0$ for all $x \in \mathcal{S}_n$. Similarly, by **(H3)**(ii), we can assume that the support of m_0 will be contained in $\cup_{x \in \mathcal{S}_n} E_x$. Based on these considerations, setting

$$M_{n,0}(x) := m_0(E_x) \quad \forall x \in \mathcal{S}_n,$$

we have that $M_{n,0} \in \mathcal{P}(\mathcal{S}_n)$. We consider the finite MFG, written in a recursive form (see (4.4)),

$$\begin{aligned} \text{(i)} \quad & U_n(x, t_k) = \min_{p \in \mathcal{P}(\mathcal{S}_n)} \left\{ \sum_{y \in \mathcal{S}_n} p(y) \left(\frac{\Delta t_n}{q} \left| \frac{y-x}{\Delta t_n} \right|^q + U_n(y, t_{k+1}) \right) + \epsilon_n \mathcal{E}_n(p) \right\} \\ & + \Delta t_n f(x, M_n(t_k)) \quad \forall x \in \mathcal{S}_n, \quad 0 \leq k < N_n^t, \\ \text{(ii)} \quad & M_n(y, t_{k+1}) = \sum_{x \in \mathcal{S}_n} \hat{P}_n(x, y, t_k) M_n(x, t_k) \quad \forall y \in \mathcal{S}_n, \quad 0 \leq k < N_n^t, \\ \text{(iii)} \quad & M_n(x, 0) = M_{n,0}(x), \quad U_n(x, T) = g(x, M_n(T)) \quad \forall x \in \mathcal{S}_n, \end{aligned} \tag{4.27}$$

where for all $x \in \mathcal{S}_n$, $0 \leq k \leq N_n^t - 1$, $\hat{P}_n(x, \cdot, t_k) \in \mathcal{P}(\mathcal{S}_n)$ is given by

$$\hat{P}_n(x, \cdot, t_k) = \operatorname{argmin}_{p \in \mathcal{P}(\mathcal{S}_n)} \left\{ \sum_{y \in \mathcal{S}_n} p(y) \left(\frac{\Delta t_n}{q} \left| \frac{y-x}{\Delta t_n} \right|^q + U_n(y, t_{k+1}) \right) + \epsilon_n \mathcal{E}_n(p) \right\}. \tag{4.28}$$

Note that system (4.27) is a particular case of (4.4), with

$$c_{xy}(p, M) := \Delta t_n \left(\frac{1}{q} \left| \frac{y-x}{\Delta t_n} \right|^q + f(x, M) \right) + \epsilon_n \log(p_y).$$

Remark 4.4.2. The positive parameter ϵ_n is introduced in (4.27) in order to ensure that \hat{P}_n is well-defined, and so that assumption **(H1)** for system (4.27) is satisfied in this case. In particular, by the results in the previous sections, the fictitious play procedure converges for system (4.27) if the couplings f and g are monotone.

Remark 4.2.1 ensures the existence of at least one solution (U_n, M_n) of (4.27), with associated transition kernel \hat{P}_n given by (4.28). In order to study the asymptotic behaviour of (U_n, M_n, \hat{P}_n) , let us introduce some useful notations. We set $\mathcal{K}_n := \mathcal{K}_{\mathcal{S}_n, \mathcal{T}_n}$ (see Definition 4.2.1) and, given $x \in \mathcal{S}_n$ and $t \in \mathcal{T}_n$, we denote by $\Gamma_{x,t}^{\mathcal{S}_n, \mathcal{T}_n} \subseteq \Gamma_t$ the set of continuous functions $\gamma : [t, T] \rightarrow \mathbb{R}^d$ such that $\gamma(t) = x$ and for each $1 \leq k \leq m$ with $t_k \in \mathcal{T}_n \cap (t, T]$, we have that $\gamma(t_k) \in \mathcal{S}_n$ and the restriction of γ to the interval $[t_{k-1}, t_k]$ is affine. Given $P \in \mathcal{K}_n$ let us define $\xi_P^{x,t,n} \in \mathcal{P}(\Gamma)$ by

$$\xi_P^{x,t,n} := \sum_{\gamma \in \Gamma_{x,t}^{\mathcal{S}_n, \mathcal{T}_n}} p_P^{x,t,n}(\gamma) \delta_\gamma, \quad \text{where} \quad p_P^{x,t,n}(\gamma) := \prod_{t_k \in \mathcal{T}_n \cap [t, T]} P(\gamma(t_k), \gamma(t_{k+1}), t_k). \tag{4.29}$$

For later use, note that, recalling (4.29), equation (4.27)(i) is equivalent to

$$\begin{aligned} U_n(x, t_k) = \min_{P \in \mathcal{K}_n} \left\{ \mathbb{E}_{\xi_P^{x,t_k,n}} \left(\Delta t_n \sum_{k'=k}^{N_n^t-1} \left[\frac{1}{q} \left| \frac{y-x}{\Delta t_n} \right|^q + f(\gamma(t_{k'}), M_n(t_{k'})) \right] + g(\gamma(T), M_n(T)) \right) \right. \\ \left. + \epsilon_n \mathbb{E}_{\xi_P^{x,t_k,n}} \left(\sum_{k'=k}^{N_n^t-1} \sum_{y \in \mathcal{S}_n} P(\gamma(t_{k'}), y, t_{k'}) \log P(\gamma(t_{k'}), y, t_{k'}) \right) \right\}, \end{aligned} \tag{4.30}$$

for all $x \in \mathcal{S}_n$ and $k = 0, \dots, N_n^t - 1$. Finally, let us define $\xi_n \in \mathcal{P}(\Gamma)$ by

$$\xi_n := \sum_{x \in \mathcal{S}} M_{n,0}(x) \xi_{\hat{P}_n}^{x,0,n}.$$

Notice that, by definition, $M_n(t) = e_t \# \xi_n$ for all $t \in \mathcal{T}_n$. We extend $M_n : \mathcal{T}_n \rightarrow \mathcal{P}_1(\mathbb{R}^d)$ to $M_n : [0, T] \rightarrow \mathcal{P}_1(\mathbb{R}^d)$ via the formula

$$M_n(t) := e_t \# \xi_n \quad \text{for all } t \in [0, T]. \quad (4.31)$$

4.4.1 Convergence analysis

We now study the limit behaviour of the solutions (U_n, M_n) in (4.27), and of the associated sequence (ξ_n) , as $n \rightarrow \infty$. In the remainder of this section, for a given Borel measurable function $L : \Gamma \rightarrow \mathbb{R}$ we set $\mathbb{E}_{\xi_n}(L) := \int_{\Gamma} L(\gamma) d\xi_n(\gamma)$, provided that the integral is well-defined.

Let us prove first two simple but useful results.

Lemma 4.4.1. *Suppose that $\epsilon_n = O\left(\frac{1}{N_n^t \log(N_n^s)}\right)$. Then, there exists a constant $C > 0$, independent of n , such that*

$$\sup_{x \in \mathcal{S}_n, t \in \mathcal{T}_n} |U_n(x, t)| \leq C, \quad (4.32)$$

$$\mathbb{E}_{\xi_n} \left(\int_0^T |\dot{\gamma}(t)|^q dt \right) \leq C. \quad (4.33)$$

Proof. Let us first prove (4.32). Since the cardinality of \mathcal{S}_n is equal to $(2(N_n^s)^2 + 1)^d$, we have that

$$\left(\frac{1}{(2(N_n^s)^2 + 1)^d}, \dots, \frac{1}{(2(N_n^s)^2 + 1)^d} \right) = \operatorname{argmin} \left\{ \sum_{x \in \mathcal{S}_n} p_x \log p_x ; p \in \mathcal{P}(\mathcal{S}_n) \right\}.$$

Hence, our assumption over ϵ_n implies the existence of $\hat{C} > 0$, independent of n , such that

$$\left| \epsilon_n \mathbb{E}_{\xi_P^{x,t,n}} \left(\sum_{k'=k}^{N_n^t-1} \sum_{y \in \mathcal{S}_n} P(\gamma(t_{k'}), y, t_{k'}) \log P(\gamma(t_{k'}), y, t_{k'}) \right) \right| \leq \hat{C} \quad \forall P \in \mathcal{K}_n. \quad (4.34)$$

Thus, the lower bound is a direct consequence of the uniform bounds for f and g in (4.25). In order to obtain the upper bound, choose $P \in \mathcal{K}_n$ in the right hand side of (4.30) such that $P(x, x, t_{k'}) = 1$ for all $k' = k, \dots, N_n^t - 1$. The bound in (4.25) implies that

$$U_n(x, t_k) \leq C(T + 1) + \hat{C},$$

and so (4.32) follows. In order to prove (4.33), note that the boundedness of f and g and the bound (4.32) imply the existence of $\hat{C}_1 > 0$, independent of n , such that

$$\mathbb{E}_{\xi_n} \left(\int_0^T |\dot{\gamma}(t)|^q dt \right) = \mathbb{E}_{\xi_n} \left(\Delta t_n \sum_{k=0}^{N_n^t-1} \left| \frac{\gamma(t_{k+1}) - \gamma(t_k)}{\Delta t_n} \right|^q \right) \leq \hat{C}_1.$$

The result follows. □

Lemma 4.4.2. *Let $C > 0$. Then the set*

$$\Gamma_C := \left\{ \gamma \in W^{1,q}([0, T]; \mathbb{R}^d) \mid |\gamma(0)| \leq C \text{ and } \int_0^T |\dot{\gamma}(t)|^q dt \leq C \right\},$$

is a compact subset of Γ .

Proof. Let (γ_n) be a sequence in Γ_C . Then, for all $0 \leq s \leq t \leq T$, the Hölder inequality yields

$$|\gamma_n(t) - \gamma_n(s)| \leq \int_s^t |\dot{\gamma}_n(t')| dt' \leq C^{1/q} (t-s)^{1/q'}. \quad (4.35)$$

Thus,

$$|\gamma_n(t)| \leq |\gamma_n(0)| + |\gamma_n(t) - \gamma_n(0)| \leq C + C^{1/q} T^{1/q'}. \quad (4.36)$$

As a consequence of (4.35)-(4.36) and the Arzelà-Ascoli theorem we have existence of $\gamma \in \Gamma$ such that, up to some subsequence, $\gamma_n \rightarrow \gamma$ uniformly in $[0, T]$. Moreover, since $\dot{\gamma}_n$ is bounded in $L^q((0, T); \mathbb{R}^d)$ and the function $L^q((0, T); \mathbb{R}^d) \ni \eta \mapsto \int_0^T |\eta(t)|^q dt \in \mathbb{R}$ is convex and continuous, and hence, weakly lower semicontinuous, we have the existence of $\bar{\eta} \in L^q((0, T); \mathbb{R}^d)$ such that, up to some subsequence, $\dot{\gamma}_n \rightarrow \bar{\eta}$ weakly in $L^q((0, T); \mathbb{R}^d)$ and $\int_0^T |\bar{\eta}(t)|^q dt \leq \liminf_{n \rightarrow \infty} \int_0^T |\dot{\gamma}_n(t)|^q dt \leq C$. By passing to the limit in the relation

$$\gamma_n(t) = \gamma_n(0) + \int_0^t \dot{\gamma}_n(s) ds \quad \forall t \in [0, T],$$

we get that

$$\gamma(t) = \gamma(0) + \int_0^t \bar{\eta}(s) ds \quad \forall t \in [0, T],$$

and, hence, $\gamma \in W^{1,q}([0, T]; \mathbb{R}^d)$, with $\dot{\gamma} = \bar{\eta}$ a.e. in $[0, T]$, $|\gamma(0)| \leq C$ and $\int_0^T |\dot{\gamma}(t)|^q dt \leq C$. Therefore, $\gamma \in \Gamma_C$ and so the set Γ_C is compact. \square

As a consequence of the previous results we easily obtain a compactness property for (ξ_n) .

Proposition 4.4.1. *Suppose that $\epsilon_n = O\left(\frac{1}{N_n^t \log(N_n^s)}\right)$. Then, the sequence (ξ_n) is a relatively compact subset of $\mathcal{P}(\Gamma)$ endowed with the topology of narrow convergence.*

Proof. By Prokhorov's theorem it suffices to show that (ξ_n) is tight, i.e. we need to prove that for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subseteq \Gamma$ such that $\sup_{n \in \mathbb{N}} \xi_n(\Gamma \setminus K_\varepsilon) \leq \varepsilon$. Given $\varepsilon > 0$, the bound (4.33) and the Markov's inequality yield

$$\xi_n \left(\left\{ \gamma \in \Gamma \mid \gamma \in W^{1,q}((0, T); \mathbb{R}^d) \text{ and } \int_0^T |\dot{\gamma}(t)|^q dt > \frac{C}{\varepsilon} \right\} \right) \leq \varepsilon \quad \forall n \in \mathbb{N}. \quad (4.37)$$

On the other hand, by **(H)**(ii), there exists $c_0 > 0$ such that for ξ_n -almost every $\gamma \in \Gamma$ we have $|\gamma(0)| \leq c_0$. By Lemma 4.4.2 and (4.37), the set $K_\varepsilon := \Gamma_{C_\varepsilon}$ with $C_\varepsilon := \max\{c_0, C/\varepsilon\}$, satisfies the required properties. \square

Now, we study the compactness of the collection of marginal laws, with respect to the time variables, in the space $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$.

Proposition 4.4.2. *Suppose that $\epsilon_n = O\left(\frac{1}{N_n^t \log(N_n^s)}\right)$. Then, there exists $C > 0$ such that for all $n \in \mathbb{N}$ we have:*

$$\int_{\mathbb{R}^d} |x|^q dM_n(t)(x) = \mathbb{E}_{\xi_n} (|\gamma(t)|^q) \leq C \quad \forall t \in [0, T], \quad (4.38)$$

$$d_1(M_n(t), M_n(s)) \leq C|t-s|^{1/q'} \quad \forall t, s \in [0, T]. \quad (4.39)$$

As a consequence, $M_n \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ for all $n \in \mathbb{N}$ and the sequence (M_n) is a relatively compact subset of $\mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$.

Proof. By definition, for all $t \in [0, T]$ we have that

$$\mathbb{E}_{\xi_n}(|\gamma(t)|^q) \leq 2^{q-1} \mathbb{E}_{\xi_n} \left(|\gamma(0)|^q + T^{q/q'} \int_0^T |\dot{\gamma}(t)|^q dt \right) \leq C, \quad (4.40)$$

for some constant $C > 0$, independent of n . In the second inequality above we have used that m_0 has compact support and (4.33). This proves (4.38). In order to prove (4.39), by definition of d_1 , we have that $d_1(M_n(t), M_n(s)) \leq d_q(M_n(t), M_n(s))$ and, setting $\rho_n := (e_t, e_s) \# \xi_n \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$,

$$\begin{aligned} d_q^q(M_n(t), M_n(s)) &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^q d\rho_n(x, y) = \int_{\Gamma} |\gamma(t) - \gamma(s)|^q d\xi_n(\gamma) \\ &\leq |t - s|^{q/q'} \int_{\Gamma} \int_0^T |\dot{\gamma}(t)|^q dt d\xi_n(\gamma) = |t - s|^{q/q'} \mathbb{E}_{\xi_n} \left(\int_0^T |\dot{\gamma}(t)|^q dt \right) \leq C |t - s|^{q/q'}, \end{aligned}$$

from which (4.39) follows.

Finally, relation (4.38) implies that for all $t \in [0, T]$ the set $\{M_n(t) ; n \in \mathbb{N}\}$ is relatively compact in $\mathcal{P}_1(\mathbb{R}^d)$ (see [8, Proposition 7.1.5]) and (4.39) implies that the family (M_n) is equicontinuous in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$. Therefore, the last assertion in the statement of the proposition follows from the Arzelà-Ascoli theorem. \square

Suppose that $\epsilon_n = O(1/(N_n^t \log(N_n^s)))$ and let $\xi^* \in \mathcal{P}(\Gamma)$ be a limit point of (ξ_n) (by Proposition 4.4.1 there exists at least one) and, for notational convenience, we still label by $n \in \mathbb{N}$ a subsequence of (ξ_n) narrowly converging to ξ^* . By Proposition 4.4.2, we have that (M_n) converges to $m(\cdot) := e_{(\cdot)} \# \xi^*$ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$. We now examine the limit behaviour of the corresponding optimal discrete costs (U_n) . In Proposition 4.4.3 we prove that (U_n) converges, in a suitable sense, to a viscosity solution of

$$\begin{aligned} -\partial_t u + \frac{1}{q'} |\nabla u(x, t)|^{q'} &= f(x, m(t)) \quad x \in \mathbb{R}^d, \quad t \in (0, T), \\ u(x, T) &= g(x, m(T)) \quad x \in \mathbb{R}^d. \end{aligned} \quad (4.41)$$

Classical results imply that under **(H)**(i) equation (4.41) admits at most one viscosity solution (see e.g. [48, Theorem 2.1]). In [15, Proposition 1.3 and Remark 1.1] the existence of a viscosity solution u is proved, as well the following representation formula: for all $(x, t) \in \mathbb{R}^d \times (0, T)$

$$u(x, t) = \inf \left\{ \int_t^T \left[\frac{1}{q} |\dot{z}(s)|^q + f(z(s), m(s)) \right] ds + g(z(T), m(T)) \mid z \in W^{1,q}([0, T]; \mathbb{R}^d), \quad z(t) = x \right\}. \quad (4.42)$$

Standard arguments using the expression (4.42) show that u is continuous in $\mathbb{R}^d \times [0, T]$ (see e.g. [15, Theorem 2.1]).

Remark 4.4.3. Definition 4.4.1 can thus be rephrased as follows: $\xi^* \in \mathcal{P}_{m_0}(\Gamma)$ is a MFG equilibrium for (4.24) if $[0, T] \ni t \mapsto m(t) := e_t \# \xi^*$ belongs to $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ and for ξ^* -almost all γ we have that

$$u(\gamma(0), 0) = \int_0^T \left[\frac{1}{q} |\dot{\gamma}(t)|^q + f(\gamma(t), m(t)) \right] dt + g(\gamma(T), m(T)),$$

where u is the unique viscosity solution to (4.41).

In order to prove the convergence of U_n to u , we will need the following auxiliary functions

$$U^*(x, t) := \limsup_{\substack{n \rightarrow \infty \\ S_n \ni y \rightarrow x \\ T_n \ni s \rightarrow t}} U_n(y, s), \quad U_*(x, t) := \liminf_{\substack{n \rightarrow \infty \\ S_n \ni y \rightarrow x \\ T_n \ni s \rightarrow t}} U_n(y, s) \quad \forall x \in \mathbb{R}^d, \quad t \in [0, T]. \quad (4.43)$$

By (4.4.1), the functions U^* and U_* are well defined if $\epsilon_n = O(1/(N_n^t \log(N_n^s)))$. In some of the next results, we will need to assume a stronger hypothesis on ϵ_n , namely $\epsilon_n = o(1/(N_n^t \log(N_n^s)))$, which will allow us to eliminate the entropy term in the limit.

Before proving the convergence of the value functions, we will need a preliminary result.

Lemma 4.4.3. *Assume that $\epsilon_n = O\left(\frac{1}{N_n^t \log(N_n^s)}\right)$. Then,*

(i) U^* and U_* are upper and lower semicontinuous, respectively.

(ii) If in addition, $\epsilon_n = o\left(\frac{1}{N_n^t \log(N_n^s)}\right)$, we have that $U^*(x, T) = U_*(x, T) = g(x, m(T))$ for all $x \in \mathbb{R}^d$.

Proof. The proof of assertion (i) is the same than the proof of [14, Chapter V, Lemma 1.5]. Let us prove (ii). For $n \in \mathbb{N}$, let $x^n \in \mathcal{S}_n$, $t^n \in \mathcal{T}_n$ and $k : \mathbb{N} \rightarrow \mathbb{N}$ such that $t^n = t_{k(n)}$ (recall that $\mathcal{T}_n = \{0, t_1, \dots, t_{N_n^t}\}$). Because of our assumption on ϵ_n , we can write

$$U_n(x^n, t^n) = \sum_{\gamma \in \Gamma_{x^n, t^n}^{\mathcal{S}_n, \mathcal{T}_n}} p_{\hat{P}_n}^{x^n, t^n}(\gamma) \left(\sum_{k=k(n)}^{N_n^t-1} \Delta t_n \left[\frac{1}{q} \left| \frac{\gamma(t_{k+1}) - \gamma(t_k)}{\Delta t_n} \right|^q + f(\gamma(t_k), M_n(t_k)) \right] + g(\gamma(T), M_n(T)) \right) + o(1), \quad (4.44)$$

where we recall that $p_{\hat{P}_n}^{x^n, t^n}$ is defined in (4.29). Using the definition of U_n and arguing as in the proof of Lemma 4.4.1, we have that

$$\sum_{k=k(n)}^{N_n^t-1} \Delta t_n f(\gamma(t_k), M_n(t_k)) = O(T - t_{k(n)}),$$

$$U_n(x^n, t^n) \leq g(x^n, M_n(T)) + O(T - t^n) + o(1).$$

Therefore, if $x^n \rightarrow x \in \mathbb{R}^d$ and $t^n \rightarrow T$, we have

$$\limsup_{n \rightarrow \infty} U_n(x^n, t^n) \leq g(x, m(T)),$$

from which we deduce that $U^*(x, T) \leq g(x, M(T))$ for all $x \in \mathbb{R}^d$. Next, for every $\gamma \in \Gamma_{x^n, t^n}^{\mathcal{S}_n, \mathcal{T}_n}$ we have

$$|\gamma(T) - x_n|^q \leq \left(\sum_{k=k(n)}^{N_n^t-1} |\gamma(t_{k+1}) - \gamma(t_k)| \right)^q \leq (N_n^t - k(n))^{q-1} \sum_{k=k(n)}^{N_n^t-1} |\gamma(t_{k+1}) - \gamma(t_k)|^q,$$

which implies that

$$\sum_{k=k(n)}^{N_n^t-1} \frac{\Delta t_n}{q} \left| \frac{\gamma(t_{k+1}) - \gamma(t_k)}{\Delta t_n} \right|^q \geq \frac{\Delta t_n}{q(N_n^t - k(n))^{q-1}} \left| \frac{\gamma(T) - x_n}{\Delta t_n} \right|^q = \frac{1}{q(T - t^n)^{q-1}} |\gamma(T) - x_n|^q. \quad (4.45)$$

Thus, setting $p_{T,y}^{x^n, t^n} := \xi_{\hat{P}_n}^{x^n, t^n}(\{\gamma \in \Gamma \mid \gamma(T) = y\})$, equation (4.44) and the last inequality above yield

$$\begin{aligned} U_n(x^n, t^n) &\geq \sum_{y \in \mathcal{S}_n} p_{T,y}^{x^n, t^n} \left(\frac{|y - x^n|^q}{q(T - t^n)^{q-1}} + g(y, M_n(T)) \right) + O(T - t^n) + o(1) \\ &\geq \min_{y \in \mathcal{S}_n} \left\{ \frac{|y - x^n|^q}{q(T - t^n)^{q-1}} + g(y, M_n(T)) \right\} + O(T - t^n) + o(1). \end{aligned} \quad (4.46)$$

Suppose that y_n^* minimizes the “min” term in the last line above. By definition, we have

$$\frac{|y_n^* - x^n|^q}{q(T - t^n)^{q-1}} \leq g(x^n, M_n(T)) - g(y_n^*, M_n(T)) \leq C |y_n^* - x^n|,$$

where the last inequality follows from (4.25). As a consequence, we get that $|y_n^* - x^n| = O(T - t^n)$ and so $\frac{|y_n^* - x^n|^q}{(T - t^n)^{q-1}} \rightarrow 0$ as $n \rightarrow \infty$. Therefore, as $n \rightarrow \infty$,

$$\min_{y \in \mathcal{S}_n} \left\{ \frac{|y - x^n|^q}{q(T - t^n)^{q-1}} + g(y, M_n(T)) \right\} = \frac{|y_n^* - x^n|^q}{q(T - t^n)^{q-1}} + g(y_n^*, M_n(T)) \rightarrow g(x, m(T)).$$

By (4.46), this implies that

$$\liminf_{n \rightarrow \infty} U_n(x^n, t^n) \geq g(x, m(T)),$$

from which we deduce that $U_*(x, T) \geq g(x, m(T))$. The result follows. \square

Now, we prove the convergence of the sequence (U_n) . The argument of the proof uses some ideas from the theory of approximation of viscosity solutions (see e.g. [17]).

Proposition 4.4.3. *Assume that, as $n \rightarrow \infty$, $N_n^t/N_n^s \rightarrow 0$ and $\epsilon_n = o\left(\frac{1}{N_n^t \log(N_n^s)}\right)$. Then, $U^* = U_* = u$, where u is given by (4.42), or equivalently, where u is the unique continuous viscosity solution to (4.41). As a consequence, for every compact set $Q \subseteq \mathbb{R}^d$ we have that*

$$\sup_{(x,t) \in (\mathcal{S}_n \cap Q) \times \mathcal{T}_n} |U_n(x, t) - u(x, t)| \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (4.47)$$

Proof. Let us prove that U^* is a viscosity subsolution of equation (4.41). Let $\phi \in \mathcal{C}^1(\mathbb{R}^d \times [0, T])$ and $(x^*, t^*) \in \mathbb{R}^d \times (0, T)$ be such that (x^*, t^*) is a local maximum of $U^* - \phi$ on $\mathbb{R}^d \times (0, T)$.

By standard arguments in the theory of viscosity solutions (see e.g. [14, Chapter II]), we may assume that ϕ is bounded as well as its time and space derivatives and that (x^*, t^*) is a strict global maximum of $U^* - \phi$. Arguing as in the proof of [14, Chapter V, Lemma 1.6], we can show the existence of a sequence (x^n, t^n) in $\mathcal{S}_n \times \mathcal{T}_n$ such that $(x^n, t^n) \rightarrow (x^*, t^*)$, $U_n(x^n, t^n) \rightarrow U^*(x^*, t^*)$ and $U_n - \phi$ has maximum at (x^n, t^n) in the set $(\mathcal{S}_n \times \mathcal{T}_n) \cap B_\delta$, where $B_\delta := \{(x, t) \in \mathbb{R}^d \times (0, T) ; |x - x^*| + |t - t^*| \leq \delta\}$ and $\delta > 0$ is such that $B_\delta \subseteq \mathbb{R}^d \times (0, T)$.

Now, let $\xi \in C^\infty(\mathbb{R}^d \times [0, T])$ be such that $0 \leq \xi \leq 1$, $\xi(x, t) = 0$ if $(x, t) \in B_{\frac{\delta}{2}}$ and $\xi(x, t) = 1$ if $(x, t) \in (\mathbb{R}^d \times (0, T)) \setminus B_\delta$. Then, using that U_n and ϕ are bounded, we can choose $M > 0$ large enough such that, setting $\bar{\phi} := \phi + M\xi$, the function $U_n - \bar{\phi}$ has maximum in $\mathcal{S}_n \times \mathcal{T}_n$ at the point (x^n, t^n) . Note that $\partial_t \bar{\phi}(x^*, t^*) = \partial_t \phi(x^*, t^*)$ and $\nabla \bar{\phi}(x^*, t^*) = \nabla \phi(x^*, t^*)$.

As in the proof of Lemma 4.4.3, let $k : \mathbb{N} \rightarrow \mathbb{N}$ be such that $t^n = t_{k(n)}$. Since $U_n(x^n, t^n)$ satisfies

$$U_n(x^n, t^n) = \min_{p \in \mathcal{P}(\mathcal{S}_n)} \sum_{y \in \mathcal{S}_n} p(y) \left(\frac{\Delta t_n}{q} \left| \frac{y - x^n}{\Delta t_n} \right|^q + \Delta t_n f(x^n, M_n(t^n)) + U_n(y, t_{k(n)+1}) \right) + \epsilon_n \mathcal{E}_n(p),$$

and $U_n(y, t_{k(n)+1}) - U_n(x^n, t^n) \leq \bar{\phi}(y, t_{k(n)+1}) - \bar{\phi}(x^n, t^n)$ for all $y \in \mathcal{S}_n$, we have that

$$\begin{aligned} 0 &\leq \min_{p \in \mathcal{P}(\mathcal{S}_n)} \sum_{y \in \mathcal{S}_n} p(y) \left(\frac{\Delta t_n}{q} \left| \frac{y - x^n}{\Delta t_n} \right|^q + \Delta t_n f(x^n, M_n(t_{k(n)})) + \bar{\phi}(y, t_{k(n)+1}) - \bar{\phi}(x^n, t_{k(n)}) \right) + \epsilon_n \mathcal{E}_n(p), \\ &\leq \min_{y \in \mathcal{S}_n} \left\{ \frac{\Delta t_n}{q} \left| \frac{y - x^n}{\Delta t_n} \right|^q + \Delta t_n f(x^n, M_n(t_{k(n)})) + \bar{\phi}(y, t_{k(n)+1}) - \bar{\phi}(x^n, t_{k(n)}) \right\} + \epsilon_n \mathcal{E}_n(p), \end{aligned} \quad (4.48)$$

where the second inequality follows from the first one by taking for each $y \in \mathcal{S}_n$ the vector $p \in \mathcal{P}(\mathcal{S}_n)$ defined as $p(z) = 1$ iff $z = y$. Dividing by Δt_n and recalling that $\epsilon_n = o\left(\frac{1}{N_n^t \log(N_n^s)}\right)$, we get

$$0 \leq f(x^n, M_n(t^n)) + \min_{y \in \mathcal{S}_n} \left\{ \frac{1}{q} \left| \frac{y - x^n}{\Delta t_n} \right|^q + \frac{\bar{\phi}(y, t_{k(n)+1}) - \bar{\phi}(x^n, t_{k(n)})}{\Delta t_n} \right\} + o(1),$$

and so, taking liminf,

$$0 \leq f(x^*, m(t^*)) + \liminf_n \min_{y \in \mathcal{S}_n} \left\{ \frac{1}{q} \left| \frac{y - x^n}{\Delta t_n} \right|^q + \frac{\bar{\phi}(y, t_{k(n)+1}) - \bar{\phi}(x^n, t_{k(n)})}{\Delta t_n} \right\}, \quad (4.49)$$

where we have used that $M_n \rightarrow m$ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$. Let us study the second term in the right hand side above. For fixed n , let y_n^* be such that

$$y_n^* \in \operatorname{argmin}_{y \in \mathcal{S}_n} \left\{ \frac{1}{q} \left| \frac{y - x^n}{\Delta t_n} \right|^q + \frac{\bar{\phi}(y, t_{k(n)+1}) - \bar{\phi}(x^n, t_{k(n)})}{\Delta t_n} \right\},$$

or equivalently, setting $\alpha_n^* := \frac{y_n^* - x^n}{\Delta t_n}$,

$$\frac{1}{q} |\alpha_n^*|^q + \frac{\bar{\phi}(x^n + \Delta t_n \alpha_n^*, t_{k(n)+1}) - \bar{\phi}(x^n, t_{k(n)})}{\Delta t_n} \leq \frac{1}{q} \left| \frac{y - x^n}{\Delta t_n} \right|^q + \frac{\bar{\phi}(y, t_{k(n)+1}) - \bar{\phi}(x^n, t_{k(n)})}{\Delta t_n}, \quad (4.50)$$

for all $y \in \mathcal{S}_n$. By taking $y = x^n$ in the expression above and using that $\partial_t \bar{\phi}$ and $\nabla \bar{\phi}$ are bounded, we obtain that the sequence (α_n^*) is bounded. Let α^* be a limit point of this sequence and consider a subsequence of (α_n) , still indexed by n , such that $\alpha_n^* \rightarrow \alpha^*$. The condition $N_n^t/N_n^s \rightarrow 0$ implies that for any $\alpha \in \mathbb{R}^d$ we can find a sequence (y^n) in \mathcal{S}_n such that $\frac{y^n - x^n}{\Delta t_n} \rightarrow \alpha$ as $n \rightarrow \infty$. Taking $y = y^n$ in (4.50) and passing to the limit yields

$$\frac{1}{q} |\alpha^*|^q + \nabla \phi(x^*, t^*) \cdot \alpha^* \leq \frac{1}{q} |\alpha|^q + \nabla \phi(x^*, t^*) \cdot \alpha \quad \forall \alpha \in \mathbb{R}^d. \quad (4.51)$$

which implies that

$$\frac{1}{q} |\alpha^*|^q + \nabla \phi(x^*, t^*) \cdot \alpha^* = -\frac{1}{q'} |\nabla \phi(x^*, t^*)|^{q'}.$$

Passing to the limit in (4.49) gives

$$-\partial_t \phi(x^*, t^*) + \frac{1}{q'} |\nabla \phi(x^*, t^*)|^{q'} \leq f(x^*, m(t^*)),$$

which proves that U^* is a subsolution to (4.41). Similarly, we can prove that U_* is a supersolution to (4.41). Assumption **(H3)**(i) ensures a comparison principle for (4.41) (see [48, Theorem 2.1]). Therefore, since $U^*(\cdot, T) = U_*(\cdot, T)$ by Lemma 4.4.3(ii), we have that $U^* = U_* = u$ as announced. Using this result, the proof of (4.47) is identical to the proof of [14, Chapter V, Lemma 1.9]. \square

We have now all the elements to prove the main result in this article. We will need an additional assumption over f and g .

There exists $C > 0$ and a modulus of continuity $\omega : [0, +\infty) \rightarrow [0, +\infty)$ such that for $h = f, g$

$$|h(x, m) - h(x, m')| \leq C(1 + |x|^q) \omega(d_1(m, m')) \quad \forall x \in \mathbb{R}^d, m, m' \in \mathcal{P}_1(\mathbb{R}^d). \quad (4.52)$$

Theorem 4.4.1. *Suppose that **(H3)** and (4.52) hold and, as $n \rightarrow \infty$, $N_n^t/N_n^s \rightarrow 0$ and $\epsilon_n = o\left(\frac{1}{N_n^t \log(N_n^s)}\right)$. Then, the following assertions hold true:*

- (i) *There exists at least one limit point ξ^* of (ξ_n) , with respect to the narrow topology in $\mathcal{P}(\Gamma)$, and every such limit point is a MFG equilibrium for (4.24).*
- (ii) *Consider any converging subsequence of $(\xi_{n'})$ of (ξ_n) , with limit $\xi^* \in \mathcal{P}(\Gamma)$, and let $(U_{n'}, M_{n'})$ be the associated solutions to (4.27). Denote by u be the unique viscosity solution to (4.41) with $m(t) := e_t \# \xi^*$ for all $t \in [0, T]$. Then, the sequence $(M_{n'}) \subseteq C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$, defined by (4.31), converge to m in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ and (4.47) holds for $(U_{n'})$ and u .*

Proof. Assertion (ii) is a straightforward consequence of the first assertion and Proposition 4.4.3, hence, we only need to prove (i). The existence of at least one limit point ξ^* of (ξ_n) is a consequence of Proposition 4.4.1. Let us still index by n a subsequence of (ξ_n) narrowly converging to ξ^* . By Proposition

4.4.2, we have that $m(\cdot) := e_{(\cdot)} \sharp \xi^*$ is the limit in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ of M_n . By definition of ξ_n and our condition over ϵ_n , we have

$$\mathbb{E}_{\xi_n} \left(\int_0^T \left[\frac{1}{q} |\dot{\gamma}(t)|^q + f(\gamma([t]_{\mathcal{T}_n}), M_n([t]_{\mathcal{T}_n})) \right] dt + g(\gamma(T), M_n(T)) \right) + o(1) = \sum_{x \in \mathcal{S}_n} U_n(x, 0) M_{n,0}(x), \quad (4.53)$$

where $[t]_{\mathcal{T}_n}$ is the greatest element in \mathcal{T}_n not larger than t . Using that the support of $M_{n,0}$ is uniformly bounded and relation (4.47) in Proposition 4.4.3, we easily get that the right hand side above converges to $\int_{\mathbb{R}^d} u(x, 0) dm_0(x) = \mathbb{E}_{\xi^*}(u(\gamma(0), 0))$, where u is the unique viscosity solution to (4.41). On the other hand, arguing as in the proof of Lemma 4.4.2 we obtain that the mapping

$$\Gamma \ni \gamma \mapsto \begin{cases} \int_0^T \frac{1}{q} |\dot{\gamma}(t)|^q dt, & \text{if } \gamma \in W^{1,q}([0, T]; \mathbb{R}^d), \\ +\infty & \text{otherwise,} \end{cases}$$

is lower semicontinuous. Therefore, by [8, Lemma 5.1.7] and (4.33), we have

$$\mathbb{E}_{\xi^*} \left(\int_0^T \frac{1}{q} |\dot{\gamma}(t)|^q dt \right) \leq \liminf_n \mathbb{E}_{\xi_n} \left(\int_0^T \frac{1}{q} |\dot{\gamma}(t)|^q dt \right) < \infty, \quad (4.54)$$

which, together with (4.38), implies that the support of ξ^* is contained in $W^{1,q}([0, T]; \mathbb{R}^d)$. By assumption **(H3)**(i), for all $k = 0, \dots, N_n^t - 1$ we have that

$$\left| \mathbb{E}_{\xi_n} \left(\int_{t_k}^{t_{k+1}} [f(\gamma(t_k), M_n(t_k)) - f(\gamma(t), M_n(t_k))] dt \right) \right| \leq C \mathbb{E}_{\xi_n} \left(\int_{t_k}^{t_{k+1}} |\gamma(t) - \gamma(t_k)| dt \right). \quad (4.55)$$

Since $\gamma(t) = \gamma(t_k) + \dot{\gamma}(t)(t - t_k)$ for ξ_n -almost all γ and all $t \in (t_k, t_{k+1})$, the bound (4.33) gives

$$\mathbb{E}_{\xi_n} \left(\int_{t_k}^{t_{k+1}} |\gamma(t) - \gamma(t_k)| dt \right) = \Delta t_n (\Delta t_n)^{\frac{1}{q'}} \left[\mathbb{E}_{\xi_n} \left(\int_0^T |\dot{\gamma}(t)|^q dt \right) \right]^{\frac{1}{q}} \leq C (\Delta t_n)^{1 + \frac{1}{q'}},$$

for some constant $C > 0$. Thus, by (4.55),

$$\mathbb{E}_{\xi_n} \left(\int_0^T f(\gamma([t]_{\mathcal{T}_n}), M_n([t]_{\mathcal{T}_n})) dt \right) = \mathbb{E}_{\xi_n} \left(\int_0^T f(\gamma(t), M_n([t]_{\mathcal{T}_n})) dt \right) + o(1).$$

The relation above and (4.52) yield

$$\begin{aligned} \mathbb{E}_{\xi_n} \left(\int_0^T f(\gamma([t]_{\mathcal{T}_n}), M_n([t]_{\mathcal{T}_n})) dt \right) &= \mathbb{E}_{\xi_n} \left(\int_0^T f(\gamma(t), m(t)) dt \right) \\ &\quad + C \left(1 + \sup_{t \in [0, T]} \mathbb{E}_{\xi_n}(|\gamma(t)|^q) \right) \sup_{t \in [0, T]} d_1(M_n([t]_{\mathcal{T}_n}), m(t)) \\ &\quad + o(1) \\ &= \mathbb{E}_{\xi_n} \left(\int_0^T f(\gamma(t), m(t)) dt \right) + o(1), \end{aligned} \quad (4.56)$$

where, in the last equality, we have used (4.38) and the fact that $M_n \rightarrow m$ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$. Analogously,

$$\mathbb{E}_{\xi_n}(g(\gamma(T), M_n(T))) = \mathbb{E}_{\xi_n}(g(\gamma(T), m(T))) + o(1). \quad (4.57)$$

Therefore, passing to the limit $n \rightarrow \infty$ in (4.53) and using (4.54), (4.56) and (4.57), we get

$$\mathbb{E}_{\xi^*} \left(\int_0^T \left[\frac{1}{q} |\dot{\gamma}(t)|^q + f(\gamma(t), m(t)) \right] dt + g(\gamma(T), m(T)) \right) \leq \mathbb{E}_{\xi^*}(u(\gamma(0), 0)). \quad (4.58)$$

Using that, by definition,

$$u(\gamma(0), 0) \leq \int_0^T \left[\frac{1}{q} |\dot{\gamma}(t)|^q + f(\gamma(t), m(t)) \right] dt + g(\gamma(T), m(T)) \quad \forall \gamma \in W^{1,q}([0, T]; \mathbb{R}^d),$$

inequality (4.58) implies that for ξ^* -almost all γ we have that

$$u(\gamma(0), 0) = \int_0^T \left[\frac{1}{q} |\dot{\gamma}(t)|^q + f(\gamma(t), m(t)) \right] dt + g(\gamma(T), m(T)),$$

i.e. ξ^* is a MFG equilibrium for (4.24) (see Remark 4.4.3). □

References

- [1] Y. Achdou, G. Barles, H. Ishii, and G. L. Litvinov. *Hamilton-Jacobi equations: approximations, numerical analysis and applications*. Springer, 2013. [10](#)
- [2] Y. Achdou, F. J. Buera, J.-M. Lasry, P.-L. Lions, and B. Moll. Partial differential equation models in macroeconomics. *Phil. Trans. R. Soc. A*, 372(2028):20130397, 2014. [30](#)
- [3] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games: numerical methods for the planning problem. *SIAM Journal on Control and Optimization*, 50(1):77–109, 2012. [10](#)
- [4] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games: convergence of a finite difference method. *SIAM Journal on Numerical Analysis*, 51(5):2585–2612, 2013. [10](#)
- [5] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: Numerical methods. *SIAM Journal on Numerical Analysis*, 48(3):1136–1162, 2010. [10](#)
- [6] S. R. Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–684, 1994. [29](#), [30](#)
- [7] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields and applications. *Inventiones mathematicae*, 158(2):227–260, 2004. [50](#)
- [8] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008. [48](#), [51](#), [62](#), [72](#), [79](#), [88](#), [92](#)
- [9] J.-P. Aubin and H. Frankowska. *Set-valued analysis*. Springer Science & Business Media, 2009. [56](#)
- [10] P. Auer, N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002. [9](#)
- [11] R. J. Aumann. Markets with a continuum of traders. *Econometrica: Journal of the Econometric Society*, 32(1/2):39–50, 1964. [3](#)
- [12] R. J. Aumann. Existence of competitive equilibria in markets with a continuum of traders. *Econometrica: Journal of the Econometric Society*, 34:1–17, 1966. [3](#)
- [13] M. Bardi. Explicit solutions of some linear-quadratic mean field games. *Networks and heterogeneous media*, 7(2):243–261, 2012.
- [14] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia. [89](#), [90](#), [91](#)
- [15] M. Bardi and F. Da Lio. On the Bellman equation for some unbounded control problems. *NoDEA Nonlinear Differential Equations Appl.*, 4(4):491–510, 1997. [88](#)

- [16] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*. Springer, 1994. [10](#)
- [17] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991. [90](#)
- [18] M. Benaïm. Dynamics of stochastic approximation algorithms. *Séminaire de Probabilités, XXXIII*, 1709:1–68, 1999. [7](#)
- [19] M. Benaïm, J. Hofbauer, and S. Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 44(1):328–348, 2005. [5](#), [7](#)
- [20] M. Benaïm, J. Hofbauer, and S. Sorin. Stochastic approximations and differential inclusions, part ii: Applications. *Mathematics of Operations Research*, 31(4):673–695, 2006. [5](#), [7](#)
- [21] J.-D. Benamou and G. Carlier. Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *Journal of Optimization Theory and Applications*, 167(1):1–26, 2015. [10](#)
- [22] J.-D. Benamou, G. Carlier, and F. Santambrogio. Variational mean field games. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 141–171. Birkhäuser/Springer, Cham, 2017. [84](#)
- [23] A. Blanchet and G. Carlier. From Nash to Cournot–Nash equilibria via the Monge–Kantorovich problem. *Phil. Trans. R. Soc. A*, 372(2028):20130398, 2014. [4](#)
- [24] A. Blanchet and G. Carlier. Remarks on existence and uniqueness of Cournot–Nash equilibria in the non-potential case. *Mathematics and Financial Economics*, 8(4):417–433, 2014. [4](#), [22](#), [54](#), [58](#)
- [25] A. Blanchet and G. Carlier. Optimal transport and Cournot–Nash equilibria. *Mathematics of Operations Research*, 41(1):125–145, 2015. [4](#), [22](#)
- [26] G. W. Brown. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951. [iii](#), [v](#), [5](#), [30](#), [54](#), [75](#)
- [27] S. Cacace, F. Camilli, and C. Marchi. A numerical method for Mean Field Games on networks. *ESAIM: Mathematical Modelling and Numerical Analysis*, 51(1):63–88, 2017. [10](#)
- [28] F. Camilli, E. Carlini, and C. Marchi. A model problem for mean field games on networks. *arXiv preprint arXiv:1402.2467*, 2014.
- [29] F. Camilli and C. Marchi. Stationary mean field games systems defined on networks. *SIAM Journal on Control and Optimization*, 54(2):1085–1103, 2016.
- [30] F. Camilli and F. Silva. A semi-discrete in time approximation for a model 1st order-finite horizon mean field game problem. *Networks and heterogeneous media*, 7(2):263–277, 2012. [10](#)
- [31] P. Cannarsa and R. Capuani. Existence and uniqueness for Mean Field Games with state constraints. *Preprint*, 2017. [84](#)
- [32] P. Cannarsa and C. Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, volume 58. Springer Science & Business Media, 2004. [10](#), [13](#), [14](#), [15](#), [16](#), [17](#), [43](#), [47](#), [49](#), [50](#), [67](#)
- [33] P. Cardaliaguet. Notes on mean field games. Technical report, 2010. [10](#), [13](#), [17](#), [75](#), [80](#), [84](#)

- [34] P. Cardaliaguet. Long time average of first order mean field games and weak KAM theory. *Dynamic Games and Applications*, 3(4):473–488, 2013. [10](#), [41](#), [43](#), [46](#), [50](#)
- [35] P. Cardaliaguet. Weak solutions for first order mean field games with local coupling. In *Analysis and geometry in control theory and its applications*, pages 111–158. Springer, 2015. [10](#), [31](#), [43](#)
- [36] P. Cardaliaguet, G. Carlier, and B. Nazaret. Geodesics for a class of distances in the space of probability measures. *Calculus of Variations and Partial Differential Equations*, pages 1–26, 2012. [43](#)
- [37] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. The master equation and the convergence problem in mean field games. *arXiv preprint arXiv:1509.02505*, 2015. [10](#), [19](#), [30](#), [31](#), [33](#), [75](#)
- [38] P. Cardaliaguet, P. J. Graber, A. Porretta, and D. Tonon. Second order mean field games with degenerate diffusion and local coupling. *Nonlinear Differential Equations and Applications NoDEA*, 22(5):1287–1317, 2015. [10](#), [31](#)
- [39] P. Cardaliaguet and S. Hadikhanloo. Learning in mean field games: The fictitious play. *ESAIM: Control, Optimisation and Calculus of Variations*, 23(2):569–591, 2017. [54](#), [75](#), [76](#), [84](#)
- [40] E. Carlini and F. J. S. Álvarez. A semi-Lagrangian scheme for a degenerate second order mean field game system. *arXiv preprint arXiv:1404.5932*, 2014. [10](#)
- [41] E. Carlini and F. J. Silva. A fully discrete semi-Lagrangian scheme for a first order mean field game problem. *SIAM Journal on Numerical Analysis*, 52(1):45–67, 2014. [10](#), [76](#)
- [42] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. *SIAM Journal on Control and Optimization*, 51(4):2705–2734, 2013. [10](#), [30](#)
- [43] N. Cesa-Bianchi. Analysis of two gradient-based algorithms for on-line regression. In *Proceedings of the tenth annual conference on Computational learning theory*, pages 163–170. ACM, 1997. [9](#)
- [44] N. Cesa-Bianchi and G. Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006. [5](#), [7](#), [9](#)
- [45] M. G. Crandall, L. C. Evans, and P.-L. Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 282(2):487–502, 1984. [14](#)
- [46] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992. [10](#)
- [47] M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277(1):1–42, 1983. [14](#)
- [48] F. Da Lio and O. Ley. Convex Hamilton-Jacobi equations under superlinear growth conditions on data. *Appl. Math. Optim.*, 63(3):309–339, 2011. [88](#), [91](#)
- [49] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Inventiones mathematicae*, 98(3):511–547, 1989. [10](#), [50](#)
- [50] L. C. Evans. *Partial differential equations*. American Mathematical Society, Providence, R.I., 2010. [14](#)
- [51] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006. [10](#), [15](#)

- [52] D. P. Foster and R. V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1-2):40–55, 1997. [10](#)
- [53] D. Fudenberg and D. K. Levine. *The theory of learning in games*. MIT press, 1998. [5](#), [30](#), [54](#), [75](#)
- [54] D. Fudenberg and J. Tirole. Game theory. Technical report, The MIT press, 1991. [1](#)
- [55] D. A. Gomes, J. Mohr, and R. R. Souza. Discrete time, finite state space mean field games. *Journal de mathématiques pures et appliquées*, 93(3):308–328, 2010. [iii](#), [v](#), [10](#), [18](#), [24](#), [25](#), [75](#), [76](#), [77](#), [78](#), [82](#)
- [56] D. A. Gomes, E. A. Pimentel, and V. Voskanyan. *Regularity theory for mean-field game systems*. SpringerBriefs in Mathematics. Springer, [Cham], 2016. [75](#)
- [57] D. A. Gomes and J. Saúde. Mean field games models—a brief survey. *Dyn. Games Appl.*, 4(2):110–154, 2014. [75](#)
- [58] O. Guéant. Mean field games equations with quadratic hamiltonian: a specific approach. *Mathematical Models and Methods in Applied Sciences*, 22(09):1250022, 2012. [10](#), [31](#)
- [59] O. Guéant, J.-M. Lasry, and P.-L. Lions. Mean field games and applications. *Paris-Princeton lectures on mathematical finance 2010*, pages 205–266, 2011. [10](#), [30](#), [75](#)
- [60] J. Hannon. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957. [7](#)
- [61] S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000. [10](#)
- [62] S. Hart and A. Mas-Colell. A general class of adaptive strategies. *Journal of Economic Theory*, 98(1):26–54, 2001. [10](#)
- [63] S. Hart and A. Mas-Colell. A reinforcement procedure leading to correlated equilibrium. In *Economics Essays*, pages 181–200. Springer, 2001. [10](#)
- [64] S. Hart and A. Mas-Colell. Regret-based continuous-time dynamics. *Games and Economic Behavior*, 45(2):375–394, 2003. [10](#)
- [65] J. Hofbauer and W. H. Sandholm. Stable games and their dynamics. *Journal of Economic theory*, 144(4):1665–1693, 2009. [4](#), [5](#), [54](#), [55](#), [58](#), [60](#)
- [66] M. Huang, P. E. Caines, and R. P. Malhamé. Individual and mass behaviour in large population stochastic wireless power control problems: centralized and nash equilibrium solutions. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 1, pages 98–103. IEEE, 2003. [30](#), [53](#)
- [67] 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. *Communications in Information & Systems*, 6(3):221–252, 2006. [10](#), [29](#), [30](#), [53](#), [75](#)
- [68] A. Kalai and S. Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005. [8](#)
- [69] A. C. Kizilkale and P. E. Caines. Mean field stochastic adaptive control. *IEEE Transactions on Automatic Control*, 58(4):905–920, 2013.

- [70] O. A. Ladyzhenskaia, V. A. Solonnikov, and N. N. Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988. [41](#), [53](#)
- [71] R. Laraki, J. Renault, and S. Sorin. *Bases Mathématiques de la théorie des jeux*. Editions de l'Ecole Polytechnique, 2013. [1](#)
- [72] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. i—le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006. [4](#), [10](#), [29](#), [75](#)
- [73] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. ii—horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343(10):679–684, 2006. [4](#), [10](#), [29](#), [53](#), [54](#), [58](#), [68](#), [75](#)
- [74] J.-M. Lasry and P.-L. Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007. [iii](#), [v](#), [10](#), [13](#), [18](#), [19](#), [29](#), [30](#), [31](#), [34](#), [35](#), [41](#), [53](#), [54](#), [75](#), [76](#), [84](#)
- [75] P.-L. Lions. *Cours au Collège de France*. www.college-de-france.fr, 2007-2008. [10](#), [75](#)
- [76] A. Mas-Colell. On a theorem of Schmeidler. *Journal of Mathematical Economics*, 13(3):201–206, 1984. [3](#), [22](#), [54](#), [56](#)
- [77] P. Mertikopoulos. Learning in concave games with imperfect information. *arXiv preprint arXiv:1608.07310*, 2016. [5](#), [10](#), [23](#), [54](#)
- [78] K. Miyasawa. On the convergence of the learning process in a 2 x 2 non-zero-sum two-person game. Technical report, Economic Research Program, Princeton University, Research Memorandum, 1961. [5](#), [30](#), [54](#)
- [79] D. Monderer and L. S. Shapley. Fictitious play property for games with identical interests. *Journal of economic theory*, 68(1):258–265, 1996. [5](#), [6](#), [30](#), [31](#), [35](#), [39](#), [40](#), [54](#)
- [80] D. Monderer and L. S. Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996. [iii](#), [v](#), [2](#), [18](#), [31](#)
- [81] J. Nash. Non-cooperative games. *Annals of mathematics*, 54:286–295, 1951. [2](#)
- [82] A. Nemirovskii, D. B. Yudin, and E. R. Dawson. *Problem complexity and method efficiency in optimization*. Wiley, 1983. [9](#), [54](#)
- [83] M. Nourian, P. E. Caines, R. P. Malhamé, and M. Huang. Mean field lqg control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation. *IEEE Transactions on Automatic Control*, 57(11):2801–2816, 2012.
- [84] M. J. Osborne. *An introduction to game theory*, volume 3. Oxford university press New York, 2004. [1](#)
- [85] M. J. Osborne and A. Rubinstein. *A course in game theory*. MIT press, 1994. [1](#)
- [86] A. Porretta. Weak solutions to Fokker-Planck equations and mean field games. *Arch. Ration. Mech. Anal*, 216(1):1–62, 2015. [10](#)
- [87] J. Robinson. An iterative method of solving a game. *Annals of mathematics*, 54:296–301, 1951. [5](#), [30](#), [54](#)
- [88] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, 33(3):520–534, 1965. [4](#), [54](#)

- [89] D. Schmeidler. Equilibrium points of nonatomic games. *Journal of statistical Physics*, 7(4):295–300, 1973. [3](#), [22](#), [54](#), [56](#)
- [90] S. Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012. [7](#), [8](#), [54](#), [63](#)
- [91] L. S. Shapley. Some topics in two-person games. *Advances in game theory*, 52(1-29):1–2, 1964. [30](#)
- [92] C. Villani. *Topics in Optimal Transportation*. Vol. 58 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2003. [79](#)
- [93] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton university press, 1944. [1](#)
- [94] P. Young and S. Zamir. *Handbook of game theory*, volume 4. Elsevier, 2014. [5](#)

Résumé

Les jeux à champ moyen (MFG) sont une classe de jeux différentiels dans lequel chaque agent est infinitésimal et interagit avec une très grande population d'agents. Dans cette thèse, nous soulevons la question de la formation effective de l'équilibre MFG. En effet, le jeu étant très complexe, il est irréaliste de supposer que les agents peuvent réellement calculer la configuration d'équilibre. Cela semble indiquer que si la configuration d'équilibre se présente, c'est parce que les agents ont appris à jouer au jeu. Donc, la question principale est de trouver des procédures d'apprentissage dans les jeux à champ moyen et d'analyser leurs convergences vers un équilibre. Nous nous sommes inspirés par des schémas d'apprentissage dans les jeux statiques et avons essayé de les appliquer à notre modèle dynamique de MFG. Nous nous concentrons particulièrement sur les applications de *fictitious play* et *online mirror descent* sur différents types de jeux à champ moyen: Potentiel, Monotone et Discret.

Mots Clés

Jeux à champ moyen, contrôle optimale, fictitious play, algorithm de descente miroir.

Abstract

Mean Field Games (MFG) are a class of differential games in which each agent is infinitesimal and interacts with a huge population of other agents. In this thesis 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. Hence the main question is to find learning procedures in mean field games and investigating if they converge to an equilibrium. We have inspired from the learning schemes in static games and tried to apply them to our dynamical model of MFG. We especially focus on *fictitious play* and *online mirror descent* applications on different types of mean field games; those are either Potential, Monotone or Discrete.

Keywords

Mean field games, optimal control, fictitious play, online mirror descent algorithm.