

# Discrete Mean Field Games: Existence of Equilibria and Convergence

Josu Doncel, Nicolas Gast, Bruno Gaujal

► **To cite this version:**

Josu Doncel, Nicolas Gast, Bruno Gaujal. Discrete Mean Field Games: Existence of Equilibria and Convergence. *Journal of Dynamics and Games, AIMS*, 2019, 6 (3), pp.1-19. 10.3934/jdg.2019016 . hal-01277098v2

**HAL Id: hal-01277098**

**<https://hal.inria.fr/hal-01277098v2>**

Submitted on 3 Sep 2019

**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.

# Discrete Mean Field Games: Existence of Equilibria and Convergence

Josu Doncel<sup>a</sup>, Nicolas Gast<sup>b,c</sup>, Bruno Gaujal<sup>b,c</sup>

<sup>a</sup>*University of the Basque Country UPV/EHU, Spain*

<sup>b</sup>*Univ. Grenoble Alpes, F-38000 Grenoble, France*

<sup>c</sup>*Inria*

---

## Abstract

We consider mean field games with discrete state spaces (called discrete mean field games in the following) and we analyze these games in continuous and discrete time, over finite as well as infinite time horizons. We prove the existence of a mean field equilibrium assuming continuity of the cost and of the drift. These conditions are more general than the existing papers studying finite state space mean field games. Besides, we also study the convergence of the equilibria of  $N$ -player games to mean field equilibria in our four settings. On the one hand, we define a class of strategies in which any sequence of equilibria of the finite games converges weakly to a mean field equilibrium when the number of players goes to infinity. On the other hand, we exhibit equilibria outside this class that do not converge to mean field equilibria and for which the value of the game does not converge. In discrete time this non-convergence phenomenon implies that the Folk theorem does not scale to the mean field limit.

---

## 1. Introduction

Mean field games have been introduced by Lasry and Lions [34] as well as Huang, Caines and Malhamé [30] to model interactions between a large number of strategic agents (players) and have had a large success ever since. Since the seminal work in [32, 33, 34, 30], a large variety of papers have been investigating mean field games. Most of the literature concerns continuous state spaces and describes a mean field game as a coupling between a Hamilton-Jacobi-Bellman equation with a Fokker-Planck equation (see for example [28, 7, 9, 24, 10, 25, 22, 23, 3]). Here, we are interested in studying mean field games with a finite number of states and finite number of actions per player. In this case, the analog of the Hamilton-Jacobi-Bellman equation is the *Bellman equation* and the discrete version of the Fokker-Planck equation is the *Kolmogorov equation*.

Finite state space mean field games in *discrete time* (*a.k.a.* with synchronous players) were previously studied in [20]. In their work, the strategy of the players is the probability matrix of the Kolmogorov equation. This implies that each player can choose her dynamics independently of the state of the others: the behavior of players is only coupled via their costs. In that case, the Kolmogorov equation becomes linear.

Finite state space mean field games in *continuous time* (*a.k.a.* with asynchronous players) have also been previously analyzed in [21, 27, 5, 13]. In their model, the players also control completely the transition rate matrix so that the dynamics are again linear once the actions of the players are given. Again, players do not interact with each other directly in these models, but only through their costs.

The models we study here, both in the synchronous and asynchronous cases cover non-linear dynamics: We consider that the players do not have the power to choose the rate matrix and that their actions only have a limited effect on their state. Here, the transition rate matrix may depend not only on the actions taken by the player, but also on the population distribution of the system. This introduces an explicit interaction between the players (and not just through their costs). This non-linear dynamics is called *the relaxed case* in [14]. We claim that the model with explicit interactions covers several natural phenomena such as information/infection propagation or resource congestion where the cost but also the state dynamics of a player depend on the state of the all the others. This type of behavior is classical in systems with a large number of interacting objects [6] and cannot be handled using previous mean field game models. For instance, in the classical SIR (Susceptible, Infected, Recovered) infection model [39], the rate of infection of one individual depends on the proportion of individuals already infected. Similarly, in a model of congestion one player cannot typically use a resource if it is already used to full capacity.

We show that the only requirement needed to guarantee the existence of a Mean Field Equilibrium in mixed strategies is that the cost is continuous with respect to the population distribution (convexity is not needed). This result nicely mimics the conditions for existence of a Nash equilibrium in the simpler case of static population games (see [36]). The existence of a mean field equilibrium in mixed strategies has been previously shown by [31, 12] in the diffusion case. In [27] the existence of a Mean Field Equilibrium is proven under the assumption that the cost of a player is strictly convex w.r.t. her strategy and in [21] the authors also consider uniformly convex functions. These conditions are rather strong because they are not satisfied in the important case of linear and/or expected costs. In [14] existence of a Nash equilibrium is also proved under mere continuity assumptions and with a compact action space (more general than the simplex, used here). However, the main difference between the two approaches is the type of mean field limit that is used. In [14], the trajectories of the states of the players are considered while we

only consider the state at time  $t$ . The first approach uses arguments in line with the propagation of chaos while the second one is closer to the work in [4, 38]. While the convergence of trajectories is a more refined convergence than the point-wise convergence in general, this is useless here. Indeed, for mean field games, costs are associated to states and actions and not to trajectories. Therefore, the point-wise mean field approach is sufficient. Another difference with [14] is that an additional assumption about the uniqueness of the argmin is needed in some parts of the convergence proof as well as for existence (in the feedback case). This is not the case here, so both papers do not cover the exact same set of games.

As in most existence proofs, our proof is based on a version of the fixed point theorem of Kakutani in infinite dimension (see for example [13] where such an extended version of the fixed point theorem is used in a mean field game model with minor and major players). Here however, we do not consider the best response operator but the evolution of the population distribution instead, as in [14]. Out of the four cases (asynchronous/synchronous, finite/infinite horizons), we mainly detail the asynchronous player case for which we prove this existence of a mean field equilibrium in an infinite horizon with discounted costs. We also show, more briefly, how these results can be extended to a finite horizon or to a finite or infinite time horizon in the synchronous-player case.

Our second contribution concerns convergence of finite games to mean field limits. Different authors have studied the convergence of  $N$ -player games equilibria to mean field equilibria, *e.g.* [29, 1, 37, 38]. The type of strategies considered in these paper is different from ours: they consider that the strategy of a player only depends on her internal state (these are called *stationary policies* in [38]), whereas here we allow time dependence in these policies. The model in [38] does include state dynamics that depend on the population distribution but only considers stationary strategies that do not depend on time, hence cannot depend on the population dynamics.

In all four combinations (finite / infinite horizon, synchronous / asynchronous), a mean field equilibrium is always an  $\epsilon$ -approximation of an equilibrium of a corresponding game with a finite number  $N$  of players, where  $\epsilon$  goes to 0 when  $N$  goes to infinity. This is the discrete pending result to similar results in continuous games [11]. However, we show also that not all equilibria for the finite version converge to a Nash equilibrium of the mean field limit of the game. We provide several counterexamples to illustrate this fact. They are all based on the following idea: The “tit for tat” principle allows one to define many equilibria in repeated games with  $N$  players. However, when the number of players is infinite, the deviation of a single player is not visible by the population that cannot punish him in retaliation for her deviation. This implies that while the games with  $N$  players may have many equilibria, as stated by the folk theorem, this may not be the case for the limit game. This fact is well-known for large repeated games (see examples of *Anti-folk*

*Theorems* in [35, 2]). However, up to our knowledge, these results have not yet been investigated in the mean field game framework.<sup>1</sup>

Finally, our four models of dynamic games do not face the issue of the order of play, nor partial information. Thus, we avoid two difficulties of dynamic games: the information structure of each player and the existence of a value [15]. In our case, all players are similar, so the order of play is irrelevant, and we only consider the full information case: players know the strategy of the other players and the current global state (more details on this are given in Section 3.2).

The rest of the article is organized as follows. We introduce mean field games with explicit interactions in continuous time in Section 2 where we mainly focus on the infinite horizon with discounted costs. We describe the evolution of the state of the players, the cost function as well as the best response operator. In both cases (finite and infinite horizon), we prove the existence of an equilibrium. We show in Section 3 that this equilibrium is an approximation of an equilibrium for the game with a finite number of players. Finally, we study an example of an  $N$ -player game inspired from the prisoner's dilemma whose equilibria are not always equilibria for the limit mean field game. We focus on the synchronous case in Section 5 (where players all play at the same time). In this case,  $N$ -player games can be seen as classical stochastic games in discrete time. We derive the mean field limit dynamics and the existence of an equilibrium. Here counter-examples of equilibria for finite games that do not go to the limit are easier to find. Indeed, the folk theorem applies and all equilibria based on retaliation cannot be equilibria at the limit.

## 2. Discrete Mean Field Games in Continuous Time

### 2.1. Notations and Definitions

A discrete mean field game  $\mathcal{G}$  is a tuple  $\mathcal{G} = (\mathcal{E}, \mathcal{A}, \{Q_a\}, \mathbf{m}_0, \{c_a\}, \beta)$ , where  $\mathcal{E}$  is the state space,  $\mathcal{A}$  the action set,  $\{Q_a\}$  the transition rate matrices,  $\mathbf{m}_0$  the initial state,  $\{c_a\}$  the cost functions and  $\beta \in \mathbb{R}$  a discount factor.

The game is described as follows.

**State and action sets.** We consider a *population* made of an infinite number of homogeneous *players* that evolve in continuous time. Each player has a finite state space denoted by  $\mathcal{E} = \{1, \dots, E\}$  and a finite action set  $\mathcal{A} = \{1, \dots, A\}$ .

We denote by  $\mathcal{P}(\mathcal{A})$  (resp.  $\mathcal{P}(\mathcal{E})$ ) the set of probability measures over  $\mathcal{A}$  (resp.  $\mathcal{E}$ ). Since  $\mathcal{A}$  is finite,  $\mathcal{P}(\mathcal{A})$  is the simplex of dimension  $A$ .

---

<sup>1</sup>An extended abstract discussing our counterexample in the continuous time model with infinite horizon was presented in [16].

**Set of strategies.** A *mixed strategy* (or strategy for short) is a measurable function  $\pi : \mathcal{E} \times \mathbb{R}^+ \rightarrow \mathcal{P}(\mathcal{A})$ , that associates to each state  $i \in \mathcal{E}$  and each time  $t \geq 0$  a probability measure  $\pi_i(t) \in \mathcal{P}(\mathcal{A})$  on the set of possible actions. We also denote by  $\pi_{i,a}(t)$  the probability that, at time  $t$ , a player in state  $i$  takes the action  $a$ , under strategy  $\pi$ . For all  $t \geq 0$  and all  $i \in \mathcal{E}$ , we have  $\sum_{a \in \mathcal{A}} \pi_{i,a}(t) = 1$ . The set of all possible strategies is denoted by  $\mathcal{S}$ .

We say that a strategy is *pure* if, for all state  $i$  and all  $t \in \mathbb{R}$ , there exists an action  $a \in \mathcal{A}$  such that  $\pi_{i,a}(t) = 1$  and  $\pi_{i,a'}(t) = 0$  for all  $a' \neq a$ .

The set  $\mathcal{S}$  is a bounded subset of the Hilbert space of the functions  $\mathcal{E} \times \mathbb{R}^+ \rightarrow \mathbb{R}^{\mathcal{A}}$  equipped with the inner product the exponentially weighted inner product :  $\langle f, g \rangle = \int_0^\infty f(g)g(t)e^{-\beta t} dt$ . This shows that  $\mathcal{S}$  is weakly compact, where the weak topology is defined as follows: a sequence of policy  $\pi^n$  converges to a policy  $\pi$  if for any bounded function  $g$ :

$$\lim_{n \rightarrow \infty} \int_0^\infty \pi^n(t)g(t)e^{-\beta t} dt = \int_0^\infty \pi(t)g(t)e^{-\beta t} dt. \quad (1)$$

**Rate matrices.** We denote by  $\mathbf{m}^\pi(t) \in \mathcal{P}(\mathcal{E})$  the *population distribution* at time  $t$ . As the state space is finite,  $\mathbf{m}^\pi(t)$  is a vector whose  $i$ -th component,  $m_i^\pi(t)$ , is the proportion of players in state  $i$  at time  $t$ . The evolution over time of the population distribution is driven by rate matrices:  $\{Q_a(\mathbf{m}^\pi(t))\}_{a \in \mathcal{A}}$ . By definition,  $Q_{ija}(\mathbf{m}^\pi(t))$  is the rate at which a player in state  $i$  moves to state  $j$  when choosing action  $a$ , when the population distribution is  $\mathbf{m}^\pi(t)$ . Note that by definition,  $\sum_{j \in \mathcal{E}} Q_{ija}(\mathbf{m}^\pi(t)) = 0$  for all  $i$  and  $a$  and  $Q_{ija}(\mathbf{m}^\pi(t))$  is non-negative for all  $j \neq i$  and all  $a$ .

In the following, we assume that for all  $i, j, a$ ,  $Q_{ija}(\mathbf{m})$  is Lipschitz-continuous in  $\mathbf{m}$  with constant  $L$ .

The initial condition is  $\mathbf{m}^\pi(0) = \mathbf{m}_0$ . For  $t \geq 0$ , the population distribution  $\mathbf{m}^\pi(t)$  is the solution of the following differential equation, that depends on the strategy  $\pi$ : for  $j \in \mathcal{E}$

$$\dot{m}_j^\pi(t) = \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} m_i^\pi(t) Q_{ija}(\mathbf{m}^\pi(t)) \pi_{i,a}(t). \quad (2)$$

The rationale behind this differential equation is that all players in state  $i$  use the action  $a \in \mathcal{A}$  and move to state  $j$  with rate  $Q_{ija}(\mathbf{m}^\pi(t))$ .

If the strategy  $\pi_i(t)$  is not continuous in time, the differential equation (2) may not be well-posed at time-points where  $\pi_i$  is not continuous. The existence of a continuous solution for (2) is guaranteed by the Carathéodory's Existence Theorem. The Lipschitz condition on  $Q$  further implies that this solution is essentially unique

because any solution of (2) must be a fixed point of

$$m_j^\pi(t) = m_{j,0} + \int_0^t \left( \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} m_i^\pi(u) Q_{ija}(\mathbf{m}^\pi(u)) \pi_{i,a}(u) \right) du. \quad (3)$$

In anticipation, the same properties (existence and uniqueness of the solution of the ODE) hold for the differential equation (4).

**Remark 1** (Explicit interactions). *In this model, the rate matrix  $Q_{ija}(\mathbf{m}^\pi(t))$  depends explicitly on the population distribution: the rate to go from state  $i$  to state  $j$  under action  $a$  depends on how the whole population is distributed among the states of the system. Other mean field models, such as [20], only consider the special case where  $Q_{ija}(\mathbf{m}^\pi(t))$  is constant:  $Q_{ija}(\mathbf{m}^\pi(t)) = Q_{ija}$ . This restricts the population dynamics given in (2) to linear dynamics.*

**Cost function.** We now concentrate on a particular player, that we call Player 0. Player 0 chooses her own strategy  $\pi^0 : \mathbb{R}^+ \times \mathcal{E} \rightarrow \mathcal{P}(\mathcal{A})$ . We denote by  $\mathbf{x}^{\pi^0}(t) \in \mathcal{P}(\mathcal{E})$  the probability distribution of Player 0 when Player 0 uses strategy  $\pi^0$  against a population who has distribution  $\mathbf{m}$ . For a given state  $i \in \mathcal{E}$ ,  $x_i^{\pi^0, \mathbf{m}}(t)$  denotes the probability for Player 0 to be in state  $i$  at time  $t$ . The distribution  $\mathbf{x}^{\pi^0, \mathbf{m}}$  evolves over time according to the following differential equation: for  $j \in \mathcal{E}$

$$\dot{x}_j^{\pi^0, \mathbf{m}}(t) = \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} x_i^{\pi^0}(t) Q_{ija}(\mathbf{m}(t)) \pi_{i,a}^0(t). \quad (4)$$

If Player 0 is in state  $i$  and takes an action  $a$ , it suffers from an *instantaneous cost*  $c_{i,a}(\mathbf{m}(t))$ , that depends on the population distribution at time  $t$ . We assume that the cost is always continuous in  $\mathbf{m}$ . Given a population distribution  $\mathbf{m}$  and the strategy of Player 0  $\pi^0$ , we define the discounted cost of Player 0 as

$$W(\pi^0, \mathbf{m}) = \int_0^\infty \left( \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} x_i^{\pi^0, \mathbf{m}}(t) c_{i,a}(\mathbf{m}(t)) \pi_{i,a}^0(t) e^{-\beta t} \right) dt, \quad (5)$$

where  $\beta > 0$  is the discount factor.

We also introduce the notation  $V(\pi^0, \pi)$  that represents the discounted cost of Player 0 when the population plays a strategy  $\pi$ :

$$V(\pi^0, \pi) = W(\pi^0, \mathbf{m}^\pi).$$

**Best response.** The *best response* to  $\pi$  of Player 0 is to choose a strategy  $\pi^0 \in \mathcal{S}$  that minimizes her discounted cost (5) when the rest of the population plays strategy

$\pi$ . For a given population strategy  $\pi$ , we denote the set of best responses of Player 0 to  $\pi$  by  $BR(\pi)$ . This set is the set of strategies that minimizes her discounted cost:

$$BR(\pi) := \arg \min_{\pi^0 \in \mathcal{S}} V(\pi^0, \pi). \quad (6)$$

Note that the best response function is well defined or, in other words, that the “argmin” is reached for some strategy in Equation (6). To prove that, we will later prove in Section 2.3 that the function  $V$  is continuous for the weak topology. As  $\mathcal{S}$  is weakly compact, this shows that the minimum in  $\pi^0$  is attained.

**Proposition 1.** *The function  $V$ , defined in Equation (5) is continuous in  $\pi^0$  and  $\pi$  (for the weak-topology on  $\mathcal{S}$ ).*

**Mean field equilibrium.** We then define a mean field equilibrium as a strategy  $\pi^{MFE}$  such that when the population strategy is  $\pi^{MFE}$ , a selfish Player 0 would also choose the same strategy  $\pi^{MFE}$  as her best response.

**Definition 1** (Mean Field Equilibrium). *A strategy  $\pi$  is called a mean field equilibrium if it is a fixed point for the best response function, i.e.,*

$$\pi^{MFE} \in BR(\pi^{MFE}). \quad (7)$$

*A mean field equilibrium is pure if it is a pure strategy.*

The rationale behind this definition is when one considers that the population is formed by players that each take selfish decisions. As the population is homogeneous, each player best response is the same as Player 0. In other words, for a given population strategy  $\pi$ , all the rational players of the populations (or players) choose the strategy  $BR(\pi)$ . As in classical games, a mean field equilibrium is a situation where no player has incentive to deviate unilaterally from the common strategy.

## 2.2. Existence of Mean Field Equilibrium

We now show that, under very general assumptions, all discrete mean field games admit a mean field equilibrium. As for classical games, these equilibria are not necessarily *pure*. As most proof on existence of equilibria, our proof relies on a generalization of Kakutani fixed point theorem to infinite dimensional spaces. However, the classical approach consisting of showing that the best response function  $BR(\pi)$  is a Kakutani map does not work here when the cost function is not strictly convex. Therefore, in our approach we focus on the state of the game instead of the best response function.

As mentioned before, the differential equations (2), (4) and the cost equation (5) are all well defined under our running Assumption (A1):



(A1) The rate function  $\mathbf{m} \mapsto Q_{ija}(\mathbf{m})$  is Lipschitz-continuous in  $\mathbf{m}$ . The cost function  $\mathbf{m} \mapsto c_{i,a}(\mathbf{m})$  is continuous in  $\mathbf{m}$ .

In particular, this assumption implies that the costs and the rates are all bounded by a finite value.

**Theorem 1.** *Any discrete mean field game  $\mathcal{G}$  whose rate and cost satisfy Assumption (A1) admits a mean field equilibrium.*

Note that in general, the best response function  $\pi \mapsto BR(\pi)$  is neither continuous nor hemi-continuous in general under (A1). In particular, the best response set  $BR(\pi)$  may not be a convex set. This makes difficult the application of the classical fixed point theorems on the best response function. As a result, our proof will formulate the fixed point problem in an alternative manner by considering a fixed point in  $\mathbf{m}$ .

### 2.3. Proofs

#### 2.3.1. Proof of Proposition 1

For a strategy  $\pi$ , the function  $\mathbf{m}^\pi$  satisfies the differential equation (2). As  $\mathbf{m}^\pi(t)$  lives in a compact and the functions  $Q$  are continuous, the right-hand side of this differential equation is bounded. This shows that there exists a constant  $L'$  such that for any strategy  $\pi$ , the function  $\mathbf{m}^\pi$  is Lipschitz-continuous with constant  $L'$ . Similarly the function  $\mathbf{x}^{\pi^0}$  is also Lipschitz-continuous with constant  $L'$ .

Let  $\mathcal{M}$  be the set of functions from  $\mathbb{R}^+$  to  $\mathcal{P}(\mathcal{E})$  that are Lipschitz-continuous with constant  $L'$ . We equip this set with the exponentially weighted  $L_\infty$ -norm :

$$\|\mathbf{m} - \mathbf{m}'\| = \sup_{i \in \mathcal{E}, t \geq 0} |m_i(t) - m'_i(t)| e^{-\beta t}.$$

By the Arzela-Ascoli theorem,  $\mathcal{M}$  is a compact space.

To prove that  $V$  is continuous in  $\pi$  and  $\pi^0$ , it suffices to show that the mapping  $\pi \mapsto \mathbf{m}^\pi$  is continuous (for the weak topology) and that the mapping  $(\pi^0, \mathbf{m}) \mapsto x^{\pi^0, \mathbf{m}}$  is continuous. To prove the continuity of  $\mathbf{m}^\pi$ , let  $\pi_n$  be a sequence of strategy that converges to a strategy  $\pi$ . As  $\mathcal{M}$  is compact, there exists a function  $\mathbf{m}$  and a subsequence of  $m^{\pi_n}$  that converges to  $\mathbf{m}$ . Moreover, we have :

$$m_j(t) = m_{j,0} + \lim_{n \rightarrow \infty} \int_0^t \left( \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} m_i^{\pi_n}(u) Q_{ija}(\mathbf{m}^{\pi_n}(u)) (\pi_n)_{i,a}(u) \right) du \quad (8)$$

$$= m_{j,0} + \int_0^t \left( \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} m_i(u) Q_{ija}(\mathbf{m}(u)) \pi_{i,a}(u) \right) du, \quad (9)$$

where the convergence holds because  $\pi_n$  converges weakly to  $\pi$  and  $\mathbf{m}^{\pi_n}$  converges uniformly on all compact to  $\mathbf{m}$ .

Equation (9) shows that the function  $\mathbf{m}$  is equal to the function  $\mathbf{m}^\pi$ . This shows that  $\pi \rightarrow \mathbf{m}^\pi$  is continuous in  $\pi$  which implies that  $V$  is continuous in  $\pi$ .

The proof that  $(\pi^0, \mathbf{m}) \mapsto x^{\pi^0, \mathbf{m}}$  is continuous is very similar to the above proof and we therefore omit it.

### 2.3.2. Proof of Theorem 1

Recall that for a given population distribution  $\mathbf{m} \in \mathcal{M}$ , the cost of a strategy  $\pi^0$  is defined as

$$W(\pi^0, \mathbf{m}) = \int_0^\infty \left( \sum_{i,a} x_i(t) \pi_{i,a}^0(t) c_{i,a}(\mathbf{m}(t)) e^{-\beta t} \right) dt, \quad (10)$$

$$\text{where } \mathbf{x} \text{ satisfies (for all } j \in \mathcal{E}\text{): } \dot{x}_j(t) = \sum_{i,a} x_i(t) Q_{ija}(\mathbf{m}(t)) \pi_{i,a}^0(t). \quad (11)$$

We now define the function  $\Phi : \mathcal{M} \rightarrow 2^{\mathcal{M}}$  as the best response to a population distribution  $\mathbf{m}$ . It is a mapping that associates to a population distribution  $\mathbf{m} \in \mathcal{M}$ , the set of all state distributions that can be induced by an optimal policy:

$$\Phi(\mathbf{m}) = \left\{ \mathbf{x}^{\pi^0} \text{ such that } \pi^0 \in \arg \min_{\pi \in \mathcal{S}} W(\pi, \mathbf{m}) \right\}. \quad (12)$$

In the remainder of the proof, for all  $\mathbf{m} \in \mathcal{M}$ ,  $\Phi(\mathbf{m})$  is well defined and non empty (*i.e.*, the minimum is attained), is convex and compact. Moreover, we will also show that the function  $\Phi(\cdot)$  is upper-semicontinuous. As  $\mathcal{M}$  is compact [8, Prop. 11.11], this shows that  $\Phi(\cdot)$  satisfies the conditions of the fixed point theorem given in [26, Theorem 8.6] and therefore has a fixed point  $\mathbf{m}^*$ . By the definition of  $\Phi$ , this implies that there exists a strategy  $\pi^0$  that is a best-response to  $\mathbf{m}^*$ , which implies that  $\pi^0$  is a mean field equilibrium.

**Definition of  $\Phi(\mathbf{m})$**  – It can be shown that  $W$  is continuous (by using a reasoning similar to the one for  $V$  (Proposition 1)). This shows that there exists  $\pi^0$  that attains the minimum on the right hand side of Equation (12), which shows that  $\Phi(\mathbf{m})$  is well defined and non-empty.

**Compactness of  $\Phi(\mathbf{m})$**  – Let us consider the following optimization problem:

$$\min_{\mathbf{x}, \mathbf{z}} \int_0^\infty \left( \sum_{i,a} z_{i,a}(t) c_{i,a}(\mathbf{m}(t)) e^{-\beta t} \right) dt \quad (13)$$

$$\text{such that } \mathbf{z} \text{ satisfies } \begin{cases} \sum_a z_{j,a}(t) = x_j(t) & \forall j \in \mathcal{E}, \\ z_{j,a}(t) \geq 0, & \forall j \in \mathcal{E}, \forall a \in \mathcal{A}, \\ \dot{x}_j(t) = \sum_{i,a} z_{i,a}(t) Q_{ija}(\mathbf{m}(t)) & \forall j \in \mathcal{E}. \end{cases} \quad (14)$$

The above problem is a linear problem, which implies that the set of optimal solutions is convex and compact. Let us show that the set of optimal solution of the optimization problem (13) is  $\Phi(\mathbf{m})$ . To show this, let us remark that the constraints (11) are equivalent to the constraints (14) by replacing the variables  $x_i(t)\pi_{i,a}^0(t)$  by  $z_{i,a}(t)$ . Then, the constraint  $\pi \in \mathcal{S}$  of (11), that corresponds to  $\pi^0(t) \in \mathcal{P}(\mathcal{A})$ , is replaced with  $z_{i,a}(t) \geq 0$  and  $\sum_a z_{i,a}(t) = x_i(t)$ .

**Upper-semi continuity of  $\Phi$ .** To prove that  $\Phi$  is upper-semi continuous, let us show that the graph of  $\mathbf{m} \mapsto \Phi(\mathbf{m})$  is closed. Let  $\mathbf{m}_n \in \mathcal{M}$  and  $\mathbf{x}_n \in \Phi(\mathbf{m}_n)$  be two sequences such that  $\lim_{n \rightarrow \infty} \mathbf{m}_n = \mathbf{m}_\infty$  and  $\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}_\infty$ . We want to show that  $\mathbf{x}_\infty \in \Phi(\mathbf{m}_\infty)$ .

As  $W$  is continuous, for all  $\mathbf{x}_n \in \Phi(\mathbf{m}_n)$ , there exists a strategy  $\pi_n$  that minimizes  $W(\pi, \mathbf{m}_n)$  and such that  $\mathbf{x}_n = \mathbf{x}^{\pi_n, \mathbf{m}_n}$ . As the set  $\mathcal{S}$  is weakly compact, this sequence of strategies has a subsequence that converges weakly to a strategy  $\pi_*$ . Moreover, we have:

- As  $W$  is continuous,  $\pi_*$  minimizes  $W(\pi, \mathbf{m}_\infty)$ . This shows that  $\mathbf{x}^{\pi_*, \mathbf{m}_\infty} \in \Phi(\mathbf{m}_\infty)$ .
- The solution of (11) is continuous in  $\pi$  and  $\mathbf{m}$ , which shows that  $\mathbf{x}_\infty = \mathbf{x}^{\pi_*, \mathbf{m}_\infty}$ .

Combining these two facts shows that  $\mathbf{x}_\infty \in \Phi(\mathbf{m}_\infty)$  which implies that the graph of  $\Phi$  is closed.  $\square$

**Remark 2.** *The continuity assumption (A1) is tight in the following sense:*

1- *If the rate  $Q$  is not Lipschitz-continuous in  $\mathbf{m}$ , then the evolution of the population is not well defined, in the sense that the evolution equation (2) may have several solutions or no solution at all.*

2- *There exist games with non-continuous cost functions that do not admit any mean field equilibrium. For example, consider the following mean field game:*

$$\mathcal{G} = \left( \mathcal{E} = \{1, 2\}, \mathcal{A} = \{a, b\}, Q_a = 0, Q_b = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, m(0) = (1, 0) \right) \quad (15)$$

$$c_a(m_1, m_2) = 0, c_b(m_1, m_2) = \begin{cases} -1 & \text{if } m_2 \leq 1/2 \\ 1 & \text{otherwise} \end{cases}, \beta. \quad (16)$$

Assume that this game has a mean field equilibrium and let denote by  $m(t)$  the state at equilibrium. By definition of  $Q_a$  and  $Q_b$ ,  $m_2(t)$  is a non-decreasing function. Hence, let  $\tau = \sup\{t : m_2(t) \leq 1/2\}$  (note that  $\tau \in [\ln 2; +\infty) \cup \{+\infty\}$ ). It should be clear that the best response of Player 0 to any state function  $m$  is the policy  $\pi^{(\tau)}$  that consists in playing “b” until  $\tau$  and “a” after  $\tau$ . However, such a policy is never a mean field equilibrium: under the policy  $\pi^{(\tau)}$ ,  $m_2(t) = 1 - e^{-\min(t, \tau)}$ , which means that  $\sup\{t : m_2(t) \leq 1/2\} \in \{\ln 2, +\infty\}$ . None of the policies  $\pi^{(\ln 2)}$  or  $\pi^{(+\infty)}$  is an equilibrium: the policy  $\pi^{(\ln 2)}$  is the best response to  $\pi^{(+\infty)}$  and vice-versa.

### 3. Convergence of Finite Games to Mean Field Games

Mean field games are often presented as a limit of a sequence of finite games as the number  $N$  of players goes to infinity. In this section, we investigate positive and negative results that link finite games and mean field games.

#### 3.1. Markov Game with $N$ Exchangeable Players

To any discrete mean field game  $\mathcal{G} = (\mathcal{E}, \mathcal{A}, \{Q_a\}, \mathbf{m}_0, \{c_a\}, \beta)$ , one can associate a stochastic  $N$ -player game  $\mathcal{G}^N = (N, \mathcal{E}, \mathcal{A}, \{Q_a\}, \mathbf{m}_0, \{c_a\}, \beta)$  as follows. The finite stochastic game  $\mathcal{G}^N$  has the same state and action spaces  $\mathcal{E}, \mathcal{A}$ , the same rate matrices  $Q_a$ , the same cost functions  $c_a$ , the same discount factor  $\beta$ , and the same initial state as  $\mathcal{G}$ . The time evolution of the finite game is as follows. At any time  $t$ , each player (say Player  $n$ ) chooses a (randomized) action  $A_n(t) \in \mathcal{P}(\mathcal{A})$ .

We consider a mean field interaction model between the players, which means that the behavior of one object only depends on the states of the other objects through the proportion of objects that are in a given state. To be more precise, we denote by  $\mathbf{M}(t) \in \mathcal{P}(\mathcal{E})$  the population distribution of the system at time  $t$ . As the set  $\mathcal{E}$  is finite,  $\mathbf{M}(t)$  is a vector with  $|\mathcal{E}|$  components and for all  $i \in \mathcal{E}$ ,  $M_i(t)$  is the fraction of players that have state  $i$  at time  $t$ :

$$M_i(t) = \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{\{X_n(t)=i\}}.$$

The state of one player (say Player  $n$ ) follows a continuous time Markov chain whose rate varies over time. The only dependence between players is through the rate that depends on the population distribution.

More precisely, the evolution of the state of Player  $n$ , under  $\mathcal{F}_t$ , the natural filtration of the process, satisfies for all  $k \in \mathbb{N}$  and all states  $i \neq j$ ,

$$\mathbb{P}(X_n(t+dt) = j | X_n(t) = i, \mathbf{M}(t) = \mathbf{m}, A_n(t) = a, \mathcal{F}_t) = Q_{ija}(\mathbf{m})dt + o(dt), \quad (17)$$

where  $A_n(t)$  is the action taken by Player  $n$  at time  $t$ .

At any time  $t$ , Player  $n$  suffers an instantaneous cost that is a function of her state  $X_n(t)$ , the action that she takes  $A_n(t)$  and the population distribution  $\mathbf{M}(t)$ . We write this instantaneous cost  $c_{X_n(t), A_n(t)}(\mathbf{M}(t))$ .

The objective of Player  $n$  is to choose a strategy  $\pi^n$  from some set of admissible strategies  $\Pi$ , in order to minimize her expected discounted cost, knowing the strategies of the others. As before, the discount factor is denoted by  $\beta$ . Given a strategy  $\pi^n \in \Pi$  used by Player  $n$  and a strategy  $\pi \in \Pi$  used by all the others, we denote by  $V^N(\pi^n, \pi)$  the expected discounted cost of Player  $n$ :

$$V^N(\pi^n, \pi) = \mathbb{E} \left[ \int e^{-\beta t} c_{X_n(t), A_n(t)}(\mathbf{M}^\pi(t)) dt \left| \begin{array}{l} A_n \text{ is chosen w.r.t. } \pi^n \\ A_{n'} \text{ is chosen w.r.t. } \pi \ (\forall n' \neq n) \end{array} \right. \right].$$

A Nash equilibrium for this game is a strategy  $\pi$  such that Player  $n$  does not have another admissible strategy that leads to a lower cost. This notion depends naturally on the set of admissible strategies.

**Definition 2** (Equilibrium of the  $N$  player game). *For a given set of strategies  $\Pi$ , a strategy  $\pi \in \Pi$  is called a symmetric equilibrium in  $\Pi$  if for any strategy  $\pi^n \in \Pi$ :*

$$V^N(\pi, \pi) \leq V^N(\pi^n, \pi).$$

We will also use the notion of  $\varepsilon$ -equilibrium:

**Definition 3** ( $\varepsilon$ -equilibrium of the  $N$  player game). *For a given set of strategies  $\Pi$ , a strategy  $\pi \in \Pi$  is called an  $\varepsilon$ -symmetric equilibrium in  $\Pi$  if for any strategy  $\pi^n \in \Pi$ :*

$$V^N(\pi, \pi) \leq V^N(\pi^n, \pi) + \varepsilon.$$

### 3.2. Subsets of Admissible Strategies

In a *full information* setting,  $A_n(t)$  is a (possibly random) function of the values  $X_{n'}(t')$  up to time  $t' \leq t$  and all actions taken in the past  $A_{n'}(t')$ , for  $t' < t$  and for  $n' \in \{1 \dots N\}$ . Such a strategy is, however, hard to analyze. Therefore, in the following, we will consider two natural subclasses for the set of admissible strategies, depending on the information available to the players:

- (Markov) – A strategy  $\pi$  is called a *Markov strategy* if it induces a choice of  $A_n(t)$  that is a (possibly random) measurable function of only  $t$ ,  $\mathbf{M}(t)$  and  $\mathbf{X}(t)$ :

$$\mathbb{P}(A_n(t) = a \mid \mathcal{F}_t) = \pi_{a, X_n(t)}(t, \mathbf{M}(t)).$$

This definition is motivated by the fact that, as indicated by Equation (17), the behavior of one object depends on the others only through the value  $\mathbf{M}(t)$ . This implies that when all the other players use a Markov strategy, the set of Markov strategies is dominant among the set of full-information strategies: there exists a full-information best response for Player  $n$  that is a Markov strategy. Furthermore, any Markov game admits a Markovian Nash equilibrium (see [17]).

- (Local) – A strategy  $\pi$  is a *local strategy* if the choice of the action only depends on the player's internal state and on the time.

$$\mathbb{P}(A_n(t) = a \mid \mathcal{F}_t) = \pi_{a, X_n(t)}(t).$$

If a player uses a local strategy, its actions may depend on time, hence may track the law of the population  $\mathbf{M}(t)$  (but not  $\mathbf{M}(t)$  itself). Also notice that a local strategy is not necessarily stationary because of its dependence on time.

### 3.3. Nash Equilibria Limits

The next theorem provides a relation between local equilibria of finite games and mean field equilibria of the limit mean field game. In particular, it shows that mean field equilibria are a good approximation of local equilibria. However, as we will show later, this result does not hold for Markovian equilibria.

**Theorem 2.** *Consider a finite stochastic game  $\mathcal{G}^N$ , with  $N$  players and assume that (A1) holds for its rate matrices  $Q_a$  and its cost functions  $c_a$ . Then:*

- (i) *Let  $\pi$  be a mean field equilibrium of the associated mean field game  $\mathcal{G}$ . There exists  $N_0$  such that for all  $N \geq N_0$ ,  $\pi$  is a local  $\varepsilon$ -equilibrium of the  $N$  player game.*
- (ii) *Let  $(\pi^N)_{N \in \mathbb{N}}$  be a sequence of local strategies such that  $\pi^N$  is an  $\varepsilon_N$ -equilibrium for the  $N$  player game, with  $\varepsilon_N \rightarrow 0$ . Then any sub-sequence of the sequence  $(\pi^N)$  has a sub-sequence that converges weakly to a mean field equilibrium of  $\mathcal{G}$ .*

*Proof.* First,  $V^N(\pi^n, \pi)$  converges to  $V(\pi^n, \pi)$  uniformly in  $(\pi^n, \pi)$ . Uniform convergence follows from Theorem 3.3.2 in [38] (The theorem is stated for stationary strategies, but local strategies as defined here are equivalent to stationary strategies, as defined in [38]).

Thus, for any  $\varepsilon$ , there exists  $N_0$  such that  $N \geq N_0$  implies that  $|V^N(\pi^n, \pi) - V(\pi^n, \pi)| \leq \varepsilon/2$ . Hence, if  $\pi$  is a mean field equilibrium, this implies that for any local strategy  $\pi^n$ :

$$V^N(\pi, \pi) \leq V(\pi, \pi) + \frac{\varepsilon}{2} \leq V(\pi^n, \pi) + \frac{\varepsilon}{2} \leq V^N(\pi^n, \pi) + \varepsilon.$$

This shows (i).

For (ii), if  $\pi^N$  is a sequence of local strategies, then any sub-sequence has a sub-sequence that converge weakly to some local strategy  $\pi^\infty$ . As  $V(\pi^n, \pi)$  is continuous in  $\pi^n$  and  $\pi$  (for the weak topology), this implies that  $V(\pi^\infty, \pi^\infty) \leq V(\pi^n, \pi^\infty)$  for all local strategy  $\pi^n$ .  $\square$

### 3.4. Markov Equilibria May Not Converge to Mean Field Equilibria

We now show that Theorem 2-(ii) does not generalize to Markov strategies. The following example was first presented in [16]. The main ingredient used to construct the following counterexample, is the “tit-for-tat” principle. This principle can be used to construct equilibria for any  $N$ -player game but cannot be used in mean field games. This approach has been used in repeated game papers (see for example the examples in [35], further generalized [2]). Up to our knowledge, this type of behavior has not yet been described in the mean field game framework.

Let us consider a mean field version of the classical prisoner’s dilemma. The state space of a player is  $\mathcal{E} = \{C, D\}$  (that stand for Cooperate and Defect) and the action set is the same  $\mathcal{A} = \mathcal{E}$ . At each time step, one player is chosen. If she selects an action  $a \in \mathcal{A}$ , her state becomes  $a$  at the next time step.

The instantaneous cost of a Player  $n$  depends on her state  $i$  and on the mean field  $m$ :

$$c_{i,a}(m) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

At each time step, this cost function corresponds to a matching game where a player plays against a randomly assigned opponent and suffers a cost that corresponds to the following matrix:

	C	D
C	1,1	3,0
D	0,3	2,2

The strategy  $D$  dominates the strategy  $C$ . This implies that playing  $D$  is the unique mean field equilibrium. Indeed, the expected cost (given by (5)) of a Player 0 that has a state vector  $x$  while the mean field is  $m(t)$  is

$$\begin{aligned} & \int_0^\infty [x_C(t)(m_C(t) + 3m_D(t))(\pi_{CC}^0(t) + \pi_{CD}^0(t)) + x_D(t)2m_D(t)(\pi_{DC}^0(t) + \pi_{DD}^0(t))]e^{-\beta t} dt \\ &= \int_0^\infty [x_C(t) + 2m_D(t)]e^{-\beta t} dt, \end{aligned}$$

by using the fact that  $\pi_{CC}^0(t) + \pi_{CD}^0(t) = \pi_{DC}^0(t) + \pi_{DD}^0(t) = 1$  and  $x_C(t) + x_D(t) = m_C(t) + m_D(t) = 1$ .

It should be clear that this cost is minimized when  $x_C$  is minimal, which occurs when the strategy is to choose action  $D$  regardless of the current state. This shows that the only mean field equilibrium is when all players choose action  $D$ .

Let us now consider the game with  $N$  players and consider the following Markov strategy:

$$\pi^N(m) = \begin{cases} C & \text{if } m_C = 1 \\ D & \text{if } m_C < 1 \end{cases}$$

and let us show that for  $\beta < 1$  and  $N$  large,  $\pi^N$  is a Markov Nash equilibrium.

Assume that all players, except Player  $n$ , play the strategy  $\pi^N$  and let us compute the best response of Player  $n$ . It should be clear that if at time 0,  $m_C < 1$ , then the best response of Player  $n$  is to play  $D$ . On the other hand, if  $m_C = 1$ , then:

- If Player  $n$  applies  $\pi^N$ , she will suffer a cost  $\int \exp(-\beta t) dt = 1/\beta$ .
- If Player  $n$  deviates from  $\pi^N$  and chooses the action  $D$ , all players will also deviate after that time. This implies that  $m_D(t) \approx 1 - \exp(-t)$  and that the player  $n$  will suffer a cost approximately equal to  $\int_0^\infty (x_C(t) + 2 - 2e^{-t})e^{-\beta t} dt \geq 2/(\beta(\beta + 1))$  when  $N$  is large.

When  $\beta < 1$ , then  $2/(\beta(\beta + 1)) > 1/\beta$ , so that Player  $n$  has no incentive to deviate from the strategy  $\pi^N$  and that therefore,  $\pi^N$  is a Nash equilibrium. We also observe that for this example, the value of the finite game does not converge to the one of the mean field game.

In conclusion to this section, one can argue that this counter-example should not be surprising because, in mean field games, punishment is possible against a fraction on the population that deviates but is not possible against individual deviation, because it is not seen in the population distribution.

As a final remark, as in the case of repeated games, the continuity with respect to  $m$  (not true here) is critical for convergence (see [35]).

#### 4. Finite Horizon Case

Let us now consider mean field games over a finite time horizon  $T$ . These games are similar to games with discounted costs, previously defined, but they only run for a finite duration  $T$ . As in the discounted case, the evolution over time of the population distribution  $\mathbf{m}^\pi$  is given by (2) and the evolution of Player 0's distribution is given by (4).



Given the population strategy  $\pi$  and Player 0 strategy  $\pi^0$ , the expected cost of Player 0 for the finite horizon case is defined as follows:

$$V(\pi^0, \pi) = \int_0^T \left( \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} x_i(t) c_{i,a}(\mathbf{m}^\pi(t)) \pi_{i,a}^0(t) \right) dt. \quad (18)$$

In the literature, similar models have been studied, considering continuous time finite state space mean field games with finite horizon. The authors in [21] consider uniformly convex cost functions and in [27] cost functions are assumed to be strictly convex. In our model, we assume that the costs are continuous in the population distribution. It can also be observed that the instantaneous cost of Player 0 is linear in  $\pi^0$ . Therefore, the model we study in this work is not covered by these papers.

We define the notion of mean field equilibrium for the finite horizon case as in the discounted case, by replacing the cost function (5) by (18). Then, the proof of the existence Theorem 1 applies *mutatis mutandis* to show the existence of a mean field equilibrium in this case: Any continuous time mean field game over a finite horizon that satisfies Assumption (A1) has a mean field equilibrium.

#### 4.1. Convergence to a Mean Field Equilibrium

The construction of a counter example of convergence with an infinite time horizon given in §3.4 cannot be directly adapted to the finite horizon case. In the finite-horizon version of the game defined in §3.4, the strategy  $\pi^N$  is not a Nash equilibrium for the  $N$ -player game because at the last time-slot, the best response of Player  $n$  to any strategy is to play  $D$ . By induction on the number of time-slots, the only Nash equilibrium of the  $N$ -player game is when all players play  $D$ , which coincides with the mean field equilibrium.

Yet, a counter-example also exists for finite-time horizon. The essential idea is to start with a matrix game with two pure Nash equilibria instead of one as in the previous example. Let us consider the following cost matrix:

	C	D	P
C	1,1	3,0	4,0
D	0,3	2,2	4,3
P	0,4	3,4	3,3

The setting is similar to the previous example: the action set is equal to the state  $\mathcal{E} = \mathcal{A} = \{C, D, P\}$  and at each time step, one player is chosen. If she selects an action  $a \in \mathcal{A}$ , then her state becomes  $a$  at the next time step. This game can be viewed as a generalization of the prisoner's dilemma with an additional Nash-equilibrium  $P$  (which stands for "punish"). It can be shown using a similar path as

in the previous section that, when  $T$  is large enough, the following time-dependent Markovian<sup>2</sup> strategy is a Nash equilibrium:

$$\pi^N(m, t) = \begin{cases} C & \text{if } t < 1 \text{ and } m_c = 1; \\ D & \text{if } t \geq 1 \text{ and } m_P = 0; \\ P & \text{otherwise.} \end{cases} \quad (19)$$

In the above strategy, the state  $P$  is used as a stick to punish people from deviating from the imposed strategy. In this case, nobody has an incentive to deviate from this strategy at the last step because  $D$  is also a Nash equilibrium.

The mean field game has only two equilibria: The whole population always plays  $D$ , or the whole population always plays  $P$ . These equilibria are also equilibria for the finite-game. Yet, they both have a larger cost than the strategy of Equation (19). This leads us to say that the value of the game does not converge: the asymptotic cost of the strategy (19) is strictly smaller than the cost of any of the mean field equilibria.

## 5. Synchronous Players

As explained in the previous section, mean field games in continuous time appear naturally as the limit of  $N$ -player asynchronous games as  $N$  goes to infinity. In these asynchronous games with  $N$  players, only one player changes state at the same time. However, there are other situations where it is more natural to consider synchronous games in which, at each time step, all players take an action.

### 5.1. Synchronous $N$ -Player Games with Exchangeable Players

Here we consider a finite synchronous game  $\mathcal{G}_s^N = (N, \mathcal{E}, \mathcal{A}, \{P_a\}, \mathbf{M}_0, \{c_a\}, \beta)$  with  $N$  identical players with several differences from the model used in Section 3.1, the main one being the replacement of the rate matrices by stochastic matrices. As before, each Player  $n$  has an internal state  $X_n(t)$  that belongs to a finite state space  $\mathcal{E}$  ( $\mathbf{X}(t) = (X_0(t), \dots, X_{N-1}(t))$ ) and chooses an action from a finite action space  $\mathcal{A}$ . The main difference with the previous asynchronous model is that at each time step  $t \in \mathbb{Z}^+$ , all players choose an action  $A_n(t) \in \mathcal{A}$  simultaneously. We assume that, a player in state  $i$  who chooses action  $a$  goes to state  $j$  with probability  $P_{ija}(\mathbf{X}(t))$  and that, given  $\mathbf{X}(t)$ , the evolution of all players are independent. Furthermore, we assume that the players are *exchangeable*, *i.e.* for any permutation  $\sigma$  of the  $N$  players,  $P_{ija}(X_0(t), \dots, X_{N-1}(t)) = P_{ija}(X_{\sigma(0)}(t), \dots, X_{\sigma(N-1)}(t))$ . The fact that

---

<sup>2</sup>When the time horizon is finite, it is natural to consider Markovian strategies that depend on time.

all players are exchangeable implies that the dependence in  $\mathbf{X}(t)$  can be replaced by a dependence on the population distribution  $\mathbf{M}(t)$ . More precisely, for any vector state  $\mathbf{x}, \mathbf{y} \in \mathcal{E}^N$  and any action vector  $\mathbf{a} \in \mathcal{A}^N$ , one can write:

$$\mathbb{P}(\mathbf{X}(t+1) = \mathbf{j} | \mathbf{X}(t) = \mathbf{i}, \mathbf{A}(t) = \mathbf{a}, \mathcal{F}_t) = \prod_{n=1}^N P_{i_n j_n a_n}(\mathbf{M}(t)), \quad (20)$$

where  $\mathcal{F}_t$  is the natural filtration of the game up to time  $t$ ,  $\mathbf{m}$  is the population distribution of  $\mathbf{x}$  and  $\forall i, j \in \mathcal{E}, \forall a \in \mathcal{A}, P_{ija}(\mathbf{m})$  forms a stochastic matrix, continuous in  $\mathbf{m}$ .

The instantaneous cost at time  $t$  depends on actions and state at time  $t - 1$ , symmetric in all players, so it can be written as a function of the population distribution:  $c_{X_n(t), A_n(t)}(\mathbf{M}(t))$ , and a discount factor  $\delta$  at each time step. Given a strategy  $\pi^0$  used by Player 0 and a strategy  $\pi$  used by all the others, the expected cost of Player  $n$  is:

$$V^N(\pi^0, \pi) = \mathbb{E} \left[ (1 - \delta) \sum_{t=0}^{\infty} \delta^t c_{X_0(t), A_0(t)}(\mathbf{M}^\pi(t)) \begin{array}{l} A_0 \text{ is chosen w.r.t. } \pi^0 \\ A_{n'} \text{ is chosen w.r.t. } \pi \text{ if } n' \neq 0 \end{array} \right]. \quad (21)$$

## 5.2. Corresponding Mean field Game

Synchronous games also admit mean field game limits. To construct this limit, let us consider a strategy  $\pi$  such that  $\pi_{i,a}(\mathbf{m})$  is the probability for a player to choose action  $a$  given that she is in state  $i$  and that  $\mathbf{M}(t) = \mathbf{m}$ . Assume that  $\mathbf{M}(0)$  converges in probability to some  $\mathbf{m}(0)$  as  $N$  goes to infinity and that all players except Player 0 apply a strategy  $\pi$  that is continuous in  $\mathbf{m}$ . As shown in Theorem 1 in [19] (up to differences in notations, the mean field model in [19] is the same as Equation (20)), the population distribution  $\mathbf{M}^\pi(t)$  converges (in probability) to a deterministic quantity  $\mathbf{m}^\pi(t)$  as  $N$  goes to infinity.  $\mathbf{m}^\pi(t)$  is defined by

$$m_j^\pi(t+1) = \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} m_i^\pi(t) P_{i,j,a}(\mathbf{m}^\pi(t)) \pi_{i,a}(\mathbf{m}(t)). \quad (22)$$

We denote by  $\pi^0$  the strategy of Player 0. The probability that Player 0 is in state  $j \in \mathcal{E}$  evolves over time according to the following equation:

$$x_j(t+1) = \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} x_i(t) P_{i,j,a}(\mathbf{m}^\pi(t)) \pi_{i,a}^0(\mathbf{m}(t)). \quad (23)$$

In this case, the cost of Player 0, given by (21) becomes

$$V(\pi^0, \pi) = (1 - \delta) \sum_{t=0}^{\infty} \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} \delta^t x_i(t) c_{i,a}(\mathbf{m}^\pi(t)) \pi_{i,a}^0(\mathbf{m}(t)).$$

As the evolution of  $m$  is deterministic, for any closed loop strategy  $\pi_{i,a}(\mathbf{m}(t))$  and any initial condition  $\mathbf{m}(0)$ , there exists an open-loop strategy  $\pi_{i,a}(t)$  that leads to the same values for  $\mathbf{m}^\pi(t)$  and the same cost. Hence, for the mean field model, one can replace any state-dependent strategy  $\pi(\mathbf{m}(t))$  in the above equations by a time-dependent strategy  $\pi(t)$ .

Player 0 chooses the strategy that minimizes her expected cost. When Player 0 does so, we say it uses the best response to the mass strategy  $\pi$ .

$$BR(\pi) = \arg \min_{\pi^0} V(\pi^0, \pi).$$

A strategy is said to be a mean field equilibrium if it is a fixed point for the best response function, that is,

$$\pi^{MFE} \in BR(\pi^{MFE}).$$

One of the difficulties of the analysis of continuous time mean field game is that the elements under consideration (the population distribution, the population strategy, Player 0 strategy...) are continuous functions of time. In the discrete time case, the model gets significantly simplified since all the elements are vectors. Hence, the proof of the existence of a mean field equilibrium for continuous-time mean field game (Theorem 1) can be adapted to show that the following result.

**Theorem 3** (Mean Field Equilibrium Existence for Synchronous Games). *Any synchronous mean field game with discounted cost that satisfies Assumption (A1) for  $P$  and  $c$  respectively, has a mean field equilibrium.*

*Sketch of proof.* We first observe that the set of discrete-time open-loop policies is a compact and convex set. Thus, to finish the proof, we need to show that the best response function has a closed graph and it is convex. The former condition is true since the set of open-loop policies belongs to a finite dimensional space and from the continuity assumptions (A1). The last condition can be shown using the same arguments as in the proof of Theorem 1.  $\square$

### 5.2.1. An Important Special Case: Repeated Games

The classical repeated games with discounted costs and with identical players form a subclass of synchronous games, as defined here. To see this, let us first consider a static  $N$ -player matrix game  $G$  with symmetric cost:  $u(a_1, \dots, a_N)$  is the instantaneous cost of any player when the players use actions  $a_1, \dots, a_N$  respectively. Furthermore, we assume that  $u(a_1, \dots, a_N) = u(a_{\sigma_1}, \dots, a_{\sigma_N})$ , for any permutation  $\sigma$  of  $\{1, \dots, N\}$ . The players repeat the matrix game infinitely often and their cost

under strategy  $\pi^1, \dots, \pi^N$  is the discounted sum of the costs:

$$V^N(\pi^1, \pi^2, \dots, \pi^N) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t u(\pi^1(t), \pi^2(t), \dots, \pi^N(t)). \quad (24)$$

These games fit in our framework: The state of a player is merely her current action ( $\mathbf{X}(t) = \mathbf{A}(t)$ ) and the evolution of the state becomes trivial: Under state  $x = a$  and selecting action  $b$ , the next state does not depend on the other players and becomes  $b$  with probability one:  $P_{ab}(b, \mathbf{M}(t)) = 1$ . The cost of one player at each stage corresponds to an immediate cost  $c_{X_n(t), A_n(t)}(\mathbf{M}(t)) = u(\mathbf{X}(t))$  since the cost  $u$  only depends on the population distribution by symmetry. As for the total cost of a player, (24) coincides with (21), as long as all players in the same state use the same strategy.

### 5.3. The Folk Theorem Does Not Scale

The relation between equilibria of  $N$ -player games with their mean field limits is also complex in the discrete time case.

Let us first focus on results that concern the performance of mean field equilibria in the  $N$ -player game. The situation is almost similar to the continuous time case and resembles Theorem 2 (i) in the sense that if  $\pi$  is a mean field equilibrium, then under assumption (A1), there exists  $N_0$  such that for all  $N \geq N_0$ ,  $\pi$  is a local  $\varepsilon$ -equilibrium of the  $N$ -player game. The proof of this is essentially similar to the proof of Theorem 2.

Let us now consider the Nash equilibria of the  $N$ -player game. The situation is very different from the continuous time case because the state of all the players can change in one time unit in the discrete time while in continuous time, state can only change in small steps, one player at a time.

This has several consequences on the nature of equilibria under both models. As mentioned before, the Nash equilibria in the continuous time case may depend on the initial population distribution, but this is not the case here, so that there is more latitude for designing equilibria.

Let us consider the particular case of repeated games, introduced in Section 5.2.1. For this type of games, the set of equilibria can be characterized using the Folk Theorem for repeated games.

**Theorem 4** (Folk theorem, adapted from Theorem A in [18])). *Let  $G$  be a symmetric matrix game, and let  $V^*$  be the cost under the strategy that repeats the Nash equilibrium of the static game  $G$ . Then for any compatible<sup>3</sup> cost  $V$  smaller than  $V^*$ ,*

---

<sup>3</sup>In this context, a compatible cost is a cost that can be attained by at least one strategy.

there exists a discount factor  $\delta \in (0, 1)$  such that  $V^*$  is the cost of an equilibrium of the discounted repeated game.

Actually, for any  $V < V^*$ , the construction of an equilibrium whose cost is  $V$  is based on the “tit for tat” principle. We claim that none of these equilibria scale at the mean field limit. Let us consider the following example for a static game. Each player only has two strategies,  $D$  and  $C$ . If all players play  $D$ , the cost is  $-1$ . If all players play  $C$ , the cost is  $-2$ . If some players play  $D$  and others play  $C$ , then, all the players who play  $C$  get  $-2M_C$  while the players who play  $D$  get  $-3M_C - M_D$ , where  $M_C$  and  $M_D$  are the proportions of players playing  $C$  and  $D$  respectively. These costs correspond to the average costs obtained by a player in a matching game against a random opponent.

The unique Nash equilibrium of the static game is strategy  $(D, D, \dots, D)$ . The cost of the corresponding repeated game is  $(1 - \delta) \sum_t -\delta^t = -1$ .

Let us now consider the following strategy (denoted  $\pi^N$  in the following) for all players: Play  $D$  for  $k$  rounds then play  $C$  as long as every-other player has followed the same pattern, else play  $D$  forever. The cost of this strategy is between  $-1$  and  $-2$ :

$$(1 - \delta) \left( \sum_{t=0}^{k-1} -\delta^t + \sum_{t=k}^{\infty} -2\delta^t \right) = -1 - \delta^k.$$

The strategy  $\pi^N$  is an equilibrium of the finite game if  $\delta$  is large enough. Indeed, no player wants to deviate in the first  $k$  rounds, because her cost would increase: In the rounds after  $k$ , a deviation provides an immediate cost advantage, at the cost of being punished until the end of time, so that a larger enough  $\delta$  makes this non-profitable.

Let us now consider the mean field game setting. If the whole population uses the strategy  $\pi^N$  and if Player 0 uses the same strategy her cost becomes

$$\begin{aligned} V(\pi^N, \pi^N) &= (1 - \delta) \sum_{t=0}^{\infty} \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} \delta^t x_i(t) c_{i,a}(\mathbf{m}^\pi(t) \pi_{i,a}^0(\mathbf{m}(t))), \\ &= (1 - \delta) \left( \sum_{t=0}^{k-1} -\delta^t + \sum_{t=k}^{\infty} -2\delta^t \right) \\ &= -1 - \delta^k. \end{aligned}$$

However in the mean field setting, the best response of Player 0 to  $\pi^N$  is not  $\pi^N$  but the strategy  $\pi^D$  where she plays  $D$  all the time. Indeed in this case her total

cost becomes

$$\begin{aligned} V(\pi^D, \pi^N) &= (1 - \delta) \left( \sum_{t=0}^{k-1} -\delta^t + \sum_{t=k}^{\infty} -3\delta^t \right) \\ &= -1 - 2\delta^k. \end{aligned}$$

This shows that  $\pi^N$  is not a mean field equilibrium and a “free rider” player can take advantage of the fact that the population will not act against her.

#### 5.4. Finite Horizon Case

We now focus on the mean field games when objects evolve in discrete time over a finite horizon, 0 to  $T$ . In this case, the population distribution  $\mathbf{m}^\pi$  is defined by (22), which depends on the strategy of the mass  $\pi$ . We assume that Player 0 can choose her own strategy  $\pi^0$ . The expected cost of Player 0 is

$$V(\pi^0, \pi) = \sum_{t=0}^T \sum_{i \in \mathcal{E}} \sum_{a \in \mathcal{A}} x_i(t) c_{i,a}(\mathbf{m}^\pi(t)) \pi_{i,a}^0(\mathbf{m}(t)),$$

where  $x_i(t)$  is the probability that Player 0 is in state  $i$  at time  $t$ . The evolution of  $x_i(t)$  over time is described in (23).

Player 0 uses best response to a given population strategy  $\pi$ , which means that she selects the strategy  $\pi^0$  that minimizes her expected cost. We are interested in proving the existence of a mean field equilibrium which consists of finding a strategy that is a fixed-point for the best response function. In Section 5.2, we showed this for the discounted case. In the finite horizon case, the vectors have finite size and, as a consequence, it is immediate to show, using the same arguments of those required for the proof of Theorem 3, that any discrete time mean field game with finite horizon cost such that  $P$  and  $c$  satisfy Assumption (A1) has a mean field equilibrium. Again, the proof mimics the proof of the analog Theorem 1 in continuous time over a finite horizon.

## 6. Conclusions

In this article, we generalize the framework of discrete-space mean field games to the cases of non-convex costs and explicit interactions. They hit a good compromise between tractability (existence of an equilibria) and modelization power (including propagation and congestion behaviors). This model consists of a finite state space mean field game where the transition rates of the objects and the cost function of a generic object depend not only on the actions taken but also on the population

distribution. We also show that there exists a sub-class of Nash equilibria for  $N$ -player games that converge to mean field equilibria when the number of players goes to infinity. Outside of this class, and in particular for all equilibria using the “tit for tat” principle over which the Folk theorem is based, the convergence does not hold.

For future work, we are interested in finding conditions ensuring the uniqueness of the mean field equilibrium. We believe that monotony assumptions similar to assumptions in [21] are required to prove the existence of a unique mean field equilibrium in this model. On the other hand, another interesting open question concerns the convergence of  $N$ -players equilibria to mean field equilibria when the number of player grows. We believe that there exist many  $N$ -player games for which the only limiting equilibria are mean field equilibrium, for example when players have incomplete information about the game. It would be interesting to characterize the sub-class of strategies where convergence to mean field equilibria holds. Obviously, this class includes all local strategies (no information) and excludes some Markovian ones (full information).

- [1] S. Adlakha, R. Johari, and G. Y. Weintraub. Equilibria of dynamic games with many players: Existence, approximation, and market structure. *Journal of Economic Theory*, 2015.
- [2] N. I. Al-Najjar and R. Smorodinsky. Large nonanonymous repeated games. *Games and Economic Behavior*, 37:26–39, 2001.
- [3] D. M. Ambrose. Strong solutions for time-dependent mean field games with non-separable hamiltonians. *Journal de Mathématiques Pures et Appliquées*, 113:141 – 154, 2018.
- [4] R. Basna, A. Hilbert, and V. N. Kolokoltsov. An epsilon-nash equilibrium for non-linear markov games of mean-field-type on finite spaces. *Commun. Stoch. Anal*, 8(4):449–468, 2014.
- [5] E. Bayraktar and A. Cohen. Analysis of a finite state many player game using its master equation. *arXiv preprint arXiv:1707.02648*, 2017.
- [6] M. Benaïm and J.-Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11):823–838, 2008.
- [7] A. Bensoussan, J. Frehse, and P. Yam. *Mean field games and mean field type control theory*. Springer, 2013.
- [8] K. C. Border. *Fixed point theorems with applications to economics and game theory*. Cambridge university press, 1989.



- [9] 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] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. *SIAM Journal on Control and Optimization*, 51(4):2705–2734, 2013.
- [11] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. *SIAM J. Control Optim.*, 51(4):2705–2734, 2013.
- [12] R. Carmona, D. Lacker, et al. A probabilistic weak formulation of mean field games and applications. *The Annals of Applied Probability*, 25(3):1189–1231, 2015.
- [13] R. Carmona and P. Wang. Finite state mean field games with major and minor players. *arXiv preprint arXiv:1610.05408*, 2016.
- [14] A. Cecchin and M. Fischer. Probabilistic approach to finite state mean field games. *Applied Mathematics & Optimization*, Mar 2018.
- [15] P. Dasgupta and E. Maskin. The existence of equilibrium in discontinuous economic games, i: Theory. *Review of Economic Studies*, 53(1):1–26, 1986.
- [16] J. Doncel, N. Gast, and B. Gaujal. Are mean-field games the limits of finite stochastic games? *SIGMETRICS Perform. Eval. Rev.*, 44(2):18–20, Sept. 2016.
- [17] A. M. Fink. Equilibrium in a stochastic  $n$ -person game. *J. Sci. Hiroshima Univ. Ser. A-I Math.*, 28(1):89–93, 1964.
- [18] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, 1986.
- [19] N. Gast and B. Gaujal. A mean field approach for optimization in discrete time. *Discrete Event Dynamic Systems*, 21(1):63–101, 2011.
- [20] 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.
- [21] D. A. Gomes, J. Mohr, and R. R. Souza. Continuous time finite state mean field games. *Applied Mathematics & Optimization*, 68(1):99–143, 2013.
- [22] D. A. Gomes and E. Pimentel. Time-dependent mean-field games with logarithmic nonlinearities. *SIAM Journal on Mathematical Analysis*, 47(5):3798–3812, 2015.

- [23] D. A. Gomes, E. Pimentel, and H. Sánchez-Morgado. Time-dependent mean-field games in the superquadratic case. *ESAIM: Control, Optimisation and Calculus of Variations*, 22(2):562–580, 2016.
- [24] D. A. Gomes and E. A. Pimentel. Regularity for mean-field games systems with initial-initial boundary conditions: The subquadratic case. In *Dynamics, Games and Science*, pages 291–304. Springer, 2015.
- [25] D. A. Gomes, E. A. Pimentel, and H. Sánchez-Morgado. Time-dependent mean-field games in the subquadratic case. *Communications in Partial Differential Equations*, 40(1):40–76, 2015.
- [26] A. Granas and J. Dugundji. *Fixed point theory*. Springer Science & Business Media, 2013.
- [27] O. Guéant. Existence and uniqueness result for mean field games with congestion effect on graphs. *Applied Mathematics & Optimization*, 72(2):291–303, 2014.
- [28] O. Guéant, J.-M. Lasry, and P.-L. Lions. Mean field games and applications. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Mathematics*, pages 205–266. Springer Berlin Heidelberg, 2011.
- [29] M. Huang. Mean field stochastic games with discrete states and mixed players. In *Game Theory for Networks*, pages 138–151. Springer, 2012.
- [30] M. Huang, R. Malhame, and P. Caines. Large population stochastic dynamic games: Closed-loop mckean vlasov systems and the nash certainty equivalence principle. *Communications in Information and Systems*, 6(3):221–252, 2006. Special issue in honor of the 65th birthday of Tyrone Duncan.
- [31] D. Lacker. A general characterization of the mean field limit for stochastic differential games. *Probability Theory and Related Fields*, 165(3), Aug 2016.
- [32] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. i–le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006.
- [33] 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.
- [34] J.-M. Lasry and P.-L. Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.
- [35] H. Sabourian. Anonymous repeated games with a large number of players and random outcomes. *JOURNAL OF ECONOMIC THEORY*, 51:92–110, 1990.
- [36] W. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, 2010.

- [37] H. Tembine. Mean field stochastic games: convergence, q/h-learning and optimality. In *American Control Conference (ACC), 2011*, pages 2423–2428. IEEE, 2011.
- [38] H. Tembine, J.-Y. L. Boudec, R. El-Azouzi, and E. Altman. Mean field asymptotics of markov decision evolutionary games and teams. In *Game Theory for Networks, 2009. GameNets' 09. International Conference on*, pages 140–150. IEEE, 2009.
- [39] Z. Wang, C. T. Bauch, S. Bhattacharyya, A. d’Onofrio, P. Manfredi, M. Perc, N. Perra, M. Salathé, and D. Zhao. Statistical physics of vaccination. *Physics Reports*, 664:1–113, 2016.