Mean-Field Games with Explicit Interactions

Josu Doncel, Nicolas Gast, Bruno Gaujal

Mean-Field Games with Explicit Interactions☆

Josu Doncel[a,b], Nicolas Gast[a,b], Bruno Gaujal[a,b]

[a] *Univ. Grenoble Alpes, F-38000 Grenoble, France*

[b] *Inria*

Abstract

We introduce the mean-field games with explicit interactions. This model is a finite-state space mean-field game where the evolution and the cost function of the individual players depend not only on the actions taken, but also on the population distribution. We analyze these games in continuous and discrete time, over finite as well as infinite time horizons. We show the existence of a mean-field equilibrium in this type of games using an adapted version of Kakutani fixed point theorem. Besides, we also study the convergence of the equilibria of N-player games to mean-field equilibria. We define classes of strategies over which any equilibrium converges to a mean-field equilibrium when the number of players goes to infinity. We also exhibit equilibria outside this class that do not converge to mean-field equilibria. In discrete time the same non-convergence phenomenon implies that the Folk theorem does not scale to the mean field limit. Finally, we construct a mean-field game with explicit interaction to study vaccination strategies over an SIR infection propagation model and compute its mean field equilibrium is almost closed form. We also compare the Nash equilibrium with a centrally optimal strategy and show that, in all but degenerated cases, the equilibrium does not coincide with the optimal solution. We design a pricing mechanism that force the equilibrium to coincide with an optimal vaccination strategy.

1. Introduction

The history of game theory has been marked by milestone theorems concerning the existence of an equilibrium: The minimax theorem of Von Neumann [29], that shows the existence of an equilibrium for finite zero sum games; the famous result of Nash [24], showing the existence of an equilibrium in finite games; the folk theorem (see for example [8]) that characterizes Nash equilibria in repeated games or

☆This work is supported by the European FP7 Quanticol Project, 600708.

the mean-field game solution, introduced by Lasry and Lions [23], that shows the existence of equilibria in dynamic games with an infinite number of players.

In this paper, we introduce a variant of mean-field games and prove the existence of an equilibrium. We also show how this equilibrium is essentially different from the equilibria obtained when the number of players is finite.

In general, mean-field games study the strategic decision making in the presence of a huge number of rational indistinguishable players. An important assumption in this type of games is that no individual player action can affect the dynamics of the system (a.k.a. the population) and conversely, each player reacts to the distribution of the population over all player's states and not on the state of each individual player. The mean-field games considered here differ from previous definitions in one crucial point: the state dynamics of a player as well as her strategy depends on the aggregate state of the other players. This is called *explicit interaction* in the following. Other (milder) differences with the model of Lasry and Lions include the fact that we consider finite state spaces over a finite and/or infinite time horizon and the fact that players may play simultaneously and/or asynchronously. We claim that this 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 others. This type of behavior is classical in systems with a large number of interacting objects [2] and cannot to handled using previous mean-field game models. In the infection model, the rate of infection of one individual depends on the proportion of individual already infected. In the congestion case, one player can barely use a resource if it is already heavily loaded.

Mean-field games with finite state spaces have been studied in [10] for the discrete time case and in [11, 16] for example, in the continuous time case. Up to our knowledge, in all these cases, the dynamics of one player does not depend on the state of the other players. Also, the existence of a Nash equilibrium in these papers is proved for strictly convex costs. This is a strong assumption because it does not cover the case of average costs.

Here, we consider general mean-field games with finite state spaces and with explicit dependence on the population distribution. We show that such games admit a Nash equilibrium under mild continuity assumptions by applying the Kakutani fixed point theorem for infinite dimension spaces. The main idea of the proof is not to consider the best response operator directly but some sort of aggregate operator instead, that satisfies all the conditions needed to apply the fixed point theorem.

We prove this existence of a mean-field equilibrium in several case: over a finite or an infinite time horizon (with discounted costs) and under discrete time or continuous time dynamics.

In all four combinations, we prove that a mean-field equilibrium is an ϵ-approximation of an equilibrium of a corresponding game with a finite number N of players, where ϵ goes to 0 when N goes to infinity. However, conversely, not all equilibria for the finite version converge to a Nash equilibria of the mean-field limit of the game. We provide several counter-examples 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 seems to have been overlooked so far.

The model we analyze in this article is very general and we believe that a large variety of models fit in our framework. As an illustrative typical example, we present a vaccination strategy in an infection propagation model that we analyze thoroughly in Section 5. This is a typical example of a mean filed game with explicit interaction because the players become infected according to a mean-field dynamic: the rate of infection of one player at time t is a function of the proportion of infected players, $I(t)$.

The rest of the article is organized as follows.

Section 2 introduces mean-field games with explicit interactions in continuous time case with infinite horizon and discounted costs as well as finite horizon. 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 also 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 a N-player game inspired from the prisoner's dilemma whose equilibria are not for the limit mean-field game.

Section 4 considers the discrete time case. In this case, N-player games can be seen as classical repeated games. The mean-field limit dynamics is derived in discrete time and the existence of an equilibrium is proved similarly as in the continuous time case. Here counter-examples of equilibria for finite games that do not go to the limit are even more rampant. Indeed, the folk theorem applies and all equilibria based on retaliation cannot be equilibria at the limit.

Finally Section 5 is dedicated to the study of one mean-field game, namely a vaccination game against an infection propagation. In this SIRV model, the players can take 4 states: Susceptible, Infected, Recovered or Vaccinated. We consider that each player uses a a vaccination strategy whenever it is susceptible. The cost includes a vaccination cost as well as a unit time cost of being infected. The strategy of a generic object is to choose her vaccination strategy so as to minimize her expected cost over time. We show that there exists a mean-field equilibrium that is pure (deterministic) and of threshold type. More precisely, the mean-field equilibrium consists in vaccinating with maximum rate from time 0 up to some critical time after which the player does not vaccinate any longer. We also formulate the centralized optimization problem, where the goal is to find the population vaccination strategy so that the total cost is minimized. We show that the solution of a centralized problem (or global optimum) is also of threshold type. We observe numerically that, except for trivial cases, the mean-field equilibrium does not coincide with a global optimum. We also present a pricing mechanism incentivating individual objects to follow the global optimum strategy. We have developed simulations that provide numerical evidence that the vaccination cost must be subsidized to encourage selfish individuals to vaccinate optimally.

2. Continuous Time

2.1. Notations and Definitions

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 $\mathcal{S} = \{1, \dots, S\}$ and a finite action set $\mathcal{A} = \{1, \dots, A\}$.

A mixed *strategy* (or strategy for short) is a measurable function $\pi : \mathcal{S} \times \mathbb{R}^+ \to \mathcal{P}(\mathcal{A})$, that associates to each state $i \in \mathcal{S}$ and each time $t \geq 0$ a probability measure $\pi_i(t)$ on the set of possible actions, where $\mathcal{P}(\mathcal{A})$ is the set of probability measures over $\mathcal{A}$ (as $\mathcal{A}$ is finite, $\mathcal{P}(\mathcal{A})$ is the simplex). 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 π. For all $t \geq 0$ and all $i \in \mathcal{S}$, we have $\sum_{a \in \mathcal{A}} \pi_{i,a}(t) = 1$. We say that a strategy is deterministic (or 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 $a' \neq a$.

We denote by $\mathbf{m}^\pi(t) \in \mathcal{P}(\mathcal{S})$ the *population distribution* at time t. As the state space is finite, $\mathbf{m}^\pi(t)$ is a vector whose ith component, $m_i^\pi(t)$, is the proportion of players in state i at time t. We assume that the initial condition $\mathbf{m}^\pi(0) = \mathbf{m}_0$ is fixed. For $t \geq 0$, the population distribution $\mathbf{m}^\pi$ is the solution of the following differential equation, that depends on the strategy π: for $j \in \mathcal{S}$

$$\dot{m}_j^\pi(t) = \sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} m_i^\pi(t) Q_{ija}(\mathbf{m}^\pi(t)) \pi_{i,a}(t). \tag{1}$$

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

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 [10], 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* (1) *to linear dynamics.*

We now concentrate on a particular player, that we call player 0. Player 0 chooses her own strategy $\pi^0 : \mathbb{R} \times \mathcal{S} \to \mathcal{P}(\mathcal{A})$. We denote by $\mathbf{x}^{\pi^0}(t) \in \mathcal{P}(\mathcal{S})$ the probability distribution of Player 0 when player 0 uses strategy π^0 against a population who plays strategy π. The dependence of the state of player 0 on π is kept implicit to avoid heavy notations. For a given state $i \in \mathcal{S}$, $x_i^{\pi^0}(t)$ denotes the probability for Player 0 to be in state i at time t. The distribution $\mathbf{x}^{\pi^0}$ evolves over time according to the following differential equation: for $j \in \mathcal{S}$

$$\dot{x}_j^{\pi^0}(t) = \sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} x_i^{\pi^0}(t) \cdot Q_{ija}(\mathbf{m}^\pi(t)) \cdot \pi_{i,a}^0(t). \tag{2}$$

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

$$V(\pi^0, \pi) = \int_0^\infty \left(\sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} x_i^{\pi^0}(t) \cdot c_{i,a}(\mathbf{m}^\pi(t)) \cdot \pi_{i,a}^0(t) \cdot e^{-\beta t} \right) \, dt, \tag{3}$$

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

Example 1 (Epidemic model)**.** *In Section 5 we analyze an epidemic model. The state space of the players is $\mathcal{S} = \{$ susceptible (S), infected (I), recovered (R) or vaccinated (V)$\}$. The propagation of the infection follows a uniform contact process: An arbitrary player gets in contact with the other players according to a Poisson process, with rate γ. If the other player is infected, then infection is transmitted to the player. As for recovery, each player recovers with rate γ. A player can only take actions when she is in the susceptible state. It strategy consists in choosing (with rate τ), a vaccination probability at time t: $\pi(t) = p_v(t)$, and the set of pure actions is reduced to two points $\{v, \neg v, v\}$ (vaccinate or not). In that case, the dynamics of the population distribution is driven by Q with*

$$\begin{aligned}
Q_{S,I,\neg v}(\mathbf{m}^\pi(t)) &= Q_{S,I,v}(\mathbf{m}^\pi(t)) = \gamma m_I^\pi(t), \\
Q_{S,V,v}(\mathbf{m}^\pi(t)) &= \tau, \\
Q_{S,V,\neg v} &= 0, \\
Q_{I,R,v}(\mathbf{m}^\pi(t)) &= Q_{I,R,\neg v}(\mathbf{m}^\pi(t)) = \rho.
\end{aligned}$$

All the other entries of Q are null. Note that Q depends explicitly on the population distribution: the rate of infection depends on the proportion of infected players.

Player 0 is an independent player that chooses her own vaccination strategy π^0. The strategy π^0 does not modify $\mathbf{m}^\pi(t)$. In this model, the cost a player 0 is decomposed into a fixed vaccination cost plus an infection cost for the duration of her infection.

The goal of Player 0 is to choose a strategy π^0 that minimizes her discounted cost (3) when the rest of the population plays a strategy π. For a given population strategy π, we denote the best-responses of Player 0 to π by $BR(\pi)$. This set is the set of strategies that minimizes her discounted cost:

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

We then define a mean-field equilibrium as a strategy π^{MFE} such that when the population strategy is π^{MFE}, a selfish player 0 would also choose the same strategy π^{MFE}:

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

$$\pi^{MFE} \in BR(\pi^{MFE}). \tag{4}$$

A mean-field equilibrium is said to be pure if it is a deterministic 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 that of Player 0. In other words, for a given population strategy π, all the rational players of the populations (or players) choose the strategy $BR(\pi)$. A mean-field equilibrium is a situation where no player has incentive to deviate unilaterally from her strategy. In the case of the epidemic model of Section 5, a vaccination strategy π^{MFE} is a mean-field equilibrium if it coincides with the vaccination chosen by a selfish Player 0.

2.2. Existence of Mean-Field Equilibrium

We now show that, under very general assumptions, these games have a mean-field equilibrium. As for classical games, these equilibria are not necessarily *pure*. In the epidemic model that we study in Section 5, we show that there exists a pure mean-field equilibrium. Our proof relies on a generalization of Kakutani fixed point theorem to infinite dimensional spaces. The main difficulty of the proof is that the best-response function BR(π) is not a Kakutani map because we did not assume the cost function to be strictly convex (contrary to [12] for example).

Our technical condition is essentially to ensure that the differential equations (1), (2) and the cost equation (3) are well defined, which is guaranteed by Assumption (A1):

(A1) The function $\mathbf{m} \mapsto Q_{ija}(\mathbf{m})$ is Lipschitz-continuous in $\mathbf{m}$. The 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 both bounded by a finite value.

Theorem 1. *Any continuous time mean-field game with discounted cost that satisfies Assumption (A1) has a mean-field equilibrium.*

Proof. The best-response function $\pi \mapsto BR(\pi)$ is neither continuous nor hemi-continuous in general. As a result, we formulate the fixed point problem in an alternative manner by considering a fixed point in $\mathbf{m}$.

By Assumption (A1), the function $Q_{ija}(\mathbf{m})$ is continuous in $\mathbf{m}$. As $\mathbf{m} \in \mathcal{P}(\mathcal{S})$ leaves in a compact space, this shows that $Q_{ija}(\mathbf{m})$ is bounded by some constant L. Let $\mathcal{M}$ be the set of continuous functions from $\mathbb{R}^+$ to $\mathcal{P}(\mathcal{M})$ that are Lipschitz-continuous with constant LS. As a result, $\mathcal{M}$ includes the set of solutions of the differential equation (1). We equip $\mathcal{M}$ with the norm $\|\mathbf{m} - \mathbf{m}'\| = \sup_{i\in\mathcal{S},t\geq 0} |m_i(t) - m'_i(t)|\, e^{-\beta t}$. By the Arzela-Ascoli theorem, $\mathcal{M}$ is a compact space.

For a given population distribution $\mathbf{m} \in \mathcal{M}$, we say that a strategy π^0 is feasible for $(\mathbf{x}, \mathbf{m})$ if $\mathbf{x}$ satisfies (2). We then define

$$Y(\mathbf{x}, \mathbf{m}) = \inf_{\pi^0 \text{ feasible for } (\mathbf{x},\mathbf{m})} \int_0^\infty \left(\sum_{i,a} x_i(t)\pi^0_{i,a}(t)c_{i,a}(\mathbf{m}(t))e^{-\beta t} \right) dt, \tag{5}$$

with the convention that $Y(\mathbf{x}, \mathbf{m}) = +\infty$ if there exists no strategy feasible for $(\mathbf{x}, \mathbf{m})$. This definition is also valid for the functions $\mathbf{m} \in \mathcal{M}$ that do not satisfies (1) for any strategy π.

We now define $\phi : \mathcal{M} \to \mathcal{M}$ such that, for all $\mathbf{m} \in \mathcal{M}$,

$$\phi(\mathbf{m}) = \arg\min_{\mathbf{x}\in\mathcal{M}} Y(\mathbf{x}, \mathbf{m}) \tag{6}$$

Note that for all $\mathbf{x} \in \phi(\mathbf{m})$, there exists a strategy such that π^0 is feasible for $(\mathbf{x}, \mathbf{m})$.

The quantity $\phi(\mathbf{m})$ is the best-response of Player 0 to the population $\mathbf{m}$. Hence, if we show that $\phi(\cdot)$ has fixed point $\mathbf{m} \in \phi(\mathbf{m})$, there exists a strategy π^0 such that π^0 is feasible for $(\mathbf{m}, \mathbf{m})$. This implies that $\mathbf{m}$ satisfies (1) from which we conclude that there exists a mean-field equilibrium for the mean-field games with explicit interactions.

We first show that the function $\phi(\cdot)$ satisfies the following properties.

Lemma 1. *(a) For all* $\mathbf{m} \in \mathcal{M}$*,* $\phi(\mathbf{m})$ *is convex,*

(b) The graph of $\mathbf{m} \mapsto \phi(\mathbf{m})$ *is closed, that is: for any sequences* $\mathbf{m}_n \in \mathcal{M}$ *and* $\mathbf{x}_n \in \phi(\mathbf{m}_n)$ *such that* $\lim_{n\to\infty} \mathbf{m}_n = \mathbf{m}$ *and* $\lim_{n\to\infty} x_n = x$*, we have* $x \in \phi(\mathbf{m})$*.*

Proof. (a) Let $\mathbf{x}$ be a trajectory. Replacing the quantity $x_i(t)\pi^0_{i,a}(t)$ by $z_{i,a}(t)$ in (5), $Y(\mathbf{x}, \mathbf{m})$ can be written as

$$Y(\mathbf{x}, \mathbf{m}) = \min_{\mathbf{z}} \int_0^\infty \left(\sum_{i,a} z_{i,a}(t) c_{i,a}(\mathbf{m}(t)) e^{-\beta t} \right) dt \tag{7}$$

$$\text{where } \mathbf{z} \text{ satisfies } \begin{cases} \sum_a z_{j,a}(t) = x_j(t) \forall j \in \mathcal{S}, \\ z_{j,a}(t) \geq 0, \forall j \in \mathcal{S} \forall a \in \mathcal{A}, \\ \dot{x}_j(t) = \sum_{i,a} z_{i,a}(t) Q_{ija}(\mathbf{m}(t)) \forall j \in \mathcal{S}. \end{cases} \tag{8}$$

Let $\mathbf{x}_1 \in \phi(\mathbf{m})$ and $\mathbf{x}_2 \in \phi(\mathbf{m})$ and let $\mathbf{x}_3 = \alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$ some some $\alpha \in (0, 1)$. We show that $\mathbf{x}_3 \in \phi(\mathbf{m})$. There is $\mathbf{z}_1$ and $\mathbf{z}_2$ that minimize (7) for $\mathbf{x}_1$ and $\mathbf{x}_2$ and satisfies the above constraints. As the constraints are linear, it follows immediately that $\mathbf{z}_3 = \alpha \mathbf{z}_1 + (1-\alpha)\mathbf{z}_2$ satisfies the same constraints and also minimizes (7).

(b) Since $\mathbf{x}_n \in \phi(\mathbf{m}_n)$, there exists a strategy π_n such that $\mathbf{x}_n$ verifies (2) and minimizes $Y(\mathbf{x}_n, \mathbf{m}_n)$. We note that $\pi_n \in L^2(\mathbb{R})$, which is a reflexive space. Therefore, we know that the sequence π_n has a subsequence that converges weakly to π. Using the continuity assumptions (A1), then $\mathbf{x} \in \phi(\mathbf{m})$.

□

Lemma 1 shows that ϕ has convex and compact values. Moreover, the function $\phi(\cdot)$ is upper-semicontinuous since it satisfies the property of Lemma 1(b) and $\mathcal{M}$ is compact [4, Prop. 11.11]. Hence, $\phi(\cdot)$ is satisfies the conditions of [13, Theorem 8.6.] and therefore it has a fixed point. □

2.3. Finite Horizon Case

The above formulation generalizes directly to a finite-horizon setting where each player tries to minimize her cost over a finite horizon T. As in the discounted case, the evolution over time of the population distribution $\mathbf{m}^\pi$ is given by (1) and the evolution of Player 0's distribution is given by (2).

Given the population strategy π and Player 0 strategy π^0, the expect cost of Player 0 for the finite horizon case is defined as follows:

$$V(\pi^0, \pi) = \int_0^T \left(\sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} x_i(t) c_{i,a}(\mathbf{m}^\pi(t)) \pi^0_{i,a}(t) \right) dt. \tag{9}$$

In the literature, there are similar models considering continuous time finite state space mean-field games with finite horizon. For example, the authors in [11] consider uniformly convex cost functions and in [14] strictly convex cost functions. In our model, we assume that the costs are continuous but, as it can be observed, the instantaneous cost of Player 0 is linear in π^0. Therefore, the model we study in this work is not covered by these models.

We define the notion of mean-field equilibrium for the finite horizon case as in the discounted case, by replacing the cost function (3) by (9). Then, the proof of Theorem 1 applies *mutatis mutandis* to show the existence of a mean-field equilibrium in this case.

Theorem 2. *Any continuous time mean-field game with finite horizon cost that satisfies Assumption (A1) has a mean-field equilibrium.*

3. Convergence of Finite Games to Mean-field Game

In this section, we will show that mean-field equilibria are the limits of a sub-class of Nash-equilibra when the number of players goes to infinity.

3.1. Mean-field Markov Game with N Interacting players

Our model is motivated by a game with N players that evolve in continuous time. Each player becomes active according to a Poisson process, independently of the others. When a player becomes active, it chooses randomly among the others a finite number of players. Each of these players then takes an action that influence the payoff of the various players and the evolution of their internal state.

When observing the system only when the players interact, this model can be viewed as a discrete-time model. Players can only take decisions to maximize their expected payoff over an infinite time-horizon at discrete times in $\mathcal{T}_N = \{0, 1/N, 2/N, \dots\}$. Each player has an internal state that evolve over time. The state of player n at time t is denoted by $X_n(t)$. The set of possible states of a player is finite and is denoted by $\mathcal{S}$. The global description of the system at time t is $\mathbf{X}(t) = (X_1(t) \dots X_N(t)) \in \mathcal{S}^N$.

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

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

At each time step $t \in \mathcal{T}_N$, a finite revision set $R(t)$ of players is defined, according to an IID process over all the sets of players. Only the players in $R(t)$ can revise their action at time t. This action is fixed until the next time this player is chosen again. We assume that for all $i \neq j \in \mathcal{S}$, there exists a function $Q_{ij} : \mathcal{A} \times \mathcal{P}(\mathcal{S}) \to \mathbb{R}^+$ that is Lipschitz-continuous in m and such that, under $\mathcal{F}_t$, the natural filtration of the process,

$$\mathbb{P}\left(X_n(t+\frac{1}{N}) = j \middle| X_n(t) = i, n \in R(t), \mathbf{M}(t) = \mathbf{m}, A_n(t) = a, \mathcal{F}_t\right) = \frac{1}{\mathbb{E}[|R(t)|]} Q_{ij}(a, \mathbf{m}), \quad (10)$$

where $\mathcal{F}_t$ is the natural filtration of the process.

At time $t \in \mathcal{T}_N$, the 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 π^n from some set of admissible strategies Π, in order to minimize her expected discounted payoff, knowing the strategies of the others. As before, the discount factor is denoted β. 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 $V(\pi^n, \pi)$ the expected discounted payoff of player n:

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

An equilibrium for this game is a strategy π such that a player has no another admissible strategy that leads to a higher payoff. This notion depends naturally on the set of admissible strategies.

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

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

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

3.2. Possible set of admissible strategies

In a *full information* setting, $A_n(t)$ is a (possibly random) functions 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 describe. Therefore, in the following, we will consider three possible restrictions for the set of admissible strategies:

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

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

  This definition is motivated by the fact that, as indicated by Equation (10), 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 strategy is dominant among the set of full-information strategy: there exists a full-information best-response for player n that is a Markov strategy.

- (Local) – We call a strategy π a *local strategy* if the choice of the action only depends on the player's internal state and on the time.

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

  Note that we allow this strategy to depend on time because $\mathbf{M}(t)$ evolve with time.

3.3. Technical assumptions and limiting regimes

In addition to Assumption 1 that ensures the continuity of the functions $Q_{ij}(a, m)$ and $c_{i,a}(m)$ in m, we will also assume Assumption 2:

(A2) the number of players that are selected at each point in $\mathcal{T}_N$ has a bounded second moment, *i.e.*, that there exists $B < \infty$ such that for all $t \in \mathcal{T}$: $\mathbb{E}\left[|R(t)|^2\right] < B$.

The next theorem provides an equivalence between local equilibria and mean-field equilibria. 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 equilibrium.

Theorem 3. *Assume that (A1) and (A2) hold. Then:*

(i) Let π be a mean-field equilibrium. There exists N_0 such that for all $N \geq N_0$, π is a local ε-equilibrium of the N player game.

(ii) If $(\pi^N)_N$ is a sequence of local strategies such that π^N is a local-equilibrium for the N player game, then, any sub-sequence of the sequence $(\pi^N)_N$ has a sub-sequence that converges weakly to a mean-field equilibrium.

Proof. The main technical difficulty of the proof is to show that for any local strategies π^n, π, $V^N(\pi^n, \pi)$ converges to $V(\pi^n, \pi)$ uniformly in π, π'. Uniform convergence follows from Theorem 3.3.2 in [28]. Indeed, local strategies as defined here are equivalent to stationary strategies, as defined in [28].

Thus, for any ε, there exists N_0 such that $N \geq N_0$ implies that $\left|V^N(\pi^n, \pi) - V(\pi^n, \pi)\right| \leq \varepsilon/2$. Hence, if π is a mean-field equilibrium, this implies that for any local strategy π^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 π^N is a sequence of local strategies, then any sub-sequence has a sub-sequence that converge weakly to some local strategy π^∞. As $V(\pi^n, \pi)$ is continuous in π^n and π (for the weak topology), this implies that $V(\pi^\infty) \leq V(\pi^n, \pi)$ for all local strategy π^n. $\square$

3.4. Markov equilibria do not converge to MFE

We now show that this theorem does not generalize to Markov strategies. The main ingredient used to construct the following counterexample, is the "tit-for-tat" principle. This idea can be used to define equilibria for any N-player game but cannot be used in mean-field games. 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.

We consider a mean-field version of the classical prisoner's dilemma. The state space of a player is $\mathcal{S} = \{C, D\}$ (that stand for Cooperate and Defect) and the action set is the same $\mathcal{A} = \mathcal{S}$. 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:

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

This cost function corresponds to at each time step, a player plays her strategy 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, by using that $x_C(t) + x_D(t) = m_C(t) + m_D(t) = 1$, the expected cost (given by (3)) of a Player 0 that has a state vector x while the mean-field is $m(t)$ is

$$\int_0^\infty [x_C(t)m_C(t) + 3x_C(t)m_D(t) + 2x_D(t)m_D(t)]e^{-\beta t}dt$$
$$= \int_0^\infty [x_C(t) + 2m_D(t)]e^{-\beta t}dt.$$

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:

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

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

Assume that all players, except player n, play the strategy π 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 $m_C = 1$ and if player n is picked, then:

- If player n applies π, she will suffer a cost $\frac{1}{N}\sum_{i=0}^{\infty} e^{-\beta i/N} \approx \int \exp(-\beta t)dt = 1/\beta$.
- If player n deviates from π at this time step and chooses the action D, all players will also deviate after the next time step. This implies that $m_D(t) = 1 - \exp(-t)$ 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.

This shows that when $\beta < 1$, player n has no incentive to deviate from the strategy π and that therefore, π is a Nash equilibrium.

3.4.1. Finite-time horizon

In the finite-horizon case, the above strategy π 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 is when all player play D, which coincide with the MFE.

Yet, a similar example also exist for finite-time horizon. Let us consider the following payoff matrix:

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

The setting is similar to the previous example: the action set is equal to the state state $\mathcal{S} = \mathcal{A} = \{C, D, P\}$ 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 that, when T is large enough, the following strategy is a Nash equilibrium:

- if $t < 1$, play C if $m_C = 1$, play P otherwise
- if $t \geq 1$, play D if $m_P = 0$, play P otherwise.

Here, 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 P is also a Nash equilibrium.

4. Discrete Time

As explained in the previous section, mean-field games in continuous time appears naturally as the limit of N-player asynchronous games as N goes to infinity. In asynchronous games with N players, only a small number of the players change state at each time-step. However, in many games, it is often more natural to consider synchronous games in which, at each time step, all players take an action.

4.1. Synchronous N-player game

We consider a game with N identical players with several differences from the model used in Section 3.1. Each player n has an internal state $X_n(t)$ that belongs to a finite state space $\mathcal{S}$ ($\mathbf{X}(t) = (X_n(t), \dots, X_N(t))$ and chooses an action from a finite action space $\mathcal{A}$. The main difference with the 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 x who chooses action a goes to state y with probability $P_{xy}(a, X(t))$ and that given $X(t)$, the evolution of all players are independent. Furthermore, the fact that all players are interchangeable 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 $\mathbf{x}, \mathbf{y} \in \mathcal{S}^N$ and any action vector $\mathbf{a} \in \mathcal{A}^N$, we have:

$$\mathbb{P}\left(\mathbf{X}(t+1) = \mathbf{y} | \mathbf{X}(t) = \mathbf{x}, \mathbf{A}(t) = \mathbf{a}, \mathcal{F}_t\right) = \prod_{n=1}^{N} P_{x_n y_n}(a_i, \mathbf{m}), \tag{11}$$

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 $P_{xy}(a, \mathbf{m})$ forms a stochastic matrix, continuous of $\mathbf{m}$.

We consider an instantaneous cost at time t, that 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 δ at each time step. Given a strategy π^n used by player n and a strategy π used by all the others, the expected cost of player n is:

$$V^N(\pi^n, \pi) = \mathbb{E}\left[(1-\delta) \sum_{t=0}^{\infty} \delta^t c_{X_n(t), A_n(t)}(\mathbf{M}^\pi(t)) \,\middle|\, \begin{array}{l} A_n \text{ is chosen w.r.t. } \pi^n \\ A_{n'} \text{ is chosen w.r.t. } \pi \ (n' \neq n) \end{array} \right]. \tag{12}$$

4.1.1. A particular case: Repeated games

The classical repeated games with discounted costs and with identical players can be defined as follows. Let us first consider a static N-player matricial game G with symmetric payoff: $u(a_1, \dots, a_N)$ is the payoff 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 σ of $\{1, \cdots, N\}$. The players repeat the matricial game infinitely often and their reward under strategy $\pi^1, \cdots, \pi^N$ is the discounted sum of the payoffs:

$$V^N(\pi^1, \pi^2, \cdots, \pi^N) = (1-\delta) \sum_{t=0}^{\infty} \delta^t u(\pi^1(t), \pi^2(t), \cdots, \pi^N(t)). \tag{13}$$

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 $x = b$ and selecting action a, the next state does not depend on the other players and becomes a with probability one: $P_{ba}(a, \mathbf{M}(t)) = 1$. As for the payoff u (or cost) of one player at each stage, it corresponds to an immediate cost

$c_{X_n(t),A_n(t)}(\mathbf{M}(t)) = -u(\mathbf{X}(t))$ since the payoff u only depends on the population distribution by symmetry. As for the total reward of a player, (13) coincides with (12), as long as all players in the same state use the same strategy.

4.2. Mean-field limit

Let us consider a strategy π 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 $\mathbf{m}(0)$ as N goes to infinity and that all players except player n apply a strategy π that is continuous in $\mathbf{m}$. As shown in Theorem 1 in [9] (up to differences in notations, the mean-field model in [9] is the same as Equation (11)), 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

$$m_j^\pi(t+1) = \sum_{i\in\mathcal{S}}\sum_{a\in\mathcal{A}} m_i^\pi(t)\cdot P_{i,j,a}(\mathbf{m}^\pi(t))\cdot \pi_{i,a}(\mathbf{m}(t)). \tag{14}$$

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

$$x_j(t+1) = \sum_{i\in\mathcal{S}}\sum_{a\in\mathcal{A}} x_i(t)\cdot P_{i,j,a}(\mathbf{m}^\pi(t))\cdot \pi^0_{i,a}(\mathbf{m}(t)). \tag{15}$$

In this case, the cost of Player 0, given by (12) becomes

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

As the evolution of m is deterministic, for any close 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 all 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, which depends as well on the strategy π. When Player 0 does so, we say it does the best-response to the mass strategy π.

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

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 the 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 4 (Mean-Field Equilibrium Existence in Discrete Time and Discounted Case)**.** *Any discrete time 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 Lemma $1(b)$. □

4.3. The Folk Theorem does not scale

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

Let us first focus on results that concerns the performance of a mean-field equilibria in the N-player game. The situation is almost similar to the continuous time case and resembles Theorem 3 (i)

Theorem 5. *Let π be a mean-field equilibrium. There exists N_0 such that for all $N \geq N_0$, π is a local ε-equilibrium of the N-player game.*

Proof. The proof is essentially similar to the proof of Theorem 3. □

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 of 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 4.1.1. For this type of games, the set of equilibria can be characterized using the folk theorem.

Theorem 6 (Folk theorem, adapted from Th. A in [8])**.** *Let G be a symmetric matricial game, and let V^* be the reward under the strategy that repeats the Nash equilibrium of the static game G. Then any feasible reward V larger than V^* is the reward of an equilibrium of the repeated game if the discount factor β is large enough.*

Actually, for any $V > V^*$, the construction of an equilibrium whose reward is V is based on the "reward and punishment" 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 payoff is 1. If all players play C, the payoff is 2. If some players play D and others play C, then, all the players who play C get 0 while the players who play D get 3.

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

Let us now consider the following strategy (called π^* 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 for ever. The reward of this strategy is between 1 and 2:

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

The strategy π^* is an equilibrium if δ is large enough. Indeed, no player wants to deviate in the first k rounds, because her gain would decrease from 1 to 0 and then stay at 1. In the rounds after k, a deviation provides an immediate payoff advantage of 3 instead of 1, at the cost of being punished until the end of time, so that as a larger enough δ makes this non-profitable.

Let us now consider the mean-field limit of this game when the whole population uses the strategy π^*.

If one player uses the same strategy her reward becomes

$$\begin{aligned}
V(\pi^*, \pi^*) &= (1-\delta)\sum_{t=0}^{\infty}\sum_{i\in\mathcal{S}}\sum_{a\in\mathcal{A}} \delta^t x_i(t)\cdot c_{i,a}(\mathbf{m}^{\pi}(t)\cdot \pi^0_{i,a}(\mathbf{m}(t)),\\
&= (1-\delta)(\sum_{t=0}^{k-1}\delta^t + \sum_{t=k}^{\infty} 2\delta^t)\\
&= 1+\delta^k.
\end{aligned}$$

However her best response to π^* is not π^* but the strategy π^D where she plays D

all the time. Indeed her total reward becomes

$$V(\pi^D, \pi^*) = (1-\delta)(\sum_{t=0}^{k-1} \delta^t + \sum_{t=k}^{\infty} 3\delta^t)$$
$$= 1 + 2\delta^k.$$

Therefore, π^* is not a mean-field equilibrium.

4.4. Finite Horizon Case

We now focus on the mean-field games when objects evolve in discrete time time over a finite horizon, 0 to T. For this case, the population distribution $\mathbf{m}^\pi$ is defined by (14), which depends on the strategy of the mass π. We assume that Player 0 can choose her own strategy π^0. The expected cost of Player 0 is

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

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 (15).

Player 0 does best-response to a given population strategy π, which means that she selects the strategy π^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 4.2, 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 4, the desired result.

Theorem 7 (Mean-Field Equilibrium Existence in Discrete Time and Finite Horizon Case)**.** *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 2 in continuous time over a finite horizon.

5. Illustrative Example: Epidemic Model with Vaccinations

In this section, we study an epidemic model with vaccinations. As we will see, this model satisfies the conditions of Theorem 2 and, as a consequence, it has a mean-field equilibrium. For this particular case, we show the existence of a pure

Nash equilibrium and compare the performance of this equilibrium with a centralized allocation.

We consider a population of homogeneous players that evolve in continuous time from time 0 to T. The state-space of a player is $\{S, I, R, V\}$, which respectively stand for susceptible, infected, recovered and vaccinated. We denote by $m_S(t)$, $m_I(t)$, $m_R(t)$ and $m_V(t)$ the proportion of the population that is, respectively, susceptible, infected, recovered and vaccinated at time t.

The dynamics of one player is a Markov process that can be described as follows. A player encounters other players with rate γ. If the initial player is susceptible and the encounter is infected, the first player becomes infected. An infected player recovers at rate ρ. We also consider that susceptible population can get vaccinated with strategy π, where $\pi(t) \in [0, \tau]$. We consider that $\tau < \infty$. Once an player is vaccinated or recovered, her state does not change. The Markovian behavior of a player is displayed in Figure 1.

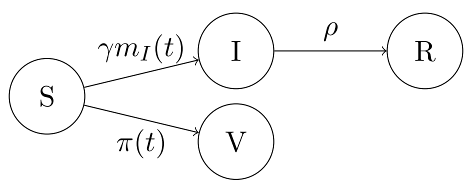


Figure 1: The dynamics of a player in the epidemic model. A player has four possible states: S (susceptible), I (infected), R (recovered) and V (vaccinated).

We are interested in the analysis of this epidemic model at the mean field limit. When the number of players goes to infinity, the dynamics of the population is given by the following system of differential equations that is typical of population dynamics studied in [18] for the non-controlled case. This system corresponds to Equation (1) using the form of the rate matrix Q given in Example 1. When all players apply the same strategy π, this system of differential equations is

$$\begin{cases} \dot{m}_S^\pi(t) = -\gamma m_S^\pi(t) m_I^\pi(t) - \pi(t) m_S^\pi(t) \\ \dot{m}_I^\pi(t) = \gamma m_S^\pi(t) m_I^\pi(t) - \rho m_I^\pi(t) \\ \dot{m}_R^\pi(t) = \rho m_I^\pi(t) \\ \dot{m}_V^\pi(t) = \pi(t) m_S^\pi(t) \end{cases}$$

In [20] the authors develop an approximation of this epidemic model, based on independence assumptions, and characterize the solution of the corresponding mean-field game. In the following subsection, we show that the mean-field game corresponding to this model is tractable and can be analyzed rigorously without making approximations.

5.1. Mean-Field Game

We focus on a particular player, that we call Player 0. Player 0 chooses her vaccination strategy π^0, where $\pi^0(t) \in [0,\tau]$. Let $x_i^{\pi^0}(t)$ be the probability that Player 0 is at time t in state i, where $i \in \{S, I, R, V\}$. The quantity $x^{\pi^0}(t)$ satisfies the following system of differential equations:

$$\begin{cases} \dot{x}_S^{\pi^0,\pi}(t) = -\gamma x_S^{\pi^0}(t) m_I^{\pi}(t) - \pi(t) x_S^{\pi^0}(t) \\ \dot{x}_I^{\pi^0,\pi}(t) = \gamma x_S^{\pi^0}(t) m_I^{\pi}(t) - \rho x_I^{\pi^0}(t) \\ \dot{x}_R^{\pi^0,\pi}(t) = \rho x_I^{\pi^0}(t) \\ \dot{x}_V^{\pi^0,\pi}(t) = \pi(t) x_S^{\pi^0}(t) \end{cases}$$

Using the foregoing notations, the expected individual cost of Player 0 is defined as follows:

$$V(\pi^0, \pi) = \int_0^T (c_V \pi^0(t) x_S^{\pi^0}(t) + c_I x_I^{\pi^0}(t)) dt,$$

where c_V is the vaccination cost and c_I is the unit time cost of being infected.

The rate at which susceptible population becomes infected is linear in the proportion of infected population. Further, the rest of the rates and instantaneous costs do not depend on the population distribution. As a result, the conditions of Assumption (A1) are satisfied in this model, which implies the existence of a mean-field equilibrium. In the following, we show that it can be also characterized.

We call the *best response to* π and denote by $\mathrm{BR}(\pi)$ the set of vaccination strategies that minimize the cost of Player 0 for a given the population strategy π. We model this minimization problem as a Continuous Time Markov Decision Process with finite horizon T [25]. We uniformize the Continuous Time Markov Decision Process and let $\mu \geq \max\{\gamma + \tau,\ \rho\}$ be the uniformization constant. We denote by $J_S(t)$ (resp. $J_I(t)$) the optimal cost in the susceptible state (resp. the infected state) at time t and thus, we write the optimality equations of the associated discrete time Markov Decision Process as follows: for all $t = 0, \ldots, T - \frac{1}{\mu}$,

$$J_S(t) = \inf_{\pi^0(t) \in [0,\tau]} \left[G(t) + \frac{\pi^0(t)}{\mu} \left(c_V - J_S(t + \frac{1}{\mu}) \right) \right], \tag{16}$$

$$J_I(t) = \frac{c_I}{\mu} + J_I(t + \frac{1}{\mu}) \left(1 - \frac{\rho}{\mu} \right), \tag{17}$$

where $J_S(T) = J_I(T) = 0$ and

$$G(t) = \left(1 - \frac{\gamma I(t)}{\mu} \right) J_S(t + \frac{1}{\mu}) + \frac{\gamma I(t)}{\mu} J_I(t + \frac{1}{\mu}). \tag{18}$$

We know that $\mathrm{BR}(\pi)$ coincides with the solution of this optimality equations. We aim to show that it is of threshold type. Prior to that, we present the following results.

Lemma 2. *The functions J_S and J_I satisfy that:*

(a) $J_I(t)$ is decreasing with t.

(b) $J_I(t) \geq J_S(t)$ for all t.

(c) $J_S(t)$ is decreasing with t when $\mathrm{BR}(\pi)(t) = 0$.

Proof. (a) We prove that $J_I(t)$ is decreasing with t by an induction on t. First, we observe that $J_I(T - \frac{1}{\mu}) = c_I/\mu > J_I(T) = 0$. Assume now that $J_I(t)$ is decreasing with t for all $t = t_0 + \frac{1}{\mu}, \ldots, T$. Using the induction assumption and that $J_I(t_0 + \frac{1}{\mu})$ and $\left(1 - \frac{\rho}{\mu}\right)$ are positive, we obtain that $J_I(t_0) > J_I(t_0 + \frac{1}{\mu})$ and the proof finishes.

(b) We show this result by induction on t. We first observe that $J_S(T - \frac{1}{\mu}) < J_I(T - \frac{1}{\mu})$. Assume now that this holds for all $t = t_0 + \frac{1}{\mu}, \ldots, T$. The proof ends if we show that $J_S(t_0) < J_I(t_0)$. First, from (16), it follows that $J_S(t_0) \leq G(t)$, where $G(t)$ is as defined in (18). From the induction assumption that states that $J_S(t_0 + \frac{1}{\mu}) < J_I(t_0 + \frac{1}{\mu})$, we have that $G(t)$ is less or equal than $J_I(t_0 + \frac{1}{\mu})$. Finally, from Lemma 2*(a)*, we know that $J_I(t)$ is decreasing with t and the desired result follows.

(c) First, we observe that if $\mathrm{BR}(\pi)(t) = 0$, then $J_S(t) = G(t)$, where $G(t)$ is as defined in (18). Besides, using Lemma 2*(b)*, we obtain that $G(t)$ is greater or equal than $J_S(t + \frac{1}{\mu})$. Therefore, we have shown that, if $\mathrm{BR}(\pi)(t) = 0$, $J_S(t) \geq J_S(t + \frac{1}{\mu})$.

□

We say that a strategy π^0 is a threshold strategy (or a strategy with threshold t_0). if there exists $t_0 \in [0, T]$ such that

$$\pi^0(t) = \begin{cases} \tau & \text{if } t \leq t_0 \\ 0 & \text{if } t > t_0 \end{cases}$$

We are now ready for the following proposition.

Proposition 1. *For any population strategy π, there exists a best-response π^0 that is a threshold strategy.*

Proof. Let π^0 be a best-response strategy. We first notice that from (16), that $\pi^0(t) = 0$ if $J_S(t) < c_V$ and $\pi^0(t) = \tau$ if $J_S(t) > c_V$. We now distinguish two cases.

The first case is if $J_S(0) < c_V$ which implies $\pi^0(0) = 0$. According to Lemma 2*(c)*, $J_S(t)$ is never higher than c_V, which implies that
$pi^0(t) = 0$ for all t.

We now study the case $J_S(0) > c_V$. Since $J_S(T) = 0$, $J_S(t)$ must be less than c_V at some time. Let t_0 be the value such that $c_V > J_S(t_0 + \frac{1}{\mu})$ for the first time. Then $\pi^0(t) = \tau$ for all $t < t_0$. Moreover, we have that $\pi^0(t_0) = 0$ and, from Lemma 2(c), it follows $\pi^0(t) = 0$ for all $t \geq t_0$. $\square$

This Proposition 1 implies that, for a strategy π with threshold t, there exists a best-response that is a threshold strategy. Let us denote by $t_{BR(t)}$ the threshold of such a best-response strategy to the strategy with threshold t. We show the following relation that the threshold of the population and of Player 0 satisfy.

Lemma 3. *The threshold $t_{BR(t)}$ decreases when t increases.*

Proof. We first observe that if t increases, then the number of vaccinated population increases, which implies that the number of infected population $m_I^\pi(t)$ decreases. From Lemma 2*(b)*, we know that $J_I(t) \geq J_S(t)$ for all t. Thus, in (16) $m_I(t)$ is multiplied by $J_I(t + \frac{1}{\mu}) - J_S(t + \frac{1}{\mu})$, which is positive. Therefore, if the number of infected population $m_I(t)$ decreases then $J_S(t)$ also decreases. Finally, from (16), we have that if $J_S(t)$ decreases, then $t_{BR(t)}$ also decreases. $\square$

A mean-field equilibrium for this epidemic model is a vaccination strategy π^{MFE} such that when the population chooses the vaccination strategy π^{MFE}, a selfish Player 0 would also choose the same vaccination strategy π^{MFE}. Hence, a strategy with threshold t is a mean-field equilibrium if $t_{BR(t)} = t$. From Lemma 3, we have that the thresholds of both strategies meet in a unique point, which gives the following proposition.

Proposition 2. *There exists a pure mean-field equilibrium that is a threshold strategy.*

We recall that, in Theorem 1, we show the existence of a mean-field equilibrium. In addition, for this epidemic model, we prove not only the existence but also that this equilibrium is pure and is a threshold strategy. This simplifies the numerical computation of a mean-field equilibrium which can be done by solving a fixed point equation for the threshold.

5.2. Centralized Control Strategy

We focus on a centralized control problem where the goal is to minimize the expected cost of the population. The authors in [19] study centralized strategies in this SIR model and show that its solution is the unique viscosity solution of an Hamilton-Jacobi-Bellman equation. We show that the centralized control problem is tractable and characterize its solution.

We denote by $C(\pi)$ the cost incurred in the system by the population vaccination strategy π, i.e.,

$$C(\pi) = \int_0^T \left(c_I m_I(t) + c_V \pi(t) m_S(t)\right) \, dt.$$

The global optimum of the problem is the population strategy that minimizes the total cost and let

$$\pi^{opt} \in \arg\min_{\pi} C(\pi).$$

As for the case of mean-field equilibrium, a global optimum is a threshold strategy.

Proposition 3. *The strategy that minimizes the total cost is a threshold strategy.*

Proof. For a given the population strategy π and $\epsilon > 0$, we define $u_0(t) = \pi(t) m_S(t)$ and

$$u_1(t) = \begin{cases} u_0(t) & \text{if } t < t_0, \\ u_0(t) - \epsilon & \text{if } t \in [t_0, t_0 + \delta), \\ u_0(t) + \epsilon & \text{if } t \in [t_0 + \delta, t_0 + 2\delta), \\ u_0(t) & \text{if } t \in [t_0 + 2\delta, T], \end{cases}$$

satisfying that $\int_0^T u_0(v) dv = \int_0^T u_1(v) dv$. We show that $u_0(t)$ is an improving strategy comparing with $u_1(t)$, which means that the incurred cost by $u_0(t)$ is less that the system cost under $u_1(t)$, that is: if $m_S^0(t)$ and $m_I^0(t)$ are the proportion of susceptible and infected population under strategy $u_0(t)$ and $m_S^1(t)$ and $m_I^1(t)$ be the proportion of susceptible and infected population under strategy $u_1(t)$, then

$$\int_0^T \left(c_I m_I^0(t) + c_V u_0(t)\right) \, dt < \int_0^T \left(c_I m_I^1(t) + c_V u_1(t)\right) \, dt. \tag{19}$$

Since $\int_0^T u_0(v) dv = \int_0^T u_1(v) dv$, this inequality holds if $m_I^0(t) < m_I^1(t)$, for all t. Let $m_I^0(t_0) = m_I^1(t_0) = i_0$ and $m_S^0(t_0) = m_S^1(t_0) = s_0$. We divide the proof in three parts: *(A)* $m_I^1(t_0 + \delta) > m_I^0(t_0 + \delta)$, *(B)* $m_I^1(t_0 + 2\delta) > m_I^0(t_0 + 2\delta)$ and *(C)* $m_I^1(t) > m_I^0(t)$, for $t = t_0 + 2\delta$.

(A) From the definition of $u_1(t)$, it follows that $\dot{m}_I^0(t_0) = \dot{m}_I^1(t_0)$ and $\dot{m}_S^1(t_0) = \dot{m}_S^0(t_0) + \epsilon$. Besides, for the second derivative, we have that $\ddot{m}_I^1(t_0) = \ddot{m}_I^0(t_0) + \gamma\epsilon i_0$ and $\ddot{m}_S^1(t_0) = \ddot{m}_S^0(t_0) - \gamma\epsilon i_0$.

As a result, using the Taylor expansion, we obtain that

$$m_I^1(t_0+\delta) = m_I^0(t_0+\delta) + \frac{\delta^2}{2}\ \gamma\ \epsilon\ i_0 + O(\delta^3),$$

and $m_S^1(t_0+\delta) = m_S^0(t_0+\delta) + \delta\ \epsilon - \frac{\delta^2}{2}\ \gamma\ \epsilon\ i_0 + O(\delta^3)$.

(B) The first derivatives of the proportion of infected and susceptible population at $t_0+\delta$ satisfy that $\dot{m}_I^1(t_0+\delta) = \dot{m}_I^0(t_0+\delta) + \delta\ \gamma\ \epsilon\ i_0 + O(\delta^2)$ and $\dot{m}_S^1(t_0+\delta) = \dot{m}_S^0(t_0+\delta) - \epsilon - \delta\ \gamma\ \epsilon\ i_0 + O(\delta^2)$. For the second derivatives, we have that $\ddot{m}_S^1(t_0+\delta) = \ddot{m}_S^0(t_0+\delta) + \gamma\epsilon i_0 + O(\delta)$ and $\ddot{m}_I^1(t_0+\delta) = \ddot{m}_I^0(t_0+\delta) - \gamma\epsilon i_0 + O(\delta)$.

From these expressions, it results that

$$m_S^1(t_0+2\delta) = m_S^0(t_0+2\delta) - \delta^2\epsilon\gamma i_0 + O(\delta^3), \tag{20}$$

and

$$m_I^1(t_0+2\delta) = m_I^0(t_0+2\delta) + \delta^2\epsilon\gamma i_0 + O(\delta^3). \tag{21}$$

(C) We show this result by induction on t. First, we note that (20) and (21) are satisfied. Then, we assume that at $t' > t_0+2\delta$, $m_I^1(t') = m_I^0(t') + \gamma\epsilon i_0\delta^2 + O(\delta^3)$ and $m_S^1(t') = m_S^0(t') + O(\delta^2)$ are satisfied. To finish the proof, we need to show that these equations hold for $t'+\delta$.

The first derivative of the infected population at t' satisfies that $\dot{m}_I^1(t') = \dot{m}_I^0(t') + O(\delta^2)$ and the second derivative $\ddot{m}_I^1(t') = \ddot{m}_I^0(t') + O(\delta^2)$. For the susceptible population, we obtain that $\dot{m}_S^1(t') = \dot{m}_S^0(t') + O(\delta^2)$ and $\ddot{m}_S^1(t') = \ddot{m}_S^0(t') + O(\delta^2)$.

Finally, using the previous expressions, the Taylor expansion gives

$$m_I^1(t'+\delta) = m_I^0(t'+\delta) + \gamma\epsilon i_0\delta^2 + O(\delta^3),$$

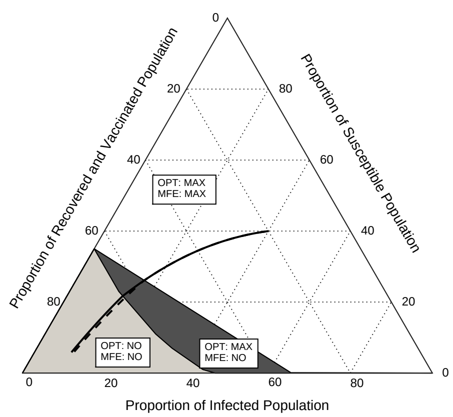
and $m_S^1(t'+\delta) = m_S^0(t'+\delta) + O(\delta^3)$. This finishes the proof.

From this result, we have that if we displace an infinitesimal part of area of $b(t)S(t)$ to the left the system cost decreases. Therefore, proceeding recursively, we conclude that the population strategy of minimal cost is a threshold strategy. $\square$

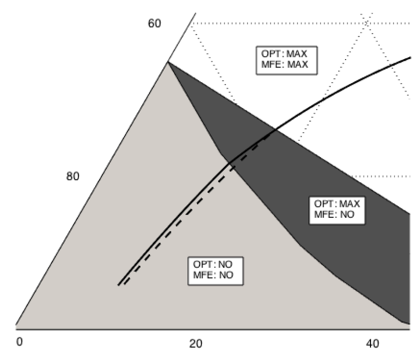
5.3. Numerical comparisons

We know that there exists a mean-field equilibrium and a global optimum that are of threshold type. Hence, we now compare the equilibrium obtained in Proposition 1 with the global optimum of Proposition 3. For this purpose, we consider the following system parameters: $\rho = 36.5$, $\gamma = 73$, $\tau = 10$, $T = 0.3$, $c_I = 36.5$ and $c_V = 0.5$

Remark 2. *The authors in [20] consider the previous system parameters except for the infection cost that is* 1, *instead of* 36.5. *It is easy to see that if the cost of infection is c_I times ρ both models coincide. The authors, using their approximation, obtain that the cost for the mean-field equilibrium and in the global optimum are, respectively,* 0.55 *and* 0.53. *Using our approach, the resulted cost is* 0.542 *for the mean-field equilibrium and* 0.524 *for the global optimum.*



(a) Population dynamics.



(b) Population dynamics (zoomed).

Figure 2: Population dynamics under the equilibrium strategy (dashed line) and the global optimum strategy (solid line). Three zones are displayed: *(i)* in the white region, the global optimum and the equilibrium vaccinate with maximum rate; *(ii)* in the dark gray region, the global optimum vaccinates with maximum rate, while the equilibrium does not vaccinate; and *(iii)* in the light gray region, neither the global optimum nor the equilibrium vaccinates. $I(0) = S(0) = 0.4$.

We computed the optimal strategies and of the mean-field equilibrium. The results are reported in Figure 2a, in a simplex-plot. In this figure, the simplex set is divided in three regions that represent the decision taken by both policies at time 0, as a function of the initial state. In the white region, both strategies vaccinate at maximum rate. In the dark gray region, the strategy of the global optimum is to vaccinate at maximum rate and the strategy of the equilibrium is not to vaccinate. In the light gray region, the strategy of the equilibrium and the global optimum is not to vaccinate.

We also plot the trajectories corresponding to both strategies when the initial proportion of infected population and of susceptible population are both equal to 0.4 at time 0. In Figure 2a (see Figure 2b for a zoomed figure), we plot with solid line the behavior of the equilibrium vaccination strategy and with a dashed line the behavior of the global optimum. We observe when at the beginning, both strategies consist in vaccinating at maximum rate. After some time, the equilibrium strategy is not to vaccinate, whereas the global optimum strategy does not change the strategy. After that, there is another instant where the global optimum strategy do not vaccinate and, therefore, no strategy vaccinates.

In this simulation we observe that the threshold at which the equilibrium changes her strategy is smaller than for the global optimum. We have compared these thresholds over a large number of simulations with different parameters and the obtained results say that in all but degenerated cases, the thresholds do not coincide, so that the price of anarchy of this model is never equal to 1.

5.4. Mechanism Design

In Figure 3, we compare the thresholds of the optimal strategy with the one of the mean-field equilibrium strategy. For a given c_V and fixed the rest of the parameters, we denote by $t^{opt}(c_V)$ (resp. $t^{eq}(c_V)$) the threshold of the global optimum strategy (resp. equilibrium strategy). It can be shown that in both cases, the thresholds are decreasing in c_V: the more costly is the vaccination, the less people vaccinate (for the globally optimal situation or for the mean-field equilibrium).

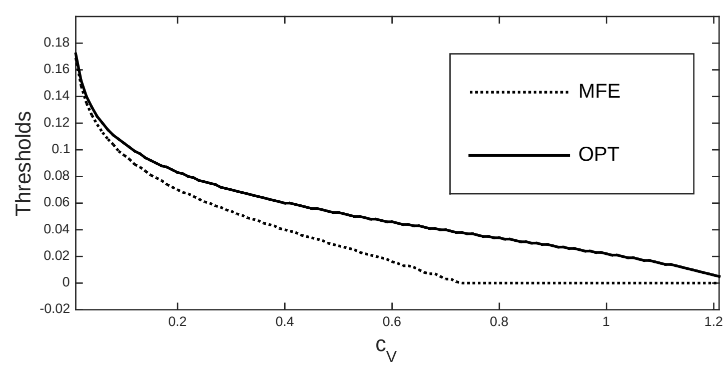


Figure 3: Comparison of the threshold of the equilibrium and of the global optimum when c_V varies from 0.01 to 1.21.

Figure 3 confirms that the threshold decrease with c_V and also shows that the thresholds are never equal for this range of parameters. This figure also suggests that the optimal threshold is always larger than the mean-field equilibrium threshold. This fact was already observed in [20] and suggests that, if the vaccination decisions are let to individual, then vaccination should be subsidized, by removing a cost p to the vaccination cost in the equilibrium so that both thresholds coincide, *i.e.*,

$$t^{eq}(c_V - p) = t^{opt}(c_V).$$

For example, with the same parameters as in the simulation of Figure 2, we observe that, for $c_V = 0.8$, the threshold of the global optimum is 0.034, while the threshold of the equilibrium is 0. As it can be seen, the threshold of the equilibrium is 0.034 when $c_V = 0.45$. This simulation shows that, to encourage selfish individuals to vaccinate optimally, vaccination should be subsidized in order to reduce its cost of $p = 0.35$.

6. Related Work

Sometimes population games as defined in [26] are also called mean field games because they are mean-field limits of *static* games. Here, we only discuss the mean limit of dynamic games, as originally introduced by Lasry and Lions [23] or by Huang, Malhame and Caines [17].

Since the seminal work in [21, 22, 23, 17], a large variety of papers have been investigating mean-field games. Most of the literature perform an analysis of these games based on a coupling of a Hamilton-Jacobi-Bellman equation as well as on the Fokker-Planck equation (see for example [15, 3, 5, 12, 6]). Here, we are interested in studying mean-field games with a finite number of states and finite number of actions. In this case, the analog of the Hamilton-Jacobi-Bellman equation is the Bellman equation (1) and the discrete version of the Fokker-Planck equation is the Kolmogorov equation (14).

In this article, we analyze continuous as well as discrete time mean-field games.

Finite state space mean-field games in discrete time has received less attention. It was previously studied in [10] where the strategy of the players is the rate at which they change between states. In our terms, this corresponds to the case where the mixed action space is the set of all bounded rate matrices and $Q_{i,j,a} = a$. This gives each player the power to decide her dynamics independently of the state of the others.

Continuous time finite state space mean-field games have been previously studied in [11, 14]. In this models, the players also control completely the transition rate matrix. Our model is more general. It considers that the players may not have such a power and their actions only have a limited effect on their state. Here, the transition rates Q may depend not only on the action taken by the player, but also on the population distribution of the system. We claim that this scenario is rather common in systems such as epidemic or belief propagation and the diffusion of information, as detailed in Section 5, but also in other cases such as resource allocation, where a player cannot use a resource already utilized by others.

Since our mean-field game model is a strict generalization of these previous models, the existence of a mean-field equilibrium can be seen as corollaries of our main theorem. Furthermore, another important difference of our work with respect to these continuous time finite state space models concerns the cost functions. In fact, in [14] it is assumed that the cost of a player is strictly convex on her strategy and in [11] the authors consider uniformly convex functions. We note that these models do not cover, for example, linear costs, which are important for applications. In our approach, the only requirement on the cost is continuity with respect to the population distribution (see Assumption (A1)).

Different authors have studied the convergence of N-player games equilibria to mean-field equilibria, *e.g.* [16, 1, 27, 28]. Their model is different from ours since they consider that the strategy of a player only depends on her internal state (called *stationary policies* in [28]). Here we allow time dependence to these policies. The model in [28] does include state dynamics that depending on the population distribution but only considers stationary strategies that do not depend on time, hence cannot depend on the population dynamics.

Finally, our four models of dynamic games with a finite number of players do not face the issue of the order of play. Thus, we avoid two difficulties of dynamics games: the information structure of each player and the existence of a value [7]. 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 their current state). Then, the continuity assumptions of the cost function and of the rate matrices are enough to ensure the existence of an equilibria.

7. Conclusions

In this article, we introduce mean field games with 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 grows. 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 [11] 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 large. We aim to characterize the largest sub-class of strategies where convergence holds. Obviously, this class includes all local strategies and excludes some Markovian ones.

References

[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] M. Benaim and J.-Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11):823–838, 2008.

[3] A. Bensoussan, J. Frehse, and P. Yam. *Mean field games and mean field type control theory*. Springer, 2013.

[4] K. C. Border. *Fixed point theorems with applications to economics and game theory*. Cambridge university press, 1989.

[5] 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.

[6] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. *SIAM Journal on Control and Optimization*, 51(4):2705–2734, 2013.

[7] P. Dasgupta and E. Maskin. The existence of equilibrium in discontinuous economic games, i: Theory. *Review of Economic Studies*, 53(1):1–26, 1986.

[8] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, 1986.

[9] N. Gast and B. Gaujal. A mean field approach for optimization in discrete time. *Discrete Event Dynamic Systems*, 21(1):63–101, 2011.

[10] 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.

[11] D. A. Gomes, J. Mohr, and R. R. Souza. Continuous time finite state mean field games. *Applied Mathematics & Optimization*, 68(1):99–143, 2013.

[12] 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.

[13] A. Granas and J. Dugundji. *Fixed point theory.* Springer Science & Business Media, 2013.

[14] O. Guéant. Existence and uniqueness result for mean field games with congestion effect on graphs. *Applied Mathematics & Optimization*, 72(2):291–303, 2014.

[15] 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.

[16] M. Huang. Mean field stochastic games with discrete states and mixed players. In *Game Theory for Networks*, pages 138–151. Springer, 2012.

[17] 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.

[18] T. G. Kurtz. *Approximation of population processes*, volume 36. SIAM, 1981.

[19] L. Laguzet and G. Turinici. Global optimal vaccination in the sir model: properties of the value function and application to cost-effectiveness analysis. *Mathematical biosciences*, 263:180–197, 2015.

[20] L. Laguzet and G. Turinici. Individual vaccination as nash equilibrium in a sir model with application to the 2009–2010 influenza a (h1n1) epidemic in france. *Bulletin of Mathematical Biology*, 77(10):1955–1984, 2015.

[21] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. i–le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006.

[22] 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.

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

[24] J. Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[25] M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons, 2014.

[26] W. Sandholm. *Population Games and Evolutinary Dynamics.* MIT Press, 2010.

[27] H. Tembine. Mean field stochastic games: convergence, q/h-learning and optimality. In *American Control Conference (ACC), 2011*, pages 2423–2428. IEEE, 2011.

[28] 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.

[29] J. von Neuman. Zur theorie der gesellschaftspiel. *Mathematische Annalen*, 100:295–320, 1928.