Bayesian Learning in Mean Field Games

Eran Shmaya, Bruno Ziliotto

▶ **To cite this version:**

Eran Shmaya, Bruno Ziliotto. Bayesian Learning in Mean Field Games. 2024. hal-04568449

HAL Id: hal-04568449
https://hal.science/hal-04568449v1

Preprint submitted on 15 Nov 2024

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.

arXiv:2401.17696v1 [math.OC] 31 Jan 2024

Bayesian Learning in Mean Field Games

Eran Shmaya * Bruno Ziliotto†

February 1, 2024

Abstract

We consider a mean-field game model where the cost functions depend on a fixed parameter, called *state*, which is unknown to players. Players learn about the state from a a stream of private signals they receive throughout the game. We derive a mean field system satisfied by the equilibrium payoff of the game and prove existence of a solution under standard regularity assumptions. Additionally, we establish the uniqueness of the solution when the cost function satisfies the monotonicity assumption of Lasry and Lions at each state.

Introduction

Mean field games (MFGs, see [19, 20, 21] and [16]) feature a continuum of non-atomic players who control their own positions, and optimize a cost function that depends on their position, their speed, and the distribution positions of the other players. These games have been intensively studied in the last two decades (see e.g. [10, 8]), and found a wide range of applications. In most of these works, players are assumed to know perfectly the variables determining the cost function. This is in contrast with Game Theory literature, where models with incomplete information have been intensively studied in a wide variety of frameworks (see e.g. [2], [24, Chapter III], [23, Chapter 10] and [13]). Incomplete information is typically on some variable affecting the cost function and the dynamics. Such a variable may be static or dynamic, and players may acquire information on it through the observation of a stream of private or public signals, or through the observation of the other's actions. As far as incomplete information is concerned, MFG literature is rather scarce. Models where players are uninformed of their own state have been studied in [27, 28, 12, 3], where players do not know the average position of the other players [11, 6], and where one major player has a private information and discloses it strategically through her control to a population of small players [4].

*Stony Brook University, United States
†CEREMADE, CNRS, PSL Research Institute, Paris Dauphine University, France.

Contribution of the paper In this paper, we introduce a MFG model where cost functions depend on a fixed parameter called *state*. Players do not know the exact value of the state, and receive a stream of private signals along the game. A major issue is the procedure by which the players learn the state of nature from their signals. We adopt the Bayesian inference approach, which is the standard approach in game theory: the players update their beliefs about the state using Bayes' formula. In a contemporary paper, [4] also study a MFG model with an unknown state. A major difference between their model and ours is that in their model all players observe the same signal, whereas in our model players observe idiosyncratic signals, conditionally independent on the states. The model in [4] reduces to MFG with common noise. In contrast, our model does not reduce to MFG with common noise and a different system of equations is needed to describe the equilibrium.

Incorporating Bayesian inference in mean field games introduces a challenge because the posterior beliefs about the parameter live in an infinite dimensional space. The work [4], following [2], assumes a finite state space, so beliefs are in a finite dimensional simplex. Our approach is different. The tractability of our model relies on the assumption that the state and the signals are jointly normally distributed, so all posteriors beliefs are also normal, and are determined by the conditional expectation of the states. Moreover, our signaling process is such that future information depends on the state alone and not on current information. This assumption on the signal process makes the posterior belief a Markov process and simplifies Bayesian updating.

We derive a PDE system that describes the equilibrium solution of the game, composed with a Hamilton-Jacobi equation and a family of Fokker-Planck equations, indexed by the state. The space variable of each player in this system is a pair (z, x), where x represents the position and z is a variable that aggregates the information of the player about the state. Such a system is not a classic MFG system. Indeed, the state appears explicitly as a parameter in the Fokker-Planck equation, while in the Hamilton-Jacobi equation, it appears only through the belief of players. This reflects the fact that the signal dynamics is driven by the state, while players controls are driven by the information variable z. Hence, one can not apply directly existence and uniqueness result to this system. Under standard regularity assumptions, we prove existence of a solution, using the classic method of Shauder fixed-point theorem.

We then turn to the question of the uniqueness of the solution of our system. A well-known condition under which a standard MFG system has a unique solution is the monotonicity condition of Lions and Lasry [21], which can be intuitively interpreted by the fact that players have a preference for being far away from each other. Assuming that the monotonicity condition is satisfied "state-by state", we prove uniqueness of our system. This is rather surprising, since the state is not observed by players; one may have expected instead a monotonicity condition expressed in terms of the distribution of the state given players information, as it is the case for instance in [6], in a different setting.

Related game-theoretic models The mean-field game literature parallels the game theoretic literature on large games with anonymous players. Here anonymity means that each players' payoff depends on opponents' actions and types only through their empirical distribution, but not on the identity of the opponent that played each action. The literature starts with static games that model one-shot interaction. A canonical model in this literature, the so called *continuum of players* or *non-atomic* approach [26, 22], assumes that idiosyncratic randomizations corresponding to different players actions completely wash out in the aggregate, so that the outcome of the game is fully deterministic. The non-atomic model can also allow some idiosyncratic characteristics of the players, or types in the game-theoretic terminology, again under the assumption that there is no aggregate uncertainty about the distribution of types. There are many results that relate the asymptotic outcomes of anonymous games with N players, where N goes to infinity with the outcome of the corresponding non-atomic model (see e.g. [9]).

A seminal paper of Green [14] combined large games with (discrete-time) dynamic games. Green's paper and most of its follow-up study games with perfect information, in which the fundamentals of the game are known to the players. Green studied repeated games, in which the players play the same game at every period. Jovanovic [17], Bergin and Bernhadt [5], and Hoppenhayn [15] studied large games in which the payoff depends on the history of the game and stochastic games with some aggregated uncertainty.

There is a huge literature on Bayesian learning in game theory (See Young [29, Chapter 7]). Most of these papers are about games with finite number of players. Kalai and Shmaya [18] seem to be the first paper that studies dynamic non-atomic games in which the fundamental of the game are not known initially. Their model also includes Bayesian learning of the fundamentals. In addition to the fact that their model is in discrete time, a major difference between the model in this paper and their model is that the signals in Kalai and Shmaya's paper are commonly observed by all players so at every period all players hold the same belief about the state of nature, whereas in our paper beliefs are idiosyncratic.

Paper outline Section 1 describes the game model, while Section 2 establishes the corresponding MFG system and relates it to the state of the art. Section 3 proves existence of a classical solution of the MFG system, under standard regularity assumptions. Section 4 proves uniqueness under a monotonicity assumption.

Notations Throughout the paper, the notation $\mathcal{N}(m, v)$ stands for the normal distribution with mean m and variance v. The n-dimensional torus, defined as the quotient of $\mathbb{R}^n$ by $\mathbb{Z}^n$, is denoted by $\mathbb{T}^n$. All finite-dimensional spaces under consideration will be equipped with the Borelian σ-algebra. Given a measurable set A, the notation $\mathscr{P}(A)$ stands for the probability distributions over A.

1 Model

1.1 A control problem

The MFG system considered in this paper is interpreted as a game with a continuum of players, where each player faces the control problem described by the following elements:

- A random *state* S in $\mathbb{R}$, unknown to the player, with distribution $\mathscr{N}(0,1)$.
- A signalling process (Z_t) for the player, that is a process on $\mathbb{R}$ that follows the stochastic differential equation (SDE)
$$dZ_t = Sdt + \sigma dB_t, \tag{1.1}$$
with $\sigma > 0$.
- A position X_t lying in the n-dimensional torus $\mathbb{T}^n$, whose trajectory is controlled by the player, according to the equation
$$dX_t = \alpha_t dt + \sigma' dB'_t, \tag{1.2}$$
where (B_t) and (B'_t) are independent Brownians, and the control variable $\alpha = (\alpha_t)_{t\geq 0}$ is measurable with respect to the filtration generated by (X_t, Z_t). When we want to emphasize the dependence of the player's signal and position process on α we also use the notation (X^α_t, Z^α_t). The position space is taken as the torus for simplification purpose.
- A *position distribution* $\rho_{s,t} \in \mathscr{P}(\mathbb{T}^n)$, with the interpretation that $\rho_{s,t}$ is the distribution density over positions at time t, conditional on the state being s,
- A *total cost* that is a function of the control $(\alpha_t)_{t\geq 0}$ used by the player, defined by
$$\mathbb{E}\left(\int_0^T C(S, X_t, \alpha_t) + F(S, X_t, \rho_{S,t})\ \mathrm{d}t\right) + G(S, X_T, \rho_{S,T}), \tag{1.3}$$
where the *cost functions* C,F, and G satisfy some regularity assumption that will be specified in the next section, and the expectation is taken over the random state S and the random path X_t of the player. The additive structure of the flow cost, with one component that does not depend on the population distribution, and a second component that does not depend on the control, is standard. Compared to the classic MFG model, the new feature is that the cost functions also depend on the unknown state.

  We assume that the flow cost is not observed by the player, so that all information about the state comes from the signaling process. Section 1.3 provides examples where such a situation arises. The fact that players do not observe their flow cost is a common feature of

several classic incomplete information models, such as Partially Observable Markov Decision Processes and stochastic games with signals [24].

1.2 Equilibrium

A *game distribution* is a mapping $m = m : \mathbb{R} \times [0,T] \times \mathbb{T}^n \times \mathbb{R} \to \mathbb{R}$ such that $m_{s,t}$ is a probability density over $\mathbb{T}^n \times \mathbb{R}$ for every s,t. The game distribution will have two related interpretations:

- $m_{s,t}$ is the empirical distribution over positions and signals at time t if the state is s.
- $m_{s,t}$ is the law of distribution of the (random) position and signal of a player at time t if the state is s.

The distribution $m_{s,t}$ is an *equilibrium* if there exists a control α such that:

- $m_{s,t} = \mathscr{L}(X^{\alpha}_{s,t}, Z^{\alpha}_{s,t})$ is the probability density function of $X_{s,t}, Z_{s,t}$.
- α minimizes the single player cost (1.3) where $\rho_{s,t}$ is the marginal of $m_{s,t}$ over positions.

The first item captures the assumption that, while each player's trajectory is stochastic due to the randomness of his own Brownian, the population's trajectory is deterministic conditional on the state. In game theoretic terminology, this assumption is called *no aggregated uncertainty*. Note, however, that in our environment, the population dynamic is still random because it depends on the random state. The second item captures the idea that the players take the population dynamic as given and ignore their own impact on it. In game theoretic terminology, players are *outcome takers* [14].

Both of this assumptions make sense when the number of players is large. Indeed, the MFG equilibrium, sometimes called *a non-atomic equilibrium* in game theoretic literature [26, 22], approximates a game with many players as follows: Consider a game with N players, each player observes a signals and controls his positions according to (1.1) and (1.2) where the Brownian components are independent between players. The flow cost to the players depends on the empirical distribution of the players' positions $\frac{1}{N}\sum_i \delta_{x^i_t}$ where x^i_t is the position of player i at time t. The MFG equilibrium strategies given by the control α are ϵ-Nash equilibrium in the finite player game for sufficiently large N.

As we mentioned, most of the game-theoretic literature focuses on discrete-time games. The continuous-time modeling is also an approximation of discrete-time players when players take action and receive signals in short intervals, or *frequently* in game theoretic terminology [25]. More explicitly, in the *discrete-time game with round duration* Δ, at the beginning of each period each player receives a signal from distribution $\mathscr{N}(s, \sigma^2\Delta)$, where s is the realization of the state, and signals are i.i.d. over players and rounds. At the beginning of every round, each player is at some position x_k, chooses a control α_k and his position moves to $x_{k+1} = x_k + \alpha\Delta + \mathscr{N}(0, \sigma'^2\Delta)$. At the

end of the round, the player receives payoff $\Delta\left(C(S, x_k, \alpha_k) + F(S, x_k, \rho_k)\right)$ where ρ_k is the empirical distribution of positions of all players at round k. The game is played for T/Δ rounds. The MFG equilibrium becomes a standard ϵ-Nash equilibrium in the discrete game for sufficiently small Δ.

1.3 Applications

Our model is a basis for future theoretical developments and applications, including the situations that we mention now.

- *Product differentiation.* consider firms that sell the same type of product (e.g. cellulars). They have a time period T to develop their product. Their position represents the characteristics of the product (size, quality, price...). The state S represents the ideal characteristics of the consumer, and only affects the terminal cost. At time T, the product is put on the market. Their aim is that the product characteristics are not far from the ideal of the consumer, and moreover, that their product is far from the one of the other firms. The terminal cost reflects these preferences, and can be taken for instance as $G(s, x, \rho) = |x - s|^2 - \int |x - y|^2 \rho(\ \mathrm{d}y)$. The flow cost is only a function of the position and the control.

- *Portfolio management.* Consider investors that put funds in some investment scheme during some time period, (life insurance, hedge fund), such that the return interest rate is only known at the end of the period. The state S represents the interest rate, and the position of each player represents the composition of their portfolio.

2 Mean field game equations

2.1 The HJB equation

Fix the population density $\rho_{s,t}$ over $\mathbb{T}^n$ at time t if the state is s. In this section we derive the solution to the single-player control problem from Section 1.1.

The tractability of our model relies on the fact that the aggregated signal Z_t at time t is a sufficient statistic for S given the signals Z_τ at times $\tau \in [0, t]$. This sufficiency, combined with the fact that the evolution of position X_t in (1.2) depends on the signal but not directly on the state, imply that the signal Z_t at time t is also a sufficient statistic for S given the process (Z_τ, X_τ) at times $\tau \in [0, t]$. This allows us to reduce the player's problem to a standard control problem with the state variable (z, x).

We first describe the player's evolution of beliefs about the state, which is a standard argument in Bayesian statistics. We denote by $\mathscr{N}(\mu, \sigma^2)$ the normal law of mean μ and variance σ^2, and by φ_{μ,σ^2} its probability density function.

Lemma 2.1. *For all $t \in \mathbb{R}_+$,*

(i) The law of $Z_t|S = s$ is $\mathcal{N}(st, \sigma^2 t)$

(ii) The law of Z_t is $\mathcal{N}(0, \sigma^2 t + t^2)$.

(iii) The law of $S|Z_t = z$ is $\mathcal{N}(r_t(z), {\sigma_t}^2)$, where

$$r_t(z) = \frac{z}{t+\sigma^2} \quad \text{and} \quad \sigma_t^2 = \frac{\sigma^2}{\sigma^2 + t} \tag{2.1}$$

Proof. The first item follows from (1.1). The second and third items follow from the fact that

$$\varphi_{0,1}(s)\varphi_{st,\sigma^2 t}(z) = \varphi_{0,\sigma^2 t+t^2}(z)\varphi_{r_t(z),\sigma_t^2}(s). \tag{2.2}$$

□

Recall that our cost functions F, C, G depend on the state that is not observed by the players. Let $\tilde{F}(t,z,x,\rho), \tilde{C}(t,z,x,\alpha), \tilde{G}(z,x,\rho)$ be the player's corresponding expected cost given the signal z at time t. More explicitly

$$\begin{aligned}
\tilde{F}(t,z,x,\rho) &= \int \varphi_{r_t(z),\sigma_t^2}(s) F(s,x,\rho_{s,t})\ \mathrm{d}s, \\
\tilde{C}(t,z,x,\alpha) &= \int \varphi_{r_t(z),\sigma_t^2}(s) C(s,x,\alpha)\ \mathrm{d}s, \text{ and} \\
\tilde{G}(z,x,\rho) &= \int \varphi_{r_T(z),\sigma_T^2}(s) G(s,x,\rho_{s,T})\ \mathrm{d}s
\end{aligned} \tag{2.3}$$

By replacing the integrand in (1.3) with its conditional expectation at time τ and then using the sufficiency of the signal process we can rewrite the player's cost function as

$$\mathbb{E}\int_0^T \left(\tilde{C}(t, Z_t, X_t, \alpha_t) + \tilde{F}(t, Z_t, X_t, \rho_{S,t})\ \mathrm{d}t\right) + \tilde{G}(t, Z_t, X_T, \rho_{S,T}), \tag{2.4}$$

In addition, it follows from Lemma 2.1 that the signal process of a single player evolves according the SDE.

$$\mathrm{d}Z_t = r_t(Z_t)\mathrm{d}t + \sigma^2 \mathrm{d}B_t. \tag{2.5}$$

Therefore, the control problem given by (2.4) is a standard control problem, with state z, x and value function

$$u(t,z,x) = \inf_{\alpha_t} \mathbb{E}\int_t^T \left(\tilde{F}(r, Z_r, X_r, \rho_{S,t}) + \tilde{C}(r, Z_r, X_r, \alpha_t)\right)\ \mathrm{d}r + \tilde{G}(T, Z_T, X_T, \rho_{S,T}) \tag{2.6}$$

with $X_t = x$ and $Z_t = z$ and evolution given by (1.2) and (2.5).

For $(t,z,x) \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^m$ and $p \in \mathbb{R}^n$, let

$$H(t,z,x,p) := \sup_{\alpha} \left\{ -\alpha \cdot p - \tilde{C}(t,z,x,\alpha) \right\} \tag{2.7}$$

The Hamilton-Jacobi-Bellman equation for u is

$$-D_t u + H(t,z,x,D_x u) - D_z u \cdot r_t(z) - \frac{\sigma^2}{2}\Delta_z u - \frac{\sigma'^2}{2}\Delta_x u = \tilde{F}(t,z,x,m), \tag{2.8}$$

with the terminal condition

$$u(T,z,x) = \tilde{G}(z,x). \tag{2.9}$$

Standard verification results (for example [7, Lemma 3.1.6]) imply that if $u(t,x,z)$ is a classical solution to (2.8) then u is the value function of (2.6) and the optimal control at (t,z,x) is $-D_p H(t,z,x,D_x u)$.

2.2 The Fokker-Planck Equation

Assume each agent chooses the control $\alpha(t,z,x)$, which is continuous in t and Holder continuous in z,x. Then for every $s \in \mathbb{R}$ the SDE

$$\mathrm{d}Z_t = s dt + \sigma dB_t \mathrm{d}X_t = \alpha_t dt + \sigma' dB'_t. \tag{2.10}$$

has a unique solution and the density $m(s,t,z,x)$ of (Z_t, X_t) at time t is a weak solution for the Fokker-Planck equation (See [7, Lemma 3.1.3])

$$\partial_t m + \mathrm{div}_x(\alpha(t,z,x)m) + s \cdot \mathrm{div}_z\, m - \frac{\sigma^2}{2}\Delta_z m - \frac{\sigma'^2}{2}\Delta_x m = 0 \tag{2.11}$$

with the initial condition

$$m(s,0,z,x) = \bar{\rho}(x)\delta_0(z). \tag{2.12}$$

Here $\bar{\rho} \in \Delta(\mathbb{T}^n)$ is the initial density of positions, and we assume that the signal process starts with 0 for all players (See [7, Section 3.1.2])

Recall that the signal Z_t at time t is a sufficient statistic for S given the process (Z_τ, X_τ) at times $\tau \in [0,t]$. The following lemma establishes the implication of sufficiency on the solution to the Fokker-Planck equations (2.11): The solution factorizes to two terms, similar to the classical Fisher–Neyman factorization theorem. In the lemma, τ does not depend on s. Note that Z_t/t is an unbiased estimator for S, which explains the drift coefficient that multiplies $div_z \tau$ in the PDE of τ. We will use the lemma in our proofs of existence and uniqueness of solutions to the MFG system.

Lemma 2.2. *The solution to (2.11) is of the form* $m(s,t,z,x) = \varphi_{st,\sigma^2 t}(z)\tau(t,z,x)$ *where* $\tau(t,z,x)$ *is the solution to*

$$\partial_t \tau + \operatorname{div}_x(\alpha(t,z,x)\tau) + \frac{z}{t}\operatorname{div}_z \tau - \frac{\sigma^2}{2}\Delta_z \tau - \frac{\sigma'^2}{2}\Delta_x \tau = 0,$$

with the initial condition

$$\tau(0,z,x) = \bar{\rho}(x).$$

Proof. Note first that $\varphi_{st,\sigma^2 t}$ is the solution to

$$\partial_t \varphi + s\operatorname{div}_z \varphi - \frac{\sigma^2}{2}\Delta_z \varphi = 0,$$

with the initial condition

$$\varphi(0,z) = \delta_0(z)\ .$$

Assume φ and τ satisfy these equations and let $m = \varphi \cdot \tau$. Then

$$\operatorname{div}_x \alpha m = \varphi \operatorname{div}_x \alpha\tau, \text{ and } \operatorname{div}_z m = \tau \operatorname{div}_z \varphi + \varphi \operatorname{div}_z \tau$$

and

$$\Delta_x m = \varphi\Delta_x\tau \text{ and } \Delta_z m = \tau\Delta_z\varphi + 2\operatorname{div}_z \varphi \operatorname{div}_z \tau + \varphi\Delta_z\tau$$

Therefore

$$\begin{aligned}\partial_t m =& \tau\partial_t\varphi + \varphi\partial_t\tau = \\ & \tau(-s\operatorname{div}_z \varphi + \frac{\sigma^2}{2}\Delta_z\varphi) + \varphi(-\operatorname{div}_x(\alpha\tau) - \frac{z}{t}\operatorname{div}_z \tau + \frac{\sigma^2}{2}\Delta_z\tau + \frac{\sigma'^2}{2}\Delta_x\tau) = \\ & -\operatorname{div}_x(\alpha m) - s\operatorname{div}_z(m) + \frac{\sigma^2}{2}\Delta_z m + \frac{\sigma'^2}{2}\Delta_x m.\end{aligned}$$

where the last equality uses the fact that

$$\sigma^2 \operatorname{div}_z \varphi = (s - \frac{z}{t})\varphi.$$

□

2.3 The MFG system

Definition 2.3. An *equilibrium* is a pair (u, m) that satisfies the following system in the classical sense:

$$\begin{cases} -D_t u + H(t,z,x,D_x u) - D_z u \cdot r_t(z) - \frac{\sigma^2}{2}\Delta_z u - \Delta_x u = \tilde{F}(t,z,x,m) \\ \partial_t m + \mathrm{div}_x(-D_p H(t,z,x,D_x u)m) + s \cdot \mathrm{div}_z\, m - \frac{\sigma^2}{2}\Delta_z m - \frac{\sigma'^2}{2}\Delta_x m = 0 \\ u(T,z,x) = \tilde{G}(z,x) \\ m(s,0,z,x) = \bar{\rho}(x)\delta_0(z) \end{cases}$$

Remark 2.4. The above equations do not constitute a standard MFG system. Indeed, the parameter s appearing in the Fokker-Planck equation is not known to the players. An interesting question is whether such a system can be rewritten with a new set of variables, that are known to the players. This type of transformation is classical in discrete-time zero-sum games,and the new state variables incorporate the *beliefs* of players about the unknown parameters of the game (see [1] for the 1-Player case and [24, Chapter IV] for the 2-Player case). In the MFG context, this has been exploited in [6, 4]. In our setting, such variables could be the triple (z, x, μ), where μ is a probability measure representing the belief of a player over the set of population distributions. This would yield to a MFG system written on an infinite dimensional state space.

2.4 Comparison to previous models

We now compare our model to the two standard models of mean-field games, first without common noise and then with common noise.

Recall our players' signaling structure (1.1). For two distinct players i, j their signaling processes $\mathrm{d}Z_t^i$ and $\mathrm{d}Z_t^j$ are stochastically dependent since their drift is governed by the same random variable S. Under fixed control functions $\alpha^i(t,z,x)$ and $\alpha^j(t,z,x)$ of the players, their positions X_t^i and X_t^j are therefore also stochastically dependent.

The canonical model of mean-field games without common noise can be viewed as our model with a fixed state. In this model, unlike our model, for a fixed control functions, the positions of two players are stochastically independent because they are governed by independent noises.

In stochastic games with common noise the positions of the players is governed by idiosyncratic brownian motion and also by some common brownian noise: If we denote by Y_t^i and Y_t^j the positions of two players under fixed controls $\alpha^i(t,y)$ and $\alpha^j(t,y)$ then

$$\mathrm{d}Y_t^i = \alpha^i(t, Y_t^i) + \mathrm{d}B_t^i + \mathrm{d}B_0^i$$

where B_t^i are the idiosyncratic noises and B_t^0 is the common noise. Because of the common noise, the players' positions are also stochastically dependent, as in our model. However, the nature of the dependence is different: For a fixed $\bar{t}$, the future process $(Y_t^i)_{t\geq\bar{t}}$ of player i is conditionally

independent of the history $(Y_t^j)_{t\leq \bar{t}}$ of player j given the history $(Y_t^i)_{t\leq \bar{t}}$ of player i. In other words, player i, who knows his history up to time $\bar{t}$, would not learn anything new about his future from observing the history of his opponent. This property does not hold in our model, since player i could gain more information about the state S by observing the trajectory of player j.

The difference between our model and these two canonical model can also be seen from the structure of the mean field equations. In the canonical model without common noise there is a single Fokker-Planck equation. We have a paramatrized set of Fokker-Planck equation and their joint solution has to be compatible with the HJB equation through the expected cost function $\tilde{F}$ in (2.3). Note that for a fixed s, our m_s is not a solution to the corresponding HJB equation with cost F, hence our system does not reduce to standard MFG equations. The forward equation of the canonical model with common noise is a stochastic equation governed by the common noise component, while we have a parametrized family of standard Fokker Plank partial differential equations.

In motivation and modeling approaches, our paper is probably most similar to Bertucci's model [6] of mean field games with incomplete information. In his model, the initial distribution of players positions is random, unknown to the players, and plays a similar role to our unknown state S. In particular, his HJB equation also involves an expected cost function under players' belief at that time. In addition to the difference in modeling the information flow, the results are different: In our model uniqueness is obtained under standard monotonicity assumptions on the primitives, while Bertucci shows that this does not hold in his model.

3 Existence

In this section, we provide sufficient conditions under which the MFG system has a classical solution, that are largely inspired by [7].

1. The functions F and G are continuous on $\mathbb{R} \times \mathbb{T}^d \times \Delta(\mathbb{T}^d)$.

2. The functions $F(s,\cdot,m)$ and $G(s,\cdot,m)$ are in $C^{1+\beta}(\mathbb{T}^d)$ and $C^{2+\beta}(\mathbb{T}^d)$ for some $\beta \in (0,1)$, with norm that is bounded by a sub-exponential function of s, uniformly w.r.t. m.

3. The function $F(s,\cdot,\pi)$ and $G(s,\cdot,\pi)$ are bounded in $C^{1+\beta}(\mathbb{T}^d)$ and $C^{2+\beta}(\mathbb{T}^d)$ (for some $\beta \in (0,1)$) uniformly in s and π.

4. The cost function C is continuously differentiable in x, α and $C(s,x,\alpha)$ has exponential growth in s, and $\alpha \to C(s,x,\alpha)$ is strongly convex, uniformly in s and x.

5. The cost function satisfies the following growth condition: there exists $A > 0$, for all (s,x,α),

$$D_x C(s,z,x,\alpha) \cdot p \geq -A(1+|\alpha|^2)$$

Theorem 3.1. *Under Assumptions 1-4, the MFG system has a classical solution.*

Before we prove Theorem 3.1, we explain how the assumptions on the cost function translate to properties of the Hamiltonian. The function $\tilde{C}$ enjoys the same properties of C. Consequently, the maximum $\alpha^*(t,z,x,p)$ is attained in (2.7) and the Hamiltonian is differentiable w.r.t. p and x with

$$\begin{aligned} D_pH &= -\alpha^*(t,z,x,p) \\ D_{z,x}H &= -D_{z,x}\tilde{C}(t,z,x,\alpha^*(t,z,x,p)) \end{aligned} \tag{3.1}$$

In addition, D_pH is globally Lipschitz, where the Lipschitz constant is uniform in (t,z,x) and depends only on the strong convexity constant of C [30].

The global Lipschitz property property on D_pH and (3.1) implies that $\alpha^*(t,z,x,p)$ has a linear growth in p. Together with the growth condition on $\tilde{C}$ and (3.1) again it implies a growth condition on H: there exists $B>0$, for all (s,x,α),

$$D_{z,x}H(t,z,x,\alpha)\cdot p \geq -B(1+|p|^2).$$

Proof. Consider the space $C^0(\mathbb{R}\times[0,T],\mathscr{P}(\mathbb{T}^d))$ of all possible position distribution μ, with the interpretation that $\mu(s,t)$ is the empirical distribution of players' positions at time t if state is s. The assumptions on F and G imply that $\tilde{F}(t,z,x,\mu)$ and $\tilde{G}(z,x,\mu)$ are bounded in $C^{1+\beta}$ and $C^{2+\beta}$ uniformly w.r.t. μ,t and that the maps $\mu\mapsto((t,z,x)\to\tilde{F}(t,z,x,\mu))$ and $\mu\mapsto((z,x)\to\tilde{G}(z,x,\mu))$ from $C^0(\mathbb{R}\times[0,T],\mathscr{P}(\mathbb{T}^d))$ to $C^0([0,T]\times\mathbb{T}^d\times\mathbb{R})$ are continuous.

For a large $C>0$ and, let $\mathscr{C}$ be the set of elements μ in $C^0(\mathbb{R}\times[0,T],\mathscr{P}(\mathbb{T}^d))$ such that

$$d_1(\mu(s,t),\mu(s,t'))\leq C|t-t'|^{1/2}, \text{ and} \tag{3.2}$$

$$d_1(\mu(s,t),\mu(s',t))\leq C|s-s'|^{1/2}. \tag{3.3}$$

Then $\mathscr{C}$ is a convex closed subset of $C^0(\mathbb{R}\times[0,T],\mathbb{T}^d)$, which we equip with the topology of compact convergence. It is compact, thanks to Arzela-Ascoli theorem.

Let $\mu\in\mathscr{C}$. We associate $\nu=\Psi(\mu)$ in the following way: Let u be the unique solution to

$$\begin{cases} -D_tu+H(t,z,x,D_xu)-D_zu\cdot r_t(z)-\Delta_xu-\Delta_zu=\tilde{F}(t,z,x,\mu) \\ u(z,x,T)=\tilde{G}(t,z,x,\mu(T)) \end{cases} \tag{3.4}$$

The existence of a solution follows from [7, Corollary A.2.2], the growth condition on H and the regularity of $\tilde{F}(t,z,x,\mu)$ and $\tilde{G}(z,x,\mu)$. Then we define $m: S\times[0,T]\times\mathbb{R}^n\times\mathbb{R}$ such as for any

$s \in \mathbb{R}$, $m(s,.)$ is the solution of the Fokker-Planck equation

$$\begin{cases} \partial_t m + \mathrm{div}_x(-D_p H(t,z,x,D_x u)m) + s \cdot \mathrm{div}_z\, m - \frac{\sigma^2}{2}\Delta_z m - \frac{\sigma'^2}{2}\Delta_x m = 0 \\ m(.,0) = m_0 \end{cases} \tag{3.5}$$

Last, we define ν such that: for any s,t,x, $\nu(s,t,x)$ is the marginal of $m(s,t,x,.)$. We now claim that $\nu \in \mathscr{C}$. The global Lipschitz assumption on H and the classical $\frac{1}{2}$-Holder estimate on the Fokker Plank equation [7, Lemma 3.1.4] imply (3.2). Lemma 2.2 implies (3.3).

Finally, the maps Ψ is continuous. Indeed, let $\mu_n \in \mathscr{C}$ converge to μ_∞ and let (u_n, m_n) be the corresponding solutions. Then the maps $(t,z,x) \mapsto F(t,z,x,\mu_n)$ and $(t,z,x) \mapsto G(t,z,x,\mu_n)$ converge uniformly to $(t,z,x) \mapsto F(t,z,x,\mu_\infty)$ and $(t,z,x) \mapsto G(t,z,x,\mu_\infty)$ thanks to the continuity properties of $\tilde{F}$ and $\tilde{G}$. Moreover, as the RHS of the HJ for u_n is bounded in $C^{1+\beta,1+\beta/2}$, then the u_n are uniformly bounded in $C^{2+\beta,1+\beta/2}$, hence converges in $C^{2,1}$ to the unique solution u_∞ of the HJB with RHS $F(\cdot,\mu_\infty)$. Then the m_n are solutions of a linear equation with uniformly Holder continuous coefficient, hence are uniformly bounded in $C^{2+\beta,1+\beta/2}$. Then the m_n converge in $C^{1,2}$ to the unique solution of the FP equation corresponding to u_∞. This implies converges of ν_n to ν_∞ in $\mathscr{C}$.

By Schauder fixed point Theorem the map Ψ admits a fxed point μ. The corresponding u, m are an equilibrium.

Together with the assumptions on H, this implies that the Hamilton-Jacobi equation (3.4) has a classical smooth solution. □

4 Uniqueness

We made the following assumptions:

1. The function $\alpha \to C(.,\alpha)$ has a Lipschitz gradient, where the Lipschitz constant is uniform in (s,x). This implies that H is strongly convex in p, uniformly in (t,z,x).

2. We assume that F satisfies the monotonicity condition [21] for every s: the cost functions F and G satisfy

$$\int \left(\rho^1(x) - \rho^2(x)\right)\left(F(s,x,\rho^1) - F(s,x,\rho^2)\right)\, \mathrm{d}x \geq 0, \text{and}$$
$$\int \left(\rho^1(x) - \rho^2(x)\right)\left(G(s,x,\rho^1) - G(s,x,\rho^2)\right)\, \mathrm{d}x \geq 0.$$

   for every $s \in \mathbb{R}^m$ and every $\rho^1 \neq \rho^2 \in \mathscr{P}(\mathbb{R}^n)$.

Theorem 4.1. *Under this assumption, if (u^1, m_1) and (u^2, m_2) are two solutions. Then $u^1 = u^2$ and $m_1(s,t,x)z = m_2(s,t,x)$ for almost every s,t.*

Proof. Fix $s \in \mathbb{R}$. Set $\bar{u} = u_1 - u_2$ and $\bar{m} = m_1 - m_2$. Then

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{T}^d\times\mathbb{R}} \bar{u}\bar{m} &= \int_{\mathbb{T}^d\times\mathbb{R}} (\partial_t \bar{u})\bar{m} + \bar{u}(\partial_t \bar{m}) \\ &= \int_{\mathbb{T}^d\times\mathbb{R}} (-\Delta_{z,x}\bar{u} + H(t,z,x,D_x u_1) - H(t,z,x,D_x u_2) - D_z\bar{u}\cdot r_t(z) - \tilde{F}(t,z,x,m_1) + \tilde{F}(t,z,x,m_2 \\ &+ \int_{\mathbb{T}^d\times\mathbb{R}} \bar{u}\Delta_{z,x}\bar{m} - \langle D_x\bar{u}, m_1 D_p H(t,z,x,D_x u_1) - m_2 D_p H(t,z,x,D_x u_2)\rangle + s\bar{m}\cdot D_z\bar{u},\end{aligned}$$

where the last term is obtained after integration by parts. Note that

$$\int_{\mathbb{T}^d\times\mathbb{R}} -(\Delta\bar{u})\bar{m} + \bar{u}(\Delta\bar{m}) = 0.$$

We now rewrite the remaining terms in H in the following way

$$\begin{aligned}&\int_{\mathbb{T}^d\times\mathbb{R}} (H(t,z,x,D_x u_1) - H(t,z,x,D_x u_2))\bar{m} - \langle D_x\bar{u}, m_1 D_p H(t,z,x,D_x u_1) - m_2 D_p H(t,z,x,D_x u_2)\rangle = \\ &\quad - \int_{\mathbb{T}^d\times\mathbb{R}} m_1 \left(H(t,z,x,D_x u_2) - H(t,z,x,D_x u_1) - \langle D_p H(t,z,x,D_x u_1), D_x u_2 - D_x u_1\rangle\right) \\ &\quad - \int_{\mathbb{T}^d\times\mathbb{R}} m_2 \left(H(t,z,x,D_x u_1) - H(t,z,x,D_x u_2) - \langle D_p H(t,z,x,D_x u_2), D_x u_1 - D_x u_2\rangle\right) \\ &\quad \leq - \int_{\mathbb{T}^d\times\mathbb{R}} \frac{m_1+m_2}{2C} |D_x u_1 - D_x u_2|^2.\end{aligned}$$

where the inequality follows from the uniform convexity assumption on H.

We now claim that from the assumption on F it follows that

$$\int \varphi_{0,1}(s)\, \mathrm{d}s \int_{\mathbb{T}^d\times\mathbb{R}} \bar{m}(s,t,z,x)(\tilde{F}(t,z,x,m_1) - \tilde{F}(t,z,x,m_2))\, \mathrm{d}s\, \mathrm{d}x\, \mathrm{d}z \geq 0$$

For every t. Indeed, by Lemma 2.2 m_i have the form $m_i(s,t,z,x) = \varphi_{st,\sigma^2 t}\tau_i(x|z)$ and therefore

$$\bar{m}(s,t,z,x) = \varphi_{st,\sigma^2 t}\bar{\tau}(x|z) \tag{4.1}$$

with $\bar\tau = \tau_1 - \tau_2$. Therefore,

$$\int \varphi_{0,1}(s)\bar m(s,t,z,x)(\tilde F(t,z,x,m_1) - \tilde F(t,z,x,m_2))\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}z =$$
$$\int \varphi_{0,1}(s)\bar m(s,t,z,x)\varphi_{r_t(z),\sigma_t^2}(s')\left(F(x,\hat m_1(s',t,\cdot),s') - F(x,\hat m_2(s',t,\cdot),s')\right)\,\mathrm{d}s'\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}z =$$
$$\int \varphi_{0,1}(s)\varphi_{st,\sigma^2 t}(z)\bar\tau(x|z)\varphi_{r_t(z),\sigma_t^2}(s')\left(F(x,\hat m_1(s',t,\cdot),s') - F(x,\hat m_2(s',t,\cdot),s')\right)\,\mathrm{d}s'\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}z =$$
$$\int \varphi_{0,\sigma^2 t+t^2}(z)\varphi_{r_t(z),\sigma_t^2}(s')\bar\tau(x|z)\left(F(x,\hat m_1(s',t,\cdot),s') - F(x,\hat m_2(s',t,\cdot),s')\right)\,\mathrm{d}s'\,\mathrm{d}x\,\mathrm{d}z =$$
$$\int \varphi_{0,1}(s')\varphi_{s't,\sigma^2 t}(z)\bar\tau(x|z)\left(F(x,\hat m_1(s',t,\cdot),s') - F(x,\hat m_2(s',t,\cdot),s')\right)\,\mathrm{d}s'\,\mathrm{d}x\,\mathrm{d}z =$$
$$\int \varphi_{0,1}(s')\,\mathrm{d}s'\left(\int \left(\hat m_1(s',t,x) - \hat m_2(s',t,x)\right)\left(F(x,\hat m_1(s',t,\cdot),s') - F(x,\hat m_2(s',t,\cdot),s')\right)\,\mathrm{d}x\right) \geq 0$$

The first equality follows from (2.3). The second from (4.1). The third and fourth from (2.2). The fifth from (4.1). The inequality follows from the state-wise monotonicity assumption.

Similarly, from the monotonicity assumption on G it follows that

$$\int_{\mathbb{R}\times\mathbb{T}^d\times\mathbb{R}} \varphi_{0,1}(s)\bar m(s,t,z,x)(\tilde G(z,x,m_1) - \tilde G(z,x,m_2))\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}z \geq 0. \tag{4.2}$$

Finally, for the remaining terms with $D_z\bar u$

$$\int \varphi_{0,1}(s)\bar m(s,t,z,x)D_z\bar u\,(-r_t(z)+s) = \mathbb{E}\int_x \tau(x|t,Z)D_Z\bar u(t,z,x)\cdot(S - \mathbb{E}(S|Z)) = 0$$

where the expectation is taken w.r.t. $S\sim\mathscr{N}(0,1)$ and $Z|S=s\sim\mathscr{N}(s,\sigma^2 t)$.

Putting all the estimates together we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \varphi_{0,1}(s)\mathrm{d}s\int_{\mathbb{T}^d\times\mathbb{R}} \bar u\bar m \leq -\int \varphi_{0,1}(s)\mathrm{d}s\int_{\mathbb{T}^d\times\mathbb{R}} \frac{m_1+m_2}{2C}|D_x u_1 - D_x u_2|^2$$

We integrate this inequality on the time interval $[0,T]$ to obtain

$$\int \varphi_{0,1}(s)\mathrm{d}s\int_{\mathbb{T}^d\times\mathbb{R}} \bar u(T)\bar m(T) - \bar u(0)\bar m(0) \leq -\int_0^T\int_{\mathbb{T}^d\times\mathbb{R}} \varphi_{0,1}(s)\frac{m_1+m_2}{2C}|D_x u_1 - D_x u_2|^2 \tag{4.3}$$

Note that $\bar m(0) = m_1(0) - m_2(0) = 0$ while, as $\bar U(T) = \tilde G(t,z,x,m_1(T)) - \tilde G(t,z,x,m_2(T))$,

$$\int \varphi_{0,1}(s)\mathrm{d}s\int_{\mathbb{T}^d\times\mathbb{R}} \bar u(T)\bar m(T) = \int_{\mathbb{T}^d\times\mathbb{R}}\int_{\mathbb{R}\times\mathbb{T}^d\times\mathbb{R}} \varphi_{0,1}(s)\bar m(s,t,z,x)(\tilde G(z,x,m_1) - \tilde G(z,x,m_2))\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}z \geq 0$$

thanks to (4.2). So the LHS of (4.3) is nonnegative, while the RHS is nonpositive, which implies that both sides must vanish. Therefore $D_x u_1 = D_x u_2$ in $\{m_1 > 0\}\cup\{m_2 > 0\}$. As a conse-

quence m_2 actually solves the same equation as m_1 (with the same drifts $D_pH(t,z,x,D_xu_1) = D_pH(t,z,x,D_xu_2)$: hence $m_1 = m_2$. Then, in turn, u_1 and u_2 solve the same HJB equation, so that $u_1 = u_2$. □

Acknowledgments

The authors are grateful to Charles Bertucci and Pierre Cardaliaguet for valuable discussions that helped improve this paper. This work was supported by the French Agence Nationale de la Recherche (ANR) under reference ANR-21-CE40-0020 (CONVERGENCE project). Part of this work was done during a 1-year visit of Bruno Ziliotto to the Center for Mathematical Modeling (CMM) at University of Chile in 2023, under the IRL program of CNRS.

References

[1] Karl J Astrom et al. Optimal control of markov processes with incomplete state information. *Journal of mathematical analysis and applications*, 10(1):174–205, 1965.

[2] Robert J Aumann, Michael Maschler, and Richard E Stearns. *Repeated games with incomplete information*. MIT press, 1995.

[3] Dirk Becherer, Christoph Reisinger, and Jonathan Tam. Mean-field games of speedy information access with observation costs. *arXiv preprint arXiv:2309.07877*, 2023.

[4] Philippe Bergault, Pierre Cardaliaguet, and Catherine Rainer. Mean field games in a stackelberg problem with an informed major player. *arXiv preprint arXiv:2311.05229*, 2023.

[5] James Bergin and Dan Bernhardt. Anonymous sequential games with aggregate uncertainty. *Journal of Mathematical Economics*, 21(6):543–562, 1992.

[6] Charles Bertucci. Mean field games with incomplete information. *arXiv preprint arXiv:2205.07703*, 2022.

[7] Pierre Cardaliaguet. Notes on mean field games. Technical report, Technical report, 2010.

[8] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. *The master equation and the convergence problem in mean field games:(ams-201)*. Princeton University Press, 2019.

[9] Guilherme Carmona and Konrad Podczeck. Approximation and characterization of nash equilibria of large games. *Economic Theory*, 73(2-3):679–694, 2022.

[10] René Carmona, François Delarue, et al. *Probabilistic theory of mean field games with applications I-II*. Springer, 2018.

[11] Philippe Casgrain and Sebastian Jaimungal. Mean field games with partial information for algorithmic trading. *arXiv preprint arXiv:1803.04094*, 2018.

[12] Dena Firoozi and Peter E Caines. epsilon-nash equilibria for major–minor lqg mean field games with partial observations of all agents. *IEEE Transactions on Automatic Control*, 66(6):2778–2786, 2020.

[13] Francoise Forges. Repeated games of incomplete information: non-zero-sum. *Handbook of game theory with economic applications*, 1:155–177, 1992.

[14] Edward J Green. Continuum and finite-player noncooperative models of competition. *Econometrica: Journal of the Econometric Society*, pages 975–993, 1984.

[15] Hugo A Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150, 1992.

[16] Minyi Huang, Roland P Malhamé, and Peter E Caines. Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle. 2006.

[17] Boyan Jovanovic and Robert W Rosenthal. Anonymous sequential games. *Journal of Mathematical Economics*, 17(1):77–87, 1988.

[18] Ehud Kalai and Eran Shmaya. Large strategic dynamic interactions. *Journal of Economic Theory*, 178:59–81, 2018.

[19] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. i–le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006.

[20] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. ii–horizon fini et contrôle optimal. *Comptes Rendus. Mathématique*, 343(10):679–684, 2006.

[21] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.

[22] Andreu Mas-Colell. On a theorem of schmeidler. *Journal of Mathematical Economics*, 13(3):201–206, 1984.

[23] Michael Maschler, Shmuel Zamir, and Eilon Solan. *Game theory*. Cambridge University Press, 2020.

[24] Jean-François Mertens, Sylvain Sorin, and Shmuel Zamir. *Repeated games*, volume 55. Cambridge University Press, 2015.

[25] Yuliy Sannikov and Andrzej Skrzypacz. The role of information in repeated games with frequent actions. *Econometrica*, 78(3):847–882, 2010.

[26] David Schmeidler. Equilibrium points of nonatomic games. *Journal of statistical Physics*, 7:295–300, 1973.

[27] Nevroz Sen and Peter E Caines. Mean field game theory with a partially observed major agent. *SIAM Journal on Control and Optimization*, 54(6):3174–3224, 2016.

[28] Nevroz Sen and Peter E Caines. Mean field games with partial observation. *SIAM Journal on Control and Optimization*, 57(3):2064–2091, 2019.

[29] H Peyton Young. *Strategic learning and its limits*. OUP Oxford, 2004.

[30] Xingyu Zhou. On the fenchel duality between strong convexity and lipschitz continuous gradient. *arXiv preprint arXiv:1803.06573*, 2018.