On non-unique solutions in mean field games

Bruce Hajek and Michael Livesay
Department of Electrical and Computer Engineering and Coordinated Science Laboratory
University of Illinois, Urbana, IL 61801, USA
Email: {b-hajek, mlivesa2}@illinois.edu

***Abstract*—The theory of mean field games is a tool to understand noncooperative dynamic stochastic games with a large number of players. Much of the theory has evolved under conditions ensuring uniqueness of the mean field game Nash equilibrium. However, in some situations, typically involving symmetry breaking, non-uniqueness of solutions is an essential feature. To investigate the nature of non-unique solutions, this paper focuses on the technically simple setting where players have one of two states, with continuous time dynamics, and the game is symmetric in the players, and players are restricted to using Markov strategies. All the mean field game Nash equilibria are identified for a symmetric follow the crowd game. Such equilibria correspond to symmetric ϵ-Nash Markov equilibria for N players with ϵ converging to zero as N goes to infinity.**

In contrast to the mean field game, there is a unique Nash equilibrium for finite N. It is shown that fluid limits arising from the Nash equilibria for finite N as N goes to infinity are mean field game Nash equilibria, and evidence is given supporting the conjecture that such limits, among all mean field game Nash equilibria, are the ones that are stable fixed points of the mean field best response mapping.

I. Introduction and Related Work

The theory of mean field games was initiated independently by Huang, Caines, and Malhamé [4] and Lasry and Lions [5]. The setting of Huang et al. is linear quadratic Gaussian (LQG) control and the setting of Lasry and Lions is continuous state Markov diffusion processes. The work of Gomes, Mohr, and Souza [3] translates much of the theory of [5] into the context of continuous time finite state Markov processes. The LQG and finite state settings are technically simpler than the setting of continuous state Markov processes. All three of these works impose assumptions implying uniqueness of solutions to the mean field game equations.

The paper [4] establishes ϵ-Nash equilibrium properties for strategy profiles consisting of the decentralized individual control laws that result as responses to the collective mass trajectory. Condition H1 of [4] is a key to guaranteeing uniqueness of the mean field equations, In particular, for the other parameters fixed, the value of r in the term for control cost, ru^2, should not be too small. In essence, condition H1 restricts the level of coupling among the players. The mean field game (MFG) equations are expressed as a fixed point of an operator $\mathcal{T}$ in [4]. Proposition 4.5 of [4] states that the fixed point for $\mathcal{T}$ is globally attracting under condition H1 in the paper. Section VI of [4] illustrates a cost gap between individual and global based controls. This is an example of the fact that the social welfare at a Nash equilibrium in game theory does not need to equal the maximum social welfare achievable if the players were to cooperate.

The paper [3] studies the continuous-time, finite state version of mean field game theory. Assumption 3, p. 110, gives a monotonicity condition that ensures uniqueness of solutions to the mean field game equations. Proposition 4 of [3], on the existence of a mean field game Nash equilibrium is proved by using Brouwer's fixed point theorem applied to the map $\theta \mapsto \xi(\theta)$, which is analogous to the map $\mathcal{T}$ of [4]. The domain of ξ is the set $\mathcal{F}$ of uniformly Lipschitz continuous functions on the interval $[0, T]$.

In contrast, multiple solutions of the mean field equations naturally arise in [7], where synchronization of coupled oscillators requires solutions that depart from the incoherence solution. The setup is similar to the discrete-state setting we consider in that it is in continuous time, the players are coupled through their running costs, and players can take actions depending on their own states and on the states of the other players. But the setup in [7] is different in that the state space is continuous – specifically it is the unit circle, and the focus is on infinite horizon average cost. The running cost for player i, $c(\theta_i, \theta_{-i}) = \frac{1}{n}\sum_j (1/2)\sin^2((\theta_i - \theta_j)/2)$, is join the crowd type; it is smaller if the states are closer together. It is similar to flocking of birds or synchronization of fireflies. The separate Brownian motions of different players tend to make them drift apart, and it requires cost for them to try to stick together. If the coefficient R for the cost is large enough it is not worth the players trying to stick close together, and for the MFG limit they will stay uniformly distributed over the circle (i.e. the incoherence solution). As R crosses below some critical value R_c, the incoherence solution still exists but it becomes unstable and additional solutions appear. We find an equivalent phenomena for the simpler discrete state model in this paper. In addition, our setting is considerably simpler than that of [7], allowing us to examine the stability of the mean field map $\mathcal{T}$ for a finite time horizon.

Some of the same issues addressed in this paper are addressed in a different way in [1].

II. Problem formulation

The model we adopt is almost a special case of the model of [4]. We consider $N+1$ players with each having state space $\{0,1\}$. The state $(i(t) : 0 \leq t \leq T)$ of a given player evolves as a controlled Markov process with predictable control α_t, such that the jump probabilities of the state process are given by

$$P(i(t+h) = 1-i | i(t) = i) = (\alpha_t + \eta)h + o(h)$$

for $h > 0$. The parameter $\eta \geq 0$ represents a background jump rate, so if $\eta > 0$ then the process has minimum jump rate η. The background jumping is similar in spirit to the Brownian motions that work against coherence of the coupled oscillators in [7]. The objective function of the reference player is to select (α_t) to solve

$$\min_{\alpha} E\left[\int_0^T c(i(t), \theta_t, \alpha_t)dt + \psi(i(T), \theta_T)\right],$$

where θ_t is the fraction of other players in state 0 at time t. The running costs are assumed to have the form $c(i,\theta,\alpha) = f(i,\theta) + \frac{\alpha^2}{2}$, such that the residence costs per unit time, $f(0,\theta)$ and $f(1,\theta)$, and terminal costs, $\psi(0,\theta), \psi(1,\theta)$, are all bounded, and uniformly Lipschitz continuous in θ.

a) Hamilton Jacobi Bellman (HJB) equation for $N+1$ player system: A state feedback control for a given player is a nonnegative function $(\alpha(i,n,t))$ such that $i \in \{0,1\}$ represents the current state of the player, $n \in \{0,\dots,N\}$ represents the number of other player in state 0, and $t \in [0,T]$. Suppose the reference player uses a state feedback control $(\alpha(i,n,t))$, and the other N players use state feedback control $(\beta(i,n,t))$. Then $(i(t), n(t))_{0\leq t\leq T}$ forms a controlled Markov process on $\{0,1\} \times \{0,1,\dots,N\}$, where $i(t)$ represents the state of the reference player and $n(t)$ represents the number of other players in state 0. The transition rates are as follows:[1]

transition	rate
$(i,n) \to (1-i,n)$	$\alpha(i,n,t) + \eta$
$(i,n) \to (i,n+1)$	$\gamma^+(i,n,t)$ $= (N-n)(\beta(1,n+1-i,t) + \eta)$.
$(i,n) \to (i,n-1)$	$\gamma^-(i,n,t)$ $= n(\beta(0,n-i,t) + \eta)$.

[1]If $j \neq i$ then i itself is one of the "other players" for player j.

Denote the cost-to-go function for the reference player by $u(i,n,t)$. The HJB equations for it are:

$$\begin{aligned} -\dot{u}(i,n,t) &= f(i,n) - \frac{((\alpha^*(i,n,t))^2}{2} \\ &+ \eta(u(1-i,n,t) - u(i,n,t)) \\ &+ \gamma^+(i,n,t)(u(i,n+1,t) - u(i,n,t)) \\ &+ \gamma^-(i,n,t)(u(i,n-1,t) - u(i,n,t)), \quad (1) \\ u(i,n,T) &= \psi(i,n) \quad (2) \end{aligned}$$

where the corresponding control policy is

$$\alpha^*(i,n,t) = (u(i,n,t) - u(1-i,n,t))_+. \quad (3)$$

The HJB equations (1)-(3) can be viewed in two different ways.

- For policy β of the other N players fixed, (1) - (3) determine the best response policy for the reference player. i.e. $\alpha^* = BR(\beta)$.
- To find a symmetric Nash equilibrium, replace $\alpha(\cdot,\cdot,t)$ and $\beta(\cdot,\cdot,t)$ by $\alpha^*(\cdot,\cdot,t)$ in the definition of $\gamma^{\pm}$ and (1)- (3). This yields a $2(N+1)$ dimensional ode with terminal boundary condition and Lipschitz continuous right hand side that uniquely determines the functions $(u(i,n,t))$ and, hence also, the feedback control law α^*. The strategy profile such that all $N+1$ players use α^* is a Markov perfect Nash equilibrium, because α^* is determined backwards from the terminal condition yielding a best response for any interval of the form $[t,T]$. Moreover, the Markov perfect equilibrium is the unique Nash equilibrium among all Markov type (i.e. state feedback) strategy profiles, because the similar HJB equations for a more detailed model description with state space $\{0,1\}^{N+1}$ still has a unique solution and it is necessarily invariant under permutation of the players.

b) Mean field game equilibria and map: A mean field game Nash equilibrium for the finite horizon problem with initial value $\overline{\theta}$ is any solution $(\theta_t, u(i,t))$ to the following equations.[2]

$$\begin{aligned} \dot{\theta}_t &= (1-\theta_t)((u(1,t) - u(0,t))_+ + \eta) \\ &\quad - \theta_t((u(0,t) - u(1,t))_+ + \eta) \quad (4) \\ -\dot{u}(i,t) &= f(i,\theta_t,t) - \eta(u(i,t) - u(1-i,t)) \\ &\quad - \frac{((u(i,t) - u(1-i,t))_+)^2}{2} \quad (5) \\ \theta_0 &= \overline{\theta}, \quad u(i,T) = \psi(i,\theta_T). \quad (6) \end{aligned}$$

Note that the boundary conditions (6) include both initial and terminal values. The mean field equations (4)-(6) can

[2]Note the double use of notation "u." We write $u(i,t)$ for u associated with mean field game solutions and $u(i,n,t)$ for u associated with the $N+1$ player Markov perfect equilibrium.

be written as a fixed point equation, $\theta = \mathcal{T}(\theta)$, where $\mathcal{T}$ maps a collective mass trajectory $(\theta_t : 0 \le t \le T)$ to another trajectory. It is determined by first computing the decentralized individual control laws for the players. Then by the uniform law of large numbers [2], if each of the players follows the same decentralized individual control law, their state processes will be independent and the empirical average of such processes will converge to an expected $\widetilde{\theta}$ that is the output collective mass trajectory. More concretely, $\mathcal{T}(\theta)$ is defined as follows. First, cost-to-go functions $(u(i,t))$ are determined by the HJB terminal value problem for a single player, in response to the collective mass trajectory θ.

$$-\dot{u}(i,t) = f(i,\theta_t) - \frac{((u(i,t)-u(1-i,t))_+)^2}{2} - \eta(u(i,t)-u(1-i,t)) \quad (7)$$

$$u(i,T) = \psi(i,\theta_T). \quad \text{boundary condition at } T \quad (8)$$

Then $\widetilde{\theta}_t$, the probability a single player using the decentralized state-feedback control $\alpha_t(i,t) = (u(i,t)-u(1-i,t))_+$ is in state 0 at time t, is determined by the initial value problem (Kolmogorov forward equation):

$$\begin{aligned} \dot{\widetilde{\theta}}_t &= (1-\widetilde{\theta}_t)((u(1,t)-u(0,t))_+ + \eta) \\ &\quad - \widetilde{\theta}_t((u(0,t)-u(1,t))_+ + \eta) \\ \widetilde{\theta}_0 &= \overline{\theta} \qquad \text{boundary condition at } 0 \end{aligned}$$

Motivated by the law of large numbers, $\widetilde{\theta}$ is defined to be the new collective mass trajectory, i.e. $\widetilde{\theta} = \mathcal{T}(\theta)$.

The mean field game equations (4) and (5), with the addition of an average cost per unit time term κ on the right-hand side of (5) correspond to an *infinite horizon game* for average cost per unit time. (See [3], Section 2.12, p. 117.) In that case the value functions $u(i,t)$ represent realative cost to go. The boundary conditions (6) are replaced by the condition that θ be constant in time or be periodic.

c) Fluid limits of Markov perfect equilibrium: As noted in the introduction, there can be multiple mean field game Nash equilibria, even for a finite horizon problem with given boundary conditions. A mean field game Nash equilibrium $(\theta_t, u(i,t))$ yields a decentralized player strategy $\alpha_t(i,t) = (u(i,t)-u(1-i,t))_+$. For finite N, the strategy profile such that every player uses $(\alpha_t(i,t))$ is easily seen to be an ϵ-Nash equilibria such that $\epsilon \to 0$ as $N \to \infty$. For details see the appendix of the full version of this paper.[3]

However, for finite N there is a unique Markov perfect Nash equilibrium strategy profile, so for a given initial condition, the distribution of the finite N system is uniquely determined. It is natural, therefore, to single out collective mass trajectories that arise as limits of the mass trajectories for Markov perfect equilibria.

Definition II.1. *Let $n^N(t)$ denote the number of players in state 0 at time t under the unique symmetric Markov perfect equilibrium for the $N+1$ player game, and for some initial condition depending on N. Then $\theta = (\theta_t : 0 \le t \le T)$ is a* fluid limit Markov perfect trajectory *(FLMP trajectory) if for some sequence of initial states with $\lim_{N\to\infty} \frac{n^N(0)}{N} \to \theta_0$, the following holds for any $\epsilon > 0$,*

$$\lim_{N\to\infty} \mathbb{P}\left[\left|\frac{n^N(t)}{N+1} - \theta_t\right| < \epsilon \text{ for } 0 \le t \le T\right] = 1. \quad (9)$$

Proposition 1. *Suppose $\eta > 0$. An FLMP trajectory is a mean field game Nash equilibrium.*

See Appendix A for a proof. We conjecture the proposition is also true for $\eta = 0$, but a change of probability measure argument in the proof breaks down if $\eta = 0$. Proposition 1 raises the question of how to identify which mean field Nash equilibria are FLMP trajectories.

d) Contributions of the paper: Proposition 1 is new and its proof extends to the general setting of [3]. It shows that the search for FLMP trajectories can be limited to the mean field game Nash equilibria. The next contribution of this paper is to identify all of the MFG equilibria for a natural special case of the two state model called follow the crowd. This model is analogous to the model of synchronization of oscillators game [7], but considerably simpler, so we can identify the finite horizon solutions as well as the infinite horizon ones. The third contribution is to offer the following conjecture, and give evidence for it:

Conjecture 1. *The FLMP trajectories are the stable fixed points of the MFG mapping $\mathcal{T}$.*

A similar type of conjecture is implicit in [7] based on a notion of stability for constant, long-term average cost infinite horizon solutions, called linear asymptotic stability. The paper [7] identifies the critical cost threshold at which the incoherence solution becomes unstable. In addition to giving evidence for Conjecture 1 in the setting of finite horizon games, we also show that the results of [7] for constant, long-term average cost infinite horizon solutions, carry over to the setting of two state Markov processes. For the infinite horizon framework, we show asymptotic stability of certain fixed points for the nonlinear dynamics in Section III-C, and an appendix in the full version of this paper gives an analysis based on the notion of linear asymptotic stability introduced in [7]. Additional results are given in the appendix of the full version of this paper, including, for contrast, a similar analysis for an avoid the crowd model with

[3]See full version at arXiv.org.

unique mean field game solutions, and a description of a partial differential equation (PDE) (given for more general model in [3]) that can be considered to be an extension of the notion of mean field game.

III. MFG EQUILIBRIA FOR FOLLOW THE CROWD

The *follow the crowd* model corresponds to the following cost per time spent in state i:

$$f(i,\theta) = |1-\theta-i| = \begin{cases} 1-\theta & i=0 \\ \theta & i=1 \end{cases}$$

In particular, if $\theta > 1/2$ (more than half of the other players in state 0), then state 0 has smaller cost per unit time than state 1.

Letting $y = u_1 - u_0$, $x = 2\theta - 1$, the mean field equations (4)- (6) can be written as:

$$\begin{aligned} \dot{x} &= y - x|y| - 2\eta x \\ -\dot{y} &= x - \tfrac{1}{2}y|y| - 2\eta y \end{aligned} \tag{10}$$

with the boundary conditions $x_0 = 2\overline{\theta} - 1$ and $y_T = \psi\left(1, \frac{1+x_T}{2}\right) - \psi\left(0, \frac{1+x_T}{2}\right)$. Once a solution (x,y) to (10) is found for the finite horizon problem over $[0,T]$, a corresponding solution (u_0, u_1, θ) to the mean field game equations can be found by simply integrating (4)- (5) because the righthand sides of (4)- (5) are determined by (x_t, y_t).

A useful fact is that the equations (10) form a Hamiltonian system, for the Hamiltonian function H:

$$H(x,y) = \frac{x^2 - 4\eta xy + y^2 - xy|y|}{2}. \tag{11}$$

In other words, (10) has the form $\dot{x} = H_y$ and $\dot{y} = -H_x$, where H_x and H_y represent partial derivatives of H. Consequently, the value of H is constant along the solutions of (10), because $\frac{dH(x_t,y_t)}{dt} = \langle \nabla H, \binom{H_y}{-H_x} \rangle \equiv 0$, so the trajectories trace out level contours of H. This model is a special case of potential mean field games defined in [3], Section 5, for which Hamiltonians exist.

Contour maps of H are shown in Fig. 1 for various values of η. For small values of x, y the quadratic terms in H dominate the cubic term, and for $\eta < 1/2$, constant $x^2 - 4\eta xy + y^2$ gives elliptical orbits of x, y, in the clockwise direction.

A. Finite time horizon mean MFG solutions

For the finite horizon mean field game with zero terminal cost (i.e. terminal boundary condition $y_T = 0$), and initial state $x_0 = 0$, correspond to paths that begin on the y axis (so the initial condition $x_0 = 0$ is satisfied) and end on the x axis. One solution is $(x_t, y_t) \equiv (0,0)$ for $0 \le t \le T$. Let $\phi = \arctan\left(\frac{x}{y}\right)$ denote the angle

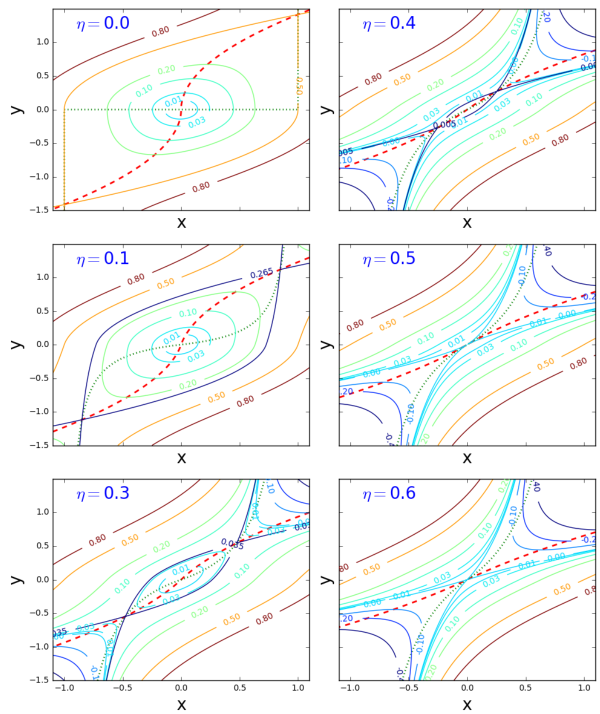

Fig. 1: Contour plot of H for several values of η. Dashed lines are the zero sets of H_x, and dotted lines are the zero sets of H_y. The intersections of dotted and dashed lines are the critical points of H (i.e. solutions to $\nabla H = 0$.)

of x, y from the positive x axis. The angular velocity of (x,y) is given by

$$\dot{\phi} = \frac{\dot{y}x - y\dot{x}}{x^2+y^2} = -1 + \frac{\frac{3}{2}xy|y| + 4\eta xy}{x^2+y^2} \tag{12}$$

It is negative along the y axis, indicating clockwise motion. If $\eta \ge 1/2$ then $\dot{\phi} > 0$ along the line $x = y$, indicating that $y = 0$ is never reached. Thus, if $\eta \ge 1/2$, the trajectory $(0,0)$ is the only MFG equilibrium.

If $\eta < 1/2$ then $\dot{\phi} < 0$ for (x,y) in a neighborhood of the origin, indicating clockwise movement. Moreover, for ϕ fixed, $\dot{\phi}$ is an increasing function of the distance of (x,y) to the origin (decreasing angular speed because angular velocity is negative). Thus, the time for (x,y) to traverse a contour across the first quadrant is increasing in y_0. for $y_0 > 0$. As $y_0 \to \mathbf{0}$ the dynamics is given, to first order, by the MFG linearized about $(0,0)$, given by

$$\begin{aligned} \dot{x} &= y - 2\eta x \\ -\dot{y} &= x - 2\eta y \end{aligned} \tag{13}$$

with solution of the form (setting $x_0 = 0$ and $y_0 > 0$):

$$\begin{aligned} x_t &= \sin\left(\sqrt{1-4\eta^2}\; t\right) \\ y_t &= \sin\left(\sqrt{1-4\eta^2}\; t + \arccos(2\eta)\right) \end{aligned}$$

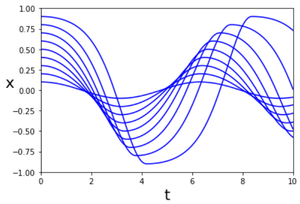

Fig. 2: Several solutions with various terminal values of x run backwards in time, for follow the crowd dynamics with $\eta = 0$.

The time it takes the linear system to traverse the first quadrant is $T_c(\eta) \triangleq \frac{\pi - \arccos(2\eta)}{\sqrt{1-4\eta^2}}$. Hence, as $y_0 \to 0$, the traversal time for the quadrant converges to $T_c(\eta)$. Thus, for $\eta < 1/2$ and $T \le T_c(\eta)$, $(0,0)$ is the unique solution to the MFG. For $T > T_c(\eta)$ there is one more solution that remains in, and traverses, the first quadrant, and the negative of that solution remains in, and traverses, the third quadrant. For T large enough there are solutions that traverse contours of H through three quadrants, five quadrants, and so on. A similar radial velocity analysis for the pair $(y, \dot{y})$ (see appendix of full version of the paper) establishes that the entire periods of the dynamical system are increasing with amplitude, as illustrated in Fig. 2. Since the dynamics is symmetric under rotation by π, we conclude that for any odd number k, starting on the positive y axis, the time required to rotate through k quadrants is increasing in the initial condition y_0. Therefore, as T increases from 0, the number of solutions starts at one and jumps up by two when T crosses times of the form $T_c + k\pi/(\sqrt{1-4\eta^2})$ for $k \ge 1$. Equivalently, the number of solutions is $1 + 2\left\lceil \frac{(T-T_c)\sqrt{1-4\eta^2}}{\pi} \right\rceil$.

B. Infinite horizon constant or periodic MFG solutions

The equilibrium points of the dynamics (10) are the critical points of the Hamiltonian function (i.e. $\nabla H = 0$), and are given as follows. If $0 \le \eta < 0.5$, $(0,0)$ is an equilibrium point and there are also exactly two nonzero equilibrium points, given by $\pm\overline{P}$, where

$$\overline{P} = \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix} \triangleq \begin{pmatrix} 1 - \eta^2 - \eta\sqrt{2+\eta^2} \\ \sqrt{2+\eta^2} - 3\eta \end{pmatrix}. \tag{14}$$

If $\eta \ge 0.5$, $(0,0)$ is the unique equilibrium point.

Regarding infinite horizon periodic solutions, examination of H and the equations for angular velocity, (12) and similar equation for angle of $(y, \dot{y})$, lead to the following conclusions. If $0 \le \eta < 0.5$, there is a two-dimensional family of periodic solutions that can be indexed by the peak amplitude of x (ranges over $(0, \overline{x})$) and phase. The period of the solutions increases continuously over $(2\pi/\sqrt{1-4\eta^2}, \infty)$ as the peak amplitude of x increases over $(0, \overline{x})$. If $\eta \ge 0.5$, there are no periodic solutions of (10).

C. Infinite horizon convergent transient MFG solutions, and the asymptotically stable constant solutions

Consider the initial value problem over $t \in [0, \infty)$ with some initial condition (x_0, y_0) and dynamics (10). First, suppose $0 \le \eta < 0.5$. For any initial condition (x_0, y_0) such that $x_0 \ne 0$, one of four cases holds: x_t is periodic with a positive period, x converges to $\overline{P}$, x converges to $-\overline{P}$, or x_t exits $[-1, 1]$ in finite time. The following categorize the convergent solutions such that x_t remains in $[-1, 1]$.

- For any initial value of $x_0 \in (-\overline{x}, \overline{x})$, there exist two corresponding initial values of y_0 such that the solution of the initial value problem satisfies (i) $x_t \in [-1, 1]$ for all t and (ii) the solution converges to a limit as $t \to \infty$. For the smaller value of y_0 the limit is $-\overline{P}$ and for the larger value of y_0 the limit is $\overline{P}$. The value of the larger y_0 for example is such that the contour of H through (x_0, y_0) contains $\overline{P}$.
- For an initial value $x_0 \in [-1, -\overline{x}]$ there exists a unique value of y_0 such that the solution of the initial value problem satisfies $x_t \in [-1, 1]$ for all t. That solution converges to $-\overline{P}$ as $t \to \infty$.
- Similarly, for an initial value $x_0 \in [\overline{x}, 1]$ there exists a unique value of y_0 such that the solution of the initial value problem satisfies $x_t \in [-1, 1]$ for all t. That solution converges to $\overline{P}$ as $t \to \infty$.

Second, suppose $\eta \ge 0.5$. For any $x_0 \in [-1, 1]$, there is a unique value of y_0, such that the solution of the initial value problem for (10) satisfies $x_t \in [-1, 1]$ for all t. Furthermore, y_0 has the same sign as x_0, and the solution converges to $(0,0)$ as $t \to \infty$. The value of y_0 is the root of $H(x_0, y_0) = 0$ (for x_0 fixed) that is closer to zero.

The above observations give a sense in which $\pm\overline{P}$ is an asymptotically stable equilibrium point of the dynamics (10) if $0 \le \eta < 0.5$, and $(0,0)$ is an asymptotically stable equilibrium point if $\eta \ge 1/2$. This sense of stability is not the usual definition of (Lyapunov) stability because we ask, for given x_0, whether there exists an associated value of y_0 giving the desired convergence. The asymptotically stable limit points are saddlepoints of H.

As mentioned above, a related definition of stability, called linear asymptotic stability, is formulated in [7]. That definition and the results of [7] for it are translated to the model of this paper in the appendix in the full version of this paper.

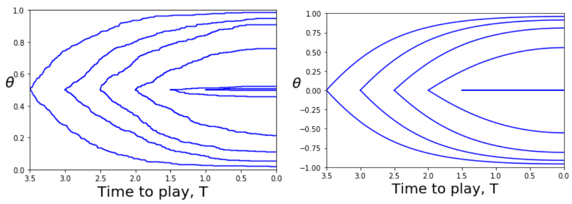

Fig. 3: On the left is a set of realizations of the $N+1$-player game with 400 players and various time to play, with initially 200 players in each state. On the right, are the MFG solutions believed to be the FLMP trajectories. Both are for follow-the-crowd game with $\eta = 0$.

IV. Evidence for Conjecture 1

In order to explore whether Conjecture 1 is true, it is natural to explore two sides of the question. One side is to identify the FLMP trajectories. Numerically that can be done by solving the $2(N+1)$ dimensional HJB equation for the system with $N+1$ players to find the strategy $\alpha^*(i,n,t)$ players use for the Markov perfect equilibrium with $N+1$ players, and then either simulating the corresponding occupancy process through Monte Carlo simulation of $N+1$ players independently using that policy, or solving the Kolmogorov forward equations to find the marginal distribution, mean and variance of the number of players in state 0 vs. time.

The other side is to identify the stable fixed points of $\mathcal{T}$. Two ways to explore which fixed points of $\mathcal{T}$ are stable are to either numerically investigate the orbit trajectories as $\mathcal{T}$ is repeatedly applied to some initial trajectory, or to examine the linearization of $\mathcal{T}$ about a fixed point–this is the Gateaux derivative and it can be expressed as an integral operator. The eigenvalues can be computed numerically, and in rare cases, analytically. By abuse of notation, we use $\mathcal{T}$ to denote the mean field map as a mapping $T(x) \mapsto \widetilde{x}$ obtained by the change of coordinates $x = 2\theta - 1$.

a) Numerical identification of FLMP trajectories: For the symmetric follow the crowd model, numerical analysis strongly and consistently indicates which MFG solutions are FLMP trajectories. We find that for $\eta \leq 1/2$ they coincide with the unique MFG equilibrium – namely, the (0,0) trajectory over $[0,T]$. And for $\eta > 1/2$ there are two FLMP trajectories. Namely, the one that traverses the first quadrant in the x-y plane once, and the negative of it, which traverses the third quadrant in the x-y plane once. In particular, the solutions that wind around the origin through three or more quadrants do not appear to be FLMP solutions. See Fig. 3 for illustration. For less symmetric examples it is less obvious where the *bifurcation curve* is that separates FLMP solutions that converge to a point closer to 1, or converge to a point closer to 0. The bifurcation curve often coincides with a line or curve of indifference for the $N+1$ player game with a large number of players, corresponding to upcrossings of zero by the mapping $n \mapsto u_1(0,n,t) - u_0(0,n,t)$. This is illustrated in Fig. 4.

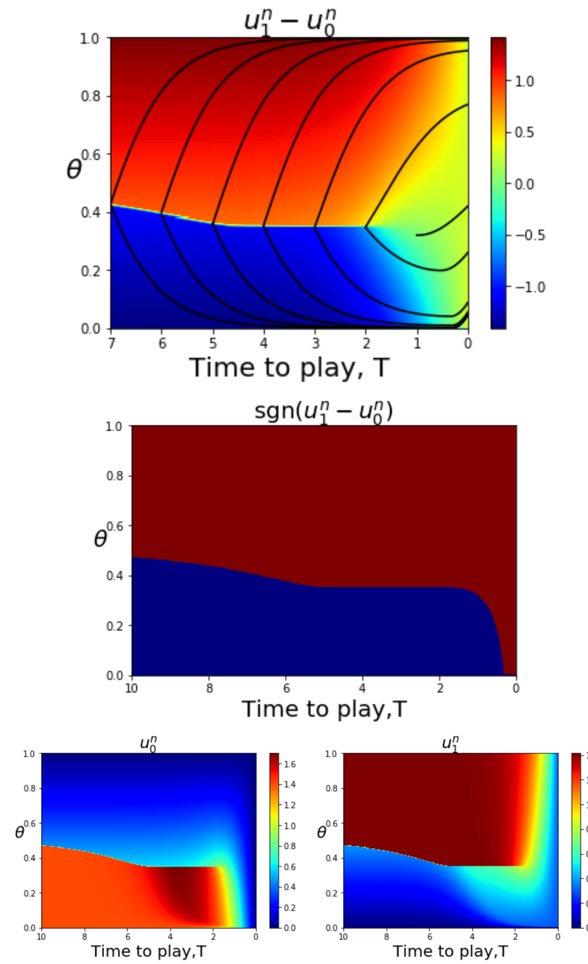

Fig. 4: Heat maps for cost-to-go functions for follow the crowd, $f(i,\theta) = |\theta - (1-i)|$, with $N = 400$, $T = 10$, $\eta = 0$, and asymmetric terminal cost: $\psi(1) = 0.3$ and $\psi(0) = 0$. The MFG equilibrium trajectories beginning at the bifurcation curve are overlaid onto the heat map of $u_1 - u_0$ in the top figure.

b) Examination of orbits of $\mathcal{T}$: Recall that the fixed points of $\mathcal{T}$ are the collective mass trajectories $(\theta_t : 0 \leq t \leq T)$ of mean field Nash equilibria. To numerically investigate the stability of fixed points of $\mathcal{T}$ we generated sequences of iterates of trajectories $(\theta^n)_{n\geq 0}$ defined by $\theta^{n+1} = \mathcal{T}(\theta^n)$, where the initial point θ^0 is a perturbation of a fixed point. Figure 5 shows such sequences of iterates such that the initial trajectory is a perturbation of one of the two MFG Nash equilibria that cross zero one time, for the follow the crowd game and time horizon $T = 20$. In both instances, the iterates converged to one of the two equilibria with no zero crossings.

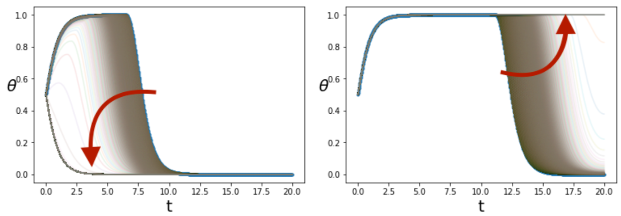

Fig. 5: Iterates $(\theta^n)_{0\le n\le 10000}$ for two different initial trajectories that are perturbations of a single-cross MFG Nash equilibrium, which is indicated by a thick blue line.

However, overall we found it difficult to numerically verify that a given solution is not a stable fixed point. On one hand, some MFG solutions that we don't expect to be stable, such as the trajectory that crosses zero once, numerically appear to be asymptotically stable for a very small basin of stability. On the other hand, we have found perturbations of MGF solutions that also numerically appear to be asymptotically stable, indicating numerical artifacts are possible.

c) Linearization of $\mathcal{T}$ about $(0,0)$: Given a fixed point $\bar{x} = \mathcal{T}(\bar{x})$, the Gateaux derivative $d\mathcal{T}_X(\bar{x}, x)$, or the directional derivative of $\mathcal{T}$ at $\bar{x}$ in the direction x, is obtained by linearizing $\mathcal{T}$ about $\bar{x}$. This is particularly simple if $\bar{x}$ is the zero trajectory. (Linearization about a nonzero trajectory is given in the full version of this paper.) In that case, the linearized MFG equations are:

$$\begin{aligned} \dot{x} &= y - 2\eta x \\ -\dot{y} &= x - 2\eta y \end{aligned} \tag{15}$$

Given (x_u), $\widetilde{x} = d\mathcal{T}_X(\bar{x}, x) = \mathcal{L}_2\mathcal{L}_1 x$, where $\mathcal{L}_1$ and $\mathcal{L}_2$ are linear operators defined as:

$$\begin{aligned} y_s = (\mathcal{L}_1 x)_s &= \int_s^T e^{-2\eta(T-u)} x_u du \\ \widetilde{x}_t = (\mathcal{L}_2 y)_t &= \int_0^t e^{-2\eta(t-s)} y_s ds \end{aligned}$$

These expressions can be combined to yield

$$x_t = \int_0^T K(t,u) x_u du$$

where $K(t,u) = e^{-2\eta(t\vee u)} \sinh(2\eta(t\wedge u))/2\eta$ for $\eta > 0$ and $K(t,u) = t \wedge u$ for $\eta = 0$. In other words, the Gateaux derivative is the integral operator with kernel K.

If $\eta = 0$, $K(t,u) = t \wedge u$, which is the covariance of Brownian motion, which has a well known Mercer series expansion. The eigenvalues of K are $\lambda_n = \left(\frac{2T}{(2n+1)\pi}\right)^2$ with corresponding eigenfunctions $h_n(t) = \sin\left(\frac{(2n+1)\pi t}{2T}\right)$ for $n \ge 0$. In particular, the largest eigenvalue is $\lambda_0 = \left(\frac{2T}{\pi}\right)^2$, and $\lambda_0 \le 1$ if and only if $T \le T_c(0) = \pi/2$, where $T_c(\eta)$ is the critical time horizon for the appearance of multiple MFG equilibria.

Here is an upper bound on the maximum eigenvalue of K for $\eta > 0$. The mappings $\mathcal{L}_1$ and $\mathcal{L}_2$ are both bounded operators in the supremum norm: $\|y\|_\infty \le c(\eta,T)\|x\|_\infty$, with operator norm $c(\eta,T) = \int_0^T e^{-2\eta t} dt = \frac{1-e^{-2\eta T}}{2\eta}$. Thus, the Gateaux derivative is also a bounded operator in the supremum norm with operator bound $c^2(\eta,T)$.[4] Hence, if $\eta \ge 1/2$, the linearized mapping is a contraction in the L^∞ norm for all $T > 0$. If $\eta < 1/2$ it is a contraction if T is small enough that $\frac{1-e^{-2\eta T}}{2\eta} < 1$.

For $\eta > 0$ we conjecture the largest eigenvalue of K is greater than one precisely when there is a nonzero MFG equilibrium, namely, when $T > T_c(\eta) \triangleq \frac{\pi - \arccos(2\eta)}{\sqrt{1-4\eta^2}}$. We numerically found the largest eigenvalue of the matrix approximation of the kernel, $(K(iT/n, jT/n)T/n)_{i,j\in[n]}$ for $n = 10^3$ for $\eta \in (0, 0.499)$ and T near T_c, and the calculations match the conjecture well.

References

[1] A. Cecchin, P.D. Pra, M. Fischer, and G. Pelino. On the convergence problem in mean field games: a two state model without uniqueness. arXiv 1810.05492, October 2018.

[2] E. Giné and J. Zinn. Some limit theorems for empirical processes. *The Annals of Probability*, 12:929–989, 1984.

[3] Diogo A Gomes, Joana Mohr, and Rafael Rigão Souza. Continuous time finite state mean field games. *Applied Mathematics & Optimization*, 68(1):99–143, 2013.

[4] Minyi Huang, Peter E Caines, and Roland P Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria. *IEEE Transactions on Automatic Control*, 52(9):1560–1571, 2007.

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

[6] Jan H. Van Schuppen and Eugene Wong. Transformation of local martingales under a change of law. *The Annals of Probability*, 2(5):879–888, 1974.

[7] H. Yin, P.G. Mehta, S.P. Meyn, and U.V. Shanbhag. Synchronization of coupled oscillators in a game. *IEEE Transaction on Automatic Control*, 57(4):920–935, 2012.

Appendix A
Proof of Proposition 1

This section proves Proposition 1, that if $\eta > 0$, FLMP trajectories are mean field game equilibria. The proof is given after some initial notation is given and two lemmas are proved. Let $(\theta_t)_{0\le t\le T}$ be an FLMP trajectory and let $(i^N(0), n^N(0))_{N\ge 1}$ be a corresponding sequence of initial conditions as in the definition of FLMP trajectory. For $N \ge 1$, let $((i(t), n(t)) : 0 \le t \le T)$ denote the controlled Markov process for $N+1$ players resulting

[4] A somewhat tighter bound is given by $\|\widetilde{x}\|_\infty \le \widetilde{c}(\eta,T)\|x\|_\infty$, where $\widetilde{c}(\eta,T) = \max_t \int_0^T K(t,s)ds$, but the expression for $\widetilde{c}(\eta,T)$ is complicated.

for initial state $(i^N(0), n^N(0))$, when all players use the unique policy $(\alpha^*(i,n,t))$ for the Markov perfect equilibrium for $N+1$ players. Since the functions $f(i,\theta,t)$ and $\psi(i,\theta)$ are bounded, for T fixed, the cost to go functions $u(i,n,t)$ determined by the HJB equations (1)- (2) are uniformly bounded for all N, i, n, and $t \in [0,T]$. Therefore, the policy α^*, determined by (3), is also uniformly bounded. Select Γ_1 such that $(\alpha^*(i,n,t)) \leq \Gamma_1$ for all N, i, n, and $t \in [0,T]$. Suppose also that Γ_1 is large enough that $\alpha(i,t) \leq \Gamma_1$ for all i,t for any decentralized policy $\alpha(i,t)$ resulting by responding to a deterministic collective mass trajectory.

Consider the following variation of the Markov perfect equilibrium. Suppose the reference player switches from using α^* to some other policy, $\beta^*(i,n,t)$, such that $\beta^*(i,n,t) \leq \Gamma_1$ and $t \mapsto \beta^*(i,n,t)$ is continuous for all (i,n). Let P denote the original probability distribution for the process $(i(t), n(t))_{0\leq t\leq T}$ and let $\widetilde{P}$ denote the probability distribution of $(i(t), n(t))_{0\leq t\leq T}$ when the reference player switches to policy β^*.

Lemma 1. *(Insensitivity of FLMP trajectory to one player switching policies) The following holds for any $\epsilon > 0$,*

$$\lim_{N\to\infty} \widetilde{\mathbb{P}}\left[\left|\frac{n^N(t)}{N+1} - \theta_t\right| < \epsilon \text{ for } 0 \leq t \leq T\right] = 1. \quad (16)$$

Lemma 2. *Let P and $\widetilde{P}$ be probability distributions on the same measurable space $(\Omega, \mathcal{F})$ such that $\widetilde{P} << P$ (i.e. $\widetilde{P}$ is absolutely continuous with respect to P) and let $\frac{d\widetilde{P}}{dP}$ denote the Radon-Nikodym derivative. Suppose $E_P\left[\left(\frac{d\widetilde{P}}{dP}\right)^p\right]^{1/p} \leq c$ for some $p > 1$ and c. Let $q > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then for any event A, $\widetilde{P}(A) \leq cP(A)^{1/q}$.*

Proof of Lemma 2. By Hölder's inequality,

$$\begin{aligned}\widetilde{P}(A) &= \int_\Omega \frac{d\widetilde{P}}{dP}\mathbf{1}_{\{A\}}dP \\ &\leq c\left(\int_\Omega \mathbf{1}^q_{\{A\}}dP\right)^{1/q} = cP(A)^{1/q}\end{aligned}$$

□

Proof of Lemma 1 . Since P and $\widetilde{P}$ only differ by the change in the policy for player 1, the Radon-Nikodym derivative $\frac{d\widetilde{P}}{dP}$ can be written explicitly as follows. Let $(Y_t)_{0\leq t\leq T}$ denote the number of jumps of the state of the reference player during $[0,t]$. Then, by standard theory of change of probability measure for point processes (Girsanov type result for point processes, see [6], Theorem 4.1 for example), $\widetilde{P} << P$ and the Radon-Nikodym derivative is given by

$$\frac{d\widetilde{P}}{dP} = \exp\left(\int_0^T \ln\left(\frac{\beta^* + \eta}{\alpha^* + \eta}\right)dY_t - \int_0^T (\beta^* - \alpha^*)dt\right)$$

where β^* is short for $\beta^*(i(t-), n(t-), t)$, α^* is short for $\alpha^*(i(t-), n(-), t)$ and η is the fixed positive background jump rate.

Note that for $p > 1$, the expression for the Radon-Nikodym derivative to the p^{th} power can be written as a product

$$\left(\frac{d\widetilde{P}}{dP}\right)^p = \frac{d\widetilde{\widetilde{P}}}{dP}e^{\int_0^T (\beta^*+\eta)^p - (\alpha^*+\eta)^p - p(\beta^*-\alpha^*)dt} \leq \frac{d\widetilde{\widetilde{P}}}{dP}\Gamma_2$$

where $\widetilde{\widetilde{P}}$ is a probability measure corresponding to a similar Radon-Nikodym derivative with a factor p in front of the log term, and $\Gamma_2 = \exp\left[T\left((\Gamma_1+\eta)^p + p\Gamma_1\right)\right]$. Thus, $E_P\left[\left(\frac{d\widetilde{P}}{dP}\right)^p\right] \leq \Gamma_2$. Lemma 1 thus follows from Lemma 2 with A equal to the complement of the event in (16). □

Proof of Proposition 1. Consider the Markov perfect equilibrium for large N. In view of Lemma 1, if the reference player deviates from using α^*, the normalized process $n(t)/N$ for the rest of the population still follows θ arbitrarily closely as $n \to \infty$. Thus, an asymptotically optimal policy for the reference player to switch to is the optimal response to deterministic collective mass trajectory θ. Furthermore, it implies $u(n,i,t) - u(i,t)$ converges to zero uniformly in n and $t \in [0,T]$, where $u(n,i,t)$ is associated with the $N+1$ player MP equilibrium, and $u(i,t)$ is the cost-to-go for the single reference player responding to the deterministic mass trajectory θ. It follows that all players in the $N+1$ game are asymptotically effectively using the same policy as the alternate policy of the reference player. (in other words, $(u(i,n,t) - u(1-i,n,t))_+ \approx (u(i,t) - u(1-i,t))_+$). Thus, the corresponding fluid limit is the same as the mean limit for the reference player with random initial state equal to 0 with probability $n^M(0)/n$. □