2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Hybrid mean field game dynamics in large population

Hamidou Tembine

Abstract— **We consider finite populations of interacting players with different types and finite action set per type. Under suitable conditions we derive the mean field game dynamics which can be deterministic or stochastic depending on how the system behave with the time-scales. Connection between mean field game dynamics and evolutionary game dynamics are established. Considering different revision protocols for each player, we derive an hybrid mean field game dynamics which offers the possibility of elimination of non-Nash rest points and give nice convergence properties in potential games and stable games.**

I. Introduction

Population games and mean field game dynamics provide a framework for describing strategic interactions among large number of players. Originally formulated to explain complicated stochastic large system behaviors, mean field dynamics become a fundamental technique in the field of large-scale systems. Mean field game dynamics covers a large class of game dynamics known in evolutionary game theory [3], [5], [9], [2], [4].

Traditionally, predictions of behavior and outcome in game theory are based on some notion of equilibrium, typically Cournot equilibrium (Cournot, 1838), Bertrand equilibrium (Bertrand, 1883), conjectural variation (Bowley, 1924), min-max equilibrium (von Neumann 1928), Stackelberg solution (Stackelberg, 1934), Nash equilibrium (Nash, 1950), Wardrop equilibrium (Wardrop, 1952) or some refinement and/or extensions thereof. Most of these solution concepts require the assumption of equilibrium knowledge, which assume that that each player correctly anticipates how the other players will react. The equilibrium knowledge assumption is too strong and is difficult to justify in particular in context with large populations of players. As an alternative to the equilibrium approach, the evolutionary game approach proposes an explicitly dynamic updating choice, a model in which players myopically update their behavior in response to their current strategic environment which is given by mean field game dynamics. This dynamic procedure does not assume the automatic coordination of players' actions and beliefs, and it can derive many players' actions and transition rates. These procedures are specified formally by defining a revision of actions called revision protocol [4] which is obtained from the transition kernels of the mean field process. A revision protocol takes current payoffs and the system state as arguments; its outputs are conditional switch rates which describe how frequently players in some class playing action who are considering switching strategies switch to another action, given the current expected payoff vector and population state. This revision of actions is flexible enough to incorporating a wide variety of paradigms, including ones based on learning, imitation, adaptation, optimization, etc. The revision of actions describe the procedures players follow in adapting their behaviors in the dynamic evolving environment such as evolving networks (Internet traffic, flow control etc).

In this paper, we introduce and analyze hybrid mean field game dynamics in large population and present a simulation framework based on the mean field process.

Our contribution can be summarized as follows. We propose an explicit interaction model in large populations and present generic convergence results. We derive not only deterministic mean field game dynamics but also stochastic mean field game dynamics. We illustrate limit cycles and non-commutativity of the mean field process. We show that a major class of evolutionary game dynamics that are used in *evolutionary game theory* can be obtained from the mean field game dynamics. This allows us to address the question of: *What happens if the players use different revision protocols and different learning schemes ?* We derive hybrid mean field game dynamics and give some convergence properties.

The paper is organized as follows. In Section II we present the mean field model. In section III we present the derivation of mean field game dynamics and their connection to evolutionary game dynamics. Hybrid mean field game dynamics are analyzed in Section IV. Section V concludes the paper.

II. The setting

Consider a system consisting of large populations of players. Each player has a finite number of actions. At each stage, a random set of players interact. The actions of all the interacting players determine together the instantaneous payoffs and the probability transitions to the next actions. We study the convergence of the Markov process with variable set of interacting players when the total number of possible players grow without bound. The limiting games are equivalent to population games. Time $t \in \mathbb{N}$ is discrete. There is a set of resources those states are represented by $S_t^n \in \mathcal{S}$ (finite). There are n players ($n \geq 2$). For every player j, $\mathcal{X}$ is its own state space. An individual state has two components: the type of the player and the current action. The type is a constant during the game. The state of player j at time t is denoted by $X_{j,t}^n = (\theta_j, A_{j,t}^n)$ where θ_j is the type. The set of possible states is finite. $\mathcal{A}_j$ may include other parameters, such as, space location, current direction and so on. The individual state of player j at time

H. Tembine is with Ecole Supérieure d'Electricité, SUPELEC, France. E-mail: tembine@ieee.org

t is denoted by $X^n_{j,t}$. For each player j, $\mathcal{A}_j(s,\theta_j)$ is the set actions that are available to player j. We assume that the set $\mathcal{A}_j(s,\theta_j)$ depends only the type θ_j. The action of player j at time t is $A^n_{j,t}$. The *global state* of the system at time t is $(S^n_t, X^n_t) = (S_t, X^n_{1,t}, ..., X^n_{n,t})$. Denote by $A^n_t = (A^n_{1,t}, \ldots, A^n_{n,t})$ the action profile at time t. The system (S^n_t, X^n_t) is Markovian once the action profile A^n_t are drawn under Markovian strategies. The players are coupled not only via their instantaneous payoff function $\tilde{u}(S^n_t, X^n_t)$ but also via the state evolution X^n_t i.e the evolution of $X^n_{j,t}$ depends on the states and the actions of the other players.

Define M^n_t to be the current population profile i.e

$$M^n_{x,t} = \frac{1}{n}\sum_{j=1}^{n} 1\!\!1_{\{X^n_{j,t}=x\}}. \quad (1)$$

At each time t, M^n_t is in the finite set $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}^{|\mathcal{X}|}$, and $M^n_{x,t}$ is the fraction of players who belong to population of individual state x. For a subset $\tilde{X} \subseteq \mathcal{X}$, define $M^n_t(\tilde{X}) := \frac{1}{n}\sum_{j=1}^n 1\!\!1_{\{X^n_{j,t}\in\tilde{X}\}}$.

Strategies and random set of interacting players: At time slot t, an ordered list $\mathcal{B}^n_t$, of players in $\{1, 2, \ldots, n\}$, without repetition, is selected randomly as follows. First we draw a random number of players k_t such that $\mathbb{P}(|\mathcal{B}^n_t| = k \mid M^n_t = m) =: J^n_k(m)$ where the distribution $J^n_k(m)$ is given for any n, $m \in \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}^{|\mathcal{X}|}$. Second, we set $\mathcal{B}^n_t$ to an ordered list of k_t players drawn uniformly at random among the $n(n-1)...(n-k_t+1)$ possible ones.

Each player such that $j \in \mathcal{B}^n_t$ takes part in a one-shot interaction at time t, as follows. First, each selected player $j \in \mathcal{B}^n_t$ has opportunity to revise its action $a_{j,t} \in \mathcal{A}(s)$.

Denoting the current set of interacting players $\mathcal{B}^n_t = \{j_1, \ldots, j_k\}$. Given the actions $a_{j_1}, ..., a_{j_k}$ drawn by the k players, we draw a new set of individual states $(x'_{j_1}, ..., x'_{j_k})$ and resource state s' with probability $L^n_{s;s'}(k,m)$, where a is the vector of the selected actions by the interacting players.

We assume that the transition kernel L^n is invariant by any permutation of the index of the players within the same type. This implies in particular that the players are only distinguishable through their individual state. Moreover, the process M^n_t is Markovian under Markovian strategies. Denote by $\bar{w}^n_{s,s'}(m)$ be the marginal transition probability between the resource states. Given any vector m of $\Delta(\mathcal{X})$, the resource state generates an irreducible Markov process with limiting invariant measure $w_s(m)$. Then, we can simplify the analysis by fixing the resource state $S_t = s$ without losing generality.

III. Mean field convergence of population games

In this section, we present two main mean field convergence results. We provide a general convergence result of the mean field to a stochastic differential equation. Let $\mathcal{F}^n_t = \sigma(A^n_{t'},\ t' \leq t)$ be the filtration generated by the sequence of states and actions up to t. The evolution of the system depends on the decision of the interacting players. Given a history $h_t = (S_0, A^n_0, \ldots, S_t = s, A^n_t)$. X^n_{t+1} evolves according to the transition probability

$$L^n(x'; x, s) = \mathbb{P}\left(X^n_{t+1} = x' \mid h_t\right)$$

The term $L^n(x'; x, s)$ is the transition kernel on $\mathcal{X}^n$. Let $x^n = (x^n_1, \ldots, x^n_n)$ such that $\frac{1}{n}\sum_{j=1}^n \delta_{x^n_j} = m$ and define

$$\mathcal{L}^n(m'; m, s) = \sum_{\substack{(x'_1,\ldots,x'_n) \\ \frac{1}{n}\sum_{j=1}^n \delta_{x'_j} = m'}} L^n(x'; x, s).$$

The system evolves according to the kernel

$$\begin{aligned} &\mathcal{L}^n(m'; m, s) \\ &:= \mathbb{P}(M^n_{t+1} = m' \mid M^n_t = m, S_t = s) \\ &= \mathbb{P}(M^n_{t+1} = m' \mid \tilde{h}_t) \end{aligned}$$

where $\tilde{h}_t = (S_{t'}, A^n_{t'},\ t' \leq t, S_t = s, X^n_t = x^n)$, such that $\frac{1}{n}\sum_{j=1}^n \delta_{x^n_j} = m$. The term $\mathcal{L}^n(m'; m, s)$ corresponds to the projected kernel of L^n.

A. Deterministic mean field limit

Theorem 1: Let $\mathcal{M}^d_n = \{m \mid nm \in \mathbb{N}^d\}$. Suppose that

D0: For every s, the function $w_s(m)$ is continuously differentiable in m.

D1: $\exists\ 0 < \delta_n, \epsilon_n \searrow 0$, and a continuously differentiable function $f:\ \mathbb{R}^d \times \mathcal{S} \longrightarrow \mathbb{R}^d$ such that

$$\lim_n \sup_{\|m\|\leq 1} \| \frac{f^n(m,s)}{\delta_n} - f(m,s) \| = 0,$$

where $x \in \mathcal{X}$ and $f^n_x(m,s) =$

$$\int_{m' \in \mathcal{M}^d_n} 1\!\!1_{\|m'-m\|\leq 2}(m'_x - m_x)\mathcal{L}^n(dm'; m, s),$$

D2:

$$\sup_n \frac{1}{\delta_n}\int_{m'\in\mathcal{M}^d_n} \| m' - m \| \mathcal{L}^n(dm'; m, s) < +\infty$$

D3: $\lim_n \frac{1}{\delta_n}\int_{m'\in\mathcal{M}^d_n} 1\!\!1_{\|m'-m\|>\epsilon_n} \| m' - m \| \mathcal{L}^n(dm'; m, s) = 0$,

D4: $M^n_0 = m^n_0$ converges to $m_0 \in \Delta(\mathcal{X})$.

Then, for all $\epsilon > 0, T < +\infty$,

$$\lim_n \mathbb{P}\left(\sup_{t\in[0,T]} \| M^n_{\frac{t}{\delta_n}} - m_t[m_0] \| > \epsilon\right) = 0,$$

where $m_t[m_0]$ is the unique solution of the ordinary differential equation $\dot{m}_t = \tilde{f}(m_t)$ starting from $m_0 \in \Delta(\mathcal{X})$ where $\tilde{f}(m_t) := \sum_{s\in\mathcal{S}} w_s(m_t) f(m_t, s)$.

Proof: We check that the assumptions D0-D4 are sufficient conditions for the application of Kurzt (1970) which gives the announced result. ∎

How these assumptions can be checked? The assumption D1 demands that as n grows large, the expected changes per time unit $\frac{f^n}{\delta_n}$ converge uniformly to a Lipschitz continuous vector field f. Lipschitz continuity of f, w ensure the existence and uniqueness of solutions of the mean field game dynamics $\dot{m}_t = \tilde{f}(m_t), m(0) = m_0$ The assumption D2 requires that the expected absolute changes per time unit is bounded. The assumption D3 demands that jumps larger than ϵ_n make vanishing contributions to the motion of the processes, where ϵ_n is a sequence that converges to zero. D0 and D4 are respectively regularity assumptions and initialization conditions.

Consequently, under the vanishing scaling assumptions δ_n, ϵ_n and the hypothesis D0-D4, one has a deterministic approximation of the random process M^n and the deterministic trajectory is described by the ODE.

As we can see some of the assumptions in the above theorem may not be satisfied in wide range of applications in large populations. The assumptions are satisfied when the second moment of number of players that change their individual states in one time slot are bounded in expectation. However, when there are simultaneous and many local interactions as large population games, the second moment may not be finite when the size of the population goes to infinity. Then, a natural question is to ask is: what will happens if the second moment condition is not satisfied?

In the next section, we will partially answer to this question by proposing a mean field convergence to stochastic differential equation called *noisy mean field limit*.

B. Stochastic mean field limit

Below we provide sufficient conditions on the transition kernels $\mathcal{L}^n$ and time-scaling δ_n to get a weak convergence of the process M_t^n.

C0: For every $s \in \mathcal{S}$, $w_s(m)$ is continuously differentiable in m.

C1: There exists $\delta_n \searrow 0$ and continuous mapping $a:\ \mathbb{R}^d \times \mathcal{S} \longrightarrow \mathbb{R}^{d\times d}$ such that $(x, x', s) \in \mathcal{X}^2 \times \mathcal{S}$,

$$\lim_n \sup_{\|m\|\leq 1} \parallel \frac{a^n(m,s)}{\delta_n} - a(m,s) \parallel = 0,$$

where $(x, x', s) \in \mathcal{X}^2 \times \mathcal{S}$, $a^n_{x,x'}(m,s) =$

$$\int_{m'} \mathbb{1}_{\|m'-m\|\leq 2}(m'_x - m_x)(m'_{x'} - m_{x'})\mathcal{L}^n(dm'; m, s),$$

and the third moment is finite. Denote by

$$\tilde{a}_{x,x'}(m) = \sum_{s\in\mathcal{S}} w_s(m) a_{x,x'}(m,s)$$

C2: There exists a continuous mapping $f:\ \mathbb{R}^d \times \mathcal{S} \longrightarrow \mathbb{R}^d$ such that $\forall s \in \mathcal{S}$,

$$\lim_n \sup_{\|m\|\leq 1} \parallel \frac{f^n(m,s)}{\delta_n} - f(m,s) \parallel = 0,$$

C3: For all $\epsilon > 0$; $\forall s \in \mathcal{S}$,

$$\lim_n \frac{1}{\delta_n} \int_{m' \in \mathbb{R}^d} \mathbb{1}_{\|m'-m\|>\epsilon} \mathcal{L}^n(dm'; m, s) = 0,$$

C3': $\forall s \in \mathcal{S}$,

$$\sup_{m\in\mathbb{R}^d} \sup_{n\geq 1} \left[\parallel \frac{a^n(m,s)}{\delta_n} \parallel + \parallel \frac{f^n(m,s)}{\delta_n} \parallel \right] < \infty,$$

Theorem 2: Assume $C0 - C3$. Then, for any test function ϕ, generator $\frac{1}{\delta_n}\mathcal{L}^n\phi(m,s) \longrightarrow \mathcal{L}\phi(m)$ for any m where $\mathcal{L}\phi(m) = \sum_x \tilde{f}_x(m)\frac{\partial}{\partial m_x}\phi(m) + \frac{1}{2}\sum_{x,x'} \tilde{a}_{x,x'}(m)\frac{\partial^2}{\partial m_x \partial m_{x'}}\phi(m)$.

Moreover, if the function $\tilde{a}(.)$ and $\tilde{f}(.)$ have the property that for each $m \in \mathbb{R}^d$, the martingale problem for a and f has exactly one solution π_m starting from m. Then, $\pi_{n,m} \longrightarrow \pi_m$ as $\delta_n \searrow 0$ uniformly in m where $\pi_{n,m}$ is the law of interpolated process from M_t^n. In addition, if $C3'$ holds then the martingale problem has a unique solution.

A detailed proof of the Theorem is given in [6]. See also [7] for a generalization. This result provides a mean field convergence to a solution of stochastic differential equation with drift f and diffusion term a which is reported in the following corollary:

Corollary 1: Suppose that $M_0^n \longrightarrow \mu_0$ in law where μ_0 is a probability measure. Under B0-B3', the process M_t^n converges in law to a solution of the stochastic differential equation (SDE) given by $d\tilde{m}_t = \tilde{f}(\tilde{m}_t)dt + \tilde{\sigma}(\tilde{m}_t)d\mathbb{B}_t$ where $\tilde{\sigma}\tilde{\sigma}^t = \tilde{a}$, and $\mathbb{B}$ is a standard Brownian motion (a Wiener process).

Proof: This result follows from Theorem 2 and the convergence of $\frac{1}{\delta_n}\mathcal{L}^n\phi(m,s) \longrightarrow \mathcal{L}\phi(m,s)$ using the tightness properties of the processes. ■

C. Derivation of evolutionary game dynamics

The first theorem provides a deterministic *mean field game dynamics* for multiple-type population games. The theorem **??** is a generalization of the deterministic *evolutionary game dynamics* based on revision protocols under the form

$$\dot{m}_{x,t} = \sum_{x'\in\mathcal{X}} \mathcal{L}_{x'x}(m_t) m_{x',t} - m_{x,t} \sum_{x'\in\mathcal{X}} \mathcal{L}_{xx'}(m_t), \quad (2)$$

which can be obtained from the drift limit f for $\mathcal{B}_t^n = \{j_l\}$. Theorem 2 derives *stochastic evolutionary game dynamics from mean field game dynamics*.

Lemma III-C.1: Let $\mathcal{L}_{xx'}(m_t) = m_{x',t}\max\{0, u_{x'}(m_t) - u_x(m_t)\}$, Then mean field game dynamics in (2) becomes the replicator dynamics [5], [3] $\dot{m}_{x,t} = m_{x,t}\left[u_x(m_t) - \sum_{x'\in\mathcal{X}} m_{x',t}u_{x'}(m_t)\right]$

Proof:

$$\begin{aligned}\dot{m}_{x,t} &= \sum_{x'\in\mathcal{X}} \mathcal{L}_{x'x}(m_t)m_t(x') - m_{x,t}\sum_{x'\in\mathcal{X}}\mathcal{L}_{xx'}(m_t) \\ &= \sum_{x'\in\mathcal{X}} m_{x,t}\max\{0, u_x(m_t) - u_{x'}(m_t)\}m_{x',t} \\ &\quad -m_{x,t}\sum_{x'\in\mathcal{X}} m_{x',t}\max\{0, u_{x'}(m_t) - u_x(m_t)\}\end{aligned}$$

which can be written as

$$\begin{aligned}\dot{m}_{x,t} &= m_{x,t}\sum_{x'\in\mathcal{X}} m_{x',t}\,(\max\{0,u_x(m_t)-u_{x'}(m_t)\} \\ &\quad -\max\{0,u_{x'}(m_t)-u_x(m_t)\}) \\ &= m_{x,t}\sum_{x'\in\mathcal{X}} m_{x',t}\,(u_x(m_t)-u_{x'}(m_t)) \\ &= m_{x,t}\left(u_x(m_t)[\sum_{x'\in\mathcal{X}} m_{x',t}]-\sum_{x'\in\mathcal{X}} m_{x',t}u_{x'}(m_t)\right) \\ &= m_{x,t}\left[u_x(m_t)-\sum_{x'\in\mathcal{X}} m_{x',t}u_{x'}(m_t)\right]\end{aligned}$$

where we have used the fact that $\sum_{x'\in\mathcal{X}} m_{x',t}=1$, and $\max\{0,u_x(m_t)-u_{x'}(m_t)\}-\max\{0,u_{x'}(m_t)-u_x(m_t)\}=u_x(m_t)-u_{x'}(m_t)$. ■

D. Equilibrium state

At limiting in population size, we associate to each generic player with state x a payoff function $u_x:\ \mathbb{R}^d\longrightarrow\mathbb{R}$. Denote by $u(.)=(u_x(.))_{x\in\mathcal{X}}$. The pair $(\mathcal{X},u)$ defines a population game and the mean field limit can be seen as the population state or population profile. We say that a population profile m^* is an equilibrium if for all $x\in\mathcal{X}$, one has

$$m^*_x>0\Longrightarrow u_x(m^*)=\max_{x'\in\mathcal{X}}\ u_{x'}(m^*).$$

This definition is equivalent to the variational inequality problem: find m^* such that $\langle u(m^*),m^*-m\rangle\geq 0,\ \forall m\in\Delta(\mathcal{X})$. Hence the following result follows

Theorem 3: Assume that u is continuous. Then, the population game has at least one equilibrium.

For the proof, we use existence of solution of the variational inequality problem under continuity argument. Another alternative is to write it in form of a Brouwer fixed point map.

We say that the population game with payoffs $u(.)$ is a full potential population game if there is a continuously differentiable mapping $\mathcal{V}_p:\ \mathbb{R}^d\longrightarrow\mathbb{R}$ such that

$$\frac{\partial}{\partial m_x}\mathcal{V}_p(m)=u_x(m).$$

The mean field game dynamics generated by the rate of transition $\mathcal{L}$ is positively correlated if $\tilde{f}(m)\neq 0\Longrightarrow\langle\tilde{f}(m),u(m)\rangle>0$.

Proposition 1 (Convergence): Global convergence holds in potential population games under positive correlation.

The proof of this result follows from the fact that the potential function provides also a Lyapunov function.

Now, we present a convergence result for population games with monotone payoffs (also called stable games). We say that the population game has monotone payoffs[1] if $\langle m-m',u(m)-u(m')\rangle\leq 0,\ \forall m,m'$. This class of population games includes zero-sum games, potential concave games, population games with two pure strategies etc.

[1]Notice that, an operator T is said monotone if $\langle Tm-Tm',m-m'\rangle\geq 0$. Here, we consider the opposite inequality.

Consider the class of rate transitions: $\tilde{\beta}_{x'x}(m,r)=\xi_x[\max(u_x(m)-u_{x'}(m),0)]$ where $\xi:\ \mathbb{R}\longrightarrow\mathbb{R}_+$ is a non-decreasing function.

Proposition 2 (convergence): Under the mean field game dynamics generated by the transition $\tilde{\beta}$, global convergence holds in any population games with monotone payoffs.

Proof: A proof can be obtained from [8]. ■

It is known that specific payoff structures such as the structure of potential, stable, and supermodular games, makes evolutionary justifications of the equilibrium prediction. However, once we move beyond these particular classes of population games, it is not clear how often convergence will occur. In addition, most of the games do not have these specific properties. We provide examples (fig. 1) that counterbalance the convergence approach by investigating non-convergence of mean field game dynamics, describing situations in which cycling offer the best predictions of long run behavior.

IV. Mean field game dynamics with hybrid learning schemes

We now study how to combine various learning schemes based on mean field game dynamics. Different learning and adaptive algorithms have been studied in the literature. In most of the analysis, the players have to follow the same rule of learning, they have to learn in the same way. We now ask the following question:

What happens if the players have different learning schemes?

We propose learning schemes in which player use less information about the other players, less memory on the history and do not need to use the same learning scheme. Below we list some mean field game dynamics that we borrow from evolutionary game dynamics [4]. They are in the form

$$\dot{m}^\theta_{a,t}=\sum_{\bar{a}} m^\theta_{\bar{a},t}\rho^\theta_{\bar{a},a}(m_t)-m^\theta_{a,t}\sum_{\bar{a}}\rho^\theta_{a,\bar{a}}(m_t)$$

where θ denotes a type and ρ is the migration rate between actions i.e $\rho^\theta_{a,\bar{a}}(m_t)=\mathcal{L}_{(\theta,a),(\theta,\bar{a})}(m_t)$.

- *excess payoff dynamics:* Brown-von Neumann-Nash dynamics is one of well-known excess payoff dynamics. The *Brown-von Neumann-Nash dynamics* is obtained for $\rho^{1,\theta}_{a',a}(m_t)=\max(0,u^\theta_a(m_t)-\sum_{\bar{a}} m^\theta_{\bar{a},t}u^\theta_{\bar{a}}(m_t))$. The set of the rest points of the BNN dynamics is exactly the set of (Nash) equilibria.
- *Imitation of neighbors:* the *replicator dynamics* is obtained for $\rho^{2,\theta}_{a',a}(m_t)=m^\theta_{a,t}\max(0,u^\theta_a(m_t)-u^\theta_{a'}(m_t))$. The set of the rest points of the replicator dynamics contains the set of equilibrium states but it can be must bigger since it is known that the replicator dynamics may not lead to equilibria. Typically, the corners are rest points and the faces of the simplex are invariant but may not be an equilibrium.
- *Boltzmann-Gibbs dynamics* also called *smooth best response* dynamics or logit dynamics is obtained for $\rho^{4,\theta}_{a',a}(m_t)=\frac{e^{\frac{u^\theta_a(m_t)}{\epsilon}}}{\sum_{\bar{a}} e^{\frac{u^\theta_{\bar{a}}(m_t)}{\epsilon}}}$, $\epsilon>0$. Using the logit map, it can

be shown that the time average of the replicator dynamics is a perturbed solution of the best reply dynamics. The rest points of the Smooth dynamics are approximated equilibria.

• *Pairwise comparison dynamics:* for example the generalized Smith dynamics is obtained for $\rho^{5,\theta}_{a',a} = \max(0, u^{\theta}_a(m_t) - u^{\theta}_{a'}(m_t))^{\gamma},\ \gamma \geq 1$. The set of rest points of the generalized Smith dynamics is exactly the set of equilibria. In [8] it is proved that under this class of dynamics, global convergence holds in stable games and in potential games etc. Extension to evolutionary game dynamics with *migration between location of players* and application to hybrid power control in wireless communications can be found in [8].

• *Best response dynamics* is obtained for ρ^6 equal to the best reply to m_t. The set of rest points of the best response dynamics is exactly the set of equilibria. More details on best response dynamics can be found in [1].

• *ray-projection dynamics* is a myopic adaptive dynamic in which a subpopulation grows when its expected payoff is less than the ray-projection payoff of all the other classes. It is obtained $\rho^{3,\theta}_{a',a}(m_t) = \Lambda^{\theta}_a(m_t)$ where Λ is a continuous map. Notice that the replicator dynamics, best response dynamics and logit dynamics can be obtained as a particular case of the ray-projection dynamics.

A. Hybrid mean field game dynamics

Consider a population in which the players can adopt different learning schemes in $\{\rho^1, \rho^2, \rho^3, \ldots\}$ (finite, say with size κ). Then, based on the composition of population and the use of each learning scheme we build an hybrid mean field game dynamics. The incoming and the outgoing flow are expressed in term of the weighted combination of different learning schemes picked from the set $\{\rho^1, \rho^2, \rho^3, \ldots\}$ with probability $\{\tilde{\lambda}^j\}_j$. Then, the hybrid learning is obtained for the rule $\sum_j \tilde{\lambda}^j \rho^j$.

Define the property *weighted equilibrium stationarity* (WES) as follows:

(WES) *Every rest point of the hybrid game dynamics generated by the weighted payoff is a weighted equililibrium and every constrained weighted equilibrium is a rest point of the dynamics.*

Note that this property is not satisfied by the well-known replicator dynamics as it is known that the replicator dynamics may not lead to (Nash) equilibria (see the figure 1). We have the following result:

Theorem 4: Suppose that all the dynamics in the support of $\tilde{\lambda}$ satisfy the positive correlation. Then (i) the resulting hybrid mean field game dynamics satisfy also the positive correlation, (ii) The Nash equilibria are rest points of the hybrid dynamics.

Proof: Denote by $\tilde{f}^j$ the drift limit generated by the learning scheme ρ^j. Then, the hybrid mean field game dynamics is given by $\dot{m}_t = \sum_{j=1}^{\kappa} \tilde{\lambda}^j \tilde{f}^j(m_t) =: \tilde{f}(m_t)$. Let $\tilde{f}(m) \neq 0$. This means that there exists at least one index k for which $\tilde{f}^k(m) \neq 0$. Since $\tilde{f}^k$ is positively correlated $\langle \tilde{f}^k(m), u(m) \rangle > 0$ and the others are positive or zero $\langle \tilde{f}^{j'}(m), u(m) \rangle \geq 0$. Combining together, $\tilde{f}(m) \neq 0 \Longrightarrow \langle \sum_j \tilde{\lambda}^j \tilde{f}^j(m), u(m) \rangle > 0$. Under positive correlation the Nash equilibria are rest points. ■

Theorem 5: Let $\tilde{\lambda}^j$ the proportion of players that adopt the learning scheme ρ^j. If all the learning schemes contained in the support of $\tilde{\lambda} = (\tilde{\lambda}^1, \ldots, \tilde{\lambda}^{\kappa}) \in \mathbb{R}^{\kappa}_+$ satisfy the property (WES) then, the resulting hybrid mean field game dynamics generated by these learning schemes satisfies also the weighted equilibrium stationarity property.

Proof: Let m be an equilibrium state. Then, from the property (WES), m is a rest point of all the dynamics in the support of $\tilde{\lambda}$. We need that prove that any rest point of the combined dynamics in the support of $\tilde{\lambda}$ is an equilibrium state.

Suppose that it is not the case. Then, there exists at least one j such that m is not rest point of the dynamics generated by the learning scheme ρ^j. But the mean field game dynamics of ρ^j satisfies (WES). This means that m is not an equilibrium state which is a contradiction. We conclude any rest point of hybrid mean field game dynamics is an equilibrium state. This completes the proof. ■

B. How to eliminate the rest points which are not equilibria?

Consider that the family of learning schemes generated by $\rho^{\gamma,\theta}_{a',a} = \max(0, u^{\theta}_a(m_t)) - u^{\theta}_{a'}(m_t))^{\gamma},\ \gamma \geq 1$. It is easy to see that this family satisfies the property (WES). We deduce that if the population is constituted of 99% of players use a learning scheme via ρ^{γ} and 1% of the population use a replicator-based learning scheme then the resulting hybrid dynamics satisfy the property (WES). We conclude that every rest point of the replicator dynamics which is a non-Nash equilibrium will be eliminated using this new hybrid dynamics. This says that, the players can learn in a bad way but if the fraction of good learners is non-zero then the rest points of resulting hybrid dynamics will be equilibria.

C. Non-commutative diagram

We examine the double limits $\lim_n \lim_t M^n_t[m_0]$ and $\lim_t \lim_n M^n_t[m_0]$. Denote by $\varpi^n[m_0]$ the limiting behavior of $\lim_t M^n_t[m_0]$.

- Convergence/nonconvergence of $m_t[m_0]$ as t goes to infinity?
- Convergence/nonconvergence of $\varpi^n[m_0]$ as n goes to ∞?
- Under which conditions, the two limits coincide (if they exist)?
- If the dynamics do not converge, is there connection between the time average of the orbits of the ordinary differential equation (ODE) $\dot{m} = \tilde{f}(m)$ starting from m_0, and the omega-limit of $\varpi^n[m_0]$?.

Note that the uniqueness of a stationary point of the ODE does not imply convergence to this stationary point As illustrated in figure 1, the double limit need not to be commutative i.e $\lim_n \lim_t M^n_t \neq \lim_t \lim_n M^n_t$. This phenomenon is in part due to the fact that the stationary distribution of the process ω^n is unique under irreducibility conditions but the dynamics can lead to a limit cycle. As a consequence, many techniques based on stationary regime (such as fixed-point equation techniques, limiting of frequencies state-actions approaches in sequence of stochastic games, replica methods, interacting-particle systems, statistical independence in large-scale interaction etc) need some justification. The non-commutativity phenomenon suggests to be careful about the use of stationary population state equilibria as the outcome prediction and the analysis of equilibrium payoffs since this equilibrium may not be played. Limit cycles are sometimes more appropriate than the stationary equilibrium approach.

In the next subsection we illustrate the non-commutativity of the diagram by simulating the process M^n.

D. Simulation replicator+smith dynamics

We consider the rock-scissor-paper (RSP) game in large population with size 8000. The individual state of a player can be in the $\{r, p, s\}$ i.e the actions. The transition probabilities between the states are given the system state and the payoff functions which depend on the other players. Thus, the transition of an individual depends on the state of the other players. The rock beats scissor which beats paper which beats rock. The payoff of the winner is $+1$ and the payoff of the looser is -1. The transitions are payoff-dependent: $\mathcal{L}_{xx'}(m) = m_{x'} \max(0, u_{x'}(m) - u_x(m))$. The number of players n is fixed to 8000. We consider an hybrid learning. The player can adopt ρ^2 or ρ^5 with the parameter $\gamma = 1$ and with probability $\tilde{\lambda} = (1/2, 1/2)$. The learning scheme ρ^2 leads to replicator dynamics and the learning scheme ρ^5 leads to Smith dynamics. In the top figure 1, we plot the deterministic mean field limit and its time average trajectory. The middle figure 1 is a ternary plot the deterministic mean field limit and its time average trajectory. The bottom figure 1 is a simulation of the mean field process M^{8000} and its time average trajectory. The evolutionary RSP game has a unique equilibrium $m^* = \frac{1}{3}(1, 1, 1)$ which is unstable as illustrated in the three figures.

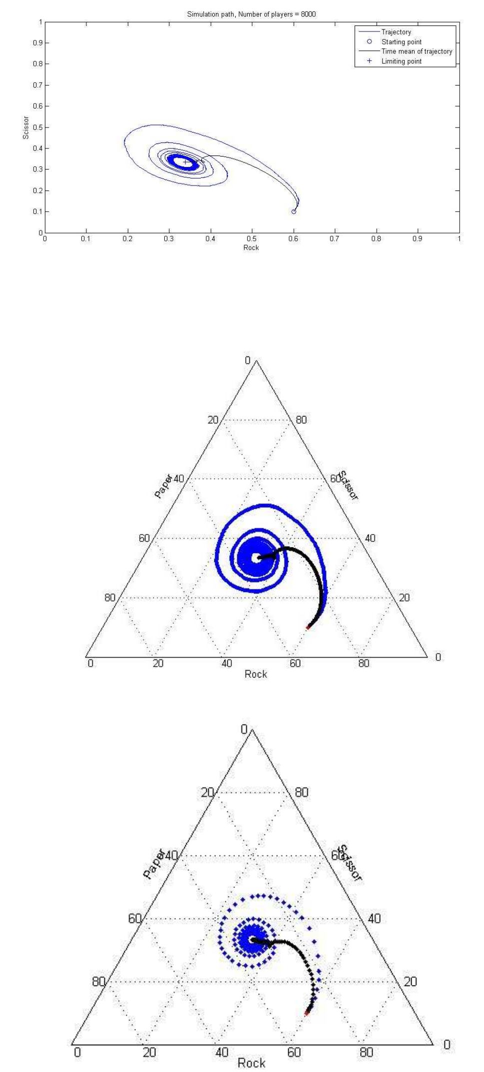

Fig. 1. Top: ODE ternary plot RSP. Middle: ODE ternary plot, Bottom: Mean of simulations of M^{8000}.

V. Conclusion

We have studied large population games under hybrid mean field game dynamics. We have shown both convergence and limit cycling behavior of the dynamics. An interesting direction that we leave for future work is the derivation of the hybrid and heterogeneous dynamics when some populations use heterogeneous learning with diffusion term and some others use deterministic dynamics with different time-scales.

References

[1] I. Gilboa and A. Matsui. Social stability and equilibrium. *Econometrica*, 59:859–867, 1991.
[2] J. Hofbauer and K. Sigmund. Evolutionary games and population dynamics. *Cambridge University Press*, 1998.
[3] J. Maynard-Smith. Evolution and the theory of games. *Cambridge University Press, Cambridge, U.K.*, 1982.
[4] W. H. Sandholm. Population games and evolutionary dynamics. *MIT Press*, 2010.
[5] Taylor and Jonker. Evoltionarily stable strategies and game dynamics. *Mathematical Bioscience*, 40:145–156, 1978.
[6] H. Tembine. Mean field stochastic games: Simulation, dynamics and network applications. *Supelec*, 2010.
[7] H. Tembine. Mean field stochastic games: convergence, q/h learning, optimality. *American Control Conference*, 2011.
[8] H. Tembine, E. Altman, R. ElAzouzi, and W. H. Sandholm. Evolutionary game dynamics with migration for hybrid power control in wireless communications. *47th SIAM/IEEE CDC*, December 2008.
[9] J. Weibull. Evolutionary game theory. *MIT Press*, 1995.