2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Mean Field Difference Games: McKean-Vlasov Dynamics

H. Tembine and M. Huang

***Abstract*— We study a class of mean field stochastic games in discrete time and continuous state space. Each player has its own individual state evolution described by a stochastic difference equation which depends not only on the control of the corresponding player but also on the states of the other players. Considering the specific structure of aggregate drift and diffusion terms, we use classical asymptotic indistinguishability properties to prove a mean field convergence in distribution. The methodology is extended to multiple classes of players, each class satisfying the asymptotic indistinguishability property, and a propagation of chaos result is obtained over the hull trajectory. Finally, we derive combined backward-forward equations that characterize the mean field equilibria for finite horizon problems.**

I. INTRODUCTION

The central issue in mean field decision problems is the development of low complexity solutions so that each player may implement a strategy based on local information in large populations. For models with mean field coupling, recent advances have been made in effectively addressing the high dimensionality issue [7], [9]. In [9], the Nash certainty equivalence (NCE) methodology has been developed for controlled McKean-Vlasov dynamics, where the key idea is to break the large population game into localized optimal control problems via specifying a consistency relationship between the individual strategies and the aggregate population effect. A very appealing feature of the resulting solutions, as asymptotic Nash equilibria, is that each player's strategy depends only on its own state and some aggregate quantities which may be learnable and may be calculated off-line under given population initial conditions. For related works using similar ideas in mean field decision problems, see [12], [24], [23], [2], [22]. Applications to synchronization of coupled oscillators can be found in [26], [25], and economic and finance in [6]. The authors in [4], [11] applied mean field games to crowd and pedestrian dynamics. Numerical methods for solving backward-forward equations for specific state dynamics and payoff functions can be found in [1].

In this paper, we study a discrete time mean field stochastic game with multiple classes of players that are coupled via their individual dynamics and their payoff functions. By mean field approximation and normalization of time, we derive the limiting continuous time equations. A related method is developed in [17], [21], [18] for discrete time mean field Markov decision problems.

This work was started when the first author was visiting McGill University and Ecole Polytechnique de Montréal, Canada in August 2010. We would like to thank Prof. Peter Caines and Prof. Roland Malhamé for their comments on an earlier version of this paper.

H. Tembine is with Ecole Supérieure d'Electricité, Supelec, France. tembine@ieee.org.

M. Huang is with School of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada. mhuang@math.carleton.ca.

Our main contributions can be summarized as follows. We prove a mean field convergence of difference games under a specific averaging structure. We characterize the mean field limit of the empirical measures as a solution of Fokker-Planck-Kolmogorov equations. The limiting individual state evolution is characterized by McKean-Vlasov equation. Then, we formulate the limiting game as *differential population game* in which each player optimizes its expected long-term payoff under individual stochastic dynamics and mean field limit dynamics.

The mean field solutions are obtained by identifying a consistency relationship between the individual-mass interaction such that in the population limit each individual optimally responds to the mass effect and these individual strategies also collectively produce the same mass effect presumed initially. This leads to a coupled system of *forward-backward equations*.

The rest of the paper is organized as follows. Section II presents the model description. In Section III we study the convergence to mean field limit. Section IV focuses on mean field equilibria and Section V concludes the paper.

We summarize some of the notations in Table I.

TABLE I
SUMMARY OF NOTATIONS

Symbol	Meaning
Θ	the set of classes
θ_j	class of player j
f	drift function (finite dimension)
σ	diffusion coefficient
$x_j^n(k)$	state of player j at time t (discrete)
$\tilde{x}_j^n(k)$	re-scaled state of player j
$\tilde{x}_j(t)$	limit of the state process of player j
$\bar{x}_j(t)$	solution of the McKean-Vlasov equation
$z_j^n(t)$	interpolated re-scaled state process of player j at time t or continuous time solution of SDE
$z_j(t)$	limit of state process $z_j^n(t)$
$F^n(t,w)$	cumulative distribution function at time t

II. MODEL DESCRIPTION: DIFFERENCE GAMES

Consider a population of players with size $n \geq 2$. The discrete time is indexed by the set

$$\mathbb{T}_n = \delta_n \mathbb{Z}_+ = \{0, \delta_n, 2\delta_n, 3\delta_n, \ldots\},$$

where $\delta_n > 0$. Let $\Theta = \{1, 2, \ldots, K\}$ be the set of classes, and $1 \leq K < \infty$ is the total number of classes. The class

978-1-61284-799-3/11/$26.00 ©2011 IEEE

of player j is denoted by $\theta_j \in \Theta$. Each player j has its own state $x_j^n(t) \in \mathcal{X} = \mathbb{R}$ at time t. We define the population profile at time t as the random measure

$$M^n(t) = \sum_{j=1}^{n} \bar{\omega}_j^n \delta_{x_j^n(t)}, \tag{1}$$

where $\bar{\omega}_j^n > 0$ is the individual weight of player j in a population of size n. The population profile process is given by $M^n = \sum_{j=1}^n \bar{\omega}_j^n \delta_{x_j^n}$ where $x_j^n = (x_j^n(t))_{t\in\mathbb{T}_n}$. We define the cumulative distribution function

$$F^n(t,w) = \sum_{j=1}^{n} \bar{\omega}_j^n \mathbb{1}_{\{x_j^n(t)\leq w\}}.$$

We normalize the weight such that $\sum_{j=1}^n \bar{\omega}_j^n = 1$. Then, $M^n(t)$ is a probability measure over $\mathcal{X}$ for each fixed t, and M^n is probability measure-valued process. We restrict our attention to the case where the weight is uniform $\bar{\omega}_{ij}^n = \frac{1}{n}$. The case of different weights and combination of major and minor players can be found in [8], [22], [17] for linear-quadratic Gaussian (LQG) mean field games. See also [3] for more general stochastic differential games with major and minor players. In this paper, the individual dynamics have the following form:

$$\begin{aligned} x_j^n(k+1) &= x_j^n(k) + \Phi_{\theta_j}^n(k, x_j^n(k), u_j^n(k), x_1^n(k)\ldots, x_n^n(k)) \\ &\quad + \Psi_{\theta_j}^n(k, x_j^n(k), u_j^n(k), x_1^n(k), \ldots, x_n^n(k)) \\ &\quad \times \left(\mathbb{B}_j^n(k+1) - \mathbb{B}_j^n(k)\right), \end{aligned} \tag{2}$$

$$x_j^n(0) = x_j,$$
$$j \in \{1,2,\ldots,n\},$$

where $u_j^n(t) \in \mathcal{U}_{\theta_j}$ is the control of player j at time t. $\Phi_{\theta_j}^n(t,.)$, $\Psi_{\theta_j}^n(t,.)$ are measurable and uniformly Lipschitz continuous functions, $\mathbb{B}_j^n(t)$ are mutually independent Brownian motions (Wiener processes) defined over a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

For the clarity of the presentation, we take an averaging structure in (2),

$$\Phi_{\theta_j}^n(k,.) = \delta_n \sum_{i=1}^{n} \omega_{ij}^n f_{\theta_j}(x_j^n(k), u_j^n(k), x_i^n(k)),$$

$$\Psi_{\theta_j}^n(k,.) = \sum_{i=1}^{n} \omega_{ij}^n \sigma_{\theta_j}(x_j^n(k), u_j^n(k), x_i^n(k)),$$

where $\sigma_{\theta_j}(.) \geq \sigma_* > 0$ and $\sigma_{\theta_j}(.)$ is bounded, differentiable and Lipschitz with Lipschitz constant $L_\sigma \leq c$, and $\omega_{ij}^n > 0$ is a strictly positive weight (the relative weighted of i to the player j). Then the individual dynamics of player j are given by

$$\begin{aligned} x_j^n(k+1) &= x_j^n(k) + \delta_n \sum_{i=1}^{n} \omega_{ij}^n f_{\theta_j}(x_j^n(k), u_j^n(k), x_i^n(k)) \\ &+ \sum_{i=1}^{n} \omega_{ij}^n \sigma_{\theta_j}(x_j^n(k), u_j^n(k), x_i^n(k)) \left(\mathbb{B}_j^n(k+1) - \mathbb{B}_j^n(k)\right), \end{aligned} \tag{3}$$

$$x_j^n(0) = x_j,$$

where the drift $f_{\theta_j}(x,u,w)$ and the noise coefficient $\sigma_{\theta_j}(x,u,w)$ are defined on $\mathcal{X} \times \mathcal{U}_{\theta_j} \times \mathcal{X}$. We re-scale the processes with step size δ_n which plays the role of intensity of interaction. For $k \in \mathbb{N}, \tilde{x}_j^n(k) := x_j^n(k\delta_n), \tilde{u}_j^n(k) := u_j^n(k\delta_n), \tilde{M}^n(k) := M^n(k\delta_n), t_k^n := k\delta_n$. The functions F^n and $\tilde{F}^n$ are defined similarly. For $k\delta_n \leq t' < (k+1)\delta_n$, define the interpolated process in continuous time as

$$\begin{aligned} z_j^n(t') &= \tilde{x}_j^n(k) + \frac{(t'-k\delta_n)}{\delta_n}(\tilde{x}_j^n(k+1) - \tilde{x}_j^n(k)) \\ &= x_j^n(k\delta_n) + \frac{(t'-k\delta_n)}{\delta_n}(x_j^n((k+1)\delta_n) - x_j^n(k\delta_n)). \end{aligned}$$

Note that $z_j^n(t_k^n) = x_j^n(t_k^n) = \tilde{x}_j^n(k)$.

Lemma 1: Under the above formulation, the individual dynamics can be written in the following form:

$$\begin{aligned} \tilde{x}_j^n(k+1) &= \tilde{x}_j^n(k) + \delta_n \int_w f_{\theta_j}(\tilde{x}_j^n(k), \tilde{u}_j^n(k), w) \tilde{M}_{j,k}^n(dw) \\ &\quad + \int_w \sigma_{\theta_j}(\tilde{x}_j^n(k), \tilde{u}_j^n(k), w) \tilde{M}_{j,k}^n(dw) \\ &\quad \times \left(\mathbb{B}_j^n(k+1) - \mathbb{B}_j^n(k)\right) \end{aligned}$$

$$\tilde{x}_j^n(0) = \tilde{x}_j,$$

where

$$\tilde{M}_{j,k}^n = \sum_{i=1}^{n} \omega_{ij}^n \delta_{\tilde{x}_i^n(k)}.$$

Proof: Consider the stochastic difference equation:

$$\left\{ \begin{array}{c} x_j^n(t_{k+1}^n) = x_j^n(t_k^n) + \delta_n \sum_{i=1}^n \bar{\omega}_{ij}^n f_{\theta_j}(x_j^n(t_k^n), u_j^n(t_k^n), x_i^n(t_k^n)) \\ + \sum_{i=1}^n \bar{\omega}_{ij}^n \sigma_{\theta_j}(x_j^n(t_k^n), u_j^n(t_k^n), x_i^n(t_k^n)) \left(\mathbb{B}_j^n(t_{k+1}^n) - \mathbb{B}_j^n(t_k^n)\right) \\ x_j^n(0) = x_j, \end{array} \right.$$

Using the above re-scaling process, one gets

$$\left\{ \begin{array}{c} \tilde{x}_j^n(k+1) = \tilde{x}_j^n(k) + \delta_n \sum_{i=1}^n \bar{\omega}_{ij}^n f_{\theta_j}(\tilde{x}_j^n(k), \tilde{u}_j^n(k), \tilde{x}_i^n(k)) \\ + \sum_{i=1}^n \bar{\omega}_{ij}^n \sigma_{\theta_j}(\tilde{x}_j^n(k), \tilde{u}_j^n(k), \tilde{x}_i^n(k)) \left(\mathbb{B}_j^n(k+1) - \mathbb{B}_j^n(k)\right) \\ \tilde{x}_j^n(0) = x_j, \end{array} \right.$$

Now, we use the fact that

$$\int_{w\in\mathcal{X}} \phi(w) M_{t_k^n}^n(dw) = \sum_{i=1}^{n} \bar{\omega}_i^n \phi(x_i^n(t_k^n)),$$

which is equivalent to

$$\int_{w\in\mathcal{X}} \phi(w) \ \tilde{M}_k^n(dw) = \sum_{i=1}^{n} \bar{\omega}_i^n \phi(\tilde{x}_i^n(k)),$$

the above system of stochastic difference equation can be written as

$$\left\{ \begin{array}{c} \tilde{x}_j^n(k+1) = \tilde{x}_j^n(k) + \delta_n \int_w f_{\theta_j}(\tilde{x}_j^n(k), \tilde{u}_j^n(k), w) \times \\ \left[\sum_{i=1}^n \bar{\omega}_{ij}^n \delta_{\tilde{x}_i^n(k)}\right](dw) \\ + \int_w \sigma_{\theta_j}(\tilde{x}_j^n(k), \tilde{u}_j^n(k), w) \times \\ \left[\sum_{i=1}^n \bar{\omega}_{ij}^n \delta_{\tilde{x}_i^n(k)}\right](dw) \left(\mathbb{B}_j^n(k+1) - \mathbb{B}_j^n(k)\right) \\ \tilde{x}_j^n(0) = x_j, \end{array} \right.$$

Denote

$$\tilde{M}_{j,k}^n = \sum_{i=1}^{n} \omega_{ij}^n \delta_{\tilde{x}_i^n(k)}.$$

Then

$$\left\{\begin{array}{c}\tilde{x}_j^n(k+1) = \tilde{x}_j^n(k) + \delta_n \int_w f_{\theta_j}(\tilde{x}_j^n(k), \tilde{u}_j^n(k), w)\tilde{M}_{j,k}^n(dw) \\ + \int_w \sigma_{\theta_j}(\tilde{x}_j^n(k), \tilde{u}_j^n(k), w)\tilde{M}_{j,k}^n(dw)\left(\mathbb{B}_j^n(k+1) - \mathbb{B}_j^n(k)\right) \\ \tilde{x}_j^n(0) = x_j.\end{array}\right.$$

■

III. MEAN FIELD CONVERGENCE: MAIN RESULTS

In this section we focus on the mean field convergence of the process M^n and the characterization of its limit. We will work on a σ-field $\mathcal{F}_t$ generated by the Brownian motion up to time t. We provide three main mean field convergence results.

(R0) We show the existence of the limit

$$\tilde{\mu}_{j,k} := \lim_n \sum_{i=1}^n \bar{\omega}_{ij}^n \delta_{\tilde{x}_i^n(k)},$$

for the weight $\bar{\omega}_{ij}^n = \frac{1}{n}$ for any control that preserves the asymptotic indistinguishability in law.

(R1) By use of a vanishing time step-size, i.e., $\delta_n \to 0$, we derive a continuous time equation when the population size $n \to +\infty$. We characterize the stochastic limit as a solution to the macroscopic stochastic differential McKean-Vlasov equation for which we have sufficient conditions that ensure uniqueness of the trajectory.

$$d\tilde{x}_j(t) = \left[\int_w f_{\theta_j}(\tilde{x}_j(t), \tilde{u}_j(t), w)\tilde{\mu}_{j,t}(dw)\right] dt + \left[\int_w \sigma_{\theta_j}(\tilde{x}_j(t), \tilde{u}_j(t), w)\tilde{\mu}_{j,t}(dw)\right] d\mathbb{B}_j(t),$$
$$\tilde{x}_j(0) = x_j.$$

(R3) For a non-vanishing step-size satisfying $\delta_n \to \delta > 0$, we characterize the limit as a discrete version of the McKean-Vlasov equation. We show that the dynamics given in Lemma 1 converge to

$$\tilde{x}_j(k+1) = \tilde{x}_j(k) + \delta \int_w f_{\theta_j}(\tilde{x}_j(k), \tilde{u}_j(k), w)\tilde{\mu}_{j,k}(dw) + \int_w \sigma_{\theta_j}(\tilde{x}_j(k), \tilde{u}_j(k), w)\tilde{\mu}_{j,k}(dw)\left(\mathbb{B}_j(k+1) - \mathbb{B}_j(k)\right)$$
$$\tilde{x}_j(0) = x_j.$$

When choosing $\bar{\omega}_{ij}^n = \frac{1}{n}$, the measure $\{\tilde{\mu}_{j,t}\}_j$ has the same law as represented by $(\tilde{\mu}_{\bar{\theta},t})_{\bar{\theta}\in\Theta}$ which is a solution of the Fokker-Planck-Kolmogorov equation

$$\frac{\partial}{\partial t}\bar{\mu}_{\bar{\theta},t}(\bar{x}) + \frac{\partial}{\partial \bar{x}}\left[\bar{f}_{\bar{\theta},t}(\bar{x}, \bar{u}, \bar{\mu}_t)\bar{\mu}_{\bar{\theta},t}\right] = \frac{1}{2}\frac{\partial^2}{\partial \bar{x}^2}\left[\bar{\sigma}^2_{\bar{\theta},t}(\bar{x}, \bar{u}, \bar{\mu})\bar{\mu}_{\bar{\theta},t}\right],$$
$$\bar{\theta} \in \Theta,\ \mu_0(\bar{x}) \ \text{ fixed},$$

where

$$\bar{f}_{\bar{\theta},t}(\bar{x}, \bar{u}, \bar{\mu}) = \int_w f_{\bar{\theta}}(\bar{x}, \bar{u}, w)\ \bar{\mu}_t(dw), \tag{4}$$

$$\bar{\sigma}_{\bar{\theta},t}(\bar{x}, \bar{u}, \bar{\mu}) = \int_w \sigma_{\bar{\theta}}(\bar{x}, \bar{u}, w)\ \bar{\mu}_t(dw). \tag{5}$$

Cost functional criterion: Following the same lines, the weak convergence of the total payoff function can be established: $L_{j,T}^n(u_j, u_{-j}, M^n, x_0) =$

$$\mathbb{E}\Big(g(M_T^n) + \sum_{s\in\mathbb{T}_n, s\le T}\sum_{i=1}^n \omega_{ij}^n c_{\theta_j}^n(x_j^n(s), u_j^n(s), x_i^n(s)) \mid x_0\Big)$$

where $c_{\theta_j}^n = \delta_n c_{\theta_j}$ is an integrable function relatively to the measure M_t^n, g is a terminal payoff (regular) and $u_{-j} = (u_{j'})_{j'\neq j}$. We assume that the functions $c_\theta(.)$ are Lipschitz continuous with Lipschitz constant L_c. The convergence comes from the fact $\sum_{i=1}^n \omega_{ij}^n c_{\theta_j}(x_j^n(s), u_j^n(s), x_i^n(s))$ can be written as

$$\int c_{\theta_j}(x_j^n(s), u_j^n(s), w)M_s^n(dw)$$

which converges to

$$\bar{c}_{\theta_j}(x_j(s), u_j(s), \mu_s) := \int_w c_{\theta_j}(x_j(s), u_j(s), w)\mu_s(dw)$$

(by weak convergence of M^n to μ). Assuming A1-A2, we express the finite horizon payoff in terms of the norm $\| \mu_s - M_s^n \|$. Since we work on a compact interval $[0,T]$, we have the convergence of individual payoff function to

$$L_j(u_j, \mu, x_0) = \mathbb{E}\Big(g(\mu_T) + \int_0^T \bar{c}_{\theta_j}(x_j(s), u_j(s), \mu_s)\ ds\ \Big) \tag{6}$$

We prove the result (R0) for specific weights via the asymptotic indistinguishability per class. The proof of the result (R2) is similar to (R1). The proof can be easily extended to the case where $\sigma^n = \sigma + o(\frac{1}{n})$. The details of the proof of (R1) follows from the following theorem:

Theorem 1: Assume

A1. $f(.)$ is a Lipschitz continuous function with Lipschitz constant L_f and $\sigma_{\theta_j}(x,u,w)$ defined on $\mathcal{X}\times\mathcal{U}_{\theta_j}\times\mathcal{X}$ is a Lipschitz continuous function with Lipschitz constant L_f. There exists $\sigma_* > 0$ such that $\sigma(x,u,w) \ge \sigma_* > 0$ for all (x,u,w).

A2. The law of the initial measure $\tilde{\mu}_0(w)$ has a continuous density bounded by $c'e^{-d'|w|^2}$ for some positive constants c' and d'.

Let $\tilde{F}^n$ be defined as in Section II. Then for any $T < +\infty$, there exists $c_T > 0$ such that

$$\mathbb{E} \parallel \tilde{F}^n(t_k^n, .) - \tilde{F}(t_k^n, .) \parallel_1 \le c_T\left[\parallel \tilde{F}_0 - \tilde{F}_0^n \parallel_1 + \frac{1}{\sqrt{n}} + \sqrt{\delta_n}\right],$$

where $\tilde{F}(t,.,.)$ (for each class) is the solution of

$$\frac{\partial}{\partial t}\bar{F}_{\bar{\theta},t}(\bar{x}) + \left[\int_w f_{\bar{\theta},t}(\bar{x}, \bar{u}, w)\frac{\partial}{\partial w}\bar{F}_{\bar{\theta},t}(w)dw\right]\frac{\partial}{\partial \bar{x}}\bar{F}_{\bar{\theta},t}(\bar{x}) = \frac{1}{2}\frac{\partial}{\partial \bar{x}}\left[\left(\int_w \sigma_{\bar{\theta}}(\bar{x}, \bar{u}, w)\partial_w \bar{F}_{\bar{\theta},t}(w)dw\right)^2 \partial_{\bar{x}}\bar{F}_{\bar{\theta},t}(\bar{x})\right]$$
$$\bar{\theta} \in \Theta,\ \bar{\mu}_0(\bar{x}) \ \text{ fixed}.$$

Note that the last equation of Theorem 1 is obtained by integration of the Fokker-Planck-Kolmogorov equation.

A. Propagation of chaos

Next, we focus on a methodology of proving a mean field convergence of the population profile process M^n based on asymptotic indistinguishability leading to the propagation of chaos properties[1] (also called *the decoupling property at the limit*) of stochastic difference games with discrete time state processes. We use the works in [14], [16], [15] to prove the mean field convergence of a class of mean field stochastic dynamic games with multiple classes of large populations. We further establish a connection to a *controlled de Finetti Theorem* which states that *for any fixed control law (can be a function of the mean field limit), an infinite indistinguishable (or exchangeable) sequence $\{x_j\}_j$ is a mixture of i.i.d. sequences* and next show the propagation of chaos property [13], [5].

Definition 1: We say that a process $x^n = (x_1^n, \ldots, x_n^n)$ satisfies the indistinguishability property[2] if the law of x^n is invariant by permutation of its n components over the index set $\{1, \ldots, n\}$.

Example 1: As an example, the system (3) satisfies the indistinguishability property (for the full population) if all the functions f_{θ_j} reduces to the same function f and the controls have the same law. In particular, for the same value of θ_j, the system (3) becomes asymptotically indistinguishable.

By defining the class of $\bar{\theta}$ as the set of players such that $\theta_j = \theta$, i.e $\{j \in \{1, \ldots, n\}, \mid \theta_j = \theta\}$, the system (3) satisfies *an asymptotic indistinguishability per class property*. This means that the (asymptotic) law of x^n (resp., $\bar{x}^n$) for a given u^n is invariant by permutation within the same class $\bar{\theta}$.

We are within the framework of [14], [5] which states that under the asymptotic indistinguishability per class, the convergence in law of the process M^n to μ and the propagation of chaos (for the hull trajectory) are equivalent. Under uncontrolled large populations, [16] has established the "propagation of chaos" - that is, if the individual initial behaviors are approximately independent and identically distributed, then their behaviors are also asymptotically approximately independent on any finite time interval and described by a common stochastic process where the asymptotic is taken in the population size.

Note that in general the chaoticity property may not hold in the stationary regime. In particular two randomly picked players in the infinite population may not be independent in the stationary regime, i.e., the statistical independence property may not hold (they can be correlated via the payoffs and/or the individual states). We mention a particular case where the rest point m^* (stationary distribution in t) of the Fokker-Planck-Kolmogorov (FPK) equation is related to the δ_{m^*}-chaoticity. If the FPK has a unique global attractor m^* for any initial condition and any control then the propagation of chaos property holds for the measure δ_{m^*}. Beyond this particular case, one may have multiple rest points but also the double limit $\lim_n \lim_t M^n(t)$ may differ from $\lim_t \lim_n M^n(t)$ leading to a non-commutative diagram. Thus, an in-depth study of the dynamical system is required if one wants to analyze a performance metric in the stationary regime. An example of different double limits is provided in [19].

Denote by μ^n the law of $x^n = (x_1^n, \ldots, x_n^n)$:

$$\mu^n = \mathcal{L}(x_1^n, \ldots, x_n^n)$$

and by $\mu^{n,1}$ the marginal of μ^n relatively to the first component, i.e., the canonical projection via the mapping $(x_1^n, \ldots, x_n^n) \longmapsto x_1^n$.

Lemma 2 ([20]): (i) The law of x_j^n is $\mathbb{E}M^n$ where $\mathbb{E}$ denotes of the expectation operator for the random measure. (ii) The process $\tilde{x}^n$ is asymptotically indistinguishable within each class. The population profile process M^n converges to μ in law if and only if

$$\lim_n \int \left(\prod_{l=1}^{L} \phi_l(x_{j_l}^n) \right) \mu^n(dx^n) = \prod_{l=1}^{L} (\int \phi_l d\mu) \qquad (7)$$

for any fixed natural number $L \geq 2$ and a collection of measurable bounded functions $\{\phi_l\}$ and any fixed class $\bar{\theta}$. The above result says that the state law of any generic player j is $\mathcal{L}(z_j^n) = \mathbb{E}M^n$.

B. Macroscopic McKean-Vlasov equation

We extend the mean field convergence result when each individual controls its state evolution and optimizes a long-term coupling payoff in parallel. We prove that for any fixed time $t > 0$, and any fixed admissible control trajectory, the individual process $\tilde{x}_j^n(t)$ is not far from a process $\bar{x}_j(t)$ solving the *individual McKean-Vlasov equation*. Using the Monge-Kontorovich distance, we show that $(\bar{x}_j^n(t))_{t\in[0,T]}$ converges to $(\bar{x}_j(t))_{t\in[0,T]}$ for any $T < +\infty$. Recall that, given two measures μ, ν, the Monge-Kontorovich distance (also called Wasserstein distance) between μ and ν is

$$\mathcal{W}_1(\mu, \nu) = \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}|X - Y|.$$

The Monge-Kontorovich distance metrizes the weak topology.

Lemma 3: Under A1-A2, the McKean-Vlasov system

$$\begin{aligned} d\bar{x}_{\bar{\theta}}(t) &= \int_w f_{\bar{\theta}}(\bar{x}_{\bar{\theta}}(t), u_{\bar{\theta}}(t), w)\mu_t(dw)dt \\ &\quad + \int_w \sigma_{\bar{\theta}}(\bar{x}_{\bar{\theta}}(t), u_{\bar{\theta}}(t), w)\mu_t(dw)d\mathbb{B}(t) \qquad (8) \\ \bar{x}(0) &= q \end{aligned}$$

has a unique solution.

Proof: From the assumptions A1-A2, the standard assumptions [10] ensuring uniqueness of the solution of SDE (8) are satisfied. ■

Let $\tilde{x}^n(t)$ be the random vector $(\tilde{x}_1^n(t), \tilde{x}_2^n(t), \ldots, \tilde{x}_n^n(t))$. Then, $\tilde{x}^n(t)$ satisfies

$$\tilde{x}^n(t+1) = \tilde{x}^n(t) + \Sigma^n(t)(\mathbb{B}(t+1) - \mathbb{B}(t)) + \delta_n G^n(t),$$

[1] This follows the works by de Finetti (1931), Hewitt & Savage (1955), Aldous (1983), Sznitman (1991), Graham (2000), Tanabe (2006), McDonald (2007) etc.

[2] It is also called exchangeability property. Note that this property does not means that the players identical.

where G^n is the n-dimensional vector with j-th component

$$G_j^n = \sum_{i=1}^{n} \omega_{ij}^n f_{\theta_j}(\tilde{x}_j^n(t), u_j^n(t), \tilde{x}_i^n(t)).$$

The term Σ^n is defined similarly.

Let $\tilde{\mu}_t^n$ be the law of $\tilde{x}^n(t) = (\tilde{x}_1^n(t), \tilde{x}_2^n(t), \ldots, \tilde{x}_n^n(t))$ generated by the individual dynamics. Then $\tilde{\mu}_t^n$ is characterized by *Fokker-Planck-Kolmogorov* forward equation. We construct independent processes $\bar{x}_j(t)$ satisfying $\bar{\mathbb{B}}_j = \tilde{\mathbb{B}}_j$ and

$$\| \bar{x}(0) - \tilde{x}^n(0) \| \leq \frac{c}{\sqrt{n}} \rightarrow 0$$

when $n \rightarrow +\infty$.

Proposition 1: Let $\tilde{\mathbb{B}}_j = \bar{\mathbb{B}}_j$. For any t and a control $u_j(t)$, there exists $\tilde{c}_t > 0$ such that

$$\mathbb{E}\left(\| \tilde{x}_j^n(t) - \bar{x}_j(t) \|\right) \leq \frac{\tilde{c}_t}{\sqrt{n}}.$$

Moreover, for any $T < \infty$, there exists $c_T > 0$ such that

$$\mathcal{W}_1\left(\mathcal{L}((\tilde{x}_j^n(t))_{t\in[0,T]}), \mathcal{L}((\bar{x}_j(t))_{t\in[0,T]})\right) \leq \frac{c_T}{\sqrt{n}}.$$

Proof: This is direct consequence of the Theorem 1. ■

As a corollary of the Proposition 1, one gets the mean field convergence of $(\tilde{x}_j^n(t))_t$ by considering the distance

$$d_\infty := \sum_{l\in\mathbb{N}} \frac{1}{2^l} d_l(\tilde{x}^n, \tilde{x}),$$

where

$$d_l(\tilde{x}^n, \tilde{x}) = \sup_{t\in[-l,l]} d(\tilde{x}^n(t), \tilde{x}(t)).$$

IV. Mean field equilibria

A. Characterization of mean field equilibria

Combining (6) and the individual dynamics, one may formulate the following problem:

$$\inf_{u_j} \mathbb{E}\left(g(x_{j,T}) + \int_0^T \bar{c}_{\theta_j}(x_j(s), u_j(s), \mu_s)\, ds\right),$$

subject to (8) where μ_s is the mean field limit at time s. We say that the control law u^* is an ϵ-best response to μ if for any j and any adapted control law u_j, one has,

$$L_{j,T}^n(u_j^*, u_{-j}^*, \mu, x_0) \geq L_{j,T}^n(u_j,, u_{-j}^*, \mu, x_0) - \epsilon.$$

Using a similar argument as in Theorem 1 and good initial estimations and step-size, it can be shown that the rate of convergence is in the order of $n^{-\frac{1}{2}}$. Hence, we have that for any $\epsilon > 0$ there exist a population size n_ϵ such that for all $n \geq n_\epsilon$,

$$\begin{aligned} L_{j,T}^n(u_j^*, u_{-j}^*, \mu, x_0) \geq\ & L_j(u_j, \mu, x_0) \\ & + O(\frac{1}{\sqrt{n}} + \sqrt{\delta_n} + \epsilon_0); \end{aligned}$$

which is in order of

$$L_{j,T}^n(u_j, u_{-j}^*, \mu, x_0) + O(\frac{1}{\sqrt{n}} + \sqrt{\delta_n} + \epsilon_0),$$

where ϵ_0 is the initial error gap. Assuming that $\epsilon_0,\ \delta_n, \frac{1}{\sqrt{n}}$ are all small enough such that each component is at most ϵ for n sufficiently large, we have 3ϵ-equilibria from the mean field equilibria.

Hamilton-Jacobi-Bellman-Fleming optimality at the limit is the following: If there exists a twice continuously differentiable function $v_{j,t}(x_j)$ such that:

$$\begin{aligned} -\partial_t v_{j,t}(x_j) = \inf_{u_j} \Big\{ & \bar{c}_{\theta_j}(x_j, u_j, \mu_t) \\ & + (\bar{f}_t(x_j, u_j, \mu_t)\partial_x v_{j,t}(x_j)) \\ & + \frac{1}{2}\bar{\sigma}_{\theta_j}^2(x_j, u_j, \mu_t)\partial_{xx}^2 v_j(x_j) \Big\}, \\ v_{j,T}(x_j) = g(x_j), & \end{aligned}$$

then $v_{j,t}$ is an optimal payoff associated to the best-response to the mean field μ. This means that the control law solving the backward HJBF gives an ϵ-best response to the finite system for n sufficiently large.

Definition 2: An adapted control law u^* is an ϵ-mean field equilibrium if u^* is an ϵ-best response to μ^* and μ^* is generated by the individual control law u^* (i.e. if every player j plays u_j^* then the mean field limit is μ^* and conversely, the control u_j^* is an ϵ-best response relatively to the payoff of player j.)

The optimal feedback control criterion leads to *Hamilton-Jacobi-Bellman-Fleming equation combined with Fokker-Planck-Kolmogorov equation and macroscopic McKean-Vlasov version of limiting individual dynamics*. Thus, we have the following coupled system of backward-forward equations:

$$\begin{cases} -\partial_t v_{j,t}(x_j) = \inf_{u_j} \left\{ \bar{c}_{\theta_j}(x_j, u_j, \mu_t) \right. \\ \qquad\qquad +(\bar{f}_t(x_j, u_j, \mu_t)\partial_x v_{j,t}(x_j)) \\ \qquad\qquad \left. +\frac{1}{2}\bar{\sigma}_{\theta_j}^2(x_j, u_j, \mu_t)\partial_{xx}^2 v_j(x_j) \right\}, \\ v_{j,T}(x_j) = g(x_j), \\ d\bar{x}_{\bar{\theta}}(t) = \int_w f_{\bar{\theta}}(\bar{x}_{\bar{\theta}}(t), u_{\bar{\theta}}(t), w)\mu_t(dw)dt \\ \qquad + \int_w \sigma_{\bar{\theta}}(\bar{x}_{\bar{\theta}}(t), u_{\bar{\theta}}(t), w)\mu_t(dw)d\mathbb{B}(t), \\ \bar{x}(0) = q, \\ \frac{\partial}{\partial t}\mu_{\bar{\theta},t} + \frac{\partial}{\partial x}\left[\bar{f}_{\bar{\theta},t}(x, u, \mu_t)\mu_{\bar{\theta},t}\right] \\ \qquad = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\bar{\sigma}_{\bar{\theta},t}^2(x, u, \mu_t)\mu_{\bar{\theta},t}\right], \\ \bar{\theta} \in \Theta,\ \mu_0(.) \in \Delta(\mathcal{X}). \end{cases}$$

A natural question is now the existence of solutions to the above system. This is a backward forward system. In general such a system need not have a solution. Also, uniqueness cannot be guaranteed in general. We leave the detailed analysis of existence and uniqueness issues for specific structures of f and σ for future work.

V. Conclusion and future work

We have studied mean field stochastic difference games and established a mean field convergence to controlled stochastic equations of McKean-Vlasov type under suitable conditions. Using Itô-Dynkin's formula, we derived a mean field HJBF and FPK equations characterizing *mean field equilibria*.

We have proved the mean field convergence only for specific weights $\bar{\omega}_{ij}^{n}$. The convergence remains an open question for general coefficients as well as the case of discontinuous mapping (namely the drift and the variance). Another direction is the existence and uniqueness of a solution to the mean field backward-forward equation (of McKean-Vlasov type) under specific drift, noise and payoff functions.

REFERENCES

[1] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: Numerical methods. *SIAM Journal on Numerical Analysis*, 48(3):1136–1162, 2010.

[2] S. Adlakha, R. Johari, G. Weintraub, and A. Goldsmith. Oblivious equilibrium for large-scale stochastic games with unbounded costs. *Proceedings of the IEEE Conference on Decision and Control*, 2008.

[3] R. Buckdahn. Stochastic differential games with a major player and a large number of minor players. 2010.

[4] C. Dogbe. Modeling crowds by the mean-field limit approach. *Mathematical and Computer Modelling*, 52:1506–1520, 2010.

[5] C. Graham. Chaoticity for multiclass systems and exchangeability within classes. *J. Appl. Probab.*, 45(4):1196–1203, 2008.

[6] O Guéant, Jean-Michel Lasry, and Pierre-Louis Lions. Mean field games and applications. *Springer: Paris-Princeton Lectures on Mathematical Finance, Eds. René Carmona, Nizar Touzi*, 2010.

[7] M. Huang, P.E. Caines, and R.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.

[8] Minyi Huang. Large-population lqg games involving a major player: The nash certainty equivalence principle. *SIAM J. Control Optim.*, 48:3318–3353, 2010.

[9] 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. *Commun. Inf. Syst.*, 6(3):221–252, 2006.

[10] P. Kotolenez and T. Kurtz. Macroscopic limits for stochastic partial differential equations of mckean-vlasov type. *Probab. Theory Relat. Fields*, pages 189–222, 2010.

[11] A. Lachapelle and M.-T. Wolfram. On a mean field game approach modeling congestion and aversion in pedestrian crowds. 2011.

[12] J.M. Lasry and P.L. Lions. Mean field games. *Japan. J. Math.*, 2:229–260, 2007.

[13] D.R. McDonald. Lecture notes on mean field convergence. *Dept. of Mathematics, University of Toronto*, 2007.

[14] A. S. Sznitman. Topics in propagation of chaos. *In P.L. Hennequin, editor, Springer Verlag Lecture Notes in Mathematics 1464, Ecole d'Eté de Probabilités de Saint-Flour XI (1989)*, pages 165–251, 1991.

[15] D. Talay and O. Vaillant. A stochastic particle method with random weights for the computation of statistical solutions of mckean-vlasov equations. *Annals Applied Probability*, 13(1):140–180, 2003.

[16] Y. Tanabe. The propagation of chaos for interacting individuals in a large population. *Mathematical Social Sciences*, 51:125–152, 2006.

[17] H. Tembine. Large-scale games in large-scale systems. *Valuetools*, May 2011.

[18] H. Tembine. Mean field stochastic games: convergence, q/h learning, optimality. *American Control Conference, ACC 2011, San Francisco, California, US*, 2011.

[19] H. Tembine. Mean field stochastic games: Simulation, dynamics and applications. *Notes, 350 pages, Supelec*, August 2011.

[20] H. Tembine, Minyi Huang, Roland P. Malhamé, and Peter E. Caines. Mean field convergence of stochastic dynamic games. *Unpublished manuscript*, August 2010.

[21] H. Tembine, J. Y. Le Boudec, R. ElAzouzi, and E. Altman. Mean field asymptotic of markov decision evolutionary games and teams. *in the Proc. of GameNets*, May 2009.

[22] B.-C. Wang and J.-F. Zhang. Distributed control of multi-agent systems with random parameters and a major agent. *Automatica*, 2011.

[23] Bingchang Wang and Ji-Feng Zhang. Mean field games for large population stochastic multi-agent systems with markov jumps. *Proceedings of the 29th Chinese Control Conference, 4572-4577, Beijing, China*, July 29-31 2010.

[24] G. Y. Weintraub, L. Benkard, and B. Van Roy. Oblivious equilibrium: A mean field approximation for large-scale dynamic games. *Advances in Neural Information Processing Systems*, 18, 2005.

[25] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag. Learning in mean-field oscillator games. *In the Proceedings of the IEEE Conference on Decision and Control, Atlanta*, December 2010.

[26] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag. Synchronization of coupled oscillators is a game. *in Proc. of American Control Conference*, 2010.