Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Cooperative Mean-Field Type Games ⋆

Abdoul Karim Cissé * **Hamidou Tembine** **

* *GrIIsG - ISG International Business School, Paris, France*
** *SRI - Center for Uncertainty Quantification in Computational Science & Engineering CEMSE, KAUST (e-mail: tembine@ieee.org).*

Abstract: In the standard formulation of a game, a player's payoff function depends on the states and actions of all the players. Yet, real world applications suggest to consider also a functional of the probability measure of states and actions of all the players. In this paper, we consider cooperative mean-field type games in which the state dynamics and the payoffs depend not only on the state and actions but also on their probability measure. We establish stochastic maximum principle and provide a time-dependent payoff allocation procedure for coalitions. The allocated payoff considers not only fairness property but also the cost of making the coalition. Finally, time consistency and subgame perfectness solution concept equations are established.

Keywords: Mean-field type, cooperation, stochastic maximum principle, subgame perfection

1. INTRODUCTION

Modeling the joint interactive behavior of players in a finite or infinite dynamical systems has been one of the major challenges. This is usually addressed with dynamic game theory. In the standard formulation of a game, a player's payoff function depends on all the states and actions of all the players. Yet, recent observations suggest to consider also a functional of the probability measure of states and actions of all the players. This leads to the introduction of Mean-Field Type Games which constitutes a class of games in which the payoff functions depend not only on the states and actions of the players but also on the probability distribution of states and actions. In this setup, the number of players is not necessarily large and the influence of a single player strategy can have a big impact on the states and the mean field term. We examine the possibility for players to cooperate in a such a scenario. One of the fundamental element in the theory of cooperative games is the formulation of the optimal behavior for the players. Player behavior (control action and imputations) satisfying specific optimality behaviors then constitutes a solution of the game. In other words, a solution concept of a dynamic cooperative game is produced by a set of optimality principles such as dynamic bargaining solution and payoff allocation procedure, Yeung and Petrosyan (2006).

Substantial progress have been done in the last decade in mean field games in the non-cooperative setup. However, very little is known about *cooperative mean-field games*. The main reason is that the classical optimality equation used in Cooperative Stochastic Differential Games (see for example Petrosjan (1977); Yeung and Petrosyan (2006)) are not valid when the mean-field is involved. As we will see in Theorem 1, the mean-field of actions plays an important role in the optimality equations Tembine (2012).

⋆ The authors would like to thank Prof. Boualem Djehiche (KTH) for his comments.

Altruism and cooperation are fascinating research areas. One has attempted to claim that the players are better off they all work cooperatively. However, we are often observing very strange behaviors that are far from cooperation. So, if cooperation is answer, what is the question and why these strange behaviors?

Let us consider a simple example with two players. Assume that if they work together (jointly) they will be able to get $v(\{12\})$. Player 1 gets $v(\{1\})$ if he or she works alone and Player 2 gets $v(\{2\})$. From these three numbers, it is not clear why these players should work together.

To formalize it in terms of their interest, we introduce a cost of making a coalition, $c(\{12\}) \geq 0$ which is the cost incured when both players pool their effort (it includes information exchanging cost, coalition creation cost, etc). While this cost is often neglected in the literature, it may be important in many setups.

Thus, a necessary condition for possible cooperation between the players is

$$v(\{12\}) - c(\{12\}) > v(\{1\}) + v(\{2\}).$$

Then, the next question is: what will be their payoff if they cooperate? To answer to this question, we need to know how to share the outcome of the cooperation. It is clear that allocating the equal share $\frac{1}{2}[v(\{12\}) - c(\{12\})]$ to each player is not necessary appropriate since it can be less than $\max(v(\{1\}), v(\{2\}))$. Thus, the allocation has to be done in a more clever way.

Cooperative game-theoretic solutions such as Bargaining, Core, Shapley, Nucleolus dealt with such problems. When stochasticity and time-dependency are involved, the solution concepts require a careful adaptation. In addition, if the payoff function and the state dynamics are of mean-field type, the optimality equations need to be established (see Result 1 and 2).

In this paper we consider cooperative mean-field type games for state processes governed by a controlled stochastic differential equation of mean-field type which is also called McKean-Vlasov type equation, in the sense that the coefficients of the state dynamics are allowed to depend on the state of the process as well as its distribution in general, so that we need to introduce two adjoint processes.

The general cooperative game of mean-field type of interest can be formulated as a collective reaction to the evolutive states, actions and distribution of state and actions of the players. The fact that the payoff functional (grand coalition value) may be nonlinear with respect to the conditional expectation, makes the cooperative game problem time inconsistent in the sense that Bellman's optimality principle, based on applying the law of iterated conditional expectations on the payoff functional, does not hold. A way to solve this cooperative game problem is to derive an extended version of the Dynamic Programming Principle and extended Hamilton-Jacobi-Bellman equation which turns out to be very complex.

We establish a payoff allocation procedure for a class of cooperative mean-field type games where the distribution enters not only through states, but also through the control actions. Our results are based on the recent development of stochastic maximum principle of such systems.

To the best to our knowledge, time-consistent allocation procedure in cooperative mean-field type games have not been established in earlier works, is entirely new and, is fundamentally different from the existing results in mean-field type control case Buckdahn et al. (2011); Andersson and Djehiche (2010); Bensoussan et al. (2011).

The remainder of this paper is organized as follows. In next section we introduce the grand coalition mean-field type game. In Section 3 we present the time-dependent allocation and the stochastic maximum principle associated the cooperative outcome. In Section 4 we propose subgame perfectness solution concept via spike variation techniques.

2. PROBLEM STATEMENT

Let $T > 0$, $t \in \mathcal{T} = [0,T]$, $\mathcal{B}$ be a standard Brownian motion on $\mathcal{T}$ defined on a given filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$. We consider a cooperative mean-field type game given by $\mathcal{G} = (\mathcal{N}, \mathcal{T}, c(.,\mathcal{T}), v(.,\mathcal{T}))$, where $\mathcal{N} = \{1, 2\ldots, n\}$, is the set of players, $n \geq 1$, $v(.,\mathcal{T}) : 2^{\mathcal{N}} \to \mathbb{R}$ is the value associated to a coalition and $c(.,\mathcal{T}) : 2^{\mathcal{N}} \to \mathbb{R}$ is a cost of coalition making. The grand coalition problem $v = v(\mathcal{N}, \mathcal{T})$ is given by

$$\begin{cases} \sup_{(a_1,\ldots,a_n)} \mathbb{E}\left[g_0(x(T), \mathbb{E}\xi(x(T))) + \int_{t\in\mathcal{T}} r_0(t,x,m,a)\ dt\right], \\ \text{subject to} \\ dx(t) = b(t,x(t),m(t),a(t))dt + \sigma(t,x(t),m(t),a(t))d\mathcal{B} \\ + \int_\Theta \bar{\gamma}(t,x(t_-),m(t_-),a(t_-),\theta)\bar{N}(dt,d\theta), \\ x(0) = x_0 \in \mathcal{X} \subseteq \mathbb{R}^d, \\ m(t) = \mathbb{E}\phi(t,x(t),a(t)) \in \mathbb{R}^{d_\phi},\ t \in \mathcal{T} \end{cases} \quad (1)$$

where $\mathcal{N}$ is the set of players (assumed to be finite), $t \in \mathcal{T}$, $x(t) \in \mathcal{X} \subseteq \mathbb{R}^d$ is a d-dimensional state vector, $d \geq 1$; $a_i(t) \in \mathcal{A}_i$, is the control of the player i at time t with $\mathcal{A} = \prod_{i\in\mathcal{N}} \mathcal{A}_i \subseteq \mathbb{R}^k$, the control action set $\mathcal{A}_i$ is non-empty, $\mathcal{B}$ is a standard Brownian motion in $\mathbb{R}^l$. $m(t) = \mathbb{E}\phi(t,x(t),a(t)) \in \mathbb{R}^{d_\phi}$, is the aggregative term which uses the probability distribution of $(x(t),a(t))$ at time t, d_ϕ is the dimension of ϕ. The aggregate function is $\phi : \mathcal{T} \times \mathcal{X} \times \mathcal{A} \to \mathbb{R}^{d_\phi}$.

Specification of the payoffs: The instantaneous payoff function is $r_0 = \sum_{i\in\mathcal{N}} r_i : \mathcal{T} \times \mathcal{X} \times \mathbb{R}^{d_\phi} \times \mathcal{A} \to \mathbb{R}$, the terminal aggregate function is $\xi : \mathcal{X} \to \mathbb{R}^{d_\xi}$. $g_0 = \sum_i g_i : \mathcal{X} \times \mathbb{R}^{d_\xi} \to \mathbb{R}$ is the terminal payoff.

Specification of the state dynamics: The drift coefficient function is $b : \mathcal{T} \times \mathcal{X} \times \mathbb{R}^{d_\phi} \times \mathcal{A} \to \mathbb{R}^d$, the diffusion coefficient function is $\sigma : \mathcal{T} \times \mathcal{X} \times \mathbb{R}^{d_\phi} \times \mathcal{A} \to \mathbb{R}^{d\times l}$.

Specification of the jump process: The term $\bar{N}$ is a Poisson martingale measure with characteristic $\mu(d\theta)$. $\bar{N}$ is independent of $\mathcal{B}$ and the measure μ is a $\sigma-$finite measure over Θ. The function $\bar{\gamma} : \mathcal{T} \times \mathcal{X} \times \mathbb{R}^{d_\phi} \times \mathcal{A} \times \Theta \to \mathbb{R}^k$. The filtration $\mathcal{F}_t$ is the one generated by the union of events from $\mathcal{B}$ or $\bar{N}$ up time t.

We will first provide the results for $\bar{\gamma} = 0$ the case of jump is in Section 4.1.

Definition 1. We refer to mean-field type game, any game in which the instantaneous payoffs and/or the state dynamics coefficient functions involve the distributions of states or the distribution of actions or the joint distributions of state-actions.

From Definition 1, a mean-field type game can be static or dynamic in time. The classical variance reduction payoff such as mean-variance payoff is an example of mean-field type.

Remark 1. Mean-field games are usually analyzed for individual states. Here we propose a shared state setup because of its interesting applications.

Definition 2. An admissible control is an $\mathcal{F}-$adapted and square integrable process with values in a non-empty subset $\mathcal{A}$. Denote by $\mathbb{A}$ the class of admissible controls. Any admissible control that solves (1) is called optimal cooperative strategy.

Our objective is to find or to characterize (through a Stochastic Maximum Principle) an optimal cooperative strategy $a^* \in \arg\max_a v(.)$, $a \in \mathbb{A}$.

2.1 On the well-posedness of the problem

We will make the following assumptions in this paper.

Hypothesis H1: The functions b, σ, r, g are twice continuously differentiable with the respect to (x,m). Moreover, b, σ, r, g and all their first and second derivatives with the respect to (x,m) are continuous in (x,m,a) and bounded.

Lemma 2. (Existence). : Under H1, for each control $a(.) \in \mathcal{A}$, the state dynamics admits a unique strong solution, $x(t) := x^a(t)$.

The proof follows from Theorem 3.1 of Buckdahn et al. (2009).

Next, we focus on how the cooperation outcome can be allocated between the players in a dynamic setting.

2.2 Payoff allocation

Noncooperative equilibrium plays the role of a benchmark in a cooperative game, i.e., gives what players could secure for themselves if there is no agreement, i.e., $v(\{i\}, \mathcal{T}) =$

$$\begin{cases} \sup_{a_i} \mathbb{E}\left[g_i(x(T), \mathbb{E}\xi(x(T))) + \int_{t\in\mathcal{T}} r_i(t,x,m,a)\ dt\right], \\ \text{subject to} \\ dx(t) = b(t,x(t),m(t),a(t))dt + \sigma(t,x,m,a)d\mathcal{B} \\ x(0) = x_0 \in \mathcal{X} \subseteq \mathbb{R}^d, \\ m(t) = \mathbb{E}\phi(t,x(t),a(t)) \in \mathbb{R}^{d_\phi} \\ t \in \mathcal{T} \end{cases} \quad (2)$$

Let $R_{c,i}$ be the total before-side-payment cooperative payoff of player i i.e.,

$$R_{c,i} = \mathbb{E}\left[g_i(x_c^*(T), \mathbb{E}\xi(x_c^*(T))) + \int_{t\in\mathcal{T}} r_i(t,x_c^*,m_c^*,a_c^*)\ dt\right].$$

One has $\sum_{i\in\mathcal{N}} R_{c,i} = v(\mathcal{N},\mathcal{T})$. However, one needs to find a better way to share the payoff $v(\mathcal{N},\mathcal{T}) - c(\mathcal{N},\mathcal{T})$. This leads to the introduction of the notion of imputation, i.e., a vector profile $\gamma_1,\ldots,\gamma_n$ such that $\sum_{i\in\mathcal{N}'}\gamma_i \geq v(\mathcal{N}',\mathcal{T}) - c(\mathcal{N}',\mathcal{T})$ for any $\mathcal{N}' \subset \mathcal{N}$, and $\sum_{i\in\mathcal{N}}\gamma_i = v(\mathcal{N},\mathcal{T}) - c(\mathcal{N},\mathcal{T})$

By virtue of mean-field type joint optimization, the sum of individual payoffs under cooperation is greater or equal to its noncooperative counterpart, i.e.,

$$\sum_i R_{c,i} \geq \sum_i v(\{i\},\mathcal{T}).$$

Thus, the dividend of cooperation (without coalition making cost) is $DC = \sum_i R_{c,i} - \sum_i v(\{i\},\mathcal{T}) \geq 0$.

Thus, the dividend of cooperation (with coalition making cost) to be distributed among the players is $DC - c(\mathcal{N},\mathcal{T})$.

As a first consequence, it is clear if the coalition making cost is too high (compared to the game coalition value) then there is no reason for the players to form coalition.

Therefore for cooperation purpose we require the positivity of $DC - c(\mathcal{N},\mathcal{T})$. Using a cooperative game approach yields individual payoffs for the whole interval $\mathcal{T}$. The selected imputation has, by definition, the property that each player's payoff in the cooperative game is higher or equal to what she would get in a noncooperative game played on the same time interval.

Let $\gamma_i(s) = \gamma(\{i\},[s,T])$ be the cooperative payoff-to-go after side payment for player i at position $[s,T], 0 < s < T$ of the game. This is the amount that player i will actually pocket. One way sharing the payoff is to use a dynamical Shapley value.

Dynamical Shapley Value The allocated payoff to player i is

$$\gamma_i = \sum_{\mathcal{N}'\subset\mathcal{N}, i\notin\mathcal{N}'} \frac{|\mathcal{N}'|!(|\mathcal{N}| - |\mathcal{N}'| - 1)!}{|\mathcal{N}|!} [(v-c)(\mathcal{N}'\cup\{i\},[s,T]) - (v-c)(\mathcal{N}',[s,T])]. \quad (3)$$

For two players case, the payoffs of the players are

$$\gamma_1 = v(\{1\}) + \frac{1}{2}\left[v(\{12\}) - c(\{12\}) - v(\{1\}) - v(\{2\})\right] \quad (4)$$

$$\gamma_2 = v(\{2\}) + \frac{1}{2}\left[v(\{12\}) - c(\{12\}) - v(\{1\}) - v(\{2\})\right] \quad (5)$$

Definition 3. A cooperative solution is time consistent if, at any position $s \in [0,T]$ the cooperation solution payoff to go $\gamma(\{i\},[s,T]) \geq v(\{i\},[s,T])$ where the deviating payoff $v(\{i\},[s,T])$ is computed along the state trajectory of the cooperative state trajectory $x_c^*(t)$.

This notion of time consistency and its implementation in cooperative differential games was introduced in Petrosjan (1977). A stringer notion of time consistency is that the cooperative payoff-to-go dominates the noncooperative payoff-to-go for any state $x(s),\ s \in \mathcal{T}$. This is called subgame consistent solution.

Dynamical proportional allocation Though one of the most commonly used allocation principles is the dynamical Shapley value, however, in the case when players may be asymmetric in their powers and sizes of payoffs, equal imputation of cooperative gains may not be totally agreeable to asymmetric player. To overcome this, one can suggest the allocation principle in which the players, shares of the gain from cooperation are proportional to the relative sizes of their expected deviating payoffs. Thus, a proportional time-consistent solution is given by

$$\begin{aligned}\gamma_i(s) &= \frac{v(\{i\},[s,T])}{\sum_i v(\{i\},[s,T])}\sum_i \gamma(\{i\},[s,T]) \\ &= \frac{v(\{i\},[s,T])}{\sum_i v(\{i\},[s,T])}[v(\mathcal{N},[s,T]) - c(\mathcal{N},[s,T])]. \quad (6)\end{aligned}$$

Assuming these quantities are positive, one gets

$$\gamma(\{i\},[s,T]) \geq v(\{i\},[s,T]), \forall s,$$

(individual rationality) and

$$\sum_i \gamma(\{i\},[s,T]) = v(\mathcal{N},[s,T]) - c(\mathcal{N},[s,T])$$

(efficiency at any time).

Then one needs t find the appropriate instantaneous distribution of the payoff γ_i, i.e. find function $\hat{\gamma}_i$ such that

$$\gamma_i(0) = \hat{\gamma}_i(T) + \int_0^T \hat{\gamma}_i(s)\ ds.$$

Next we establish necessary conditions under H1. These conditions may not be sufficient in general. However, if in addition convexity holds, these conditions become sufficient.

3. STOCHASTIC MAXIMUM PRINCIPLE

Let H be $H(t,x,m,a,p,q) = b'p + trace(\sigma' q) + r$. In order to simplify the notation, we remove the trace. This Pontryagin function introduces two adjoint processes p an q. The adjoint process p is associated with the drift f and q is associated with the diffusion coefficient σ.

First order adjoint equation: Let

$$\begin{aligned}\alpha_\phi(t) = r_x + (b_x)'p + (\sigma_x)'q \\ +(\partial\phi)'\left(\mathbb{E}[r_m] + \mathbb{E}[(b_m)'p] + \mathbb{E}[(\sigma_m)'q]\right) \quad (7)\end{aligned}$$

then $\alpha_\phi(t) \in \mathbb{R}^d$ and the first order adjoint equation is given by

$$dp = -\alpha_\phi(t)dt + q(t)d\mathcal{B}(t), \tag{8}$$
$$p(T) = g_x(T) + (\partial\xi)'\mathbb{E}[g_m(T)]. \tag{9}$$

Lemma 3. (Existence). Under assumption H1, the mean-field type first order adjoint equation admits a unique solution $\mathcal{F}-$adapted solution (p, q) such that $\mathbb{E}[\sup_{t\in\mathcal{T}} |p(t)|^2 + \int_0^T |q(t)|^2 \, dt] < +\infty$.

Second order adjoint equation: Let

$$\beta_\phi = [Pb_x + b_x'P' + P\sigma_x^2 + Q\sigma_x + \sigma_x'Q' + H_{xx}] + (\mathbb{E}[r_m'] + \mathbb{E}[P'b_m] + \mathbb{E}[\sigma_m'Q]')\,\partial^2\phi. \tag{10}$$

The second order adjoint equation is given by $dP = -\beta_\phi dt + Qd\mathcal{B}(t),\ P(T) = g_{xx} + \mathbb{E}[g_m']\partial^2\xi$.

Lemma 4. (Existence). Under assumption H1, the mean-field second order adjoint equation admits a unique solution $\mathcal{F}-$adapted solution (P, Q) such that $\mathbb{E}[\sup_{t\in\mathcal{T}} |P(t)|^2 + \int_0^T |Q(t)|^2 \, dt] < +\infty$.

3.1 Necessary condition under H1

Result 1. Let H1 holds. If (x^*, a^*) is an optimal solution of the cooperative mean-field type game then there are two pairs of processes $(p, q), (P, Q)$ that satisfy the first order and the second order adjoint equations such that

$$H(t, x^*, m^*, a^*, p, q) - H(t, x^*, m^*, a, p, q) + \frac{1}{2}P(t)\left(\sigma(t, x^*, m^*, a^*) - \sigma(t, x^*, m^*, a)\right)^2 \geq 0, \tag{11}$$

for all $a(.) \in \mathcal{A}$, almost every t and $\mathbb{P}-$almost surely where

$$\begin{cases} dp(t) = -\alpha_\phi dt + qd\mathcal{B}(t),\ p(T) = g_x(T) + (\partial\xi)'\mathbb{E}[g_m(T)], \\ dP = -\beta_\phi dt + Qd\mathcal{B}(t),\ P(T) = g_{xx} + \mathbb{E}[g_m']\partial^2\xi. \end{cases}$$

The inequality involving the Pontryagin function in Theorem 1 may be achieved at several control laws. In that case we associate and properly define a specific state trajectory to any optimal control law. Note that even if the optimal control law is unique, the optimal state trajectory may not unique (discontinuous) depending on the starting point. This leads to the so-called Skiba point. Note that Theorem 1 does not say anything about the cooperative game $\mathcal{G}$. For the cooperative mean-field type games one needs to check the stability of the coalitions and the allocation procedure which is missing in the global optimization of mean-field type. The cooperative strategy obtained in Theorem 1 may NOT be time consistent, i.e., for $t' > t$,

$$a^*(t_0, x_0, .) \neq \arg\max_a v(\mathcal{N}, x', [t', T]).$$

In order to characterize the payoff of the deviating players $v(\{i\}, \mathcal{T})$, one needs to understand the non-cooperative game solution, which necessary condition under H1 is given by the following system:

If (x_i^*, a_i^*) is an equilibrium solution of the non-cooperative mean-field type game then there are $2n$ pairs of processes $(p_i, q_i), (P_i, Q_i),\ i \in \mathcal{N}$ that satisfy the first order and the second order system of adjoint equations such that

$$H_i(t, x^*, m^*, a^*, p_i, q_i) - H_i(t, x^*, m^*, a, p_i, q_i) + \frac{1}{2}P_i(t)\left(\sigma(t, x^*, m^*, a^*) - \sigma(t, x^*, m^*, a)\right)^2 \geq 0, \tag{12}$$

for all $a_i(.) \in \mathcal{A}_i$, almost every t and $\mathbb{P}-$almost surely where

$$\begin{cases} dp_i(t) = -\alpha_{i,\phi}dt + q_i d\mathcal{B}(t), \\ p_i(T) = g_{i,x}(T) + (\partial\xi)'\mathbb{E}[g_{i,m}(T)], \\ \alpha_{i,\phi} = r_{i,x} + (b_x)'p_i + (\sigma_x)'q_i \\ \quad +(\partial\phi)'\left(\mathbb{E}[r_{i,m}] + \mathbb{E}[(b_m)'p_i] + \mathbb{E}[(\sigma_m)'q_i]\right) \\ dP_i = -\beta_{i,\phi}dt + Q_i d\mathcal{B}(t),\ P_i(T) = g_{i,xx} + \mathbb{E}[g_{i,m}']\partial^2\xi \\ \beta_{i,\phi} = [P_i b_x + b_x'P_i' + P_i\sigma_x^2 + Q_i\sigma_x + \sigma_x'Q_i' + H_{i,xx}] \\ \quad + \left(\mathbb{E}[r_{i,m}'] + \mathbb{E}[P_i'b_m] + \mathbb{E}[\sigma_m'Q_i]'\right)\partial^2\phi. \end{cases}$$

In the next section we focus on time-consistent and subgame consistent solution.

3.2 Aggregative term as the vector of mean states and mean actions

We consider a particular case where $\phi(t, x(t), a(t)) = (x(t), a(t))$, i.e., the aggregative term is exactly the vector of mean of states and mean actions. The payoff has the form of $r(t, x(t), \mathbb{E}x(t), \mathbb{E}a(t), a(t))$. Similar for the coefficient functions b, σ. This means that $d_\phi = 2$.

$\alpha_\phi(t)$ becomes $r_x + (b_x)'p + (\sigma_x)'q + \mathbb{E}[r_m] + \mathbb{E}[(b_m)'p] + \mathbb{E}[(\sigma_m)'q]$.

The following result follows from Theorem 1

Corollary 5. Let H1 holds. If (x^*, a^*) is an optimal solution of the cooperative mean-field type game then there are two pairs of processes $(p, q), (P, Q)$ that satisfy the first order and the second order adjoint equations, such that

$$H(t, x^*, m^*, a^*, p, q) - H(t, x^*, m^*, a, p, q) + \frac{1}{2}P(t)\left(\sigma(t, x^*, m^*, a^*) - \sigma(t, x^*, m^*, a)\right)^2 \geq 0 \tag{13}$$

for all $a(.) \in \mathcal{A}$, almost every t and $\mathbb{P}-$almost surely where

$$\begin{cases} dp = -(b_x p + \mathbb{E}[b_m p] + \sigma_x q + \mathbb{E}[\sigma_m q] + r_x + \mathbb{E}[r_m])dt \\ \quad +qd\mathcal{B}(t), \\ p(T) = g_x(T) + \mathbb{E}[g_m(T)], \\ dP = -[2b_x P + \sigma_x^2 P + 2\sigma_x Q + H_{xx}]dt + Qd\mathcal{B}(t), \\ P(T) = g_{xx}. \end{cases}$$

Example 1. (Variance Reduction). Let

$-r(t, x(t), \mathbb{E}x(t), \mathbb{E}a(t), a(t)) = \langle Qx, x\rangle + \langle\hat{Q}\mathbb{E}x, \mathbb{E}x\rangle + \langle Ra, a\rangle + \langle\hat{R}\mathbb{E}a, \mathbb{E}a\rangle$.

$-g(T, x(T), \mathbb{E}x(T)) = \langle Gx(T), x(T)\rangle + \langle\hat{G}\mathbb{E}x(T), \mathbb{E}x(T)\rangle$.

$b(t, x(t), \mathbb{E}x(t), \mathbb{E}a(t), a(t)) = Ax + \hat{A}\mathbb{E}x + Ba + \hat{B}\mathbb{E}a$.

$\sigma(t, x(t), \mathbb{E}x(t), \mathbb{E}a(t), a(t)) = A_1x + \hat{A}_1\mathbb{E}x + B_1a + \hat{B}_1\mathbb{E}a$, where $Q, \hat{Q}, \hat{R}, \hat{R}, A, \hat{A}, A_1, \hat{A}_1, B_1, \hat{B}_1$ are deterministic matrix-valued functions chosen appropriately with the dimension of the control process $a(t)$.

If $R, Q, G\ \hat{R}, \hat{Q}, \hat{G}$ are symmetric then our necessary condition above becomes

$$\begin{cases} 2Ra + 2\hat{R}\mathbb{E}a + B'p + \hat{B}'\mathbb{E}p + B_1'q + \hat{B}_1'\mathbb{E}q = 0 \\ dp = -[2Qx + A'p + A_1'q + 2\hat{Q}\mathbb{E}x + \hat{A}'\mathbb{E}p + \hat{A}_1'\mathbb{E}q]dt \\ \quad +qd\mathcal{B}(t) \\ \frac{1}{2}p(T) = Gx + \hat{G}\mathbb{E}[x]. \end{cases}$$

The optimal cooperative strategy under this procedure may depend on the initial state x_0.

Mean of actions as aggregative term For the particular case where $\phi(t, x(t), a(t)) = a(t)$, the aggregative term is exactly the mean of actions. The payoff has the form $r(t, x(t), \mathbb{E}a(t), a(t))$. Similar for the coefficients b, σ. One gets the payoff model used in Kamgarpour and Tembine (2013) in a smart grid context.

Mean of states as aggregative term For the particular case where $\phi(t, x(t), a(t)) = x(t)$, the aggregative is exactly the mean of states. The payoff has the form $r(t, x(t), \mathbb{E}x(t), a(t))$. Similar for the coefficients b, σ. One gets the model in Buckdahn et al. (2011).

4. TOWARDS A CREDIBLE COOPERATION

For dynamic games, an additional and stringent condition on the solutions is required: The specific optimality principle must remain optimal at any instant of time throughout the game duration along the optimal state trajectory. This condition is known as dynamic stability or time consistency.

In the context of mean-field type games, the notion of time consistency is crucial since the initial distribution of states and starting time influences naturally the Kolmogorov forward equation. A cooperative solution is subgame consistent if an extension of the cooperative strategy to a situation with a later starting time and to any possible state brought about by the prior optimal behavior of the players remains optimal. Subgame consistent is a stronger notion of time consistency. In the presence of stochastic elements, subgame consistency is required in a credible cooperative solution. In the field of cooperative mean-field type games, little research has been published to date on subgame consistent solutions.

If the set $\mathcal{A}$ is not necessarily convex, but a general separable complete metric space (Polish space), Pontryagin's approach suggests the following perturbation method called spike variation. The approach is well-adapted to subgame perfection in games. Fix $(s, x) \in [0, T] \times \mathcal{X}$ and define the control law a_ϵ as the spike variation of $\hat{a}$ over the set $[s, s + \epsilon],\ \epsilon > 0$ i.e.,

$$a_\epsilon(t) = a(t)1\!\!1_{[s,s+\epsilon]}(t) + \hat{a}(t)1\!\!1_{\mathcal{T}\setminus[s,s+\epsilon]}(t),$$

where a is an arbitrary admissible control and $1\!\!1_{[s,s+\epsilon]}$ is the indicator function over the set $[s, s + \epsilon]$.

Definition 4. Let R be objective to be maximized in v. We say that $\hat{a}$ is a subgame perfect cooperative strategy under spike variation if for any t_0, x_0, a,

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \left[R([t_0, T], x_0, \hat{a}) - R([t_0, T], x_0, a_\epsilon)\ \right] \geq 0.$$

Note that a subgame perfect cooperative strategy under spike variation is in particular a time consistent solution.

The key difference here is that the solution that we are looking for, should NOT depend on the initial data (when and where we started).

Let H^{t_0,x_0} be the Pontryagin function associated with the random variable x that starts from x_0 at $t_0 \in \mathcal{T}$. $H^{t_0,x_0}(t, x, m, a, p, q) = bp^{t_0,x_0} + \sigma q^{t_0,x_0} + r^{t_0,x_0}$, where the notation r^{t_0,x_0} is obtained from r when the aggregate term is $m = \mathbb{E}_{t_0,x_0}\ \phi(t, x, a)$.

Let

$$\alpha_\phi^{t_0,x_0}(t) = r_x^{t_0,x_0} + (b_x)'p^{t_0,x_0} + (\sigma_x)'q^{t_0,x_0} + (\partial\phi)'\left(\mathbb{E}[r_m^{t_0,x_0}] + \mathbb{E}[(b_m)'p^{t_0,x_0}] + \mathbb{E}[(\sigma_m)'q^{t_0,x_0}]\right)$$

then the first order adjoint adjoint equation is given by

$$dp^{t_0,x_0} = -\alpha_\phi^{t_0,x_0}(t)dt + q^{t_0,x_0}d\mathcal{B}(t), \quad (14)$$

$$p^{t_0,x_0}(T) = g_x^{t_0,x_0}(T) + (\partial\xi)'\mathbb{E}_{t_0,x_0}[g_m^{t_0,x_0}(T)]. \quad (15)$$

Under assumption H1, the mean-field type first order adjoint equation admits a unique solution $\mathcal{F}-$adapted solution $(p^{t_0,x_0}, q^{t_0,x_0})$ such that $\mathbb{E}[\sup_{t\in[t_0,T]} |p^{t_0,x_0}(t)|^2 + \int_{t_0}^T |q^{t_0,x_0}(t)|^2\ dt] < +\infty$.

Let

$$\beta_\phi^{t_0,x_0} = P^{t_0,x_0}b_x + b_x'P'^{t_0,x_0} + P^{t_0,x_0}\sigma_x^2 + Q^{t_0,x_0}\sigma_x^{t_0,x_0} + \sigma_x'Q'^{t_0,x_0} + H_{xx}^{t_0,x_0} + \left(\mathbb{E}[r_m'^{t_0,x_0}] + \mathbb{E}[P'^{t_0,x_0}b_m] + \mathbb{E}[\sigma_m'Q^{t_0,x_0}]'\right)\partial^2\phi.$$

The second order adjoint equation is given by

$$dP^{t_0,x_0} = -\beta_\phi^{t_0,x_0}dt + Q^{t_0,x_0}d\mathcal{B}(t),$$

$$P^{t_0,x_0}(T) = g_{xx}^{t_0,x_0} + \mathbb{E}_{t_0,x_0}[g_m'^{t_0,x_0}]\partial^2\xi.$$

Under assumption H1, the mean-field second order adjoint equation admits a unique solution $\mathcal{F}-$adapted solution $(P^{t_0,x_0}, Q^{t_0,x_0})$ such that $\mathbb{E}[\sup_{t\in[t_0,T]} |P^{t_0,x_0}(t)|^2 + \int_{t_0}^T |Q^{t_0,x_0}(t)|^2\ dt] < +\infty$.

Result 2. Let H1 holds. If $(\hat{x}, \hat{a})$ is an optimal solution of the cooperative mean-field type game then there are two pairs of processes $(p, q), (P, Q)$ that satisfy the first order and the second order adjoint equations, such that

$$H^{t,\hat{x}(t)}(t, \hat{x}, \hat{m}, \hat{a}, p^{t,\hat{x}(t)}, q^{t,\hat{x}(t)}) - H^{t,\hat{x}(t)}(t, \hat{x}, \hat{m}, a, p^{t,\hat{x}(t)}, q^{t,\hat{x}(t)}) + \frac{1}{2}P^{t,\hat{x}(t)}(t)\left(\sigma(t, \hat{x}, \hat{m}, \hat{a}) - \sigma(t, \hat{x}, \hat{m}, a)\right)^2 \geq 0,$$

for all $a(.) \in \mathcal{A}$, almost every t and $\mathbb{P}-$almost surely.

The proof of Theorem 2 follows similar steps as for Theorem 1 conditioning on $t_0 = t, x_0 = x$. Note that now the cooperative solution $\hat{a}$ does not depend on the initial data x_0 and can therefore be used for any $t \in [0, T]$ in any subgame.

Example 2. (Mean-Variance Approach). In the line of the variance reduction example 1 above, we consider a terminal payoff g as $\mathbb{E}x(T) - \lambda\ var(x(T)),\ \lambda > 0$ and $Q, \hat{Q}, R, \hat{R}$ are zeros, which is obviously of mean-field type. The state dynamic is control-dependent and is obtained by setting $\hat{A}, \hat{B}, \hat{A}_1, \hat{B}_1$ to zero. By Theorem 2, one has

$$(*) \begin{cases} B_1'q + B'p = 0 \\ dp = -Apdt + qd\mathcal{B} \\ p(T) = 1\!\!1 - \lambda[x(T) - \mathbb{E}_{t_0,x_0}x(T)]. \end{cases}$$

The final form of $p(T)$ gives an incentive to look for a solution in the following form:

$$p^{t_0,x_0}(t) = \eta_1(t) - \eta_2(t)[x(t) - \mathbb{E}_{t_0,x_0}x(t)],$$

where the terminal values are $\eta_1(T) = \mathbb{1}$ and $\eta_2(T) = \lambda I$. By conditioning on t_0, x_0, it is clear that

$$p^{t_0,x_0}(t_0) = \eta_1(t_0),$$

which is independent of x_0. Using system (*), one deduces $q^{t_0,x_0}(t_0) = -(B_1')^{-1}B'\eta_1(t_0)$. By Ito's formula applied to p, and arranging the coefficient functions, one gets,

$$(\dot{\eta}_1 + A\eta_1) - (\dot{\eta}_2 + \eta_2 A + A\eta_2)[x(t) - \mathbb{E}_{t_0,x_0}x(t)] = 0,$$

and $q = -\eta_2 B_1 \hat{a}$.

$$\begin{cases} \dot{\eta}_1 + A\eta_1 = 0, \ \eta_1(T) = \mathbb{1} \\ \dot{\eta}_2 + \eta_2 A + A\eta_2 = 0, \ \eta_2(T) = \lambda I \end{cases}$$

The optimal cooperative control action is

$$\hat{a}(t) = -(\eta_2 B_1)^{-1}q = (\eta_2 B_1)^{-1}(B_1')^{-1}B'\eta_1(t),$$

which is time-consistent and independent on which state the system starts at time t_0 and is a subgame perfect optimal cooperative strategy.

The unidimensional case,

$$(**) \begin{cases} \eta_1(t) = e^{A(T-t)} \\ \eta_2(T) = \lambda e^{2A(T-t)} \\ p^{t,x}(t) = e^{A(T-t)} \\ q^{t,x}(t) = -(B_1')^{-1}B'e^{-A(T-t)} \\ \hat{a}(t) = \frac{1}{\lambda}B_1^{-2}B'e^{A(T-t)}. \end{cases}$$

Note that, now, the optimal control-action $\hat{a}(t)$ is independent of the initial data x_0.

4.1 With jump process

In this section we reconsider the state process with non-zero jump term $\bar{\gamma}$. In this case the Pontryagin function will be modified to $\bar{H} = H + \int_\Theta trace[\bar{\gamma}\mu(d\theta)\bar{q}]$ where $\bar{q}$ is an adjoint process associated to the jump rate coefficient,

$$\bar{\alpha}_\phi = \alpha_\phi + \int_\Theta \bar{\gamma}_x\mu(d\theta)\bar{q} + (\partial\phi)\mathbb{E}[\int_\Theta \bar{\gamma}_m\mu(d\theta)\bar{q}]$$

and the adjoint process dp will be modified with an extra term $+\int_\Theta \bar{q}\tilde{N}(dt, d\theta)$

5. CONCLUDING REMARKS

We have presented cooperative mean-field type games and derived necessary condition for optimality under assumption H1. These optimality equations apply for non-convex domain as well as non-convex Hamiltonian functions. The maximum principle involves two quasi-linear mean-field backward stochastic differential equation for which the existence and uniqueness are established under H1. However, the conditions we provided here are NOT sufficient in general and non-transferable payoff is not considered.

REFERENCES

Andersson, D. and Djehiche, B. (2010). A maximum principle for relaxed stochastic control of linear sde's with application to bond portfolio optimization. *Math. Methods in Operations Research, 72(2), 273-310.*

Bensoussan, A., Sung, K.C.J., Yam, S.C.P., and Yung, S.P. (2011). Linear-quadratic mean field games.

Buckdahn, R., Djehiche, B., and Li, J. (2011). A general stochastic maximum principle for sdes of mean-field type. *Applied Math. and Optimization, 64(2), 197-216.*

Buckdahn, R., Li, J., and Peng, S. (2009). Mean-field backward stochastic differential equations and related partial differential equations. *Stoch. Process. Appl. 119(10), 3133-3154.*

Kamgarpour, M. and Tembine, H. (2013). A bayesian mean field game approach to supply demand analysis of the smart grid. *In Proceedings International Black Sea Conference on Communications and Networking.*

Petrosjan, L.A. (1977). Stability of the solutions in differential games with several players. russian) vestnik leningrad. univ. *no. 19 Mat. Mat. Meh. Astronom. vyp. 4, 4652, 147.*

Tembine, H. (2012). Distributed strategic learning for wireless engineers. *CRC Press, Taylor & Francis, 496 pages, ISBN: 1439876371.*

Yeung, D.W. and Petrosyan, L.A. (2006). Cooperative stochastic differential games. *Springer Series in Operations Research and Financial Engineering, Springer pp. 242.*