

Risk-Sensitive Mean-Field Stochastic Differential Games

Hamidou Tembine* Quanyan Zhu** Tamer Başar**

* *École Supérieure d'Électricité (SUPÉLEC), France. (e-mail: tembine@ieee.org)*

** *Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL, USA. (e-mail: {zhu31, basar1}@illinois.edu)*

Abstract:

In this paper, we study a class of risk-sensitive mean-field stochastic differential games. Under regularity assumptions, we use results from standard risk-sensitive differential game theory to show that the mean-field value of the exponentiated cost functional coincides with the value function of a Hamilton-Jacobi-Bellman-Fleming (HJBF) equation with an additional quadratic term. We provide an explicit solution of the mean-field best response when the instantaneous cost functions are log-quadratic and the state dynamics are affine in the control. An equivalent mean-field risk-neutral problem is formulated and the corresponding mean-field equilibria are characterized in terms of backward-forward macroscopic McKean-Vlasov equations, Fokker-Planck-Kolmogorov equations and HJBF equations.

Keywords: Risk-Sensitive Games, Mean-Field Analysis, Stochastic Differential Games.

1. INTRODUCTION

Most formulations of mean-field models such as anonymous sequential population games Jovanovic and Rosenthal (1988); Bergin and Bernhardt (1992), mean-field (MF) stochastic controls Huang et al. (2003, 2007); Yin et al. (2010), mean-field optimization, mean-field teams Tembine et al. (2009), mean-field stochastic games Weintraub et al. (2005); Adlakha et al. (2008); Tembine et al. (2009); Tembine (2011), mean-field stochastic differential games Lasry and Lions (2007); Guéant et al. (2010); Tembine et al. (2010) have been of risk-neutral type where the cost (or payoff, utility) functions to be minimized (or to be maximized) are the expected values of the stage-additive loss functions.

Not all behavior, however, can be captured by risk-neutral cost functions. One way of capturing risk-seeking or risk-averse behavior is by exponentiating instantaneous loss functions before expectation (see Başar (1999); Jacobson (1973) and the references therein).

The particular risk-sensitive mean-field stochastic differential game that we consider in this paper involves an exponential in the long-term cost function. This approach was first taken by Jacobson (1973), when considering the risk-sensitive Linear-Quadratic-Gaussian (LQG) problem with state feedback. Jacobson demonstrated a link between the exponential cost criterion and deterministic linear-quadratic differential games. He showed that the risk-sensitive approach provides a method for varying the robustness of the controller and noted that in the case of no risk, or risk-neutral case, the well known LQR

* This work was supported in part by Grant AFOSR MURI FA 9550-10-1-0573.

solution would result (see, for follow-up work on risk-sensitive stochastic control problems with noisy state measurements, Whittle (1981); Bensoussan and van Schuppen (1985); Pan and Başar (1996)).

In this paper, we examine the risk-sensitive stochastic differential game in the context of large population of players. We first present a mean-field stochastic differential game model where the players are coupled not only via their risk-sensitive cost functionals but also via their states. The main coupling term is the mean-field process also called *occupancy process* or *population profile process*.

Our contribution can be summarized as follows. Using a particular structure of state dynamics, we derive the mean-field limit of the individual state dynamics leading to the non-linear controlled macroscopic McKean-Vlasov equation; see Kotolenez and Kurtz (2010). Combining this together with the convergence of the risk-sensitive cost functional, we arrive at the mean-field optimality principle, and their compatibility with the density distribution are obtained using the Fokker-Planck-Kolmogorov forward equation. The mean-field equilibria are characterized by coupled backward-forward equations which may not have a solution in general (a simple example is provided in section 4.2). Explicit solution of the Hamilton-Jacobi-Bellman equation is provided for the affine-exponentiated-Gaussian mean-field problem. Finally, an equivalent risk-neutral mean-field problem is formulated and the solution of the mean-field response problem is explicitly given.

The rest of the paper is organized as follows. In Section 2, we present the model description. We provide an overview of the mean-field convergence result in Section 3. In Section 4, we present the risk-sensitive mean-field stochastic

differential game formulation and its equivalences. In Section 5, we illustrate with a numerical example. Section 6 concludes the paper.

2. THE PROBLEM SETTING

We consider a class of n -person stochastic differential games, where Player j 's individual states evolve according to the Itô stochastic differential equations (S):

$$(S) \left\{ \begin{array}{l} dx_j^n(t) = \frac{1}{n} \sum_{i=1}^n f_t(x_j^n(t), u_j^n(t), x_i^n(t)) dt \\ + \frac{\sqrt{\epsilon}}{n} \sum_{i=1}^n \sigma_t(x_j^n(t), u_j^n(t), x_i^n(t)) d\mathbb{B}_j(t), \\ x_j^n(0) = x_{j,0} \in \mathcal{X} \subseteq \mathbb{R}^k, \quad k \geq 1, j \in \{1, \dots, n\}, \end{array} \right.$$

where x_j^n is the k -dimensional state of player j ; $u_j^n(t) \in \mathcal{U}_j$, is the control of player j at time t with \mathcal{U}_j being a subset of p_i -dimensional Euclidean space \mathbb{R}^{p_i} ; $\mathbb{B}_j(t)$ are mutually independent standard Brownian motions in \mathbb{R}^k ; and ϵ is a small positive parameter, which will play a role in the analysis in the later sections.

We have to specify the nature of information that players are allowed in the choice of their control at each time. A *state-feedback strategy* for Player j is a mapping $\tilde{u}_j : \mathbb{R}_+ \times (\mathbb{R}^k)^n \rightarrow \mathcal{U}_j$, whereas an *individual state-feedback strategy* for Player j is a mapping $\bar{u}_j : \mathbb{R}_+ \times \mathbb{R}^k \rightarrow \mathcal{U}_j$. Hence the latter involves only the self state of a player, whereas the former involves the entire nk -dimensional state vector. The individual strategy sets in either case have to be chosen in such a way that the resulting system of stochastic differential equations (S) admits a unique solution when the players pick their strategies independently; furthermore, the sets should be invariant under concatenation of controls and translation of time. We denote by \bar{U}_j the set of such admissible control laws $\bar{u}_j : [0, T] \times \mathbb{R}^k \rightarrow \mathcal{U}_j$ for Player j ; a similar set can be defined for state-feedback strategies.

Now, normally, when we have a cost function for Player j which depends also on the state variables of the other players, either directly, or implicitly through the coupling of the state dynamics (as in (S)), then any state feedback Nash equilibrium solution will surely depend not only on self states but also on the entire state vector, i.e., it will not be in the set $\bar{U}_j, j = 1, \dots, n$. However, what we are interested in this paper is the solution in the high-population regime (i.e., as $n \rightarrow \infty$) in which case the dependence on other players' states will be through the distribution of the player states. Hence each player will respond (in an optimal cost minimizing manner) to the behavior of the mass population and not to behaviors of individual players. Validity of this property will be established later in Section 3 of the paper, but in anticipation of this, we first introduce the quantity

$$m_t^n = \frac{1}{n} \sum_{j=1}^n \delta_{x_j^n(t)}, \quad (1)$$

as an empirical measure of the collection of states of the players, where δ is a Dirac measure on the state space. This then enables us to introduce the long-term cost function of Player j (to be minimized by him) in terms of only the self

variables (x_j and u_j) and $m_t^n, t \geq 0$, where the latter can be viewed as an exogenous process (not directly influenced by Player j). This cost function will be of the risk-sensitive type, and given by, for each $t \in [0, T]$, and each m that initializes m_t^n at t :

$$L_j(\bar{u}_j, m_{[t,T]}^n; t, x, m) =$$

$$\delta \log \mathbb{E} \left(e^{\frac{1}{\delta} [g(x_T) + \int_t^T c_s(x_s, u_s, m_s^n) ds]} \mid x_j(t) = x, m_t^n = m \right),$$

where $\delta > 0$ is the risk-sensitivity index and $m_{[t,T]}^n$ denotes the process $\{m_s^n, t \leq s \leq T\}$. This cost function is called the risk-sensitive cost functional or the exponentiated integral cost, see Jacobson (1973); Whittle (1981); Bensoussan and van Schuppen (1985); Başar (1999). We assume the following standard conditions on f_t, g_T, c_t .

- (i) f_t, c_t are C^1 in (t, x, u, m) ; g is C^2 in x ; c_t, g are non-negative;
- (ii) The entries of the matrix σ_t are C^2 and $\sigma_t \sigma_t'$ is strictly positive;
- (iii) $f_t, \partial_x f_t, c_t, \partial_x c_t$ are uniformly bounded; $g, \partial_x g$ are uniformly bounded;
- (iv) \mathcal{U}_j is closed and bounded;
- (v) $\bar{u}_j : [0, T] \times \mathbb{R}^k \rightarrow \mathcal{U}_j$ is piecewise continuous in t and Lipschitz in x .

The stochastic differential game problem with logarithm of the expected *exponentiated integral cost* function is called the risk-sensitive stochastic differential game. Note that if x and m^n are deterministic, then the above cost function reduces to the standard additive cost function.

With the dynamics (S) in the large population regime (see later (SM) in the next section) and cost functionals as introduced, we seek an individual state-feedback non-cooperative Nash equilibrium $\{\bar{u}_i^*, i = 1, \dots, n\}$, satisfying the set of inequalities

$$L_j(\bar{u}_j^*, m_{[0,T]}^n; 0, x_{j,0}, m) \leq L_j(\bar{u}_j, m_{[0,T]}^n; 0, x_{j,0}, m) \quad (2)$$

for all $u_j \in \bar{U}_j, j = 1, 2, \dots, n$, or the strongly time-consistent individual state-feedback equilibrium,

$$L_j(\bar{u}_j^*, m_{[t,T]}^n; t, x_j, m) \leq L_j(\bar{u}_j, m_{[t,T]}^n; t, x_j, m) \quad (3)$$

for all $x_j \in \mathcal{X}, t \in [0, T], u_j \in \bar{U}_j, j = 1, 2, \dots, n$.

3. MEAN-FIELD ANALYSIS: AN OVERVIEW

3.1 Mean field representation

The system (S) can be written into a measure representation by the formula

$$\int \phi(w) \left[\sum_{i=1}^n \bar{\omega}_i \delta_{x_i} \right] (dw) = \sum_{i=1}^n \bar{\omega}_i \phi(x_i)$$

where $\delta_z, z \in \mathcal{X}$ is a Dirac measure on set \mathcal{X} , ϕ is a measurable bounded function defined on the state space and $\bar{\omega}_i \in \mathbb{R}$. Then, the system (S) reduces to the system

$$\left\{ \begin{array}{l} dx_j^n = \left(\int_w f_t(x_j^n(t), u_j^n(t), w) \left[\frac{1}{n} \sum_{i=1}^n \delta_{x_i^n(t)} \right] (dw) \right) dt + \\ \sqrt{\epsilon} \left(\int_w \sigma_t(x_j^n(t), u_j^n(t), w) \left[\frac{1}{n} \sum_{i=1}^n \delta_{x_i^n(t)} \right] (dw) \right) d\mathbb{B}_j(t), \\ x_j^n(0) = x_{j,0} \in \mathbb{R}^k, \quad k \geq 1, j \in \{1, \dots, n\}, \end{array} \right.$$

which, by (1), is equivalent to

$$(SM) \begin{cases} dx_j^n(t) = \left(\int_w f_t(x_j^n(t), u_j^n(t), w) m_t^n(dw) \right) dt \\ + \sqrt{\epsilon} \left(\int_w \sigma_t(x_j^n(t), u_j^n(t), w) m_t^n(dw) \right) d\mathbb{B}_j(t), \\ x_j^n(0) = x_{j,0} \in \mathbb{R}^k, \quad k \geq 1, j \in \{1, \dots, n\}. \end{cases}$$

As we will see in the next sections, the system (SM) can be seen as a controlled interacting particles representation of a macroscopic McKean-Vlasov equation where m_t^n represents the discrete density of the population. Next, we focus on the convergence of the population profile process m^n .

3.2 Mean field convergence

The above system has the structure satisfying the asymptotic indistinguishability conditions of Tanabe (2006) under suitable controls. This is equivalent to the existence of a random measure μ such that the system is μ -chaotic, i.e.

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

for any fixed natural number $L \geq 2$ and a collection of measurable bounded functions $\{\phi_l\}_{1 \leq l \leq L}$. Following the indistinguishability property, one has that the law of $x_j^n = (x_j^n(t), t \geq 0)$ is $\mathbb{E}m^n$. The same result is obtained by proving the weak convergence of the individual state dynamics to a macroscopic McKean-Vlasov equation. Then, for an initial i.i.d. condition and under the controls u^* , the solution of the state dynamics generates an indistinguishable sequence and the weakly convergence of the population profile process m^n to μ is equivalent to the μ -chaoticity.

These processes depend implicitly on the strategies used by the players. Note that the control \bar{u}^* may depend on time t , the value of the individual state $x_j(t)$ and the mean-field process m_t . The weak convergence of the process m^n implies the weak convergence of its marginal m_t^n and one can characterize the distribution of m_t by the Fokker-Planck-Kolmogorov (FPK) equation:

$$\begin{aligned} \partial_t m_t + D_x^1 \left(m_t \int_w f_t(x, u^*(t), w) m_t(dw) \right) \\ = \frac{\epsilon}{2} D_{xx}^2 \left(m_t \left(\int_w \sigma_t'(x, u^*(t), w) m_t(dw) \right) \right. \\ \left. \left(\int_w \sigma_t(x, u^*(t), w) m_t(dw) \right) \right). \end{aligned} \quad (4)$$

Here $f_t(\cdot) \in \mathbb{R}^k$ which we denote by $(f_{k',t}(\cdot))_{1 \leq k' \leq k}$. Let

$$\underline{\sigma}_t[x, u^*(t), m_t] = \int_w \sigma_t(x, u^*(t), w) m_t(dw),$$

$\Gamma_t(\cdot) := \underline{\sigma}_t(\cdot) \underline{\sigma}_t'(\cdot)$ is a square matrix with dimension $k \times k$. The term $D_x^1(\cdot)$ denotes

$$\sum_{k'=1}^k \frac{\partial}{\partial x_{k'}} \left(m_t \int_w f_{k',t}(x, u^*(t), w) m_t(dw) \right),$$

and the last term on $D_{xx}^2(\cdot)$ is

$$\sum_{k''=1}^k \sum_{k'=1}^k \frac{\partial^2}{\partial x_{k'} \partial x_{k''}} (m_t \Gamma_{k'k'',t}(\cdot)).$$

In the one-dimensional case, the terms D^1, D^2 reduce to the divergence div and the Laplacian operator Δ , respectively.

It is important to notice that the existence of a unique rest point (distribution) in FPK does not automatically imply that the mean-field converges to the rest point when t goes to infinity. This is because the rest point may not be stable.

In statistical mechanics, convergence to an independent and identically distributed system is called *chaoticity*, and the fact that chaoticity at the initial time implies chaoticity at further times is called propagation of chaos. In general the chaoticity property may not hold. We need to mention a particular case where the rest point m^* is related to the δ_{m^*} -chaoticity. If the mean-field dynamics has a unique global attractor m^* 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 $\lim_n \lim_t m_t^n$ may differ from $\lim_t \lim_n m_t^n$ leading a non-commutative diagram. Thus, a deep study of the dynamical system is required if one wants to analyze a performance metric for a stationary regime.

4. RISK-SENSITIVE BEST RESPONSE TO MEAN-FIELD

In this section, we present the risk-sensitive mean-field results. We first provide an overview of the optimality criterion for a given mean-field trajectory $m^n = (m^n(s), s \geq 0)$. A mean-field best-response strategy of a generic player j is a measurable mapping \bar{u}_j^* satisfying: $\forall \bar{u}_j \in \bar{\mathcal{U}}_j$,

$$L_j^n(\bar{u}_j^*, m_{[0,T]}^n, 0, x_{j,0}, m) \leq L_j^n(\bar{u}_j, m_{[0,T]}^n, 0, x_{j,0}, m).$$

Let $v_j^n(t, x_j, m) = \inf_{u_j} L_j^n(u_j, m_{[0,T]}^n, t, x_j, m)$. The next proposition establishes the Hamilton-Jacobi-Bellman-Fleming (HJBF) equation of the risk-sensitive cost function satisfied by a regular optimal value function of a generic player. The main difference from the standard HJBF is the presence of the term $\frac{\epsilon}{2\delta} \|\sigma_t \partial_{x_j} v_j^n\|^2$.

Proposition 1. Suppose that the trajectory of m_t^n is given. If v_j^n is twice continuously differentiable, then v_j^n is solution of the HJBF equation

$$\begin{aligned} \partial_t v_j^n + \inf_{u_j} \left\{ \partial_{x_j} v_j^n \cdot f_t + \frac{\epsilon}{2} \text{tr}(\sigma_t \sigma_t' \partial_{x_j x_j}^2 v_j^n) \right. \\ \left. + \frac{\epsilon}{2\delta} \|\sigma_t \partial_{x_j} v_j^n\|^2 + c_t \right\} = 0, \end{aligned} \quad (5)$$

$$v_j^n(T, x_j) = g(x_j)$$

Moreover, any strategy satisfying

$$\begin{aligned} u_j^n(t) \in \arg \min_{u_j} \left\{ \partial_{x_j} v_j^n \cdot f_t + \frac{\epsilon}{2} \text{tr}(\sigma_t \sigma_t' \partial_{x_j x_j}^2 v_j^n) \right. \\ \left. + \frac{\epsilon}{2\delta} \|\sigma_t \partial_{x_j} v_j^n\|^2 + c_t \right\} \end{aligned}$$

constitutes a best response strategy to the mean-field m^n .

The next proposition provides an explicit solution to the affine-quadratic-exponentiated cost-Gaussian mean-field game

Proposition 2. Suppose $\sigma(t, x) = \sigma_t := \sigma(t)$ and

$$f_t(x_j, u_j, m) = \bar{f}(t, x_j, m) + B(t, x_j, m)u_j,$$

$$c_t(x_j, u_j, m) = \bar{c}(t, x_j, m) + \|u_j\|^2,$$

then the optimal control of Player j is $u_j^{n,*} = -\frac{1}{2} B \partial_{x_j} v_j^n$.

Proposition 3. (Explicit optimal cost, Başar (1999)). Consider the risk-sensitive mean-field stochastic game described in Section 2 with $\bar{f} = A_t x$, $c_t = x' Q_t x$, $Q_t \geq 0$, $g(x) = x' Q_T x$, $Q_T \geq 0$, with continuous matrix (in time). Then, the solution (whenever it exists) is given by $v_j^n(t, x_j) = x' z_t x + \epsilon \int_t^T \text{tr}(z_s \sigma_s \sigma_s') ds$, where z_s is the nonnegative definite solution of the generalized Riccati differential equation

$$\dot{z}_t + A_t' z_t + z_t A_t + Q_t - z_t \left(B_t B_t' - \frac{1}{\gamma^2} \sigma_t \sigma_t' \right) z_t = 0, \quad z_T = Q_T,$$

where $\gamma = (\frac{\delta}{2\epsilon})^{1/2}$ and the optimal response strategy is

$$u_j^*(t) = -B_t' z_t x. \quad (6)$$

Using Proposition 3, one has the following result for any given trajectory $(m_t^n)_{t \geq 0}$.

Proposition 4. If c_t is in the form $c_t = x'(Q_t - \Lambda_t(m_t^n))x$, i.e., a function of m_t^n , then the generalized Riccati equation becomes

$$\dot{z}_t^* + A_t' z_t^* + z_t^* A_t + Q_t - \Lambda_t(m_t^n) - z_t^* \left(B_t B_t' - \frac{1}{\gamma^2} \sigma_t \sigma_t' \right) z_t^* = 0.$$

and $v_j^n(t, x_j) = x' z_t^* x + \epsilon \int_t^T \text{tr}(z_s^* \sigma_s \sigma_s') ds$.

4.1 Risk-sensitive McKean-Vlasov equation

Since the control used by the players influence the mean-field limit via the state dynamics, we need to characterize the evolution of the mean-field limit as a function of the controls. The law of m_t is the solution of the Fokker-Planck-Kolmogorov equation given by (4) and the individual state dynamics follows the so-called macroscopic McKean-Vlasov equation

$$\begin{cases} d\bar{x}_j(t) = \left(\int_w f_t(\bar{x}_j(t), u_j^*(t), w) m_t(dw) \right) dt \\ + \sqrt{\epsilon} \left(\int_w \sigma_t(\bar{x}_j(t), u_j^*(t), w) m_t(dw) \right) d\mathbb{B}_j(t). \end{cases}$$

In order to obtain an error bound we introduce the following notions: Given two measures μ, ν the Monge-Kantorovich metric (also called Wasserstein metric) between μ and ν is

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

In other words, let $E(\mu, \nu)$ be the set of probability measures \mathbb{P} on $\mathcal{X}^{\mathbb{R}} \times \mathcal{X}^{\mathbb{R}}$ such that the image of \mathbb{P} under the projection on the first factor (resp. on the second factor) is μ (resp. ν). Then,

$$\mathcal{W}_1(\mu, \nu) = \inf_{\mathbb{P} \in E(\mu, \nu)} \int \int |z - z'| \mathbb{P}(dz, dz')$$

This is known as a distance (it can be checked that the separation, the triangular inequality and positivity properties are satisfied) and it metrizes the weak topology of $\mathcal{X}^{\mathbb{R}}$.

Then, the following holds: For any t , if the control law $u_j^*(t)$ is used then there exists $\tilde{y}_t > 0$ such that

$$\mathbb{E}(\|x_j^n(t) - \tilde{x}_j(t)\|) \leq \frac{\tilde{y}_t}{\sqrt{n}}.$$

Moreover, for any $T < \infty$, there exists $C_T > 0$ such that $\mathcal{W}_1(\mathcal{L}((x_j^n(t))_{t \in [0, T]}), \mathcal{L}((\tilde{x}_j(t))_{t \in [0, T]})) \leq \frac{C_T}{\sqrt{n}}$.

The last inequality says that the error bound is at most of $O(\frac{1}{\sqrt{n}})$ for any fixed compact interval.

Convergence of the risk-sensitive mean-field cost: Using the fact that M^n converges weakly to m under suitable controls $(u_j^n(s), s \geq 0) \rightarrow (u_s, s \geq 0)$ when n goes to infinity, one can derive a weak convergence of the risk-sensitive cost function as stated in the following Proposition.

Proposition 5. (Cost at the limit). The risk-sensitive cost functional $L_j^n(\bar{u}_j^n, m^n[t, T], t, x, m)$ converges to $L_j(\bar{u}_j, m_{[t, T]}, 0, x, m)$ given by

$$\delta \log \mathbb{E} \left(e^{\frac{1}{\delta} [g(x_j(T)) + \int_t^T c_s(x_j(s), u_j(s), m_s) ds]} \Big| x_j(t) = x, m_t = m \right).$$

The proof follows from the weak convergence of m^n and the regularity of the functions c and g . Using this lemma, we are able to construct the mean-field optimal control at the limit. Given $\{m_s\}_{s \in [t, T]}$, we aim to minimize $L_j(\bar{u}_j, m_{[t, T]}, 0, x, m)$ subject to the state-dynamics constraints.

4.2 Risk-sensitive FPK-McV equations

The mean-field optimality criterion leads to *HJBF backward equation combined with FPK equation and macroscopic McKean-Vlasov version of the limiting individual dynamics*, i.e.,

$$\left\{ \begin{aligned} dx_j(t) &= \left(\int_w f_t(x_j(t), u_j^*(t), w) m_t(dw) \right) dt \\ &+ \sqrt{\epsilon} \left(\int_w \sigma_t(x_j(t), u_j^*(t), w) m_t(dw) \right) d\mathbb{B}_j(t), \\ x_j(0) &= x_{j,0} = x \\ \partial_t v_j + \inf_{u_j} \left\{ \partial_{xx} v_j \cdot f_t + \frac{\epsilon}{2} \text{tr}(\sigma_t \sigma_t' \partial_{xx}^2 v_j) \right. \\ &\quad \left. + \frac{\epsilon}{2\delta} \|\sigma_t \partial_x v_j\|^2 + c_t \right\} = 0, \\ x_j &:= x; v_j(T, x) = g(x) \\ \partial_t m_t + D_x^1 \left(m_t \int_w f_t(x, u^*, w) m_t(dw) \right) \\ &= \frac{\epsilon}{2} D_{xx}^2 \left(m_t \left(\int_w \sigma_t'(x, u^*, w) m_t(dw) \right) \right. \\ &\quad \left. \left(\int_w \sigma_t(x, u^*, w) m_t(dw) \right) \right) \\ &\quad m_0(\cdot) \text{ fixed.} \end{aligned} \right.$$

Then, the question of existence of a solution to the above system arises. This is a backward-forward system. Very little is known about the existence of a solution of such a system. In general a solution may not exist. Next we provide a non-solvability example.

4.3 The backward-forward boundary problem may not have a solution

There are many examples of systems of backward-forward equations which do not admit solutions. Here is a very simple one: $\dot{v} = m$, $\dot{m} = -v$, $m(0) = m_0$; $v_T = -m_T$.

It is obvious that the coefficients of this pair of backward-forward differential equations are all uniformly Lipschitz. However, we claim that depending on T , this may not

be solvable for $m_0 \neq 0$. We can easily show that for $T = k\pi + 3\pi/4$ (k , a nonnegative integer), the above two-point boundary value problem does not admit a solution for any $m_0 \neq 0$ and it admits infinitely many solutions for $m_0 = 0$.

Following the same ideas, one can show that the system of SDEs $dv = mdt + \sigma dB_t$, $dm = -vdt + \nu dB_t$, with the initial conditions: $m(0) = m_0 \neq 0$; $v_T = -m_T$, and $T = 7\pi/4$ has no solution.

4.4 Risk-sensitive mean-field equilibria

Theorem 6. Consider a risk-sensitive mean-field stochastic differential game as formulated above. Assume that the $\sigma_t(\cdot) = \sigma(t)$ and there exists a unique pair (u^*, m^*) such that

(i) The coupled backward-forward PDEs

$$\begin{cases} \partial_t v^* + \inf_{u_j} \left\{ \partial_x v^* \cdot f_t^* + \frac{\epsilon}{2} \text{tr}(\sigma_t \sigma_t' \partial_{xx}^2 v^*) \right. \\ \quad \left. + \frac{\epsilon}{2\delta} \|\sigma_t \partial_x v\|^2 + c_t^* \right\} = 0, \\ v(T, x) = g(x), m_0^*(x) \text{ fixed} \\ \partial_t m_t^* + D_x^1 \left(m_t^* \int_w f_t^*(x, u^*, w) m_t^*(dw) \right) \\ \quad = \frac{\epsilon}{2} D_{xx}^2 \left(m_t^* \left(\int_w \sigma_t' m_t^*(dw) \right) \left(\int_w \sigma_t m_t^*(dw) \right) \right) \end{cases}$$

admit a pair a bounded nonnegative solutions v^*, m^* ;

(ii) $u_t^*(x)$ minimizes $\partial_x f_t + c_t$.

Under these conditions, the pair (u^*, m^*) is a mean-field equilibrium and $L_j(t, u^*, m^*) = v^*$.

If $c_t = x'(Q_t - \Lambda_t(m_t^n))x$, then any convergent subsequence of optimal control law $u_j^{\alpha(n)}$ leads to a best strategy of m . Moreover, for any $\epsilon' > 0$ there exists $n_{\epsilon'}$ such that for all $n \geq n_{\epsilon'}$, u^* is an ϵ -Nash equilibrium of the differential game with size n .

This result can be extended to multiple classes of players.

Sensitivity of the perturbation ϵ : We scale the parameters δ, ϵ and γ such that $\delta = 2\epsilon\gamma^2$. The PDE given from Proposition 1 becomes

$$\begin{cases} \partial_t v + \inf_u \left\{ \partial_x v \cdot f_t^* \right. \\ \quad \left. + \frac{\epsilon}{2} \text{tr}(\sigma_t \sigma_t' \partial_{xx}^2 v) + \frac{1}{4\gamma^2} \|\sigma_t \partial_x v\|^2 + c_t \right\} = 0, \\ v(T, x) = g(x). \end{cases}$$

When the parameter ϵ goes to zero one gets a deterministic PDE. This situation captures the large deviation limit:

$$\begin{cases} \partial_t v + \inf_u \left\{ \partial_x v \cdot f_t^* + \frac{1}{4\gamma^2} \|\sigma_t \partial_x v\|^2 + c_t^* \right\} = 0, \\ v(T, x) = g(x). \end{cases}$$

4.5 Equivalent stochastic mean-field problem

In this subsection, we formulate an equivalent $(n+1)$ -player game in which the state dynamics of the n players are given by

$$(SM) \begin{cases} dx_j^n(t) = \left(\int_w f_t(x_j^n(t), u_j^n(t), w) m_t^n(dw) \right) dt \\ \quad + \sigma_t \zeta_t + \sqrt{\epsilon} \sigma_t d\mathbb{B}_j(t), \\ x_j^n(0) = x_{j,0} \in \mathbb{R}^k, k \geq 1, j \in \{1, \dots, n\}, \end{cases}$$

where ζ_t is the control parameter of the “fictitious” $(n+1)$ -th player. We define the risk-neutral cost function of the n players as $\tilde{L}_j^n(\bar{u}_j, \bar{\zeta}, x_j^n, m^n; t, x, m)$

$$\begin{aligned} &= \mathbb{E} \left(g(x_{j,T}^n) + \int_t^T c_s(x_j^n(s), u_j^n(s), m_s^n) ds \right. \\ &\quad \left. - \gamma^2 \int_t^T \|\zeta_s\|^2 ds \mid x_j(t) = x, m_t^n = x \right) \end{aligned}$$

Every player $j \in \{1, 2, \dots, n\}$ minimizes L_j by taking the worst over the feedback strategy $\bar{\zeta}$ of player $n+1$ which is piecewise continuous in t and Lipschitz in x_j . Let $\tilde{v}_j^n = \inf_{\bar{u}_j} \sup_{\bar{\zeta}} \tilde{L}_j^n(\bar{u}_j, \bar{\zeta}, x_j^n, m^n, t, x_j, m_0)$. We refer to this as *robust mean-field game*.

Proposition 7. Under the regularity assumptions (i-v), \tilde{v}_j^n satisfies the Hamilton-Jacobi-Isaacs (HJI) equation

$$\begin{aligned} \inf_u \sup_{\zeta} \left\{ \partial_t \tilde{v}_j^n + \partial_{x_j} \tilde{v}_j^n (f_t + \sigma_t \zeta_t) + c_t - \gamma^2 \|\zeta_t\|^2 \right. \\ \left. + \frac{\epsilon}{2} \text{tr}(\partial_{x_j x_j}^2 \tilde{v}_j^n \sigma_t \sigma_t') \right\} = 0, \tilde{v}_j^n(T, x_j) = g(x_j). \end{aligned}$$

The proof of Proposition 7 can be obtained by using results from Başar (1999). Since the dependence on u and ζ above are separable, the Isaacs condition holds i.e., $\inf \sup = \sup \inf$ and, hence the function \tilde{v}_j^n is solution of the partial differential equation:

$$\begin{aligned} -\partial_t \tilde{v}_j^n &= \partial_{x_j} \tilde{v}_j^n \cdot f_t + c_t + \frac{1}{4\gamma^2} \|\sigma_t' \partial_{x_j} \tilde{v}_j^n\|^2 \\ &\quad + \frac{\epsilon}{2} \text{tr}(\partial_{x_j x_j}^2 \tilde{v}_j^n \sigma_t \sigma_t') \end{aligned} \quad (7)$$

Hence, the two PDEs: (7) and the one given in Lemma 1 are identical. Moreover, the optimal cost and the optimal control law of the two problems are the same. The robust mean-field LQG game satisfies the properties of Theorem 6. Hence, the following result follows:

Proposition 8. Consider the risk-sensitive mean-field stochastic game with the logarithm of expectation of the exponentiated-integral cost functional. Then, the risk-sensitive mean-field stochastic game is equivalent to a risk-neutral mean-field game with cost functional \tilde{L}_j^n (resp. \tilde{L}_j at the limit).

5. NUMERICAL ILLUSTRATION

In this section, we illustrate the risk-sensitive mean-field game with a numerical example. We let Player j 's state evolution be described by a decoupled stochastic differential equation (SE)

$$dx_j^n = u_j dt + \sqrt{\epsilon} \sigma d\mathbb{B}_j(t) \quad (8)$$

The risk-sensitive cost functional is given by

$$\begin{aligned} L_j(\bar{u}_j, m^n; t, x_j, m) &= \delta \log \mathbb{E} \left\{ \exp \left[\frac{1}{\delta} \left(Q_j(x_j^n)^2 \right. \right. \right. \\ &\quad \left. \left. + \int_0^T (q_j - \mathbb{E}(m_t^n))(x_j^n)^2(t) + \bar{u}_j^2(t) dt \right) \right] \right\}, \end{aligned} \quad (9)$$

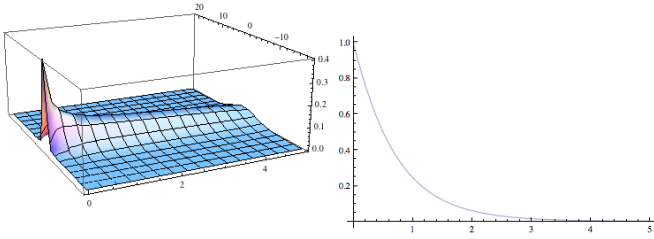


Fig. 1. The evolution of distribution $m_t^*(x), 0 \leq t \leq 5, -19 \leq x \leq 21$. Fig. 2. Mean value $\mathbb{E}(m_t^*)$ as a function of time, $0 \leq t \leq 5$.

where δ, Q_j, q_j are positive parameters. The optimal strategy of Player j has the form of

$$\bar{u}_j^*(t) = -z_t x, \quad (10)$$

where z_t is a solution to the Riccati equation

$$\dot{z}_t + q_j - \mathbb{E}(m^n) - z_t^2(1 - \sigma_j^2/\gamma^2) = 0, \quad (11)$$

with boundary condition $z_T = Q_j$. An explicit solution is given by

$$z_t = -\frac{\sqrt{q_j - M}}{\sqrt{L}} \tan \left[\sqrt{L} \sqrt{q_j - M} (t - T) + \arctan \left(\frac{\sqrt{L} Q}{\sqrt{q_j - M}} \right) \right], 0 \leq t \leq T, \quad (12)$$

where $L := 1 - \sigma_j^2/\gamma^2$ and $M := \mathbb{E}(m^n)$. The FPK-McV equation reduces to

$$\partial_t m_t^* + \partial_x m_t^* z_t x = \frac{\epsilon}{2} \sigma_j^2 \partial_{xx}^2 m_t^*. \quad (13)$$

We set the parameters as follows: for all $j, q_j = 1.2, Q_j = 0.1, \delta = 100,000, \sigma = 2.0, T = 5$ and $\epsilon = 5.0$. Let $m_0^*(x)$ be a normal distribution $\mathcal{N}(1, 1)$ and for every $0 \leq t \leq T, m_t^*$ vanishes at the infinities. In Figure 1, we show the evolution of the distribution m_t^* and in Figures 2 and 3, we show the mean and the variance of the distribution which affects the optimal strategies in (10). The optimal linear feedback z_t is illustrated in Figure 4. We can observe that the mean value $\mathbb{E}(m_t^*)$ monotonically decreases from 1.0 and hence the unit cost on state is monotonically increasing. As the state cost increases, the control effort becomes relatively cheaper and therefore we can observe an increment in the magnitude of z_t . However, when the mean value goes beyond 1.08, we observe that the control effort reduces to avoid undershooting in the state.

6. CONCLUDING REMARKS

We have studied risk-sensitive mean-field stochastic differential games with state dynamics given by an Itô stochastic differential equation and exponentiated cost function. An interesting direction that we leave for future work is the extension of some of these results to the time average risk-sensitive cost functional criterion.

REFERENCES

Adlakha, S., Johari, R., Weintraub, G., and Goldsmith, A. (2008). Oblivious equilibrium for large-scale stochastic games with unbounded costs. *Proc. IEEE CDC, Cancun, Mexico*, 5531–5538.
Başar, T. (1999). Nash equilibria of risk-sensitive nonlinear stochastic differential games. *J. of Optimization Theory and Applications*, 100(3), 479–498.

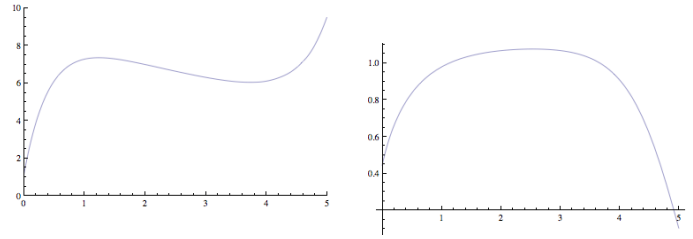


Fig. 3. Variance of the distribution m_t^* as a function of time, $0 \leq t \leq 5$. Fig. 4. z_t as a function of time, $0 \leq t \leq T$.

Bensoussan, A. and van Schuppen, J. (1985). Optimal control of partially observable stochastic systems with an exponential-of-integral performance index. *SIAM J. Control and Optimization*, 23, 599–613.
Bergin, J. and Bernhardt, D. (1992). Anonymous sequential games with aggregate uncertainty. *J. Mathematical Economics*, 21, 543–562.
Guéant, O., Lasry, J.M., and Lions, P.L. (2010). Mean field games and applications. *Springer: Paris-Princeton Lectures on Mathematical Finance*, Eds. René Carmona, Nizar Touzi.
Huang, M., Caines, P.E., and Rolan P., M. (2007). Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ -Nash equilibria. *IEEE Trans. Automat. Control*, 52, 1560–1571.
Huang, M.Y., Caines, P.E., and Malhamé, R.P. (2003). Individual and mass behaviour in large population stochastic wireless power control problems : Centralized and Nash equilibrium solution. *IEEE Conference on Decision and Control, HA, USA*, 98 – 103.
Jacobson, D. (1973). Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Trans. Automat. Contr.*, 18(2), 124–131.
Jovanovic, B. and Rosenthal, R.W. (1988). Anonymous sequential games. *Journal of Mathematical Economics*, 17, 77–87.
Kotolenez, P. and Kurtz, T. (2010). Macroscopic limits for stochastic partial differential equations of McKean-Vlasov type. *Probability theory and related fields*, 146(1), 189–222.
Lasry, J. and Lions, P. (2007). Mean field games. *Japan. J. Math.*, 2, 229–260.
Pan, Z. and Başar, T. (1996). Model simplification and optimal control of stochastic singularly perturbed systems under exponentiated quadratic cost. *SIAM J. Control and Optimization*, 34(5), 1734–1766.
Tanabe, Y. (2006). The propagation of chaos for interacting individuals in a large population. *Mathematical Social Sciences*, 51, 125–152.
Tembine, H. (2011). Mean field stochastic games: Convergence, q/h-learning and optimality. *In Proc. American Control Conference (ACC), San Francisco, California, USA*.
Tembine, H., Lasaulce, S., and Jungers, M. (2010). Joint power control-allocation for green cognitive wireless networks using mean field theory. *Proc. 5th IEEE Intl. Conf. on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, 1–5.
Tembine, H., Le Boudec, J.Y., ElAzouzi, R., and Altman, E. (2009). Mean field asymptotic of markov decision evolutionary games and teams. *Proc. International Conference on Game Theory for Networks (GameNets), Istanbul, Turkey, May 13-15, 2009.*, 140–150.
Weintraub, G.Y., Benkard, L., and Van Roy, B. (2005). Oblivious equilibrium: A mean field approximation for large-scale dynamic games. *Advances in Neural Information Processing Systems*, 18.
Whittle, P. (1981). Risk-sensitive linear quadratic gaussian control. *Advances in Applied Probability*, 13, 764–777.
Yin, H., Mehta, P.G., Meyn, S.P., and Shanbhag, U.V. (2010). Synchronization of coupled oscillators is a game. *Proc. American Control Conference (ACC), Baltimore, MD*, 1783–1790.