

# Linear-Quadratic Risk-Sensitive Mean Field Games

Jun Moon and Tamer Başar

**Abstract**—In this paper, we consider linear-quadratic risk-sensitive mean field games (LQRSMFGs). Each agent strives to minimize an exponentiated integral quadratic cost or risk-sensitive cost function, which is coupled with other agents via a mean field term. By invoking the *Nash certainty equivalence principle*, we first obtain a robust decentralized control law for each agent to construct a mean field system. We then provide appropriate conditions under which the mean field system admits a unique deterministic function that approximates the mean field term with arbitrarily small error when the number of agents, say  $N$ , goes to infinity. We also show the closed-loop system stability, and prove that the set of  $N$  robust decentralized control laws possesses an  $\epsilon$ -Nash equilibrium property. Moreover, we show that  $\epsilon$  can be taken to be arbitrarily close to zero as  $N \rightarrow \infty$ , but our  $\epsilon$  bound is weaker than its linear-quadratic mean field game (LQMFG) counterpart due to risk-sensitivity in the present case. Finally, we discuss two different limiting cases, and show that one of these is equivalent to the corresponding LQMFG.

## I. INTRODUCTION

The problem of decentralized control of multi-agent systems has been studied extensively in recent years to analyze the dynamic behavior of agents under different performance criteria. Mean field games are known to be one such subclass, where there is a large number of agents who are coupled with each other through a mean field term that describes the average behavior of all the agents [1]–[3]. A variety of applications on mean field games were considered for engineering, economics, computer science, and social sciences; see [1], [2], [4], [5], and the references therein.

In mean field games, similar to other noncooperative games, the main objective is to obtain a characterization of Nash equilibria under the assumption of rationality for each player [3]. The complexity issue, however, is one to be dealt with, since it increases with the number of players and the dimension of the state space [6]. In addition to complexity, another issue arises because of the decentralized nature of the information available to each agent. That is, the agents may not share all the information, but have access to only local information [3], [6], [7].

In [3] and [8], the Nash certainty equivalence (NCE) principle was proposed for linear and nonlinear mean field games, respectively, to overcome these issues. More precisely, the mean field game was analyzed through two steps: 1) solving a single-agent optimal control problem by replacing the mean field coupling term with a known deterministic

function, and 2) constructing a mean field system to estimate the mass behavior with this deterministic function. It was shown that under some conditions, the overall effect of each agent on the mass behavior can be captured by this deterministic function asymptotically when the number of agents, say  $N$ , becomes infinity. Moreover, the set of all agents' individual optimal control laws possesses an  $\epsilon$ -Nash equilibrium property, where  $\epsilon$  can be chosen to be arbitrarily close to zero as the number of agents, say  $N$ , becomes arbitrarily large. Similar results were obtained in [7], [9] and [10] where a NCE-like methodology was used, but for different problem settings. Other classes of mean field games under the NCE principle were considered in [11]–[14].

The class of risk-sensitive mean field games considered in [15] and [16] is different from those considered in [1], [3], [6] and [7] in the sense that the former involves an exponentiated integral cost. The one-agent version is the risk-sensitive control problem, which was first introduced in [17] to capture risk-averse behavior (see [18] and the references therein). In [17], also the relationship of the LQ risk-sensitive control problem to a stochastic LQ zero-sum differential game was established. Later, the infinite-horizon risk-sensitive control problem was considered in [19], and risk-sensitive differential nonzero-sum games were studied in [18].

In this paper, we study linear-quadratic risk-sensitive mean field games (LQRSMFGs). In [15] and [16], the general class of risk-sensitive mean field games was considered for the finite-horizon and uniform agent case. This paper extends those results to the infinite-horizon problem and the heterogeneous agent case. By invoking the NCE principle, we first obtain a robust decentralized control law for each agent to construct the mean field system. We then provide appropriate conditions under which the mean field system admits a unique deterministic fixed point that actually approximates the mean field term. Moreover, we analyze closed-loop system stability, and show the consistency between the fixed point and the mean field coupling when  $N$  goes to infinity.

We also discuss an  $\epsilon$ -Nash equilibrium property of the set of  $N$  robust decentralized control laws, and show that  $\epsilon$  can be picked to be arbitrarily close to zero as  $N \rightarrow \infty$ , but our  $\epsilon$  bound is weaker than its linear-quadratic mean field game (LQMFG) counterpart, due to risk sensitivity. Finally, we discuss two different limiting cases, and show that one of these is equivalent to the corresponding LQMFG in [7].

The structure of the paper is as follows. We formulate the problem in Section II. The result of a single agent robust tracking problem is given in Section III. In Sections IV and V, we solve the LQRSMFG and provide the optimality anal-

Jun Moon is supported in part by the Fulbright Commission.  
This research was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) MURI grant FA9550-10-1-0573.  
All authors are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801; emails: {junmoon2, basarl}@illinois.edu.

ysis of the decentralized control laws, respectively. Section VI discusses the two different limiting cases. A numerical example is provided in Section VII. We end the paper with the concluding remarks of Section VIII.

## II. PROBLEM FORMULATION

The stochastic differential equation (SDE) describing the state of agent  $i$ ,  $1 \leq i \leq N$ , is given by

$$dx_i = A_i x_i dt + B_i u_i dt + \sqrt{\mu} D_i dW_i, \quad (1)$$

where  $x_i \in \mathbb{R}^n$  is the state;  $u_i \in \mathbb{R}^m$  is the control input;  $W_i$  is a  $p$ -dimensional standard Brownian motion;  $\mu > 0$  is a noise intensity parameter;  $A_i$ ,  $B_i$ , and  $D_i$  are time-invariant matrices with appropriate dimensions. The set of initial states  $\{x_i(0) : 1 \leq i \leq N\}$  are independent, and have zero mean and given covariances. We also assume that  $\{W_i : 1 \leq i \leq N\}$  are independent, and are also independent of  $\{x_i(0) : 1 \leq i \leq N\}$ .

In (1),  $A_i$ ,  $B_i$  and  $D_i$  are parameters of each agent's SDE that can be identical or different depending on the game at hand. We assume that the triplet  $(A_i, B_i, D_i)$  constitutes the system parameter  $\theta_i \in \Theta$ , that is,  $\theta_i := (A_i, B_i, D_i) \in \Theta$  where  $\Theta$  is the set of values of  $\theta_i$ . Note that  $\theta_i$  is a vector that stands for the ordered  $(A_i, B_i, D_i)$  with stacked column vectors.

**Assumption 1:** For the system parameters of the first  $N$  agents, it is assumed that we have the following empirical distribution:

$$F_N(\theta) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{\theta_i \leq \theta\}}, \quad \theta \in \Theta,$$

where  $\mathbb{1}_{\{\cdot\}}$  is the indicator function and  $\theta_i \leq \theta$  means componentwise inequality for the two vectors  $\theta_i$  and  $\theta$ . We further assume that there is a probability distribution  $F(\theta)$  such that  $F_N(\theta)$  converges weakly to  $F(\theta)$  as  $N \rightarrow \infty$ , i.e.,  $\lim_{N \rightarrow \infty} F_N(\theta) = F(\theta)$ , and the support of  $F(\theta)$ ,  $\Theta$ , is compact.  $\square$

The risk-sensitive cost function for agent  $i$  is given by

$$J_i^N(u_i, u_{-i}) = \limsup_{T \rightarrow \infty} \frac{\delta}{T} \log \mathbb{E} \left\{ e^{\frac{1}{\delta} \int_0^T c_i(x_i, f_N, u_i, t) dt} \right\}, \quad (2)$$

where  $u_{-i} := (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$ ,  $\delta$  is a risk sensitive index, and

$$c_i(x_i, f_N, u_i, t) := \|x_i(t) - f_N(t)\|_Q^2 + \|u_i(t)\|_R^2, \quad (3)$$

where  $\|z\|_S^2 := z^T S z$  for some vector  $z$  and matrix  $S \geq 0$ ;  $Q \geq 0$ ;  $R > 0$ ; and the coupling term,  $f_N(t)$ , interprets the average behavior (as measured by their states) of the first  $N$  agents, which can be written as

$$f_N(t) = \frac{1}{N} \sum_{i=1}^N x_i(t). \quad (4)$$

In this paper, we consider only the risk averse problem; hence  $\delta > 0$ . We denote the disturbance attenuation parameter by

$\gamma := \sqrt{\delta/2\mu}$ . The significance and relevance of this will be clarified in Section III.

Let  $\mathcal{F}_i^t$  be the  $\sigma$ -algebra generated by  $x_i(s)$  for  $s \leq t$  and  $\mathcal{F}^t := \{\mathcal{F}_i^t, 1 \leq i \leq N\}$  be the  $\sigma$ -algebra generated by states of all agents for  $s \leq t$ . We define the following two sets of admissible controls:

$$\mathcal{U}_i^{dec} = \{u_i : u_i(t) \text{ is adapted to } \mathcal{F}_i^t\} \quad (5)$$

$$\mathcal{U}_i^{cen} = \{u_i : u_i(t) \text{ is adapted to } \mathcal{F}^t\}. \quad (6)$$

Notice that while the admissible control in  $\mathcal{U}_i^{dec}$  is a function of agent  $i$ 's local state information, the admissible control in  $\mathcal{U}_i^{cen}$  is a function of full state information on all agents. Therefore, we have  $\mathcal{U}_i^{dec} \subset \mathcal{U}_i^{cen}$ . We call the former a set of decentralized controllers and the latter a set of centralized controllers.

We next introduce an approximate equilibrium solution for linear-quadratic risk-sensitive mean field games (LQRSM-FGs) formulated in this section.

**Definition 1:** The set of controllers  $\{u_i \in \mathcal{U}_i^{cen}, 1 \leq i \leq N\}$  constitutes an  $\epsilon$ -Nash equilibrium with respect to the cost functions  $\{J_i^N, 1 \leq i \leq N\}$ , if there exists  $\epsilon \geq 0$  such that for any  $i$ ,  $1 \leq i \leq N$ ,

$$J_i^N(u_i, u_{-i}) \leq \inf_{v_i \in \mathcal{U}_i^{cen}} J_i^N(v_i, u_{-i}) + \epsilon. \quad \square$$

In what follows, we first obtain the robust decentralized controller for a single agent, and provide appropriate conditions under which the mean field term (4) can be approximated by the deterministic function with arbitrarily small error. We also provide the  $\epsilon$ -Nash equilibrium property of the individual robust decentralized controller for the LQRSMFG in the large population regime, i.e., the case when  $N \rightarrow \infty$ .

**Remark 1:** By applying the Taylor expansion to (2), and letting  $\delta \rightarrow \infty$ , one arrives at the following risk-neutral cost function:

$$J_i^N(u_i, u_{-i}) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left\{ \int_0^T c_i(x_i, f_N, u_i, t) dt \right\}, \quad (7)$$

where  $c_i(x_i, f_N, u_i, t)$  is defined in (3). This is the cost function of the LQ mean field game considered in [7].  $\square$

## III. ROBUST TRACKING CONTROL

We consider the SDE

$$dx = A x dt + B u dt + \sqrt{\mu} D dW, \quad (8)$$

and the risk-sensitive cost function

$$\bar{J}(u, g) = \limsup_{T \rightarrow \infty} \frac{\delta}{T} \log \mathbb{E} \left\{ e^{\frac{1}{\delta} \int_0^T \bar{c}(x, g, u, t) dt} \right\} \quad (9)$$

$$\bar{c}(x, g, u, t) := \|x(t) - g(t)\|_Q^2 + \|u(t)\|_R^2,$$

where  $g \in \mathcal{C}_n^b$  is a known deterministic function in the set of  $n$ -dimensional continuous and bounded vector-valued functions on  $[0, \infty)$  with the norm  $\|g\|_\infty \triangleq \sup_{t \geq 0} \|g(t)\|$  where  $\|\cdot\|$  denotes 2-norm of vectors. Note that  $\mathcal{C}_n^b$  is a Banach space [3]. The minimization of (9) can be seen as a

robust tracking problem with respect to the given reference signal  $g(t)$  [20].

**Theorem 1:** Consider the risk-sensitive control problem with the SDE (8) and the cost function (9). Suppose that  $(A, B)$  is controllable and  $(A, Q^{1/2})$  is observable. Suppose that  $\gamma := \sqrt{\delta/2\mu}$  is fixed such that with  $\gamma$ , there is a positive definite matrix  $P$  that solves the following generalized algebraic Riccati equation (GARE):

$$A^T P + P A + Q - P(BR^{-1}B^T - \frac{1}{\gamma^2}DD^T)P = 0. \quad (10)$$

Then:

(i) The robust tracking controller can be written as

$$\bar{u}(t) = -R^{-1}B^T P x(t) - R^{-1}B^T s(t), \quad (11)$$

where  $s(t)$  satisfies the following differential equation:

$$\frac{ds(t)}{dt} = -H^T s(t) + Qg(t), \quad (12)$$

where  $H := A - BR^{-1}B^T P + \frac{1}{\gamma^2}DD^T P$ .

(ii)  $H$  and  $G := A - BR^{-1}B^T P$  are Hurwitz.

(iii) The differential equation (12) has a unique solution in  $\mathcal{C}_n^b$  with the initial condition  $s(0) = -\int_0^\infty e^{H^T s} Qg(s)ds$ , which can be written as

$$s(t) = -\int_t^\infty e^{-H^T(t-s)} Qg(s)ds. \quad (13)$$

(iv) The minimum cost is

$$\begin{aligned} \bar{J}(\bar{u}, g) \\ = \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left( g^T(s) Qg(s) - s^T(s) B R^{-1} B^T s(s) \right. \\ \left. + \frac{1}{\gamma^2} s^T(s) D D^T s(s) \right) ds + \mu \text{Tr}(P D D^T), \end{aligned}$$

where  $\text{Tr}(\cdot)$  is the trace operator.

**Proof.** Parts (i) and (iv) can be shown by solving the Hamilton-Jacobi-Bellman equation in [16] and [18] with the value function  $V(x, t) = \|x(t)\|_{P(t, T)}^2 + 2x^T(t)s(t) + q(t)$ , where  $P(t, T)$  can be obtained by solving the Riccati differential equation in Chapter 4 of [20],  $s(t)$  satisfies the differential equation in (12), and  $q(t)$  satisfies the following differential equation:

$$\begin{aligned} -\frac{dq(t)}{dt} &= -s^T(t) B R^{-1} B^T s(t) + \frac{1}{\gamma^2} s^T(t) D D^T s(t) \\ &+ g^T(t) Qg(t) + \mu \text{Tr}(P D D^T). \end{aligned}$$

Then the existence of the solution of (10) guarantees that there is no conjugate point of  $P(t, T)$  for all  $t$  and  $T$ , which also implies that  $\lim_{T \rightarrow \infty} P(t, T) = P$  and part (iv) [20]. Part (ii) can be shown by using an argument as in [20], since the robust tracking problem in this section is equivalent to the stochastic zero-sum differential game in [20] (see also [18], [19]). To prove part (iii), we first note that the solution of (12) can be written as

$$s(t) = e^{-H^T t} s(0) + \int_0^t e^{-H^T(t-s)} Qg(s)ds.$$

Then it is easy to show that  $s(0)$  in (iii) admits a unique function in  $\mathcal{C}_n^b$  [3], [7]. This completes the proof.  $\square$

**Remark 2:** (i) Define

$$\gamma^* := \inf\{\gamma > 0 : P > 0, P \text{ solves (10)}\}. \quad (14)$$

Then by definition, if  $\gamma^*$  is finite, for any  $\gamma > \gamma^*$ , the GARE in (10) always has a positive definite solution  $P$ . The value of  $\gamma^*$  is known as the optimum disturbance attenuation level of the  $H^\infty$  control problem, and any  $\gamma > \gamma^*$  determines the level of the disturbance attenuation [20].

(ii) If  $\gamma^*$  is finite, then since  $H$  and  $G$  are Hurwitz for any  $\gamma > \gamma^*$ , there exist positive constants  $\rho_1, \rho_2, \eta_1$ , and  $\eta_2$  such that  $\|e^{Ht}\| \leq \rho_1 e^{-\eta_1 t}$  and  $\|e^{Gt}\| \leq \rho_2 e^{-\eta_2 t}$  where  $\|\cdot\|$  denotes 2-norm of matrices.  $\square$

#### IV. LINEAR-QUADRATIC RISK-SENSITIVE MEAN FIELD GAMES

In this section, we study the LQRSMFG in the large population regime. We address particularly the computation of the mean field coupling term with arbitrarily small error by invoking the Nash certainty equivalence principle in [3].

##### A. Mean Field System

In this subsection, we construct a mean field system to arrive at the exact model of the mean field term in (4) in the large population regime.

We modify the optimum disturbance attenuation level in (14) for the continuum agent of  $\theta \in \Theta$ :

$$\begin{aligned} \gamma_\theta^* &:= \inf\{\gamma > 0 : P(\theta) > 0, A^T(\theta)P(\theta) + P(\theta)A(\theta) + Q \\ &- P(\theta)(B(\theta)R^{-1}B^T(\theta) - \frac{1}{\gamma^2}D(\theta)D^T(\theta))P(\theta) = 0\}. \end{aligned}$$

Let  $\bar{x}_\theta(t) = \mathbb{E}\{x_\theta(t)\}$ , and suppose  $\gamma > \gamma_\theta^*$  is finite. By substituting the robust decentralized controller (11) with  $\theta = \theta_i \in \Theta$  into (1), and taking expectation, the resulting system can be written as

$$\begin{aligned} \bar{x}_\theta(t) &= \int_0^t e^{G(\theta)(t-\tau)} B(\theta) R^{-1} B^T(\theta) \\ &\times \left( \int_\tau^\infty e^{-H^T(\theta)(\tau-s)} Qg(s)ds \right) d\tau. \end{aligned} \quad (15)$$

Now, by using (15) and the empirical distribution of the agents introduced in Assumption 1, the mean field system can be written as

$$\begin{aligned} \mathcal{T}(g)(t) &:= \int_\Theta \left\{ \int_0^t e^{G(\theta)(t-\tau)} B(\theta) R^{-1} B^T(\theta) \right. \\ &\times \left. \left( \int_\tau^\infty e^{-H^T(\theta)(\tau-s)} Qg(s)ds \right) d\tau \right\} dF(\theta), \end{aligned} \quad (16)$$

where  $dF(\theta)$  denotes the measure induced by the distribution function  $F(\theta)$ . Note that in (15) and (16), we use  $\theta$  instead of  $i$  to emphasize the point that we have a continuum of agents under the parameter set and its empirical distribution in Assumption 1. Further note that the operator (16) describes the average behavior of all the agents within  $\Theta$ ; hence it must

be consistent with the mean field coupling term in (4) when  $N$  is sufficiently large.

**Assumption 2:** (i)  $(A(\theta), B(\theta))$  and  $(A(\theta), Q^{1/2})$  are controllable and observable, respectively, for all  $\theta \in \Theta$ .

(ii)  $\gamma_\theta^*$  is finite and  $\gamma > \gamma_\theta^*$  for all  $\theta \in \Theta$ .

(iii) We have

$$\begin{aligned} & \|R^{-1}\| \|Q\| \int_{\Theta} \|B(\theta)\|^2 \left( \int_0^\infty \|e^{G(\theta)\tau}\| d\tau \right) \\ & \times \left( \int_0^\infty \|e^{H(\theta)\tau}\| d\tau \right) dF(\theta) < 1. \end{aligned} \quad \square$$

**Theorem 2:** Suppose that Assumptions 1 and 2 hold. Then:

(i) The operator (16) is in  $\mathcal{C}_n^b$  for all  $g \in \mathcal{C}_n^b$ .

(ii) There is a unique  $g^* \in \mathcal{C}_n^b$  such that  $g^* = \mathcal{T}(g^*)$  where  $g^*(t)$  is the approximated mean field coupling term in (4).

**Proof.** Part (i) can be shown by following an argument similar to that in [7]. For part (ii), it can be shown that

$$\begin{aligned} & \|\mathcal{T}x - \mathcal{T}y\|_\infty \\ & \leq \|x - y\|_\infty \|R^{-1}\| \|Q\| \int_{\Theta} \|B(\theta)\|^2 \left( \int_0^\infty \|e^{G(\theta)\tau}\| d\tau \right) \\ & \times \left( \int_0^\infty \|e^{H(\theta)\tau}\| d\tau \right) dF(\theta). \end{aligned}$$

Then by Assumption 2(iii) and the Banach fixed point theorem, we have the desired result.  $\square$

### B. Stabilization and Approximation Performance

The next theorem shows stability of the closed-loop system of (1) under the robust decentralized controller in (11). We rewrite the robust decentralized controller in Theorem 1 for agent  $i$ :

$$u_i^*(t) = -R^{-1}B_i^T P_i x_i^*(t) - R^{-1}B_i^T s_i(t), \quad (17)$$

where  $s_i(t)$  is (13) with  $\theta = \theta_i$  and under  $g^*(t)$  in Theorem 2. We denote the system state under the decentralized controller in (17) by  $x_i^*(t)$ . We also denote the mean field coupling term when each agent is under the robust decentralized controller in (17) by  $f_N^*$ .

**Theorem 3:** Suppose that Assumptions 1 and 2 hold. Then, the individual closed-loop system is stable in the time-average sense:

$$\sup_{N \geq 1} \max_{1 \leq i \leq N} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left\{ \int_0^T \|x_i^*(t)\|^2 dt \right\} < \infty.$$

**Proof.** We have the following relation:

$$\begin{aligned} & \mathbb{E}\{\|x_i^*(t)\|^2\} \\ & \leq 4\mathbb{E}\{\|e^{G_i t} x_i(0)\|^2\} \\ & \quad + 4\left\| \int_0^t e^{G_i(t-\tau)} B_i R^{-1} B_i^T s_i(\tau) d\tau \right\|^2 \\ & \quad + 4\mathbb{E}\left\{ \left\| \sqrt{\mu} \int_0^t e^{G_i(t-\tau)} D_i dW_i(\tau) \right\|^2 \right\} \\ & = 4T_1(t) + 4T_2(t) + 4T_3(t). \end{aligned}$$

Now, since  $G_i$  is Hurwitz,  $\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T T_1(t) dt = 0$ . Moreover, by following an argument similar to that in [7], there exists a constant  $M \geq 0$  such that  $\|s_i\|_\infty \leq M$ . Then, it can be shown that  $T_2^{1/2}(t) \leq \|B_i\|^2 \|R^{-1}\| M \rho_2 / \eta_2$  for all  $t$ . Hence  $\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T T_2(t) dt < \infty$ . For the last part, note that

$$T_3(t) = \mu \text{Tr} \left( \int_0^t e^{G_i(t-\tau)} D_i D_i^T e^{G_i^T(t-\tau)} d\tau \right) \triangleq \text{Tr}(\Xi_i(t)),$$

where  $\Xi_i(0) = 0$ . Then, we know that  $\Xi_i(t)$  is nondecreasing, and converges to  $\Xi_i \geq 0$  as  $t \rightarrow \infty$ , since  $G_i$  is Hurwitz. Hence,  $T_3(t) \leq \text{Tr}(\Xi_i)$  for all  $t$  and  $\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T T_3(t) dt \leq \text{Tr}(\Xi_i) < \infty$ . Finally, since  $\Theta$  is compact by Assumption 1, we have the desired result.  $\square$

We now show the consistency between  $g^*(t)$  and the mean field term  $f_N^*(t)$  in the large population regime.

**Theorem 4:** Suppose that Assumptions 1 and 2 hold. Then the following three results hold:

$$\begin{aligned} & \text{(i)} \quad \lim_{N \rightarrow \infty} \mathbb{E} \left\{ \int_0^T \|f_N^*(t) - g^*(t)\|^2 dt \right\} = 0 \\ & \text{(ii)} \quad \lim_{N \rightarrow \infty} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left\{ \int_0^T \|f_N^*(t) - g^*(t)\|^2 dt \right\} = 0. \end{aligned}$$

**Proof.** Consider the relation

$$\begin{aligned} & \mathbb{E} \left\{ \int_0^T \|f_N^*(t) - g^*(t)\|^2 dt \right\} \\ & \leq 2\mathbb{E} \left\{ \int_0^T \left\| f_N^*(t) - \frac{1}{N} \sum_{i=1}^N \bar{x}_i^*(t) \right\|^2 dt \right\} \end{aligned} \quad (18)$$

$$+ 2T \sup_{t \geq 0} \left\| \frac{1}{N} \sum_{i=1}^N \bar{x}_i^*(t) - g^*(t) \right\|^2. \quad (19)$$

Let  $e_i^*(t) = x_i^*(t) - \bar{x}_i^*(t)$  and  $\Lambda_i(t) := \mathbb{E}\{e_i^*(t)(e_i^*(t))^T\}$ . Then the SDE for  $e_i^*(t)$  can be written as

$$de_i^* = G_i e_i^* dt + \sqrt{\mu} D_i dW_i.$$

Hence, for (18), we have

$$\mathbb{E} \left\{ \int_0^T \left\| \frac{1}{N} \sum_{i=1}^N e_i^*(t) \right\|^2 dt \right\} = \frac{1}{N^2} \text{Tr} \left( \int_0^T \sum_{i=1}^N \Lambda_i(t) dt \right),$$

where the equality follows due to the fact that the SDE of each agent is independent under the robust decentralized control law. It can be shown that  $\Lambda_i(t)$  satisfies the following Lyapunov equation:

$$\dot{\Lambda}_i(t) = G_i \Lambda_i(t) + \Lambda_i(t) G_i^T + \mu D_i D_i^T.$$

Since  $G_i$  is Hurwitz for all agents, the above Lyapunov equation is monotonically nondecreasing, and converges to  $\Lambda_i > 0$  as  $t \rightarrow \infty$  for all  $i$ . Hence there exists a positive definite matrix  $\Lambda$  independent of  $N$  such that

$$\mathbb{E} \left\{ \int_0^T \left\| f_N^*(t) - \frac{1}{N} \sum_{i=1}^N \bar{x}_i^*(t) \right\|^2 dt \right\} \leq \frac{T}{N} \text{Tr}(\Lambda).$$

Therefore (18) converges to zero as  $N \rightarrow \infty$  for all  $T \geq 0$ . The convergence of (19) follows from Theorems 5.1 in [7] and the Helly-Bray theorem [21]. Part (ii) immediately follows from part (i). This completes the proof.  $\square$

## V. OPTIMALITY ANALYSIS: THE $\epsilon$ -NASH EQUILIBRIUM PROPERTY

In this section, we discuss an  $\epsilon$ -Nash equilibrium property of the set of  $N$  robust decentralized control laws of (17),  $\{u_i^* : 1 \leq i \leq N\}$ .

Let

$$J_i^N(u_i^*, u_{-i}^*) = \limsup_{T \rightarrow \infty} \frac{\delta}{T} \log \mathbb{E} \left\{ e^{\frac{1}{\delta} \int_0^T c_i(x_i^*, f_N^*, u_i^*, t) dt} \right\} \quad (20)$$

$$J_i^N(u_i, u_{-i}^*) = \limsup_{T \rightarrow \infty} \frac{\delta}{T} \log \mathbb{E} \left\{ e^{\frac{1}{\delta} \int_0^T c_i(x_i, f_N^{-i*}, u_i, t) dt} \right\}, \quad (21)$$

where  $c_i(x_i^*, f_N^*, u_i^*, t)$  is (3) when all agents use the robust decentralized control law (17), and  $c_i(x_i, f_N^{-i*}, u_i, t)$  is (3) when agent  $i$  is under the full state feedback controller  $u_i \in \mathcal{U}_i^{cen}$ , while other agents are still under the robust decentralized controller (17).

We now present the main result of this section.

**Theorem 5:** Consider the LQRSMFG. Suppose that Assumptions 1 and 2 hold. Then, the set of  $N$  robust decentralized control laws of (17),  $\{u_i^* : 1 \leq i \leq N\}$ , constitutes an  $\epsilon$ -Nash equilibrium for the LQRSMFG. That is, for any  $i$ ,  $1 \leq i \leq N$ ,

$$J_i^N(u_i^*, u_{-i}^*) \leq \inf_{u_i \in \mathcal{U}_i^{cen}} J_i^N(u_i, u_{-i}^*) + \epsilon_N,$$

where  $\epsilon_N \rightarrow 0$  as  $N \rightarrow \infty$ .

**Proof.** For the proof, we consider the controllers in (5) and (6) that guarantee the second moment stability of each system in the almost sure (a.s.) sense ( $\|x_i(T)\|^2 = o(T)$  a.s. and  $\int_0^T \|x_i(t)\|^2 dt = O(T)$  a.s.) to guarantee boundedness of (21). We also use the result:  $\lim_{N \rightarrow \infty} \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \|f_N^*(t) - g^*(t)\|^2 dt = 0$  a.s., which was shown in [7]. Then by using the Cauchy-Schwarz inequality and the fact that exponentiating, taking expectation and logarithm, and then taking the limit do not change the direction of the inequality, we can establish the result. The detailed proof is omitted due to space limitation.  $\square$

**Remark 3:** The  $\epsilon$ -Nash equilibrium for the LQRSMFG can be achieved when each agent behaves individually optimal under the robust decentralized controller.  $\square$

## VI. LIMITING BEHAVIORS OF THE LINEAR QUADRATIC RISK-SENSITIVE MEAN FIELD GAME

In this section, we discuss some implications of the results presented above as the design parameters approach specific limits. In particular, we are interested in two types of limiting processes. The first is known as the large deviation limit (small noise limit) and the second is the risk-neutral limit.

### A. Small Noise Limit

We consider the small noise limit of the LQRSMFG in which the noise intensity parameter  $\mu$  decreases to zero. If we take  $\mu \rightarrow 0$ , then the SDE for agent  $i$  now becomes the following ordinary differential equation:

$$\frac{dx_i(t)}{dt} = A_i x + B_i u_i. \quad (22)$$

Moreover, the risk-sensitive objective function in (2) heavily penalizes any large deviation of the running cost. Then we have the following result:

**Proposition 1:** Consider the LQRSMFG in Section II with  $\mu \rightarrow 0$  and  $\delta \rightarrow 0$  such that  $\gamma = \sqrt{\delta/2\mu}$  is fixed and positive. Suppose that Assumptions 1 and 2 hold. Then:

- (i) The decentralized controller for each agent is (17).
- (ii) The decentralized controller in (i) stabilizes the closed-loop system of (22) in the time-average sense.
- (iii) The set of  $N$ -decentralized controllers in (i) constitutes an  $\epsilon$ -Nash equilibrium for the LQRSMFG, and  $\epsilon$  converges to zero as  $N \rightarrow \infty$ .

### B. Linear-Quadratic Mean Field Game ( $\delta \rightarrow \infty$ )

Now, consider the case when  $\delta \rightarrow \infty$  for a fixed  $\mu$ . As mentioned in Remark 1, when  $\delta \rightarrow \infty$ , the risk-sensitive cost function in (2) can be written as (7). This imposes a smaller weight on the large deviation of the running cost. Then we have the following result:

**Proposition 2:** Consider the LQRSMFG in Section II with  $\delta \rightarrow \infty$  for a fixed  $\mu$ . Suppose that Assumptions 1 and 2 hold. Then:

- (i) The optimal decentralized control law for each agent can be written as

$$u_i^*(t) = -R^{-1} B_i^T Z_i x_i(t) - R^{-1} B_i^T r_i(t)$$

where  $Z_i$  is a positive definite matrix that is a solution of the following ARE:

$$A_i^T Z_i + Z_i A_i + Q - Z_i B_i R^{-1} B_i^T Z_i = 0,$$

and  $r_i(t)$  is

$$r_i(t) = - \int_t^\infty e^{-G_i^T(t-s)} Q g(s) ds.$$

- (ii) The decentralized control law in (i) stabilizes the SDE (1) in the time-average sense.
- (iii) The set of  $N$ -decentralized controllers in (i) constitutes an  $\epsilon$ -Nash equilibrium for the LQRSMFG. Moreover,  $\epsilon$  can be made arbitrarily small as  $N \rightarrow \infty$ .  $\square$

Note that Proposition 2 immediately follows from the results of the LQMFG in [7], since  $\delta \rightarrow \infty$  corresponds to  $\gamma \rightarrow \infty$ . This implies that the LQRSMFG has the same limiting behavior of  $\delta$  as the case of one-player risk-sensitive control problem discussed in [20].

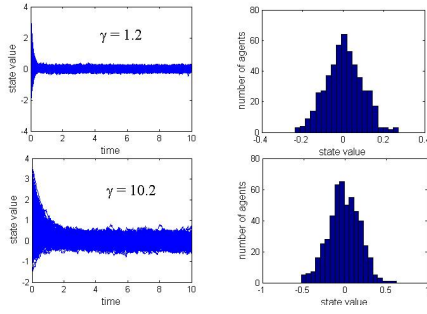


Fig. 1. State trajectories of  $N = 500$  and histograms at  $t = 10$  when  $\gamma = 1.2$  (top) and  $\gamma = 10.2$  (bottom).

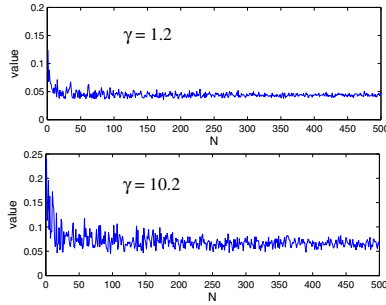


Fig. 2. The plot of Theorem 4(ii) with respect to  $N$  when  $\gamma = 1.2$  and  $\gamma = 10.2$ .

## VII. NUMERICAL EXAMPLE

We present in this section a numerical example on the LQRSMFG. Assume that each agent is uniform with  $A = 1.1$ ,  $B = D = Q = R = 1$  and  $\mu = 0.1$ . The initial condition is normally distributed with zero mean and unit variance. We have  $\gamma^* = 1$ . When  $\gamma = 1.2$ , it can be checked that Assumption 2(iii) holds, and we have  $g^*(t) = \bar{x}(t) = e^{\lambda t} \bar{x}(0)$  and  $s(t) = \frac{2}{\lambda + H} e^{\lambda t}$  where  $H = -1.48$  and  $\lambda = -1.8$ . The state trajectories of 500 agents and histograms at  $t = 10$  when  $\gamma = 1.2$  and  $\gamma = 10.2$  are shown in Fig. 1. The trajectories of the agents are more concentrated around the mean field value when  $\gamma = 1.2$  since the risk-sensitive control for each agent also takes into account attenuation of disturbance [20]. Figure 2 shows the plot of the approximation result of  $g^*$  under Theorem 4(ii) with respect to  $N$ . Note that each plot shows convergence approximately at the rate of  $O(1/N)$ , and also shows that a smaller value of  $\gamma$  leads to a better approximation performance due to the robustness property of the risk-sensitive control.

## VIII. CONCLUDING REMARKS

In this paper, we have studied linear-quadratic risk-sensitive mean field games (LQRSMFGs). We have shown that the mean field system provides the best approximation to the mean field coupling term under appropriate conditions. We have also shown stabilizability of the individual decentralized control law. Moreover, we have proven that the set of  $N$ -decentralized robust decentralized control laws constitutes an  $\epsilon$ -Nash equilibrium for the LQRSMFG. We

have shown that  $\epsilon$  can be made arbitrarily close to zero when  $N \rightarrow \infty$ . Finally, we have established the equivalence of LQRSMFG and LQMFG when the risk-sensitivity parameter goes to infinity, which implies that the  $\epsilon$ -Nash equilibrium for the LQRSMFG also features robustness for the individual SDEs.

## REFERENCES

- [1] J. M. Lasry and P. L. Lions, "Mean field games," *Jap. J. Math.*, vol. 2, no. 1, pp. 229–260, 2007.
- [2] G. Y. Weintraub, C. L. Benkard, and B. Van Roy, "Markov perfect industry dynamics with many firms," *Econometrica*, vol. 76, no. 6, pp. 1375–1411, 2008.
- [3] M. Huang, P. Caines, and R. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized  $\epsilon$ -Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.
- [4] M. Huang, P. Caines, and R. P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and nash equilibrium solutions," in *Proc. 42nd IEEE Conference on Decision and Control*, 2003, Dec. 2003, pp. 98–103.
- [5] H. Yin, P. Mehta, S. Meyn, and U. Shanbhag, "Synchronization of coupled oscillators is a game," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 920–935, 2012.
- [6] S. Adlakha, R. Johari, G. Weintraub, and A. Goldsmith, "Oblivious equilibrium for large-scale stochastic games with unbounded costs," in *Proc. 47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 5531–5538.
- [7] T. Li and J.-F. Zhang, "Asymptotically optimal decentralized control for large population stochastic multiagent systems," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1643–1660, 2008.
- [8] M. Huang, R. P. Malhamé, and P. Caines, "Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle," *Communications in Information and Systems*, vol. 6, no. 3, pp. 221–252, 2006.
- [9] C. Ma, T. Li, and J. Zhang, "Linear quadratic decentralized dynamic games for large population discrete-time stochastic multi-agent systems," *J. Sys. Sci. and Math. Scis.*, vol. 27, no. 3, pp. 464–480, 2007, (in Chinese).
- [10] A. Bensoussan, K. Sung, S. Yam, and S. Yung, "Linear-quadratic mean field games," (available at <http://www.sta.cuhk.edu.hk/scpy/Preprints>).
- [11] B. Wang and J. Zhang, "Mean field games for large-population multiagent systems with Markov jump parameters," *SIAM Journal on Control and Optimization*, vol. 50, no. 4, pp. 2308–2334, 2012.
- [12] A. C. Kizilkale and P. Caines, "Mean field stochastic adaptive control," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 905–920, 2013.
- [13] M. Huang, "Large-population LQG games involving a major player: The Nash certainty equivalence principle," *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 3318–3353, 2010.
- [14] M. Huang, P. Caines, and R. Malhamé, "Social optima in mean field LQG control: Centralized and decentralized strategies," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1736–1751, July 2012.
- [15] H. Tembine, Q. Zhu, and T. Başar, "Risk-sensitive mean-field stochastic differential games," in *Proceedings of the 2011 IFAC World Congress*, Milan, Italy, Aug. 2011, pp. 3222–3227.
- [16] —, "Risk-sensitive mean field games," *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 835–850, Apr. 2014.
- [17] D. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 124–131, Apr. 1973.
- [18] T. Başar, "Nash equilibria of risk-sensitive nonlinear stochastic differential games," *Journal of Optimization Theory and Applications*, vol. 100, no. 3, pp. 479–498, 1999.
- [19] T. Runolfsson, "The equivalence between infinite-horizon optimal control of stochastic systems with exponential-of-integral performance index and stochastic differential games," *IEEE Transactions on Automatic Control*, vol. 39, no. 8, pp. 1551–1563, Aug. 1994.
- [20] T. Başar and P. Bernhard, *H<sup>∞</sup> Optimal Control and Related Minimax Design Problems*, 2nd ed. Boston, MA: Birkhäuser, 1995.
- [21] G. R. Shorack, *Probability for Statisticians*. Springer, 2000.