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

Hybrid Risk-Sensitive Mean-Field Stochastic Differential Games with Application to Molecular Biology

Quanyan Zhu, Hamidou Tembine and Tamer Başar

***Abstract*— We consider a class of mean-field nonlinear stochastic differential games (resulting from stochastic differential games in a large population regime) with risk-sensitive cost functions and two types of uncertainties: continuous-time disturbances (of Brownian motion type) and event-driven random switching. Under some regularity conditions, we first study the best response of the players to the mean field, and then characterize the (strongly time-consistent Nash) equilibrium solution in terms of backward-forward macroscopic McKean-Vlasov (MV) equations, Fokker-Planck-Kolmogorov (FPK) equations, and Hamilton-Jacobi-Bellman (HJB) equations. We then specialize the solution to linear-quadratic mean-field stochastic differential games, and study in this framework the optimal transport of the GlpF transmembrane channel of *Escherichia coli*, where glycerol molecules (as players in the game) choose forces to achieve optimal transport through the membrane. Simulation studies show that GlpF improves the glycerol conduction more in a higher periplasmic glycerol concentration, which is consistent with observations made in the biophysics literature.**

I. INTRODUCTION

In recent years, game theory has been used to understand and design complex systems with many interacting independent decision-making agents in changing environments. Several papers have focused on games with a large number of players, and different approaches to large population games have been proposed. In [1] and [2], evolutionary games are used to model the multiple accesses in a population of transmitters and receivers in wireless communication systems. In [3], randomized sampling methods are used to solve large zero-sum matrix games to obtain security policies against an adversary. In [4], the equilibrium of a Stackelberg game with a single service provider and a large number of users is studied by taking the limit of the solution in analytic form. In [5], large population minority games are studied from a statistical mechanics viewpoint in the context of financial markets. In [6], mobile subscribers are modeled as a continuum, and it is shown that a threshold-type Wardrop equilibrium exists as a result of competitions for users between two base stations. In [10], the concept of oblivious equilibrium is proposed for the analysis of a class of large-scale stochastic games.

In this paper, we use a mean-field approach to study a class of large population hybrid risk-sensitive stochastic differential games. In most formulation of mean-field stochastic differential games (SDGs) such as [7] and [8], the cost functions to be minimized are expected values of the stage-additive loss functions. Such risk-neutral type of games cannot capture all the behaviors of agents in an uncertain and adversarial environment. In [14], a class of risk-sensitive mean-field SDGs with an exponentiated long-term cost function is considered and the corresponding mean-field equilibrium is characterized in terms of backward-forward macroscopic McKean-Vlasov equations, Fokker-Planck Kolmogorov equations and Hamilton-Jacobi-Bellman-Fleming (HJBF) equations. In this paper, we extend the risk-sensitive SDGs to a class of hybrid games that involve two types of uncertainties. One is the additive disturbances modeled by a zero-mean continuous-time Brownian motion process and the other one is a discrete uncertainty that is modeled by random switching between different structural states. Such a hybrid structure of uncertainties has been considered earlier in [22] for risk-neutral single-person linear-quadratic control problems. The worst-case design problem for the Markovian switching systems under a linear-quadratic cost criterion has been studied in [15]. In [17], minimax controllers of such switching systems have been characterized under the sampled-data perfect-state information structure.

In the first part of this paper, we examine the risk-sensitive SDGs with the hybrid structure of uncertainties in the context of a large population of players. We first present a general mean-field SDG model where the players are coupled through their risk-sensitive cost functionals as well as through their states. We provide the mean-field optimality principle and discuss compatibility with the density distribution obtained using generalized McKean-Vlasov and Fokker-Planck-Kolmogorov equations. We investigate the special linear-quadratic case and characterize the equilibrium solutions in terms of solutions to Riccati equations.

In the second part of the paper, we discuss an application in molecular biology based on the linear-quadratic framework. More specifically, we establish a hybrid stochastic differential equation model for the aquaglyceroporin GlpF transmembrane channel of *Escherichia coli* that facilitates the uptake of glycerol by the cell. We model the potential mean force as a control variable and the viewpoint whereby in which nature allows glycerol molecules to choose a force to achieve optimal transport through the membrane. Using the hybrid SDG framework, we observe in our simulations the phenomenon that GlpF improves the glycerol conduction more in a higher periplasmic glycerol concentration, which agrees with the observations made in several papers in the biophysics literature.

This work was supported in part by an AFOSR MURI, under grant number FA9550-10-1-0573. Q. Zhu and T. Başar are with ECE Department and CSL, University of Illinois at Urbana-Champaign, 1308 W. Main St., Urbana, IL, 61801, USA. E-mail: {zhu31,basar1}@illinois.edu. H. Tembine is with Telecommunication Department, Supelec, Gif-sur-Yvette, France. E-mail: tembine@ieee.org.

The mean-field game framework has been applied to many areas of research. In [12] and [9], the authors have treated pedestrians as individual entities and proposed a dynamical crowd evolution model using mean-field differential game theory in the framework of [8]. In [11], the consensus problem was studied based on stochastic mean-field Nash certainty equivalence (NCE) framework dispensing with the global evolution of the population. In [13], SDGs are used to understand the phase transition of a large population of interacting oscillators. Ours here seems to be the first attempt to use mean-field framework in biological applications.

The rest of the paper is organized as follows. In Section II, we present the hybrid risk-sensitive SDG model and discuss mean-field convergence and the general equilibrium principle. In Section III, we specialize the results to a special case of linear-quadratic differential games. In Section IV, we discuss an application of the hybrid SDG model in molecular biology and illustrate with numerical examples. Section V concludes the paper.

II. Problem Statement

We consider a class of n-person stochastic differential games subject to two types of uncertainties: a continuous-time Brownian motion that models the additive disturbances and a discrete stochastic uncertainty that models the random switching between different structural states. Denote by $\mathcal{N} := \{1, 2, \cdots, n\}$ the set of n players. A player Pi's individual physical state $x_i \in \mathcal{X}, i \in \mathcal{N}$, evolves according to the following Itô stochastic differential equation (SDE)

$$dx_i(t) = \frac{1}{n}\sum_{j=1}^{n} f(t, \theta_i(t), x_i^n(t), u_i^n(t), x_j^n(t), \theta_j(t))dt$$
$$+\frac{\sqrt{\epsilon}}{n}\sum_{j=1}^{n} \sigma(t, \theta_i(t), x_i^n(t), u_i^n(t), x_j^n(t), \theta_j(t))dB_i(t), \quad (1)$$
$$x_i^n(0) = x_{i,0} \in \mathcal{X} \subseteq \mathbb{R}^k,$$
$$\theta_i(0) = \theta_{i,0} \in \Theta,\ t \geq 0,\ i \in \mathcal{N}.$$

where $x_i^n(t) \in \mathcal{X}_i \subseteq \mathbb{R}^k$ is the k-dimensional state of Pi in a population of size n; $u_i^n \in \mathcal{U}_i \in \mathbb{R}^p$ is the p-dimensional control of player i in a population of size n; $B_i(t)$ are mutually independent standard Brownian motion processes in $\mathbb{R}^k$; and $\epsilon \geq 0$ is a small positive parameter that indicates the extent of stochastic influence from the Brownian motion; $\theta_i(t), \theta_{i,0} \in \Theta$ is a structural state taking values in a finite state space $\Theta := \{1, 2, \cdots, S\}$.

The process $\theta_i(t)$ is a Markov jump process with right-continuous sample paths, with initial states $\theta_{i,0}$ according to initial distribution π_0, and with rate matrix $\Lambda = \{\lambda_{ss'}\}_{s,s'\in\Theta}$; $\lambda_{ss'} \in \mathbb{R}_+$ are the transition rates such that for $s \neq s', \lambda_{ss'} \geq 0$, and $\lambda_{ss} = -\sum_{s'\neq s}\lambda_{ss'}$ for $s \in \Theta$. More precisely, for $h > 0$,

$$\text{Prob}\{\theta(t+h) = s'|\theta(t) = s\} = \begin{cases} \lambda_{ss'} + o(h) & s' \neq s, \\ 1 + \lambda_{ss} + o(h), & s' = s. \end{cases} \quad (2)$$

Note that all θ_j have the same transition law Λ, which is independent of the actions of the players and their states. We consider an individual state-feedback strategy for Pi, i.e., Pi chooses a control action $u_i = \mu_i(t, x)$, where $\mu_i : \mathbb{R}_+ \times \mathbb{R}^k \times \Theta \to \mathcal{U}_i$ is a feedback strategy. Let $\mathbb{U}_i^F$ be the class of such strategies. When the number of players is large, the cost of Pi does not explicitly depend on the actions of the other players but rather coupled through the distribution of the agent states. Let $m^n(t, x, \theta) = \frac{1}{n}\sum_{i=1}^{n}\delta_{x_i^n(t),\theta_i(t)}$ be an empirical measure of the states of the players, where $\delta_z, z \in \mathcal{X} \times \Theta$, is a Dirac measure on $\mathcal{X} \times \Theta$.

The goal of Pi, $i \in \mathcal{N}$, is to minimize his long-term cost functional

$$L_i^n(u_i, m^n; t, x, \theta, m) = \delta \log \mathbb{E}\left(\exp \tfrac{1}{\delta}\left[g(x(T), \theta(T))\right.\right.$$
$$+\int_t^T c(t', x_i(t'), u_i(t'), m^n(t'), \theta_i(t'))dt'$$
$$\left.\left. \mid x_i(t) = x, m^n(t, \cdot) = m, \theta_i(t) = \theta\right]\right), \quad (3)$$

where $\delta > 0$ is the risk-sensitivity index. We assume the following regularity conditions on $c, g,$ and f.

- (**A**1): f is piece-wise continuous in t, and Lipschitz continuous in (x, u, m) for each fixed sample path of θ, with probability one.
- (**A**2): g is continuous in x, and c is jointly continuous in (t, x, u, m) for each fixed sample path of θ, with probability one.
- (**A**3): The entries of matrix σ are C^2 and $\sigma\sigma'$ is strictly positive for each fixed sample path of θ, with probability one.
- (**A**4): $f, \partial_x f, c, \partial_x c, g, \partial_x g$ are uniformly bounded for each fixed sample path of θ, with probability one.
- (**A**5): $\mathcal{U}_i$ is closed and bounded.
- (**A**6): $u_i : [0, T] \times \mathbb{R}^k \times \Theta \to \mathcal{U}_i$ is piece-wise continuous in t and Lipschitz continuous in x_i for every $\theta \in \Theta$.

The system dynamics of Pi, (1), can be further written as

$$dx_i(t) = \left(\int_{w\in\mathcal{X}\times\Theta} f(t, \theta_i(t), x_i^n(t), u_i^n(t), w)m^n(t, dw)\right)dt+$$
$$\sqrt{\epsilon}\left(\int_{w\in\mathcal{X}\times\Theta} \sigma(t, \theta_i(t), x_i^n(t), u_i^n(t), w)m^n(t, dw,)\right)dB_i(t),$$
$$x_i^n(0) = x_{i,0} \in \mathcal{X} \subseteq \mathbb{R}^k, \theta_i(0) = \theta_{i,0} \in \Theta, t \geq 0, i \in \mathcal{N}. \quad (4)$$

The stochastic differential game problem with dynamics (1), (2) and cost functionals (3) is called a hybrid risk-sensitive stochastic differential game with perfect-state measurements. We seek an individual state-feedback non-cooperative Nash equilibrium $\mathbf{u}^* := \{u_i^*, i = 1, 2, \cdots, n\}$, satisfying the set of inequalities:

$$L_i(u_i^*, m^n; 0, x_{i,0}, m, \theta_{i,0}) \leq L_i(u_i, m^n; 0, x_{i,0}, m, \theta_{i,0}),$$

for all $u_i \in \mathbb{U}_i^F$, $i \in \mathcal{N}$, or the *strongly time-consistent* individual state-feedback Nash equilibrium,

$$L_i(u_i^*, m^n; t, x_i, m, \theta_i) \leq L_i(u_i, m^n; t, x_i, m, \theta_i), \quad (5)$$

for all $x_i \in \mathcal{X}, \theta_i \in \Theta, t \in [0, T)$, $u_i \in \mathbb{U}_i^F, i \in \mathcal{N}$.

A. FPK Equations

In [14], we have shown that the system (4) has the structure satisfying the asymptotic indistinguishability conditions of [18]. This leads to the existence of a random measure μ such that the system is μ-chaotic. Under an i.i.d. condition of the players and the controls $\mathbf{u}^*$, the solution of the state dynamics generates an indistinguishable sequence and weak convergence of the population profile process m^n to μ is equivalent to μ-chaocity. The weak convergence of the process m^n to m allows one to characterize the distribution m by the Fokker-Planck-Kolmogorov (FPK) equation. From [19] and [20], one has the generalized FPK equations:

$$\begin{aligned}\partial_t m(x,s,t) - \sum_{s'\in\Theta, s'\neq s} \lambda_{s's} m(x,s',t) + m(x,s,t)\sum_{s'\neq s}\lambda_{ss'} \\ +D_x^1\left(m(x,s,t)\int_w f(t,s,x,u^*(t),w)m(dw,t)\right) \\ = \frac{\epsilon}{2}D_{xx}^2\left(m(x,s,t)\left(\int_w \sigma'(t,s,u^*(t),w)m(dw,t)\right)\right. \\ \left.\cdot\left(\int_w \sigma'(t,s,u^*(t),w)m(dw,t)\right)\right), \\ s\in\Theta,\ x\in\mathcal{X},\ t\geq 0,\end{aligned}$$

Here $D_x^2(\cdot) := \sum_{k''=1}^k\sum_{k'=1}^k \frac{\partial^2}{\partial x_{k'}\partial x_{k''}}(m\Gamma_{k'k''})$, where $\Gamma = \sigma\sigma'$ is a $k\times k$ dimensional matrix and $\Gamma_{k'k''}$ indicates the (k',k'') entry of the matrix Γ; and $D_x^1(\cdot) := \sum_{k'=1}^k \frac{\partial}{\partial x_{k'}}\left(m\int_w f_{k'}(t,s,x,u^*,w)m(dw,s,t)\right)$; $f_{k'}$ denotes k'-th component of k-dimensional vector-valued $f := [f_{k''}]_{1\leq k''\leq k}$.

B. Hybrid Risk-Sensitive Best Response to Mean-Field

Consider the risk-sensitive stochastic differential game defined in (1), (2) and (3). Let $v_i^n(t,x_i,m,\theta)$ denote the value function associated with this differential game, i.e., $v_i^n(t,x_i,m,\theta) = \inf_{u_i\in\mathbb{U}_i^F} L_i^n(u_i,m^n,0,t,x_i,m,\theta)$. For the simplicity of the results, we assume f and σ to be one-dimensional and hence the differential operators D^1 and D^2 reduce to divergence div and the Laplacian operator Δ. In addition, we make the following assumption on v_i.

(**A**7): The differential game defined in this subsection has value functions $v_i, i\in\mathcal{N}$, for every initial time t, state $x(t)$ and structure $\theta(t)$, which is jointly continuously differentiable in (t,x) and twice continuously differentiable in x.

Theorem 1: Under Assumptions (**A**1)-(**A**7), suppose the trajectory of m^n is given. Then, v_i^n satisfies the Hamilton-Jacobi-Bellman-Fleming (HJBF) equation

$$\begin{aligned}\partial_t v_i^n(t,x_i,m,s) + \inf_{u_i\in\mathbb{U}_i^F}\{\partial_{x_i}v_i^n(t,x_i,m,s)\cdot f \\ +\sum_{s'\in\Theta}\lambda_{ss'}v_i^n(t,x_i,m,s') + \frac{\epsilon}{2}\mathrm{tr}(\sigma_t\sigma_t'\partial_{x_ix_i}^2 v_i^n(t,x_i,m,s)) \\ +\frac{\epsilon}{2\delta}\|\sigma\partial_{x_i}v_i^n(t,x_i,m,s)\|^2 + c\Big\} = 0. \\ v_i^n(T,x_i,m,s) = g(x_i(T),s),\quad s\in\Theta. \quad (6)\end{aligned}$$

Moreover, any strategy derived from

$$\begin{aligned}u_i^n(t) \in \arg\min_{u_i\in\mathbb{U}_i^F}\{\partial_{x_i}v_i^n(t,x_i,m,s)\cdot f \\ +\sum_{s'\in\Theta}\lambda_{ss'}v^i(t,x_i,m,s') + \frac{\epsilon}{2}\mathrm{tr}(\sigma_t\sigma_t'\partial_{x_i}^2 v_i^n(t,x_i,m,s)) \\ +\frac{\epsilon}{2\delta}\|\sigma\partial_{x_i}v_i^n(t,x_i,m,s)\|^2 + c\Big\} \quad (7)\end{aligned}$$

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

We can let $\delta\to\infty$ and obtain the risk-neutral best-response strategies to the mean field as stated below.

Corollary 1: Under Assumptions (**A**1)-(**A**7), suppose the trajectory of m^n is given. Then, v_i^n is a solution of the Hamilton-Jacobi-Bellman-Fleming (HJBF) equation

$$\begin{aligned}\partial_t v_i^n(t,x_i,m,s) + \inf_{u_i\in\mathbb{U}_i^F}\{\partial_{x_i}v_i^n(t,x_i,m,s)\cdot f + c+ \\ \textstyle\sum_{s'\in\Theta}\lambda_{ss'}v_i^n(t,x_i,m,s') + \frac{\epsilon}{2}\mathrm{tr}(\sigma_t\sigma_t'\partial_{x_ix_i}^2 v_i^n(t,x_i,m,s))\} \\ = 0. \quad (8) \\ v_i^n(T,x_i,m,s) = g(x_i(T),s),\quad s\in\Theta \quad (9)\end{aligned}$$

Moreover, any strategy derived from

$$\begin{aligned}u_i^n(t) \in \arg\min_{u_i\in\mathbb{U}_i^F}\{\partial_{x_i}v_i^n(t,x_i,m,s)\cdot f \\ +\sum_{s'\in\Theta}\lambda_{ss'}v^i(t,x_i,m,s') + \frac{\epsilon}{2}\mathrm{tr}(\sigma_t\sigma_t'\partial_{x_i}^2 v_i^n(t,x_i,m,s)) + c\}\end{aligned}$$

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

Consider the following $(n+1)$-person game with the set of players $\mathcal{N}\cup\{n+1\}$, in which the state dynamics of Pi, $i\in\mathcal{N}$, are given by

$$\begin{aligned}dx_i(t) = \left(\int_{w\in\mathcal{X}} f(t,\theta_i(t),x_i^n(t),u_i^n(t),w)dm^n(t,w,\theta_i)\right)dt \\ +\sigma(t)\xi(t) + \sqrt{\epsilon}\sigma(t)dB_i(t), x_i^n(0) = x_{i,0}\in\mathcal{X}\subseteq\mathbb{R}^k, \\ \theta_i^n(0) = \theta_{i,0},\ t\geq 0,\ i\in\mathcal{N}, \\ (10)\end{aligned}$$

where the player $n+1$ is a fictitious player who controls the parameter $\xi(t)\in\Xi^F$ to play against other players, where Ξ^F is the admissible set of feedback strategies for the $(n+1)$st player. Assume that $\xi : [0,T]\times\mathbb{R}^k\times\Theta\to\Xi^F$ is piecewise continuous in t and Lipschitz continuous in x for every $s\in\Theta$. Define the risk-neutral cost function for Pi, $i\in\mathcal{N}$, by

$$\begin{aligned}\hat{L}_i^n(t,u_i,\xi(t),m^n;t,x,\theta,m) = \mathbb{E}\left[g(x(T),\theta(T))\right. \\ +\int_t^T \hat{c}(t',x_i(t'),u_i(t'),\xi(t'),m^n(t'),\theta(t'))dt' \\ \left.\Big| x_i(t) = x, m^n(t,\cdot) = m, \theta_i(t) = \theta\right], \quad (11)\end{aligned}$$

where the cost $\hat{c}$ is separable in $\xi(t)$ and is given by

$$\hat{c} = c(t,x_i(t),u_i(t),m^n(t),\theta(t)) - \gamma^2\|\xi(t)\|^2. \quad (12)$$

Each player Pi seeks to find a robust controller under the uncertainty of the strategies chosen by player $n+1$, i.e, Pi

minimizes the worst case of $\hat{L}_i^n$ over the feedback strategy $\hat{\xi} \in \Xi_i^F$ of player $n+1$, i.e.,

$$\hat{v}_i^n = \inf_{\hat{u}_i \in \mathbb{U}_i^F} \sup_{\hat{\xi} \in \Xi^F} \hat{L}_i^n(\hat{u}_i, \hat{\xi}, m^n; t, x_i, \theta, m). \quad (13)$$

Then, we can establish an equivalence of the game described above with the hybrid risk-sensitive game in Section II.

Corollary 2: Consider the hybrid risk-neutral mean-field game described by (11) and (13) where each player seeks a robust (min-max) control against the $(n+1)$-th player. The robust hybrid risk-neutral mean-field stochastic game is equivalent to the hybrid risk-sensitive mean-field stochastic game defined by (3) and (4) under the regularity assumptions.

C. Coupled HJBF and Generalized FPK-McV Equations

The players choose an optimal controller to react to the mean-field and the mean-field of the game is influenced by the controls used by the players. In [14], we have shown that under suitable controls $(u_i^n(t), t \geq 0) \to u(t), t \geq 0)$ as $n \to \infty$, one can derive a weak convergence of the risk-sensitive cost function as stated in the following.

Theorem 2 ([14]): The risk-sensitive cost functional $L_i^n(u_i^n, m^n, t, x_i, \theta_i, m_0)$ converges to $L_i(u_i, m, 0, x_i, \theta_i, m_0)$ given by

$$\begin{aligned} L_i(u_i, m; t, x_i, \theta_i, m_0) = \delta \log \mathbb{E}\Bigg(&\exp \frac{1}{\delta}\Big[g(x(T), \theta(T)) \\ &+ \int_t^T c(t, x(t), u(t), m(t), \theta(t))dt \\ &\Big| \; x_i(t) = x_i, m(t, \cdot) = m_0, \theta_i(t) = \theta_i\Big]\Bigg), \end{aligned} \quad (14)$$

The mean-field response system requires solving a coupled HJBF backward equation combined with the generalized FPK equation as follows. Assuming that the population is homogeneous, we can drop the index i.

$$\begin{cases} dx(t) = \left(\int_{w\in\mathcal{X}} f(t,\theta(t),x(t),u(t),w)dm(t,w,\theta)\right)dt + \\ \sqrt{\epsilon}\left(\int_{w\in\mathcal{X}} \sigma(t,\theta(t),x(t),u(t),w)m^n(t,w,\theta)dt\right)dB_i(t), \\ \\ x(0) = x_0 \in \mathcal{X} \subseteq \mathbb{R}^k, t \geq 0, i \in \mathcal{N}. \\ \\ \partial_t m(x,\theta,t) - \sum_{s'\in\Theta} \lambda_{s's} m(x,s',t) + m(x,s,t)\sum_{s'\neq s}\lambda_{ss'} \\ +D_x^1(m(x,\theta,t)\int_w f(t,\theta,x,u^*(t),w)m(dw,\theta,t)) \\ = \frac{\epsilon}{2}D_{xx}^2\left(m(x,\theta,t)\left(\int_w \sigma'(t,\theta,u^*(t),w)m(dw,\theta,t)\right)\right. \\ \left.\cdot\left(\int_w \sigma'(t,\theta,u^*(t),w)m(dw,\theta,t)\right)\right), \\ \\ m(x,s,0) = m^0(x,s), \; s\in\Theta, x\in\mathcal{X}. \\ \\ \partial_t v(t,x,m,s) + \inf_{u\in\mathbb{U}^F}\{\partial_x v(t,x,m,s)\cdot f \\ + \sum_{s'\in\Theta}\lambda_{ss'}v(t,x,m,s') + \frac{\epsilon}{2}\mathrm{tr}(\sigma_t\sigma_t'\partial_{xx}^2 v(t,x,m,s)) \\ + \frac{\epsilon}{2\delta}\|\sigma\partial_x v(t,x,m,s)\|^2 + c\} = 0. \\ \\ v(T,x,m,s) = g(x(T),s), s\in\Theta. \end{cases}$$

III. Linear-Quadratic Case

In this section, we consider the special case of the linear-quadratic problem. We let the functions defined in Section II take the following forms:

$$\begin{aligned} f(t,\theta_i,x_i,u_i,m) &= A(t,\theta_i(t))x_i + B(t,\theta_i(t))u_i; \\ \sigma(t,\theta_i,x_i,u_i,m) &= \sigma(t,\theta_i(t)); \\ c(t,\theta_i,x_i,u_i,m) &= |x_i|^2_{Q(t,\theta_i(t),m)} + |u_i|^2_{R(t,\theta_i(t),m)}; \\ g(x_i) &= |x_i(T)|^2_{Q_T(\theta_i(t))}. \end{aligned}$$

We make the following assumptions for the problem formulation for the linear-quadratic case.

(**A**8): $A(t,s)$, $B(t,s)$, $D(t,s)$, $Q(t,s,m)$, $R(t,s,m)$ are piece-wise continuous in t for each $s \in \Theta$.

(**A**9): Matrix functions $Q(t,s,m), R(t,s,m)$ are positive definite for all $t \in [0,T]$, $s \in \Theta$, for a given distribution m.

(**A**10): The Markov chain θ is irreducible.

A. Best Response to Mean-Field

For linear-quadratic games, we can use Riccati equations to characterize the best response to the mean field [15].

1) Finite-Horizon: For the finite horizon problem under perfect state measurements, the value function v, whenever it exists, is given by

$$\begin{aligned} v(t,\theta,x,m) &= x'Z(t,\theta(t),m)x \\ &+ \epsilon\int_t^T \mathrm{tr}\left(Z(\tau,\theta(\tau),m)\sigma(\tau)\sigma'(\tau)\right)d\tau, \end{aligned} \quad (15)$$

where $Z(t,\theta(t),m), \theta(t) \in \Theta, t \in [0,T]$ is the nonnegative definite solution of the generalized Riccati differential equation (GRDE)

$$\begin{aligned} &-\dot{Z}(t,s,m) = \\ &A'(t,s)Z(t,s,m) + Z(t,s,m)A(t,s) - Z'(t,s,m)\cdot \\ &\left(B(t,s)R^{-1}(t,s,m)B'(t,s) - \frac{1}{\gamma^2}\sigma(t,s)\sigma'(t,s)\right)Z(t,s,m) \\ &\quad + Q(t,s,m) + \sum_{s'\in\Theta}\lambda_{ss'}Z(t,s',m); \quad Z(T,s) = Q_T(s), \end{aligned} \quad (16)$$

where $\gamma = \sqrt{\frac{\delta}{2\epsilon}}$ and the strategy for Pi in best response to the mean field is

$$u_i^*(t,s,m) = -R^{-1}(t,s)B'(t,s)Z(t,s,m)x_i(t). \quad (17)$$

The associated Fokker-Planck equation is given by

$$\begin{aligned} &\partial_t m(x,s,t) + m(x,s,t)\sum_{s'\neq s}\lambda_{ss'} - \sum_{s'\in\Theta}\lambda_{ss'}m(x,s',t) = \\ &\mathrm{div}\left(m(x,s,t)\cdot(A(t,\theta(t))x + B(t,\theta(t))u^*(t,s,m))\right) \\ &\quad + \frac{\epsilon}{2}\Delta(m(x,s,t)\sigma(t,s)\sigma'(t,s)), \end{aligned} \quad (18)$$

with initial condition $m(x,s,0) = m^0(x,s)$. The above discussion is now captured by the following theorem.

Theorem 3: Consider the linear-quadratic hybrid risk-sensitive mean-field stochastic differential game of this subsection. Let Assumptions (**A**8)-(**A**10) hold. Then, the game admits a best response strategy to the mean field, given by

(17), where Z is the nonnegative solution to (GRDE) (16). The finite-horizon mean-field equilibrium is a solution to the coupled equations (18) and (17).

2) Infinite-Horizon Case: We take A, B, D, Q, R, Λ to be time-invariant and $Q_T(\cdot) = 0$ and π_0 to be the stationary distribution of the Markov chain, and require $x(t) \to 0$ as $t \to \infty$. The following two assumptions are needed for the ensuing analysis.

(**A**11): The pair $(A(\theta(t)), B(\theta(t)))$ is *stochastically stabilizable* (See Definition 1 in [22]).

(**A**12): The pair $(A(s), Q(s))$ is observable for each $s \in \Theta$.

We introduce the following set of coupled generalized algebraic Riccati equations (GAREs)

$$A'(s)Z(s) + Z(s)A(s) - Z(s)\Big(B(s)R^{-1}(s,m)B(s) - \frac{1}{\gamma^2}\sigma(s)\sigma'(s)\Big)Z(s) + Q(s,m) + \sum_{s'=1}^{S}\lambda_{ss'}Z(s'); \quad i = 1, \cdots, s. \tag{19}$$

Then, we have the following counterpart of Theorem 3.

Theorem 4: Consider the infinite-horizon linear-quadratic hybrid mean-field stochastic differential game with perfect state measurements, as defined in this subsection. Let Assumptions (**A**8)-(**A**12) hold. The best-response strategy for Pi to the mean field is given by

$$u_i^*(t) = \mu_i^*(t, x(t), \theta(t)) = -R^{-1}(\theta(t))B'(\theta(t))Z(\theta(t))x(t), \tag{20}$$

where $Z(\cdot)$ is a positive definite solution (GARE) to (19). The closed-loop linear system given by

$$\dot{x} = \Big(A(\theta(t)) - \Big(B(\theta(t))R^{-1}(\theta(t))B'(\theta(t)) - \frac{1}{\gamma^2}\sigma(\theta(t))\sigma'(\theta(t))\Big)Z(\theta(t))\Big)x(t) \tag{21}$$

is mean-square stable, i.e., $\lim_{t\to\infty}\mathbb{E}\{|x(t)|^2\} = 0$. In addition, the infinite-horizon mean-field equilibrium is a solution to the coupled equations (18) and (20).

IV. Application to Molecular Biology

Cell membranes act as a protective permeability barrier for preserving the internal integrity of the cell. Cell metabolism requires controlled molecular transport across the cell membrane. The *Escherichia coli* glycerol uptake facilitator (GplF) is an aquaglycerol channel protein, which transports both water and glycerol molecules, but excludes charged solutes, e.g. protons. It is a very efficient glycerol uptake system that allows rapid growth of the bacteria in glycerol solution at a concentration of 5 μM or in a nutrient-poor condition of 1 μM [28], [29]. The recent discovery of the structure of GlpF at a resolution of 2.2 Å has revealed important features of the transport mechanism. The channel walls match the hydrophilic and hydrophobic sides of glycerol, so that glycerol can be dehydrated without an energetic penalty. The entrance to the channel contains a narrow (4 Å wide) region that is thought to function as a selectivity filter. The potential of mean force (PMF) that guides the transport of glycerol

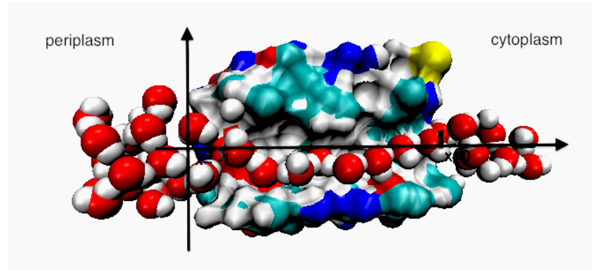

Fig. 1. Section through the glycerol conduction pathway in GlpF. The periplasmic side is taken as the origin marked with $x = -1$ and the cytoplasmic side is marked with $x = 0$.

through the channel is highly asymmetric reflecting the atomic structure of GlpF – a large (10 Å wide) hydrophilic vestibule extends into the water on the external side of the membrane [27]. In [26], the authors show that no biological functions could be attributed to asymmetry. Conduction rate of a single glycerol does not depend on the orientation of GlpF, and GlpF appears to conduct equally well in both directions under physiological conditions.

In [23], it has been demonstrated that the asymmetry of GlpF furnishes active glycerol transport through a ratchet-like mechanism under realistic physiological conditions. The ratchet effect refers to the generation of directed motion of Brownian particles in a spatially periodic and asymmetric ratchet potential in the presence of non-equilibrium fluctuations and/or externally applied time-periodic force with zero-mean. In this section, we use stochastic large population games to explain the dependence of the glycerol transport on function of the structure of GlpF. We adopt a simplified view as in [23] to model the glycerol as Brownian particles subject to thermal fluctuations and an on-and-off stochastic driving force. The periplasmic glycerol concentration has an effect on the conduction rate. We model the potential mean force as a control variable u and adopt the viewpoint whereby nature allows glycerol molecules to choose a force to achieve optimal transport along the membrane. In our simulations, we have observed the same phenomena as in [26], i.e., GlpF improves the glycerol conduction more in a higher periplasmic glycerol concentration. Recent interests in synthetic biology and protein designs have inspired us to study dependence of the transport on different system parameters such as risk sensitivity as well as weighting parameters [30], [31].

Aquaglyceroporin is a subclass of aquaporin water channels, which organize as tetramers in cell membranes, with each monomer forming a functionally independent water pore. The side view of GlpF tetramer is shown in Fig. 2, surrounded by water. In Fig. 3, we show the top view of GlpF tetramer embedded in a lipid membrane and surrounded by water. The tetrameric structure is universal to the entire family of aquaporins. Each monomer in an aquaporin tetramer functions independently as depicted in Fig. 4.

Here, we consider a fluctuation-driven molecular transport through an aquaglyceroporin GlpF membrane channel as depicted in Fig. 1. We let $x_i \in [-1, 0]$ denote the position of molecule i with $x_i = -1$ denoting the periplasmic side

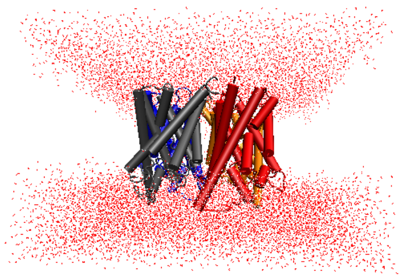

Fig. 2. Side view of GlpF tetramer surrounded by water.

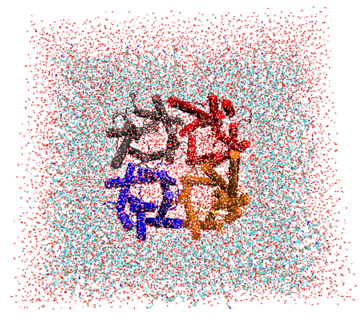

Fig. 3. Top view of GlpF tetramer embedded in a lipid membrane and surrounded by water.

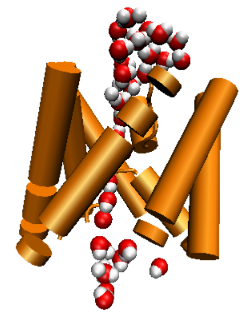

Fig. 4. Molecules inside the GlpF monomer.

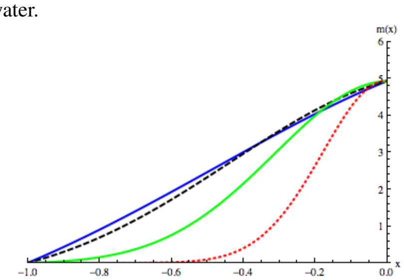

Fig. 5. Stationary distribution $m(x)$ with different Q values: $Q = 1$ (blue), $Q = 10$ (dotted black), $Q = 100$ (green), $Q = 1000$ (dotted red).

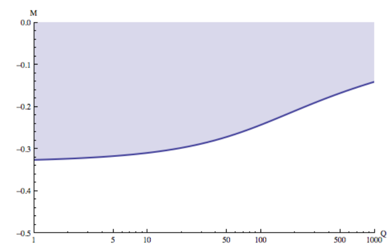

Fig. 6. Steady-state mean value for different values of Q in a linear-log plot in risk-neutral case.

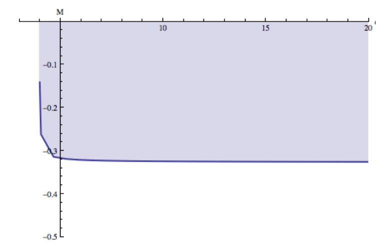

Fig. 7. Steady-state mean value M for different values of δ at $Q = 1000$.

and $x_i = 0$ cytoplasmic side. The molecule is driven by a stochastic driving force $F_{RTN}(t)$, i.e.,

$$dx_i(t) = (u_i + F_{RTN})dt + \sqrt{2}dB(t), \tag{22}$$

where F_{RTN} randomly switches between the two force levels F_0 and $-F_0$, $F_0 \in [0, 8]$pN, at independent, exponentially distributed times. The distribution function of the switching time $P_F(t) = 1 - \exp(-t/T_0)$, where T_0 is the mean switching time that is often chosen within the range from 10^{-9}s to 10^{-2}s. Let s_+ denote the state where F_0 is applied and s_- be the state where $-F_0$ is applied. Let $\overline{m} := m(x, s_+, t) + m(x, s_-, t)$ and $\underline{m} := m(x, s_+, t) - m(x, s_-, t)$. We are interested in the steady-state probability density. Hence, we can drop the time dependence of m, assuming the time limit exists, and study the corresponding stationary Fokker-Planck equations. By adding and subtracting two FPEs associated with each state, we obtain the following

$$\begin{aligned} -\partial_x \overline{m}(x) + u(x) \cdot \overline{m}(x) + F_0 \underline{m}(x) &=: J, \\ -\partial_x^2 \underline{m}(x) + \partial_x \left(u(x)\underline{m}(x) + F\overline{m}(x) \right) + \frac{2}{T_0}\underline{m}(x) &= 0, \end{aligned}$$

where J is the steady-state flux constant and $\int_{-1}^{0} \overline{m}(x)dx = 1$, $\int_{-1}^{0} \underline{m}(x)dx = 0$ and the boundary condition $\overline{m}(-1, t) = \overline{m}_p(t)$. From the above two coupled equations, we arrive at an ordinary differential equation model in terms of $\overline{m}(x)$:

$$-\partial_x^3 \overline{m}(x) + 2u(x)\partial_x^2 \overline{m}(x) + \left(F^2 - u^2(x) + \frac{2}{T_0} \right) \partial_x \overline{m} - \frac{2u(x)}{T_0}\overline{m} + \frac{2J}{T_0} = 0.$$

The nature of transmembrane transport is to conduct molecules from $x = 1$ to $x = 0$. It is shown in [25] that GlpF is an optimized mechanism for glycerol conduction under physiological conditions. As nature seeks to use minimum energy to achieve optimal transport through the membrane a, we construct an objective functional in terms of state energy and transport work as follows:

$$J_i = \delta \log \mathbb{E} \left(\frac{1}{\delta} \exp \left[\int_0^\infty |x_i(t)|^2_{Q(\theta, m)} + |u_i(t)|^2 dt \right] \right), \tag{23}$$

where $Q(\theta, m) := Q(\theta(t)) - M$ and $M := \mathbb{E}(m)$ models the interactions between the molecules in the transport channel of GlpF. More molecules at a particular site x will lead to a relatively higher cost on the control effort to transport the molecules. The infinite-horizon GAREs for state s_+ and s_- are

$$2F_0 z_+ - (1 - 4/\delta)z_+^2 + (Q_+ - M) - \frac{1}{T_0}(z_+ - z_-) = 0,$$
$$-2F_0 z_- - (1 - 4/\delta)z_-^2 + (Q_- - M) + \frac{1}{T_0}(z_+ - z_-) = 0.$$

Letting $Q_+ = Q_- = Q$, $\underline{z} = z_+ - z_-$ and $\overline{z} = z_+ + z_-$, we arrive at a set of equations in terms of $\underline{z}$ and $\overline{z}$ as follows:

$$F\underline{z} + Q - M = \frac{1}{2}(1 - 4/\delta) \left(\frac{\overline{z}_+^2 + \underline{z}_-^2}{2} \right), \tag{24}$$

$$F\overline{z} + Q - M = \frac{1}{2}(1 - 4/\delta)\overline{z}\underline{z} + \frac{1}{T}\underline{z}. \tag{25}$$

In our simulations, we chose $F_0 = 5$pN and $T_0 \in 10^{-5}$s. We solved the coupled equations and obtained the steady state distribution $m(x)$ which is the same for both states s_+ and s_-. In Fig. 5, we show the normalized stationary distribution $m(x)$ for different values of Q. A higher value of Q suggests that the periplasmic concentration is high and it is relatively easier for glycerol conduction. In Fig. 6, we show the mean value of the molecule positions as a function of parameter Q in a log-linear plot. We observe that the mean value grows almost exponentially when $Q > 50$ and the transport is more efficient for higher values of Q or higher periplasmic concentration glycerol. This result is in agreement with the biophysical observations made in [25] and [26], where the GlpF vestibule leads to a higher 75% conduction rate in a 10-μM periplasmic glycerol concentration than a 45% conduction rate in a 10 mM one.

In Fig. 7, we show the effect of risk sensitivity parameter δ on transport. We can view the effect of δ as an indicator of robustness. A lower value of δ suggests a higher demand of robustness in the conduction. We observe that with less need for robustness, the transport becomes less efficient, i.e., the mean value of the molecules at the steady state are farther away from the cytoplasm.

V. Conclusion

In this paper, we have studied a class of hybrid risk-sensitive mean-field stochastic differential games in which the players are coupled through their states and the cost functionals are exponentiated and coupled as well. We have discussed the application of the linear-quadratic differential games to the GlpF transmembrane channel of *Escherichia coli* and have shown that the SDGs can provide meaningful explanations to the transport phenomena and biological regulations for a large population of molecules. One of the interesting future directions would be the extension of the results to systems with output feedback and imperfect or noisy output measurements. Another direction of future work would be to connect synthetic biology with large population games and seek the possibility of designing proteins to achieve efficient mechanisms in transport and chemotaxis.

References

[1] Q. Zhu, H. Tembine and T. Başar, "Evolutionary game for hybrid additive white gaussian noise multiple access control," in *Proc. of IEEE Globecom*, Honolulu, Hawaii, Dec. 2009, pp. 1–6 .

[2] Q. Zhu, H. Tembine, T. Başar, "A constrained evolutionary Gaussian multiple access channel game," in *Proc. of Intl. Conf. on Game Theory for Networks (GameNets)*, Istanbul, Turkey, 13-15 May 2009, pp. 403–410.

[3] S. Bopardikar, A. Borri, J. Hespanha, M. Prandini and A. D. Di Benedetto, "Randomized sampling for large zero-sum games," in *Proc. of the 49th IEEE Conf. on Decision and Control (CDC)*, Dec. 2010, pp. 7675–7680.

[4] T. Başar and R. Srikant, "A Stackelberg network game with a large number of followers," *J. Optimization Theory and Applications*, vol. 115, no. 3, pp. 479–490, Dec. 2002.

[5] A.C.C. Coolen, *The Mathematical Theory of Minority Games: Statistical Mechanics of Interacting Agents*, Oxford University Press, 2005.

[6] G S. Kasbekar, E. Altman and S. Sarkar, "A hierarchical spatial game over licensed resources", in *Proc. of Intl. Conf. on Game Theory for Networks (GameNets)*, Istanbul, Turkey, 13-15, May 2009, pp. 70 – 79.

[7] M. Huang, P.E. Caines, and R.P. Malhame. "Large-population cost-coupled LQG problems with non-uniform agents: individual-mass behavior and decentralized epsilon-Nash equilibria," *IEEE Trans. on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, Sept. 2007.

[8] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, pp. 229–260, 2007.

[9] A. Lachapelle and M. T. Wolfram, "On a mean field game approach modeling congestion and aversion in pedestrian crowds," working paper, 2011, `http://events.chairefdd.org/?p=388`.

[10] S. Adlakha, R. Johari, G. Weintraub, and A. Goldsmith, "Oblivious equilibrium for large-scale stochastic games with unbounded costs," in *Proc. of the IEEE Conf. on Decision and Control (CDC)*, Cancun, Mexico, pp. 5531–5538, 2008.

[11] M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang, "A solution to the consensus problem via stochastic mean field control," in *Proc. of the 2nd IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys'10)*, Sept., 2010.

[12] C. Dogbé, "Modeling crowd dynamics by the mean-field limit approach," *Mathematical and Computer Modelling*, vol. 52, 2010, pp. 1506–1520.

[13] H. Yin, P. G. Mehta, S. P. Meyn and U. V. Shanbhag, "Synchronization of coupled oscillators is a game," in *Proc. of American Control Conf. (ACC)*, Baltimore, MD, 2010.

[14] H. Tembine, Q. Zhu and T. Başar, "Risk-sensitive mean-field stochastic differential games," in *Proc. of 18th IFAC World Congress*, Aug. 28 - Sept. 2, 2011, Milano, Italy.

[15] Z. Pan and T. Başar, "H-infinity control of Markovian jump systems and solutions to associated piecewise-deterministic differential games," in *New Trends in Dynamic Games and Applications* (G. J. Olsder, ed.), pp. 61–94, Boston, MA: Birkhäuser, 1995.

[16] W.H. Fleming and H.M. Soner, *Controlled Markov Processes and Viscosity Solutions*, vol. 25, Springer, Berlin, 1993.

[17] T. Başar, "Minimax control of switching systems under sampling," *Systems and Control Letters*, vol. 25, no. 5, pp. 315–325, Aug. 1995.

[18] Y. Tanabe, "The propagation of chaos for interacting individuals in a large population", *Mathematical Social Sciences*, vol. 51, pp. 125–152.

[19] V. Kontorovich and V. Lyandres,"Dynamic systems with random structure: an approach to the generation of nonstationary stochastic processes," *Journal of the Franklin Institute*, vol. 336, pp. 939–954, 1999.

[20] J. Bect, "A unifying formulation of the Fokker-Planck-Kolmogorov equation for general stochastic hybrid systems," in *Proc. of the 17th IFAC World Congress*, Seoul, Korea, July 6-11, pp. 15480–15485, 2008.

[21] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd edition, Classics in Applied Mathematics, SIAM, Philadelphia, 1999.

[22] Y. Ji, H. J. Chizeck, "Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control," *IEEE Trans. on Automatic Control*, vol. 35, no.7, pp. 777–788, July 1990.

[23] I. Kosztin and K. Schulten, "Fluctuation-driven molecular transport through an asymmetric membrane channel," *Phys. Rev. Lett.*, vol. 93, no. 23, pp. 238102, Nov. 2004.

[24] F. Zhu, E. Tajkhorshid and K. Schulten, "Collective diffusion model for water permeation through microscopic channels", *Phys. Rev. Lett.*, vol. 93, no. 22, pp. 224501, Nov. 2004.

[25] M. O. Jensen, E. Tajkhorshid, and K. Schulten, "The mechanism of glycerol conduction in aquaglyceroporins," *Structure*, vol. 9, 2001, pp. 1083–1093.

[26] D. Lu, P. Grayson, and K. Schulten, "Glycerol conductance and physical asymmetry of the Escherichia coli glycerol facilitator GlpF," *Biophysical Journal*, vol. 85, pp. 2977–2987, 2003.

[27] D. Fu, A. Libson, J. W. Miercke, C. Weitzman, P. Nollert, J. Krucinski, and R. M. Stroud, "Structure of a glycerol conducting channel and the basis for its selectivity," *Science*, vol. 290, pp. 481–486, 2000.

[28] E. Mitscherlich and E. H. Marth, *Microbial Survival in the Environment*, Springer-Verlag, Berlin, Germany, 1984.

[29] F. C. Neidhart (Ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ASM Press, Washington, DC.

[30] D. Endy, "Foundations for engineering biology," *Nature*, vol. 438, no. 24, Nov. 2005.

[31] R. Grunberg and L. Serrano, "Strategies for protein synthetic biology," *Nucleic Acids Research*, vol. 38, no. 8, 2010.