Large-Population Cost-Coupled LQG Problems With Nonuniform Agents: Individual-Mass Behavior and Decentralized ε-Nash Equilibria

Minyi Huang, *Member, IEEE*, Peter E. Caines, *Fellow, IEEE*, and Roland P. Malhamé, *Member, IEEE*

***Abstract*—We consider linear quadratic Gaussian (LQG) games in large population systems where the agents evolve according to nonuniform dynamics and are coupled via their individual costs. A state aggregation technique is developed to obtain a set of decentralized control laws for the individuals which possesses an ε-Nash equilibrium property. A stability property of the mass behavior is established, and the effect of inaccurate population statistics on an isolated agent is also analyzed by variational techniques.**

***Index Terms*—Cost-coupled agents, large-scale systems, linear quadratic Gaussian (LQG) systems, Nash equilibria, noncooperative games, state aggregation.**

I. Introduction

THE control and optimization of large-scale complex systems is of importance due to their ubiquitous appearance in engineering, industrial, social, and economic settings. These systems are usually characterized by features such as high dimensionality and uncertainty, and the system evolution is associated with complex interactions among its constituent parts or subsystems. Techniques for dealing with various large-scale systems include model reduction, aggregation, and hierarchical optimization, just a few examples being [1], [7], [22], [28], and [29].

In many social, economic, and engineering models, the individuals or agents involved have conflicting objectives and it is natural to consider optimization based upon individual payoffs or costs. This gives rise to noncooperative game theoretic approaches partly based upon the vast corpus of relevant work within economics and the social sciences [6], [11], [12]; for recent engineering applications, see [9] and [10]. Game theoretic approaches may effectively capture the individual interest seeking nature of agents; however, in a large-scale dynamic model this approach results in an analytic complexity which is in general prohibitively high. We note that the so-called evolutionary games which have been used to treat large population dynamic models at reduced complexity [6], [12] are useful mainly for analyzing the asymptotic behavior of the overall system, and do not lead to a satisfactory framework for the dynamic quantitative optimization of individual performance.

In this paper, we investigate the optimization of large-scale linear quadratic Gaussian (LQG) control systems wherein many agents have similar dynamics and will evolve independently when state regulation is not included. To facilitate our exposition the individual cost based optimization shall be called the *LQG game*. In this framework, each agent is weakly coupled with the other agents only through its cost function. The study of such large-scale cost-coupled systems is motivated by a variety of scenarios, for instance, dynamic economic models involving agents linked via a market [11], [21], and wireless power control [15], [16]. LQ and LQG games have been considered by other authors, see, e.g., [3], [23]–[25], and [27], where the coupled Riccati equations play an important role for studying feedback Nash equilibria. In general, it is very difficult to compute solutions to these equations, even if they exist. For dynamic games with weak coupling in dynamics and costs, extensive effort has been devoted to numerical approximations, see, e.g., [25], [30]. In contrast with such existing research, our concentration is on games with large populations. We analyze the ε-Nash equilibrium properties for a control law by which each individual optimizes its cost function using *local information* depending upon its own state and the average state of all other agents taken together, hereon referred to as "the mass." In this setup, a given agent knows its own dynamics, and the information concerning other agents is available in a statistical sense as described by a randomized parametrization for agents' dynamics across the population.

Due to the particular structure of the individual cost, the mass formed by all other agents impacts any given agent as a nearly deterministic quantity. In response to any known mass influence, a given individual will select its localized control strategy to minimize its own cost. In a practical situation the mass influence cannot be assumed known *a priori*. It turns out, however, that this does not present any difficulty for applying the individual-mass interplay methodology in a population limit framework. In the noncooperative game setup, an overall rationality assumption for the population, to be characterized later, implies the potential of achieving a stable predictable mass behavior in the following sense: If some deterministic mass behavior were

Manuscript received October 4, 2005; revised August 23, 2006 and November 1, 2006. Recommended by Associate Editor J. Hespanha. This work was supported in part by the Australian Research Council (ARC) and by the Natural Sciences and Engineering Research Council of Canada (NSERC). Part of M. Huang's work was performed at the Department of Electrical and Electronic Engineering, The University of Melbourne, Australia, during 2004-2005.

M. Huang is with the School of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: mhuang@math.carleton.ca).

P. E. Caines is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada (e-mail: peterc@cim.mcgill.ca).

R. P. Malhamé is with the Department of Electrical Engineering, École Polytechnique de Montréal, Montreal, QC H3C 3A7, Canada (e-mail: roland.malhame@polymtl.ca).

Digital Object Identifier 10.1109/TAC.2007.904450

to be given, rationality would require that each agent synthesize its individual cost based optimal response as a *tracking* action. Thus the mass trajectory corresponding to rational behavior would guide the agents to collectively generate the trajectory which, individually, they were assumed to be reacting to in the first place.

Indeed, if a mass trajectory with this fixed point property existed, if it were unique, and, furthermore, if each individual had enough information to compute it, then rational agents who were assuming all other agents to be rational would anticipate their collective state of agreement and select a control policy consistent with that state. Thus, in the context of our game problem, we make the following rationality assumption which forms the basis for each agent's optimization behavior: Each agent is rational in the sense that it both: (i) optimizes its own cost function; and (ii) assumes that all other agents are being simultaneously rational when evaluating their competitive behavior.

In fact, the resulting situation is seen to be that of a Nash equilibrium holding between any agent and the mass of the other agents. This equilibrium then has the rationality and optimality interpretations but we underline that these rationality hypotheses are not employed in the mathematical derivation of the results.

It is worth noting that the large population limit formulation presented here is relevant to economic problems involving (mainly static) models with a large number or a continuum of agents; see, e.g., [13]. Instead of a direct continuum population modeling, we induce a probability distribution on a parameter space via empirical statistics; this approach avoids certain measurability difficulties arising in continuum population modeling [19]. Furthermore, we develop *state aggregation* using the population empirical distribution, and our approach differs from the well-known aggregation techniques based upon time-scales [26], [28].

The paper is organized as follows: in Section II we introduce the dynamically independent and cost-coupled systems. Section III gives preliminary results on linear tracking. Section IV contains the individual and mass behavior analysis via a state aggregation procedure. In Section V we establish the ε-Nash equilibrium property of the decentralized individual control laws. Section VI illustrates a cost gap between the individual and global cost based controls, and Section VII addresses the effect of inaccuracy in population statistics.

II. Dynamically Independent and Cost-Coupled Systems

We consider an n dimensional linear stochastic system where the evolution of each state component is described by

$$dz_i = (a_i z_i + b_i u_i)dt + \sigma_i dw_i, \quad 1 \le i \le n, \ t \ge 0 \quad (2.1)$$

where $\{w_i, 1 \le i \le n\}$ denotes n independent standard scalar Wiener processes. The initial states $z_i(0)$ are mutually independent and are also independent of $\{w_i, 1 \le i \le n\}$. In addition, $E|z_i(0)|^2 < \infty$ and $b_i \neq 0$. Each state component shall be referred to as the state of the corresponding individual (also to be called an agent or a player).

We investigate the behavior of the agents when, apart from the impact of feedback, they only interact with each other through coupling terms appearing in their individual cost functions:

$$J_i(u_i, v_i) \triangleq E \int_0^\infty e^{-\rho t} \left[(z_i - v_i)^2 + r u_i^2\right] dt. \quad (2.2)$$

We term this type of model a *dynamically independent and cost-coupled system*. For simplicity of analysis, we assume in this paper

$$b_i = b > 0, \quad 1 \le i \le n.$$

In particular, we assume the cost-coupling to be of the following form for most of our analysis:

$$v_i = \gamma \left(\frac{1}{n} \sum_{k \neq i}^{n} z_k + \eta \right)$$

and we study the large-scale system behavior in the dynamic noncooperative game framework. Evidently, the linking term v_i gives a measure of the average effect of the mass formed by all other agents. Here we assume $\rho, r, \gamma, \eta > 0$, and throughout the paper z_i is described by the dynamics (2.1).

A. A Production Output Planning Example

This production output adjustment problem is based upon the early work [2] where a quadratic nonzero-sum game was considered for a static duopoly model with linear price decrease when the total production level increases. A noncooperative game was also formulated by Lambson for output adjustment in a large dynamic market via nonlinear payoff and price models [21]. Here we study a dynamic model consisting of many players. Our formulation differs from Lambson's dynamic optimization in that we introduce dynamics into the adjustment and further assume that the rate of change of the production output level incurs an explicit cost.

Consider n firms $F_i, 1 \le i \le n$, supplying the same product to the market. First, let x_i be the production level of firm F_i and suppose x_i is subject to adjustment by the following model:

$$dx_i = u_i dt + \sigma_i dw_i, \quad t \ge 0 \quad (2.3)$$

which is a special form of (2.1). Here u_i denotes the action of increasing or decreasing the production level x_i, and $\sigma_i dw_i$ denotes uncertainty in the change of x_i.

Second, we assume the price of the product is given by

$$p = \overline{\eta} - \overline{\gamma} \left(\frac{1}{n} \sum_{i=1}^{n} x_i \right) \quad (2.4)$$

where $\overline{\eta}$, $\overline{\gamma} > 0$. In (2.4) the overall production level $Q \triangleq \sum_{i=1}^n x_i$ is scaled by the factor $1/n$. A justification for scaling is that this may be used to model the situation when an increasing number of firms distributed over different areas join together to serve an increasing number of consumers. In fact, (2.4) may be regarded as a simplified form of a more general price model introduced by Lambson for many agents producing the same goods [21].

We now assume that firm F_i adjusts its production level x_i by referring to the current price of the product. Indeed, an increasing price calls for more supplies of the product to consumers and a decreasing price for less.

We seek a production level which is approximately in proportion to the price that the current market provides, i.e.

$$x_i \approx \beta p = \beta \left[\overline{\eta} - \overline{\gamma} \left(\frac{1}{n} \sum_{i=1}^{n} x_i \right) \right] \tag{2.5}$$

where $\beta > 0$ is a constant. Based upon (2.5), we introduce a penalty term $\{x_i - \beta[\overline{\eta} - \overline{\gamma}((1/n)\sum_{i=1}^{n} x_i)]\}^2 \triangleq (x_i - v_i^0)^2$. On the other hand, in the adjustment of x_i, the control u_i corresponds to actions of shutting down or restarting production lines, or even the construction of new ones; these may further lead to hiring or laying off workers [5]. Each of these actions will incur certain costs to the firm; for simplicity we denote the instantaneous cost of the adjustment by ru_i^2, where $r > 0$. We now write the infinite horizon discounted cost for firm F_i as follows:

$$J_i^x \left(u_i, v_i^0\right) = E \int_0^\infty e^{-\rho t} \left[\left(x_i - v_i^0\right)^2 + r u_i^2 \right] dt \tag{2.6}$$

where $\rho > 0$ and we use the superscript in J_i^x to indicate that the associated dynamics are (2.3). Due to the penalty on the change rate u_i, this situation may be regarded as falling into the framework of smooth planning [5], [20]. Here obviously $v_1^0 = \cdots = v_n^0$. Notice that

$$v_i^0 = \beta \left[\overline{\eta} - \overline{\gamma} \left((1/n) \sum_{i=1}^{n} x_i \right) \right]$$

in this example and

$$v_i = \gamma \left((1/n) \sum_{k \neq i}^{n} z_k + \eta \right)$$

in the generic case (2.2) share the common feature of taking an average over a mass.

III. The Preliminary Linear Tracking Problem

To begin with, for large n, assume

$$z_{-i}^* \triangleq \gamma \left((1/n) \sum_{k \neq i}^{n} z_k + \eta \right)$$

in Section II is approximated by a *deterministic* continuous function z^* defined on $[0, \infty)$. For a given z^*, we construct the *individual cost* for the ith player as follows:

$$J_i(u_i, z^*) = E \int_0^\infty e^{-\rho t} \left\{ [z_i - z^*]^2 + r u_i^2 \right\} dt. \tag{3.1}$$

And for this cost we shall consider the general tracking problem with bounded z^*. For minimization of J_i, the admissible control set is taken as $\mathcal{U}_i \triangleq \{u_i | u_i$ is adapted to the σ-algebra $\sigma(z_i(0), w_i(s), s \le t)$, and $E \int_0^\infty e^{-\rho t} (z_i^2 + u_i^2)\, dt < \infty\}$. The set $\mathcal{U}_i$ is nonempty due to controllability of (2.1). Define $C_b[0, \infty) \triangleq \{x \in C[0, \infty), |x|_\infty < \infty\}$, where $|x|_\infty = \sup_{t \ge 0} |x(t)|$, for $x \in C[0, \infty)$. Under the norm $|\cdot|_\infty$, $C_b[0, \infty)$ is a Banach space [32].

Let Π_i be the positive solution to the algebraic Riccati equation

$$\rho \Pi_i = 2 a_i \Pi_i - \frac{b^2}{r} \Pi_i^2 + 1. \tag{3.2}$$

It is easy to verify that $-a_i + (b^2 \Pi_i / r) + (\rho/2) > 0$. Denote

$$\beta_1 = -a_i + \frac{b^2}{r} \Pi_i, \quad \beta_2 = -a_i + \frac{b^2}{r} \Pi_i + \rho. \tag{3.3}$$

Clearly, $\beta_2 > (\rho/2)$. Propositions (3.1)–(3.3) may be proved by an algebraic approach [4], [14], [17].

Proposition 3.1: Assume (i) $E|z_i(0)|^2 < \infty$ and $z^* \in C_b[0, \infty)$; (ii) $\Pi_i > 0$ is the solution to (3.2) and $\beta_1 = -a_i + (b^2/r)\Pi_i > 0$; and (iii) $s_i \in C_b[0, \infty)$ is determined by the differential equation

$$\rho s_i = \frac{ds_i}{dt} + a_i s_i - \frac{b^2}{r} \Pi_i s_i - z^*. \tag{3.4}$$

Then the control law

$$\widehat{u}_i = -\frac{b}{r} (\Pi_i z_i + s_i) \tag{3.5}$$

minimizes $J_i(u_i, z^*)$, for all $u_i \in \mathcal{U}_i$. □

Proposition 3.2: Suppose assumptions (i)-(iii) in Proposition 3.1 hold and $q \in C_b[0, \infty)$ satisfies

$$\rho q = \frac{dq}{dt} - \frac{b^2}{r} s_i^2 + (z^*)^2 + \sigma_i^2 \Pi_i. \tag{3.6}$$

Then the cost for the control law (3.5) is given by $J_i(\widehat{u}_i, z^*) = \Pi_i E z_i^2(0) + 2 s(0) E z_i(0) + q(0)$. □

Remark: In Proposition 3.1, assumption (i) ensures that J_i has a finite minimum attained at some $u_i \in \mathcal{U}_i$. Assumption (ii) means that the resulting closed-loop system has a stable pole. □

Remark: We point out that s_i in Proposition 3.1 may be uniquely determined only utilizing its boundedness, and it is unnecessary to specify the initial condition for (3.4) separately. Similarly, after $s_i \in C_b[0, \infty)$ is obtained, q in Proposition 3.2 can be uniquely determined from its boundedness. We state these facts in Proposition 3.3. □

Proposition 3.3: Under the assumptions of Proposition 3.1, there exists a unique initial condition $s_i(0) \in \mathbb{R}$ such that the associated solution s_i to (3.4) is bounded, i.e., $s_i \in C_b[0, \infty)$. And moreover, for the obtained $s_i \in C_b[0, \infty)$, there is a unique initial condition $q(0) \in \mathbb{R}$ for (3.6) such that the solution $q \in C_b[0, \infty)$. □

IV. Competitive Behavior and Continuum Mass Behavior

For control synthesis, each agent is assumed to be rational in the sense that it both optimizes its own cost and its strategy is based upon the assumption that the other agents are rational. Due to the specific cost structure, under the rationality assumption it is possible for the agents to achieve mutual anticipation and approximate the linking term v_i by a purely deterministic process z^*, and as a result, if deterministic tracking is employed by the ith agent, its optimality loss will be negligible in large population conditions. Hence, over the large population, all agents would tend to adopt such a tracking based control strategy if an approximating z^* were to be given.

However, we stress that the rationality notion is only used to construct the aggregation procedure, and all the theorems in the paper will be based solely upon their mathematical assumptions.

A. State Aggregation

Assume $z^* \in C_b[0,\infty)$ is given, and $s_i \in C_b[0,\infty)$ is a solution to (3.4). For the ith agent, after applying the optimal tracking based control law (3.5), the closed loop equation is

$$dz_i = \left(a_i - \frac{b^2}{r}\Pi_i\right) z_i dt - \frac{b^2}{r} s_i dt + \sigma_i dw_i. \tag{4.1}$$

Denoting $\overline{z}_i(t) = Ez_i(t)$ and taking expectation on both sides of (4.1) yields

$$\frac{d\overline{z}_i}{dt} = \left(a_i - \frac{b^2}{r}\Pi_i\right)\overline{z}_i - \frac{b^2}{r}s_i \tag{4.2}$$

where the initial condition is $\overline{z}_i|_{t=0} = Ez_i(0)$. We further define the population average of means (simply called population mean) as $\overline{z}^{(n)} \triangleq (1/n)\sum_{i=1}^n \overline{z}_i$.

So far, the individual reaction is easily determined if a mass effect z^* is given *a priori*. Here one naturally comes up with the important question: how is the deterministic process z^* chosen? Since we wish to have $z^*(t) \approx z^*_{-i} = \gamma\left((1/n)\sum_{k\neq i}^n z_k + \eta\right)$, for large n it is reasonable to express z^* in terms of the population mean $\overline{z}^{(n)}$ as

$$z^*(t) = \gamma(\overline{z}^{(n)}(t) + \eta). \tag{4.3}$$

As n increases, accuracy of the approximation of z^*_{-i} by z^* given in (4.3) is expected to improve.

Our analysis below will be based upon the observation that the large population limit may be employed to determine the effect of the mass of the population on any given individual. Specifically, our interest is in the case when a_i, $i \geq 1$, is "adequately randomized" in the sense that the population exhibits certain statistical properties. In this context, the association of the value a_i, $i \geq 1$ and the specific index i plays no essential role, and the more important fact is the frequency of occurrence of a_i on different segments in the range space of the sequence $\{a_i, i \geq 1\}$.

Within this setup, we assume that the sequence $\{a_i, i \geq 1\}$, has an empirical distribution function $F(a)$, for which a more detailed specification will be stated in Section V. For the Riccati equation (3.2), when the coefficient a is used in place of a_i, we denote the corresponding solution by Π_a. Accordingly, we express $\beta_1(a)$ and $\beta_2(a)$ when a and Π_a are substituted into (3.3). Straightforward calculation gives

$$\Pi_a = \left(\frac{b^2}{r}\right)^{-1}\left[a - \frac{\rho}{2} + \sqrt{\left(a-\frac{\rho}{2}\right)^2 + \frac{b^2}{r}}\right]$$

$$\beta_1(a) = -\frac{\rho}{2} + \sqrt{\left(a-\frac{\rho}{2}\right)^2 + \frac{b^2}{r}} \tag{4.4}$$

$$\beta_2(a) = \frac{\rho}{2} + \sqrt{\left(a-\frac{\rho}{2}\right)^2 + \frac{b^2}{r}}. \tag{4.5}$$

1) Example 4.1: For the set of parameters: $a = 1$, $b = 1$, $\sigma = 0.3$, $\rho = 0.5$, $\gamma = 0.6$, $r = 0.1$, $\eta = 0.25$, we have $\Pi_a = 0.4$, $\beta_1(a) = 3$, $\beta_2(a) = 3.5$. □

To simplify the aggregation procedure we assume zero mean for initial conditions of all agents, i.e., $Ez_i(0) = 0$, $i \geq 1$. The above analysis suggests we introduce the equation system:

$$\rho s_a = \frac{ds_a}{dt} + as_a - \frac{b^2}{r}\Pi_a s_a - z^* \tag{4.6}$$

$$\frac{d\overline{z}_a}{dt} = \left(a - \frac{b^2}{r}\Pi_a\right)\overline{z}_a - \frac{b^2}{r}s_a \tag{4.7}$$

$$\overline{z} = \int_{\mathcal{A}} \overline{z}_a dF(a) \tag{4.8}$$

$$z^* = \gamma(\overline{z} + \eta). \tag{4.9}$$

In the above, each individual equation is indexed by the parameter a. For the same reasons as noted in Proposition 3.3, here it is unnecessary to specify the initial condition for s_a. Equation (4.7) with $\overline{z}_a|_{t=0} = 0$ is based upon (4.2). Hence $\overline{z}_a$ is regarded as the expectation given the parameter a in the individual dynamics. Also, in contrast to the arithmetic average for computing $\overline{z}^{(n)}$ appearing in (4.3), (4.8) is derived by use of an empirical distribution function $F(a)$ for the sequence of parameters $a_i \in \mathcal{A}$, $i \geq 1$, with the range space $\mathcal{A}$. Equation (4.9) is the large population limit form for the equality relation (4.3). With a little abuse of terminology, we shall conveniently refer to either z^* being affine in $\overline{z}$, or in some cases $\overline{z}$ itself as the mass trajectory.

Remark: For the sake of simplicity, the integration in (4.8) assumes a zero mean initial condition for all agents. In the more general case with nonzero $Ez_i(0)$, we may introduce a joint empirical distribution $F_{a,z}$ for the two dimensional sequence $\{(a_i, Ez_i(0)), i \geq 1\}$. Then the function in (4.7) is to be labelled by both the dynamic parameter a and an associated initial condition, and furthermore, the integration in (4.8) is computed with respect to $F_{a,z}$. In this paper we only consider the zero mean $Ez_i(0)$ to avoid notational complication. □

We introduce the assumptions:

(H1) For $\beta_1(a)$, $\beta_2(a)$ defined by (4.4)–(4.5), $\beta_1(a) > 0$ holds for all $a \in \mathcal{A}$, and $\int_{\mathcal{A}}(M/(\beta_1(a)\beta_2(a)))dF(a) < 1$, where $M = (b^2\gamma/r)$, $\mathcal{A}$ is an interval containing all a_i, $i \geq 1$, and $F(a)$ is the empirical distribution function for $\{a_i, i \geq 1\}$, which is assumed to exist. □

(H2) All agents have zero mean initial condition, i.e., $Ez_i(0) = 0$, $i \geq 1$. □

We state a sufficient condition to ensure $\beta_1(a) > 0$ for $a \in (-\infty,\infty)$. The proof is trivial and is omitted.

Proposition 4.2: If $b^2 > (r\rho^2/4)$, then $\beta_1(a) > 0$ for all $a \in (-\infty,\infty)$. □

Remark: Under **(H1)**, we have $-\beta_2(a) < -\beta_1(a) < 0$ where $-\beta_1(a)$ is the stable pole of the closed-loop system for the agent with parameter a and $|\beta_1(a)|$ measures the stability margin. The ratio $(M/(\beta_1(a)\beta_2(a)) = b^2\gamma/[r\beta_1(a)(\beta_1(a)+\rho)]$ depends upon the stability margin and the linking parameter γ. Notice that **(H1)** holds for the system of n uniform agents with parameters specified in Example 4.1. □

Given $z^* \in C_b[0,\infty)$, Proposition 3.3 implies that (4.6) leads to the bounded solution

$$s_a(t) = -e^{\beta_2(a)t}\int_t^\infty e^{-\beta_2(a)\tau} z^*(\tau)d\tau \triangleq \mathcal{T}_1 z^*. \tag{4.10}$$

By expressing $\overline{z}_a(t)$ and $\overline{z}(t)$ in terms of s_a, we obtain from (4.9)

$$\begin{aligned} z^*(t) = &\frac{\gamma b^2}{r}\int_{\mathcal{A}}\int_0^t\int_s^\infty e^{-\beta_1(a)(t-s)} \\ &\cdot e^{-\beta_2(a)(\tau-s)} z^*(\tau)d\tau ds dF(a) + \gamma\eta \\ \triangleq &(\mathcal{T}z^*)(t). \end{aligned} \tag{4.11}$$

The theorems below may be proved by following the methods in [14] and [17].

Theorem 4.3: Under (**H1**), we have (i) $\mathcal{T}x \in C_b[0,\infty)$, for any $x \in C_b[0,\infty)$, and (ii) $\mathcal{T}: C_b[0,\infty) \to C_b[0,\infty)$ has a unique fixed point which is uniformly Lipschitz continuous on $[0,\infty)$. □

Theorem 4.4: Under (**H1**)–(**H2**), the equation system (4.6)–(4.9) admits a unique bounded solution. □

B. The Virtual Agent and Policy Iteration

We proceed to investigate asymptotic properties of the interaction between the individual and the mass, and the formulation shall be interpreted in the large population limit (i.e., an infinite population) context. Assume each agent is assigned a cost according to (3.1), i.e.

$$J_i(u_i, z^*) = E\int_0^\infty e^{-\rho t}\left\{[z_i - z^*]^2 + r u_i^2\right\} ds, \quad i \geq 1. \tag{4.12}$$

We now introduce a so-called *virtual agent* to represent the mass effect and use $z^* \in C_b[0,\infty)$ to describe the behavior of the virtual agent. Here the virtual agent acts as a passive player in the sense that z^* appears as an exogenous function of time to be tracked by the agents.

After each selection of the set of individual control laws, a new z^* will be induced as specified below; subsequently, the individual shall consider its optimal policy (over the whole time horizon) to respond to this updated z^*. Thus, the interplay between a given individual and the virtual agent may be described as a sequence of virtual plays which may be employed by the individual as a calculation device to eventually learn the mass behavior. In the following policy iteration analysis in function spaces, we take the virtual agent as a *passive leader* and the individual agents as *active followers*.

Now, we describe the iterative update of an agent's policy from its *policy space*. For a fixed iteration number $k \geq 0$, suppose that there is *a priori* $z^{*(k)} \in C_b[0,\infty)$. Then by Proposition 3.1 the optimal control for the ith agent using the cost (4.12) with respect to $z^{*(k)}$ is given as $u_i^{(k+1)} = -(b/r)\left(\Pi_i z_i + s_i^{(k+1)}\right)$ where $s_i^{(k+1)} \in C_b[0,\infty)$ is given by

$$\rho s_i^{(k+1)} = \frac{ds_i^{(k+1)}}{dt} + a_i s_i^{(k+1)} - \frac{b^2}{r}\Pi_i s_i^{(k+1)} - z^{*(k)}. \tag{4.13}$$

By Proposition 3.3, the unique solution $s_i^{(k+1)} \in C_b[0,\infty)$ to (4.13) may be represented by

$$s_i^{(k+1)} = -e^{\beta_2 t}\int_t^\infty e^{-\beta_2\tau} z^{*(k)}(\tau)d\tau. \tag{4.14}$$

Subsequently, the control laws $\left\{u_i^{(k+1)}, i \geq 1\right\}$ produce a mass trajectory $\overline{z}^{(k+1)} = \int_{\mathcal{A}} \overline{z}_a^{(k+1)} dF(a)$, where

$$\frac{d\overline{z}_a^{(k+1)}}{dt} = -\beta_1(a)\overline{z}_a^{(k+1)} - \frac{b^2}{r}s_a^{(k+1)} \tag{4.15}$$

with initial condition $\overline{z}_a^{(k+1)}|_{t=0} = 0$ by (**H2**). Notice that (4.15) is indexed by the parameter $a \in \mathcal{A}$ instead of all $i's$. Then the virtual agent's state (function) z^* corresponding to all $u_i^{(k+1)}$, $i \geq 1$, is determined as $z^{*(k+1)} = \gamma(\overline{z}^{(k+1)} + \eta)$. From the above and using the operator introduced in (4.11), we get the recursion for $z^{*(k)}$ as

$$z^{*(k+1)} = \mathcal{T}z^{*(k)}$$

where $z^{*(k+1)}|_{t=0} = \gamma(\overline{z}^{(k+1)}|_{t=0} + \eta) = \gamma\eta$ for all k.

By the sequential adjustments of the individual strategies in response to the virtual agent, we induce the mass behavior by a sequence of functions $z^{*(k)} = \mathcal{T}z^{*(k-1)} = \mathcal{T}^k z^{*(0)}$. We note that the rule for the individual strategy selection here is comparable to the well-known best response map for static game models [11]. The next proposition, which may be proved by Theorems 4.3–4.4, establishes that as the population grows, a statistical mass equilibrium exists and it is globally attracting.

Proposition 4.5: Under (**H1**)–(**H2**), $\lim_{k\to\infty} z^{*(k)} = z^*$ for any $z^{*(0)} \in C_b[0,\infty)$, where z^* is determined by (4.6)–(4.9). □

C. Explicit Solution With Uniform Agents

For a system of uniform agents, a solution to the state aggregation equation system may be explicitly calculated where $F(a)$ degenerates to point mass and (4.8) is no longer required. Omitting the subscript a for s_a involved, the equation system (4.6)–(4.9) reduces to

$$\rho s = \frac{ds}{dt} + as - \frac{b^2}{r}\Pi_a s - z^* \tag{4.16}$$

$$\frac{d\overline{z}}{dt} = \left(a - \frac{b^2}{r}\Pi_a\right)\overline{z} - \frac{b^2}{r}s \tag{4.17}$$

$$z^* = \gamma(\overline{z} + \eta). \tag{4.18}$$

Here we shall solve the above equation system with a general initial condition $\overline{z}(0)$ which is not necessarily zero. Setting the derivatives to zero, we write a set of steady-state equations as follows:

$$\begin{cases} \gamma\overline{z}_\infty - z^*_\infty = -\gamma\eta \\ z^*_\infty - \left(a - \frac{b^2}{r}\Pi_a - \rho\right)s_\infty = 0 \\ \left(a - \frac{b^2}{r}\Pi_a\right)\overline{z}_\infty - \frac{b^2}{r}s_\infty = 0. \end{cases} \tag{4.19}$$

It can be verified that under **(H1)** the equation (4.19) is nonsingular and has a unique solution $(s_\infty, \overline{z}_\infty, z^*_\infty)$.

Let $\widetilde{z}(t)=\overline{z}(t)-\overline{z}_\infty$ and $\widetilde{s}(t)=s(t)-s_\infty$. Here, we simply write $\beta_1(a)$, $\beta_2(a)$ as β_1, β_2. Eliminating s in (4.17) by (4.10) and (4.18), we get

$$\begin{aligned}\frac{d\widetilde{z}}{dt}=\frac{d\overline{z}}{dt}&=\left(a-\frac{b^2}{r}\Pi_a\right)\overline{z}\\&\quad+M\int_t^\infty e^{\beta_2(t-\tau)}\overline{z}(\tau)d\tau+\frac{M\eta}{\beta_2}\\&=\left(a-\frac{b^2}{r}\Pi_a\right)\widetilde{z}+M\int_t^\infty e^{\beta_2(t-\tau)}\widetilde{z}(\tau)d\tau. \quad (4.20)\end{aligned}$$

Differentiating both sides of (4.20) gives

$$\frac{d^2\widetilde{z}}{dt^2}=-\beta_1\frac{d\widetilde{z}}{dt}+\beta_2M\int_t^\infty e^{\beta_2(t-\tau)}\widetilde{z}(\tau)d\tau-M\widetilde{z}(t)$$

which combined again with (4.20) yields

$$\frac{d^2\widetilde{z}}{dt^2}-\rho\frac{d\widetilde{z}}{dt}+(M-\beta_1\beta_2)\widetilde{z}=0. \quad (4.21)$$

The characteristic equation of (4.21) is $\lambda^2-\rho\lambda+(M-\beta_1\beta_2)=0$ with two distinct eigenvalues: $\lambda_1=\left(\rho-\sqrt{\rho^2+4(\beta_1\beta_2-M)}\right)/2<0$, $\lambda_2=\left(\rho+\sqrt{\rho^2+4(\beta_1\beta_2-M)}\right)/2>0$, where $\beta_1\beta_2-M>0$, which follows from (**H1**) for the degenerate $F(a)$. Recalling the boundedness condition for $\overline{z}$, and, hence, for $\widetilde{z}$, we have $\widetilde{z}=\widetilde{z}(0)e^{\lambda_1 t}=(\overline{z}(0)-\overline{z}_\infty)e^{\lambda_1 t}$, and we can check that $\widetilde{z}$ is indeed a solution to (4.20). Consequently, we may obtain the bounded solution for $\overline{z}$ and s. We summarize the above calculation.

Proposition 4.6: Under **(H1)**, the unique bounded solution $(\overline{z},s)$ in (4.16)–(4.17), is given by $\overline{z}(t)=\overline{z}_\infty+(\overline{z}(0)-\overline{z}_\infty)e^{\lambda_1 t}$ and $s(t)=s_\infty+(\gamma/(\beta_2-\lambda_1))(\overline{z}_\infty-\overline{z}(0))e^{\lambda_1 t}$, where $\lambda_1=\left(\rho-\sqrt{\rho^2+4(\beta_1\beta_2-M)}\right)/2<0$, $\beta_1=-a+(b^2/r)\Pi_a$, and $\beta_2=-a+(b^2/r)\Pi_a+\rho$. □

V. The Decentralized ε-Nash Equilibrium

Let $J_i(u_i,v_i(u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n))$ denote the individual cost with respect to the linking term $v_i=\gamma\left((1/n)\sum_{k\neq i}^n z_k(u_k)+\eta\right)$ for the ith player when the kth player takes control u_k, $1\le k\le n$, and n is the population size. Let (5.1), shown at the bottom of the page hold, where $z_k\left(u_k^0\right)=z_k\left(u_k^0(z^*,z_k)\right)$. Here we use u_i^0 to denote the optimal tracking based control law

$$u_i^0=-\frac{b}{r}(\Pi_i z_i+s_i) \quad (5.2)$$

where s_i and the associated z^* are derived from (4.6)–(4.9). In particular, $J_i\left(u_i^0,v_i\left(u_1^0,\ldots,u_{i-1}^0,u_{i+1}^0\ldots,u_n^0\right)\right)=J_i\left(u_i,v_i\left(u_1^0,\ldots,u_{i-1}^0,u_{i+1}^0\ldots,u_n^0\right)\right)|_{u_i=u_i^0}$. It should be noted that in the following asymptotic analysis the control law u_i^0 for the ith agent among n agents is constructed using the limit empirical distribution $F(a)$ involved in (4.8).

Within the context of a population of n agents, for any $1\le k\le n$, the kth agent's admissible control set $\mathcal{U}_k$ consists of all controls u_k adapted to the σ-algebra $\sigma(z_i(\tau),1\le i\le n,\tau\le t)$ such that the closed-loop for the n agents has a unique solution. In this setup we give the definition.

Definition 5.1: A set of controls $u_k\in\mathcal{U}_k$, $1\le k\le n$, for n players is called an *ε-Nash equilibrium with respect to the costs* J_k, $1\le k\le n$, if there exists $\varepsilon\ge0$ such that for any fixed $1\le i\le n$, we have

$$\begin{aligned}&J_i(u_i,v_i(u_1,\ldots,u_{i-1},u_{i+1}\ldots,u_n))\\&\qquad\le J_i(u_i',v_i(u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n))+\varepsilon \quad (5.3)\end{aligned}$$

when any alternative control $u_i'\in\mathcal{U}_i$ is applied by the ith player. □

If $\varepsilon=0$, Definition 5.1 reduces to the usual definition of a Nash equilibrium.

Remark: The admissible control set $\mathcal{U}_k$ is not decentralized since the kth agent has perfect information on other agents' states. In effect, such admissible control sets lead to a stronger qualification of the ε-Nash equilibrium property for the decentralized control analyzed in this section. □

For the sequence $\{a_i,i\ge1\}$, we define the empirical distribution associated with the first n agents

$$F_n(x)=\frac{1}{n}\sum_{i=1}^n 1_{(a_i<x)},\quad x\in\mathbb{R}.$$

We introduce the assumption on the asymptotic behavior of F_n.

(**H3**) There exits a probability distribution function F on $\mathbb{R}$ such that $F_n\to F$ weakly as $n\to\infty$, i.e., $\lim_{n\to\infty}F_n(x)=F(x)$ whenever F is continuous at $x\in\mathbb{R}$. □

(**H3'**) There exists a probability distribution function F such that $\lim_{n\to\infty}\sup_{x\in\mathbb{R}}|F_n(x)-F(x)|=0$. □

Remark: It is obvious that (**H3'**) implies (**H3**). Notice that if the sequence $a_1^\infty\triangleq\{a_i,i\ge1\}$ is sufficiently "randomized" such that a_1^∞ is generated by independent observations on the same underlying distribution function F, then with probability one (**H3'**) holds by the Glivenko–Cantelli theorem [8]. □

Given the distribution function F and $z^*\in C_b[0,\infty)$, from (4.6)–(4.7) it is seen that both s_a and $\overline{z}_a$ may be explicitly expressed as a function of a taken from the parameter space $\mathcal{A}$.

(**H4**) The interval $\mathcal{A}\supset\{a_i,i\ge1\}$ satisfies: (i) $\mathcal{A}$ is a closed set, $\int_{\mathcal{A}}dF(a)=1$ and $\inf_{a\in\mathcal{A}}\beta_1(a)\ge\widehat{\varepsilon}$ for a constant $\widehat{\varepsilon}>0$; and (ii) $\mathcal{A}$ is bounded. In addition, $\sup_{i\ge1}[\sigma_i^2+Ez_i^2(0)]<\infty$. □

Remark: For compact $\mathcal{A}$, the positivity condition for $\beta_1(a)$ in (**H1**) implies $\widehat{\varepsilon}$ specified above always exists. Condition (ii)

$$J_i\left(u_i,v_i\left(u_1^0,\ldots,u_{i-1}^0,u_{i+1}^0\ldots,u_n^0\right)\right)\triangleq E\int_0^\infty e^{-\rho t}\left\{\left[z_i(u_i)-\gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right]^2+ru_i^2\right\}dt \quad (5.1)$$

is only used for the performance estimate in proving Theorem 5.6, and it may be relaxed. □

Remark: The method in this section may deal with the case where $\mathcal{A}$ consists of a finite number of disjoint closed subintervals

$$\mathcal{I}_i = \left[c_l^{(i)}, c_r^{(i)}\right], \quad 1 \le i \le N$$

where

$$c_r^{(1)} < c_l^{(2)} \le c_r^{(2)} < \cdots < c_l^{(N)}.$$

This includes the special case of uniform agents. □

We give two auxiliary lemmas in a more general form without restricting $\mathcal{A}$ to be compact.

Lemma 5.2: Assume (**H1**) and (**H4**)-(i) hold. Then $\overline{z}_a(t)$ satisfies: (i) $\sup_{a\in\mathcal{A}} |\overline{z}_a|_\infty < \infty$; and (ii) $\lim_{a'\to a}\sup_t |\overline{z}_a(t) - \overline{z}_{a'}(t)| = 0$ with a vanishing rate depending only on $|a-a'|$, for a, $a' \in \mathcal{A}$. □

Lemma 5.2 may be proved by the method in [17]. Now we define

$$\epsilon_n(t) = \left|\int_{\mathcal{A}} \overline{z}_a(t)dF_n(a) - \int_{\mathcal{A}} \overline{z}_a(t)dF(a)\right|, \quad t \ge 0. \quad (5.4)$$

Lemma 5.3: Under (**H1**)–(**H3**) and (**H4**)-(i), we have $\lim_{n\to\infty}\overline{\epsilon}_n \triangleq \lim_{n\to\infty}\sup_{t\ge 0}\epsilon_n(t) = 0$, where $\epsilon_n(t)$ is defined by (5.4).

Proof: Letting $I_C = [-C, C]$ for $C > 0$, we have

$$\begin{aligned}\epsilon_n(t) = &\left|\int_{\mathcal{A}\cap I_C} \overline{z}_a(t)dF_n(a) + \int_{\mathcal{A}\cap(\mathbb{R}\setminus I_C)} \overline{z}_a(t)dF_n(a)\right.\\ &\left. - \int_{\mathcal{A}\cap I_C} \overline{z}_a(t)dF(a) - \int_{\mathcal{A}\cap(\mathbb{R}\setminus I_C)} \overline{z}_a(t)dF(a)\right|\\ \triangleq &\left|I_n^{(1)} + I_n^{(2)} - I^{(1)} - I^{(2)}\right|.\end{aligned}$$

Now for any fixed $\varepsilon > 0$, there exists a sufficiently large constant $C > 0$ such that F is continuous at $a = \pm C$ and such that

$$\begin{aligned}\left|I_n^{(2)}\right| + |I^{(2)}| \le &\sup_{a\in\mathcal{A}}\sup_t |\overline{z}_a(t)|\\ &\cdot \left[\sup_n \int_{|a|\ge C} dF_n(a) + \int_{|a|\ge C} dF(a)\right]\\ \le &\frac{\varepsilon}{2}\end{aligned}$$

where we get the last inequality by Lemma 5.2 and the tightness of $\{F_n\}$ implied by (**H3**) (see, e.g., [8, p. 276]). We write

$$\begin{aligned}\left|I_n^{(1)} - I^{(1)}\right| &= \left|\int_{\mathcal{A}\cap I_C} \overline{z}_a(t)dF_n(a) - \int_{\mathcal{A}\cap I_C} \overline{z}_a(t)dF(a)\right|\\ &= \left|\int_{I_C} \overline{z}'_a(t)dF_n(a) - \int_{I_C} \overline{z}'_a(t)dF(a)\right|\end{aligned}$$

where we make the convention that the domain of $\overline{z}_a(t)$ (as a function of a), if necessary, is extended from $\mathcal{A}$ to $\mathbb{R}$ such that properties (i) and (ii) specified in Lemma 5.2 still hold after $\mathcal{A}$ is replaced by $\mathbb{R}$. We denote the resulting function by $\overline{z}'_a(t)$ which is identical to $\overline{z}_a(t)$ on $\mathcal{A}$. For instance, in the case $\mathcal{A} = [c_1, c_2]$ with $c_2 < \infty$, we may simply set $\overline{z}'_a(t) = \overline{z}_{c_2}(t)$ when $a > c_2$. Such an extension can deal with the general case when $\mathcal{A}$ consists of a finite number of disjoint closed subintervals.

Next we combine the equicontinuity of $\overline{z}'_a(t)$ in $a \in I_C$ with $t \in [0, \infty)$, ensured by Lemma 5.2 and the above extension procedure, continuity of F at $a = \pm C$, and the standard subinterval dividing technique for the proof of Helly-Bray theorem (see [8, pp. 274,275]) to conclude that there exists a sufficiently large n_0 such that for all $n \ge n_0$,

$$\left|I_n^{(1)} - I^{(1)}\right| = \left|\int_{I_C} \overline{z}'_a(t)dF_n(a) - \int_{I_C} \overline{z}'_a(t)dF(a)\right| \le \frac{\varepsilon}{2}$$

for any fixed ε, and consequently $\lim_{n\to\infty}\sup_{t\ge 0}\epsilon_n(t) = 0$. □

In the proof of Lemma 5.3, in order to preserve the boundedness and equicontinuity (with respect to a) properties, we extend $\overline{z}_a(t)$ to $a \notin \mathcal{A}$ in a specific manner and avoid directly using (4.6)–(4.9) to calculate $\overline{z}_a(t)$, $a \notin \mathcal{A}$, even if the equation system may give a well defined $\overline{z}_a(t)$ for some $a \notin \mathcal{A}$.

Lemma 5.4: Under (**H1**)–(**H4**), for z^* determined by (4.6)–(4.9), we have

$$\begin{aligned}E\int_0^\infty e^{-\rho t}\left[z^* - \gamma\left(\frac{1}{n}\sum_{k\ne i}^n z_k\left(u_k^0\right) + \eta\right)\right]^2 dt\\ = O\left(\overline{\epsilon}_n^2 + \frac{\gamma^2}{n}\right) \quad (5.5)\end{aligned}$$

where $\overline{\epsilon}_n$ is given in Lemma 5.3 and the state $z_k(u_k^0)$ of the kth player, $k \ne i$, is generated by the control law u_k^0 given by (5.2).

Proof: In the proof we shall omit the control u_k^0 associated with z_k in various places. Obviously, we have

$$(1/n)\sum_{k=1}^n Ez_k = \int_{\mathcal{A}} \overline{z}_a dF_n(a).$$

Setting

$$\begin{aligned}\Psi &\triangleq z^* - \gamma\left((1/n)\sum_{k\ne i}^n z_k + \eta\right)\\ &= \gamma\left(\int_{a\in\mathcal{A}} \overline{z}_a dF(a) - (1/n)\sum_{k\ne i}^n z_k\right)\end{aligned}$$

we obtain from Lemma 5.3

$$\begin{aligned}E\Psi^2(t) = &\gamma^2 E\left\{\left[\int_{\mathcal{A}} \overline{z}_a dF(a) - \frac{1}{n}\sum_{k=1}^n Ez_k\right]\right.\\ &\left. + \left[\frac{1}{n}\sum_{k=1}^n Ez_k - \frac{1}{n}\sum_{k\ne i}^n z_k\right]\right\}^2\\ \le &2\gamma^2\overline{\epsilon}_n^2 + 2\gamma^2 E\left[\frac{1}{n}\sum_{k\ne i}^n (z_k - Ez_k) - \frac{Ez_i}{n}\right]^2\\ \le &2\gamma^2\overline{\epsilon}_n^2 + O\left(\frac{\gamma^2}{n}\right) \quad (5.6)\end{aligned}$$

where, by each agent's closed-loop stability $(\inf_k \beta_1(a_k) > 0)$ and $\sup_{k,t}|s_k(t)| \le \sup_{a,t}|s_a(t)| < \infty$, we can show that the higher-order term holds uniformly with respect to $t \ge 0$, and (5.5) readily follows. □

Theorem 5.5: Under (**H1**)–(**H4**), we have

$$\left| J_i\left(u_i^0, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right) - J_i(u_i^0, z^*)\right| = O\left(\overline{\epsilon}_n + \frac{\gamma}{\sqrt{n}}\right)$$

as $n \to \infty$, where $J_i(u_i^0, z^*)$ is the individual cost with respect to z^*, u_i^0 is given by (5.2), and $\overline{\epsilon}_n$ is given in Lemma 5.3. □

The proof is done by a similar decomposition technique as in proving Theorem 5.6 below and is postponed until after the proof of the latter.

Theorem 5.6: Under (**H1**)–(**H4**), the set of controls u_i^0, $1 \le i \le n$, for the n players is an ε-Nash equilibrium with respect to the costs $J_i\left(u_i, \gamma\left((1/n)\sum_{k\neq i}^n z_i(u_k)+\eta\right)\right)$, $1 \le i \le n$, i.e.

$$\begin{aligned} J_i\left(u_i^0, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_i\left(u_k^0\right)+\eta\right)\right) - \varepsilon &\le \inf_{u_i} J_i\left(u_i, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_i\left(u_k^0\right)+\eta\right)\right) \\ &\le J_i\left(u_i^0, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_i\left(u_k^0\right)+\eta\right)\right) \end{aligned} \tag{5.7}$$

where $0 < \varepsilon = O\left(\overline{\epsilon}_n + (\gamma/\sqrt{n})\right)$ with $\overline{\epsilon}_n$ given in Lemma 5.3 as $n \to \infty$, u_k^0 is given by (5.2), and $u_i \in \mathcal{U}_i$ is any alternative control which depends upon $(t, z_1, \ldots, z_n)$.

Proof: The second inequality is obviously true. We prove the first one. Consider any full state dependent $u_i \in \mathcal{U}_i$ satisfying

$$J_i\left(u_i, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right) \le J_i\left(u_i^0, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right). \tag{5.8}$$

By (**H4**), it is easy to obtain a uniform (with respect to n) upper bound for the right-hand side (RHS) of (5.8) using each agent's closed-loop stability and boundedness of the feedback gain $-(b/r)\Pi_{a_k}$ due to compactness of $\mathcal{A}$. Hence, for u_i satisfying (5.8), there exists a fixed $C_1 > 0$ independent of n such that

$$\begin{aligned} &J_i\left(u_i, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right) \\ &= E\int_0^\infty e^{-\rho t}\left\{\left[z_i(u_i) - \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right]^2 + ru_i^2\right\}dt \\ &\le C_1. \end{aligned} \tag{5.9}$$

Here and hereafter in the proof, $(z_i(u_i), u_i)$, $\left(z_k\left(u_k^0\right), u_k^0\right)$, $k \neq i$, denote the corresponding state-control pairs. For notational brevity, we may omit the associated control in $z_i(u_i)$, $z_k(u_k^0)$, $k \neq i$ and simply write z_i, z_k without causing confusion. After the controls u_k^0, $k \neq i$, are selected, all z_k, $k \neq i$, are stable since $\inf_k \beta_1(a_k) > 0$ by (**H4**); for (u_i, z_i) satisfying (5.9) there exists $C_2 > 0$ independent of n such that $E\int_0^\infty e^{-\rho t} z_i^2 dt \le C_2$ and $E\int_0^\infty e^{-\rho t}(z_i - z^*)^2 dt \le C_2$, where z^* is the same as in Lemma 5.4.

On the other hand, we have

$$\begin{aligned} &E\int_0^\infty e^{-\rho t}\left\{\left[z_i - \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right]^2 + ru_i^2\right\}dt \\ &= E\int_0^\infty e^{-\rho t} \cdot\left\{\left[(z_i - z^*) + \left(z^* - \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta\right)\right)\right]^2 + ru_i^2\right\}dt \\ &= E\int_0^\infty e^{-\rho t}\left[(z_i - z^*)^2 + ru_i^2\right]dt \\ &\quad + E\int_0^\infty e^{-\rho t}\left[z^* - \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta\right)\right]^2 dt \\ &\quad + 2E\int_0^\infty e^{-\rho t}(z_i - z^*)\left[z^* - \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta\right)\right]dt \\ &\triangleq I_1 + I_2 + I_3. \end{aligned} \tag{5.10}$$

Then we have

$$\begin{aligned} I_1 &= J(u_i, z^*) \ge J_i(u_i^0, z^*) \\ &\ge J_i\left(u_i^0, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right) - O\left(\overline{\epsilon}_n + \frac{\gamma}{\sqrt{n}}\right) \end{aligned} \tag{5.11}$$

$$I_2 = O\left(\overline{\epsilon}_n^2 + \frac{\gamma^2}{n}\right) \tag{5.12}$$

where (5.11) follows from Theorem 5.5 and (5.12) follows from Lemma 5.4. Moreover

$$\begin{aligned} |I_3| &\le 2\int_0^\infty e^{-\rho t}\left[E(z_i - z^*)^2\right]^{1/2} \cdot\left\{E\left[z^* - \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta\right)\right]^2\right\}^{1/2} dt \\ &\le 2\left[\int_0^\infty e^{-\rho t}E(z_i - z^*)^2 dt\right]^{1/2} \cdot\left\{\int_0^\infty e^{-\rho t}E\left[z^* - \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k + \eta\right)\right]^2 dt\right\}^{1/2} \\ &= O(\sqrt{I_2}) = O\left(\overline{\epsilon}_n + \frac{\gamma}{\sqrt{n}}\right). \end{aligned} \tag{5.13}$$

Hence, it follows from the above estimates that there exists $c > 0$ independent of n such that

$$J_i\left(u_i, \gamma\left((1/n)\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right) \geq$$
$$J_i\left(u_i^0, \gamma\left((1/n)\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right) - c\left(\bar{\epsilon}_n + (\gamma/\sqrt{n}\right).$$

In other words, when all the players $k \neq i$ retain their decentralized controls u_k^0 and the ith player is allowed to use a full state based control u_i, it can reduce its cost at most by $O\left(\bar{\epsilon}_n + (\gamma/\sqrt{n})\right)$. This completes the proof. □

Proof of Theorem 5.5: As in (5.10) we make the decomposition

$$\begin{aligned}
&J_i\left(u_i^0, \gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right) \\
&= E\int_0^\infty e^{-\rho t} \\
&\left\{\left[\left(z_i\left(u_i^0\right)-z^*\right)+\left(z^*-\gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right)\right]^2\right. \\
&\left.+ r(u_i^0)^2\right\}dt \\
&= J_i\left(u_i^0, z^*\right) \\
&+ E\int_0^\infty e^{-\rho t}\left[z^*-\gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right]^2 dt \\
&+ 2E\int_0^\infty e^{-\rho t}\left(z_i\left(u_i^0\right)-z^*\right) \\
&\cdot\left[z^*-\gamma\left(\frac{1}{n}\sum_{k\neq i}^n z_k\left(u_k^0\right)+\eta\right)\right] dt \\
&\triangleq J_i\left(u_i^0, z^*\right) + I_2' + I_3'.
\end{aligned} \tag{5.14}$$

Finally, similar to (5.12) and (5.13), we apply Schwarz inequality and Lemma 5.4 to obtain $|I_2'+I_3'| = O\left(\bar{\epsilon}_n + (\gamma/\sqrt{n})\right)$, and this completes the proof. □

It should be noted that the proof of Theorem 5.5 does not depend upon Theorem 5.6.

VI. A Cost Gap Between the Centralized Optimal Control and Decentralized Tracking

In this section, we compare the individual cost based control with a global cost based control and identify a cost gap. Let the global cost be defined as $J = \sum_{i=1}^n J_i$ for a system of n agents, where J_i is defined in (2.2), and we term minimization of J the centralized optimal control problem.

Throughout this section we consider a system of uniform agents under (**H1**) with $a_i = a$, $\sigma_i = \sigma$, and assume the initial state $z_i(0)$ of all agents is 0 in the two cases. We scale by n the global optimal cost (with 0 initial state for all players)

$$v(0) \triangleq \inf J\Big|_{z_i(0)=0, 1\leq i\leq n} = \inf\left(\sum_{i=1}^n J_i\right)\Big|_{z_i(0)=0, 1\leq i\leq n}$$

to get $\bar{v}_n(0) = (v(0)/n)$, and set $\bar{v}(0) = \lim_{n\to\infty}\bar{v}_n(0)$. Here $\bar{v}(0)$ may be interpreted as the optimal cost incurred per agent with identically 0 initial state. Straightforward calculation gives [14]

$$\bar{v}(0) = \frac{\gamma^2\eta^2}{\rho}\left[1 - \frac{\bar{b}^2(\gamma-1)^2}{\left(\frac{\rho}{2}+\sqrt{\bar{a}^2+(1-\gamma)^2\bar{b}^2}\right)^2}\right] + \frac{\sigma^2\left(\bar{a}+\sqrt{\bar{a}^2+\bar{b}^2}\right)}{\rho\bar{b}^2} \tag{6.1}$$

where $\bar{a} = a - (\rho/2)$ and $\bar{b} = (b/\sqrt{r})$.

For the LQG game in the large population limit, when each agent applies the optimal tracking based control law $u_i^0 = -(b/r)(\Pi_a z_i + s_i)$, let $v_i(0)$ be the resulting individual cost. Write $v_{ind}(0) = v_i(0)$ for any i since all agents are assumed to have 0 initial state.

With $s_i, z^* \in C_b[0,\infty)$ determined from Proposition 4.6, one can get a solution $q \in C_b[0,\infty)$ for (3.6) if and only if the initial condition is given by

$$q(0) = \left\{\left(\beta_2^2 - \bar{b}^2\right)\frac{1}{\rho} + \frac{2\bar{b}^2\gamma}{(\rho-\lambda_1)\beta_1}\left(\frac{\bar{b}^2}{\beta_1+\lambda_2} - \beta_2\right) + \frac{\gamma^2\bar{b}^4}{(\rho-2\lambda_1)\beta_1^2}\left[1 - \frac{\bar{b}^2}{(\beta_1+\lambda_2)^2}\right]\right\}s_\infty^2 + \frac{\Pi_a\sigma^2}{\rho} \tag{6.2}$$

where λ_1, λ_2 are given in Section IV-C and $s_\infty = \lim_{t\to\infty} s_i(t)$. And since $z_i(0) = 0$, it follows from Proposition 3.2 that $v_{ind}(0) = q(0)$.

We derive from (6.1) and (6.2) that

$$|\bar{v}(0) - v_{ind}(0)| = O(\gamma^2).$$

The *gap* between $\bar{v}(0)$ and $v_{ind}(0)$ is demonstrated in Fig. 1 where all related parameters are given in Example 4.1 but γ takes values in $[0, 0.6]$; we can show that (**H1**) holds for all $\gamma \in [0, 0.6]$. If each agent applies the global cost based optimal control all of them will have a lower cost. However this requires a strong coordination in the sense that each individual player should be restrained from taking advantage of the other agents' presumably fixed global cost based control strategies.

VII. Effect of Inaccurate Information

Within the state aggregation framework, in the case where an individual has incorrect information on the dynamics of the competing population, that individual will naturally optimize with respect to an incorrectly calculated mass trajectory. An issue of interest concerns the offset between the cost actually attained and the expected cost (as calculated by the deviant individual based upon the incorrect mass behavior). This is related to deviancy detection and also to robustness of our control design.

For simplicity, here we take an isolated agent, and consider the ideal case in which all other agents have precise population statistics, but only the ith agent has an *inaccurate* estimate of the density $p(a)$ associated with the distribution F. Let the error in the density be the function $\delta p : \mathbb{R} \to \mathbb{R}$ satisfying $\int_{\mathbb{R}} \delta p(a)da = 0$. Hence $[p + \delta p](a)$ is used in the control calculation of the

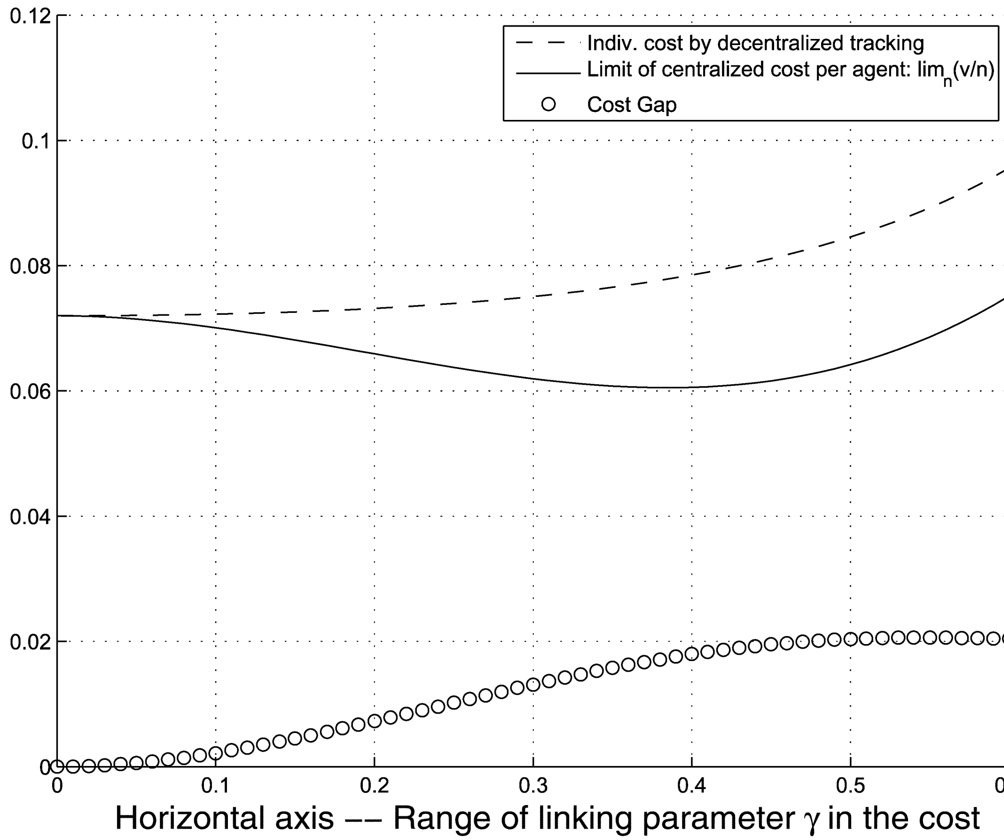

Fig. 1. (Top) Individual tracking based cost $v_{\rm ind}(0)$. (Middle) Scaled global cost $\overline{v}(0)$. (Bottom) The cost gap $|\overline{v}(0) - v_{ind}(0)|$.

ith agent; however the mass trajectory z^* computed by other agents is not affected by the perturbation δp. We consider the large population limit; then z^* is also the actually generated mass effect. In addition to **(H1)**, we consider small variations δp such that for a fixed small $\epsilon > 0$

$$(\mathbf{H1}') \quad \int_{\mathcal{A}} \frac{b^2\gamma[p+\delta p](a)}{r\beta_1(a)\beta_2(a)} da < 1-\epsilon.$$

(**H1'**) ensures well posedness of the aggregation procedure, giving an expected mass effect z_δ^*, for the ith agent. Also, we assume (**H2**) is known to the ith agent. By perturbation analysis for (4.11) and omitting higher-order error terms, we get the equation for the first-order variation δz^* with respect to z^*.

$$\begin{aligned}[\delta z^*](t) = &M \int_{\mathcal{A}}\int_0^t \int_s^\infty [\delta p](a) \\ &\cdot e^{-\beta_1(a)(t-s)} e^{-\beta_2(a)(\tau-s)} z^*(\tau) d\tau ds da \\ &+ M \int_{\mathcal{A}}\int_0^t \int_s^\infty p(a) e^{-\beta_1(a)(t-s)} \\ &\cdot e^{-\beta_2(a)(\tau-s)} [\delta z^*](\tau) d\tau ds da \end{aligned} \quad (7.1)$$

where $M = (\gamma b^2/r)$. The ith agent constructs its control law u_i^δ for optimally tracking $z_\delta^* \approx \widetilde{z}_\delta^* \triangleq z^* + \delta z^*$. Let $\widehat{u}_i$ be the optimal tracking control law with respect to z^* with cost $\widehat{v} = J_i(\widehat{u}_i, z^*)$. Following (3.1), we use $v_a = J_i(u_i^\delta, z^*)$ and $v_e = J_i(u_i^\delta, z_\delta^*)$ to denote the attained cost and expected cost, respectively.

Theorem 7.1: In addition to (**H1**)–(**H2**), assume δp satisfies (**H1'**) and denote $I(\delta p) = \int_{\mathcal{A}} (|[\delta p](a)|/(\beta_1(a)\beta_2(a))) da$. Then we have: (i) (7.1) has a unique solution in $C_b[0,\infty)$ with the bound estimate $\sup_{t\ge 0} |\delta z^*| = O(I(\delta p))$; (ii) the higher-order error estimate $\sup_{t\ge 0} |z_\delta^* - z^* - \delta z^*| = o(I(\delta p))$; and (iii) $0 \le v_a - \widehat{v} = O(I(\delta p))$, $|v_a - v_e| = O(I(\delta p))$.

Proof: We can show the first term at the RHS of (7.1) is bounded and continuous in t by expressing it as two integrals involving $p + \delta p$ and p and using (**H1**)–(**H1'**). Then, (i) follows by a fixed point method. For proving (ii), we write the integral equations corresponding to (4.11), satisfied by z_δ^* and z^*, respectively, and first show a uniform upper bound for $\sup_{t\ge 0} |z_\delta^*|$ and then $\sup_{t\ge 0} |z_\delta^* - z^*| = O(I(\delta p))$; finally we compare the two equations for $z_\delta^* - z^*$ and δz^*. For (iii), we may show the first part by checking the structural difference between $\widehat{u}_i$ and u_i^δ and using $\sup_{t\ge 0} |z_\delta^* - z^*| = O(I(\delta p))$. We may estimate $|v_a - v_e|$ by the change between the costs with respect to z^* and z_δ^*, respectively, when the same control law u_i^δ is used. □

VIII. Concluding Remarks

In a system of uniform agents, one can adopt the so-called direct approach by exploiting the coupled algebraic Riccati equations which arise in the theory of LQG dynamic games [3]. One may then find asymptotic estimates for all entries in the $n \times n$ Riccati solution matrix and hence obtain a limiting decentralized solution for the dynamic game [18]. Such asymptotic expansion based methods have been effective for weakly coupled game and team problems [30], [31]. Furthermore, it can be shown that the resulting solution is asymptotically consistent with the state aggregation based approach employed in this paper [18]. The fundamental reason that the direct approach works is that in estimating the magnitude of the elements of the large Riccati matrix only a few degrees-of-freedom are involved and hence, owing to the symmetry of the system, this number does not increase unboundedly (in fact, is constant) with respect to the number n of agents. However, this is not the case for a system with nonuniform agents due to the fact that the number of degrees-of-freedom increases combinatorially with n if the number of distinct values in $\{a_i, 1 \le i \le n\}$ increases proportionally to n. In particular, it may not be possible to get tight estimates when the a_i parameters in the system's set of coupled Riccati equations vary as a continuum.

Acknowledgment

The authors would like to thank a referee for suggesting the direct approach and its comparison with the individual-mass consistency based approach.

References

[1] A. C. Antoulas and D. C. Sorensen, Approximation of large-scale dynamical systems: An overview. Rice Univ., Houston, TX, Tech. Rep., 2001.

[2] T. Başar and Y.-C. Ho, "Informational properties of the Nash solutions of two stochastic nonzero-sum games," *J. Econ. Theory*, vol. 7, pp. 370–387, 1974.

[3] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, second ed. London, U.K: Academic, 1995.

[4] A. Bensoussan, *Stochastic Control of Partially Observable Systems*. Cambridge , U.K.: Cambridge Univ. Press, 1992.

[5] A. Bensoussan, M. Crouhy, and J. M. Proth, *Mathematical Theory of Production Planning*. Amsterdam, Netherlands: North-Holland, 1983.

[6] L. E. Blume, "Population games," in *The Economy as An Evolving Complex System II*, W. B. Arthur, S. Durlauf, and D. Lane, Eds. Reading, MA: Addison-Wesley, 1997, pp. 425–460.

[7] D. D. Šiljak, *Decentralized Control of Complex Systems*. Boston: Academic, 1991.

[8] Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, third ed. New York: Springer-Verlag, 1997.

[9] Z. Dziong and L. G. Mason, "Fair-efficient call admission for broadband networks—A game theoretic framework," *IEEE/ACM Trans. Netw.*, vol. 4, pp. 123–136, Feb. 1996.

[10] D. Famolari, N. B. Mandayam, D. Goodman, and V. Shah, "A new framework for power control in wireless data networks: Games, utility and pricing," in *Wireless Multimedia Network Technologies*, R. Ganesh, K. Pahlavan, and Z. Zvonar, Eds. Boston, MA: Kluwer, 1999, pp. 289–310.

[11] J. W. Friedman, *Game Theory with Applications to Economics*, Second ed. New York: Oxford Univ. Press, 1990.

[12] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*. Cambridge, MA: MIT Press, 1998.

[13] E. J. Green, "Continuum and finite-player noncooperative models of competition," *Econometrica*, vol. 52, no. 4, pp. 975–993, 1984.

[14] M. Huang, "Stochastic control for distributed systems with applications to wireless communications," Ph.D. dissertation, Dep. Elect. Comp. Eng., McGill Univ., Montreal, Canada, 2003.

[15] M. Huang, P. E. 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 Conf. Decision Contr.*, HI, Dec. 2003, pp. 98–103.

[16] M. Huang, P. E. Caines, and R. P. Malhamé, "Uplink power adjustment in wireless communication systems: A stochastic control analysis," *IEEE Trans. Autom. Control*, vol. 49, pp. 1693–1708, Oct. 2004.

[17] M. Huang, R. P. Malhamé, and P. E. Caines, "Nash equilibria for large-population linear stochastic systems of weakly coupled agents," in *Analysis, Control and Optimization of Complex Dynamic Systems*, E. K. Boukas and R. P. Malhamé, Eds. New York: Springer, 2005, pp. 215–252.

[18] M. Huang, R. P. Malhamé, and P. E. Caines, "Nash certainty equivalence in large population stochastic dynamic games: Connections with the physics of interacting particle systems," in *Proc. 45th IEEE Conf. Decision and Control*, San Diego, CA, Dec. 2006, pp. 4921–4926.

[19] K. L. Judd, "The law of large numbers with a continuum of i.i.d. random variables," *J. Econ. Theory*, vol. 35, pp. 19–35, 1985.

[20] M. Klein, "On production smoothing," *Manage. Sci.*, vol. 7, no. 3, pp. 286–293, 1961.

[21] V. E. Lambson, "Self-enforcing collusion in large dynamic markets," *J. Econ. Theory*, vol. 34, pp. 282–291, 1984.

[22] R. P. Malhamé and C.-Y. Chong, "Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system," *IEEE Trans. Autom. Control*, vol. 30, pp. 854–860, Sep. 1985.

[23] G. P. Papavassilopoulos, "On the linear-quadratic-Gaussian Nash game with one-step delay observation sharing pattern," *IEEE Trans. Autom. Control*, vol. 27, pp. 1065–1071, Oct. 1982.

[24] G. P. Papavassilopoulos, J. V. Medanic, and J. B. Cruz, Jr, "On the existence of Nash strategies and solutions to coupled Riccati equations in linear-quadratic Nash games," *J. Optim. Theory Appl.*, vol. 28, pp. 49–76, 1979.

[25] B. Petrovic and Z. Gajic, "The recursive solution of linear quadratic Nash games for weakly interconnected systems," *J. Optim. Theory Appl.*, vol. 56, pp. 463–477, Mar. 1988.

[26] R. G. Phillips and P. V. Kokotovic, "A singular perturbation approach to modelling and control of Markov chains," *IEEE Trans. Autom. Control*, vol. 26, pp. 1087–1094, Oct. 1981.

[27] R. S. Pindyck, "Optimal economic stabilization policies under decentralized control and conflicting objectives," *IEEE Trans. Autom. Control*, vol. 22, no. 4, pp. 517–530, Aug. 1977.

[28] S. P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*. Boston: Birkhäuser, 1994.

[29] G. Shen and P. E. Caines, "Hierarchically accelerated dynamic programming for finite-state machines," *IEEE Trans. Autom. Control*, vol. 47, pp. 271–283, Feb. 2002.

[30] R. Srikant and T. Başar, "Iterative computation of noncooperative equilibria in nonzero-sum differential games with weakly coupled players," *J. Optim. Theory Appl.*, vol. 71, pp. 137–168, Oct. 1991.

[31] R. Srikant and T. Başar, "Asymptotic solutions to weakly coupled stochastic teams with nonclassical information," *IEEE Trans. Autom. Control*, vol. 37, pp. 163–173, Feb. 1992.

[32] K. Yosida, *Functional Analysis*, sixth ed. New York: Springer-Verlag, 1980.

Minyi Huang (S'01–M'04) received the B.Sc. degree from Shandong University, Jinan, Shandong, China, in 1995, the M.Sc. degree from the Institute of Systems Science, Chinese Academy of Sciences, Beijing, China, in 1998, and the Ph.D. degree from the Department of Electrical and Computer Engineering, McGill University, Montreal, Canada, in 2003, all in the area of systems and control.

From February 2004 to March 2006, he was a Research Fellow with the Department of Electrical and Electronic Engineering, The University of Melbourne, Victoria, Australia. From April 2006 to June 2007, he was a Research Fellow with the Department of Information Engineering, Research School of Information Sciences and Engineering, The Australian National University, Canberra, Australia. He joined Carleton University, Ottawa, ON, Canada, in July 2007, where he is an Assistant Professor in the School of Mathematics and Statistics. His research interests include stochastic control and its application, game theory, coordination of multiagent systems, power control in wireless networks, and distributed signal processing in sensor networks.

Peter E. Caines (M'74–SM'83–F'86) received the B.A. degree in mathematics from Oxford University, U.K., in 1967, and the DIC and Ph.D. degrees in systems and control theory in 1970 from the University of London, Imperial College, U.K.

After periods as a postdoctoral researcher and faculty member with Stanford University, Stanford, CA, the University of California—Berkeley, the University of Toronto, Toronto, Canada, and Harvard University, Cambridge, MA, he joined McGill University, Montreal, Canada, in 1980, where he is James

McGill Professor and Macdonald Chair in the Department of Electrical and Computer Engineering. He is the author of *Linear Stochastic Systems* (New York: Wiley, 1988). His research interests lie in the areas of stochastic systems and hierarchical, hybrid, and discrete event systems.

Dr. Caines was elected to the Royal Society of Canada in May 2003.

Roland P. Malhamé (S'82–M'92) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the American University of Beirut, the University of Houston, Houston, TX, and the Georgia Institute of Technology, Atlanta, in 1976, 1978, and 1983, respectively.

After single-year stays with the University of Québec, Canada, and CAE Electronics Ltd. Montreal, Canada, he joined the École Polytechnique de Montréal in 1985, where he is Professor of Electrical Engineering. In 1994 and 2004, he spent sabbatical periods, respectively, with LSS, CNRS, and École Centrale de Paris, France. His current research interests are in stochastic control and the analysis and optimization of complex networks, in particular, manufacturing and communication networks. Past contributions include a statistical mechanics-based approach to electric load modeling. He is the director of Groupe d'Études et de Recherche en Analyse des Décisions.