Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems, Kyoto, Japan, July 24-28, 2006

Distributed Multi-Agent Decision-Making with Partial Observations: Asymptotic Nash Equilibria

Minyi Huang†, Peter E. Caines‡, and Roland P. Malhamé§

***Abstract*— We consider dynamic games in a large population of stochastic agents which are coupled by both individual dynamics and costs. These agents each have local noisy measurements of its own state. We investigate the synthesis of decentralized Nash strategies for the agents. The study for this class of large-scale systems provides interesting insights into competitive decision-making with localized information under large population conditions.**

I. Introduction and Motivation

The modeling and analysis of dynamic systems with many competing agents is of importance due to their wide appearance in socio-economic and engineering areas [11], [10], [12], [17], [1], as well as biological science [18], [20], and central issues concerning analysis and optimization of those systems include appropriate characterization of competition, temporal evolution of system behavior, information constraints, and implementation complexity of control strategies. Aiming at addressing these issues and developing a general optimization methodology, we study weakly coupled multi-agent decision-making with partial information. The underlying model and the methodology developed here will provide useful insights into understanding the behavior of systems in a wider scope with complex interactions between agents.

This kind of weak coupling in both dynamics and costs is used to model the mutual impact of agents during competitive decision-making. Specifically, cost coupling has been frequently encountered in economic theory where the agent's payoff is affected by the market condition, e.g., price, which in turn is affected by the aggregate population behavior such as the production level of all agents [17], [13]. In contrast, the dynamic coupling is used to specify an environment effect to the individual's decision-making generated by the population. For this kind of dynamic coupling involving the population effect, a simple illustrative example is the oligopoly product advertising model in which the given firm's sales rate is influenced by its own advertising expenditure (treated as a control variable) and the unsold market proportion (i.e., the market potential minus the total sales rate of all firms); for details see the bilinear Vidale-Wolfe oligopoly models examined in [8], [10]. In the engineering area, there is a similar phenomenon with Internet applications — as the aggregate consumption of all network users causes a higher congestion level, each individual user feels more difficult to acquire extra improvement in its service. In other words, the availability of network resource is responsive to the instantaneous activity of all users [5].

Although the models arising in these application areas take their different specific forms, the dynamic interaction between the individual and the mass (consisting of all others or the overall population) has a close resemblance to the generic linear model investigated in this paper. Specifically, as a key common feature to these different systems arising in economics, engineering or biology, while each agent only receives a negligible influence from any other given individual, the effect from the overall population is significant for each agent's strategy selection.

In this paper, we focus on the analysis for the linear models and develop a methodology for multi-agent competitive decision-making with local information and, in contrast to the extensive literature on linear-quadratic (LQ) or linear-quadratic-Gaussian (LQG) games (see, e.g., [22], [6], [24]), we are particularly interested in large populations. We note that games with a large or infinite population have long been a major research area in game theory [16], [3], [4], [21], [23], but traditionally most work has been based on static models. To simplify the analysis of our stochastic dynamic game, we consider a system of uniform agents, i.e., the agents are described by similar individual dynamics. Differing from our previous research [13], [14], [15], in the present system, each agent only has noisy observations of its own state. To obtain a localized control synthesis, it is critical to extract its state information from its available measurements, and it turns out this further depends on an appropriate anticipation of the collective effect of all agents.

A. The general approach and organization

In the decentralized game setting, we first introduce a local state estimation scheme for each agent, which is achieved by using a deterministic function $\bar{z}$ to approximate the mass effect, where the construction of $\bar{z}$ will form a key step in the analysis. Then by state aggregation

Work partially supported by Australian Research Council (ARC) and Natural Science and Engineering Research Council of Canada (NSERC).

†Department of Information Engineering, Research School of Information Sciences and Engineering, The Australian National University, Canberra 0200, Australia. Email: minyi.huang@rsise.anu.edu.au. This author's work was performed at Department of Electrical and Electronic Engineering, University of Melbourne, Victoria 3010, Australia.

‡Department of Electrical and Computer Engineering, McGill University, Montreal H3A 2A7, Canada. Email: peterc@cim.mcgill.ca.

§Department of Electrical Engineering, École Polytechnique de Montréal, Montreal H3C 3A7, Canada. Email: roland.malhame@polymtl.ca.

we obtain the control synthesis for the individual agents utilizing the filter output. The above localized filtering-control design relies on specifying a consistency relationship between the individual actions and the mass effect in a large population limit context. Subsequently, we examine the close-loop stable behavior of the population. Finally we present performance analysis by establishing an ε-Nash equilibrium property for the decentralized control laws of the agents.

II. The Weakly Coupled Systems

Consider an n dimensional linear stochastic system where each state component and its measurement are described by

$$dz_i = (az_i + bu_i)dt + \alpha z^{(n)}dt + \sigma dw_i, \qquad t \geq 0, \quad (1)$$
$$dy_i = cz_i dt + \bar{\sigma} dv_i, \qquad 1 \leq i \leq n, \qquad (2)$$

where $\{w_i, v_i, 1 \leq i \leq n\}$ denotes $2n$ independent standard scalar Wiener processes and $z^{(n)} = \frac{1}{n}\sum_{i=1}^{n} z_i$, $\alpha \in \mathbb{R}$. Hence, $z^{(n)}$ may be looked at as a nominal driving term imposed by the population. The Gaussian initial conditions $z_i(0)$ are mutually independent and are also independent of $\{w_i, v_i, 1 \leq i \leq n\}$. In addition, $b \neq 0$. Each state component shall be referred to as the state of the corresponding individual (also to be called an agent or a player). We note that the limiting version of equation (1) (i.e, as $n \to \infty$) may be viewed as a linear controlled version of the well-known McKean-Vlasov equation for weakly interacting diffusions [7], [19].

For simplicity of analysis, in this paper we consider a system of uniform agents in the sense that all agents share the same set of parameters (a, b, α, σ) and $(c, \bar{\sigma})$ as given in (1) and (2).

We investigate the behavior of the agents when they interact with each other through specific coupling terms appearing in their cost functions; this is displayed in the following set of *individual* cost functions which shall be used henceforth in the analysis:

$$J_i(u_i, \nu_i) \stackrel{\triangle}{=} E\int_0^\infty e^{-\rho t}[(z_i - \nu_i)^2 + ru_i^2]dt. \quad (3)$$

The objective of our work is to design the individual control strategies such that each agent's cost function is optimized in a certain sense utilizing only its local information, and we will cast the specific optimality criteria into the Nash equilibrium framework.

In particular, we assume the cost-coupling to be of the following form:

$$\nu_i = \Phi(z^{(n)}) = \Phi(\frac{1}{n}\sum_{k=1}^{n} z_k),$$

where Φ is a continuous function on $\mathbb{R}$. The linking term ν_i gives a measure of the average effect generated by the mass formed by all agents. Here we assume $\rho, r > 0$ and unless otherwise stated, throughout the paper z_i is described by the dynamics (1).

III. Competitive Decision-making with Local Information

Although the underlying system is linear, a straight application of Kalman filtering to the n dimensional system is out of question due to the information constraints for the agents. In other words, in our model there is not a central optimizer which can access all agents' outputs and then form the optimal estimate of the state vector. To overcome such a difficulty, we first formally approximate the term $z^{(n)}$ by a deterministic function $\bar{z}$ (to be determined later). Note that when a function $\bar{z}$, instead of $z^{(n)}$, appeared in the dynamics of z_i leading to uncoupled dynamics, the optimal state estimation for z_i would be given by the standard scalar Kalman filtering. Now in the large but finite population condition, it is expected that the Kalman filtering structure will still produce a satisfactory estimate when $z^{(n)}$ appears in the state equation (1) but is approximated by $\bar{z}$ when constructing the filtering equation.

It is evident that the term $z^{(n)}$ and hence $\bar{z}$ are related to the control laws of all agents. However, here we simply proceed by presuming $\bar{z}$ as a given function, and the exact procedure for determining this function will be clear after the control synthesis is described.

A. The auxiliary output regulation problem

As a preliminary step for the control design of the multi-agent system, we first introduce the following auxiliary Gaussian-Markov model

$$dz_i^0 = (az_i^0 + bu_i)dt + \alpha\bar{z}dt + \sigma dw_i, \qquad t \geq 0, \quad (4)$$
$$dy_i^0 = cz_i^0 dt + \bar{\sigma} dv_i, \qquad (5)$$

where $\bar{z} \in C_b[0, \infty)$ is given. We denote by $C_b[0, \infty)$ the set of deterministic, bounded and continuous functions on $[0, \infty)$. The noise terms have the same statistics as in the model (1)-(2). The cost function is

$$J^0(u_i) \stackrel{\triangle}{=} E\int_0^\infty e^{-\rho t}[(z_i^0 - z^*)^2 + ru_i^2]dt, \quad (6)$$

where $z^* \in C_b[0, \infty)$.

Let $\Pi > 0$ be the solution to the Riccati equation:

$$\rho\Pi = 2a\Pi - \frac{b^2}{r}\Pi^2 + 1. \quad (7)$$

And denote $\beta_1 = -a + \frac{b^2}{r}\Pi$, $\beta_2 = -a + \frac{b^2}{r}\Pi + \rho$. It is easy to check that $\beta_2 > \frac{\rho}{2}$. For Kalman filtering, we write the Riccati equation:

$$\frac{dP(t)}{dt} = 2aP(t) - c^2R^{-1}P^2(t) + Q, \qquad t \geq 0, \quad (8)$$

where $R = \bar{\sigma}^2$ and $Q = \sigma^2$. The initial condition for $P(t)$ is taken as the variance of $z_i^0(0)$. Let s be a bounded solution for the differential equation

$$\rho s = \frac{ds}{dt} + as - \frac{b^2}{r}\Pi s + \alpha\Pi\bar{z} - z^*. \quad (9)$$

Remark: It can be shown that there exists a unique initial condition $s(0)$ leading to a bounded solution s and that any other initial condition gives an unbounded solution [14]. $\square$

In fact, we have the expression [14]

$$s(t) = e^{\beta_2 t}\int_t^{\infty} e^{-\beta_2 \tau}[\alpha\Pi\bar{z}(\tau) - z^*(\tau)]d\tau \in C_b[0,\infty). \tag{10}$$

Then it is easy to check that the optimal control law is given by

$$d\hat{z}_i^0 = (a\hat{z}_i^0 + bu_i + \alpha\bar{z})dt + P(t)cR^{-1}[dy_i^0 - c\hat{z}_i^0 dt],$$
$$u_i = -\frac{b}{r}\left[\Pi\hat{z}_i^0 + s(t)\right].$$

B. The local approximate Kalman filtering

In the multi-agent system, since each agent's information is restricted to its own measurement, one cannot directly use the standard Kalman filtering. However, after introducing a structural approximation for the mass effect, the state of each agent can be estimated by use of only its local information, and the associated approximate filtering equation may be constructed by the usual Kalman filtering for a scalar model. The justification of such an approximation will be given during the closed-loop stability analysis.

For agent i, the initial condition for equation (8) is $\kappa_i \triangleq Var(z_i(0))$ and we denote the corresponding solution by $P_i(t)$. To simplify our analysis below, we first assume all κ_i are equal to the same value $\kappa > 0$, and hence the same function $P(t)$ is used for all agents.

The local filter is constructed as follows:

$$d\hat{z}_i = (a\hat{z}_i + bu_i + \alpha\bar{z})dt + P(t)cR^{-1}[dy_i - c\hat{z}_i dt], \tag{11}$$

where $\bar{z}$, used for approximating $z^{(n)}$, is to be determined. We shall call this the approximate Kalman filter since an approximation step involving $\bar{z}$ is introduced here. Note that since z_i is driven by $z^{(n)}$ in its actual model, the correcting term $dy_i - c\hat{z}_i dt$ is not an innovation process (i.e., a Wiener process). Denote the error term

$$\tilde{z}_i = z_i - \hat{z}_i.$$

The approximate Kalman filtering equation may be written in the form

$$d\hat{z}_i = (a\hat{z}_i + bu_i + \alpha\bar{z})dt + P(t)c^2R^{-1}\tilde{z}_i dt + P(t)cR^{-1}\bar{\sigma}dv_i.$$

Combining (1) and (11) gives the error equation

$$d\tilde{z}_i = \left[a - P(t)c^2R^{-1}\right]\tilde{z}_i dt + \alpha[z^{(n)} - \bar{z}]dt + \sigma dw_i - P(t)cR^{-1}\bar{\sigma}dv_i.$$

C. Feedback and closed-loop dynamics

By use of the single agent based control law in Section III-A, we proceed to formally construct the individual control laws for n agents as follows:

$$u_i^0 = -\frac{b}{r}\left[\Pi\hat{z}_i + s(t)\right], \qquad 1 \le i \le n, \tag{12}$$

which henceforth will be adopted by the individual agents with $\hat{z}_i$ determined by (11). Here the function s in (12) is to be constructed using a large population limit; and the remaining critical issue is to first determine a tracking reference trajectory z^* and then s.

To determine $s(t)$ in (12), we use the state aggregation technique within a population limit as in [13], [14], and introduce the state aggregation equation system

$$\rho s = \frac{ds}{dt} + as - \frac{b^2}{r}\Pi s + \alpha\Pi\bar{z} - z^*, \tag{13}$$
$$\frac{d\bar{z}}{dt} = a\bar{z} - \frac{b^2}{r}\left[\Pi\bar{z} + s(t)\right] + \alpha\bar{z}, \tag{14}$$
$$z^* = \Phi(\bar{z}), \tag{15}$$

where $s \in C_b[0,\infty)$. For given $\bar{z}$ and z^*, whose boundedness will be established later, there is a unique initial condition $s(0)$ yielding a bounded solution s, and hence it is unnecessary to specify $s(0)$ separately. Here (13) results form the single agent based optimal tracking once the population effect is approximated by the function $\bar{z}$. Equation (14) is based on taking expectation in the closed-loop of (1) with control law u_i^0, where we approximate $z^{(n)}$ by $\bar{z}$ and the expectation $E\hat{z}_i$ is also approximated by the same $\bar{z}$ under the condition of uniform agents.

The underlying mechanism for devising the control strategy (12) is that, in the large population limit and for *a given mass effect* $\bar{z} \in C_b[0,\infty)$, each individual will tend to take an optimal tracking action, and in turn, these individual actions will collectively generate *the same mass effect* $\bar{z}$ as described by (14) corresponding to the optimal tracking based control law. This constitutes the so-called *mutual consistency relationship* between the individual and the mass, and this notion is captured mathematically by a fixed point theorem (to be discussed below) which provides the existence of such a trajectory $\bar{z} \in C_b[0,\infty)$ and characterizes the resulting pair $(\bar{z}, s)$ as the unique bounded solution to (13)-(15).

Now we have the closed-loop equations for $\hat{z}_i$ and $\tilde{z}_i$ after the control law u_i^0 is implemented:

$$d\hat{z}_i = \left(a - \frac{b^2}{r}\Pi\right)\hat{z}_i dt + P(t)c^2R^{-1}\tilde{z}_i dt + \alpha\bar{z}dt - \frac{b^2}{r}s(t)dt + P(t)cR^{-1}\bar{\sigma}dv_i, \tag{16}$$

$$d\tilde{z}_i = \left[a - P(t)c^2R^{-1}\right]\tilde{z}_i dt + \frac{\alpha}{n}\sum_{j=1}^n \hat{z}_j dt + \frac{\alpha}{n}\sum_{j=1}^n \tilde{z}_j dt - \alpha\bar{z}dt + \sigma dw_i - P(t)cR^{-1}\bar{\sigma}dv_i. \tag{17}$$

Letting $\hat{z}_n^\dagger = \frac{1}{n}\sum_{j=1}^n \hat{z}_j$ and $\tilde{z}_n^\dagger = \frac{1}{n}\sum_{j=1}^n \tilde{z}_j$, we obtain

$$d\hat{z}_n^\dagger = \left(a - \frac{b^2}{r}\Pi\right)\hat{z}_n^\dagger dt + P(t)c^2R^{-1}\tilde{z}_n^\dagger dt - \frac{b^2}{r}s(t)dt + \frac{1}{\sqrt{n}}P(t)cR^{-1}\bar{\sigma}dv^\dagger, \tag{18}$$

$$d\tilde{z}_n^\dagger = \alpha\hat{z}_n^\dagger dt + \left[a - P(t)c^2R^{-1} + \alpha\right]\tilde{z}_n^\dagger dt - \alpha\bar{z}dt + \frac{1}{\sqrt{n}}\sigma dw^\dagger - \frac{1}{\sqrt{n}}P(t)cR^{-1}\bar{\sigma}dv^\dagger, \tag{19}$$

where $w^\dagger = n^{-1/2}\sum_{i=1}^n w_i$ and $v^\dagger = n^{-1/2}\sum_{i=1}^n v_i$ are two independent standard Wiener processes.

Remark. For specifying the mean trajectories of $\hat{z}_n^\dagger$ and $\tilde{z}_n^\dagger$, it suffices to remove the noise terms at the right hand side of (18)-(19). □

IV. STABLE POPULATION BEHAVIOR

We introduce the following assumptions:

($\mathbb{H}1$) Let $P^\dagger > 0$ denote the steady state value of the solution to the filtering Riccati equation (8). Assume $a - \frac{b^2}{r}\Pi < 0$ and the matrix

$$M = \begin{bmatrix} a - \frac{b^2}{r}\Pi & P^\dagger c^2 R^{-1} \\ \alpha & a - P^\dagger c^2 R^{-1} + \alpha \end{bmatrix} \qquad (20)$$

is strictly stable. □

($\mathbb{H}2$) The function Φ is Lipschitz continuous on $\mathbb{R}$ with a Lipschitz constant $\gamma > 0$, i.e., $|\Phi(y_1) - \Phi(y_2)| \le \gamma|y_1 - y_2|$ for all $y_1, y_2 \in \mathbb{R}$. □

($\mathbb{H}3$) $\beta_1 > 0$, and $\frac{|\alpha|}{\beta_1} + \frac{b^2(\gamma+|\alpha|\Pi)}{r\beta_1\beta_2} < 1$, where β_1, β_2 are computed from the Riccati equation (7). The constant γ is specified in ($\mathbb{H}2$). □

($\mathbb{H}4$) All agents have mutually independent Gaussian initial conditions of zero mean, i.e. $Ez_i(0) = 0$. In addition, all $Ez_i^2(0) = \kappa > 0$, $i \ge 1$. □

Remark: It is easy to verify that $a - P^\dagger c^2 R^{-1} < 0$. Hence for sufficiently small $|\alpha|$, the matrix M in ($\mathbb{H}1$) is always stable provided that $a - \frac{b^2}{r}\Pi < 0$ holds. □

Before analyzing it solution, it is of interest to note that the state aggregation equation system (13)-(15) is not affected by the partial observation situation; in other words, the case of full information will still yield the same set of equations; see [14], [15] for details. Then using the method in [14], [15] to the current special case of uniform agents, we may eliminate s in (14) and derive a fixed point equation for $\bar{z}$, for which we can establish the existence and uniqueness of a bounded solution under $(\mathbb{H}2) - (\mathbb{H}4)$. Accordingly, a unique bounded solution to (13)-(15) can be obtained. The following theorem is a direct consequence of the existence theorem in [14] for a system of non-uniform agents and perfect observations, which itself is proven by a fixed point argument.

Theorem 1: [14] Under ($\mathbb{H}2$)-($\mathbb{H}4$), the state aggregation equation system (13)-(15) admits a unique bounded solution $\bar{z} \in C_b[0,\infty)$ and $s \in C_b[0,\infty)$. □

Theorem 2: Under ($\mathbb{H}2$)-($\mathbb{H}4$) and the control strategies u_i^0 for all agents, the closed-loop system for the n agents admits a unique strong solution.

Proof. This follows by verifying the Lipschitz condition for the closed-loop system when all n agents apply the set of control laws $(u_1^0, \cdots, u_n^0)$. □

Theorem 3: Under ($\mathbb{H}1$)-($\mathbb{H}4$), there exists a constat C such that

$$\sup_{t\ge 0, 1\le i\le n} E\left[\hat{z}_i^2(t) + \tilde{z}_i^2(t)\right] \le C \qquad (21)$$

where C is independent of the population size n.

Proof. We begin by constructing the deterministic time-varying ODE system,

$$\frac{dx}{dt} = \begin{bmatrix} a - \frac{b^2}{r}\Pi & P(t)c^2R^{-1} \\ \alpha & a - P(t)c^2R^{-1} + \alpha \end{bmatrix} x. \qquad (22)$$

We denote the fundamental solution matrix for the ODE (22) by $\Phi(t, t_0)$ where $t \ge t_0 \ge 0$ and $\Phi(t_0, t_0) = I$. It can be checked that $\Phi(t, t_0)$ is exponentially stable under ($\mathbb{H}1$). By use of this fact, we establish the L_2 stability of $\hat{z}_n^\dagger$ and $\tilde{z}_n^\dagger$, i.e.,

$$\sup_{t\ge 0, 1\le i\le n} E\left(|\hat{z}_n^\dagger(t)|^2 + |\tilde{z}_n^\dagger(t)|^2\right) \le C_1, \qquad (23)$$

which combined with (17) gives

$$\sup_{t\ge 0, 1\le i\le n} E\tilde{z}_i^2(t) \le C_2. \qquad (24)$$

The constants C_1 and C_2 do not depend on n.

By virtual of (24) and (16) we further get

$$\sup_{t\ge 0, 1\le i\le n} E\hat{z}_i^2(t) \le C, \qquad (25)$$

for some constant $C > 0$ independent of n. This completes the proof. □

V. THE ASYMPTOTIC EQUILIBRIUM ANALYSIS

For the population of n agents, the agents' admissible control set $\mathcal{U}_{1,\cdots,n}$ consists of all feedback controls $(u_1, \cdots, u_n)$ adapted to the σ-algebra $\sigma(y_i(\tau), \tau \le t, 1 \le i \le n)$ (i.e., each $u_k(t)$ is a functional of $(t, y_1(\tau), \cdots, y_n(\tau)), \tau \le t$) such that a unique strong solution to the closed-loop system of the n agents exists on $[0, \infty)$. Here we only have a very general requirement for the control such that it depends on the measurements $(y_i, 1 \le i \le n)$ and is allowed to depend on all available past history as long as a solution is well defined. With the coupling in dynamics, each agent's admissible control set may be affected by the strategies taken by other agents. This is very similar to the social equilibrium scenario [2] by imposing additional constraints on individuals' choices of strategies; such a notion dates back to the early work [9] and is widely used for the analysis of Nash equilibrium in the economics literature. In our setting, the implicit constraints (in the sense that the n agents' joint strategy space $\mathcal{U}_{1,\cdots,n}$ does not decompose into the Cartesian product of n individual strategy sets) on each agent's control serve only to ensure the existence of a well-defined solution for the closed-loop system. We use u_{-k} to denote the vector of individual strategies obtained by deleting u_k in $(u_1, \cdots, u_n)$. Then the vector $(u_1, \cdots, u_k, \cdots, u_n)$ may be equivalently denoted as (u_k, u_{-k}). For a fixed u_{-k}, we induce the projection of $\mathcal{U}_{1,\cdots,n}$ to its kth component as the set $\mathcal{U}_k|_{u_{-k}} = \{u_k | (u_k, u_{-k}) \in \mathcal{U}_{1,\cdots,n}\}$. Note that $\mathcal{U}_k|_{u_{-k}}$ is not restricted to be decentralized since u_k is allowed to depend on y_i, $i \ne k$, which actually leads to a stronger characterization for the decentralized control law analyzed in this section. In this setup we give the definition.

Definition 4: A set of controls $(u_k)_{k=1}^n \in \mathcal{U}_{1,\cdots,n}$ for n players is called an *ε-Nash equilibrium* with respect to the costs $J_k, 1 \le k \le n$, if there exists $\varepsilon \ge 0$ such that for any fixed $1 \le i \le n$, we have

$$J_i(u_i, u_{-i}) \le J_i(u_i', u_{-i}) + \varepsilon, \tag{26}$$

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

Note that the costs J_k, $1 \le k \le n$, appearing in Definition 4 are deterministic quantities in the functional form of a set of individual feedback control laws depending on the system outputs $(y_1, \cdots, y_n)$.

For obtaining a desired performance estimate, the following lemma is useful. Notice that for independent random variables, it is usually easy to derive magnitude estimate of this type. When there exists dependence, the estimate is less obvious. In the following we will get the estimate by use of the stability property for the closed-loop system.

Lemma 5: Assume ($\mathbb{H}$1)-($\mathbb{H}$4) hold. Under the optimal tracking based control laws u_i^0, $1 \le i \le n$, we have

$$\sup_{t\ge 0} E\left|n^{-1}\sum_{i=1}^n(\hat{z}_i - E\hat{z}_i)\right|^2 + \sup_{t\ge 0} E\left|n^{-1}\sum_{i=1}^n(\tilde{z}_i - E\tilde{z}_i)\right|^2 = O(n^{-1}),$$

which further implies $\sup_{t\ge 0} E\left[n^{-1}\sum_{i=1}^n(z_i - Ez_i)\right]^2 = O\left(n^{-1}\right)$.

Proof. From (18) and (19), we get

$$\begin{aligned} d(\hat{z}_n^\dagger - E\hat{z}_n^\dagger) = &\left(a - \frac{b^2}{r}\Pi\right)(\hat{z}_n^\dagger - E\hat{z}_n^\dagger)dt \\ &+ P(t)c^2R^{-1}(\tilde{z}_n^\dagger - E\tilde{z}_n^\dagger)dt \\ &+ \frac{1}{\sqrt{n}}P(t)cR^{-1}\bar{\sigma}dv^\dagger, \\ d(\tilde{z}_n^\dagger - E\tilde{z}_n^\dagger) = &\ \alpha(\hat{z}_n^\dagger - E\hat{z}_n^\dagger)dt \\ &+ \left[a - P(t)c^2R^{-1} + \alpha\right](\tilde{z}_n^\dagger - E\tilde{z}_n^\dagger)dt \\ &+ \frac{1}{\sqrt{n}}\sigma dw^\dagger - \frac{1}{\sqrt{n}}P(t)cR^{-1}\bar{\sigma}dv^\dagger. \end{aligned}$$

As in proving Theorem 3, let $\Phi(t,\tau)$ denote the fundamental solution matrix to the ODE (22), which can be shown to be uniformly exponentially stable, i.e, there exists a constant $C > 0$ and $0 < \theta < 1$ such that for $0 \le \tau \le t$,

$$\|\Phi(t,\tau)\| \le C\theta^{t-\tau}. \tag{27}$$

We write $\Phi(t,\tau)$ as a 2×2 matrix to get

$$\Phi(t,\tau) = (\Phi_{ij}(t,\tau))_{i,j=1}^2.$$

Then it follows that

$$\begin{aligned} (\hat{z}_n^\dagger - E\hat{z}_n^\dagger)(t) = &\Phi_{11}(t,0)\hat{z}_n^\dagger(0) + \Phi_{12}(t,0)\tilde{z}_n^\dagger(0) \\ &+ \frac{1}{\sqrt{n}}\int_0^t [\Phi_{11}(t,\tau) - \Phi_{12}(t,\tau)]P(\tau)cR^{-1}\bar{\sigma}dv_\tau^\dagger \\ &+ \frac{1}{\sqrt{n}}\int_0^t \Phi_{12}(t,\tau)\sigma dw_\tau^\dagger \end{aligned}$$

where $\hat{z}_n^\dagger(0) = 0$ and $\tilde{z}_n^\dagger(0) = \frac{1}{n}\sum_{i=1}^n z_i(0)$. Hence this gives

$$\begin{aligned} &E|\hat{z}_n^\dagger(t) - E\hat{z}_n^\dagger(t)|^2 \\ &= O(\frac{1}{n}) + \\ &O(\frac{1}{n}\int_0^t \left|[\Phi_{11}(t,\tau) - \Phi_{12}(t,\tau)]P(\tau)cR^{-1}\bar{\sigma}\right|^2 d\tau) \\ &\quad + O(\frac{1}{n}\int_0^t |\Phi_{12}(t,\tau)\sigma|^2 d\tau). \end{aligned}$$

And it follows that

$$\sup_{t\ge 0} E|\hat{z}_n^\dagger(t) - E\hat{z}_n^\dagger(t)|^2 = O(\frac{1}{n}). \tag{28}$$

Similarly, we obtain

$$\sup_{t\ge 0} E|\tilde{z}_n^\dagger(t) - E\tilde{z}_n^\dagger(t)|^2 = O(\frac{1}{n}), \tag{29}$$

and the Lemma follows readily from (28) and (29). □

For establishing the main result given by Theorem 7, the following Lemma 6 is instrumental and quantifies the impact generated by a given agent's strategy deviation provided that a prior bound is known for its performance when other agents stick to their optimal tracking based strategies. Define

$$I(u_i, u_{-i}^0) = E\int_0^\infty e^{-\rho t}\left[z^* - \Phi(\frac{1}{n}\sum_{k=1}^n z_k)\right]^2_{(u_i, u_{-i}^0)} dt,$$

$$I(u_i^0, u_{-i}^0) = E\int_0^\infty e^{-\rho t}\left[z^* - \Phi(\frac{1}{n}\sum_{k=1}^n z_k)\right]^2_{(u_i^0, u_{-i}^0)} dt,$$

where the state trajectories inside the integral are generated by the associated control strategies, and z^*, as the tracking reference trajectory, is determined by the state aggregation equation system (13)-(15).

Lemma 6: Assume ($\mathbb{H}$1)-($\mathbb{H}$4) for the system of n agents. Let i be fixed and $u_i \in \mathcal{U}_{i|u_{-i}^0}$ be a control such that $J_i(u_i, u_{-i}^0) \le C$, for a fixed $C > 0$ independent of n, where all other agents take the controls u_j^0, $j \ne i$. Then there exists C_1 independent of n such that

$$\int_0^\infty e^{-\rho t}E(z_i^2 + u_i^2)dt \le C_1.$$

And moreover, for the above fixed i,

$$|I(u_i, u_{-i}^0) - I(u_i^0, u_{-i}^0)| = O(\frac{1}{n}). \tag{30}$$

□

By making use of Theorem 3, we can find a constant C independent of n such that $J_i(u_i^0, u_{-i}^0) \le C$, which hence ensures that the constant $C > 0$ used in Lemma 6 exists. The lemma can be proven by use of the closed-loop system of the n agents, and the main step is to show $\int_0^\infty e^{-\rho t}Ez_i^2|_{(u_i, u_{-i}^0)}dt < \infty$ when $J_i(u_i, u_{-i}^0) \le C$ holds.

Now we are in a position to state the central result of this paper. Let the individual control law u_i^0 be given by (12).

Theorem 7: Suppose ($\mathbb{H}1$)-($\mathbb{H}4$) hold for the system of n agents. Then the set of feedback strategies given by $(u_1^0, \cdots, u_n^0)$ is an ε-Nash equilibrium, where $\varepsilon \to 0$ as $n \to \infty$.

Proof. Based on Lemmas 5 and 6, this theorem can be proven by a similar method as in [12], [13], [14]. The details are omitted here. □

Recall that when specifying the ε-Nash equilibrium in Definition 4, the admissible control set for the n agents depends on centralized information. In conjunction with this definition, the implication of Theorem 7 is that, in a large but finite population, the decentralized observation dependent control is almost as good as global history information based controls; any deviating agent i cannot take essential advantage of others by utilizing full information of the system outputs, provided that all other agents stick to their optimal (output) tracking based control laws u_{-i}^0.

VI. Conclusion

In this paper we investigate distributed decision-making in a system of uniform agents coupled by their linear dynamics and individual costs, where each agent has noisy measurements of its own state. We propose a decentralized control synthesis in which each agent utilizes its local information for its control strategies. The key steps in the control design consist of local approximate Kalman filtering based on the anticipation of the collective effect of all agents, and then of an optimal tracking based control law. It is shown that the resulting set of individual control laws leads to stable population behavior and has an asymptotic Nash equilibrium property.

Acknowledgement

The authors would like to thank William McEneaney for inspiring discussions.

References

[1] E. Altman, T. Basar, and R. Srikant. Nash equilibria for combined flow control and routing in networks: asymptotic behavior for a large number of users. *IEEE Trans. Automat. Contr.*, vol. 47, pp. 917-930, June 2002.

[2] J.-P. Aubin. *Optima and Equilibria: An Introduction to Nonlinear Analysis*, 2nd ed. Springer, Berlin, 1998.

[3] R. J. Aumann. Values of markets with a continuum of traders. *Econometrica*, vol. 43, pp. 611-646. 1975.

[4] R. J. Aumann and L. S. Shapley. *Values of Non-Atomic Games*. Princeton Univ. Press, Princeton, NJ, 1974.

[5] F. Baccelli, D. Hong, and Z. Liu. Fixed point methods for the simulation of the sharing of a local loop by a large number of interacting TCP connections. INRIA Tech. Report No. 4154, France, 2001.

[6] T. Basar and G. J. Olsder. *Dynamic Noncooperative Game Theory*, 2nd ed., Academic Press, London, UK, 1995.

[7] D. A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffussions. *Stochastics*, vol. 20, pp. 247-308, 1987.

[8] K. R. Deal. Optimizing advertising expenditures in a dynamic duopoly. *Operations Research*, vol. 27, no. 4, pp. 682-692, 1979.

[9] G. Debreu. A social equilibrium existence theorem. *Proc. Nat. Acad. Sci. USA* , vol. 38, pp. 886-893, 1952.

[10] G. M. Erickson. Differential game models of advertsing competition. *Europ. J. Oper. Res.*, vol. 83, pp. 431-438, 1995.

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

[12] 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. *Proc. 42nd IEEE Conf. Decision Contr.*, Maui, Hawaii, pp. 98-103, Dec. 2003.

[13] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems: generalizations to non-uniform individuals. *Proc. the 43rd IEEE Conference on Decision and Control*, Paradise Island, Bahamas, pp. 3453-3458, December 2004.

[14] 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., Chapter 9, pp. 215-252, Springer, New York, 2005.

[15] M. Huang, R. P. Malhamé, and P.E. Caines. Nash strategies and adaptation for decentralized games involving weakly-coupled agents. *Proc. 44th IEEE CDC-ECC'05 Conference*, Seville, Spain, pp. 1050-1055, Dec. 2005.

[16] M. Ali Khan and Y. Sun, Non-cooperative games with many players. In *Handbook of Game Theory with Economic Applications*, vol. 3, R. J. Aumann and S. Hart eds., North-Holland, 2002.

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

[18] J. Maynard Smith. *Evolution and The Theory of Games*, Cambridge University Press, Cambridge, UK, 1982.

[19] H. P. McKean Jr. A calss of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci., USA*, vol. 56, pp. 1907-1911, 1966.

[20] J. M. McNamara, A. I. Houston, and E. J. Collins. Optimality models in behavioral biology. *SIAM Review*, vol. 43, no. 3, pp. 413-466, 2001.

[21] J. W. Milnor and L. S. Shapley. Values of large games II: Oceanic games. *Math. Oper. Res.*, vol. 3, pp. 290-307, 1978.

[22] G. P. Papavassilopoulos. On the linear-quadratic-Gaussian Nash game with one-step delay observation sharing pattern. *IEEE Trans. Automat. Control*, vol. 27, no. 5, pp. 1065-1071, 1982.

[23] N. Z. Shapiro and L. S. Shapley. Values of large games I: A limit theorem. *Math. Oper. Res.*, vol. 3, pp. 1-9, 1978.

[24] A. J. T. M. Weeren, J. M. Schumacher, and J. C. Engwerda. Asymptotic analysis of linear feedback Nash equilibria in nonzero-sum linear-quadratic differential games. *J. Optim. Theory Appl.*, vol. 101, no. 3, pp. 693-722, Jun. 1999.