Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

FrA09.4

Nash Certainty Equivalence in Large Population Stochastic Dynamic Games: Connections with the Physics of Interacting Particle Systems

Minyi Huang   Roland P. Malhamé   Peter E. Caines

Abstract— We consider large population dynamic games and illuminate methodological connections with the theory of interacting particle systems. Combined with the large population modelling, a Nash Certainty Equivalence (NCE) Methodology is introduced for specifying the localized strategy selection of a given agent within the Nash equilibrium setting. The NCE methodology closely parallels that found in the study of uncontrolled interacting particle systems within the framework of the McKean-Vlasov equation [19]: for both problems the solution is derived by focussing on a single generic individual at a microscopic level and analyzing its interaction with the ensemble of the other individuals of which it is itself, in a statistical sense, a representative.

I. INTRODUCTION

In the literature, there has existed little connection between the physics of interacting particle (IP) systems and stochastic dynamic game theory, and they are generally regarded as independent areas. IP systems have long been a major research focus in statistical physics with the objective of understanding complex phenomena produced by many randomly moving particles [7], [3], [18], and in a mathematical context they have also provided a rich source for stochastic analysis [9], [8], [17], [5], [6].

In contrast, stochastic dynamic game theory is concerned with a set of agents competing for individual interests in a random environment. The modelling and analysis of dynamical systems with many competing agents is of importance due to their wide appearance in socio-economic and engineering areas [12], [11], [13], [20], [1], [2], as well as biological science [21], [22]. We are particularly interested in a large population of weakly coupled agents. The weak coupling in both dynamics and costs is used to model the interaction of agents during competitive decision-making. Specifically, cost coupling is frequently encountered in economic theory [20], [14], and also in wireless network resource optimization [13]. On the other hand, dynamic coupling is used to specify an environment effect generated by the population and captures the indirect impact of other agents' activity on a specific individual, and an illustrative example is the oligopoly product advertising competition model; see, e.g., [10], [11]. A brief literature review for dynamic games with weak coupling – as described by a small parameter – can be found in [14], [15], and those existing work usually dealt with approximate numerical solutions of Nash strategies for a few players [23].

In this work, we discuss the connection between existing research on IP systems and a new optimization approach which has been developed for large population stochastic dynamic games in our recent work [13], [14], [15]. This also naturally suggests we look into methodological implications for large population dynamic games based on existing methods in physics. In [13], [14], [15], the optimization paradigm is based on the individual-mass interaction and the subsequently developed state aggregation which extracts the mass effect a given agent receives. This leads to a localized control synthesis generating decentralized ε-Nash strategies.

In order to have a streamlined presentation, we will mostly work on the game problem and discuss, when appropriate, its relation with the IP system and the general methodology developed therein. The organization of the paper is as follows. We first introduce in Section II the IP system with both finite and infinite particles, the n agent stochastic dynamic game and its large population limit form. In Section III we analyze the solution for the infinite population game where the individual strategy selection is based on the Nash Certainty Equivalence (NCE) Methodology for the search of a fixed point mass effect measure, and we further illustrate the connection of this approach with the IP system modelling via the notion of consistent pair which has played a fundamental role in the theory of infinite particle systems. In this setting, we designate the NCE Principle as the property that the resulting scheme is consistent in the sense that the prescribed control laws produce sample paths which produce the mass effect with respect to which the individual optimal control is derived. In Section IV we analyze the asymptotic Nash equilibrium property for the decentralized strategy obtained in Section III. Section V considers the special case of linear cost coupling which leads to an explicit solution to the NCE based approach. In addition, we examine the asymptotic behavior of the full information based feedback Nash equilibrium strategies, establish a consistency relationship with the NCE based approach, and discuss the limitation of the full state based scheme.

M. Huang is with Department of Information Engineering, Research School of Information Sciences and Engineering, The Australian National University, Canberra 0200 ACT, Australia. Email: `minyi.huang@rsise.anu.edu.au`

R. P. Malhamé is with Department of Electrical Engineering, École Polytechnique de Montréal, Montreal H3C 3A7, QC, Canada. Email: `roland.malhame@polymtl.ca`

P. E. Caines is with Department of Electrical and Computer Engineering, McGill University, Montreal H3A 2A7, QC, Canada. Email: `peterc@cim.mcgill.ca`

II. IP Systems and Dynamic Games with Weakly Coupled Agents

In this section we formalize three different models, i.e., the IP systems, the finite and infinite population dynamic games. For the two game problems, we only make brief descriptions here and leave further discussion in the subsequent sections.

A. The interacting particle system

In an IP system, the state evolution of an individual particle is affected by an empirical average of coupling terms with all other particles which are otherwise indistinguishable. In other words, when viewed from a given particle, only the empirical distribution information for other particles matters. Mathematically, this leads to a set of weakly coupled diffusions, each describing the motion of a single particle, where an averaging across the population produces the coupling term in the individual dynamics. Within this modelling framework, a single particle may be looked at as a generic representative of all members in the model, and the system property is largely captured by the interaction of an individual and the mass of all others.

We introduce the following dynamics [9] in the form of N coupled stochastic differential equations (SDE):

$$dx_i = \frac{1}{N}\sum_{k=1}^{N} b(x_i, x_k)dt + \sigma dw_i, \quad 1 \le i \le N, \quad (1)$$

where $b(\cdot, \cdot)$ is a function from $\mathbb{R}^2$ to $\mathbb{R}$, N is the number of particles and all x_i's are assumed to have i.i.d. initial conditions at $t = 0$. Here we assume x_i is a scalar although the modelling is also applicable to the case of vector particle states. The noises $\{w_i, 1 \le i \le N\}$ are N independent Wiener processes independent of the initial conditions $x_i(0)$, $1 \le i \le N$. Let $\varepsilon_x = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_k}$ denote the empirical measure of the particle configuration $(x_1, x_2, \cdots, x_N)$ where $\delta_\bullet$ is the Dirac measure. Then the drift term in (1) may be expressed as a function of x_i and ε_x.

For this class of particle models, one can achieve a remarkable degree of economy in the description of population dynamics, by expressing the aggregate coupling term in terms of an expectation over an individual's probability distribution function. This is based on the intuition that as the number of particles grows to infinity, there is a decoupling effect such that a single particle's statistical properties can effectively approximate the empirical distribution produced by all particles [7]. More specifically, as N tends to infinity, the individual dynamics may be written in the limiting form:

$$dx_t = b[x_t, \mu_t]dt + \sigma dw_t \quad (2)$$

which is the celebrated McKean-Vlasov equation. Here $b[x, \mu_t] = \int b(x, y)\mu_t(dy)$ for the probability distribution μ_t. This equation, as well as its variants, has been extensively studied in physics, stochastic analysis, and partial differential equations. The noise w_t may be determined in different ways. For instance, if one intends to approximate x_1 in (1) by x_t, one may set $w_t = w_1$ as the driving Brownian motion in (2)

and $x_0 = x_1(0)$. This leads us to the notion of a consistent pair for solving equation (2).

Definition 1: The pair (x_t, μ_t) is called a consistent pair if x_t is a solution to the SDE (2) and μ_t is its distribution for all $t \ge 0$. $\square$

Note that by introducing the density function $p_t(x)$ for x_t, one may recast (2) in the form of a Fokker-Planck equation whose coefficient depends on $p_t(x)$ itself. For a detailed analysis on the existence and uniqueness of a solution to (2), see [5]. For a weak convergence relation between solutions of (1) and (2), and large deviation analysis for μ_t, see [24], [9] and references therein.

B. The stochastic dynamic game

We consider an n dimensional stochastic system

$$dz_i = (az_i + bu_i)dt + \psi^{(n)}dt + \sigma dw_i, \quad 1 \le i \le n, \quad t \ge 0, \quad (3)$$

where $\{w_i, 1 \le i \le n\}$ denotes n independent standard scalar Wiener processes, $\psi^{(n)} = \frac{1}{n}\sum_{i=1}^{n}\psi(z_i)$, and ψ is a deterministic function of z_i. Hence, $\psi^{(n)}$ may be looked at as a driving term imposed by the population. The initial states $z_i(0)$ are i.i.d. and independent of $\{w_i, 1 \le i \le n\}$. In addition, $E|z_i(0)|^2 < \infty$ and $b \ne 0$. Each state component shall be referred to as the state of the corresponding individual (also to be called an agent or a player).

The individual cost for the ith agent takes the form:

$$J_i(u_i, \phi^{(n)}) \triangleq E\int_0^\infty e^{-\rho t}[(z_i - \Lambda(\phi^{(n)}))^2 + ru_i^2]dt, \quad (4)$$

where $\phi^{(n)} = \frac{1}{n}\sum_{i=1}^{n}\phi(z_i)$ with $\phi : \mathbb{R} \to \mathbb{R}$, and $\rho, r > 0$. To set the reference trajectory in a more general form, the nonlinear function $\Lambda : \mathbb{R} \to \mathbb{R}$ is introduced to act on the average term $\phi^{(n)}$. The linking term $\Lambda(\phi^{(n)})$ gives a measure of the average effect of the mass formed by all agents.

For the system configuration $z = (z_1, \cdots, z_n)$ at any given time t, define the empirical distribution $\varepsilon_z = \frac{1}{n}\sum_{i=1}^{n}\delta_{z_i}$. Then the coupling terms in the individual dynamics and costs are functionals of ε_z.

For the above system, the main objective is to seek individual control strategies and appropriately characterize their optimality, and a standard approach is to analyze Nash (equilibrium) strategies. However, within this basic solution scheme, each agent needs to have state information of all other agents, and this leads to a very high control complexity under large-population conditions. This motivates us to search for lower complexity control strategies.

C. The limiting game with an infinite population

For the n agent dynamic game, as $n \to \infty$, and following the complexity reducing strategies of statistical mechanics, one may attempt to use a distribution μ_t over a continuum to approximate the empirical distribution of $(z_1, \cdots, z_n)$. Thus we write the dynamics with the new state variable x_t:

$$dx_t = (ax_t + bu_t + \psi[\mu_t])dt + \sigma dw_t, \quad (5)$$

where the distribution of x_0 is the same as that of $z_i(0)$ and the coupling term is $\psi[\mu_t]$ defined in the form:

$$\psi[\mu_t] = \int_{\mathbb{R}} \psi(y)\mu_t(dy). \qquad (6)$$

And corresponding to (4), we set the cost function:

$$J(u, \phi[\mu]) \triangleq E \int_0^\infty e^{-\rho t} \left[(x_t - \Lambda(\phi[\mu_t]))^2 + ru_t^2 \right] dt, \quad (7)$$

where $\phi[\mu_t]$ is determined by replacing ψ by ϕ in (6). We may write $\Lambda(\phi[\mu_t]) = \Lambda \circ \phi[\mu_t]$.

We give the interpretation for the controlled system dynamics as follows. For a given feedback control u_t, as a function of (t, x_t), we look for a consistent pair (x_t, μ_t) such that the distribution of x_t in the closed-loop system coincides with μ_t for all $t \geq 0$. From this point of view, (5) may be looked at as a controlled McKean-Vlasov equation with the control performance measured by (7).

III. LOCALIZED CONTROL SYNTHESIS VIA THE NCE METHODOLOGY

This section is based on the infinite population model specified by (5)-(7), where the pair (x_t, μ_t) in (5) is interpreted as in the McKean-Vlasov equation (2). At first glance, one might think of treating this limiting problem as an optimal control problem:

(P) The objective is formulated as minimizing $J(u, \phi[\mu])$ by a control law $\hat{u}$ such that any other control generates a higher cost. Here u_t explicitly controls μ_t. ☐

However, after further thinking, one would be cautious about such an approach if the limiting control problem is intended to give an approximate solution to the finite population game problem in Section II-B. The reason is that μ_t in (5) is intended to approximate the empirical distribution ε_z in an n agent system with large n. Hence, when the given agent's strategy is adjusted, its impact on ε_z is negligible. It is seen that (P) is not consistent with such a situation.

Thus, we formulate the following approach:

(Q) By the NCE Methodology, we mean the overall game decomposition into an optimal control problem involving a mass effect measure (treated as an exogenous quantity) μ_t, $t \geq 0$, and a closed-loop McKean-Vlasov equation; these two parts are related to each other by the optimal control law derived from the former. The key step here is to construct a mutually consistent pair of (i) the mass effect and (ii) the individual strategies such that the latter not only each constitute an optimal response to the mass effect but also collectively produce that mass effect. ☐

In this setting, we designate the NCE Principle as the property that the resulting scheme is consistent in the sense that the prescribed control laws produce sample paths which produce the mass effect with respect to which the individual optimal control is derived. It is a property of this overall closed-loop behavior that each agent's optimal behavior with respect to all other agents holds in the game theoretic Nash sense. The solution to the overall problem, and hence the demonstration of the NCE Principle, or Property, relies on finding a fixed point mass effect.

A. Control synthesis based on the NCE methodology

We introduce the function space $C([0,\infty), \mathbb{R}^2) = \{\bar{f} = (\bar{f}_1, \bar{f}_2) : \bar{f}_i \in C[0,\infty), i = 1, 2\}$, and let $|\bar{f}|_\infty = \max_{i=1,2} \sup_{t \geq 0} |\bar{f}_i(t)|$. We further define $C_b([0,\infty), \mathbb{R}^2) = \{\bar{f} \in C([0,\infty), \mathbb{R}^2), |\bar{f}|_\infty < \infty\}$, which is a Banach space under the norm $|\cdot|_\infty$. The same notion $|\cdot|_\infty$ is also used for scalar function $x \in C[0,\infty)$, i.e., $|x|_\infty = \sup_{t \geq 0} |x(t)|$, and $C_b([0,\infty), \mathbb{R})$ is defined similarly.

Based on the NCE methodology, the control u_t should be optimized as if it were acting with respect to exogenous signals ψ^0 and ϕ^0 (corresponding to $\psi[\mu_t]$ and $\phi[\mu_t]$) as functions of time t. Subsequently, one may compute ψ^0 and ϕ^0 using the closed-loop dynamics based on the control which has been designed in the first step.

We set $f_1 = \psi[\mu_t]$ and $f_2 = \Lambda \circ \phi[\mu_t]$. Since f_1 and f_2 are treated as exogenous signals, it is natural for the given agent to take an optimal tracking action; see [14], [15], [16] for details about the motivation for such a control design. We introduce the scalar algebraic Riccati equation:

$$\rho \Pi = 2a\Pi - \frac{b^2}{r}\Pi^2 + 1,$$

and the unique positive solution is given by $\Pi = (\frac{b^2}{r})^{-1}[a - \frac{\rho}{2} + \sqrt{(a - \frac{\rho}{2})^2 + \frac{b^2}{r}}]$.

To begin with, we assume $(f_1, f_2) \in C_b([0,\infty), \mathbb{R}^2)$, and such a property will be established in the proof of Theorem 2. After taking $(f_1, f_2) \in C_b([0,\infty), \mathbb{R}^2)$ in the dynamics and cost, the individual's optimal tracking leads to the closed-loop equation system:

$$dx_t = (a - \frac{b^2}{r}\Pi)x_t dt - \frac{b^2}{r}s dt + f_1 dt + \sigma dw_t, \quad (8)$$

$$\rho s = \frac{ds}{dt} + as - \frac{b^2}{r}\Pi s + \Pi f_1 - f_2, \qquad (9)$$

where $s \in C_b([0,\infty), \mathbb{R})$. As long as s is bounded, as required in the optimal tracking solution, one can determine a unique s on $[0,\infty)$ and the initial condition for s does not need to be separately specified [15]. The control law associated with the above equation system (8)-(9) is

$$u_t = -\frac{b}{r}(\Pi x_t + s). \qquad (10)$$

In fact, (10) is the optimal control law for the auxiliary control problem with the dynamics and cost given by

$$dx_t = (ax_t + bu_t + f_1)dt + \sigma dw_t,$$

$$J = E \int_0^\infty e^{-\rho t}[(x - f_2)^2 + ru^2]dt.$$

Denote

$$\beta_1 = -a + \frac{b^2}{r}\Pi = -\frac{\rho}{2} + \sqrt{(a - \frac{\rho}{2})^2 + \frac{b^2}{r}}, \qquad (11)$$

$$\beta_2 = -a + \frac{b^2}{r}\Pi + \rho = \frac{\rho}{2} + \sqrt{(a - \frac{\rho}{2})^2 + \frac{b^2}{r}}. \qquad (12)$$

Note that $\beta_2 > \frac{\rho}{2}$. We may express the solution $s \in C_b([0,\infty), \mathbb{R})$ to (9) in the form $s(t) = \int_t^\infty e^{-\beta_2(\tau - t)}[\Pi f_1 - f_2](\tau)d\tau$.

Now we write the closed-loop equation for x_t:

$$dx_t = (a - \frac{b^2}{r}\Pi)x_t dt + \psi[\mu_t]dt$$
$$- \frac{b^2}{r}\Pi \int_t^\infty e^{-\beta_2(\tau - t)} \{\Pi\psi[\mu_\tau] - \Lambda \circ \phi[\mu_\tau]\} d\tau dt$$
$$+ \sigma dw_t. \qquad (13)$$

This equation may be regarded as a generalized McKean-Vlasov equation in that at time t, the right hand side depends on the future distribution of the state x_τ, $\tau \geq t$. For specifying a solution to (13), Definition 1 introduced for the standard McKean-Vlasov equation may be adapted to equation (13) in an obvious manner, and accordingly, we call (x_t, μ_t) an (individual-mass) consistent pair.

B. Existence of an individual-mass consistent pair

For obtaining existence results, we make the assumptions:
(H1) The functions Λ, ϕ, ψ are Lipschitz continuous with Lipschitz constants L_Λ, L_ϕ, L_ψ, respectively. □
(H2) The constant $\beta_1 > 0$. □
(H3) $(L_\phi \vee L_\psi) \cdot \left[\frac{1}{\beta_1} + \frac{b^2\Pi}{r\beta_1\beta_2}(\Pi + L_\Lambda)\right] < 1$. □

Theorem 2: Assuming (H1)-(H3) hold, then there exists a consistent pair (x_t, μ_t) satisfying (13).

Proof. Letting $(f_1, f_2) \in C_b([0, \infty), \mathbb{R}^2)$, there is a unique strong solution to the SDE,

$$dx_t = (a - \frac{b^2}{r}\Pi)x_t dt + f_1 dt \qquad (14)$$
$$- \frac{b^2}{r}\Pi \int_t^\infty e^{-\beta_2(\tau - t)} \{\Pi f_1 - \Lambda \circ f_2\} d\tau dt + \sigma dw_t$$

with the fixed initial condition x_0. The solution is denoted as $x_t^{f_1, f_2}$, with marginal distribution $\mu_t^{f_1, f_2}$, $t \geq 0$. Denote $\mu_\cdot^{f_1, f_2} = \Gamma_1(f_1, f_2)$, which is a mapping from $C_b([0, \infty), \mathbb{R}^2)$ to the space $\mathcal{M}$ of probability distributions $\mu_\cdot$ indexed by $t \in [0, \infty)$. For $\mu_\cdot \in \mathcal{M}$, let $\hat{f}_1(t) = \psi[\mu_t]$ and $\hat{f}_2(t) = \phi[\mu_t]$, and denote $(\hat{f}_1, \hat{f}_2) = \Gamma_2(\mu_\cdot)$.

We define $H(f_1, f_2) = (H_1(f_1, f_2), H_2(f_1, f_2)) = \Gamma_2 \circ \Gamma_1(f_1, f_2)$ for $(f_1, f_2) \in C_b([0, \infty), \mathbb{R}^2)$. In fact, $H_1(f_1, f_2)(t) = E\psi(x_t^{f_1, f_2})$, $H_2(f_1, f_2)(t) = E\phi(x_t^{f_1, f_2})$, which are functions on $[0, \infty)$. It is easy to verify that $H_i(f_1, f_2)$, $i = 1, 2$, are continuous and bounded on $[0, \infty)$ by the Lipschitz continuity of ψ and ϕ. Hence $H = \Gamma_2 \circ \Gamma_1$ gives a well defined mapping from $C_b([0, \infty), \mathbb{R}^2)$ to itself.

Now we take $(f_1, f_2) \in C_b([0, \infty), \mathbb{R}^2)$, $(g_1, g_2) \in C_b([0, \infty), \mathbb{R}^2)$ and denote the two corresponding solutions by $x_t^f = x_t^{f_1, f_2}$ and $x_t^g = x_t^{g_1, g_2}$. It follows that

$$dx_t^f - dx_t^g = -\beta_1(x_t^f - x_t^g)dt + (f_1 - g_1)dt$$
$$- \frac{b^2}{r}\Pi \int_t^\infty e^{-\beta_2(\tau - t)}\Delta_{f,g}(\tau)d\tau dt$$

where $\Delta_{f,g} = \Pi(f_1 - g_1) - (\Lambda \circ f_2 - \Lambda \circ g_2)$ and the initial condition satisfies $[x_t^f - x_t^g]|_{t=0} = 0$. It follows that

$$x_t^f - x_t^g = \int_0^t e^{-\beta_1(t-\tau)}(f_1 - g_1)(\tau)d\tau$$
$$- \frac{b^2}{r}\Pi \int_0^t \int_\theta^\infty e^{-\beta_1(t-\theta)}e^{-\beta_2(\tau-\theta)}\Delta_{f,g}(\tau)d\tau d\theta$$

which gives

$$|x_t^f - x_t^g| \leq \frac{1}{\beta_1}|f_1 - g_1|_\infty$$
$$+ \frac{b^2}{r}\Pi \cdot \frac{\Pi|f_1 - g_1|_\infty + L_\Lambda|f_2 - g_2|_\infty}{\beta_1\beta_2}. \qquad (15)$$

Hence

$$|H_1(f_1, f_2) - H_1(g_1, g_2)|_\infty = \sup_{t \geq 0}|E\psi(x_t^f) - E\psi(x_t^g)|$$
$$\leq \frac{L_\psi}{\beta_1}|f_1 - g_1|_\infty + L_\psi\frac{b^2}{r}\Pi \cdot \frac{\Pi|f_1 - g_1|_\infty + L_\Lambda|f_2 - g_2|_\infty}{\beta_1\beta_2},$$
$$|H_2(f_1, f_2) - H_2(g_1, g_2)|_\infty = \sup_{t \geq 0}|E\phi(x_t^f) - E\phi(x_t^g)|$$
$$\leq \frac{L_\phi}{\beta_1}|f_1 - g_1|_\infty + L_\phi\frac{b^2}{r}\Pi \cdot \frac{\Pi|f_1 - g_1|_\infty + L_\Lambda|f_2 - g_2|_\infty}{\beta_1\beta_2}.$$

Hence $H(f_1, f_2)$ is a contraction on $C_b([0, \infty), \mathbb{R}^2)$ (with the norm $|\cdot|_\infty$) by (H3), with a unique fixed point satisfying $H(f_1^*, f_2^*) = (f_1^*, f_2^*)$. Let $x_t^{f_1^*, f_2^*}$ be the corresponding solution to (14). Then $x_t^{f_1^*, f_2^*}$ and its marginal distribution μ_t^* give a consistent pair. We see that $\mu_\cdot^* = \Gamma_1 \circ \Gamma_2(\mu_\cdot^*)$. □

Remark: The Lipschitz constants $L_\psi, L_\phi, L_\Lambda$ provide a measure of the agent's sensitivity to the population effect in either the dynamics or costs. □

It should be noted that in the nonlinearly coupled model, an explicit construction for the operator H is in general unavailable, although it is well defined by a probabilistic calculation. However, it is of interest to mention that for the case $\psi(x) = \alpha x$ and $\phi(x) = x^2$, one may characterize the consistent pair (x_t, μ_t) by using a finite dimensional equation system involving deterministic functions on $[0, \infty)$ by a similar fixed point procedure as in [13], [14], [15]. The main reason is that if a consistent pair (x_t, μ_t) exists, the component x_t is Gaussian by (13); and therefore it can be specified by its mean and variance process on $[0, \infty)$.

To characterize uniqueness of the solution, we introduce the class: $\mathcal{C}$ consists of all $\{(x_t, \mu_t)$ which satisfies the conditions i) (x_t, μ_t) is a consistent pair to (13), ii) $\phi[\mu_t]$ and $\psi[\mu_t]$ are measurable functions of t on $[0, \infty)$, and iii) $\sup_{t \geq 0} \int_\mathbb{R} |y|\mu_t(dy) < \infty\}$.

Corollary 3: Under the conditions in Theorem 2, there is a unique solution (x_t, μ_t) to (13) in the class $\mathcal{C}$.

Proof. Assume a solution $(x_t', \mu_t') \in \mathcal{C}$. Write $f_1(t) = \psi[\mu_t']$ and $f_2(t) = \phi[\mu_t']$ and we can show that $(f_1, f_2) \in C_b([0, \infty), \mathbb{R}^2)$ since $\psi[\mu_t'] = E\psi(x_t')$ and $\phi[\mu_t'] = E\phi(x_t')$ and x_t' is given by the linear equation (13). By the proof of Theorem 2, we see that (f_1, f_2) is the unique fixed point determined therein. Thus x_t' and hence μ_t' coincide with the solution obtained in Theorem 2, and uniqueness follows. □

IV. ASYMPTOTIC NASH EQUILIBRIA FOR LARGE POPULATIONS

We now employ the control strategy obtained in the limiting system to the system of n agents described in Section II-B. We write

$$u_i = -\frac{b}{r}(\Pi z_i + s), \qquad 1 \leq i \leq n, \qquad (16)$$

based on (10), where the associated s, f_1, f_2 are determined from the fixed point procedure in Theorem 2. We first show a stabilizing property of the control law (16).

Theorem 4: Assuming (H1)-(H3), there exists a constant C independent of n such that for the closed-loop system under (16), we have $\sup_{1 \le i \le n} \sup_{0 \le t \le \infty} E z_i^2(t) \le C$. □

We can further show the following approximation result.

Lemma 5: Under (H1)-(H3) and the control law (16), as $n \to \infty$, we have

$$\sup_{0 \le t \le \infty} E \left| \frac{1}{n} \sum_{i=1}^{n} [z_i(t) - E z_i(t)] \right|^2 = O(\frac{1}{n}), \quad (17)$$

$$\sup_{0 \le t \le \infty} E \left| \frac{1}{n} \sum_{i=1}^{n} [\zeta(z_i(t)) - \zeta[\mu_t]] \right|^2 = O(\frac{1}{n}), \quad (18)$$

where the function ζ stands for ψ or ϕ. □

Now we state the asymptotic Nash equilibrium result; see e.g. [14], [15] for the definition of ε-Nash equilibrium.

Theorem 6: Assume (H1)-(H3) hold. For a system of n agents specified by (3)-(4), the set of control laws (16) is an ε-Nash equilibrium where $0 < \varepsilon \to 0$, as $n \to \infty$, i.e., the unilateral strategy change (allowed to take as a function of $(z_1, \cdots, z_n)$, as long as a closed-loop solution is well defined) of any given agent can produce a cost reduction for itself by at most $\varepsilon > 0$. □

V. EXPLICIT SOLUTIONS WITH LINEAR DYNAMICS

In this section we consider the case of linear dynamics and compute the explicit solutions for two different control schemes. To simplify the analysis, the linear coupling only appears in the costs as in [14].

We first analyze the solution based on the NCE methodology which gives a *decentralized* control law as each agent only requires its own state. Next we consider a system of n agents and compute the exact solution for the set of feedback Nash strategies (under the specific definition in [4]) and examine its asymptotic behavior as the population size n increases to infinity. For a fixed n, the resulting solution is characterized by a coupled algebraic Riccati equation and leads to *centralized* individual control laws where each agent needs to know the states of other agents. By its definition, the equilibrium property holds in a strong sense since it requires the control to possess the Nash equilibrium property for subgames (from any t_0 to ∞, with the attained state $x_i(t_0)$, $1 \le i \le n$, as the initial conditions). In the end we show that asymptotically the two different optimization approaches are equivalent; this indicates the decentralized strategy performs as effectively as the centralized one. However, it should be pointed out that for more complicated coupling or when the agents have different dynamic parameters, it is difficult if not impossible to carry out the full information based Nash strategy computation while the NCE based approach is powerful for dealing with more complicated models [15].

We note that unlike the standard optimal control case, the existence and uniqueness result for feedback Nash strategies for LQG games is far less complete. In our analysis, we assume there is a unique set of feedback Nash strategies

and our focus is on explicit calculation for the centralized solution.

A. The Nash equilibrium solution

We consider the case $\psi(z) = 0$, $\phi(z) = \gamma(z + \eta)$, and $\Lambda(z) = z$. The state aggregation equation system resulting from the NCE methodology reduces to

$$\rho s = \frac{ds}{dt} + as - \frac{b^2}{r} \Pi s - z^*, \quad (19)$$

$$\frac{d\bar{z}}{dt} = (a - \frac{b^2}{r} \Pi) \bar{z} - \frac{b^2}{r} s, \quad (20)$$

$$z^* = \phi(\bar{z}) = \gamma(\bar{z} + \eta), \quad (21)$$

where $\bar{z}$ is the mean of the state. Setting the derivatives to zero, we write a set of steady state equations as follows

$$\begin{cases} \beta_2 s(\infty) + z^*(\infty) = 0 \\ -\frac{b^2}{r} s(\infty) - \beta_1 \bar{z}(\infty) = 0 \\ \gamma \bar{z}(\infty) - z^*(\infty) = -\gamma \eta. \end{cases} \quad (22)$$

It can be verified that under the conditions in Theorem 2 we have $\Theta \triangleq \beta_2 \beta_1 - \frac{b^2 \gamma}{r} > 0$, and therefore (22) is nonsingular and has a unique solution $(s(\infty), \bar{z}(\infty), z^*(\infty))$. Denote $\lambda_1 = \frac{\rho - \sqrt{\rho^2 + 4\Theta}}{2} < 0$.

Proposition 7: [15] The unique bounded solution $(\bar{z}, s)$ in (19)-(20) is given by

$$\bar{z}(t) = \bar{z}(\infty) + (\bar{z}(0) - \bar{z}(\infty)) e^{\lambda_1 t},$$

$$s(t) = s(\infty) + \frac{\gamma - \alpha \Pi}{\beta_2 - \lambda_1} (\bar{z}(\infty) - \bar{z}(0)) e^{\lambda_1 t},$$

where $\lambda_1 < 0$ and $\beta_2 = -a + \frac{b^2}{r} \Pi + \rho$. □

Proposition 8: Assume the existence and uniqueness of a set of feedback Nash strategies for each population size $n > 1$ and the conditions in Theorem 2 hold. Then the ith agent's feedback Nash strategy is given as

$$u_i = -\frac{b}{r} (K_1 z_i + \sum_{j \ne i} K_2 z_j + s_1), \quad 1 \le i \le n,$$

where $K_1 = \Pi + O(\frac{1}{n})$, $K_2 = \frac{c_2}{n} + o(\frac{1}{n})$ and

$$c_2 = (\frac{b^2}{r})^{-1} \left[-(\beta_1 + \frac{\rho}{2}) + \sqrt{(\beta_1 + \frac{\rho}{2})^2 - \frac{b^2 \gamma}{r}} \right] < 0$$

$$s_1 = -\frac{\gamma \eta}{-a + \frac{b^2}{r} c_{10} + \rho + \frac{b^2}{r} c_2} + O(\frac{1}{n}).$$

And moreover, if we set $\bar{z}^{(n)} = \sum_{i=1}^{n} E z_i$, then $\bar{z}^{(n)} = \bar{z} + O(\frac{1}{n})$, where $\bar{z}$ is given by Proposition 7. □

Thus, asymptotically the decentralized and centralized control strategies lead to the same mass effect $\bar{z}$.

VI. DISCUSSIONS ON MORE GENERAL DYNAMICS

In the above analysis, the McKean-Vlasov equation for the dynamic game is linear in the state x_t, which is separated from the nonlinear coupling term in the drift. Such a structural property makes it possible to have explicit calculation of the individual control via optimal tracking. Moreover, with the separation of x_t from the coupling term, the mass effect

may be approximated by a deterministic function. All these features combined together simplify the existence proof for the consistent pair. For future work, there exist different ways for generalizing our methodology to more general models:

- There drift term is given in the form of summation of $\psi(x_i, x_j)$, and the control term is given as a nonlinear function of its own state and control. In general, to deal with such a situation, one needs to develop more general fixed point techniques for existence analysis.

  Also, it is of interest to study the steady state behavior of the population distribution evolution.

- The modelling of non-uniform agents in the setting of McKean-Vlasov equations. One potential approach is to re-construct the driving distribution in the McKean-Vlasov equation as the averaging of a family (indexed via a parameter set) of distributions, which mimics the aggregate behavior of multi-class interacting particles.

Below we give a heuristic description of applying the NCE methodology to the nonlinear model

$$dx_t = b[x_t, \mu_t]dt + g(x_t, u_t)dt + \sigma dw_t \qquad (23)$$

where $b[x_t, \mu_t]$ is defined as in the uncontrolled McKean-Vlasov equation (2). The range space of u_t is denoted as the set U. The cost function is given as

$$J = E \int_0^\infty e^{-\rho t}\{l_1[x_t, \mu_t] + l_2(x_t, u_t)\}dt, \qquad (24)$$

where $l_1[x, \mu] = \int_{\mathbb{R}} l_1(x, y)\mu_t(dy)$, and l_1, l_2 are deterministic functions. We assume sufficient regularity of the system (in terms of the dynamics and cost) to ensure:

(i) There exists a distribution μ_t^* such that the HJB equation

$$\rho V(x, t) = \frac{\partial V}{\partial t} + \min_{u \in U}\left\{[b[x, \mu_t^*] + g(x, u)]\frac{\partial V}{\partial x}\right.$$
$$\left. + l_1[x, \mu_t^*] + l_2(x, u)\right\} + \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial x^2} \qquad (25)$$

has a smooth solution $V(t, x)$ and the optimal control $u = u^*(t, x)$ is continuous in (t, x) and Lipschitz in x;

(ii) For the control $u_t = u^*(t, x)$ in (23), there exists a strong solution x_t with the distribution μ_t^*. We have the closed-loop equation

$$dx_t = b[x_t, \mu_t^*]dt + g(x_t, u^*(t, x_t))dt + \sigma dw_t \qquad (26)$$

which induces a Fokker-Planck equation for μ_t^*.

The HJB equation (25) results from the NCE based control synthesis when $\{\mu_t^*, t \geq 0\}$ is first presumed to be known. By the above procedure, the limiting game problem is decomposed into a stochastic control problem to be solved along μ_t^* and a specification of the density for x_t by a Fokker-Planck equation induced from the closed-loop dynamics.

ACKNOWLEDGEMENT

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] 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. Rep. No. 4154, France, 2001.

[3] R. Balescue. *Equilibrium and Nonequilibrium Statistical Mechanics*, John Wiley, New York, 1975.

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

[5] S. Benachour, B. Roynette, D. Talay and P. Vallois. Nonlinear self-stabilizing processes. I: existence, invariant probability, propagation of chaos. *Stochastic Processes Appl.*, vol. 75, no. 2, pp. 173-201, 1998.

[6] S. Benachour, B. Roynette and P. Vallois. Nonlinear self-stabilizing processes. II: convergence to invariant probability. *Stochastic Processes Appl.*, vol. 75, no. 2, pp. 203-224, 1998.

[7] D. Chandler. *Introduction to Modern Statistical Mechanics*. Oxford University Press, New York, 1987.

[8] D. A. Dawson. Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *J. Statist. Phys.*, vol. 31, no. 1, pp. 29-85, 1983.

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

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

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

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

[13] 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.

[14] 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.

[15] 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.

[16] 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 Conf.*, Seville, Spain, pp. 1050-1055, Dec. 2005.

[17] M. Kac. *Probability and Related Topics in Physical Science*, Interscience, New York, 1958.

[18] Y. L. Klimontovich. *Statistical Physics*, Harwood Academic Publ., New York, 1986.

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

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

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

[22] 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.

[23] R. Srikant and T. Basar. Iterative computation of noncooperative equilibria in nonzero-sum differential games with weakly coupled players. *J. Optim. Th. Appl.*, vol. 71, no. 1, pp. 137-168, Oct. 1991.

[24] A.-S. Sznitman. Topics in propagation of chaos. In Ecole d'Eté de Probabilitiés de Saint-Flour XIX -1989, Lect. Notes Math. 1464, pp. 165-252, Springer-Verlag, Berlin, 1991.