Social Certainty Equivalence in Mean Field LQG Control: Social, Nash and Centralized Strategies

Minyi Huang Peter E. Caines Roland P. Malhamé

***Abstract*— We study social decision problems and Nash games for a class of linear-quadratic-Gaussian (LQG) models with N decision makers possessing different dynamics. For the social decision case, the basic objective is to minimize a social cost as the sum of N individual costs containing mean field coupling, and the exact social optimum requires centralized information. Continuing from the previous work (Huang, Caines, and Malhamé, 2009 Allerton Conference), we develop decentralized cooperative optimization so that each agent only uses its own state and a function which can be computed off-line. We prove asymptotic social optimality results with general vector individual states and continuum dynamic parameters. In finding the asymptotic social optimum, a key step is to let each agent optimize a new cost as the sum of its own cost and another component capturing its social impact on all other agents. We also discuss the relationship between the socially optimal solution and the so-called Nash Certainty Equivalence (NCE) based solution presented in previous work on mean field LQG games, and for the NCE case we illustrate a cost blow-up effect due to the strength of interaction exceeding a certain threshold.**

I. Introduction

Mean field decision models have attracted extensive attention due to their significance in many domains [3], [5], [6], [7], [9], [17], [18], [22], [26], [28]. In such models, a distinctive feature is the interaction between any given agent and the average effect of the overall population of agents. In the search for decentralized optimization paradigms, game theoretic solutions have been successfully developed by different researchers [9], [10], [28], [29], [20], [21], [22], [23]; along this line, decentralized solutions may be obtained by identifying a consistency relationship between the individual-mass interaction such that in the population limit each individual optimally responds to the mass effect and these individual strategies also collectively produce the same mass effect presumed in the first place [9], [10], [13], [14], [8]. Under reasonable conditions we have shown the existence of a mass effect satisfying such a fixed point property and proven that the resulting set of decentralized individual strategies is an asymptotic Nash equilibrium. This solution property has been designated as the Nash Certainty Equivalence (NCE) principle [10], [15]. Closely related mean field approximation approaches were developed in [28], [29] using the notion of oblivious equilibria (OE) for games on industry dynamics, and OEs with unbounded costs were analyzed in [1]. The works [20], [21], [22] adopted a similar consistency based approach for mean field games, but for each finite population size a simplifying assumption was used stipulating that each agent's strategy depends only on its own driving Brownian motion. In [23], the interaction-consistency based approach was applied to models with long term average costs. A game theoretic framework was proposed in [32] for the control of coupled nonlinear oscillators, and the mean field approximation approach was applied to obtain decentralized strategies and further study phase transition of the closed-loop system.

The above game theoretic solution framework is based on the assumption that these agents are individually incentive driven and noncooperative. In [11], within the mean field modeling we studied a different situation where the agents are cooperative and seek socially optimal decisions. We note that the notion of social optima has long been a central issue in decision problems with multiple agents, and Pareto optimality is well known as one approach for characterizing social optimality [2], [24]. The goal of this paper is to study how the agents in a mean field LQG model should choose their strategies for optimizing a social objective in the social decision setting, or optimizing individual objectives in a Nash game setting. We generalize the preliminary analysis in [11] which considered a finite number of classes of agents and proved social optimality results for uniform agents with scalar states. For the social decision problem, we consider both i) centralized strategies where each agent may use the state information of all agents and ii) decentralized strategies where each agent only uses local information. Related numerical comparison of the optimized costs between the socially optimal solution and the NCE based game theoretic solution was provided in [19], where each agent assigns nonuniform cost coupling weights across the population. For stochastic differential games, cooperation issues were addressed in [31] by extending concepts such as coalition and Shapley value to dynamic models. But in general, this approach will not lead to social optima.

As a historical connection, it is worth briefly comparing the current work with classical team decision problems, where all the agents share a common cost but have different information regarding the system state and other agents' strategies, which is specified by the so-called information structure [16], [27]. Our social optimization problem with decentralized information may be viewed as a mean field generalization of team problems where each agent has a

M. Huang is with School of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada. mhuang@math.carleton.ca.
P. E. Caines is with Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada. peterc@cim.mcgill.ca.
R. P. Malhamé is with Department of Electrical Engineering, École Polytechnique de Montréal, Montreal, QC H3C 3A7, Canada. roland.malhame@polymtl.ca.

priori information but no real time information on other agents. The local forecast of the mean field effect is now a part of the control problem. This leads to the development of the Social Certainty Equivalence (SCE) methodology, whereby the fundamental idea is to first quantify the social cost change due to the control perturbation of a given agent and subsequently apply mean field approximations. We mention that a team formulation of mean field Markov decision with discount was developed in [26]. A fixed point approach was used there to show the existence of an optimum when restricted to stationary strategies. This approach ignores the transient behavior of the mean field and may lead to optimality loss.

The organization of the paper is as follows. The socially optimal control problem is formulated in Section II. The centralized solution is analyzed in Section III. Section IV presents the SCE methodology and the asymptotic optimality theorem. Section V analyzes the scalar case and provides a comparison with the NCE equation system. Section VI presents the explicit calculation of the asymptotic average social optimum, and Section VII concludes the paper.

II. The Socially Optimal Control Problem

A. Dynamics and Costs

Consider a system of N agents. The dynamics of agent i are given by the stochastic differential equation (SDE)

$$dx_i = A(\theta_i)x_i dt + Bu_i dt + DdW_i, \qquad t \geq 0. \tag{1}$$

The underlaying filtration is $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, P)$, where $(\mathscr{F}_t)_{t\geq 0}$ is a collection of non-decreasing σ-algebras. The state x_i and control u_i are, respectively, n and n_1 dimensional vectors. The initial states $\{x_i(0),\ 1 \leq i \leq N\}$, are independent and measurable on $\mathscr{F}_0$, and $E|x_i(0)|^2 < \infty$. The noise processes $\{W_i,\ 1 \leq i \leq N\}$ are n_2 dimensional independent standard Brownian motions adapted to $\mathscr{F}_t$, which are also independent of $\{x_i(0), 1 \leq i \leq N\}$. The constant matrices $A(\cdot)$, B and D all have compatible dimensions. Here θ_i denotes a dynamic parameter associated with agent i. The variability of θ_i is used to model a population of nonuniform agents. We only take $A(\cdot)$ to be dependent on θ_i for the purpose of notational simplicity. When other matrix parameters for agent i also depend on θ_i, the analysis is similar and will not be given in detail. For notational brevity the time argument for a process (x_i, u_i, etc.) is often suppressed when the value of that process at time t is used. Denote $x = [x_1^T, \dots, x_N^T]^T$ and $u = [u_1^T, \dots, u_N^T]^T$.

The individual cost for agent i, $1 \leq i \leq N$, is given by

$$\begin{aligned} J_i(u(\cdot)) =& E\int_0^\infty e^{-\rho t}\{\left[x_i - \Phi(x^{(N)})\right]^T Q\left[x_i - \Phi(x^{(N)})\right] \\ &+ u_i^T R u_i\} dt, \end{aligned} \tag{2}$$

where $\Phi(x^{(N)}) = \Gamma x^{(N)} + \eta$ and $x^{(N)} = (1/N)\sum_{i=1}^N x_i$. We call $x^{(N)}$ the mean field term. All the constant matrices or vector Γ, $Q \geq 0$, $R > 0$ and η have compatible dimensions. We use $u(\cdot)$ (or u) to denote the N individual control processes, and also call it the control of the overall system. By a slight abuse of notation sometimes we will write $u = (u_1, \dots, u_N)$. The social cost is defined as

$$J_{\text{soc}}^{(N)}(u(\cdot)) = \sum_{i=1}^N J_i(u(\cdot)).$$

The objective is for the agents to minimize $J_{\text{soc}}^{(N)}$. To achieve this, from the point of view of an individual's control selection, it is necessary to maintain a delicate balance in reducing its own cost and also taking into account the social impact of such reductions (i.e., affecting the sum of the costs of all other agents).

For the large population system, a natural way for modeling the sequence of dynamic parameters $\theta_1, \dots, \theta_N$ is to view it as being sampled from an underlying parameter space such that when $N \to \infty$, the sequence exhibits certain statistical properties; this is made precise by assumption (**A1**) below. However, we stipulate that $\{\theta_i, i \geq 1\}$ is treated as a deterministic sequence. We assume that each θ_i is in a compact set $\Theta \subset \mathbb{R}^\kappa$. For a given N, define the empirical distribution function

$$F_N(\theta) = \frac{1}{N}\sum_{i=1}^N 1_{\{\theta_i \leq \theta\}}, \tag{3}$$

where $\theta \in \mathbb{R}^\kappa$ and $\theta_i \leq \theta$ holds componentwise for the two vectors. We make the assumptions.

(**A1**) There exists a distribution function $F(\theta)$ on $\mathbb{R}^\kappa$ such that F_N converges to F weakly, i.e., for any bounded and continuous function $\varphi(\theta)$ on $\mathbb{R}^\kappa$,

$$\lim_{N\to\infty}\int \varphi(\theta)dF^{(N)}(\theta) = \int \varphi(\theta)dF(\theta). \qquad \diamondsuit$$

(**A2**) The initial states $\{x_i(0), 1 \leq i \leq N\}$ are independent, $Ex_i(0) = m_0$ for a fixed m_0 and all $i \geq 1$, and there exists $c_0 < \infty$ independent of N such that $\sup_{i\geq 1} E|x_i(0)|^2 \leq c_0$. $\diamondsuit$

(**A3**) $A(\theta)$ is a continuous matrix function of $\theta \in \Theta$, where Θ is a compact subset of $\mathbb{R}^\kappa$. $\diamondsuit$

(**A4**) For $\theta \in \Theta$, (i) the pair $[A(\theta) - (\rho/2)I, B]$ is stabilizable and (ii) the pair $[Q^{1/2}, A(\theta) - (\rho/2)I]$ is detectable. $\diamondsuit$

In the special case where $\Theta = \{1, \dots, K\}$ for some finite integer K, the empirical distribution of $\theta_1, \dots, \theta_N$ reduces to a probability mass function on Θ, denoted as $\pi^{(N)}$, and (**A1**) reduces to the convergence of $\pi^{(N)}$ to a limit π. In addition, (**A3**) is trivially true under the discrete topology of Θ where an open set is the union of singletons or is the null set. Thus (**A3**) becomes redundant.

It is possible to generalize our analysis to different initial means as long as $\{Ex_i(0), i \geq 1\}$ has a limiting empirical distribution (see related discussions in [10]).

B. Two Solutions Based on Different Information Patterns

We will study two problems for optimizing $J_{\text{soc}}^{(N)}$ according to different information patterns.

Problem I-A – Find a social solution $(u_1, \dots, u_N)$ with centralized information (SSCI), where each u_i in a feedback form is a function of $(t, x_1, \dots, x_N)$ for attaining the minimum of $J_{\text{soc}}^{(N)}$.

Problem I-B – Find a social solution $(u_1, \dots, u_N)$ with decentralized information (SSDI), where each u_i in a feedback form is a function of (t, x_i). Note that when restricted to decentralized information, the set of controls of the N agents will not in general attain the same cost as in Problem I-A. Instead, a set of decentralized strategies $\{u_i,\ 1 \le i \le N\}$ is sought such that the optimality loss with respect to Problem I-A in minimizing $J_{\rm soc}^{(N)}$ tends to zero when $N \to \infty$.

For comparison with our previous work, the following problem will also be reviewed.

Problem II – Find a competitive solution $(u_1, \ldots, u_N)$ with decentralized information (CSDI), where agent i is associated with the cost J_i and the objective is to obtain a set of ε-Nash strategies such that each u_i is a function of (t, x_i).

For a detailed account of this competitive solution framework, the reader is referred to [10].

III. The Centralized Solution

Problem I-A leads to a standard LQG control problem and the optimal control law $(\hat{u}_1, \ldots, \hat{u}_N)$ may be determined from a high dimensional algebraic Riccati equation (ARE) if the standard stabilizability and detectability conditions are satisfied [30]. We first give a reinterpretation of each individual component in $(\hat{u}_1, \ldots, \hat{u}_N)$, which will motivate the construction of decentralized strategies via a mean field approximation argument. The analysis here is similar to the person-by-person optimality characterization of team decision problems [16].

A. Centralized Optimal Control: Person-by-person Optimality

To facilitate further analysis, denote $\mathscr{F}_t^0 = \sigma(x(0), W_1(s), \ldots, W_N(s),\ \ s \le t)$ for $t \ge 0$, which is the σ-algebra generated by $x(0)$ and the Brownian motions up to time t. Denote the control set $\mathscr{U}_o := \{(u_1, \ldots, u_N) | u_i(t, \omega)$ is adapted to $\mathscr{F}_t^0, \forall i\}$, where $\omega \in \Omega$ explicitly indicates the dependence of u_i on the sample. Each $u = (u_1, \ldots, u_N) \in \mathscr{U}_o$ is called an $\mathscr{F}_t^0$-adapted control and may be viewed as a functional of $x(0)$ and the Brownian motions without being related to the state process $x(t)$. The distinction between $\mathscr{F}_t^0$-adapted controls and feedback strategies should be clear from the context.

Given $x(0)$, a very important observation for Problem I-A is that any feedback control law $u(t,x) = (u_1(t,x), \ldots, u_N(t,x))$, if continuous in (t,x) and Lipschitz continuous in x (thus ensuring a unique strong solution to the closed-loop system), naturally induces a process on $[0, \infty)$, denoted as $u(t, \omega)$ which belongs to $\mathscr{U}_o$. This is due to the fact that we may express the closed-loop solution $x(t)$ in terms of $x(0)$ and the Brownian motions. Note that under the stabilizability and detectability conditions, the optimal control law is a linear feedback control law, indeed satisfying the above continuity assumptions; the reader is referred to [4], [30], [8] for detail. The benefit of introducing the control set $\mathscr{U}_o$ is that one can fix the controls of other agents while allowing a selected agent to perturb its control.

Let the optimal control minimizing $J_{\rm soc}^{(N)}$ be denoted by $\hat{u} = (\hat{u}_1, \ldots, \hat{u}_N)$, which is now interpreted as a control from $\mathscr{U}_o$. Let $\hat{x}_i$ correspond to $\hat{u}_i$. Denote $\hat{x}^{(N)} = (1/N)\sum_{j=1}^N \hat{x}_j$ and $\hat{x}_{-i}^{(N)} = (1/N)\sum_{j \ne i, j=1}^N \hat{x}_j$. For $1 \le i \le N$, let $\mathscr{U}_{oi} := \{u_i | u_i(t, \omega)$ is adapted to $\mathscr{F}_t^0\}$. Given $u \in \mathscr{U}_o$, define $u_{-i} = (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N)$, and $\hat{u}_{-i}$ is defined similarly.

Lemma 1: [11] Assume (**A4**) and let $x(0)$ be given. Suppose that $\hat{u} \in \mathscr{U}_o$ attains the minimum of $J_{\rm soc}^{(N)}$ with the admissible control set $\mathscr{U}_o$. Then $\hat{u}_i$ is the unique optimal control for the control problem:

$$\textbf{(P0)} \qquad dx_i = A(\theta_i)x_i dt + Bu_i dt + DdW_i, \qquad t \ge 0, \quad (4)$$

$$J_i^0(u_i) = J_{\rm soc}^{(N)}(\hat{u}_1, \ldots, \hat{u}_{i-1}, u_i, \hat{u}_{i+1}, \ldots, \hat{u}_N), \quad (5)$$

where $J_i^0(u_i)$ is to be minimized with $u_i \in \mathscr{U}_{oi}$. □

Let $\hat{u}$ be as in Lemma 1. Due to the coupling of the states in $J_{\rm soc}^{(N)}$, $\hat{u}_i$ in general will depend on all W_j, $j = 1, \ldots, N$. In the following we give a reformulation of Problem (P0).

Lemma 2: [11] Assume (**A4**). In finding the optimal control $\hat{u}_i \in \mathscr{U}_{oi}$, Problem (P0) is equivalent to the optimal control problem:

$$\textbf{(P1)} \qquad dx_i = A(\theta_i)x_i dt + Bu_i dt + DdW_i, \qquad t \ge 0, \quad (6)$$

$$J_i^1(u_i) = E\int_0^\infty e^{-\rho t} L(x_i, \hat{x}_{-i}^{(N)}, u_i)(t) dt, \quad (7)$$

where $J_i^1(u_i)$ is to be minimized with $u_i \in \mathscr{U}_{oi}$, and

$$\begin{aligned} L = {} & x_i^T[(I - \Gamma/N)^T Q (I - \Gamma/N) + (N-1)/N^2 \Gamma^T Q \Gamma] x_i \\ & - 2(\Gamma \hat{x}_{-i}^{(N)} + \eta)^T Q (I - \Gamma/N) x_i \\ & - 2\{[I - (1 - 1/N)\Gamma]\hat{x}_{-i}^{(N)} - (1 - 1/N)\eta\}^T Q \Gamma x_i + u_i^T R u_i. \end{aligned} \quad (8)$$

□

Lemma 2 shows that all other agents' effect on $J_i^1(u_i)$ appears in the form of $\hat{x}_{-i}^{(N)}$. This feature is useful for finding a decentralized suboptimal control u_i in Section IV by a deterministic approximation of $\hat{x}_{-i}^{(N)}$.

B. Explicit Solutions: Uniform Agents with Scalar States

For uniform agents with scalar individual states, $A(\theta_i)$ in (1) is denoted by the same number A. To avoid triviality, suppose $B \ne 0$. Without loss of generality, we set $Q = 1$ in (2). To quantify the interaction between an individual agent and the mean field, we introduce the parametrization

$$\Gamma = \gamma, \quad \eta = \gamma \eta_0, \quad (9)$$

where γ is a parameter and η_0 is fixed. So Γ and η are scaled by the same parameter γ. If we apply the parametrization (9) to the original mean field model (1)-(2), a larger γ means stronger interaction between x_i and $x^{(N)} + \eta_0$. Now $\Phi(x^{(N)}) = \gamma(x^{(N)} + \eta_0)$. Denote $\mathbf{1}_N = [1, \ldots, 1]^T$ consisting of N ones. By rearranging the integrand of $J_{\rm soc}^{(N)}$, we write

$$J_{\rm soc}^{(N)} = E\int_0^\infty e^{-\rho t}(x^T \widehat{Q} x + 2Gx + Ru^T u + N\gamma^2\eta_0^2) dt, \quad (10)$$

where $G = \gamma\eta_0(\gamma - 1)\mathbf{1}_N^T$ and $\widehat{Q}$ is given in the form

$$\widehat{Q} = \begin{pmatrix} \alpha & \beta & \cdots & \beta \\ \beta & \alpha & \cdots & \beta \\ \vdots & & \ddots & \vdots \\ \beta & \beta & \cdots & \alpha \end{pmatrix}, \quad (11)$$

and $\alpha = (1-\frac{\gamma}{N})^2 + \frac{(N-1)\gamma^2}{N^2} = 1 + \frac{\gamma^2-2\gamma}{N}$, $\beta = \frac{\gamma^2-2\gamma}{N} = \alpha - 1$. The eigenvalues of $\widehat{Q}$ are given by

$$\lambda_1 = \alpha + (N-1)\beta = (1-\gamma)^2, \quad \lambda_2 = \ldots = \lambda_N = \alpha - \beta = 1.$$

A similar LQG control problem was briefly analyzed in [10], where the coupling term in agent i's cost J_i' is $\Phi'(x) = \gamma(1/N\sum_{j\neq i} x_j + \eta_0)$ instead of $\Phi(x^{(N)})$ and where the social cost is $J' = \sum_{i=1}^N J_i'$. A very subtle difference between the two formulations is that given any γ, the pair $[Q'^{1/2}, (A-\rho/2)I_N]$ in [10] is always observable for all sufficiently large N, where Q' appears in the quadratic term $x^T Q' x$ in J', while this is not the case when Φ is used. As it turns out, when Φ is used and $\gamma = 1$, span$\{\mathbf{1}_N\}$ is always an unobservable subspace for the system so that the state vector's component within span$\{\mathbf{1}_N\}$ is not penalized by the cost. To minimize (10), we set a deterministic initial condition $x(0) = z$ and write the optimal cost v in the form

$$v(z) = z^T P z + 2 s_1^T z + s_0. \tag{12}$$

Invoking the standard results of LQG control [4], [8], [25], we have

$$2(A-\rho/2)P - B^2R^{-1}P^2 + \widehat{Q} = 0, \tag{13}$$

$$\rho s_1 = A s_1 - B^2 R^{-1} P s_1 + \gamma\eta_0(\gamma-1)\mathbf{1}_N, \tag{14}$$

$$\rho s_0 = -B^2R^{-1}s_1^T s_1 + D^2 \mathrm{Tr}(P) + N\gamma^2\eta_0^2, \tag{15}$$

which results in $P \in \mathbb{R}^{N\times N}$ of the form

$$P = \begin{pmatrix} p & q & \cdots & q \\ q & p & \cdots & q \\ \vdots & & \ddots & \vdots \\ q & q & \cdots & p \end{pmatrix}. \tag{16}$$

We consider two cases. Let I_N be the $N\times N$ identity matrix.

Case 1: $\gamma \neq 1$ so that $\lambda_1 > 0$. Then clearly $\widehat{Q} > 0$ and the pair $[\widehat{Q}^{\frac{1}{2}}, (A-\rho/2)I_N]$ is observable, so that (13) has a unique solution $P > 0$.

Case 2: $\gamma = 1$ so that $\lambda_1 = 0$. Then $[\widehat{Q}^{\frac{1}{2}}, (A-\rho/2)I_N]$ is not fully observable. By using an orthogonal transformation Ψ such that $\Psi^T\widehat{Q}\Psi = \mathrm{Diag}(\lambda_i) =: \Lambda_{\widehat{Q}}$, from (13) we obtain

$$2(A-\rho/2)\Psi^T P \Psi - B^2R^{-1}(\Psi^T P\Psi)^2 + \Lambda_{\widehat{Q}} = 0.$$

We restrict the entry of $\Psi^T P \Psi$ at the first row and the first column to be zero, corresponding to the unobservable state in the new coordinate system. Then we may find a unique $\Psi^T P\Psi \geq 0$ of rank $N-1$, and subsequently find $P \geq 0$ to (13).

For simplicity, below we analyze Case 1 in detail. Substituting P into (13) and denoting $\overline{a} = A - \rho/2$, $\overline{b} = B/\sqrt{R}$, we obtain the following equations

$$2\overline{a}p - \overline{b}^2[p^2 + (N-1)q^2] + \alpha = 0, \tag{17}$$

$$2\overline{a}q - \overline{b}^2[2pq + (N-2)q^2] + \beta = 0. \tag{18}$$

Under the condition $P > 0$, solving (17)-(18) yields

$$p_N = \frac{\overline{a} + \sqrt{\overline{a}^2 + \overline{b}^2}}{\overline{b}^2} + \frac{\sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2} - \sqrt{\overline{a}^2+\overline{b}^2}}{N\overline{b}^2}, \tag{19}$$

$$q_N = \frac{\sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2} - \sqrt{\overline{a}^2+\overline{b}^2}}{N\overline{b}^2}, \tag{20}$$

where the subscript in p_N and q_N indicates their dependence on N. Furthermore, we obtain

$$s_1 = \frac{\gamma\eta_0(\gamma-1)}{(\rho/2) + \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2}}\mathbf{1}_N =: s_{11}\mathbf{1}_N^T, \tag{21}$$

which does not depend on N.

Proposition 3: [11] If $\gamma \neq 1$, then the pair $[(A-\rho/2)I_N, BI_N]$ is controllable, the pair $[\widehat{Q}^{1/2}, (A-\rho/2)I_N]$ is observable, and (13) has a unique solution $P > 0$ given by (16) and (19)-(20). □

Let $u^* = (u_1^*, \ldots, u_N^*)$ be the optimal control law. Then

$$u_i^* = -BR^{-1}p_N x_i - BR^{-1}q_N \sum_{k\neq i} x_k - BR^{-1}s_{11}, \tag{22}$$

where s_{11} is defined in (21). Define

$$p_\infty = \lim_{N\to\infty} p_N = \frac{\overline{a} + \sqrt{\overline{a}^2 + \overline{b}^2}}{b^2}.$$

Proposition 4: Assume $\gamma \neq 1$ and (**A2**) holds with $Ex_i(0) = m_0$, $\mathrm{Var}(x_i(0)) = \sigma_0^2$ for all i. Then the optimal social cost per agent when $N \to \infty$ is

$$\lim_{N\to\infty}(1/N)\inf_{u\in\mathscr{U}_o} J_{\mathrm{soc}}^{(N)}(u) = \sigma_0^2 p_\infty + m_0^2\frac{\overline{a} + \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2}}{b^2} + 2m_0 s_{11} + (1/\rho)\left(D^2 p_\infty + \gamma^2\eta_0^2 - \overline{b}^2 s_{11}^2\right). \tag{23}$$

□

We continue to examine the limiting dynamics of the closed-loop system when $N\to\infty$. It is sufficient to consider a given agent. By (22), agent i has the closed-loop dynamics

$$dx_i = (A - \overline{b}^2 p_N)x_i dt - \overline{b}^2 q_N \sum_{j\neq i} x_j dt - \overline{b}^2 s_{11} dt + D dW_i. \tag{24}$$

Denote the ordinary differential equation (ODE)

$$d\overline{x}_c = \left[(\rho/2) - \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2}\right]\overline{x}_c dt - \overline{b}^2 s_{11} dt,$$

where $\overline{x}_c(0) = Ex_i(0) = m_0$. It is easy to obtain the explicit expression of $\overline{x}_c$. Denote

$$h(t) = \left[\sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2} - \sqrt{\overline{a}^2 + \overline{b}^2}\right]\overline{x}_c(t) + \overline{b}^2 s_{11}.$$

Proposition 5: Denote $x^{(N)} = (1/N)\sum_{i=1}^N x_i$ in (24). Assume $\gamma \neq 1$ and let y_i satisfy

$$dy_i = (A - \overline{b}^2 p_\infty) y_i dt - h dt + D dW_i, \tag{25}$$

where $y_i(0) = x_i(0)$. Let $\hat{\rho} \in (0, \rho]$ be fixed such that $(\rho - \hat{\rho})/2 - \sqrt{\overline{a}^2 + (1-\gamma)^2\overline{b}^2} < 0$. Then

$$\sup_{t\geq 0} e^{-\hat{\rho}t}\left\{E|x^{(N)}(t) - \overline{x}_c(t)|^2 + E|x_i(t) - y_i(t)|^2\right\} = O(1/N).$$

□

IV. The Social Certainty Equivalence Methodology and Decentralized Strategies

For integer $k \geq 1$ and real number $\delta > 0$, define $C_\delta([0,\infty),\mathbb{R}^k)$ consisting of all $f \in C([0,\infty),\mathbb{R}^k)$ such that $\sup_{t\geq 0}|f(t)|e^{-\delta' t} < \infty$ for some $\delta' < \delta$. The parameter δ' may change with f.

A. The Mean Field Approximation

For large N, it is plausible to approximate $\hat{x}_{-i}^{(N)} = (1/N)\sum_{j\neq i}\hat{x}_j$ in (8) by a deterministic function $\bar{x}$. As an approximation to Problem (P1), we construct the auxiliary optimal control problem:

$$\textbf{(P2)} \qquad dx_i = A(\theta_i)x_i dt + Bu_i dt + DdW_i, \tag{26}$$
$$J^*(u_i) = E\int_0^\infty e^{-\rho t}L^*(x_i,u_i,\bar{x})dt,$$

where J^* is to be minimized and

$$L^*(x_i,u_i,\bar{x}) = x_i^T Q x_i - 2(\Gamma\bar{x}+\eta)^T Q x_i - 2[(I-\Gamma)\bar{x}-\eta]^T Q\Gamma x_i + u_i^T R u_i \tag{27}$$

as an approximation of L in (8). To ensure that J^* is finite, we restrict that $\bar{x}\in C_{\rho/2}([0,\infty),\mathbb{R}^n)$.

For $\theta\in\Theta$, denote the ARE

$$\rho\Pi_\theta = \Pi_\theta A_\theta + A_\theta^T\Pi_\theta - \Pi_\theta BR^{-1}B^T\Pi_\theta + Q, \tag{28}$$

where $A_\theta := A(\theta)$. Under (**A4**), (28) has a unique solution $\Pi_\theta \geq 0$. Denote the ODE

$$\rho s_\theta = \frac{ds_\theta}{dt} + (A_\theta^T - \Pi_\theta BR^{-1}B^T)s_\theta - [(\Gamma^T Q + Q\Gamma - \Gamma^T Q\Gamma)\bar{x} + (I-\Gamma^T)Q\eta], \tag{29}$$

which does not have a pre-specified initial condition. In fact, if $\bar{x}\in C_{\rho/2}([0,\infty),\mathbb{R}^n)$, one may use the fact that $A_\theta - BR^{-1}B^T\Pi_\theta - (\rho/2)I$ is asymptotically stable to identify a unique initial condition $s_\theta(0)$ provided that $s_\theta(t)$ is required to be within $C_{\rho/2}([0,\infty),\mathbb{R}^n)$; see [8, Lemma A.2] for related detail.

Assume that $\bar{x}$ has been given. Following the method in [4], [10], [14], [8], we may use (**A4**) and Lemma 11 to show that if $s_{\theta_i}\in C_{\rho/2}([0,\infty),\mathbb{R}^n)$ satisfies (29) after setting $\theta=\theta_i$, the optimal control law $\hat{u}_i$ for Problem (P2) is

$$\hat{u}_i = -R^{-1}B^T(\Pi_{\theta_i}x_i + s_{\theta_i}). \tag{30}$$

The closed-loop dynamics take the form

$$dx_i = A_{\theta_i}x_i dt - BR^{-1}B^T(\Pi_{\theta_i}x_i + s_{\theta_i})dt + DdW_i, \tag{31}$$

where the initial condition is $x_i(0)$.

We construct the equation system with the dynamic parameter $\theta\in\Theta$

$$\rho s_\theta = \frac{ds_\theta}{dt} + (A_\theta^T - \Pi_\theta BR^{-1}B^T)s_\theta - [(\Gamma^T Q + Q\Gamma - \Gamma^T Q\Gamma)\bar{x} + (I-\Gamma^T)Q\eta], \tag{32}$$

$$\frac{d\bar{x}_\theta}{dt} = A_\theta\bar{x}_\theta - BR^{-1}B^T(\Pi_\theta\bar{x}_\theta + s_\theta), \tag{33}$$

$$\bar{x} = \int \bar{x}_\theta dF(\theta), \tag{34}$$

where $\bar{x}_\theta(0) = m_0$ due to (**A2**) for the initial mean of all agents, and s_θ is sought within $C_{\rho/2}([0,\infty),\mathbb{R}^n)$. The first two equations are based on (29) and (31), and (34) is based on interaction consistency, i.e., the mean field assumed at the beginning should be replicated when averaging the individual closed-loop dynamics of a large number of agents. This equation system will be referred to as the Social Certainty Equivalence (SCE) equation system.

In general, the existence and uniqueness analysis of a solution to (32)-(34) may be developed using fixed point methods as in [10], [13] for NCE equation systems. After a solution is obtained, the N agents in the social optimal control problem may determine their strategies by (30).

B. The Social Optimality Theorem

We state the assumption on the existence of a solution to the SCE equation system.

(**A5**) There exists a solution $(s_\theta,\bar{x}_\theta,\bar{x},\theta\in\Theta)$ to the SCE equation system (32)-(34) such that each component of $(s_\theta,\bar{x}_\theta,\bar{x})$, as a function of t, is within $C_{\rho/2}([0,\infty),\mathbb{R}^n)$ and such that both s_θ and $\bar{x}_\theta$ are continuous in θ for each fixed $t\in[0,\infty)$. $\diamondsuit$

Let $F^{(N)}$ be the empirical distribution specified by (**A1**). Define $\varepsilon_N \geq 0$ by

$$\varepsilon_N^2 = \int_0^\infty e^{-\rho t}\left|\int \bar{x}_\theta(t)dF^{(N)}(\theta) - \int \bar{x}_\theta(t)dF(\theta)\right|^2 dt.$$

Lemma 6: [12] Suppose that (**A1**)-(**A5**) hold. Let $\hat{x}_i$ be the closed-loop solution of agent i under the SCE based control law (30) and $\hat{x}^{(N)} = (1/N)\sum_{i=1}^N\hat{x}_i$. Then

$$E\int_0^\infty e^{-\rho t}|\hat{x}^{(N)}(t) - \bar{x}(t)|^2 dt \leq C(1/N + \varepsilon_N^2),$$

where $\lim_{N\to\infty}\varepsilon_N = 0$. $\square$

The asymptotic performance of the SCE based strategies is characterized by the central result below.

Theorem 7: [12] Assume (i) (**A1**)-(**A3**), (**A4**)-(i) and (**A5**) hold; (ii) $Q>0$ and $I-\Gamma$ is nonsingular. Then the set of SCE based control laws $\{\hat{u}_i = -R^{-1}B^T(\Pi_{\theta_i}\hat{x}_i + s_{\theta_i}),\ 1\leq i\leq N\}$ has asymptotic social optimality, i.e., for $\hat{u} = (\hat{u}_1,\ldots,\hat{u}_N)$,

$$|(1/N)J_{\rm soc}^{(N)}(\hat{u}) - \inf_{u\in\mathscr{U}_o}(1/N)J_{\rm soc}^{(N)}(u)| = O(1/\sqrt{N} + \varepsilon_N),$$

where $\mathscr{U}_o$ is defined in Section III-A as a set of centralized information based controls. $\square$

Note that $Q>0$ implies (**A4**)-(ii). Hence it is ensured by (**A4**)-(i) and (ii) that (28) has a unique solution $\Pi_\theta\geq 0$ and that $A(\theta) - BR^{-1}B^T\Pi_\theta - (\rho/2)I$ is asymptotically stable.

Note that for the scalar model with uniform agents and the parametrization (9), the requirement of a nonsingular $I-\Gamma$ reduces to $\gamma\neq 1$. The invertibility of $I-\Gamma$ is used for prior integral estimates of the state process provided that the social cost is finite [12].

C. Explicit Solutions with Uniform Agents

For uniform agents (i.e., $\theta_i \equiv \theta$ and $A(\theta_i) \equiv A$), the SCE equation system (32)-(34) becomes

$$\frac{d\bar{x}}{dt} = (A - BR^{-1}B^T\Pi)\bar{x} - BR^{-1}B^T s, \qquad (35)$$
$$\frac{ds}{dt} = Q_\Gamma \bar{x} + (\rho I - A^T + \Pi BR^{-1}B^T)s + \eta_\Gamma, \qquad (36)$$

where we denote $\Pi_\theta = \Pi$, $s_\theta = s$ and $\bar{x}_\theta = \bar{x}$ by omitting the subscript and

$$Q_\Gamma = \Gamma^T Q + Q\Gamma - \Gamma^T Q\Gamma, \quad \eta_\Gamma = (I - \Gamma^T)Q\eta. \qquad (37)$$

A solution $(s, \bar{x})$ of (35)-(36) is said to be within $C_{\rho/2}([0,\infty), \mathbb{R}^{2n})$ if both s and $\bar{x}$ are within $C_{\rho/2}([0,\infty), \mathbb{R}^n)$. We have the following existence and uniqueness theorem.

Theorem 8: [12] Suppose that (**A4**) holds for $A(\theta) = A$ and that $Q_\Gamma = \Gamma^T Q + Q\Gamma - \Gamma^T Q\Gamma \geq 0$. Then the equation system (35)-(36) has a unique solution $(s, \bar{x})$ within $C_{\rho/2}([0,\infty), \mathbb{R}^{2n})$. □

D. Connection with the NCE Equation System

To compare with our past work on Problem II, we review the NCE approach for the game problem when agent i is associated with cost J_i (see e.g., [10]). To obtain decentralized strategies, this approach proceeds as follows.

First, the representative agent i approximates the coupling term $\Gamma x^{(N)} + \eta$ in (2) by a deterministic function $\Gamma\bar{x} + \eta$ and solves an optimal tracking problem. Next, the mean trajectory Ex_i for the closed-loop of agent i is determined by an ODE. Finally, the state average of all the individual agents shall replicate $\bar{x}$ initially assumed. This procedure leads to the NCE equation system

$$\rho s_\theta^\dagger = \frac{ds_\theta^\dagger}{dt} + (A_\theta^T - \Pi_\theta BR^{-1}B^T)s_\theta^\dagger - Q(\Gamma\bar{x}^\dagger + \eta), \qquad (38)$$
$$\frac{d\bar{x}_\theta^\dagger}{dt} = A_\theta \bar{x}_\theta^\dagger - BR^{-1}B^T(\Pi_\theta \bar{x}_\theta^\dagger + s_\theta^\dagger), \qquad \bar{x}_\theta^\dagger(0) = m_0, \quad (39)$$
$$\bar{x}^\dagger = \int \bar{x}_\theta^\dagger dF(\theta). \qquad (40)$$

The superscript in $(s_\theta^\dagger, \bar{x}_\theta^\dagger, \bar{x}^\dagger)$ distinguishes the solution from that of the SCE equation system. The set of strategies

$$u_i = -BR^{-1}(\Pi_{\theta_i} x_i + s_{\theta_i}^\dagger), \qquad 1 \leq i \leq N \qquad (41)$$

is an ε-Nash equilibrium [10]. The NCE and SCE equation systems differ by a different equation for $s_\theta^\dagger$.

V. The Scalar Model with Uniform Agents

Recall that for the scalar model, we set $Q = 1$ and assume $B \neq 0$, and that Γ and η are parametrized according to (9). Let $\Pi > 0$ be the solution to the ARE

$$\rho\Pi = 2A\Pi - B^2R^{-1}\Pi^2 + 1.$$

Let $\beta_1 = -A + B^2R^{-1}\Pi$ and $\beta_2 = -A + B^2R^{-1}\Pi + \rho$. Following the notation in Section III-B, denote $\bar{a} = A - \rho/2$, $\bar{b} = B/\sqrt{R}$. We have $\beta_1 = -(\rho/2) + \sqrt{\bar{a}^2 + \bar{b}^2}$ and $\beta_2 = (\rho/2) + \sqrt{\bar{a}^2 + \bar{b}^2}$.

A. Comparison of Solvability of the Two Equation Systems

Now, the SCE equation system (32)-(34) reduces to

$$\frac{ds}{dt} = \beta_2 s + (2\gamma - \gamma^2)\bar{x} + (1-\gamma)\gamma\eta_0, \qquad (42)$$
$$\frac{d\bar{x}}{dt} = -\bar{b}^2 s - \beta_1 \bar{x}. \qquad (43)$$

Note that in Theorem 8 we may calculate the explicit solution of the SCE equation system under the assumption $Q_\Gamma \geq 0$ (see [12] for detail), which translates to $\gamma \in [0,2]$ for the scalar model. The calculation here does not need that assumption. The related analysis will reveal a fundamental difference between the SCE and NCE equation systems.

The NCE equation system (38)-(40) reduces to

$$\frac{ds^\dagger}{dt} = \beta_2 s^\dagger + \gamma\bar{x}^\dagger + \gamma\eta_0, \qquad (44)$$
$$\frac{d\bar{x}^\dagger}{dt} = -\bar{b}^2 s^\dagger - \beta_1 \bar{x}^\dagger. \qquad (45)$$

For further analysis of (44)-(45), we introduce

$$\beta_2 s^\dagger(\infty) + \gamma\bar{x}^\dagger(\infty) = -\gamma\eta_0, \qquad (46)$$
$$\bar{b}^2 s^\dagger(\infty) + \beta_1 \bar{x}^\dagger(\infty) = 0. \qquad (47)$$

The notation $s^\dagger(\infty)$, $\bar{x}^\dagger(\infty)$ is only for constructing the algebraic equations. It does not necessarily mean $\bar{x}^\dagger(\infty) = \lim_{t\to\infty} \bar{x}^\dagger(t)$. Denote

$$\Delta = \beta_1\beta_2 - \bar{b}^2(2\gamma - \gamma^2) = \bar{a}^2 + (1-\gamma)^2\bar{b}^2 - \rho^2/4,$$
$$\Delta^\dagger = \beta_1\beta_2 - \bar{b}^2\gamma = \bar{a}^2 + (1-\gamma)\bar{b}^2 - \rho^2/4.$$

Note that $\Delta^\dagger = 0$ if $\gamma = \gamma_1^\dagger = 1 + \bar{a}^2/\bar{b}^2 - \rho^2/(4\bar{b}^2)$. Further denote $\gamma_2^\dagger = 1 + \bar{a}^2/\bar{b}^2$.

We study the solvability of the two equation systems (42)-(43) and (44)-(45) in terms of the interaction parameter γ such that each function is within $C_{\rho/2}([0,\infty), \mathbb{R})$.

Theorem 9: [12] Suppose $B \neq 0$.

i) If $\gamma \neq 1$ and $\Delta \neq 0$, (42)-(43) has a unique solution $(s, \bar{x}) \in C_{\rho/2}([0,\infty), \mathbb{R}^2)$ for any $\bar{x}(0)$.

ii) If $\gamma \neq \gamma_1^\dagger$, (46)-(47) has a unique solution $(s^\dagger(\infty), \bar{x}^\dagger(\infty))$. If

$$\gamma \in (-\infty,\ \gamma_1^\dagger) \cup (\gamma_1^\dagger,\ \gamma_2^\dagger), \qquad (48)$$

(44)-(45) has a unique solution $(s, \bar{x}) \in C_{\rho/2}([0,\infty), \mathbb{R}^2)$ for any $\bar{x}(0)$. If $\gamma \geq \gamma_2^\dagger$, (44)-(45) has a unique solution $(s^\dagger(t), \bar{x}^\dagger(t)) \equiv (s^\dagger(\infty), \bar{x}^\dagger(\infty))$ within $C_{\rho/2}([0,\infty), \mathbb{R}^2)$ when $\bar{x}^\dagger(0)$ coincides with $\bar{x}^\dagger(\infty)$, and otherwise there is no solution in such a function class. □

Remark: To simplify the calculation, let $x(0) = 0$ and $\eta_0 \neq 0$. We may use the solution of the NCE equation system to explicitly compute the individual cost in the population limit, and further show a cost blow-up effect. Namely, when γ approaches $\gamma_1^\dagger$ from both sides or $\gamma_2^\dagger$ from the left side, the associated individual cost approaches infinity. However, for the socially optimal solution, there is no blow-up effect for the asymptotic average social optimum $\lim_{N\to\infty}(1/N)\inf_u J_{\mathrm{soc}}^{(N)}|_{x(0)=0}$ as γ approaches $\gamma_1^\dagger$ or $\gamma_2^\dagger$. ◇

Remark: Note that if $B \neq 0$, $\gamma \neq 1$ and $\Delta \neq 0$, we may explicitly compute s in the SCE based control law $u_i = -R^{-1}B(\Pi x_i + s)$ for agent i, which leads to

$$dx_i = (A - \overline{b}^2\Pi)x_i dt - \overline{b}^2 s dt + D dW_i. \qquad (49)$$

In fact (49) coincides with the limiting dynamics (25) of the centralized optimal control problem since we may verify that $\Pi = p_\infty$ and $\overline{b}^2 s = h$ and (49) and (25) have the same initial condition. ♢

VI. The Asymptotic Social Optimum

Now we give a closed form expression of the asymptotic social optimum when $N \to \infty$ in terms of the solution of the SCE equation system. Denote the ODE

$$\rho q_\theta = \frac{dq_\theta}{dt} - s_\theta^T BR^{-1}B^T s_\theta + \mathrm{Tr}(D^T \Pi_\theta D), \qquad \theta \in \Theta.$$

By Lemma 11 there exists a unique $q_\theta(0)$ such that the resulting solution $q_\theta \in C_\rho([0,\infty),\mathbb{R})$.

Theorem 10: [12] Assume that **(A1)**-**(A5)** hold and that $\{x_i, 1 \le i \le N\}$ have the same initial mean m_0 and initial covariance $\Sigma_0 = \mathrm{Cov}(x_i(0))$. The asymptotic average social optimum is given by

$$\begin{aligned}&\lim_{N\to\infty} \inf_u (1/N) J_{\mathrm{soc}}^{(N)} \\ &= m_0^T \left[\int \Pi_\theta dF(\theta)\right] m_0 + 2m_0^T \int s_\theta(0) dF(\theta) \\ &\quad + \mathrm{Tr}\left[\Sigma_0 \int \Pi_\theta dF(\theta)\right] + \int q_\theta(0) dF(\theta) \\ &\quad + \int_0^\infty e^{-\rho t} \overline{x}^T (Q\Gamma + \Gamma^T Q - \Gamma^T Q \Gamma)\overline{x} dt + (1/\rho)\eta^T Q \eta.\end{aligned}$$

□

Remark: If we specialize Theorem 10 to uniform agents with scalar states and if $B \neq 0$, $\gamma \neq 1$, $\Delta \neq 0$, the asymptotic average social optimum may be evaluated using the expression of $(s, \overline{x})$ in Theorem 9. By lengthy but elementary calculation, we may further verify that this expression coincides with (23). ♢

VII. Conclusions

This paper develops the social certainty equivalence approach for asymptotically achieving the social optima in mean field decision models. In this solution scheme, each agent only needs to know the empirical distribution of the dynamic parameters across the population to solve an ODE system off-line and then uses its own state to construct a feedback strategy.

Appendix A: The Optimal Control Lemma

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge 0}, P)$ be an underlying filtration. Consider the controlled SDE

$$dx(t) = Ax(t)dt + Bu(t)dt + f(t)dt + DdW(t), \qquad t \ge 0, \qquad (50)$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^{n_1}$, $f \in C_{\rho/2}([0,\infty),\mathbb{R}^n)$, and $W(t)$ is an n_2 dimensional standard Brownian motion adapted to $\mathscr{F}_t$. The initial condition $x(0)$ is independent of $W(t)$ and $E|x(0)|^2 < \infty$. The admissible control set $\mathscr{U}$ consists of all controls $u(\cdot)$ adapted to $\mathscr{F}_t$ with $\int_0^\infty e^{-\rho t}|u(t)|^2 dt < \infty$. For $u(\cdot) \in \mathscr{U}$, let the cost function be given by

$$J(u(\cdot)) = E\int_0^\infty e^{-\rho t}\left[x^T(t)Qx(t) - 2g^T(t)x(t) + u^T(t)Ru(t)\right]dt, \qquad (51)$$

where $\rho > 0$, $Q \ge 0$, $R > 0$ and $g \in C_{\rho/2}([0,\infty),\mathbb{R}^n)$. The matrices A, B, D, Q and R have compatible dimensions. Denote the ARE

$$\rho\Pi = \Pi A + A^T\Pi - \Pi BR^{-1}B^T\Pi + Q. \qquad (52)$$

Denote the ODEs

$$\rho s = \frac{ds}{dt} + (A^T - \Pi BR^{-1}B^T)s + \Pi f - g, \qquad t \ge 0, \qquad (53)$$

$$\rho q = \frac{dq}{dt} - s^T BR^{-1}B^T s + 2f^T s + \mathrm{Tr}(D^T\Pi D), \qquad (54)$$

where s and q are to be sought within $C_{\rho/2}([0,\infty),\mathbb{R}^n)$ and $C_\rho([0,\infty),\mathbb{R}^n)$, respectively. The initial conditions $s(0)$ and $q(0)$ are not pre-specified.

Lemma 11: For the optimal control problem (50)-(51), assume (i) the pair $[Q^{1/2}, A-(\rho/2)I]$ is detectable and (52) has a solution $\Pi \ge 0$ such that $A - BR^{-1}B^T\Pi - (\rho/2)I$ is asymptotically stable, and (ii) both f and g are in $C_{\rho/2}([0,\infty),\mathbb{R}^n)$. Then we have

(a) there exists a unique solution $s \in C_{\rho/2}([0,\infty),\mathbb{R}^n)$ to (53);

(b) the optimal control law is given by $\hat{u}(t) = -R^{-1}B^T[\Pi x(t) + s(t)]$;

(c) there exists a unique solution $q \in C_\rho([0,\infty),\mathbb{R})$ to (54). The optimal cost is given by

$$\inf_{u\in\mathscr{U}} J(u) = J(\hat{u}) = E[x^T(0)\Pi x(0)] + 2s^T(0)Ex(0) + q(0).$$

Proof: We may show part (a) using the method in [8, Lemma A.2]. We prove part (b) by first obtaining a prior integral estimate of x (see (56)) and then using a completion of squares technique. Compared with [8], the cost integrand in (51) does not necessarily allow rewriting the term $x^TQx - 2g^Tx$ in the form $(x-g_1)^TQ(x-g_1) + h_1$ for some functions g_1 and h_1. For $u \in \mathscr{U}$, we show that the prior upper bound

$$J(u) \le C_0 \qquad (55)$$

for some constant C_0 implies

$$E\int_0^\infty e^{-\rho t}|x|^2 dt < \infty, \qquad (56)$$

where x is associated with u. We have $2|g^Tx| \le 2\sqrt{g^Tg}\sqrt{x^Tx} \le (1/\varepsilon)g^Tg + \varepsilon x^Tx$ for any $\varepsilon > 0$. Hence (55) leads to $E\int_0^\infty e^{-\rho t}[x^TQx - (1/\varepsilon)g^Tg - \varepsilon x^Tx]dt \le C_0$, which implies

$$E\int_0^\infty e^{-\rho t}(x^TQx - \varepsilon x^Tx)dt \le C_1. \qquad (57)$$

If necessary, we may apply a nonsingular linear transformation. Here without loss of generality we simply assume $A = \mathrm{Diag}[A_{11}, A_{22}]$, where all eigenvalues of A_{11} (resp., A_{22}) have a real part greater than or equal to (resp., less than)

$\rho/2$. Write $x=[x_1^T,x_2^T]^T$, where x_1 and x_2 corresponds to the sub-matrices A_{11} and A_{22}, respectively, in the dynamics. We write $Q^{1/2}=[M_1,M_2]$ so that $x^TQx=|M_1x_1+M_2x_2|^2$, where $[M_1,A_{11}]$ is observable due to the detectability of $[Q^{1/2},A-(\rho/2)I]$. By [8, Lemma A.1] and $E\int_0^\infty e^{-\rho t}u^TRudt\le C_0$, it follows that $E\int_0^\infty e^{-\rho t}|x_2|^2dt<\infty$, which combined with (57) implies $E\int_0^\infty e^{-\rho t}(x_1^TM_1^TM_1x_1-\varepsilon|x_1|^2)dt<\infty$. By the observability of $[M_1,A_{11}]$, we may further show that there exist fixed $c_1>0$ and $c_2>0$, both independent of ε, such that $E\int_0^\infty e^{-\rho t}x_1^TM_1^TM_1x_1dt\ge c_1E\int_0^\infty e^{-\rho t}|x_1|^2dt-c_2$ (see [8] for a similar argument). Hence $E\int_0^\infty e^{-\rho t}(c_1-\varepsilon)E|x_1|^2dt<\infty$. By taking a sufficiently small ε such that $c_1-\varepsilon>0$, (56) follows. The rest part of the proof of part (b) is similar to [8, Lemma A.2] and is omitted.

We prove (c) as follows. Let the initial condition of (54) be $q(0)$ to give

$$q(t)=e^{\rho t}q(0)+e^{\rho t}\int_0^t e^{-\rho\tau}(s^TBR^{-1}B^Ts-2f^Ts-\mathrm{Tr}(D^T\Pi D))d\tau.$$

We can show that $q\in C_\rho([0,\infty),\mathbb{R})$ if and only if

$$q(0)=\int_0^\infty e^{-\rho\tau}(2f^Ts+\mathrm{Tr}(D^T\Pi D)-s^TBR^{-1}B^Ts)d\tau =(1/\rho)\mathrm{Tr}(D^T\Pi D)+\int_0^\infty e^{-\rho\tau}(2f^Ts-s^TBR^{-1}B^Ts)d\tau, \quad (58)$$

where the integral converges. Next, we apply Ito's formula to $e^{-\rho t}[x(t)^T\Pi x(t)+2s(t)^Tx(t)+q(t)]$ to obtain the expression of $J(\hat{u})$, where x is the closed-loop solution under $\hat{u}$. □

Remark: If f and g are constant vectors, we may obtain the unique solutions to (53) and (54) with the required growth conditions as two constants s and q, satisfying

$$\rho s=(A^T-\Pi BR^{-1}B^T)s+\Pi f-g,$$
$$\rho q=2f^Ts-s^TBR^{-1}B^Ts+\mathrm{Tr}(D^T\Pi D). \quad \diamond$$

REFERENCES

[1] S. Adlakha, R. Johari, G. Weintraub, and A. Goldsmith. Oblivious equilibrium for large-scale stochastic games with unbounded costs. *Proc. IEEE CDC 2008*, Cancun, Mexico, pp. 5531-5538, Dec. 2008.

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

[3] C. T. Bauch and D. J. D. Earn. Vaccination and the theory of games. *Proc. Natl. Acad. Sci., U.S.A.*, vol. 101, pp. 13391-13394, Sept. 2004.

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

[5] R. Breban, R. Vardavas, and S. Blower. Mean-field analysis of an inductive reasoning game: application to influenza vaccination. *Physical Review E*, vol. 76, 2007. DOI: 10.1103/PhysRevE.76.031127.

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

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

[8] M. Huang. Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM J. Control Optim.*, vol. 48, no. 5, pp. 3318-3353, 2010.

[9] 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 CDC*, Maui, HI, pp. 98-103, Dec. 2003.

[10] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ε-Nash equilibria. *IEEE Trans Autom. Contr.*, vol. 52, pp. 1560-1571, Sep. 2007.

[11] M. Huang, P. E. Caines, and R. P. Malhamé. Social optima in mean field LQG control: centralized and decentralized strategies. *Proc. 47th Allerton Conference*, UIUC, IL, pp. 322-329, Sept. 2009.

[12] M. Huang, P. E. Caines, and R. P. Malhamé. Social optima in mean field LQG control: centralized and decentralized strategies. Under review for journal publication, submitted Mar. 2010.

[13] M. Huang, P. E. Caines, and R. P. Malhamé. The NCE (mean field) principle with locality dependent cost interactions. Submitted to *IEEE Trans. Automatic Control*, Jan. 2009 (provisionally accepted).

[14] M. Huang, R. P. Malhamé, and P. E. Caines. Nash equilibria for large-population linear stochastic systems of weakly coupled agents. In *Analysis, Contr. Optim. Complex Dyn. Syst.*, E. K. Boukas and R. P. Malhamé eds., pp. 215-252, Springer, New York, 2005.

[15] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inform. Systems*, vol. 6, no. 3, pp. 221-252, 2006.

[16] Y.-C. Ho. Team decision theory and information structures. *Proc. IEEE*, vol. 86, pp. 644-654, 1980.

[17] M. A. 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., Elsevier Science, North-Holland, Amsterdam, 2002, pp. 1761-1808.

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

[19] A. Lazarte. *Large population stochastic dynamic games with nonuniform agents and localized coupling*, Master's thesis, Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, 2009.

[20] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I – Le cas stationnaire, *C. R. Acad. Sci. Paris, Ser. I*, vol. 343, pp. 619-625, 2006.

[21] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II – Horizon fini et contrôle optimal, *C. R. Acad. Sci. Paris, Ser. I*, vol. 343, pp. 679-684, 2006.

[22] J.-M. Lasry and P.-L. Lions. Mean field games. *Japan. J. Math.*, vol. 2, no. 1, pp. 229-260, 2007.

[23] T. Li and J. F. Zhang. Asymptotically optimal decentralized control for large population stochastic multiagent systems. *IEEE Trans. Auto. Control*, vol. 53, no. 7, pp. 1643-1660, 2008.

[24] G. Owen. *Game Theory*, 3rd ed., Academic Press, San Diego, 1995.

[25] E. D. Sontag. *Mathematical Control Theory*, Springer-Verlag, New York, 1990.

[26] H. Tembine, J.-V. Le Boudec, R. El-Azouzi, and E. Altman. Mean field asymptotics of Markov decision evolutionary games and teams. *Proc. International Conference on Game Theory for Networks*, Istanbul, Turkey, pp. 140-150, May 2009.

[27] P. R. de Waal and J. H. van Schuppen. A class of team problems with discrete action spaces: optimality conditions based on multimodularity. *SIAM J. Control Optim.*, vol. 38, no. 3, pp. 875-892, 2000.

[28] G. Y. Weintraub, C. L. Benkard, and B. Van Roy. Oblivious equilibrium: a mean field approximation for large-scale dynamic games, *Advances in Neural Information Processing Systems*, MIT Press, 2005.

[29] G. Y. Weintraub, C. L. Benkard, and B. Van Roy. Markov perfect industry dynamics with many firms. *Econometrica*, vol. 76, no. 6, pp. 1375-1411, Nov. 2008.

[30] W. M. Wonham. *Linear Multivariable Control: A Geometric Approach*, 2nd ed., Springer-Verlag, New York, 1979.

[31] D. W. K. Yeung and L. A. Petrosyan. *Cooperative Stochastic Differential Games*, Springer, New York, 2006.

[32] H. Ying, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag. Synchronization of coupled oscillators is a game. *Proc. American Control Conference*, Baltimore, MD, June 2010.