

# Social Optima in Mean Field LQG Control: Centralized and Decentralized Strategies

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

**Abstract**—We study a class of linear-quadratic-Gaussian (LQG) control problems with  $N$  decision makers, where the basic objective is to minimize a social cost as the sum of  $N$  individual costs containing mean field coupling. The exact socially optimal solution (determining a particular Pareto optimum) requires centralized information for each agent and has high implementational complexity. As an alternative we subsequently exploit a mean field structure in the centralized optimal control problem to develop decentralized cooperative optimization so that each agent only uses its own state and a function which may be computed offline; the resulting set of strategies asymptotically achieves the social optimum as  $N \rightarrow \infty$ . A key feature in this scheme 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 decentralized cooperative solution and the so-called Nash Certainty Equivalence based solution presented in previous work on mean field LQG games.

**Index Terms**—Decentralized control, linear-quadratic-Gaussian (LQG) control, mean field models, Nash equilibria, Pareto optima, social optima.

## I. INTRODUCTION

**I**N decision problems with a large number of agents, mean field models have attracted extensive attention due to their significance in many domains [3], [5], [8], [9], [12], [19], [20], [23], [29], [30]. A distinctive feature of such models is the interaction between any given agent and the average effect of the overall population. In the search for decentralized optimization paradigms, game theoretic solutions have been successfully developed by different researchers [12], [13], [21]–[23], [25], [30], [31]; along this line, decentralized solutions may be obtained by identifying a consistency relationship between the individual and the mass behavior such that in the population limit each individual optimally responds to the mass effect and these individual strategies also collectively replicate the mass effect ini-

tially presumed [11]–[15]. This work rests upon the fact that under reasonable conditions one may show the existence of a mass effect satisfying such a fixed point property and prove 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 [13], [16]. Closely related mean field approximation approaches were developed in [30], [31] using the notion of oblivious equilibria (OE) for game models on industry dynamics, and OE with unbounded costs were analyzed in [1]. The works [21]–[23] 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 [25], the interaction consistency based approach was applied to LQG games with long run average costs. A significant application and development of the NCE principle appears in [34] where a game theoretic framework was proposed for the control of a large number of coupled nonlinear oscillators; the mean field approximation approach gives decentralized strategies, and phase transition phenomena are observed in the closed-loop system.

When developing the above game theoretic solution framework, the starting point is that these agents are individually incentive driven and noncooperative. In this paper, within the mean field modeling we study a different situation where the agents are cooperative and seek socially optimal decisions. The notion of social optima provides an appropriate solution concept in decision problems with multiple agents who are willing to cooperate, and Pareto optimality is a useful necessary condition of social optimality [2], [26]. In various network optimization problems involving an aggregate cost as the sum of individual costs, social optima have been widely studied, either for their intrinsic interest or as a benchmark against which to measure the efficiency loss of Nash equilibria (see, e.g., [17], [18], [27]), but in general only static models have been considered. 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. We consider both 1) centralized strategies where each agent may use the state information of all agents and 2) decentralized strategies where each agent only uses local information. To find the solution with centralized strategies, one may solve a standard LQG control problem and each agent is required to know the states of all the agents. By contrast, the solution with decentralized strategies, if computable, has much lower informational requirements. For this reason, our analysis focusses on decentralized strategies. A related numerical comparison of the optimized costs between the socially optimal solution and the NCE based game theoretic solution was provided in [24], where each agent assigns nonuniform cost coupling weights across the population. For stochastic differential games, cooperation issues

Manuscript received March 18, 2010; revised September 23, 2010; accepted January 02, 2012. Date of publication January 09, 2012; date of current version June 22, 2012. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). This paper was presented in a preliminary form at the 2009 Allerton Conference and the 2010 MTNS Symposium. Recommended by Associate Editor H. Zhang.

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

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

R. P. Malhamé is with the Department of Electrical Engineering, École Polytechnique de Montréal, Montreal, QC H3C 3A7, Canada, and also with the Group for Research In Decision Analysis (GERAD), Montreal, QC H3T 2A7, Canada (e-mail: roland.malhamé@polymtl.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TAC.2012.2183439

were addressed in [33] by extending concepts such as coalition and Shapley value to dynamic models.

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 [6], [10]. Our social optimization problem with decentralized information may be viewed as a mean field generalization of team problems where each agent has *a 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 next apply mean field approximations.

We mention that a mean field Markov team decision problem with discrete state and action spaces was studied in [29]. The authors focus on long run behaviors by considering stationary policies. Since a discounted cost and general initial conditions are used, the restriction to stationary policies may result in optimality losses with respect to general policies. By the compact parametrization of stationary policies, the problem in [29] is viewed as an abstract static optimization problem on a compact space. This approach does not give an explicit characterization of the minimizing policy. Our paper differs from the above work since it takes into account the impact of the initial states and the resulting transient behavior. Our approach is to take consistent mean field approximations to decompose the original social optimization problem into (equivalent) local optimal control problems.

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 develops the SCE methodology for obtaining decentralized strategies and proves the asymptotic optimality theorem. Section V provides an in-depth analysis of the scalar case, where a comparison with the NCE based solution is also presented. Section VI presents the explicit calculation of the asymptotic average social optimum. Section VII compares the costs of different solutions, and Section VIII concludes the paper. Finally, in this Introduction we set a convention about notation. Throughout the paper we use  $C$  to denote a generic constant not depending on  $(N, i)$ , the population size and agent index, and its value may change from place to place. For a positive semi-definite matrix  $M$ , the quadratic form  $z^T M z$  is sometimes written as  $|z|_M^2$  to make the expression more compact.

## II. 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)

$$dx_i = A(\theta_i)x_i dt + Bu_i dt + DdW_i, \quad t \geq 0. \quad (1)$$

The underlying filtered probability space is  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ , where  $(\mathcal{F}_t)_{t \geq 0}$  is a collection of non-decreasing  $\sigma$ -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  $E|x_i(0)|^2 < \infty$ . The noise processes  $\{W_i, 1 \leq i \leq N\}$  are  $n_2$ -dimensional independent standard Brownian motions adapted to  $\mathcal{F}_t$ , which are also independent of  $\{x_i(0), 1 \leq i \leq N\}$ . The constant matrices  $A(\cdot)$ ,  $B$  and  $D$  have compatible dimensions. Here,  $\theta_i$  is a dynamic parameter to model a population of nonuniform agents. For notational brevity the time argument of a process ( $x_i, u_i$ , etc.) is often suppressed when its value 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

$$J_i(u(\cdot)) = E \int_0^\infty e^{-\rho t} \left[ |x_i - \Phi(x^{(N)})|_Q^2 + u_i^T R u_i \right] dt \quad (2)$$

where  $\Phi(x^{(N)}) = \Gamma x^{(N)} + \eta$  and  $x^{(N)} := (1/N) \sum_{i=1}^N x_i$  is called the mean field term. The constant matrices or vector  $\Gamma, Q \geq 0, R > 0$  and  $\eta$  have compatible dimensions. We only take  $A(\cdot)$  to be dependent on  $\theta_i$  so that in the subsequent analysis we may keep the notation relatively light. When other constant matrices and vectors in (1)–(2) also depend on  $\theta_i$ , the methodology of this paper may be easily applied. 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 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 of 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  $N \rightarrow \infty$ , 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 all  $\theta_i$ 's are in a compact set  $\Theta \subset \mathbb{R}^\kappa$ . Define the empirical distribution function

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

where  $\theta \in \mathbb{R}^\kappa$  and  $\theta_i \leq \theta$  means that the inequality holds componentwise for the two vectors. We use the convention:  $1_{\{\theta_i \leq \theta\}} = 1$  if  $\theta_i \leq \theta$  holds, and  $1_{\{\theta_i \leq \theta\}} = 0$  otherwise. 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  $\mathbb{R}^\kappa$ ,  $\lim_{N \rightarrow \infty} \int \varphi(\theta) dF_N(\theta) = \int \varphi(\theta) dF(\theta)$ .  $\diamond$

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

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

**(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.  $\diamond$

The reader is referred to [13] for discussions on “randomized” generation of dynamic parameters from an underlying distribution so that the resulting empirical distributions converge weakly. 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  $\Theta$ , denoted as  $\pi^{(N)}$ , and (A1) reduces to the convergence of  $\pi^{(N)}$  to a limit  $\pi$ . In addition, (A3) is trivially true under the discrete topology of  $\Theta$  where an open set is the union of singletons or is the null set. Thus (A3) becomes redundant.

For simplicity, in (A2) it is assumed that all agents have the same initial mean  $m_0$ . 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 [13]).

### B. Two Solutions Based on Different Information Patterns

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

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

2) *Problem I-B:* Find a social solution  $(u_1, \dots, u_N)$  with decentralized information (SSDI), where each  $u_i$  is a function of  $(t, x_i)$ . Note that when restricted to decentralized information, one cannot in general attain the same cost as in Problem I-A. Instead, a set of decentralized strategies  $\{u_i, 1 \leq i \leq N\}$  is sought such that the optimality loss with respect to Problem I-A in minimizing  $J_{\text{soc}}^{(N)}$  tends to zero when  $N \rightarrow \infty$ .

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

3) *Problem II:* Find a competitive solution  $(u_1, \dots, u_N)$  with decentralized information (CSDI), where agent  $i$  is associated with cost  $J_i$  and the objective is to obtain a set of  $\varepsilon$ -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 [13].

To facilitate further analysis, denote  $\mathcal{F}_t^0 = \sigma(x(0), W_1(s), \dots, W_N(s), s \leq t)$  for  $t \geq 0$ , which is the  $\sigma$ -algebra generated by  $x(0)$  and the Brownian motions up to time  $t$ . Denote the control set  $\mathcal{U}_o := \{(u_1, \dots, u_N) | u_i(t, \omega) \text{ is adapted to } \mathcal{F}_t^0, \forall i\}$ , where  $\omega \in \Omega$  explicitly indicates the dependence of  $u_i$  on the sample. Each  $u = (u_1, \dots, u_N) \in \mathcal{U}_o$  is called an  $\mathcal{F}_t^0$ -adapted control and may be viewed as a functional of  $x(0)$  and the Brownian motions without being directly related to the state process  $x(t)$ . Let  $\mathcal{U}_{oi} := \{u_i | u_i(t, \omega) \text{ is adapted to } \mathcal{F}_t^0\}$ . In the subsequent exposition, the distinction between  $\mathcal{F}_t^0$ -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  $u(t, x) = (u_1(t, x), \dots, u_N(t, x))$ , 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  $\mathcal{U}_o$ . 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; see [4], [11], [32] for details.

## III. CENTRALIZED SOLUTION

For Problem I-A, we first consider a scalar model with uniform agents (i.e., agents having the same dynamic parameter), and develop the asymptotic analysis of the optimal cost and closed-loop dynamics, which will be used later to compare with the SCE approach. Problem I-A of the general case, in principle, may be treated as a standard LQG control problem and the optimal control law  $(\hat{u}_1, \dots, \hat{u}_N)$  may be determined from a high dimensional algebraic Riccati equation (ARE) if the standard stabilizability and detectability conditions are satisfied [32]. The calculation of the optimal control law is difficult due to dimensionality when  $N$  is very large. Section III-B provides an optimality interpretation of a fixed component in  $(\hat{u}_1, \dots, \hat{u}_N)$ , which will motivate the mean field approximation scheme. This optimality interpretation is similar to the person-by-person optimality characterization of team decision problems [10].

### A. 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  $a$ . Without loss of generality, we set  $Q = 1$  in (2). We write  $B = b$  and  $R = r$ . To avoid triviality, suppose  $b \neq 0$ . We introduce the parametrization

$$\Gamma = \gamma, \quad \eta = \gamma\eta_0. \quad (4)$$

So  $\Gamma$  and  $\eta$  are scaled by the same parameter  $\gamma$ . If we apply (4) to the model (1)–(2), a larger  $\gamma$  means stronger interaction between  $x_i$  and  $x^{(N)} + \eta_0$ . Now  $\Phi(x^{(N)}) = \gamma(x^{(N)} + \eta_0)$ . Denote  $\mathbf{1}_N = [1, \dots, 1]^T$  consisting of  $N$  ones. By rearranging the integrand of  $J_{\text{soc}}^{(N)}$ , we write

$$J_{\text{soc}}^{(N)} = E \int_0^\infty e^{-\rho t} \left( x^T \hat{Q} x + 2Gx + ru^T u + N\gamma^2 \eta_0^2 \right) dt \quad (5)$$

where  $G = \gamma\eta_0(\gamma - 1)\mathbf{1}_N^T$  and  $\hat{Q} := (\hat{q}_{ij})_{1 \leq i, j \leq N}$  is given by

$$\hat{q}_{ii} = \alpha, \quad i = 1, \dots, N, \quad \hat{q}_{ij} = \beta, \quad i \neq j \quad (6)$$

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

$$\lambda_1 = (1 - \gamma)^2, \quad \lambda_2 = \dots = \lambda_N = 1.$$

A similar LQG control problem was briefly analyzed in [13], 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 subtle difference exists between the two formulations. Given any  $\gamma$ , the pair  $[Q^{1/2}, (A - \rho/2)I_N]$  in [13] is always observable for all sufficiently large  $N$ , where  $Q'$  appears in the quadratic term  $x^T Q' x$  in  $J'$ . When  $\Phi$  is used and  $\gamma = 1$ ,  $\text{span}\{\mathbf{1}_N\}$  is always an unobservable subspace for the system so that the state vector's projection within  $\text{span}\{\mathbf{1}_N\}$  is not penalized by the cost. To minimize (5), we set a deterministic initial condition  $x(0) = z$  and write the optimal cost  $v$  in the form

$$v(z) = z^T P z + 2s_1^T z + s_0. \quad (7)$$

Invoking the standard results of LQG control [4], [11], [28], we have

$$2 \left( a - \frac{\rho}{2} \right) P - \left( \frac{b^2}{r} \right) P^2 + \hat{Q} = 0 \quad (8)$$

$$\rho s_1 = a s_1 - \left(\frac{b^2}{r}\right) P s_1 + \gamma \eta_0 (\gamma - 1) \mathbf{1}_N \quad (9)$$

$$\rho s_0 = - \left(\frac{b^2}{r}\right) s_1^T s_1 + D^2 \text{Tr}(P) + N \gamma^2 \eta_0^2 \quad (10)$$

which results in  $P$  of the form

$$P := (p_{ij})_{1 \leq i, j \leq N}, \quad p_{ij} = \begin{cases} p_N & \text{if } i = j \\ q_N & \text{if } i \neq j. \end{cases} \quad (11)$$

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

1) *Case 1:*  $\gamma \neq 1$  so that  $\lambda_1 > 0$ . Then clearly  $\hat{Q} > 0$  and the pair  $[\hat{Q}^{1/2}, (a - \rho/2)I_N]$  is observable, so that (8) has a unique solution  $P > 0$ .

2) *Case 2:*  $\gamma = 1$  so that  $\lambda_1 = 0$ . Then  $[\hat{Q}^{1/2}, (a - \rho/2)I_N]$  is not fully observable. By using an orthogonal transformation  $\Psi$  such that  $\Psi^T \hat{Q} \Psi = \text{Diag}(\lambda_i) =: \Lambda_{\hat{Q}}$ , from (8) we obtain

$$2 \left(a - \frac{\rho}{2}\right) \Psi^T P \Psi - \left(\frac{b^2}{r}\right) (\Psi^T P \Psi)^2 + \Lambda_{\hat{Q}} = 0.$$

We require 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 a solution  $P \geq 0$  to (8).

For simplicity, below we analyze Case 1 in detail. Substituting (11) into (8) and denoting  $\bar{a} = a - \rho/2$ ,  $\bar{b} = b/\sqrt{r}$ , we obtain the following equations:

$$2\bar{a}p_N - \bar{b}^2 [p_N^2 + (N-1)q_N^2] + \alpha = 0 \quad (12)$$

$$2\bar{a}q_N - \bar{b}^2 [2p_N q_N + (N-2)q_N^2] + \beta = 0. \quad (13)$$

Under the condition  $P > 0$ , solving (12)-(13) yields

$$p_N = \frac{\bar{a} + \sqrt{\bar{a}^2 + \bar{b}^2}}{\bar{b}^2} + \frac{\sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2} - \sqrt{\bar{a}^2 + \bar{b}^2}}{N \bar{b}^2} \quad (14)$$

$$q_N = \frac{\sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2} - \sqrt{\bar{a}^2 + \bar{b}^2}}{N \bar{b}^2}. \quad (15)$$

Furthermore, we obtain

$$s_1 = \frac{\gamma \eta_0 (\gamma - 1)}{\left(\frac{\rho}{2}\right) + \sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2}} \mathbf{1}_N =: s_{11} \mathbf{1}_N^T. \quad (16)$$

We summarize the above results as follows.

*Proposition 3.1:* If  $\gamma \neq 1$ , then the pair  $[(a - \rho/2)I_N, bI_N]$  is controllable, the pair  $[\hat{Q}^{1/2}, (a - \rho/2)I_N]$  is observable, and (8) has a unique solution  $P > 0$  given by (11) and (14)-(15).  $\square$

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

$$u_i^* = - \left(\frac{b}{r}\right) \left( p_N x_i + q_N \sum_{j \neq i} x_j + s_{11} \right). \quad (17)$$

Define  $p_\infty = (\bar{a} + \sqrt{\bar{a}^2 + \bar{b}^2})/\bar{b}^2$ . We state two propositions which are proven in Appendix B.

*Proposition 3.2:* Assume  $\gamma \neq 1$  and **(A2)** holds with  $E x_i(0) = m_0$ ,  $\text{Var}(x_i(0)) = \sigma_0^2$  for all  $i$ . Then

$$\inf_{u \in \mathcal{U}_o} J_{\text{soc}}^{(N)}(u) = \sigma_0^2 \text{Tr}(P) + m_0^2 \mathbf{1}_N^T P \mathbf{1}_N + 2N m_0 s_{11} + s_0 \quad (18)$$

and the asymptotic average social optimum (per agent) is

$$\begin{aligned} & \lim_{N \rightarrow \infty} \left(\frac{1}{N}\right) \inf_{u \in \mathcal{U}_o} J_{\text{soc}}^{(N)}(u) \\ &= \sigma_0^2 p_\infty + m_0^2 \cdot \frac{\bar{a} + \sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2}}{\bar{b}^2} + 2m_0 s_{11} \\ &+ \left(\frac{1}{\rho}\right) (D^2 p_\infty + \gamma^2 \eta_0^2 - \bar{b}^2 s_{11}^2). \end{aligned} \quad (19)$$

$\square$

We continue to examine the limiting dynamics of the closed-loop system when  $N \rightarrow \infty$ . By (17), agent  $i$  has the closed-loop dynamics

$$dx_i = (a - \bar{b}^2 p_N) x_i dt - \bar{b}^2 q_N \sum_{j \neq i} x_j dt - \bar{b}^2 s_{11} dt + D dW_i. \quad (20)$$

Let us denote the ordinary differential equation (ODE)

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

where  $\bar{x}_c(0) = m_0$ . It is easy to solve  $\bar{x}_c$  in terms of  $\bar{x}_c(0)$  and  $s_{11}$ . Set

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

*Proposition 3.3:* Denote  $x^{(N)} = (1/N) \sum_{i=1}^N x_i$  in (20). Assume  $\gamma \neq 1$  and let  $y_i$  satisfy

$$dy_i = (a - \bar{b}^2 p_\infty) y_i dt - h dt + D dW_i \quad (21)$$

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

$$\begin{aligned} & \sup_{t \geq 0} e^{-\hat{\rho} t} \left[ E |x^{(N)}(t) - \bar{x}_c(t)|^2 + E |x_i(t) - y_i(t)|^2 \right] \\ &= O\left(\frac{1}{N}\right). \end{aligned} \quad (22)$$

$\square$

For the control law (17), let  $q_N \sum_{k \neq i} x_k$  be approximated by using  $\bar{x}_c(t)$ . Next  $p_N \rightarrow p_\infty$  as  $N \rightarrow \infty$ . This gives the approximation of  $u_i^*$  by the decentralized control law

$$\begin{aligned} u_i^d = & - \left(\frac{b}{r}\right) p_\infty x_i - \left( \sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2} - \sqrt{\bar{a}^2 + \bar{b}^2} \right) \left( \frac{\bar{x}_c}{b} \right) \\ & - \left(\frac{b}{r}\right) s_{11}, \quad 1 \leq i \leq N. \end{aligned}$$

These steps show that we may construct decentralized strategies by using the limiting behavior of the centralized optimal control law. However, when a large population of nonuniform agents is considered, this procedure faces substantial mathematical difficulties.

### B. Centralized Optimal Control: Person-by-Person Optimality

We consider controls from  $\mathcal{U}_o$ . The benefit is that one can fix the controls of other agents while perturbing the control of a selected agent. Let  $J_{\text{soc}}^{(N)}$  be minimized by  $\hat{u} = (\hat{u}_1, \dots, \hat{u}_N)$ , which is interpreted as a control from  $\mathcal{U}_o$ . Let  $\hat{x}_i$  correspond to  $\hat{u}_i$ . Denote  $\hat{x}^{(N)} = (1/N) \sum_{j=1}^N \hat{x}_j$  and  $\hat{x}_{-i}^{(N)} =$

$(1/N) \sum_{j \neq i} \hat{x}_j$ . Given  $u \in \mathcal{U}_o$ , define  $u_{-i} = (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$ , and  $\hat{u}_{-i}$  is defined similarly.

**Lemma 3.4:** Assume **(A4)** and let  $x(0)$  be given. Suppose that  $\hat{u} \in \mathcal{U}_o$  minimizes  $J_{\text{soc}}^{(N)}$  with the admissible control set  $\mathcal{U}_o$ . Then  $\hat{u}_i$  is the unique optimal control for the control problem

$$\text{(P0)} \quad dx_i = A(\theta_i)x_i dt + B u_i dt + D dW_i, \quad t \geq 0 \quad (23)$$

$$J_i^0(u_i) = J_{\text{soc}}^{(N)}(\hat{u}_1, \dots, \hat{u}_{i-1}, u_i, \hat{u}_{i+1}, \dots, \hat{u}_N) \quad (24)$$

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

*Proof:* Since the integrand of  $J_{\text{soc}}^{(N)}$  is convex in  $(x, u)$  and  $R > 0$ , the existence and uniqueness of an optimal control  $\hat{u} \in \mathcal{U}_o$  holds. If  $\hat{u}_{-i}$  is fixed and  $\hat{u}_i$  is replaced by  $u_i \in \mathcal{U}_{oi}$  such that  $E \int_0^\infty 1_{\{u_i(t, \omega) \neq \hat{u}_i(t, \omega)\}}(t, \omega) dt > 0$ , then  $J_i^0(u_i) > J_i^0(\hat{u}_i)$ .  $\square$

Let  $\hat{u}$  be as in Lemma 3.4. Due to the mean field term in  $J_{\text{soc}}^{(N)}$ ,  $\hat{u}_i$  in general depends on all  $W_j$ ,  $j = 1, \dots, N$ . We give a reformulation of Problem (P0).

**Lemma 3.5:** Assume **(A4)**. In finding the optimal control  $\hat{u}_i \in \mathcal{U}_{oi}$ , Problem (P0) is equivalent to the optimal control problem

$$\text{(P1)} \quad dx_i = A(\theta_i)x_i dt + B u_i dt + D dW_i, \quad t \geq 0 \quad (25)$$

$$J_i^1(u_i) = E \int_0^\infty e^{-\rho t} L(x_i, \hat{x}_{-i}^{(N)}, u_i)(t) dt \quad (26)$$

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

$$\begin{aligned} L = & x_i^T \left[ \left( I - \frac{\Gamma}{N} \right)^T Q \left( I - \frac{\Gamma}{N} \right) + \frac{(N-1)\Gamma^T Q \Gamma}{N^2} \right] x_i \\ & - 2(\Gamma \hat{x}_{-i}^{(N)} + \eta)^T Q \left( I - \frac{\Gamma}{N} \right) x_i \\ & - 2 \left\{ \left[ I - \left( 1 - \frac{1}{N} \right) \Gamma \right] \hat{x}_{-i}^{(N)} - \left( 1 - \frac{1}{N} \right) \eta \right\}^T Q \Gamma x_i \\ & + u_i^T R u_i. \end{aligned} \quad (27)$$

*Proof:* Since  $\hat{u} \in \mathcal{U}_o$ ,  $\hat{u}_{-i}$  has been specified in advance and does not change with  $u_i \in \mathcal{U}_{oi}$ . Hence, each  $\hat{x}_j$ ,  $j \neq i$ , does not change with  $u_i$ . We may write

$$\begin{aligned} J_i(\hat{u}_1, \dots, \hat{u}_{i-1}, u_i, \hat{u}_{i+1}, \dots, \hat{u}_N) \\ = E \int_0^\infty e^{-\rho t} (Z_i + Z'_i) dt \end{aligned}$$

where

$$\begin{aligned} Z_i &= \left| \left( I - \frac{\Gamma}{N} \right) x_i \right|_Q^2 - 2(\Gamma \hat{x}_{-i}^{(N)} + \eta)^T Q \left( I - \frac{\Gamma}{N} \right) x_i + u_i^T R u_i, \\ Z'_i &= |\Gamma \hat{x}_{-i}^{(N)} + \eta|_Q^2 \end{aligned}$$

and  $Z'_i$  does not change with  $u_i$ . For  $j \neq i$ , we have

$$\begin{aligned} J_j(\hat{u}_1, \dots, \hat{u}_{i-1}, u_i, \hat{u}_{i+1}, \dots, \hat{u}_N) \\ = E \int_0^\infty e^{-\rho t} (Z_j + Z'_j) dt \end{aligned}$$

where

$$\begin{aligned} Z_j &= \left( \frac{1}{N^2} \right) x_j^T \Gamma^T Q \Gamma x_i - \left( \frac{2}{N} \right) (\hat{x}_j - \Gamma \hat{x}_{-i}^{(N)} - \eta)^T Q \Gamma x_i, \\ Z'_j &= |\hat{x}_j - \Gamma \hat{x}_{-i}^{(N)} - \eta|_Q^2 + \hat{u}_j^T R \hat{u}_j \end{aligned}$$

and  $Z'_j$  does not change with  $u_i$ . So minimizing  $J_i^0(u_i)$  in (24) is equivalent to minimizing  $E \int_0^\infty e^{-\rho t} \sum_{k=1}^N Z_k dt$ , which is equal to  $J_i^1(u_i)$  in (26).  $\square$

The cost (26) identifies the components of  $J_{\text{soc}}^{(N)}$  which are affected by  $u_i$  when  $\hat{u}_{-i}$  is given. The significance of Lemma 3.5 lies in the determination of a mean field structure where all other agents' effect on  $J_i^1(u_i)$  appears in the form of  $\hat{x}_{-i}^{(N)}$ , which does not change with  $u_i$ . This feature is useful for finding a decentralized control  $u_i$  in Section IV by a deterministic approximation of  $\hat{x}_{-i}^{(N)}$ .

#### IV. 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  $\delta' < \delta$ , where  $\delta'$  may depend on  $f$ .

##### A. 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 (27) by a deterministic function  $\bar{x}$ . To approximate Problem (P1), we construct the auxiliary optimal control problem

$$\begin{aligned} \text{(P2)} \quad dx_i &= A(\theta_i)x_i dt + B u_i dt + D dW_i \\ J^*(u_i) &= E \int_0^\infty e^{-\rho t} L^*(x_i, \bar{x}, u_i) dt \end{aligned} \quad (28)$$

where  $J^*$  is to be minimized and

$$\begin{aligned} L^*(x_i, \bar{x}, u_i) &= x_i^T Q x_i - 2(\Gamma \bar{x} + \eta)^T Q x_i \\ &\quad - 2[(I - \Gamma)\bar{x} - \eta]^T Q \Gamma x_i + u_i^T R u_i \end{aligned} \quad (29)$$

is an approximation of  $L$  in (27). To ensure that  $J^*$  is finite, we impose that  $\bar{x} \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$ .

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

$$\rho \Pi_\theta = \Pi_\theta A_\theta + A_\theta^T \Pi_\theta - \Pi_\theta B R^{-1} B^T \Pi_\theta + Q \quad (30)$$

where  $A_\theta := A(\theta)$ . Under **(A4)**, (30) has a unique solution  $\Pi_\theta \geq 0$ . Denote the ODE

$$\begin{aligned} \rho s_\theta &= \frac{ds_\theta}{dt} + (A_\theta^T - \Pi_\theta B R^{-1} B^T) s_\theta \\ &\quad - [(\Gamma^T Q + Q \Gamma - \Gamma^T Q \Gamma) \bar{x} + (I - \Gamma^T) Q \eta] \end{aligned} \quad (31)$$

which does not have a pre-specified initial condition  $s_\theta(0)$ . In fact, if  $\bar{x} \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$ , one may use the asymptotic stability of  $A_\theta - B R^{-1} B^T \Pi_\theta - (\rho/2)I$  to identify a unique  $s_\theta(0)$  provided that  $s_\theta$  is required to be within  $C_{\rho/2}([0, \infty), \mathbb{R}^n)$ ; see [11, Lemma A.2] for related details.

Assume that  $\bar{x}$  has been given. By using the method in [4], [11], [13], [15], we may show that if **(A4)** holds and if  $s_{\theta_i} \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$  satisfies (31) after setting  $\theta = \theta_i$ , the optimal control law  $\hat{u}_i$  for Problem (P2) is

$$\hat{u}_i = -R^{-1} B^T (\Pi_{\theta_i} x_i + s_{\theta_i}). \quad (32)$$

The closed-loop dynamics take the form

$$dx_i = A_{\theta_i} x_i dt - B R^{-1} B^T (\Pi_{\theta_i} x_i + s_{\theta_i}) dt + D dW_i \quad (33)$$

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

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

$$\begin{aligned} \rho s_\theta &= \frac{ds_\theta}{dt} + (A_\theta^T - \Pi_\theta B R^{-1} B^T) s_\theta \\ &\quad - [(\Gamma^T Q + Q \Gamma - \Gamma^T Q \Gamma) \bar{x} + (I - \Gamma^T) Q \eta] \end{aligned} \quad (34)$$

$$\frac{d\bar{x}_\theta}{dt} = A_\theta \bar{x}_\theta - B R^{-1} B^T (\Pi_\theta \bar{x}_\theta + s_\theta) \quad (35)$$

$$\bar{x} = \int \bar{x}_\theta dF(\theta) \quad (36)$$

where  $\bar{x}_\theta(0) = m_0$  due to **(A2)**, and  $s_\theta$  is sought within  $C_{\rho/2}([0, \infty), \mathbb{R}^n)$ . The first two equations are based on (31) and (33), and (36) is based on the consistency requirement that the mean field initially assumed should be replicated when averaging the closed-loop states of a large number of agents. This equation system is called the Social Certainty Equivalence (SCE) equation system.

The existence and uniqueness analysis of a solution to (34)–(36) may be developed using fixed point methods similar to those in [13] and [14] for NCE equation systems. For the analysis in this paper, we adopt the following assumption.

**(A5)** There exists a solution  $(s_\theta, \bar{x}_\theta, \bar{x}, \theta \in \Theta)$  to the SCE equation system (34)–(36) 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  $s_\theta$  and  $\bar{x}_\theta$  are continuous in  $\theta$  for each fixed  $t \in [0, \infty)$ .  $\diamond$

After a solution is obtained, the  $N$  agents may determine their decentralized strategies by (32).

### B. Social Optimality Theorem

We give a prior error estimate for the mean field approximation. Let  $F_N$  be specified by **(A1)**. Define  $\epsilon_N \geq 0$  by

$$\epsilon_N^2 = \int_0^\infty e^{-\rho t} \left| \int \bar{x}_\theta(t) dF_N(\theta) - \int \bar{x}_\theta(t) dF(\theta) \right|^2 dt. \quad (37)$$

*Lemma 4.1:* Suppose that **(A1)**–**(A5)** hold. Let  $\hat{x}_i$  be the closed-loop solution of agent  $i$  under the SCE based control law (32) and  $\hat{x}^{(N)} = (1/N) \sum_{i=1}^N \hat{x}_i$ . Then

$$E \int_0^\infty e^{-\rho t} \left| \hat{x}^{(N)}(t) - \bar{x}(t) \right|^2 dt \leq C \left( \frac{1}{N} + \epsilon_N^2 \right)$$

where  $\lim_{N \rightarrow \infty} \epsilon_N = 0$ .

*Proof:* For  $\bar{x}$  within  $(s_\theta, \bar{x}_\theta, \bar{x}, \theta \in \Theta)$ , there exist  $C_0$  and  $\rho' \in (0, \rho)$  such that  $|\bar{x}(t)| \leq C_0 e^{(\rho'/2)t}$  for  $t \geq 0$ . Denote  $\hat{A}_\theta := A_\theta - B R^{-1} B^T \Pi_\theta - \rho I$ , whose eigenvalues all have a real part less than  $-\rho/2$ . By using (34) and following the argument in [11, Eq. (A.9)], we obtain

$$\begin{aligned} s_\theta(t) &= -e^{-\hat{A}_\theta^T t} \\ &\quad \cdot \int_t^\infty e^{\hat{A}_\theta^T \tau} [(\Gamma^T Q + Q \Gamma - \Gamma^T Q \Gamma) \bar{x} + (I - \Gamma^T) Q \eta] d\tau. \end{aligned} \quad (38)$$

By [32, p. 54], the continuity of  $A(\theta)$  ensures that the stabilizability of the pair  $[A(\theta), B]$  depends continuously on  $\theta$ . Furthermore, by [7],  $\Pi_\theta$  depends continuously on  $[A(\theta), B, Q^{1/2}]$  and hence on  $\theta$ . Since  $\Theta$  is compact,  $\sup_{\theta \in \Theta} |\Pi_\theta| \leq C$ . In addition, there exists a fixed  $\delta > 0$  satisfying 1) all eigenvalues of  $\hat{A}_\theta$  have a real part less than  $-(\delta + (\rho/2))$ , and hence 2) for

some  $C$ ,  $\sup_{\theta \in \Theta} |e^{\hat{A}_\theta t}| \leq C e^{-(\rho+\delta)t/2}$ , where without loss of generality we assume  $\rho - \delta > \rho'$ . Subsequently, (38) leads to

$$\sup_{\theta \in \Theta} |s_\theta(t)| \leq C e^{\rho' t/2}. \quad (39)$$

Combining (39) with (35), we further obtain

$$\sup_{\theta \in \Theta} |\bar{x}_\theta(t)| \leq C e^{(\rho-\delta)t/2}. \quad (40)$$

By our convention,  $C$  may take different values. Since  $\{\hat{x}_i - \bar{x}_{\theta_i}, i \geq 1\}$  are independent processes of zero mean:

$$\begin{aligned} E |\hat{x}^{(N)} - \bar{x}|^2 &= E \left| \left( \frac{1}{N} \right) \sum_{i=1}^N (\hat{x}_i - \bar{x}_{\theta_i}) + \left( \frac{1}{N} \right) \sum_{i=1}^N (\bar{x}_{\theta_i} - \bar{x}) \right|^2 \\ &= \left( \frac{1}{N^2} \right) \sum_{i=1}^N E |\hat{x}_i - \bar{x}_{\theta_i}|^2 + \left| \left( \frac{1}{N} \right) \sum_{i=1}^N (\bar{x}_{\theta_i} - \bar{x}) \right|^2. \end{aligned}$$

Since  $\int_0^\infty e^{-\rho t} E |\hat{x}_i - \bar{x}_{\theta_i}|^2 dt \leq C$ , where  $C$  does not depend on  $(i, N)$  due to the compactness of  $\Theta$  and (39)–(40), we have

$$\left( \frac{1}{N^2} \right) \sum_{i=1}^N \int_0^\infty e^{-\rho t} E |\hat{x}_i - \bar{x}_{\theta_i}|^2 dt \leq \frac{C}{N}. \quad (41)$$

Moreover,

$$\begin{aligned} \left( \frac{1}{N} \right) \sum_{i=1}^N [\bar{x}_{\theta_i}(t) - \bar{x}(t)] &= \int \bar{x}_\theta(t) dF_N(\theta) - \int \bar{x}_\theta(t) dF(\theta) =: G_N(t). \end{aligned}$$

By weak convergence in **(A1)**,  $\lim_{N \rightarrow \infty} G_N(t) = 0$  for each  $t$  since  $\bar{x}_\theta$  is continuous in  $\theta$  by **(A5)**.

For any  $\epsilon > 0$ , by (40) there exists a sufficiently large  $T_1$  such that

$$\sup_N \int_{T_1}^\infty e^{-\rho t} |G_N(t)|^2 dt \leq \frac{\epsilon}{2}.$$

Since the sequence  $\{G_N(t)\}$  is uniformly bounded on  $[0, T_1]$  and converges to 0 for each  $t$ ,  $\lim_{N \rightarrow \infty} \int_0^{T_1} e^{-\rho t} |G_N(t)|^2 dt = 0$ . Then for each sufficiently large  $N$ ,  $\int_0^\infty e^{-\rho t} |G_N(t)|^2 dt \leq \epsilon$ . Since  $\epsilon$  is arbitrary,  $\lim_{N \rightarrow \infty} \epsilon_N^2 = \lim_{N \rightarrow \infty} \int_0^\infty e^{-\rho t} |G_N(t)|^2 dt = 0$ . Recalling (41), the lemma follows.  $\square$

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

*Theorem 4.2:* 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  $\hat{u} = (\hat{u}_1, \dots, \hat{u}_N)$ :

$$\left| \left( \frac{1}{N} \right) J_{\text{soc}}^{(N)}(\hat{u}) - \inf_{u \in \mathcal{U}_o} \left( \frac{1}{N} \right) J_{\text{soc}}^{(N)}(u) \right| = O \left( \frac{1}{\sqrt{N}} + \epsilon_N \right)$$

where  $\mathcal{U}_o$  is defined in Section II-B as a set of centralized information-based controls.

*Proof:* See Appendix C.  $\square$

For the scalar model with uniform agents and the parametrization (4), 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 (see Lemma C.1, Appendix C).

### 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 (34)–(36) becomes

$$\frac{d\bar{x}}{dt} = (A - BR^{-1}B^T\Pi)\bar{x} - BR^{-1}B^T s \quad (42)$$

$$\frac{ds}{dt} = Q_\Gamma \bar{x} + (\rho I - A^T + \Pi BR^{-1}B^T) s + \eta_\Gamma \quad (43)$$

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

$$Q_\Gamma = \Gamma^T Q + Q\Gamma - \Gamma^T Q\Gamma, \quad \eta_\Gamma = (I - \Gamma^T)Q\eta. \quad (44)$$

A solution  $(s, \bar{x})$  of (42)–(43) 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. The explicit calculation of the solution is given in the proof in Appendix B.

*Theorem 4.3:* Suppose that **(A4)** holds for  $A(\theta) = A$  and that  $Q_\Gamma = \Gamma^T Q + Q\Gamma - \Gamma^T Q\Gamma \leq 0$ . Then (42)–(43) has a unique solution  $(s, \bar{x}) \in C_{\rho/2}([0, \infty), \mathbb{R}^{2n})$ .  $\square$

### D. The Social Variational Interpretation

We describe a social variational interpretation to the SCE methodology, and this is potentially useful for generalization to nonlinear models. We assume **(A1)**–**(A4)** and illustrate the idea by starting with finite horizon. Let the dynamics of agent  $i$  be given by (1) and its cost be

$$J_i = E \int_0^T e^{-\rho t} \left[ |x_i - (\Gamma x^{(N)} + \eta)|_Q^2 + u_i^T R u_i \right] dt.$$

Define the social cost  $J_{\text{soc}}^{(N)} = \sum_{i=1}^N J_i$ .

Let  $J_{\text{soc}}^{(N)}$  be minimized by  $\hat{u} = (\hat{u}_1, \dots, \hat{u}_N)$ , which is adapted to  $\mathcal{F}_t^0$  defined in Section II-B. Let  $\hat{x}_i$  correspond to  $\hat{u}_i$ , and  $\hat{x}^{(N)} = (1/N) \sum_{j=1}^N \hat{x}_j$ . Then

$$J_i(\hat{u}_1, \dots, \hat{u}_N) = E \int_0^T e^{-\rho t} \left[ |\hat{x}_i - (\Gamma \hat{x}^{(N)} + \eta)|_Q^2 + \hat{u}_i^T R \hat{u}_i \right] dt. \quad (45)$$

Let the control variation  $\delta \hat{u}_i$  be adapted to  $\mathcal{F}_t^0$  and satisfy  $E \int_0^T |\delta \hat{u}_i|^2 dt < \infty$ . Suppose that  $\delta \hat{u}_i$  gives rise to the state variation  $\delta \hat{x}_i$ . By the linearity of the dynamics,  $E \int_0^T |\delta \hat{x}_i|^2 dt < \infty$ . Denote  $\phi_1 = \hat{x}_i - (\Gamma \hat{x}^{(N)} + \eta) + (I - \Gamma/N)\delta \hat{x}_i$ . We have

$$J_i(\hat{u}_1, \dots, \hat{u}_i + \delta \hat{u}_i, \dots, \hat{u}_N) = E \int_0^T e^{-\rho t} \left[ \phi_1^T Q \phi_1 + (\hat{u}_i + \delta \hat{u}_i)^T R (\hat{u}_i + \delta \hat{u}_i) \right] dt.$$

This gives the first variation of  $J_i$  with respect to  $\delta \hat{u}_i$  as

$$\delta \hat{J}_i(\delta \hat{u}_i) = 2E$$

$$\cdot \int_0^T e^{-\rho t} \cdot \left\{ [\hat{x}_i - (\Gamma \hat{x}^{(N)} + \eta)]^T Q \left( I - \frac{\Gamma}{N} \right) \delta \hat{x}_i + \hat{u}_i^T R \delta \hat{u}_i \right\} dt.$$

Denote  $\phi_2 = \hat{x}_j - (\Gamma \hat{x}^{(N)} + \eta) - (\Gamma/N)\delta \hat{x}_i$ . To find the cost variation of agent  $j \neq i$ , we write

$$J_j(\hat{u}_1, \dots, \hat{u}_i + \delta \hat{u}_i, \dots, \hat{u}_N) = E \int_0^T e^{-\rho t} \left( \phi_2^T Q \phi_2 + \hat{u}_j^T R \hat{u}_j \right) dt.$$

Let  $\delta \hat{J}_j(\delta \hat{u}_i)$  be the first variation of  $J_j$  with a control variation  $\delta \hat{u}_i$  at  $(\hat{x}_k, \hat{u}_k, 1 \leq k \leq N)$ . Then

$$\delta \hat{J}_j(\delta \hat{u}_i) = \left( -\frac{2}{N} \right) E \int_0^T e^{-\rho t} \left[ \hat{x}_j - (\Gamma \hat{x}^{(N)} + \eta) \right]^T Q \Gamma \delta \hat{x}_i dt.$$

Denote  $\hat{x}_{-i}^{(N)} = (1/N) \sum_{j \neq i} \hat{x}_j$ . The aggregate cost variation (over all agents except agent  $i$ ) is obtained as

$$\sum_{j \neq i} \delta \hat{J}_j(\delta \hat{u}_i) = -2E \int_0^T e^{-\rho t} \cdot \left[ \hat{x}_{-i}^{(N)} - \left( 1 - \frac{1}{N} \right) (\Gamma \hat{x}^{(N)} + \eta) \right]^T Q \Gamma \delta \hat{x}_i dt.$$

The social cost variation is  $\delta \hat{J}(\delta \hat{u}_i) := \sum_{k=1}^N \delta \hat{J}_k(\delta \hat{u}_i)$ . The zero first variational condition  $\delta \hat{J}(\delta \hat{u}_i) = 0$  combined with the approximation of  $\hat{x}^{(N)}$  and  $\hat{x}_{-i}^{(N)}$  by a deterministic function  $\bar{x}$  gives the following condition in the large population limit

$$0 = E \int_0^T e^{-\rho t} \left\{ [\hat{x}_i - (\Gamma \bar{x} + \eta)]^T Q \delta \hat{x}_i - [\bar{x} - (\Gamma \bar{x} + \eta)]^T Q \Gamma \delta \hat{x}_i + \hat{u}_i^T R \delta \hat{u}_i \right\} dt. \quad (46)$$

We write (46) in the equivalent form

$$0 = E \int_0^T e^{-\rho t} \left\{ \hat{x}_i^T Q \delta \hat{x}_i + [(\Gamma \bar{x} + \eta)^T Q (\Gamma - I) - \bar{x}^T Q \Gamma] \delta \hat{x}_i + \hat{u}_i^T R \delta \hat{u}_i \right\} dt. \quad (47)$$

The key observation is that we may reinterpret (47) as the variational condition for the optimal control problem with dynamics (1) and cost

$$J'_i = E \int_0^T e^{-\rho t} \left\{ x_i^T Q x_i + 2[(\Gamma \bar{x} + \eta)^T Q (\Gamma - I) - \bar{x}^T Q \Gamma] x_i + u_i^T R u_i \right\} dt \quad (48)$$

for which the determination of the optimal state and control pair  $(\hat{x}_i, \hat{u}_i)$  is straightforward. Subsequently, letting  $\bar{x}$  be replicated by the closed-loop dynamics of a large number of agents, we will again obtain the equation system (34)–(36) when  $T \rightarrow \infty$ . The details are omitted here. So in the linear quadratic setting this variational approach is equivalent to the approach based on the construction of  $L^*$  in (29).

### E. Connection With the NCE Equation System

To compare with our past work on Problem II, we review the NCE approach for the game problem where agent  $i$  is associated with cost  $J_i$  (see, e.g., [13]). 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  $\Gamma \bar{x}^\dagger + \eta$  and solves an auxiliary 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  $\bar{x}^\dagger$  initially assumed. This procedure leads to the NCE equation system

$$\rho s_\theta^\dagger = \frac{ds_\theta^\dagger}{dt} + (A_\theta^T - \Pi_\theta B R^{-1} B^T) s_\theta^\dagger - Q(\Gamma \bar{x}^\dagger + \eta) \quad (49)$$

$$\frac{d\bar{x}_\theta^\dagger}{dt} = A_\theta \bar{x}_\theta^\dagger - B R^{-1} B^T (\Pi_\theta \bar{x}_\theta^\dagger + s_\theta^\dagger) \quad (50)$$

$$\bar{x}^\dagger = \int \bar{x}_\theta^\dagger dF(\theta), \quad \bar{x}_\theta^\dagger(0) = m_0. \quad (51)$$

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

$$u_i = -B R^{-1} (\Pi_\theta x_i + s_\theta^\dagger), \quad 1 \leq i \leq N \quad (52)$$

is an  $\varepsilon$ -Nash equilibrium [13], where  $\varepsilon \rightarrow 0$  as  $N \rightarrow \infty$ , and an intuitive interpretation of this property is as follows. When the strategies  $(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$  are given by (52), the term  $\Gamma x^{(N)} + \eta$  in the cost (2) of agent  $i$  is nearly equal to  $\Gamma \bar{x}^\dagger + \eta$  for large  $N$  and agent  $i$  has little room to further reduce its cost by using a strategy other than (52) since (49) has been constructed to let  $x_i$  optimally track  $\Gamma \bar{x}^\dagger + \eta$ . The NCE and SCE equation systems differ by a different equation for  $s_\theta^\dagger$ . The reason is that in the socially optimal control problem, each agent must consider its social impact when choosing its strategy and cannot simply make an optimal response to a deterministic approximation of  $\Gamma x^{(N)} + \eta$ .

## V. SCALAR MODEL WITH UNIFORM AGENTS

Recall that for the scalar model, we set  $Q = 1$ ,  $(A(\cdot), B, R) = (a, b, r)$  and assume  $b \neq 0$ , and that  $\Gamma$  and  $\eta$  are parametrized according to (4). Let  $\Pi > 0$  be the solution to the ARE

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

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

### A. Comparison of Solvability of the Two Equation Systems

Now the SCE equation system (34)–(36) reduces to

$$\frac{ds}{dt} = \beta_2 s + (2\gamma - \gamma^2) \bar{x} + (1 - \gamma) \gamma \eta_0 \quad (53)$$

$$\frac{d\bar{x}}{dt} = -\bar{b}^2 s - \beta_1 \bar{x}. \quad (54)$$

Recall that the proof of Theorem 4.3 describes a procedure to solve the SCE equation system with uniform agents and it requires  $Q_\Gamma \leq 0$ , which translates here to  $2\gamma - \gamma^2 \leq 0$ . Now we drop that assumption and identify a much more general solvability condition.

The NCE equation system (49)–(51) reduces to

$$\frac{ds^\dagger}{dt} = \beta_2 s^\dagger + \gamma \bar{x}^\dagger + \gamma \eta_0 \quad (55)$$

$$\frac{d\bar{x}^\dagger}{dt} = -\bar{b}^2 s^\dagger - \beta_1 \bar{x}^\dagger. \quad (56)$$

For further analysis of (53)–(54), we introduce the set of algebraic equations

$$\beta_2 s(\infty) + (2\gamma - \gamma^2) \bar{x}(\infty) = (\gamma - 1) \gamma \eta_0 \quad (57)$$

$$\bar{b}^2 s(\infty) + \beta_1 \bar{x}(\infty) = 0. \quad (58)$$

The notation  $s(\infty)$ ,  $\bar{x}(\infty)$  is only for constructing the algebraic equations by dropping the derivative terms in (53)–(54). It does not necessarily mean  $\bar{x}(\infty) = \lim_{t \rightarrow \infty} \bar{x}(t)$ , and in fact the solution  $(s(t), \bar{x}(t))$  is allowed to be unbounded. Similarly, for (55)–(56) we introduce

$$\beta_2 s^\dagger(\infty) + \gamma \bar{x}^\dagger(\infty) = -\gamma \eta_0 \quad (59)$$

$$\bar{b}^2 s^\dagger(\infty) + \beta_1 \bar{x}^\dagger(\infty) = 0. \quad (60)$$

Denote

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

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

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  $\gamma_2^\dagger = 1 + \bar{a}^2/\bar{b}^2$ . The following lemma may be easily verified and the proof is omitted.

*Lemma 5.1:* If  $\Delta \neq 0$ , (57)–(58) has a unique solution

$$s(\infty) = \frac{\gamma(\gamma - 1)\eta_0\beta_1}{\Delta}, \quad \bar{x}(\infty) = -\frac{\gamma(\gamma - 1)\eta_0\bar{b}^2}{\Delta}.$$

□

We study the solvability of (53)–(54) and (55)–(56) in terms of the parameter  $\gamma$  such that each function is within  $C_{\rho/2}([0, \infty), \mathbb{R})$ .

*Theorem 5.2:* Suppose  $b \neq 0$ . We have the following assertions.

- 1) If  $\gamma \neq 1$  and  $\Delta \neq 0$ , then (53)–(54) has a unique solution  $(s, \bar{x}) \in C_{\rho/2}([0, \infty), \mathbb{R}^2)$  for any  $\bar{x}(0)$ .
- 2) If  $\gamma \neq \gamma_1^\dagger$ , then (59)–(60) has a unique solution  $(s^\dagger(\infty), \bar{x}^\dagger(\infty))$ . If

$$\gamma \in (-\infty, \gamma_1^\dagger) \cup (\gamma_1^\dagger, \gamma_2^\dagger) \quad (61)$$

the equation system (55)–(56) has a unique solution  $(s^\dagger, \bar{x}^\dagger) \in C_{\rho/2}([0, \infty), \mathbb{R}^2)$  for any  $\bar{x}(0)$ . If

$$\gamma \geq \gamma_2^\dagger \quad (62)$$

the equation system (55)–(56) 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.

*Proof:* 1) Let  $(s(\infty), \bar{x}(\infty))$  be given by Lemma 5.1. Denote  $\tilde{s}(t) = s(t) - s(\infty)$  and  $\tilde{x}(t) = \bar{x}(t) - \bar{x}(\infty)$ . From (53)-(54), we obtain

$$\frac{d\tilde{s}}{dt} = \beta_2 \tilde{s} + (2\gamma - \gamma^2) \tilde{x} \quad (63)$$

$$\frac{d\tilde{x}}{dt} = -\bar{b}^2 \tilde{s} - \beta_1 \tilde{x} \quad (64)$$

for which the coefficient matrix is denoted by  $M$ . Then  $M$  has eigenvalues  $\alpha_1$  and  $\alpha_2$  as

$$\alpha_{1,2} = \frac{\rho}{2} \mp \sqrt{\bar{a}^2 + (1 - \gamma)^2 \bar{b}^2}, \quad \alpha_1 < \frac{\rho}{2}. \quad (65)$$

The solution  $(\tilde{s}, \tilde{x})$  to (63)-(64) is given by  $e^{Mt}[\tilde{s}(0), \tilde{x}(0)]^T$  (as a column vector), which is in  $C_{\rho/2}([0, \infty), \mathbb{R}^2)$  if and only if  $v = [\tilde{s}(0), \tilde{x}(0)]^T$  is an eigenvector of  $M$  such that

$$Mv = \alpha_1 v. \quad (66)$$

By (66), we determine

$$s(0) = s(\infty) - (2\gamma - \gamma^2)(\beta_2 - \alpha_1)^{-1}[\bar{x}(0) - \bar{x}(\infty)]$$

where  $\beta_2 - \alpha_1 > 0$ . Subsequently, we obtain a unique solution  $(s, \bar{x}) \in C_{\rho/2}([0, \infty), \mathbb{R}^2)$ , where

$$\begin{aligned} s(t) &= s(\infty) - (2\gamma - \gamma^2)(\beta_2 - \alpha_1)^{-1}[\bar{x}(0) - \bar{x}(\infty)]e^{\alpha_1 t} \\ \bar{x}(t) &= \bar{x}(\infty) + [\bar{x}(0) - \bar{x}(\infty)]e^{\alpha_1 t}. \end{aligned} \quad (67)$$

2) If  $\gamma \neq \gamma_1^\dagger$ ,  $\Delta^\dagger \neq 0$  so that (59)-(60) has a unique solution

$$(s^\dagger(\infty), \bar{x}^\dagger(\infty)) = \left( -\frac{\gamma\beta_1\eta_0}{\Delta^\dagger}, \frac{\gamma\bar{b}^2\eta_0}{\Delta^\dagger} \right).$$

Denote  $\tilde{s}^\dagger(t) = s^\dagger(t) - s^\dagger(\infty)$  and  $\tilde{x}^\dagger(t) = \bar{x}^\dagger(t) - \bar{x}^\dagger(\infty)$ . The pair  $(\tilde{s}^\dagger, \tilde{x}^\dagger)$  satisfies a linear ODE system which has two eigenvalues  $\lambda_1$  and  $\lambda_2$  as

$$\lambda_{1,2} = \frac{\rho}{2} \mp \sqrt{\bar{a}^2 + (1 - \gamma)\bar{b}^2}.$$

If (61) holds,  $\lambda_1 < \rho/2$ . By the same method as in proving part 1), we may find a unique solution

$$s^\dagger(t) = s^\dagger(\infty) - \gamma(\beta_2 - \lambda_1)^{-1} \tilde{x}^\dagger(0) e^{\lambda_1 t} \quad (68)$$

$$\bar{x}^\dagger(t) = \bar{x}^\dagger(\infty) + \tilde{x}^\dagger(0) e^{\lambda_1 t}. \quad (69)$$

If (62) holds, then  $\Delta^\dagger \neq 0$ , and both  $\lambda_1$  and  $\lambda_2$  have a real part equal to  $\rho/2$ . The only solution for  $(\tilde{s}^\dagger, \tilde{x}^\dagger)$  within  $C_{\rho/2}([0, \infty), \mathbb{R}^2)$  is a constant solution  $(\tilde{s}^\dagger(t), \tilde{x}^\dagger(t)) \equiv (0, 0)$ , which is equivalent to  $s^\dagger(t) \equiv s^\dagger(\infty)$ ,  $\bar{x}^\dagger(t) \equiv \bar{x}^\dagger(\infty)$ . Now, if  $\bar{x}^\dagger(0) = \bar{x}^\dagger(\infty)$ , we further select the initial condition  $s^\dagger(0) = s^\dagger(\infty)$  to obtain  $(s^\dagger(t), \bar{x}^\dagger(t))$  as a constant vector. If  $\bar{x}^\dagger(0) \neq \bar{x}^\dagger(\infty)$ , there is no solution  $(s^\dagger, \bar{x}^\dagger)$  with the required growth condition.  $\square$

*Remark:* The existence and uniqueness of a bounded solution to (55)-(56) was analyzed in [13] via a fixed point approach under the additional conditions  $\beta_1 > 0$  and  $\beta_1\beta_2 > |\gamma|\bar{b}^2$ . In Theorem 5.2-2) the solution is allowed to be unbounded and we identify the maximum range for  $\gamma$ .  $\diamond$

*Remark:* If  $b \neq 0$ ,  $\gamma \neq 1$  and  $\Delta \neq 0$ , we may use (67) to compute  $s$  in the SCE based control law  $u_i = -(b/r)(\Pi x_i + s)$  for agent  $i$ , which leads to

$$dx_i = (a - \bar{b}^2\Pi)x_i dt - \bar{b}^2 s dt + DdW_i. \quad (70)$$

In fact (70) coincides with the limiting dynamics (21) of the centralized optimal control problem since we may verify that  $\Pi = p_\infty$ ,  $\bar{b}^2 s = h$ .  $\diamond$

### B. Limiting Average Costs and Blowup Effect

For the LQG game with individual costs  $J_i$ ,  $1 \leq i \leq N$ , we use the solution of (55)-(56) to compute the individual cost in the population limit. To simplify the calculation, let  $x_i(0) \equiv 0$ . We evaluate the optimal cost to the limiting control problem

$$dx_i = ax_i dt + bu_i dt + DdW_i \quad (71)$$

$$\bar{J}_i(u_i) = E \int_0^\infty e^{-\rho t} \left\{ [x_i - \gamma(\bar{x}^\dagger + \eta_0)]^2 + ru_i^2 \right\} dt \quad (72)$$

where  $\bar{x}^\dagger \in C_{\rho/2}([0, \infty), \mathbb{R})$  is to be determined by (55)-(56). Note that in the closed-loop with the NCE based control laws for all agents,  $\bar{x}^\dagger$  is an approximation of  $x^{(N)} = (1/N) \sum_{j=1}^N x_j$ .

*Lemma 5.3:* If  $x_i(0) \equiv 0$  and  $\gamma \in (-\infty, \gamma_1^\dagger) \cup (\gamma_1^\dagger, \gamma_2^\dagger)$ , the optimal cost of (71)-(72) is

$$v_{\text{ncc}}(\gamma) = \frac{\Pi D^2}{\rho} + \left( \frac{\gamma^2 \eta_0^2}{|\Delta^\dagger|^2} \right) [V_1(\gamma) + V_2(\gamma)] \quad (73)$$

where

$$\begin{aligned} V_1(\gamma) &= \frac{(\beta_2^2 - \bar{b}^2)\beta_1^2}{\rho} - \frac{2\gamma\bar{b}^2\beta_1\beta_2}{(\rho - \lambda_1)} + \frac{2\gamma\bar{b}^4\beta_1}{[(\beta_2 - \lambda_1)(\rho - \lambda_1)]}, \\ V_2(\gamma) &= \left[ \frac{\gamma^2\bar{b}^4}{(\rho - 2\lambda_1)} \right] \left[ 1 - \frac{\bar{b}^2}{(\beta_2 - \lambda_1)^2} \right]. \end{aligned}$$

*Proof:* Let  $(s^\dagger, \bar{x}^\dagger) \in C_{\rho/2}([0, \infty), \mathbb{R}^2)$  be given by (55)-(56). First, the NCE based control law  $u_i = -(b/r)(\Pi x_i + s^\dagger)$  minimizes  $\bar{J}_i(\cdot)$ . Next, we use Lemma A.1 to calculate the optimal cost  $v_{\text{ncc}}(\gamma)$  with  $x_i(0) \equiv 0$ . Denote the ODE

$$\rho q_{\text{ncc}} = \frac{dq_{\text{ncc}}}{dt} - \bar{b}^2 s^{\dagger 2} + \Pi D^2$$

where  $q_{\text{ncc}} \in C_\rho([0, \infty), \mathbb{R})$ . In view of (A.8)

$$\begin{aligned} v_{\text{ncc}}(\gamma) &= q_{\text{ncc}}(0) + \gamma^2 \int_0^\infty e^{-\rho t} (\bar{x}^\dagger + \eta_0)^2 dt \\ &= \frac{\Pi D^2}{\rho} - \bar{b}^2 \int_0^\infty e^{-\rho t} |s^\dagger|^2 dt \\ &\quad + \gamma^2 \int_0^\infty e^{-\rho t} (\bar{x}^\dagger + \eta_0)^2 dt. \end{aligned} \quad (74)$$

By substituting (68)-(69) into (74) and using  $\bar{x}(0) = 0$ , we obtain (73).  $\square$

Note that  $\gamma_1^\dagger$  is an eigenvalue of (59)-(60) and, as  $\gamma \rightarrow \gamma_1^\dagger$ ,  $\bar{x}^\dagger(\infty)$  becomes unbounded if  $\gamma_1^\dagger \eta_0 \neq 0$ . When  $\gamma$  approaches  $\gamma_2^\dagger$  from the left side,  $(s^\dagger, \bar{x}^\dagger)$  has a growth nearly equal to

$e^{(\rho/2)t}$ , which the discount rate in the cost can barely handle. A natural question is under these situations how the corresponding costs behave. The following theorem provides a classification.

**Theorem 5.4:** Suppose that  $\eta_0 \neq 0$  and  $v_{\text{ncc}}(\gamma)$  is given by (73) for  $\gamma \in (-\infty, \gamma_1^\dagger) \cup (\gamma_1^\dagger, \gamma_2^\dagger)$ . Then the following holds:

- 1) If  $\bar{a} \neq 0$ , then  $\gamma_2^\dagger > 1$  and  $\lim_{\gamma \uparrow \gamma_2^\dagger} v_{\text{ncc}}(\gamma) = \infty$ .
- 2) If  $\bar{a} = 0$ , then  $\gamma_2^\dagger = 1$  and  $\lim_{\gamma \uparrow 1} v_{\text{ncc}}(\gamma) = D^2/(\rho|\bar{b}|) + \eta_0^2(4|\bar{b}|^2 + \rho^2)/\rho^3$ .
- 3) If  $\bar{a}^2 + \bar{b}^2 \neq \rho/4$ , then  $\gamma_1^\dagger \neq 0$  and  $\lim_{\gamma \rightarrow \gamma_1^\dagger} v_{\text{ncc}}(\gamma)$  is finite.
- 4) If  $\bar{a}^2 + \bar{b}^2 = \rho/4$ , then  $\gamma_1^\dagger = 0$  and  $\lim_{\gamma \rightarrow 0} v_{\text{ncc}}(\gamma) = \Pi D^2/\rho$ .

*Proof:* Note that  $\rho - \lambda_1 \geq \rho/2$ ,  $\beta_2 - \lambda_1 \geq |\bar{b}| > 0$ , and  $\rho - 2\lambda_1 = 2\sqrt{\bar{a}^2 + (1-\gamma)\bar{b}^2}$ . We have  $\lim_{\gamma \uparrow \gamma_2^\dagger} \Delta^\dagger = -\rho^2/4$ .

We show 1). If  $\bar{a} \neq 0$ ,  $\gamma_2^\dagger = 1 + \bar{a}^2/\bar{b}^2 > 1$ . When  $\gamma \uparrow \gamma_2^\dagger$ ,  $V_1(\gamma)$  converges to a finite limit and  $V_2(\gamma) \rightarrow \infty$ . Hence, 1) follows.

For 2), if  $\bar{a} = 0$ , we have  $\gamma_2^\dagger = 1 + \bar{a}^2/\bar{b}^2 = 1$  and

$$\lim_{\gamma \rightarrow 1} V_1(\gamma) = -|\bar{b}|^3 + \frac{\rho|\bar{b}|^2}{4} + \frac{\rho^3}{16}, \quad \lim_{\gamma \rightarrow 1} V_2(\gamma) = |\bar{b}|^3.$$

Hence, 2) follows from  $\lim_{\gamma \uparrow 1} v_{\text{ncc}}(\gamma) = \Pi D^2/\rho + \eta_0^2(4|\bar{b}|^2 + \rho^2)/\rho^3$  and  $\Pi = 1/|\bar{b}|$ .

We continue to check 3) and 4). For  $\gamma_1^\dagger = 1 + \bar{a}^2/\bar{b}^2 - \rho^2/(4\bar{b}^2)$ , we have

$$V_1(\gamma_1^\dagger) + V_2(\gamma_1^\dagger) = \frac{\beta_1^2(\beta_2^2 - \bar{b}^2)}{\rho} - \frac{2\beta_1^2\beta_2^2}{\rho} + \frac{2\bar{b}^2\beta_1^2}{\rho} + \frac{\beta_1^2\beta_2^2(1 - \frac{\bar{b}^2}{\beta_2^2})}{\rho} = 0.$$

Note that  $(\beta_1, \beta_2, \rho, \bar{b})$  does not depend on  $\gamma$ , and  $\gamma = (\beta_1\beta_2 + \rho\lambda_1 - \lambda_1^2)/\bar{b}^2$ . Denote  $\Psi(\lambda_1) = (\rho - \lambda_1)(\rho - 2\lambda_1)(\beta_2 - \lambda_1)^2$ . We write  $V_1 + V_2$  as a function of  $\lambda_1$  with parameters  $(\beta_1, \beta_2, \rho, \bar{b})$  so that  $V_1(\gamma) + V_2(\gamma) = H(\lambda_1)/\Psi(\lambda_1)$ , where  $H(\lambda_1)$  is a polynomial of  $\lambda_1$  and is uniquely determined. Obviously,  $\lambda_1 = 0$  if  $\gamma = \gamma_1^\dagger$ . We have  $H(0) = [V_1(\gamma) + V_2(\gamma)]_{\gamma=\gamma_1^\dagger} \Psi(\lambda_1)_{\lambda_1=0} = 0$ . Let  $c[\lambda_1^k; H]$  be the coefficient of  $\lambda_1^k$  in  $H$ ,  $k = 1, 2, 3, \dots$ . By lengthy but elementary calculations, we may further show that  $c[\lambda_1; H] = 0$  and

$$c[\lambda_1^2; H] = \frac{\beta_2^2(2\beta_1^2\beta_2^2 + 2\rho^2\beta_1\beta_2 + \rho^4 - 2\bar{b}^2\beta_1\beta_2 - \bar{b}^2\rho^2)}{\rho}.$$

By using  $\beta_1\beta_2 = \bar{a}^2 + \bar{b}^2 - \rho^2/4$ , we may verify that  $c[\lambda_1^2; H] > 0$ .

If  $\bar{a}^2 + \bar{b}^2 \neq \rho^2/4$ , we have  $\gamma_1^\dagger \neq 0$ . In addition,  $\Delta^\dagger = -\lambda_1\lambda_2$ ,  $\lim_{\gamma \rightarrow \gamma_1^\dagger} \lambda_2 = \rho$ , and

$$\lim_{\gamma \rightarrow \gamma_1^\dagger} \left( \frac{\gamma^2}{|\Delta^\dagger|^2} \right) (V_1(\gamma) + V_2(\gamma)) = \frac{|\gamma_1^\dagger|^2 c[\lambda_1^2; H]}{[\rho^2 \Psi(0)]}.$$

Hence,

$$\lim_{\gamma \rightarrow \gamma_1^\dagger} v_{\text{ncc}}(\gamma) = \frac{\Pi D^2}{\rho} + \frac{|\gamma_1^\dagger|^2 \eta_0^2 (2\beta_1^2\beta_2^2 + 2\rho^2\beta_1\beta_2 + \rho^4 - 2\bar{b}^2\beta_1\beta_2 - \bar{b}^2\rho^2)}{\rho^5}$$

and 3) follows. Finally, if  $\bar{a}^2 + \bar{b}^2 = \rho^2/4$ , we have  $\gamma_1^\dagger = 0$ ,  $V_1(0) + V_2(0) = 0$  and  $\lim_{\gamma \rightarrow 0} \gamma^2/|\Delta^\dagger|^2 = \lim_{\gamma \rightarrow 0} \gamma^2/(\gamma^2\bar{b}^4) = 1/\bar{b}^4$ . So  $\lim_{\gamma \rightarrow 0} v_{\text{ncc}}(\gamma) = \Pi D^2/\rho$ .  $\square$

Theorem 5.4 shows that cost blowup occurs at  $\gamma_2^\dagger$  if  $\gamma_2^\dagger > 1$ , and  $\gamma_1^\dagger$  is a removable pole of the cost function. For the socially optimal solution, there is no blowup effect for the asymptotic average social optimum  $\lim_{N \rightarrow \infty} (1/N) \inf_u J_{\text{soc}}^{(N)}|_{x(0)=0}$  as  $\gamma$  approaches any given value. To show this fact, we take  $x_i(0) \equiv 0$  for all  $i$ . By (19)

$$\begin{aligned} \bar{v}(\gamma) &:= \lim_{N \rightarrow \infty} \left( \frac{1}{N} \right) \inf_u J_{\text{soc}}^{(N)}|_{x(0)=0} \\ &= \frac{\Pi D^2}{\rho} + \frac{\gamma^2 \eta_0^2}{\rho} - \frac{\frac{\gamma \eta_0 \bar{b}^2 (\gamma - 1)}{\rho}}{\left( \frac{\rho}{2} \right) + \sqrt{\bar{a}^2 + (1 - \gamma)^2 \bar{b}^2}}. \end{aligned}$$

Clearly,  $\bar{v}(\gamma)$  is continuous and bounded on  $(-\infty, 1) \cup (1, \infty)$  and  $\lim_{\gamma \rightarrow 1} v(\gamma) = (\Pi D^2 + \eta_0^2)/\rho$ .

## VI. ASYMPTOTIC AVERAGE SOCIAL OPTIMUM

Now we give a closed form expression of the asymptotic average social optimum in terms of the solution of the SCE equation system. Denote the ODE

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

By Lemma A.1 there exists a unique  $q_\theta(0)$  such that  $q_\theta \in C_\rho([0, \infty), \mathbb{R})$ . Let  $\hat{u} = (\hat{u}_1, \dots, \hat{u}_N)$  be the set of SCE-based control laws given by (32) and  $\hat{x}_i$  correspond to  $\hat{u}_i$ .

**Lemma 6.1:** Assume (A1)–(A5) and let  $J_i$  be defined by (2). Then

$$\max_{1 \leq i \leq N} \left| J_i(\hat{u}) - E \int_0^\infty e^{-\rho t} \hat{L}(\hat{x}_i, \bar{x}, \hat{u}_i) dt \right| = O\left( \frac{1}{\sqrt{N}} + \epsilon_N \right)$$

where  $\epsilon_N$  is defined by (37) and

$$\hat{L}(\hat{x}_i, \bar{x}, \hat{u}_i) = \hat{x}_i^T Q \hat{x}_i - 2(\Gamma \bar{x} + \eta)^T Q \hat{x}_i + |\Gamma \bar{x} + \eta|_Q^2 + \hat{u}_i^T R \hat{u}_i.$$

*Proof:* Let  $\hat{x}^{(N)} = (1/N) \sum_{i=1}^N \hat{x}_i$  in  $J_i(\hat{u})$  be approximated by  $\bar{x}$ . Next, by Schwarz's inequality, (C.4) and Lemma 4.1, we obtain the desired estimate.  $\square$

Let  $L^*$  be defined by (29). We have the relation

$$\hat{L}(\hat{x}_i, \bar{x}, \hat{u}_i) = L^*(\hat{x}_i, \bar{x}, \hat{u}_i) + 2[(I - \Gamma)\bar{x} - \eta]^T Q \Gamma \hat{x}_i + |\Gamma \bar{x} + \eta|_Q^2. \quad (75)$$

**Theorem 6.2:** Assume i) the conditions of Theorem 4.2 hold; ii)  $\{x_i, 1 \leq i \leq N\}$  have the same initial mean  $m_0$  and initial

covariance  $\Sigma_0 = \text{Cov}(x_i(0))$ . Then the asymptotic average social optimum is given by

$$\begin{aligned} & \lim_{N \rightarrow \infty} \inf_{u \in \mathcal{U}_o} \left( \frac{1}{N} \right) J_{\text{soc}}^{(N)} \\ &= \lim_{N \rightarrow \infty} \left( \frac{1}{N} \right) J_{\text{soc}}^{(N)}(\hat{u}) \\ &= m_0^T \left[ \int \Pi_\theta dF(\theta) \right] m_0 + 2m_0^T \int s_\theta(0) dF(\theta) \quad (=: f_{m_0, s}) \\ &+ \text{Tr} \left[ \Sigma_0 \int \Pi_\theta dF(\theta) \right] + \int q_\theta(0) dF(\theta) \quad (=: f_{\Sigma_0, q}) \\ &+ \int_0^\infty e^{-\rho t} \bar{x}^T (\Gamma^T Q + Q \Gamma - \Gamma^T Q \Gamma) \bar{x} dt + \left( \frac{1}{\rho} \right) \eta^T Q \eta. \end{aligned}$$

*Proof:* By Theorem 4.2 and Lemma 6.1, it suffices to find  $\lim_{N \rightarrow \infty} (1/N) \sum_{i=1}^N E \int_0^\infty e^{-\rho t} \hat{L}(\hat{x}_i, \bar{x}, \hat{u}_i) dt$ . Let  $J_{\text{opt}}^*(\hat{u}_i)$  be the optimal cost of Problem (P2) with its optimal control law  $\hat{u}_i$  where  $\bar{x}$  is now given by the SCE equation system. By (75)

$$\begin{aligned} & \left( \frac{1}{N} \right) \sum_{i=1}^N E \int_0^\infty e^{-\rho t} \hat{L}(\hat{x}_i, \bar{x}, \hat{u}_i) dt \\ &= \left( \frac{1}{N} \right) \sum_{i=1}^N J_{\text{opt}}^*(\hat{u}_i) \\ &+ 2E \int_0^\infty e^{-\rho t} [(I - \Gamma)\bar{x} - \eta]^T Q \Gamma \hat{x}^{(N)} dt \\ &+ \int_0^\infty e^{-\rho t} |\Gamma \bar{x} + \eta|_Q^2 dt. \end{aligned}$$

By Schwarz's inequality and Lemma 4.1

$$\begin{aligned} & \left| E \int_0^\infty e^{-\rho t} [(I - \Gamma)\bar{x} - \eta]^T Q \Gamma \hat{x}^{(N)} dt \right. \\ & \quad \left. - E \int_0^\infty e^{-\rho t} [(I - \Gamma)\bar{x} - \eta]^T Q \Gamma \bar{x} dt \right|^2 \\ & \leq C \int_0^\infty e^{-\rho t} |(I - \Gamma)\bar{x} - \eta|^2 dt \cdot E \int_0^\infty e^{-\rho t} |\hat{x}^{(N)} - \bar{x}|^2 dt \\ & \leq C \left( \frac{1}{N} + \epsilon_N^2 \right) \end{aligned} \quad (76)$$

where  $\bar{x} \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$ . By (76)

$$\begin{aligned} & 2 \lim_{N \rightarrow \infty} E \int_0^\infty e^{-\rho t} [(I - \Gamma)\bar{x} - \eta]^T Q \Gamma \hat{x}^{(N)} dt \\ & \quad + \int_0^\infty e^{-\rho t} |\Gamma \bar{x} + \eta|_Q^2 dt \\ &= \int_0^\infty e^{-\rho t} \bar{x}^T (\Gamma^T Q + Q \Gamma - \Gamma^T Q \Gamma) \bar{x} dt + \left( \frac{1}{\rho} \right) \eta^T Q \eta. \end{aligned}$$

By the continuity of  $\Pi_\theta$  and  $s_\theta(t)$  for each fixed  $t$ , we may use (A.8) to show that  $q_\theta(0)$  is continuous in  $\theta$ . By the weak convergence of  $F_N$  to  $F$ , we obtain from Lemma A.1-3) that  $\lim_{N \rightarrow \infty} (1/N) \sum_{i=1}^N J_{\text{opt}}^*(\hat{u}_i) = f_{m_0, s} + f_{\Sigma_0, q}$ , and the theorem follows.  $\square$

*Remark:* If we apply Theorem 6.2 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  $(s, \bar{x})$  in Theorem 5.2. By lengthy but elementary calculations, we may further verify that the result coincides with (19).  $\diamond$

*Remark:* In a system of uniform agents with vector states, one may solve the SCE equation system and explicitly calculate the asymptotic average social optimum. For such a model, the standard LQG solution faces substantial difficulty to evaluate the performance.  $\diamond$

## VII. COST COMPARISON OF DIFFERENT SOLUTIONS

We now provide a comparison of the costs based on the scalar model in Section V. For simplicity, we suppose  $x_i(0) \equiv 0$  for each agent, and the randomness comes only from the noises. The parameters are given by  $[a, b, D, \rho, r, \eta_0] = [1, 1, 1, 0.8, 0.25, 0.2]$ .

We first consider a population of  $N$  agents. Let  $J_{\text{soc}}^{(N)}(\hat{u}^{\text{sce}})$  be the cost of the set of SCE based control laws  $\hat{u}^{\text{sce}}$ , and  $J_{\text{soc}}^{(N)}(u^*)$  be the cost of the centralized optimal control law  $u^*$ . By Proposition 3.2 and  $x_i(0) \equiv 0$ ,  $J_{\text{soc}}^{(N)}(u^*) = s_0$ .

Under  $\hat{u}^{\text{sce}}$ , we have  $dx_i = -\beta_1 x_i dt - \bar{b}^2 s dt + D dW_i$ ,  $1 \leq i \leq N$ . By (5),  $J_{\text{soc}}^{(N)}(\hat{u}^{\text{sce}}) = \int_0^\infty e^{-\rho t} L(t, x(t)) dt$ , where

$$\begin{aligned} L(t, x) &= x^T (\hat{Q} + \bar{b}^2 \Pi^2 I) x + 2(G + \bar{b}^2 \Pi s \mathbf{1}_N^T) x \\ & \quad + N(\bar{b}^2 s^2 + \gamma^2 \eta_0^2). \end{aligned}$$

To evaluate  $J_{\text{soc}}^{(N)}(\hat{u}^{\text{sce}})$ , we consider a general initial condition  $x(t) = x = [x_1, \dots, x_N]^T$  at time  $t \geq 0$  for the  $N$  agents, and denote  $v^{\text{sce}, N}(t, x) = E \left[ \int_t^\infty e^{-\rho(\tau-t)} L(\tau, x(\tau)) d\tau \mid x(t) = x \right]$ . By

$$\begin{aligned} v^{\text{sce}, N}(t, x) &= E \left[ \int_t^{t+\delta t} e^{-\rho(\tau-t)} L(\tau, x(\tau)) d\tau \right. \\ & \quad \left. + e^{-\rho \delta t} v^{\text{sce}, N}(t + \delta t, x(t + \delta t)) \mid x(t) = x \right], \quad \delta t \geq 0 \end{aligned}$$

and Ito's formula for  $v^{\text{sce}, N}(\tau, x(\tau))$ , we obtain

$$\begin{aligned} \rho v^{\text{sce}, N} &= v_t^{\text{sce}, N} - (\beta_1 x^T + \bar{b}^2 s \mathbf{1}_N^T) v_x^{\text{sce}, N} \\ & \quad + \left( \frac{D^2}{2} \right) \text{Tr}(v_{xx}^{\text{sce}, N}) + L(t, x). \end{aligned} \quad (77)$$

By writing  $v^{\text{sce}, N}(t, x) = x^T \bar{P} x + 2h_1^T(t)x + h_0(t)$  and substituting it into (77), we determine  $\bar{P} = (\hat{Q} + \bar{b}^2 \Pi^2 I)/(\rho + 2\beta_1)$  and further obtain the ordinary differential equations for  $h_1$  and  $h_0$ , respectively, which can be uniquely solved in explicit forms subject to the growth conditions  $h_1 \in C_{\rho/2}([0, \infty), \mathbb{R}^N)$  and  $h_0 \in C_\rho([0, \infty), \mathbb{R})$ . We obtain  $J_{\text{soc}}^{(N)}(\hat{u}^{\text{sce}}) = v^{\text{sce}, N}(0, 0) = h_0(0)$  due to  $x(0) = 0$ . Fig. 1(a) shows  $(1/N) J_{\text{soc}}^{(N)}(\hat{u}^{\text{sce}})$ ,  $(1/N) J_{\text{soc}}^{(N)}(u^*)$  and the cost difference  $(1/N)(J_{\text{soc}}^{(N)}(\hat{u}^{\text{sce}}) - J_{\text{soc}}^{(N)}(u^*))$  as a measure of the optimality loss per agent, where  $\gamma = 0.95$  and  $N$  increases from 2 to 50.

Next, we compare the per agent costs  $\bar{v}_{\text{sce}}(\gamma)$  and  $v_{\text{ncc}}(\gamma)$  within the population limit based on the SCE and NCE solutions determining the mean field, respectively. By the remark after Theorem 6.2, we see that  $\bar{v}_{\text{sce}}(\gamma) = (D^2 \Pi + \gamma^2 \eta_0^2 - \bar{b}^2 s_{11}^2)/\rho$ . In fact, when  $\gamma = 1$ , Theorem 5.2-1) still holds since the above parameters ensure  $\Delta \neq 0$  and (65), and the SCE based control law still gives a cost expression as above. For the NCE based solution, we have  $\gamma_1^\dagger = 1.05$  and  $\gamma_2^\dagger = 1.09$ . Fig. 1(b) shows  $\bar{v}_{\text{sce}}(\gamma)$ ,  $v_{\text{ncc}}(\gamma)$  and  $v_{\text{ncc}}(\gamma) - \bar{v}_{\text{sce}}(\gamma)$ , where  $v_{\text{ncc}}(\gamma)$  has a blowup at  $\gamma = 1.09$ .

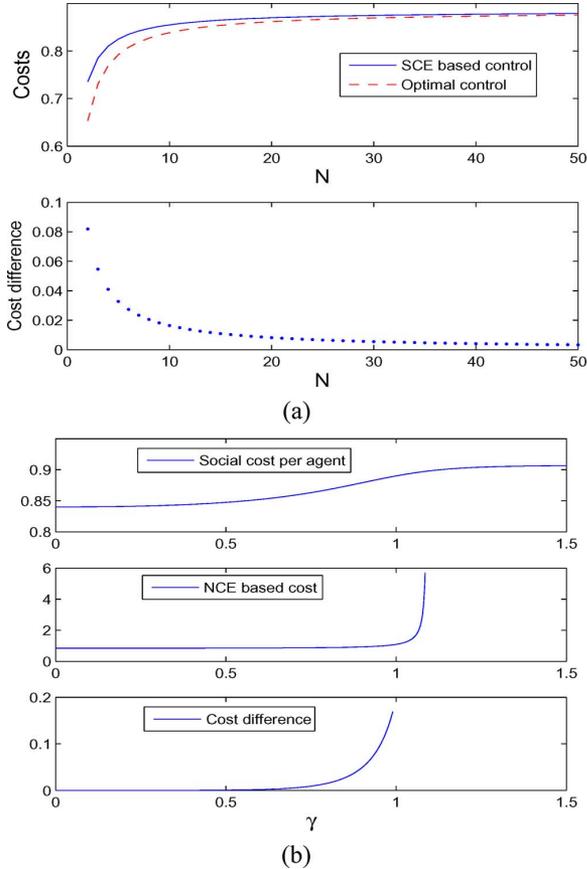


Fig. 1. (a) Costs for a finite population of  $N$  agents. (b) Costs as a function of  $\gamma$  in the population limit.

## VIII. CONCLUSION

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 distribution of the dynamic parameters across the population to solve an ODE system offline and then uses its own state to construct a feedback strategy. An interesting generalization is to consider models quantifying the level of willingness of individual agents in cooperation. This may be achieved by letting each agent optimize a cost as a convex combination of a social component and a selfish component, where the weights of the two components vary continuously over the population. This modeling will enable us to characterize the impact of cooperation willingness on the formation of mean field behavior. Also, the social variational approach has the potential to be applied to general nonlinear models to construct limiting optimal control problems for decentralized control synthesis.

## APPENDIX A

### THE OPTIMAL CONTROL LEMMA

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$  be an underlying filtered probability space. Consider

$$dx(t) = Ax(t)dt + Bu(t)dt + f(t)dt + DdW(t), \quad t \geq 0 \quad (\text{A.1})$$

where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^{m_1}$ ,  $f \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$ , and  $W(t)$  is an  $n_2$  dimensional standard Brownian motion adapted to  $\mathcal{F}_t$ . The initial condition  $x(0)$  is independent of  $W(t)$  and

$E|x(0)|^2 < \infty$ . The admissible control set  $\mathcal{U}$  consists of all controls  $u(\cdot)$  adapted to  $\mathcal{F}_t$  with  $\int_0^\infty e^{-\rho t} |u(t)|^2 dt < \infty$ . For  $u(\cdot) \in \mathcal{U}$ , let the cost function be given by

$$J(u(\cdot)) = E \int_0^\infty e^{-\rho t} (x^T Q x - 2g^T x + u^T R u) dt \quad (\text{A.2})$$

where  $\rho > 0$ ,  $Q \geq 0$ ,  $R > 0$  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

$$\rho \Pi = \Pi A + A^T \Pi - \Pi B R^{-1} B^T \Pi + Q. \quad (\text{A.3})$$

Denote the ODEs

$$\rho s = \frac{ds}{dt} + (A^T - \Pi B R^{-1} B^T) s + \Pi f - g, \quad t \geq 0 \quad (\text{A.4})$$

$$\rho q = \frac{dq}{dt} - s^T B R^{-1} B^T s + 2f^T s + \text{Tr}(D^T \Pi D) \quad (\text{A.5})$$

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 A.1:* For the optimal control problem (A.1)–(A.2), assume 1) the pair  $[Q^{1/2}, A - (\rho/2)I]$  is detectable and  $[A - (\rho/2)I, B]$  is stabilizable which together give a solution  $\Pi \geq 0$  for (A.3) such that  $A - B R^{-1} B^T \Pi - (\rho/2)I$  is asymptotically stable, and 2) both  $f$  and  $g$  are in  $C_{\rho/2}([0, \infty), \mathbb{R}^n)$ . Then we have

- 1) there exists a unique solution  $s \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$  to (A.4);
- 2) the optimal control law is given by  $\hat{u}(t) = -R^{-1} B^T [\Pi x(t) + s(t)]$ ;
- 3) there exists a unique solution  $q \in C_\rho([0, \infty), \mathbb{R})$  to (A.5). The optimal cost is given by

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

*Proof:* We may show part 1) using the method in [11, Lemma A.2]. We prove part 2) by first obtaining a prior integral estimate of  $x$  [see (A.6)] and then using a completion of squares technique. Compared with [11], the cost integrand in (A.2) does not necessarily allow rewriting the term  $x^T Q x - 2g^T x$  in the form  $(x - g_1)^T Q (x - g_1) + h_1$  for some functions  $g_1$  and  $h_1$ . For  $u \in \mathcal{U}$ , we show that  $J(u) \leq C_0$  for some prior upper bound  $C_0$  implies

$$E \int_0^\infty e^{-\rho t} |x|^2 dt < \infty \quad (\text{A.6})$$

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

$$E \int_0^\infty e^{-\rho t} (x^T Q x - \epsilon x^T x) dt \leq C_1. \quad (\text{A.7})$$

If necessary, we may apply a nonsingular linear transformation. Here without loss of generality we simply assume  $A = \text{Diag}[A_{11}, A_{22}]$ , 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  $x_2$  correspond to the sub-matrices  $A_{11}$  and  $A_{22}$ , respectively, in the dynamics. We write  $Q^{1/2} = [M_1, M_2]$  so that  $x^T Q x =$

$|M_1x_1 + M_2x_2|^2$ , where  $[M_1, A_{11}]$  is observable due to the detectability of  $[Q^{1/2}, A - (\rho/2)I]$ . By [11, Lemma A.1] and  $E \int_0^\infty e^{-\rho t} |u|^2 dt < \infty$ , it follows that  $E \int_0^\infty e^{-\rho t} |x_2|^2 dt < \infty$ , which combined with (A.7) implies  $E \int_0^\infty e^{-\rho t} (x_1^T M_1^T M_1 x_1 - \epsilon |x_1|^2) dt < \infty$ . 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  $\epsilon$ , such that  $E \int_0^\infty e^{-\rho t} x_1^T M_1^T M_1 x_1 dt \geq c_1 E \int_0^\infty e^{-\rho t} |x_1|^2 dt - c_2$  (see [11] for a similar argument). Hence,  $E \int_0^\infty e^{-\rho t} (c_1 - \epsilon) |x_1|^2 dt < \infty$ . By taking a sufficiently small  $\epsilon$  such that  $c_1 - \epsilon > 0$ , (A.6) follows. The rest part of the proof of part 2) is similar to [11, Lemma A.2] and is omitted.

We prove 3) as follows. Let the initial condition of (A.5) be  $q(0)$  to give

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

It is straightforward to show that  $q \in C_\rho([0, \infty), \mathbb{R})$  if and only if

$$\begin{aligned} q(0) &= \int_0^\infty e^{-\rho \tau} (2f^T s + \text{Tr}(D^T \Pi D) - s^T B R^{-1} B^T s) d\tau \\ &= \left(\frac{1}{\rho}\right) \text{Tr}(D^T \Pi D) + \int_0^\infty e^{-\rho \tau} (2f^T s - s^T B R^{-1} B^T s) d\tau \end{aligned} \quad (\text{A.8})$$

where the integral in (A.8) converges. Next, we apply Ito's formula to  $e^{-\rho t} (x^T \Pi x + 2s^T x + q)$  to obtain the expression of  $J(\hat{u})$ , where  $x$  is the closed-loop solution under  $\hat{u}$ .  $\square$

*Remark:* If  $f$  and  $g$  are constant vectors, we obtain the unique solutions to (A.4)-(A.5) with the required growth conditions as a constant vector  $s$  and a constant  $q$ , which in turn are determined by

$$\begin{aligned} \rho s &= (A^T - \Pi B R^{-1} B^T) s + \Pi f - g, \\ \rho q &= 2f^T s - s^T B R^{-1} B^T s + \text{Tr}(D^T \Pi D). \end{aligned}$$

$\diamond$

## APPENDIX B

*Proof of Proposition 3.2:* When  $x(0)$  is a random variable, (18) is obtained from (7) by calculating  $E v(x(0)) = E[x^T(0) P x(0) + 2s_1^T x(0) + s_0]$ . By independence of  $x_1(0), \dots, x_N(0)$

$$\begin{aligned} E[x^T(0) P x(0)] &= \text{Tr}(P E[x(0) x^T(0)]) \\ &= \text{Tr}\{P(\sigma^2 I_N + m_0^2 \mathbf{1}_N \mathbf{1}_N^T)\} \\ &= \sigma^2 \text{Tr}(P) + m_0^2 \mathbf{1}_N^T P \mathbf{1}_N \end{aligned}$$

and (18) follows readily. Denote  $P = (p_{ij})_{1 \leq i, j \leq N}$ . We have  $\text{Tr}(P) = N p_N$  and  $\mathbf{1}_N^T P \mathbf{1}_N = \sum_{i, j} p_{ij} = N p_N + (N^2 - N) q_N$ . Hence,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{[\sigma_0^2 \text{Tr}(P) + m_0^2 \mathbf{1}_N^T P \mathbf{1}_N + 2N m_0 s_{11}]}{N} \\ = \sigma_0^2 p_\infty + m_0^2 \frac{[\bar{a} + \sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2}]}{\bar{b}^2} + 2m_0 s_{11}. \end{aligned}$$

Next, we use (14)–(16) to compute  $s_0$  and subsequently  $\lim_{N \rightarrow \infty} s_0/N$ . These calculations give the expression (19).  $\square$

*Proof of Proposition 3.3:* By adding up the  $N$  equations given by (20) with  $i \leq N$  and multiplying both sides of the resulting equation by  $1/N$ , we obtain

$$\begin{aligned} dx^{(N)} &= \left[ \left(\frac{\rho}{2}\right) - \sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2} \right] x^{(N)} dt - \bar{b}^2 s_{11} dt \\ &\quad + \left(\frac{1}{N}\right) \sum_{i=1}^N D dW_i. \end{aligned}$$

Since  $(\rho/2) - \sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2} < \hat{\rho}/2$ , we may use the SDE of  $e^{-(\hat{\rho}/2)t} x^{(N)}$  to verify

$$\sup_{N \geq 1} \sup_{t \geq 0} e^{-\hat{\rho} t} E |x^{(N)}|^2 < \infty. \quad (\text{B.1})$$

Denote  $\hat{x}^{(N)}(t) = x^{(N)}(t) - \bar{x}_c(t)$  and  $\tilde{x}_\rho^{(N)}(t) = e^{-(\hat{\rho}/2)t} \hat{x}^{(N)}(t)$ . It follows that

$$\begin{aligned} d\tilde{x}_\rho^{(N)} &= \left[ \frac{(\rho - \hat{\rho})}{2} - \sqrt{\bar{a}^2 + (1-\gamma)^2 \bar{b}^2} \right] \tilde{x}_\rho^{(N)} dt \\ &\quad + e^{-(\hat{\rho}/2)t} \left(\frac{1}{N}\right) \sum_{i=1}^N D dW_i \end{aligned}$$

which gives  $\sup_{t \geq 0} e^{-\hat{\rho} t} E |x^{(N)}(t) - \bar{x}_c(t)|^2 = O(1/N)$ . By using the SDE of  $e^{-(\hat{\rho}/2)t} (x_i - y_i)$  and (B.1), we may further show  $\sup_{t \geq 0} e^{-\hat{\rho} t} E |x_i(t) - y_i(t)|^2 = O(1/N)$  since  $a - \bar{b}^2 p_\infty - \hat{\rho}/2 = (\rho - \hat{\rho})/2 - \sqrt{\bar{a}^2 + \bar{b}^2} < 0$ .  $\square$

*Proof of Theorem 4.3:* We first find a solution satisfying the growth condition. Denote the ARE

$$\bar{P} A_\rho + A_\rho^T \bar{P} - \bar{P} B R^{-1} B^T \bar{P} - Q_\Gamma = 0 \quad (\text{B.2})$$

where  $A_\rho = A - B R^{-1} B^T \Pi - (\rho/2)I$ . Equation (B.2) has a unique solution  $\bar{P} \geq 0$  since  $A_\rho$  is asymptotically stable, and furthermore,  $A_\rho - B R^{-1} B^T \bar{P}$  is also asymptotically stable (see [32]). Denote  $s = \bar{P} \bar{x} + \phi$ , where  $\phi$  is a function of  $t$ . The left-hand side of (43) becomes

$$\begin{aligned} \frac{ds}{dt} &= \bar{P} \left( \frac{d\bar{x}}{dt} \right) + \frac{d\phi}{dt} \\ &= \bar{P} \left[ A_\rho + \left(\frac{\rho}{2}\right) I \right] \bar{x} - \bar{P} B R^{-1} B^T (\bar{P} \bar{x} + \phi) + \frac{d\phi}{dt}. \end{aligned} \quad (\text{B.3})$$

The right-hand side of (43) is now expressed as

$$\begin{aligned} Q_\Gamma \bar{x} + (\rho I - A^T + \Pi B R^{-1} B^T) s + \eta_\Gamma \\ = Q_\Gamma \bar{x} + \left[ -A_\rho^T + \left(\frac{\rho}{2}\right) I \right] (\bar{P} \bar{x} + \phi) + \eta_\Gamma. \end{aligned} \quad (\text{B.4})$$

Letting the right-hand sides of both (B.3) and (B.4) be equal, we obtain

$$\frac{d\phi}{dt} = - \left[ A_\rho - B R^{-1} B^T \bar{P} - \left(\frac{\rho}{2}\right) I \right]^T \phi + \eta_\Gamma \quad (\text{B.5})$$

where each eigenvalue of  $-[A_\rho - B R^{-1} B^T \bar{P} - (\rho/2)I]$  has a real part greater than  $\rho/2$ . We may identify a unique  $\phi(0)$

such that  $\phi(t) \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$  and the resulting solution is a constant vector

$$\phi = \left[ A_\rho^T - \bar{P}BR^{-1}B^T - \left(\frac{\rho}{2}\right)I \right]^{-1} \eta_T. \quad (\text{B.6})$$

Now (42) gives

$$\frac{d\bar{x}}{dt} = \left[ A_\rho - BR^{-1}B^T\bar{P} + \left(\frac{\rho}{2}\right)I \right] \bar{x} - BR^{-1}B^T\phi \quad (\text{B.7})$$

where  $\phi$  is given by (B.6). Since  $\bar{x}(0) = m_0$  and each eigenvalue of  $A_\rho - BR^{-1}B^T\bar{P} + (\rho/2)I$  has a real part less than  $\rho/2$ , we obtain  $\bar{x}(t) \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$ . Subsequently,  $s(t) = \bar{P}\bar{x}(t) + \phi \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$ . Hence, we have obtained a solution pair  $(s, \bar{x})$  satisfying the growth condition.

To show uniqueness, suppose that  $(s', \bar{x}')$  gives another solution within  $C_{\rho/2}([0, \infty), \mathbb{R}^{2n})$ . Denote  $s' = \bar{P}\bar{x}' + \phi'$ . So  $\phi' \in C_{\rho/2}([0, \infty), \mathbb{R}^n)$ . It can be shown that  $\phi'$  satisfies (B.5) and is necessarily given by (B.6). Subsequently,  $\bar{x}'$  satisfies the same equation (B.7) with  $\bar{x}'(0) = m_0$  as  $\bar{x}$  does. Hence,  $\bar{x}' \equiv \bar{x}$ , which further implies that  $s' \equiv s$ .  $\square$

#### APPENDIX C

##### PROOF OF THEOREM 4.2

The admissible control set  $\mathcal{U}_o$  is defined as in Section II-B.

*Lemma C.1:* Suppose that  $Q > 0$  and  $I - \Gamma$  is a nonsingular matrix. For  $u \in \mathcal{U}_o$ ,  $J_{\text{soc}}^{(N)}(u) < \infty$  implies that  $E \int_0^\infty e^{-\rho t} (|x_i|^2 + |u_i|^2) dt < \infty$  for all  $1 \leq i \leq N$ .

*Proof:* Since  $R$  is nonsingular,  $J_{\text{soc}}^{(N)}(u) < \infty$  leads to  $E \int_0^\infty e^{-\rho t} |u_i|^2 dt < \infty$ . We also obtain  $E \int_0^\infty e^{-\rho t} |x_i - \Gamma x^{(N)} - \eta|_Q^2 dt < \infty$ . Since  $Q$  is nonsingular, this yields

$$E \int_0^\infty e^{-\rho t} \left| x_i - \Gamma x^{(N)} \right|^2 dt < \infty \quad (\text{C.1})$$

for each  $i$ , which further implies

$$\begin{aligned} E \int_0^\infty e^{-\rho t} \left| x^{(N)} - \Gamma x^{(N)} \right|^2 dt \\ \leq \left( \frac{1}{N} \right) \sum_{i=1}^N E \int_0^\infty e^{-\rho t} \left| x_i - \Gamma x^{(N)} \right|^2 dt < \infty. \end{aligned} \quad (\text{C.2})$$

Since  $I - \Gamma$  is nonsingular,  $E \int_0^\infty e^{-\rho t} |x^{(N)}|^2 dt < \infty$ , so that  $E \int_0^\infty e^{-\rho t} |x_i|^2 dt < \infty$  by (C.1).  $\square$

*Proof of Theorem 4.2:*

*Step 1 (Prior Upper Bounds of the Cost and State):*

Note that  $Q > 0$  implies (A4)-ii). Hence, it is ensured by (A4) that (30) has a unique solution  $\Pi_\theta \geq 0$  and that  $A(\theta) - BR^{-1}B^T\Pi_\theta - (\rho/2)I$  is asymptotically stable. Consider the closed-loop system

$$d\hat{x}_i = [A(\theta_i) - BR^{-1}B^T\Pi_{\theta_i}] \hat{x}_i dt - BR^{-1}B^T s_{\theta_i} dt + DdW_i \quad (\text{C.3})$$

where  $\hat{x}_i(0) = x_i(0)$ . By the eigenvalue distribution of  $A(\theta) - BR^{-1}B^T\Pi_\theta - \rho I$  in Lemma 4.1 and (39)–(40), there exist constants  $c_0, \rho''$  both independent of  $(i, N)$  such that

$$E \int_0^\infty e^{-\rho t} (|\hat{x}_i|^2 + |\hat{u}_i|^2) dt \leq c_0 \quad (\text{C.4})$$

$$\sup_N \sup_{1 \leq i \leq N} E |\hat{x}_i(t)|^2 = O(e^{\rho'' t}) \quad (\text{C.5})$$

where  $0 \leq \rho'' < \rho$ . By (C.4), we obtain  $(1/N)J_{\text{soc}}^{(N)}(\hat{u}) \leq C_0$  for some  $C_0$  independent of  $N$ .

It suffices to consider all  $u = (u_1, \dots, u_N) \in \mathcal{U}_o$  such that

$$\left( \frac{1}{N} \right) J_{\text{soc}}^{(N)}(u) \leq \left( \frac{1}{N} \right) J_{\text{soc}}^{(N)}(\hat{u}) \leq C_0. \quad (\text{C.6})$$

Given  $u \in \mathcal{U}_o$  satisfying (C.6), we have

$$\begin{aligned} dx_i = [A(\theta_i) - BR^{-1}B^T\Pi_{\theta_i}] x_i dt - BR^{-1}B^T s_{\theta_i} dt \\ + B [u_i + R^{-1}B^T(\Pi_{\theta_i} x_i + s_{\theta_i})] dt + DdW_i. \end{aligned} \quad (\text{C.7})$$

Denote  $\tilde{x}_i = x_i - \hat{x}_i$  and  $\tilde{u}_i = u_i + R^{-1}B^T(\Pi_{\theta_i} x_i + s_{\theta_i})$ . By Lemma C.1

$$E \int_0^\infty e^{-\rho t} (|x_i|^2 + |\tilde{x}_i|^2 + |u_i|^2 + |\tilde{u}_i|^2) dt < \infty. \quad (\text{C.8})$$

It follows from (C.3) and (C.7) that

$$d\tilde{x}_i = [A(\theta_i) - BR^{-1}B^T\Pi_{\theta_i}] \tilde{x}_i dt + B\tilde{u}_i dt \quad (\text{C.9})$$

where  $\tilde{x}_i(0) = 0$ . By (C.9) and (C.8) for  $\tilde{u}_i$ , we may use the ODE of  $e^{-(\rho/2)t}\tilde{x}_i$  to show

$$\sup_{1 \leq i \leq N} E |\tilde{x}_i(t)|^2 = O(e^{\rho t}) \quad (\text{C.10})$$

where the right-hand side is allowed to change with  $N$  since it is only shown that  $E \int_0^\infty e^{-\rho t} |\tilde{u}_i|^2 dt < \infty$ .

*Step 2 (Cost Integrand Approximation):* Note that

$$\begin{aligned} J_{\text{soc}}^{(N)}(u) = E \\ \cdot \int_0^\infty e^{-\rho t} \left\{ \sum_{i=1}^N |x_i - \Gamma x^{(N)} - \eta|_Q^2 + \sum_{i=1}^N u_i^T R u_i \right\} dt. \end{aligned} \quad (\text{C.11})$$

Denote  $\hat{x}^{(N)} = (1/N) \sum_{j=1}^N \hat{x}_j$  and  $\tilde{x}^{(N)} = (1/N) \sum_{j=1}^N \tilde{x}_j$ . The first summation in (C.11) gives

$$\begin{aligned} \sum_{i=1}^N |x_i - \Gamma x^{(N)} - \eta|_Q^2 \\ = \sum_{i=1}^N \left| \hat{x}_i - \Gamma \hat{x}^{(N)} - \eta + \tilde{x}_i - \Gamma \tilde{x}^{(N)} \right|_Q^2 \\ = \sum_{i=1}^N \left| \hat{x}_i - \Gamma \hat{x}^{(N)} - \eta \right|_Q^2 \quad (=:\zeta_1) \\ + \sum_{i=1}^N \left| \tilde{x}_i - \Gamma \tilde{x}^{(N)} \right|_Q^2 \quad (=:\zeta_2) \\ + \sum_{i=1}^N 2 \left( \hat{x}_i - \Gamma \hat{x}^{(N)} - \eta \right)^T Q \left( \tilde{x}_i - \Gamma \tilde{x}^{(N)} \right) \quad (=:\zeta_3). \end{aligned} \quad (\text{C.12})$$

Notice that  $\zeta_1$  does not change with  $u$ . Since  $\zeta_1 \geq 0$ , for  $u$  satisfying (C.6) we have

$$E \int_0^\infty e^{-\rho t} (\zeta_2 + \zeta_3) dt \leq J_{\text{soc}}^{(N)}(u) \leq C_0 N. \quad (\text{C.13})$$

By the inequality  $2\alpha^T\beta \geq -2\alpha^T\alpha - (1/2)\beta^T\beta$  for two real-valued vectors  $\alpha$  and  $\beta$ , we obtain

$$\begin{aligned} \zeta_3 &\geq -2 \sum_{i=1}^N \left| \left( \hat{x}_i - \Gamma \hat{x}^{(N)} - \eta \right)^T Q^{1/2} \right|^2 \\ &\quad - \left( \frac{1}{2} \right) \sum_{i=1}^N \left| Q^{1/2} \left( \hat{x}_i - \Gamma \hat{x}^{(N)} \right) \right|^2 \\ &= -2\zeta_1 - \left( \frac{\zeta_2}{2} \right). \end{aligned} \quad (\text{C.14})$$

It follows from (C.13)–(C.14) that  $E \int_0^\infty e^{-\rho t} [(\zeta_2/2) - 2\zeta_1] dt \leq C_0 N$ .

On the other hand, by (C.4) we may show that  $E \int_0^\infty e^{-\rho t} \zeta_1 dt \leq CN$ . So for  $u$  satisfying (C.6)

$$E \int_0^\infty e^{-\rho t} \zeta_2 dt \leq CN. \quad (\text{C.15})$$

By using (C.15), we obtain

$$\begin{aligned} &E \int_0^\infty e^{-\rho t} \left| \hat{x}^{(N)} - \Gamma \hat{x}^{(N)} \right|^2 dt \\ &= E \int_0^\infty e^{-\rho t} \left| \left( \frac{1}{N} \right) \sum_{i=1}^N \left( \hat{x}_i - \Gamma \hat{x}^{(N)} \right) \right|^2 dt \\ &\leq E \int_0^\infty e^{-\rho t} \left( \frac{1}{N} \right) \sum_{i=1}^N \left| \hat{x}_i - \Gamma \hat{x}^{(N)} \right|^2 dt \\ &\leq \left( \frac{C}{N} \right) E \int_0^\infty e^{-\rho t} \zeta_2 dt \leq C \end{aligned}$$

which further implies that for some  $C$  independent of  $N$ ,

$$E \int_0^\infty e^{-\rho t} \left| \hat{x}^{(N)} \right|^2 dt \leq C. \quad (\text{C.16})$$

By replacing  $\hat{x}^{(N)}$  by  $\bar{x}$  in  $\zeta_3$ , denote

$$\begin{aligned} \zeta_3' &= \sum_{i=1}^N 2 \left( \hat{x}_i - \Gamma \bar{x} - \eta \right)^T Q \left( \hat{x}_i - \Gamma \hat{x}^{(N)} \right) \\ &= \sum_{i=1}^N 2 \left( \hat{x}_i - \Gamma \bar{x} - \eta \right)^T Q \hat{x}_i - 2 \left( \hat{x}^{(N)} - \Gamma \bar{x} - \eta \right)^T Q \Gamma \sum_{j=1}^N \hat{x}_j \\ &= 2 \sum_{i=1}^N \hat{x}_i^T \left( Q \hat{x}_i - Q \Gamma \bar{x} - \Gamma^T Q \bar{x} + \Gamma^T Q \Gamma \bar{x} - Q \eta + \Gamma^T Q \eta \right) \\ &= (\zeta_3'') + 2 \sum_{i=1}^N \hat{x}_i^T \Gamma^T Q \left( \bar{x} - \hat{x}^{(N)} \right). \end{aligned} \quad (\text{C.17})$$

By the definition of  $\tilde{u}_i$

$$\begin{aligned} u_i &= \tilde{u}_i - R^{-1} B^T \left( \Pi_{\theta_i} x_i + s_{\theta_i} \right) \\ &= \tilde{u}_i - R^{-1} B^T \Pi_{\theta_i} \hat{x}_i - R^{-1} B^T \left( \Pi_{\theta_i} \hat{x}_i + s_{\theta_i} \right). \end{aligned}$$

It follows that

$$\sum_{i=1}^N u_i^T R u_i$$

$$\begin{aligned} &= \sum_{i=1}^N \left( \Pi_{\theta_i} \hat{x}_i + s_{\theta_i} \right)^T B R^{-1} B^T \left( \Pi_{\theta_i} \hat{x}_i + s_{\theta_i} \right) \quad (=:\zeta_4) \\ &\quad + \sum_{i=1}^N \left| \tilde{u}_i - R^{-1} B^T \Pi_{\theta_i} \hat{x}_i \right|_R^2 \quad (=:\zeta_5) \\ &\quad + 2 \sum_{i=1}^N \left( \Pi_{\theta_i} \hat{x}_i + s_{\theta_i} \right)^T B \left( R^{-1} B^T \Pi_{\theta_i} \hat{x}_i - \tilde{u}_i \right) \quad (=:\zeta_6). \end{aligned}$$

Now we write

$$\begin{aligned} &\sum_{i=1}^N \left( \left| x_i - \Gamma x^{(N)} - \eta \right|_Q^2 + u_i^T R u_i \right) \\ &= (\zeta_1 + \zeta_4) + (\zeta_2 + \zeta_5) + (\zeta_3'' + \zeta_6) + (\zeta_3 - \zeta_3''). \end{aligned}$$

*Step 3 (Estimate of Optimality Loss):* We have

$$\begin{aligned} \zeta_3'' + \zeta_6 &= \sum_{i=1}^N \left[ 2 \hat{x}_i^T \left( Q \hat{x}_i - Q \Gamma \bar{x} - \Gamma^T Q \bar{x} + \Gamma^T Q \Gamma \bar{x} - Q \eta + \Gamma^T Q \eta \right) \right. \\ &\quad \left. - 2 \left( \Pi_{\theta_i} \hat{x}_i + s_{\theta_i} \right)^T B \left( \tilde{u}_i - R^{-1} B^T \Pi_{\theta_i} \hat{x}_i \right) \right] \\ &=: \sum_{i=1}^N Y_i. \end{aligned}$$

As in [11, Eq. (A.13)], we apply Ito's formula to  $e^{-\rho t} \hat{x}_i^T \left( \Pi_{\theta_i} \hat{x}_i + s_{\theta_i} \right)$  and use (C.5), (C.10) to show that  $E \int_0^\infty e^{-\rho t} Y_i(t) dt = 0$ . Note that  $J_{\text{soc}}^{(N)}(\hat{u}) = E \int_0^\infty e^{-\rho t} (\zeta_1 + \zeta_4) dt$ . Hence,

$$\begin{aligned} J_{\text{soc}}^{(N)}(u) &= J_{\text{soc}}^{(N)}(\hat{u}) + E \int_0^\infty e^{-\rho t} (\zeta_2 + \zeta_5) dt \\ &\quad + E \int_0^\infty e^{-\rho t} (\zeta_3 - \zeta_3'') dt. \end{aligned}$$

We have

$$\zeta_3 - \zeta_3'' = 2N(\bar{x} - \hat{x}^{(N)})^T (\Gamma^T Q + Q \Gamma - \Gamma^T Q \Gamma) \hat{x}^{(N)}.$$

By (C.16), Schwarz's inequality and Lemma 4.1, we can show that for  $u$  satisfying (C.6)

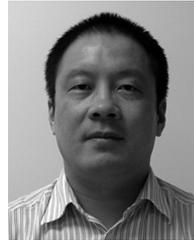
$$\left| E \int_0^\infty e^{-\rho t} (\zeta_3 - \zeta_3'') dt \right| \leq CN \left( \frac{1}{\sqrt{N}} + \epsilon_N \right)$$

where  $C$  does not depend on  $N$ . Moreover,  $E \int_0^\infty e^{-\rho t} (\zeta_2 + \zeta_5) dt \geq 0$ . The theorem follows.  $\square$

## REFERENCES

- [1] S. Adlakha, R. Johari, G. Weintraub, and A. Goldsmith, "Oblivious equilibrium for large-scale stochastic games with unbounded costs," in *Proc. 48th IEEE CDC*, Cancun, Mexico, Dec. 2008, pp. 5531–5538.
- [2] J.-P. Aubin, *Optima and Equilibria: An Introduction to Nonlinear Analysis*, 2nd ed. Berlin, Germany: Springer, 1998.
- [3] C. T. Bauch and D. J. D. Earn, "Vaccination and the theory of games," *Proc. Nat. Acad. Sci.*, vol. 101, pp. 13391–13394, Sep. 2004.
- [4] A. Bensoussan, *Stochastic Control of Partially Observable Systems*. Cambridge, U.K.: Cambridge Univ. Press, 1992.
- [5] R. Breban, R. Vardavas, and S. Blower, "Mean-field analysis of an inductive reasoning game: Application to influenza vaccination," *Phys. Rev. E*, vol. 76, no. DOI: 10.1103/PhysRevE.76.031127, 2007.

- [6] 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.
- [7] D. F. Delchamps, "Analytic feedback control and the algebraic Riccati equation," *IEEE Trans. Autom. Control*, vol. AC-29, no. 11, pp. 1031–1033, Nov. 1984.
- [8] G. M. Erickson, "Differential game models of advertising competition," *Eur. J. Oper. Res.*, vol. 83, pp. 431–438, 1995.
- [9] E. J. Green, "Continuum and finite-player noncooperative models of competition," *Econometrica*, vol. 52, no. 4, pp. 975–993, 1984.
- [10] Y.-C. Ho, "Team decision theory and information structures," *Proc. IEEE*, vol. 68, no. 6, pp. 644–654, Jun. 1980.
- [11] 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.
- [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," in *Proc. 42nd IEEE Conf. Decision Control*, Maui, HI, Dec. 2003, pp. 98–103.
- [13] M. Huang, P. E. Caines, and R. P. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized  $\varepsilon$ -Nash equilibria," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1560–1571, Sep. 2007.
- [14] M. Huang, P. E. Caines, and R. P. Malhamé, "The NCE (mean field) principle with locality dependent cost interactions," *IEEE Trans. Autom. Control*, vol. 55, no. 12, pp. 2799–2805, Dec. 2010.
- [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. New York: Springer, 2005, pp. 215–252.
- [16] 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. Inf. Syst.*, vol. 6, no. 3, pp. 221–252, 2006.
- [17] R. Johari and J. N. Tsitsiklis, "Efficiency loss in a network resource allocation game," *Math. Oper. Res.*, vol. 29, no. 3, pp. 407–435, 2004.
- [18] F. P. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, pp. 33–37, 1997.
- [19] M. A. Khan and Y. Sun, "Non-cooperative games with many players," in *Handbook of Game Theory with Economic Applications*, R. J. Aumann and S. Hart, Eds. Amsterdam, The Netherlands: North-Holland, 2002, vol. 3, pp. 1761–1808.
- [20] V. E. Lambson, "Self-enforcing collusion in large dynamic markets," *J. Econ. Theory*, vol. 34, pp. 282–291, 1984.
- [21] 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.
- [22] 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.
- [23] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japan. J. Math.*, vol. 2, no. 1, pp. 229–260, 2007.
- [24] A. Lazarte, "Large population stochastic dynamic games with nonuniform agents and localized coupling," M.S. thesis, McGill Univ., Dept. Elect. Comput. Eng., Montreal, QC, Canada, 2009.
- [25] T. Li and J.-F. Zhang, "Asymptotically optimal decentralized control for large population stochastic multiagent systems," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1643–1660, Aug. 2008.
- [26] G. Owen, *Game Theory*, 3rd ed. San Diego, CA: Academic, 1995.
- [27] T. Roughgarden and E. Tardos, "Bounding the inefficiency of equilibria in nonatomic congestion games," *Games Econ. Behavior*, vol. 47, pp. 389–403, 2004.
- [28] E. D. Sontag, *Mathematical Control Theory*. New York: Springer-Verlag, 1990.
- [29] H. Tembine, J.-Y. L. Boudec, R. El-Azouzi, and E. Altman, "Mean field asymptotics of Markov decision evolutionary games and teams," in *Proc. Int. Conf. Game Theory for Netw.*, Istanbul, Turkey, May 2009, pp. 140–150.
- [30] G. Y. Weintraub, C. L. Benkard, and B. V. Roy, "Oblivious equilibrium: A mean field approximation for large-scale dynamic games," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2005.
- [31] G. Y. Weintraub, C. L. Benkard, and B. V. Roy, "Markov perfect industry dynamics with many firms," *Econometrica*, vol. 76, no. 6, pp. 1375–1411, Nov. 2008.
- [32] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, 2nd ed. New York: Springer-Verlag, 1979.
- [33] D. W. K. Yeung and L. A. Petrosyan, *Cooperative Stochastic Differential Games*. New York: Springer, 2006.
- [34] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag, "Synchronization of coupled oscillators is a game," *IEEE Trans. Autom. Control*, vol. 57, no. 4, pp. 920–935, Apr. 2012.



**Minyi Huang** (M'03) received the B.Sc. degree from Shandong University, Jinan, Shandong, China, in 1995, the M.Sc. degree from the Institute of Systems Science, Chinese Academy of Sciences, Beijing, China, in 1998, and the Ph.D. degree from the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, in 2003, all in the area of systems and control.

He was a Research Fellow first in the Department of Electrical and Electronic Engineering, the University of Melbourne, Melbourne, Australia, from February 2004 to March 2006, and then in the Department of Information Engineering, Research School of Information Sciences and Engineering, the Australian National University, Canberra, Australia, from April 2006 to June 2007. He joined the School of Mathematics and Statistics, Carleton University, Ottawa, ON, Canada, as an Assistant Professor in July 2007, where he has been an Associate Professor since July 2011. His research interests include mean field stochastic control and dynamic games, multi-agent control and computation in distributed networks with applications.



**Peter E. Caines** (M'73–SM'83–F'86–LF'11) received the B.A. degree in mathematics from Oxford University, Oxford, U.K., in 1967 and the Ph.D. degree in systems and control theory from Imperial College, University of London, London, U.K., in 1970 under the supervision of David Q. Mayne, FRS.

He had been a Postdoctoral Researcher and Faculty Member at UMIST, Stanford, UC Berkeley, Toronto, and Harvard; he joined McGill University, Montreal, QC, Canada, in 1980, where he is James McGill Professor and Macdonald Chair. He is the author of *Linear Stochastic Systems* (Wiley, 1988), and his research interests include stochastic, mean field, and hybrid systems theory.

Dr. Caines is a Fellow of SIAM, IMA, and the CIAR, and was elected to the Royal Society of Canada in 2003. In 2000, the adaptive control paper he coauthored in the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, 1980, was recognized by the IEEE Control Systems Society as one of the 25 seminal control theory papers of the 20th century. In 2009, he received the IEEE Control Systems Society Bode Lecture Prize.

Dr. Caines is a Fellow of SIAM, IMA, and the CIAR, and was elected to the Royal Society of Canada in 2003. In 2000, the adaptive control paper he coauthored in the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, 1980, was recognized by the IEEE Control Systems Society as one of the 25 seminal control theory papers of the 20th century. In 2009, he received the IEEE Control Systems Society Bode Lecture Prize.



**Roland P. Malhamé** (M'81) received the B.S. degree from the American University of Beirut, Beirut, Lebanon, in 1976, the M.S. degree from the University of Houston, Houston, TX, in 1978, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, in 1983, all in electrical engineering.

After single-year stays at the University of Quebec (Chicoutimi), and CAE Electronics, Ltd. (Montreal), he joined in 1985 the École Polytechnique de Montréal, Montreal, QC, Canada, where he is a Professor of electrical engineering. His current research interests are in the analysis of complex systems and the development of biology-inspired schemes for their decentralized control, with applications in manufacturing, communications, smart grids, and collective dynamics. Past contributions include a statistical mechanics inspired approach to electric load modeling. From June 2005 to April 2011, he was director of GERAD (Groupe d'Études et de Recherche en Analyse des Décisions).