2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

# Mean Field (NCE) Stochastic Control: Populations of Major and Egoist-Altruist Agents

Arman C. Kizilkale and Peter E. Caines

***Abstract*—For noncooperative games the Nash Certainty Equivalence (NCE), or Mean Field (MF) methodology [1], [2] provides decentralized strategies which asymptotically yield Nash equilibria. An extension of this theory to populations of altruistic agents (defined with so-called social cost functions) and to mixed populations was carried out in [3], and a theory treating populations of egoistic agents and one or more so-called major agents was developed in [4]. In this paper we study the equilibria and the overall stability of dynamic LQG games, where (i) there is a single major agent and a large population of mixed minor agents, and (ii) the cost for each minor agent is a convex combination of its own cost and the social cost of the minor agents. We analyse the resulting equilibria, provide experimental results, and present a mean field stochastic control algorithm, which when applied by all agents in the system, gives rise to system behaviour where (i) all agents systems are $L^2$ stable, (ii) the set of controls yields an $\epsilon$-Nash equilibrium for all $\epsilon$, and (iii) if each minor agent in the system only considers the social cost, then the difference between (i) the cost observed by each minor agent and (ii) the social cost that would be observed if a centralized controller minimizes the social cost tends to zero as the population size grows to infinity.**

## I. INTRODUCTION

The optimization and control of large scale dynamic systems is both analytically and computationally complex. Usually one encounters the *curse of dimensionality*, therefore distributed or decentralized control approaches are applied. Game theory has been formulated to capture individual payoffs or costs, but even individual best response algorithms lead to few practical results, since the analytic complexity is usually very high.

The analysis of large scale dynamic systems where agents are coupled via dynamics and cost functions was presented in [1], [2], [5] where the theory of Nash Certainty Equivalence (NCE) (Mean Field (MF)) control was introduced and decentralized strategies which yield Nash equilibria were provided. It is to be noted that the dynamic large scale cost coupled structure of [2] is motivated by various scenarios, for instance those analysed in [6]–[9]. Individual control laws use local information and the average effect of all agents taken together, henceforth referred to as *the mass*. Related approaches have been independently developed in [10]–[12], a nonlinear extension using McKean-Vlasov Markov process models is presented in [13], and the analysis of an adaptive framework is presented in [14]–[16].

*Cooperative Behaviour in Mean Field LQG Control:*

The notion of social global optima (i.e., minimum summed individual costs) is a major issue in decentralized and distributed control and optimization problems, and Pareto optimality is a widely accepted characterization. In contrast to competitive behaviour studied in the NCE (MF) framework, a different situation may arise when the agents in the population seek socially optimal actions. Even though it is often only studied by static models, Pareto optimality has been extensively studied in optimization problems, either as a problem on its own, or as a tool for comparison to competitive behaviour (see e.g. [17]). In the mean field framework an LQG model is adopted in [18] for the minimization of a social cost function, and centralized and decentralized strategies are considered; in this work social certainty equivalence (SCE) methodology is introduced, where the mean field trajectory is obtained by each agent without any observations on other agents' trajectories. This cooperative game problem is then extended to a model [3] where the optimization problem of individual agents involves a cost reflecting both individual and social interests. To model this situation the weight that each agent assigns to its individual cost function and the social cost function changes continuously across the population.

*LQG Games Involving a Major Agent:*

A different dynamic game model is studied in [4], where a single major and a large population of minor agents exist, in contrast to [1], [2], where all agents are non-atomic. In cooperative game theory, games including small and large agents are denoted as *mixed games* [19]. Such games are useful in modelling markets with a dominant big corporation and several small entities.

*Major Agent vs Minor Egoistic and Altruistic Agents:*

The step in the development of MF stochastic dynamic game theory taken in this paper is the analysis of the situation where there is a population of mixed agents; namely a single major agent together with a large population of minor agents whose cost functions reflect both individual and social interest. The major agent and the minor agents are coupled in a way to be described in detail later. Under reasonable conditions on the population dynamical parameter distribution, this paper presents a mean field stochastic control algorithm which when applied by all agents in the system, gives rise to system behaviour where (i) all agents systems are $L^2$ stable, (ii) the set of controls yields an $\epsilon$-Nash equilibrium for all $\epsilon$, and (iii) if each minor agent in the system only considers the social cost, then the difference between the cost observed by each minor agent and the social cost that would be observed if a centralized controller minimizes the social cost tends to zero as the population size grows to infinity.

The organization of the paper is as follows. The dynamic game problem is formulated in Section II. The individual

Arman C. Kizilkale and Peter E. Caines are with the Department of Electrical and Computer Engineering and the Centre for Intelligent Machines, McGill University, Montreal, Canada. {arman;peterc}@cim.mcgill.ca

control is examined for the single major agent and minor agents in Section III. In order to obtain decentralized solutions, the NCE-SCE methodology is presented in Section IV. In Section V, we study the stability of the action profile, when the weight each mixed minor agent assigns to its individual cost is defined as a decision parameter. In Section VI we present the supporting simulation results and Section VII concludes the paper.

## II. STOCHASTIC DYNAMIC GAME MODEL

Following [4], we consider a large population of $N$ stochastic dynamic minor agents $\mathcal{A}_i$, $1 \le i \le N$, and a single major agent $\mathcal{A}_0$, where the individual dynamics are defined by

$$\begin{aligned} dx_0 &= [A_0 x_0 + B_0 u_0]dt + D_0 dw_0, \\ dx_i &= [A(\theta_i)x_i + B(\theta_i)u_i + Gx_0]dt + Ddw_i, \end{aligned} \tag{1}$$

$t \ge 0, 1 \le i \le N$. Here $x_0, x_i \in \mathbb{R}^n$ are the states, $u_0, u_i \in \mathbb{R}^m$ are the control inputs, $\{w_i, 0 \le i \le N\}$ denotes $(N+1)$ independent standard Wiener processes in $\mathbb{R}^r$ on a sufficiently large underlying probability space $(\Omega, \mathcal{F}, P)$ that $w$ is progressively measurable with respect to $\mathcal{F}^w \triangleq (\mathcal{F}^w_t; t \ge 0)$. Note that the major agent $\mathcal{A}_0$ affects each minor agent through its dynamics. The initial states are defined on $(\Omega, \mathcal{F}, P)$, and $\{x_i(0), 0 \le i \le N\}$ are mutually independent and also independent of $\mathcal{F}^w_\infty$; $\mathbb{E}ww^\top = \Sigma$, and $\mathbb{E}\|x(0)\|^2 < \infty$. We denote the state configuration by $x = (x_0, \cdots, x_N)^\top$, and the minor agent population average state by $x^N = (1/N)\sum_{i=1}^N x_i$.

The individual discounted cost functions for the agents $\mathcal{A}_i, 0 \le i \le N$, are given by

$$\begin{aligned} J_0^N(u_0, u_{-0}) &= \mathbb{E}\int_0^\infty e^{-\rho t}\{\|x_0 - \Phi(x^N)\|^2_{Q_0} + \|u_0\|^2_{R_0}\}dt, \\ J_i^N(u_i, u_{-i}) &= \mathbb{E}\int_0^\infty e^{-\rho t}\{\|x_i - \Psi(x^N)\|^2_{Q} + \|u_i\|^2_{R}\}dt, \end{aligned} \tag{2}$$

where we assume the cost-coupling to be of the form $\Phi(t) \triangleq H_0 x^N(t) + \eta_0$, $\eta_0 \in \mathbb{R}^n$, and $\Psi(t) \triangleq Hx_0(t) + \hat{H}x^N(t) + \eta$, $\eta \in \mathbb{R}^n$, and we use the notation $\|a-b\|^2_Q \triangleq (a-b)^\top Q(a-b)$, $Q \ge 0$, for arbitrary $a$, $b$ and $Q$. The coefficients $[A(\theta_i), B(\theta_i)] \in \mathbb{R}^{n(n+m)}$, will be called the dynamical parameters. The variability of the parameter $\theta_i$ is used to model the population of minor agents. We assume that $\theta_i$ takes values from the finite set $\mathbf{\Theta} = \{1, ..., K\}$ so that there are $K$ types of minor agents. The disturbance weight matrices $D_0, D$, the major agent weight matrix $G$, and the control action penalizing matrices $R_0, R$ are constant matrices and constitute known information for all agents in the system. The functions $u_0(\cdot)$ and $u_i(\cdot)$ are the control inputs of the agents $\mathcal{A}_0$ and $\mathcal{A}_i$, $1 \le i \le N$, respectively and $u_{-i}$ denotes the control inputs of the complementary set of agents $\mathcal{A}_{-i} = \{\mathcal{A}_j, j \ne i, 0 \le j \le N\}$.

The social cost of the minor agents for a minor population size $N$ is defined as

$$J^N_{soc}(u) = \sum_{i=1}^N J_i^N(u_i, u_{-i}). \tag{3}$$

To model the *egoism degree* of a minor individual agent and its contribution in optimizing the social cost, for agent $i$, $1 \le i \le N$, we define

$$J_i^N(u_i, u_{-i} \mid \lambda_i) = \lambda(\theta_i)J_i^N(u_i, u_{-i}) + (1-\lambda(\theta_i))J^N_{soc}(u), \tag{4}$$

where the egoism degree of each agent $\lambda(\theta_i) \in [0,1]$. In the noncooperative game problem studied in [1], [2] each agent optimizes with respect to $J_i^N(\cdot)$ alone. The agent type dependent parameter $\lambda(\theta_i)$ is a measure of the egoism degree of agent $\mathcal{A}_i$, and $1-\lambda(\theta_i)$ measures the weight it contributes to the social interest. When $\lambda(\theta_i)$ increases, the agent is more self oriented: if $\lambda(\theta_i) = 1$, the optimization behaviour of agent $\mathcal{A}_i$ is purely *egoistic*, and if $\lambda(\theta_i) = 0$, it is a purely *altruistic* behaviour.

For a given $N$, define $\mathcal{I}_k = \{i : \theta_i = k, 1 \le i \le N\}$, $N_k = |\mathcal{I}_k|$. Let $\pi^N_k = N_k/N$. The empirical distribution of $(\theta_1, ..., \theta_N)$ is given by the probability vector $\pi^N = (\pi^N_1, ..., \pi^N_K)$.

For the basic MF control problem, the following assumptions are adopted:

**A1**: All agents have mutually independently distributed initial conditions: $\{w_i, 0 \le i \le N\}$ are mutually independent and independent of the initial conditions, and $\sup_{i \ge 0}[\mathrm{Tr}\Sigma_i + \mathbb{E}\|x_i(0)\|^2] < \infty$. □

**A2**: $\mathbf{\Theta}$ is a set such that for each $k \in \mathbf{\Theta}$, $[A_k - (\rho/2)I, B_k]$ is controllable and $[Q^{1/2}, A_k - (\rho/2)I]$ is observable. □

**A3**: The cost-coupling is of the form: $\Phi(\cdot) \triangleq H_0 x^N + \eta_0$, $\eta_0 \in \mathbb{R}^n$ and $\Psi(\cdot) \triangleq Hx_0 + \hat{H}x^N + \eta$, $\eta \in \mathbb{R}^n$. □

**A4**: There exists a probability vector $\pi$ such that $\lim_{N\to\infty} \pi^N = \pi$, where $\pi = (\pi_1, ..., \pi^N_K)$ is a probability vector which gives the empirical distribution of $(\theta_1, ..., \theta_N)$. □

For the rest of the paper we assume that $\min_{1 \le k \le K} N_k \ge 1$.

## III. CONTROL ACTIONS OF INDIVIDUAL AGENTS

For the optimality analysis, we first introduce three admissible control sets. The set of control inputs $\mathcal{U}_g$, based upon the global observation control set, consists of all feedback controls adapted to $\{A_j, B_j, 0 \le j \le N; \pi; \mathcal{F}^N_t; t \ge 0\}$ and the set of control inputs $\mathcal{U}_{l,i}$, based upon the local information set of minor agent $\mathcal{A}_i$, consists of the feedback controls adapted to the set $\{A_i, A_0, B_i, B_0; \pi; \mathcal{F}_{i,t}, \mathcal{F}_{0,t}; t \ge 0\}$. The $\sigma$-field $\mathcal{F}_{i,t}$ is the increasing family of $\sigma$-fields generated by $(x_i(\tau); 0 \le \tau \le t)$, and $\mathcal{F}^N_t$ is the increasing family of the $\sigma$-field generated by the set $\{x_j(\tau); 0 \le \tau \le t, 0 \le j \le N\}$. The last set of admissible control inputs, $\mathcal{U}_{l,0}$ for the major agent $\mathcal{A}_0$ is the set of all feedback controls adapted to $\{A_0, B_0; \pi; \mathcal{F}_{0,t}; t \ge 0\}$. We classify the states $x_i$, $1 \le i \le N$, into $K$ groups. Define $z_k = \frac{1}{N_k}\sum_{i \in \mathcal{I}_k} x_i$, $1 \le k \le K$, which is the average state of the same type of agents. It is shown in [4] that for an infinite population of minor agents, the equation system that captures the dynamics can be written in the form

$$d\bar{z}_k = \sum_{j=1}^{K} \bar{A}_{k,j} z_j dt + \bar{G}_k x_0 dt + \bar{m}_k dt, \quad 1 \le k \le K,$$

where the noise term disappears since $\lim_{N\to\infty} N_k/N = \pi_k > 0$. Notice that the notation $\bar{z}$ reflects the dynamics in the limit of an infinite population, and $\bar{m}_k(t)$ denotes the offset term that appears due to $\eta_0$ and $\eta$ defined in **A3**.

### A. Control Action of the Major Agent

We first formalize the auxiliary game between the major agent and the *mass* via the approximation of the mean state $x^N$. Notice the relation $x^N = (1/N)\sum_{k=1}^{K} N_k z_k = \sum_{k=1}^{K} \pi_k^N z_k$. One can also write $x^N = (1/N)\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} x_i = \sum_{k=1}^{K} \pi_k^N (1/N_k)\sum_{i\in\mathcal{I}_k} x_i$. For large $N$, we may approximate $x^N$ by $\sum_{k=1}^{K} \pi_k \bar{z}_k$ where $\bar{z}_k \in \mathbb{R}^n$ is used to approximate $(1/N_k)\sum_{i\in\mathcal{I}_k} x_i$ for an infinite population. We denote $\bar{z} = [\bar{z}_1^\top, ..., \bar{z}_k^\top]^\top$, which is to be called the limiting mass state. This limiting process is described by the equation [4]

$$d\bar{z}(t) = \bar{A}\bar{z}(t)dt + \bar{G}x_0(t)dt + \bar{m}(t)dt, \tag{5}$$

where $\bar{z}(0) = 0_{nK\times 1}$, $\bar{A} \in \mathbb{R}^{nK\times nK}$, and $\bar{G} \in \mathbb{R}^{nK\times n}$ are constant matrices, and $\bar{m}(t)$ is a continuous function on $[0,\infty)$.

The existence of a major agent in the system alters the analysis significantly. In [2] the best response actions are calculated offline using the NCE stochastic control law. Statistical information of the dynamical parameters of the system is known by each agent, and each calculates the mass tracking trajectory offline. Therefore, the equilibrium of the system is achieved for a completely decentralized set of statistically independent agents. The rationale is that the tracking signal can be calculated offline as a deterministic process due to the fact that each agent's contribution diminishes to zero as the population size tends to infinity. Following [4], in the framework in this paper, the major agent and the mass of the minor agents are in a bilateral relation such that the mass behaviour evolves as a function of the major agent's state. As the major agent's state is a random process, the mass behaviour and major agent's state affect each other continuously.

The dynamics of the major agent $\mathcal{A}_0$ may be written in the form

$$\begin{bmatrix} dx_0 \\ d\bar{z} \end{bmatrix} = \begin{bmatrix} A_0 & 0_{nK\times n} \\ \bar{G} & \bar{A} \end{bmatrix} \begin{bmatrix} x_0 \\ \bar{z} \end{bmatrix} dt + \begin{bmatrix} B_0 \\ 0_{nK\times m} \end{bmatrix} u_0 dt + \begin{bmatrix} 0_{n\times 1} \\ \bar{m} \end{bmatrix} dt + \begin{bmatrix} D_0 dw_0 \\ 0_{nK\times 1} \end{bmatrix}, \tag{6}$$

where $\bar{z}(0) = 0_{nK\times 1}$. Note that the limiting mass state is augmented to the state of the major agent. Let $\otimes$ denote the Kronecker product of two matrices, and set $H_0^\pi = \pi \otimes H_0$. Then we define the following:

$$\mathbb{A}_0 = \begin{bmatrix} A_0 & 0_{nK\times n} \\ \bar{G} & \bar{A} \end{bmatrix}, \mathbb{B}_0 = \begin{bmatrix} B_0 \\ 0_{nK\times m} \end{bmatrix},$$

$$\mathbb{M}_0 = \begin{bmatrix} 0_{n\times 1} \\ \bar{m} \end{bmatrix}, Q_0^\pi = \begin{bmatrix} Q_0 & -Q_0 H_0^\pi \\ -H_0^{\pi\top} Q_0 & H_0^{\pi\top} Q_0 H_0^\pi \end{bmatrix},$$

and $\bar{\eta}_0 = [I_{n\times n}, -H_0^\pi]^\top Q_0 \eta_0$, and we introduce the algebraic Riccati Equation

$$\rho\Pi_0 = \Pi_0\mathbb{A}_0 + \mathbb{A}_0^\top\Pi_0 - \Pi_0\mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0 + Q_0^\pi,$$

and the ODE

$$\rho s_0 = \frac{ds_0}{dt} + (\mathbb{A}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0)^\top s_0 + \Pi_0\mathbb{M}_0 - \bar{\eta}_0,$$

Under **A1** and **A2**, in the admissible control set $\mathcal{U}_{l,0}$, the optimal control law for $\mathcal{A}_0$ is given as [4]

$$u_0 = -R_0^{-1}\mathbb{B}_0^\top \left[\Pi_0(x_0^\top, \bar{z}^\top)^\top + s_0\right]. \tag{7}$$

### B. Control Action of the Mixed Minor Agents

Non-atomic agents are continuously effected by the major agent $\mathcal{A}_0$'s trajectory, and as a result of this, $\{x_i, 1\le i\le N\}$ is correlated with the state process $x_0$ of $\mathcal{A}_0$. As described in Sec. III, the admissible control set for an agent $\mathcal{A}_i$ is $\mathcal{U}_{l,i}$; therefore, the state trajectory of the major agent is observed at each time iteration. Again for a large population approximation we use the limiting mass state $\bar{z}$ and obtain [4] the infinite population equations:

$$\begin{bmatrix} dx_i \\ dx_0 \\ d\bar{z} \end{bmatrix} = \begin{bmatrix} A_k & [G \;\; 0_{n\times nK}] \\ 0_{(nK+n)\times n} & \mathbb{A}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0 \end{bmatrix} \begin{bmatrix} x_i \\ x_0 \\ \bar{z} \end{bmatrix} dt + \begin{bmatrix} B_k \\ 0_{(nK+n)\times m} \end{bmatrix} u_i dt + \begin{bmatrix} 0_{n\times 1} \\ \mathbb{M}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top s_0 \end{bmatrix} dt + \begin{bmatrix} D dw_i \\ D_0 dw_0 \\ 0_{nK\times 1} \end{bmatrix}.$$

Define

$$\mathbb{A}_k = \begin{bmatrix} A_k & [G \;\; 0_{n\times nK}] \\ 0_{(nK+n)\times n} & \mathbb{A}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0 \end{bmatrix},$$

$$\mathbb{B}_k = \begin{bmatrix} B_k \\ 0_{(nK+n)\times m} \end{bmatrix},$$

$$\mathbb{M} = \begin{bmatrix} 0_{n\times 1} \\ \mathbb{M}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top s_0 \end{bmatrix},$$

and $\bar{\eta}_{\lambda_k} = [I_{n\times n} - (1-\lambda_k)\hat{H}, -H, -\hat{H}^\pi]^\top Q\eta$. Note that the *egoism degree* $\lambda_k$ enters the dynamics through the $\bar{\eta}_{\lambda_k}$ function. We define the cost function for a mixed minor agent as $Q_{\lambda_k}^\pi = Q_{ind}^\pi + (1-\lambda_k)Q_{soc}^\pi$, where $Q_{ind}^\pi = [I, -H, -\hat{H}^\pi]^\top Q[I, -H, -\hat{H}^\pi]$, and

$$Q_{soc}^\pi = \left[\begin{array}{c|c|c} 0 & \hat{H}^\top Q H & -\hat{H}^\top Q I^\pi + \hat{H}^\top Q \hat{H}^\pi \\ H^\top Q \hat{H} & H^\top Q H & -H^\top Q I^\pi + H^\top Q \hat{H}^\pi \\ -I^{\pi\top} Q\hat{H} + [\hat{H}^\pi]' Q\hat{H} & -I^{\pi\top} Q H + [\hat{H}^\pi]' Q H & -2I^{\pi\top} Q\hat{H}^\pi + [\hat{H}^\pi]' Q\hat{H}^\pi \end{array}\right].$$

Here $Q^{\pi}_{ind}$ is the individual cost paid by the mixed minor agent $\mathcal{A}_i$, whereas $Q^{\pi}_{soc}$ denotes its cost contribution to the social cost (3).

We introduce the algebraic Riccati Equation

$$\rho\Pi_k = \Pi_k\mathbb{A}_k + \mathbb{A}_k^\top\Pi_k - \Pi_k\mathbb{B}_k R^{-1}\mathbb{B}_k^\top\Pi_k + Q^{\pi}_{\lambda_k},$$

and the ODE

$$\rho s_k = \frac{ds_k}{dt} + (\mathbb{A}_k - \mathbb{B}_k R^{-1}\mathbb{B}k^\top\Pi_k)^\top s_k + \Pi_k\mathbb{M} - \bar{\eta}_{\lambda_k}.$$

Subject to **A1** and **A2**, in the admissible control set $\mathcal{U}_{l,i}$, the optimal control law for the mixed minor agent $\mathcal{A}_i$ is given by [4]

$$u_i^0 = -R^{-1}\mathbb{B}_0^\top\left[\Pi_k(x_i^\top, x_0^\top, \bar{z}^\top)^\top + s_k\right]. \tag{8}$$

## IV. NCE-SCE EQUATIONS

All agents in the system are assumed to be rational, i.e., each agent minimizes its own cost function (more or less cooperative depending on the egoism degree). The linking term parameters $\bar{A}_k$, $\bar{G}_k$, $\bar{m}_k$ can be obtained by each agent under this assumption. Note that due to the major agent's presence, the mass trajectory cannot be obtained offline as the mass evolves by reacting to the actions of the major agent which is continuously subject to disturbances. It is this stochasticity of the major agent which removes the possibility of an offline computation of a deterministic mass behaviour.

We partition the matrix $\Pi_k$, $1 \le k \le K$, and obtain

$$\Pi_k = \begin{bmatrix} \Pi_{k,11} & \Pi_{k,12} & \Pi_{k,13} \\ \Pi_{k,21} & \Pi_{k,22} & \Pi_{k,23} \\ \Pi_{k,31} & \Pi_{k,32} & \Pi_{k,33} \end{bmatrix}.$$

For the overall population, we may now specify the Nash certainty equivalence - social certainty equivalence (NCE-SCE) equation system:

*Definition 4.1: NCE-SCE Equation System:*

$$\begin{aligned}
\rho\Pi_0 &= \Pi_0\mathbb{A}_0 + \mathbb{A}_0^\top\Pi_0 - \Pi_0\mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0 + Q_0^{\pi}, \\
\rho\Pi_k &= \Pi_k\mathbb{A}_k + \mathbb{A}_k^\top\Pi_k - \Pi_k\mathbb{B}_k R^{-1}\mathbb{B}_k^\top\Pi_k + Q^{\pi}_{\lambda_k}, \ \ \forall k, \\
\bar{A}_k &= [A_k - B_k R^{-1}B_k^\top\Pi_{k,11}]\mathbf{e}_k - B_k R^{-1}B_k^\top\Pi_{k,13}, \ \ \forall k, \\
\bar{G}_k &= -B_k R^{-1}B_k^\top\Pi_{k,12}, \ \ \forall k, \\
\rho s_0 &= \frac{ds_0}{dt} + (\mathbb{A}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0)^\top s_0 + \Pi_0\mathbb{M}_0 - \bar{\eta}_0, \\
\rho s_k &= \frac{ds_k}{dt} + (\mathbb{A}_k - \mathbb{B}_k R^{-1}\mathbb{B}_k^\top\Pi_k)^\top s_k + \Pi_k\mathbb{M} - \bar{\eta}_{\lambda_k}, \ \ \forall k, \\
\bar{m}_k &= -B_k R^{-1}B_k^\top s_k, \ \ \forall k,
\end{aligned} \tag{9}$$

□

where $\mathbf{e}_k = [0_{n\times n}, ..., 0_{n\times n}, I_{n\times n}, 0_{n\times n}, ..., 0_{n\times n}]$ ($k_{th}$ entry), and where the identity matrix $I_{n\times n}$ is at the $k$th block. Note that $\lambda(\theta_i) = 1$, $1 \le i \le N$, ,yields the NCE equation system described in [2] and $\lambda(\theta_i) = 0$, $1 \le i \le N$, gives rise to the SCE equation system defined in [18] for an entire population of minor cooperative agents. Due to the mixed nature of the agents in the system described in this paper, we call (9) the *NCE-SCE equation system*. In order to ensure the existence of a solution to (9), *stabilizing consistency* conditions described in [4] have to be satisfied.

### A. The NCE-SCE Stochastic Control Law

Each agent solves the NCE-SCE equations offline, from which it obtains $\bar{A}$, $\bar{G}$ matrices and $\bar{m}$ vector that describe the behaviour of the mixed minor population's mass trajectory in (5). Then, at each time instant the major agent's optimal control action is given by (7), and for each minor agent, the optimal control action is given by (8).

Recall that $\mathcal{U}_g$ is defined in Sec. III as a set of centralized information based controls.

*Theorem 4.1: Major and All Altruistic Agents: Asymptotic Performance:* Let **A1-A4** hold and assume that the assumptions and *stabilizing consistency* requirements given in [4] are satisfied. Also let $\lambda(\theta_i) = 0$, for all $1 \le i \le N$. Then the set of NCE-SCE based control laws given in (7) and (8) have asymptotic social optimality, i.e., for $u^0 = (u_1^0, ..., u_N^0)$,

$$|(1/N)J^N_{soc}(u^0) - \inf_{u\in\mathcal{U}_g}(1/N)J^N_{soc}(u)| = O(1/\sqrt{N} + \bar{\epsilon}_N),$$

where $\lim_{N\to\infty}\bar{\epsilon}_N = 0$. □

The proof is similar to the proof of Theorem 2 in [3] and given in [20]. This theorem shows that as the population size tends to infinity, applying the NCE-SCE control law, the minor agents in a large population are able to increase their performance to the value that would be obtained only by a centralized controller with all the information of the system. Note that the minor agents only observe the major agent in the system, and not each other.

*Theorem 4.2: Major and Mixed Agents: NCE-SCE Equilibrium:* Let **A1-A4** hold and assume that the assumptions and *stabilizing consistency* requirements given in [4] are satisfied. The NCE-SCE stochastic control law generates a set of controls $\mathcal{U}^N_{nce-sce} \triangleq \{u_i^0; 0 \le i \le N\}$, $1 \le N < \infty$, with $u_0^0$ given in (7) and $u_i^0$, $1 \le i \le N$, given in (8) such that

(i) All agent systems $S(A_i)$, $0 \le i \le N$, are second order stable.

(ii) $\{\mathcal{U}^N_{nce-sce}; 1 \le N < \infty\}$ yields an $\epsilon$-Nash equilibrium for all $\epsilon$, i.e. for all $\epsilon > 0$, there exists $N(\epsilon)$ such that for all $N \ge N(\epsilon)$

$$\begin{aligned} J_i^N(u_i^0, u_{-i}^0 \mid \lambda_i) - \epsilon \le \inf_{u_i\in\mathcal{U}_g} J_i^N(u_i, u_{-i}^0, \mid \lambda_i) \\ \le J_i^N(u_i^0, u_{-i}^0 \mid \lambda_i). \end{aligned}$$

□

The proof follows that of Theorem 10 in [4] and is given in [20]. Note that the major agent has only statistical information about the mixed minor agents and does not observe their trajectories in the game. Also, even though mixed minor agents are allowed to observe the major agent, they do not observe each other. Therefore, the stability is obtained for a controlled system with a high degree of decentralization.

## V. INSTABILITY OF NON-EGOIST SOLUTIONS

We have shown that (8) gives the best reply of a minor agent and presented in Theorem 4.2 states that all agents are second order stable. Each minor agent's cost function is defined by (4) which is agent dependent via $\lambda(\theta_i)$. In this section we consider a different scenario: the case where the egoism degree $\lambda_i$ of an agent $\mathcal{A}_i$, $1 \le i \le N$, is not type dependent, but can be decided by each agent itself.

We define the average cost paid in the system for a finite population as $J_{ave}^N(u) = \frac{1}{N} J_{soc}^N(u)$, and the limiting average cost as, $J_{ave}(u) = \lim_{N\to\infty} \frac{1}{N} J_{soc}^N(u)$, where $J_{soc}^N(u)$ is defined in (3). We obtain the following proposition.

*Proposition 5.1:* Let **A1**-**A4** hold. Then,

(i) Each minor agent's cost function $J_i^N(u_i, u_{-i}; \lambda_i)$ is a decreasing function of $\lambda_i \in [0,1]$.

(ii) Given that all agents uniformly select the same egoism degree, i.e., $\lambda_i = \lambda \in [0,1]$, $1 \le i \le N$, the limiting average cost paid in the system $J_{ave}(u)$ is an increasing function of $\lambda \in [0,1]$. □

*Proof: Proof of (i):* The cost for each agent is written as

$$\begin{aligned} J^N(u_i, u_{-i} \mid \lambda_i) &= \lambda_i J_i^N(.) + (1-\lambda_i) J_{soc}^N(.) \\ &= \lambda_i J_i^N(.) \\ &+ (1-\lambda_i)\left(J_i^N(.) + \sum_{j\neq i} J_j^N(.)\right) \\ &= J_i^N(u_i, u_{-i}) + (1-\lambda_i)\sum_{j\neq i} J_j^N(u_j, u_{-j}). \end{aligned} \tag{10}$$

For $N \ge 1$, $J_i^N(u_i, u_{-i}) \ge 0$, and $\sum_{j\neq i} J_j^N(u_j, u_{-j}) \ge 0$. Also, minor agents do not observe other minor agents' actions and trajectories; therefore, $u_{-i}$ is independent of $\lambda_i$. As $J_i^N(u_i, u_{-i}) \ge 0$, (10) is minimized when $\lambda_i = 1$, and increasing as $\lambda_i$ tends to 0.

The proof of (ii) is given in [20]. □

These results emphasize the difference between the Nash certainty equivalence (NCE) and social certainty equivalence (SCE) frameworks. The mixed game model is equivalent to an NCE framework when all minor agents are egoistic (i.e., $\lambda_i = 1$ for all $1 \le i \le N$) and, conversely, equivalent to an SCE framework when all agents are altruistic (i.e., $\lambda_i = 0$ for all $1 \le i \le N$). Proposition 5.1 shows that even though a game where all agents in the system play altruistically gives the minimum cost per head, this action profile is not stable in the sense that it is always more profitable for an agent to set its egoism degree to 1. Therefore, one could imagine that whenever agents are allowed to successively choose their egoism degrees, all agents would eventually act selfishly, hence the NCE equilibrium would be the resulting asymptotic equilibrium of actions in the game.

## VI. SIMULATIONS

Consider a system of 100 minor agents and a single major agent. The system matrices $\{A_k, B_k, 1 \le k \le 100\}$ for the minor agents are uniformly defined as

$$A \triangleq \begin{bmatrix} -0.05 & -2 \\ 1 & 0 \end{bmatrix}, \quad B \triangleq \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

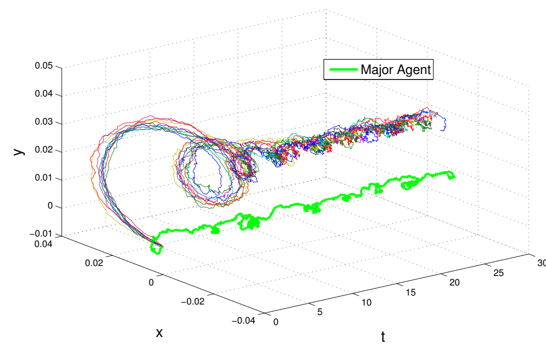


Fig. 1. State trajectories

and for the major agent we have

$$A_0 \triangleq \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_0 \triangleq \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The parameters used in the simulation are: $t_{final} = 30s$, $\Delta t = 0.025s$, $\sigma = 0.002$, $\rho = 0.01$, $\eta = [0.25, 0.25]^\top$, $\eta_0 = [0.25, 0.25]^\top$, $Q = I_{2\times2}$, $Q_0 = I_{2\times2}$, $R = 1$, $R_0 = 1$, $H = 0.6 \times I_{2\times2}$, $H_0 = 0.6 \times I_{2\times2}$, $\hat{H} = 0.6 \times I_{2\times2}$, $G = 0_{2\times2}$, and the NCE-SCE equation system is iterated 100 times. The state trajectories of a single realization can be seen for a population of all altruistic agents ($\lambda = 0$) in Fig. 1. Only 10 minor agents are displayed for clarity. In Fig. 2, we present the instantaneous loss of each agent. The same parameters as previously are used, except that we set $\sigma = 0$ for clarity. We run the experiment twice, one for $\lambda = 1$ and one for $\lambda = 0$. Here the thick graph line with squares shows the instantaneous loss paid by the major agent when the minor agents are all acting egoistically ($\lambda = 1$), and the thick graph line without squares shows the instantaneous loss when all minor agents are altruistic ($\lambda = 0$). Likewise, the thin graph line with squares shows the loss paid by a minor agent in an egoistic minor agent population and the one without squares shows the altruistic case. The effect of altruistic behaviour on performance can easily be observed. In Fig. 3, we plot the loss function of a minor agent $\mathcal{A}_i$ as a function of time. Here all the rest of the minor agents in the system apply $\lambda_{-i} = 0$, therefore they are all altruistic. This plot shows the performance with respect to altering $\lambda_i$ of agent $\mathcal{A}_i$ from 1 to 0, when all other agents are altruistic. As shown in Proposition 5.1, the cost for agent $\mathcal{A}_i$ is maximum for $\lambda_i = 0$, and minimum when $\lambda_i$ is 1. In Fig. 4, we plot the cost paid by a minor agent with respect to the *egoism degree* coefficient $\lambda$. Note that this plot shows the case where all minor agents in the system uniformly apply the same $\lambda$ parameter. As shown in Proposition 5.1, the cost increases with an increasing to $\lambda$.

## VII. CONCLUSION

This paper considers decentralized control for large population LQG dynamic games involving a major agent and a large number of minor agents where each agent optimizes with respect to a convex combination of its own cost and a

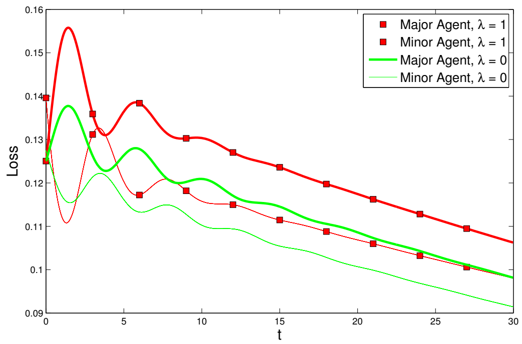


Fig. 2. Loss comparison

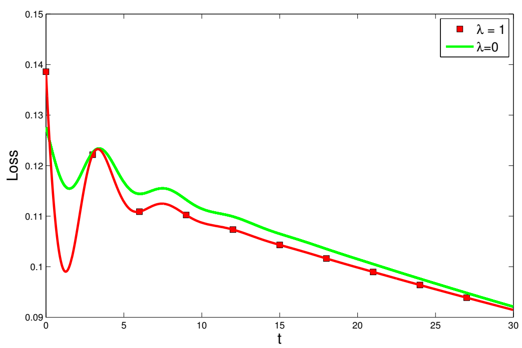


Fig. 3. Egoistic action against altruistic population

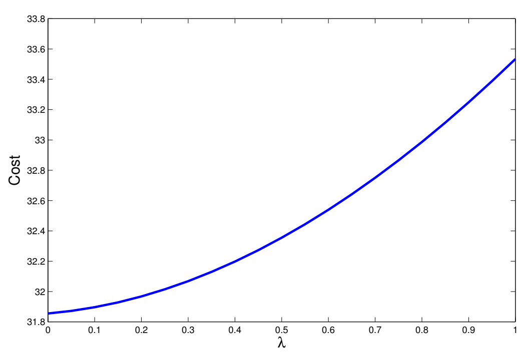


Fig. 4. Cost wrt $\lambda$ in a uniform population

social cost for the partial optimization of a social objective. A mean field approximation is used such that the game problem can be analysed in the population limit, where the aggregate effect of all minor agents is characterized by linear stochastic differential equations driven by the state of the major agent. The NCE-SCE analysis yields decentralized strategies for all agents. Decentralized control synthesis for the general case where the weight assigned to the individual cost changes continuously across the population is developed and an asymptotic Nash equilibrium theorem is proved. Then, making the egoism degree a decision action for each agent, we analyse the resulting equilibria and performance of the mass and any individual agent. We establish the result which at first sight may be thought to be paradoxical: even though a game where all agents in the system play altruistically gives the minimum cost per head, this action profile is not stable in the sense that it is always more profitable for an agent to play egoistically.

## REFERENCES


[1] 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 *42nd IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2003, pp. 98–103.

[2] ——, "Large population cost-coupled LQG problems with non-uniform agents: individual-mass behaviour and decentralized $\epsilon$ - Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, Sep. 2007.

[3] ——, "Social dynamics in mean field LQG control: Egoistic and altruistic agents," in *49th IEEE Conference on Decision and Control*, 2010, pp. 3140–3145.

[4] M. Huang, "Large-population LQG games involving a major player: the Nash Certainty Equivalence principle," *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 3318–3353, 2010.

[5] P. E. Caines, *Mean Field Stochastic Control*. Bode Lecture at the 48th IEEE Conf. on Decision and Control and 28th Chinese Control Conf., Dec. 2009. [Online]. Available: http://www.cim.mcgill.ca/~arman/Shanghai2009/Shanghai2009.zip

[6] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamic features of escape panic," *Nature*, vol. 407, pp. 487–490, Sep. 2000.

[7] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Stable flocking of mobile agents, part i: fixed topology," in *42nd IEEE Conference on Decision and Control*, Maui, Hawaii, 2003, pp. 2010–2015.

[8] Y. Liu and K. M. Passino, "Stable social foraging swarms in a noisy environment," *IEEE Transactions on Automatic Control*, vol. 49, pp. 30–44, Jan. 2004.

[9] D. J. Low, "Following the crowd," *Nature*, vol. 407, pp. 465–466, Sep. 2000.

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

[11] 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, 2008.

[12] H. Tembine, J.-Y. L. Boudec, R. El-Azouzi, and E. Altman, "Mean field asymptotics of Markov decision evolutionary games and teams," in *Int. Conf. on Game Theory for Networks (GameNets 2009)*, Istanbul, 2009, pp. 140–150.

[13] 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," *Special issue in honour of the 65th birthday of Tyrone Duncan, Communications in Information and Systems*, vol. 6, pp. 221–252, Nov. 2006.

[14] A. C. Kizilkale and P. E. Caines, "Stochastic adaptive Nash certainty equivalence control: Self-identification case," in *19th Intl. Symp. on Math. Theory of Networks and Systems (MTNS 2010)*, July 2010, pp. 2093–2099.

[15] ——, "Stochastic adaptive Nash certainty equivalence control: Population parameter distribution estimation," in *49th IEEE Conference on Decision and Control*, December 2010, pp. 6169–6176.

[16] ——, "Nash certainty equivalence (mean field) stochastic adaptive control," *IEEE Transactions on Automatic Control*, 2011, submitted.

[17] T. Roughgarden and E. Tardos, "Bounding the inefficiency of equilibria in nonatomic congestion games," *Games and Economic Behavior*, vol. 47, pp. 389–403, 2004.

[18] M. Huang, P. E. Caines, and R. P. Malhamé, "Social optima in mean-field LQG control: centralized and decentralized strategies," in *Proc. the 47th Allerton Conf. on Communication, Control and Computing*, Monticello, Illinois, Sep. 2009.

[19] S. Hart, "Values of mixed games," *International J. Game Theory*, vol. 2, pp. 69–86, 1973.

[20] A. C. Kizilkale and P. E. Caines, "Initial investigations of the emergence of coalitions in mean field stochastic systems," Department of Electrical and Computer Engineering, McGill University, Tech. Rep., 2011.