ϵ-Nash Equilibria for Partially Observed LQG Mean Field Games With a Major Player

Peter E. Caines, *Life Fellow, IEEE*, and Arman C. Kizilkale, *Member, IEEE*

***Abstract*—Huang (2010) and Nguyen and Huang (2012) solved the linear quadratic mean field systems and control problem in the case where there is a major agent (i.e. non-asymptotically vanishing as the population size goes to infinity) together with a population of minor agents (i.e. individually asymptotically negligible). The new feature in this case is that the mean field becomes stochastic and then, by minor agent state extension, the existence of ϵ-Nash equilibria together with the individual agents' control laws that yield the equilibria may be established. This paper presents results initially announced by Caines and Kizilkale (2013, 2014) where it is shown that if the major agent's state is partially observed by the minor agents, and if the major agent completely observes its own state, all agents can recursively generate estimates (in general individually distinct) of the major agent's state and the mean field, and thence generate feedback controls yielding ϵ-Nash equilibria.**

I. Introduction

MEAN Field Game (MFG) systems theory establishes the existence of approximate Nash equilibria together with the corresponding individual strategies for stochastic dynamical systems in games involving a large number of agents. The equilibria are termed ϵ-Nash equilibria and are generated by the local, limited information feedback control actions of each agent in the population, where the feedback control actions constitute the best response of each agent with respect to the precomputed behaviour of the mass of agents and where the approximation error converges to zero as the population size goes to infinity.

The determination of an approximate equilibrium and the corresponding individual agent control actions in the complex, arbitrarily large finite population case (i.e. the domain of application) is achieved by exploiting its relationship with the infinite population limit problem with its far simpler description and solution. Specifically, the solution to the infinite population problem is obtained via the mean field (MF) Hamilton-Jacobi-Bellman partial differential equation (PDE) and the (McKean-Vlasov) Fokker-Planck-Kolmogorov PDE equations which are linked to each other by the state distribution of a generic agent, otherwise known as the system's mean field. This linked pair of Hamilton-Jacobi-Bellman (HJB) and Fokker-Planck-Kolmogorov (FPK) PDEs is referred to as the Mean Field Game (MFG) equations.

The analysis of this set of problems originated in [1]–[3] (see [4]), and independently in [5], [6]. In [7] and [8] the authors analyse and solve the linear quadratic systems case where there is a major agent (i.e. non-asymptotically vanishing as the population size goes to infinity) together with a population of minor agents (i.e. individually asymptotically negligible). The new feature in this case is that the mean field becomes stochastic and then, by minor agent state extension, the existence of ϵ-Nash equilibria may be established together with the individual agents' control laws that yield the equilibria [8].

In the purely minor agent case the mean field is deterministic and this obviates the need for observations on other agents' states for the generation via recursive filtering of estimates of the global systems state or the mean field. This is a separate issue from that of an agent estimating its own state (self state for short) from partial observations on that state, see [9]. However, a new situation arises for systems with a major agent whose state is partially observed and which must in general be estimated by each agent in order to generate an equilibrium best response control.

The main result of the present paper (first announced in [10], [11]) is that subject to the assumptions that (i) the major agent's state is partially observed by the minor agents and (ii) the major agent has complete observations of its own state (aka self-observation), the minor agents can recursively generate estimates (in general individually distinct) of the major agent's state and the mean field, and thence generate feedback controls yielding ϵ-Nash equilibria.

II. Mean Field of a Parameterized Family of SDEs

In this section we first give a general definition of the terms mean field and observation dependent mean field. Then in Section III we present the dynamics and cost functions for the LQG major-minor mean field game (MM-MFG) problem in the completely observed case together with the corresponding mean field equations and the ϵ-Nash equilibirium thereom.

In the general equation scheme (1), for each member of the population of agents $\mathcal{A}_i$, $1 \le i \le N < \infty$, let x lie in the state space $\mathbb{R}^n$, u take values in $\mathbb{R}^m$, the drift take values in $\mathbb{R}^n$, and the diffusion coefficient matrix be conformable with the

Manuscript received October 7, 2015; revised June 22, 2016 and June 28, 2016; accepted November 2, 2016. Date of publication December 8, 2016; date of current version June 26, 2017. This work was supported by NSERC and AFOSR grants. Recommended by Associate Editor R. M. Jungers.

P. E. Caines is with the Department of Electrical and Computer Engineering and the Centre for Intelligent Machines, McGill University, Montreal Canada (e-mail: peterc@cim.mcgill.ca).

A. C. Kizilkale is with Automat, Montreal Canada (e-mail: arman@automat.ai).

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.2016.2637347

standard Wiener process w lying in $\mathbb{R}^r$, $\theta \in \Theta$, where Θ is a compact parameter set in $\mathbb{R}^s$ (later to be restricted to a finite set), and let the SDE have a unique strong solution:

$$dx_i = \frac{1}{N}\sum_{j=1}^{N} f_\theta(x_i, x_j, u_\theta(x_i))dt + \frac{1}{N}\sum_{j=1}^{N} g_\theta(x_i, x_j)dw_i,$$
$$0 \le t \le T,\ 1 \le i \le N < \infty. \tag{1}$$

Then the *mean field* of the population of stochastic systems (1) is defined to be the probability law

$$P^x_{\theta,t}(B) \triangleq \lim_{N\to\infty} P(x_i^{[N]}(\theta)(t) \in B), \tag{2}$$

whenever the limit for each $x_i(\theta)$ exists, and is independent of i, where $B \in \mathcal{B}(\mathbb{R}^n)$ and where $x_i^{[N]}$ denotes the state process of the ith agent subsystem.

In case a density $\mu^x_{\theta,t}(\cdot)$ exists for $P^x_{\theta,t}(\cdot)$, it will also, by abuse of language, be called the mean field of the associated system.

Consistency Property

It is an important consistency property of the formulation above that the given conditions are sufficient to imply (see [2] and [12]), as N tends to ∞, both (i) the convergence of a generic member of the sequence of solutions to the scheme (1) in L^1, and hence in probability, to the unique finite dimensional component of the solution pair (x^∞, μ^∞) for the associated McKean-Vlasov equation and (ii) the convergence in (2) of the sequence of distributions of the states of the solutions to (1) to the distribution appearing in the solution pair of the McKean-Vlasov equation (which is generated by the associated Fokker-Planck-Kolmogorov equation). It follows that the distribution component of the McKean-Vlasov solution pair is the mean field of the system (1).

Consider a population of agents $\mathcal{A}_i, 1 \le i \le N < \infty$, for which the evolution of the state of each controlled linear SDE with parameter θ on $[0;\infty)$ is given by

$$dx_i(\theta) = A_\theta x_i(\theta)dt + \bar{A}_\theta\left(\frac{1}{N}\sum_{j=1}^{N} x_j(\theta)\right)dt+$$
$$B_\theta u_i(\theta)dt + [C_\theta + \bar{C}_\theta(\frac{1}{N}\sum_{j=1}^{N} x_j(\theta))]dw_i,$$
$$0 \le t \le T, \quad 1 \le i \le N, \tag{3}$$

where the coefficient matrices are such that A_θ denotes the system state matrix, $\bar{A}_\theta$ the population empirical mean state matrix, B_θ the control input matrix, and C_θ and $\bar{C}_\theta$ the diffusion intensity matrices, the latter being a function of the population empirical mean state process, and where $u_i(\theta)$ is assumed to be affine in the state $x_i(\theta)$ of the agent $\mathcal{A}_i$ and weighted averages of the states of the agent population.

Then the set of linear stochastic systems is said to have the *Gaussian mean field*

$$P^x_{\theta,t}(\cdot) \triangleq N(\bar{x}(\theta,t), \Sigma(\theta,t)), \tag{4}$$

if

$$\lim_{N\to\infty} P(x_i^{[N]}(\theta)(t) \in B) = P^x_\theta(B), \quad \forall B \in \mathcal{B}(\mathbb{R}^n). \tag{5}$$

Exogenous Process Dependent Mean Fields

Let there be a process y which is exogenous with respect to the system (3) and jointly distributed with the initial conditions and Wiener processes of (1) on a suitably expanded probability space. Next, for each t, x, let $u_\theta(y,x)$ be affine with respect to (y,x) and hence, for each x, measurable with respect to the σ-field generated by $\{y_s, 0 \le s \le t\}$. Finally assume that y has almost surely (a.s.) bounded sample paths on $[0,T]$ for all finite T. It follows that for each N, $1 \le N < \infty$, there exist a.s. unique solutions to the family of linear stochastic differential equations (3).

Then, following (2) above, the y-dependent distribution

$$P^x_{\theta,t,y}(\cdot) \triangleq N_{\theta,y^t}(\bar{x}^y(t), \Sigma(t)), \tag{6}$$

is called the *Gaussian y-(exogenous process) dependent mean field* of the system (2) if

$$N_{\theta,y^t}(\bar{x}^y(t), \Sigma(t)) = \lim_{N\to\infty} P_{y^t}(x_i^{[N]}(\theta)(t) \in B)$$
$$\equiv \lim_{N\to\infty} P(x_i^{[N]}(\theta)(t) \in B|\mathcal{F}_{y^t}) \quad \forall B \in \mathcal{B}(\mathbb{R}^n). \tag{7}$$

III. Major-Minor Agent LQG Systems

In this section we give a succinct summary of the LQG major-minor agent MF framework together with the principal ϵ-Nash Equilibrium result. All random variables and stochastic processes are taken to be defined on the underlying probability space $(\Omega, \mathcal{F}, P)$.

Dynamics: Finite Population

Specializing to the case where $\bar{A}_\theta$ and $\bar{C}_\theta$ are set to zero for simplicity, and following [7], we consider a major agent together with a large population of N stochastic dynamic minor agents as given by

$$dx_0 = [A_0x_0 + B_0u_0]dt + D_0dw_0, \tag{8a}$$
$$dx_i = [A(\theta_i)x_i + B(\theta_i)u_i + Gx_0]dt + Ddw_i, \tag{8b}$$

$0 \le t < \infty$, $1 \le i \le N < \infty$, where $\theta_i \in \Theta$ denotes the parameter of the i-th system and the state of the major agent is subscripted by 0. Here $x_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^m$, and $w = \{w_i, 0 \le i \le N\}$ denotes $(N+1)$ independent standard Wiener processes in $\mathbb{R}^r$, where w is progressively measurable with respect to the filtration $\mathcal{F}^w \triangleq \{\mathcal{F}^w_t \subset \mathcal{F}; t \ge 0\}$. Note that the common agent $\mathcal{A}_0$ affects each minor agent through its dynamics.

Initial Conditions

The initial states $\{x_i(0), 0 \le i \le N < \infty\}$ are identically distributed and mutually independent and also independent of $\mathcal{F}^w_\infty$; $\mathbb{E}w_iw_i^\top = \Sigma$, $0 \le i \le N < \infty$, and $\mathbb{E}\|x_i(0)\|^2 \le C < \infty$, $0 \le i \le N < \infty$, with Σ and C independent of N.

Control σ-Fields

We now introduce two admissible sets of controls. The null set augmented σ-field $\mathcal{F}_{i,t}, 1 \leq i \leq N$, is defined to be the increasing family of null set augmented σ-fields generated by $(x_i(\tau);\, 0 \leq \tau \leq t)$, and by definition $\mathcal{F}_{0,t}$ is the increasing family of σ-fields generated by $(x_0(\tau);\, 0 \leq \tau \leq t)$. $\mathcal{F}_t^N$ is the increasing family of σ-fields generated by the set $\{x_j(\tau), x_0(\tau);\, 0 \leq \tau \leq t, 1 \leq j \leq N\}$. By definition the set $\mathcal{U}_0$ consists of the feedback controls adapted to the filtration $\{\mathcal{F}_{0,t};\, t \geq 0\}$. The set of control inputs $\mathcal{U}_i,\ 1 \leq i \leq N$, based upon the local information set of the minor agents $\mathcal{A}_i,\ 1 \leq i \leq N$, consists of the feedback controls adapted to $\{\mathcal{F}_{i,t} \vee \mathcal{F}_{0,t};\, t \geq 0\}$, $1 \leq i \leq N$, while $\mathcal{U}_g^N$ is adapted to $\{\mathcal{F}_t^N,\, t \geq 0\}$, $1 \leq N < \infty$.

Minor Agents' Types

The restriction is now imposed that the non-uniform minor agents are grouped into a finite number of K identical parametric types, $1 \leq K < \infty$, and hence Θ has finite cardinality K; correspondingly we adopt the notation

$$\mathcal{I}_k = \{i : \theta_i = k,\ 1 \leq i \leq N\}, \quad N_k = |\mathcal{I}_k|,$$

$1 \leq k \leq K$. Then, $\pi^N = (\pi_1^N, \ldots, \pi_K^N)$, $\pi_k^N = N_k/N$, $1 \leq k \leq K$, shall denote the empirical distribution of the parameters $(\theta_1, \ldots, \theta_N)$ sampled independently of the initial conditions and Wiener processes of the agents $\mathcal{A}_i,\ 1 \leq i \leq N$.

A1: There exists π such that $\lim_{N\to\infty} \pi^N = \pi$ a.s.

We shall denote by $x_k^{N_k}$ the empirical state average $x_k^{N_k} = \frac{1}{N_k}\sum_{j=1}^{N_k} x_j^k$, $1 \leq k \leq K$, of systems of type k, and by x^N the nK vector of empirical state averages of all K types.

Performance Functions

The individual infinite horizon performance, or cost, function for the major agent is specified by

$$J_0^N(u_0, u_{-0}) = \mathbb{E}\int_0^\infty e^{-\rho t}\left\{\left\|x_0 - \Phi(x^N)\right\|_{Q_0}^2 + \|u_0\|_{R_0}^2\right\} dt, \tag{9}$$

$$\Phi(\cdot) := H_0 x^N + \eta_0,$$

and the individual infinite horizon cost for a minor agent $\mathcal{A}_i$, $1 \leq i \leq N$, is specified as

$$J_i^N(u_i, u_{-i}) = \mathbb{E}\int_0^\infty e^{-\rho t}\left\{\left\|x_i - \Psi(x^N)\right\|_Q^2 + \|u_i\|_R^2\right\} dt, \tag{10}$$

$$\Psi(\cdot) := H_1 x_0 + H_2 x^N + \eta.$$

Finite and Infinite Population Control Laws and Dynamics

We now construct the mean field state equations for the system (8a)–(8b) subject to assumptions on the class of control inputs.

For each generic minor agent $\mathcal{A}_i$ of type k, $1 \leq k \leq K$, consider uniform (with respect to i) feedback controls $u_i^k \in \mathcal{U}_{i,L} \subset \mathcal{U}_i$, where $\mathcal{U}_{i,L}$ consists of controls which are linear time invariant functions of the following:

1) state x_i,
2) population states x_j^k, $1 \leq k \leq K$, uniformly by type,
3) major agent state x_0,
4) bounded continuous functions of time $m^k(\cdot) \in \mathbf{C}_b[0,\infty)$.

Hence,

$$u_i^k(t) = L_1^k x_i^k(t) + \sum_{l=1}^K \sum_{j=1}^{N_l} L_2^{k,l} x_j^l + L_3^k x_0(t) + m^k(t), \tag{11}$$

$0 \leq t < \infty$, for some time invariant matrices L_1^k, $L_2^{k,l}$ and L_3^k of appropriate dimension, where m^k permits the class of controls u_i^k to include the off-set functions appearing in the solution of the optimal tracking problem. It is assumed, in addition, that the coefficient matrices $L_2^{k,l}$ depend upon N_l, and satisfy $N_l L_2^{k,l} \to \bar{L}_2^{k,l}$ as $N_l \to \infty$ for all k, $1 \leq k \leq K$.

Under the stated assumptions for (8a)–(8b) and (11), unique L^2 solutions exist a.s. on $[0,t]$, $0 \leq t < \infty$, to the stochastic differential equation of the overall feedback system for any N, $1 \leq N < \infty$, where substitution of u_i^k into (8b) yields

$$\begin{aligned} dx_i^k &= (A_k + B_k L_1^k) x_i^k dt + B_k\left(\sum_{l=1}^K L_2^{k,l}\left(\sum_{j=1}^{N_l} x_j^l\right)\right) dt \\ &\quad + B_k(L_3^k x_0 + m^k)dt + G x_0 dt + D dw_i^k, \end{aligned} \tag{12}$$

$1 \leq k \leq K$. Hence the N_k systems of type k have an empirical state average which satisfies

$$\begin{aligned} S_k^{N_k}: \quad dx_k^{N_k} &= (A_k + B_k L_1^k) x_k^{N_k} dt + B_k \sum_{l=1}^K N_l L_2^{k,l} x_l^{N_l} dt \\ &\quad + B_k(L_3^k x_0 + m^k)dt + G x_0 dt \\ &\quad + D\frac{1}{N_k}\sum_{i=1}^{N_k} dw_i^k, \quad 1 \leq k \leq K. \end{aligned} \tag{13}$$

Proposition 3.1: Consider the family of systems $S_k^{N_k}$, $1 \leq k \leq K$, with deterministic inputs $\bar{m}^k \in \mathbf{C}_b[0,\infty)$ and a given sample path x_0 generated by (8a) independent of $\{w_i,\ 1 \leq i < \infty\}$. Then, subject to **A1** (i) unique L^2 solutions exist a.s. on $[0,\infty)$ to the system of SDEs $S_k^{N_k}$, $1 \leq k \leq K$, and (ii) the pointwise in time L^2 limits of the set of trajectories of $S_k^{N_k}$ as $N \to \infty$ are given by the solutions to the mean field differential equations on $[0,\infty)$

$$\begin{aligned} d\bar{x}^k &= (A_k + B_k L_1^k)\bar{x}^k dt + B_k\left(\sum_{l=1}^K \bar{L}_2^{k,l}\bar{x}^l\right) dt \\ &\quad + B_k(L_3^k x_0 + m^k)dt + G x_0 dt, \quad 1 \leq k \leq K, \end{aligned} \tag{14}$$

where $\bar{x}^k$ denotes the mean of the mean field of the k-th type subsystem, and which can be written as

$$d\bar{x}^k = A_k \bar{x}^k dt + B_k \bar{u}_k dt + G x_0 dt, \quad 1 \leq k \leq K, \tag{15}$$

where the kth type feedback control $\bar{u}_k$ is defined by

$$\bar{u}_k(t) = L_1^k \bar{x}^k(t) + \sum_{l=1}^{K} \bar{L}_2^{k,l} \bar{x}^l + L_3^k x_0(t) + m^k(t).$$

Finally, in an evident notation, the state vector $\bar{x}$ with components $[\bar{x}^1, \ldots, \bar{x}^K]$ satisfies

$$d\bar{x} = \bar{A}\bar{x}dt + \bar{G}x_0 dt + \bar{m}(t)dt, \tag{16}$$

where $\bar{x}(0)$ denotes the vector array of $[\mathbb{E}x^1(0), \mathbb{E}x^2(0), \cdots, \mathbb{E}x^K(0)]$ and $\bar{x}$ constitutes the mean value of the x_0-dependent mean field of the controlled system (8a), (8b), (11).

Proof: First, assertion (i) follows immediately from the existence, uniqueness and pointwise in time L^2 boundedness of the solution to the set of $N = \sum_{k=1}^{K} N_k$ stochastic differential equations (12). Next, to establish (ii), it will be shown that for each t, $x_k^{N_k}(t)$, $1 \le N_k \le N < \infty$, constitutes an L^2 Cauchy sequence whose limit $\bar{x}^k(t)$, $t \in [0, \infty)$, is generated by (16). For clarity and simplicity we first consider the case of a single type k only with $L_2^{k,k} = 0$. Let

$$x_{N,M} \triangleq x^N - x^M \triangleq \frac{1}{N}\sum_{i=1}^{N} x_i - \frac{1}{M}\sum_{j=1}^{M} x_j, \tag{17}$$

$$dw^N - dw^M \triangleq \frac{1}{N}\sum_{i=1}^{N} dw_i - \frac{1}{M}\sum_{j=1}^{M} dw_j. \tag{18}$$

Then in an obvious notation the solution to (13) is given by

$$x_N(t) = \Phi(t,0)x_N(0) + \int_0^t \Phi(t,\tau)(Gx_0(\tau) + m(\tau))d\tau + \frac{D}{N}\sum_{i=1}^{N}\int_0^t \Phi(t,\tau)dw_i(\tau), \tag{19}$$

and so

$$x_{N,M}(t) = \Phi(t,0)x_{N,M}(0) + \int_0^t \Phi(t,\tau)\{\frac{D}{N}\sum_{i=1}^{N} dw_i(\tau) - \frac{D}{M}\sum_{j=1}^{M} dw_j(\tau)\}.$$

Hence,

$$\mathbb{E}\|x_{N,M}(t)\|^2 \le 2\mathbb{E}\|\Phi(t,0)x_{N,M}(0)\|^2 + 2\mathbb{E}\|\int_0^t D\Phi(t,\tau)d(w^N - w^M)\|^2.$$

But

$$\mathbb{E}\|\Phi(t,0)x_{N,M}(0)\|^2 \le \|\Phi(t,0)\|^2 \mathbb{E}\|x_{N,M}(0)\|^2 \tag{20}$$

$$\le 2\|\Phi(t,0)\|^2[\frac{1}{N^2}N\mathbb{E}\|x_i(0)\|^2 + \frac{1}{M^2}M\mathbb{E}\|x_i(0)\|^2] \tag{21}$$

$$= 2\|\Phi(t,0)\|^2\left(\frac{1}{N} + \frac{1}{M}\right)C, \tag{22}$$

by $x_i(0)$, $1 \le i < \infty$, inedpendent and identically distributed with $\mathbb{E}x_i(0) = 0$, $1 \le i < \infty$, and

$$\begin{aligned}
&\mathbb{E}\|D\int_0^t \Phi(t,\tau)d(w^N - w^M)(\tau)\|^2 \\
&\le 2[\mathbb{E}\|\frac{D}{N}\sum_{i=1}^{N}\int_0^t \Phi(t,\tau)dw_i(\tau)\|^2 \\
&+ \mathbb{E}\|\frac{D}{M}\sum_{j=1}^{M}\int_0^t \Phi(t,\tau)dw_j(\tau)\|^2] \\
&\le \frac{2\|D\|^2}{N}\mathrm{tr}\int_0^t \Phi(t,\tau)\Phi^\top(\tau,t)d\tau \\
&\quad + \frac{2\|D\|^2}{M}\mathrm{tr}\int_0^t \Phi(t,\tau)\Phi^\top(\tau,t)d\tau.
\end{aligned}$$

Consequently for each t, $0 \le t < \infty$,

$$\mathbb{E}\|x_N(t) - x_M(t)\|^2 = \mathbb{E}\|x_{N,M}(t)\|^2 \to 0,$$

as $N, M \to \infty$; and so by the completeness of L^2, there exists $\bar{x}(t)$, $0 \le t < \infty$, satisfying $\mathbb{E}\|x_N(t) - \bar{x}(t)\|^2 \to 0$ as $N \to \infty$.

It follows that taking L^2 limits in (19) yields

$$\bar{x}(t) = \Phi(t,0)\bar{x}(0) + \int_0^t \Phi(t,\tau)(\bar{G}x_0(\tau) + m(\tau))d\tau, \tag{23}$$

since in L^2 as $N \to \infty$, $x_N(t) \to \bar{x}(t)$, $x_N(0) \to \bar{x}(0)$ and $(1/N)\sum_{i=1}^{N}\int_0^t \Phi(t,\tau)dw_i(\tau) \to 0$. But the ordinary differential equation (ODE) to which (23) is the solution is given by (15).

In the general multiple types case the proof follows the same path with the empirical average x^N being extended to the vector array of $[x_1^{N_1}, \ldots, x_K^{N_K}]$ and where we note that **A1** ensures that as $N \to \infty$ so does each of the subpopulation magnitudes $N_1, \ldots N_K$. In the general case, the second term on the right hand side of (13) depends upon $N_1, \ldots N_K$, and hence, the fundamental matrix of (13) depends upon the multiple index $N_1, \ldots N_K$ of the population sizes.

Consequently we now replace $\Phi(t,\tau)$ in (19)–(23) by the fundamental matrix $\Phi_{N_1,\ldots,N_K}(t,\tau)$ and observe that $\Phi_{N_1,\ldots,N_K}(t,\tau)$ converges uniformly to a limit denoted $\Phi(t,\tau)$ for t and τ lying in any compact interval. The general result now follows as in the single type case. ■

Henceforth, *infinite population case* refers collectively to *(i)* a major agent system (8a), (ii) an infinite family of agent systems $\mathcal{A}_i$, $1 \le i < \infty$, as in (8b) satisfying the (infinite) subpopulation fractions $\pi = (\pi_1, \ldots, \pi_K)$, any member of which is referred to as a generic agent (of appropriate type) and (iii) the corresponding mean field.

Major Agent's Dynamics and Tracking Problem

We now consider the major agent's state extension $x_0^{\bar{x}} \triangleq [x_0, \bar{x}]$ by the mean field and obtain the major agent's dynamics in the infinite population case as (see Sec 4.1 in [7])

$$\begin{bmatrix} dx_0 \\ d\bar{x} \end{bmatrix} = \begin{bmatrix} A_0 & 0_{nK\times n} \\ \bar{G} & \bar{A} \end{bmatrix} \begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} dt + \begin{bmatrix} B_0 \\ 0_{nK\times m} \end{bmatrix} u_0 dt + \begin{bmatrix} 0_{n\times 1} \\ \bar{m} \end{bmatrix} dt + \begin{bmatrix} D_0 & 0_{n\times nK} \\ 0_{nK\times r} & 0_{nK\times nK} \end{bmatrix} \begin{bmatrix} dw_0 \\ 0_{nK\times 1} \end{bmatrix}, \tag{24}$$

which may be written

$$dx_0^{\bar{x}} = \mathbb{A}_0 x_0^{\bar{x}} dt + \mathbb{B}_0 u_0 dt + \mathbb{M}_0 dt + \mathbb{D}_0 \begin{bmatrix} dw_0 \\ 0_{nK\times 1} \end{bmatrix}, \tag{25}$$

where

$$\mathbb{A}_0 = \begin{bmatrix} A_0 & 0_{nK\times n} \\ \bar{G} & \bar{A} \end{bmatrix}, \qquad \mathbb{B}_0 = \begin{bmatrix} B_0 \\ 0_{nK\times m} \end{bmatrix},$$
$$\mathbb{M}_0 = \begin{bmatrix} 0_{n\times 1} \\ \bar{m} \end{bmatrix}, \qquad \mathbb{D}_0 = \begin{bmatrix} D_0 & 0_{n\times nK} \\ 0_{nK\times r} & 0_{nK\times nK} \end{bmatrix}, \tag{26}$$

where

$$\bar{A} = \begin{bmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_K \end{bmatrix}, \text{ where } \bar{A}_k \in \mathbb{R}^{n\times nK}. \tag{27}$$

In the infinite population case, the individual cost for the major agent is given by

$$J_0^\infty(u_0, u_{-0}) = \mathbb{E}\int_0^\infty e^{-\rho t}\left\{\left\|x_0 - \Phi(\bar{x})\right\|_{Q_0}^2 + \|u_0\|_{R_0}^2\right\} dt, \tag{28}$$

$$\Phi(\cdot) = H_0^\pi \bar{x} + \eta_0,$$

where

$$H_0^\pi = \pi \otimes H_0 \triangleq [\pi_1 H_0,\ \pi_2 H_0,\ \cdots,\ \pi_K H_0].$$

We then have the major agent tracking problem solution:

$$\rho\Pi_0 = \Pi_0\mathbb{A}_0 + \mathbb{A}_0^\top\Pi_0 - \Pi_0\mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0 + Q_0^\pi,$$
$$\rho s_0 = \frac{ds_0}{dt} + (\mathbb{A}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0)^\top s_0 + \Pi_0\mathbb{M}_0 - \bar{\eta}_0,$$
$$u_0^\circ = -R_0^{-1}\mathbb{B}_0^\top\left[\Pi_0(x_0^\top, \bar{x}^\top)^\top + s_0\right],$$

where

$$\bar{\eta}_0 = [I_{n\times n}, -H_0^\pi]^\top Q_0 \eta_0,$$
$$Q_0^\pi = [I_{n\times n}, -H_0^\pi]^\top Q_0 [I_{n\times n}, -H_0^\pi].$$

Minor Agent's Dynamics and Tracking Problem

Similarly we introduce the minor agent's state extended by the major agent's state and the x_0 dependent mean field to obtain $x_i^{0,\bar{x}} \triangleq [x_i, x_0, \bar{x}]$. Then each minor agent's dynamics in the infinite population case is given by (see Section 4.2 in [7])

$$\begin{bmatrix} dx_i \\ dx_0 \\ d\bar{x} \end{bmatrix} = \begin{bmatrix} A_k & [G\ \ 0_{n\times nK}] \\ 0_{(nK+n)\times n} & \mathbb{A}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0 \end{bmatrix} \begin{bmatrix} x_i \\ x_0 \\ \bar{x} \end{bmatrix} dt + \begin{bmatrix} B_k \\ 0_{(nK+n)\times m} \end{bmatrix} u_i dt + \begin{bmatrix} 0_{n\times 1} \\ \mathbb{M}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top s_0 \end{bmatrix} dt + \begin{bmatrix} D & 0 \\ 0 & \mathbb{D}_0 \end{bmatrix} \begin{bmatrix} dw_i \\ \begin{bmatrix} dw_0 \\ 0_{nK\times 1} \end{bmatrix} \end{bmatrix}, \tag{29}$$

which we may write as

$$dx_i^{0,\bar{x}} = \mathbb{A}_k x_i^{0,\bar{x}} dt + \mathbb{B}_k u_i dt + \mathbb{M} dt + \mathbb{D} \begin{bmatrix} dw_i \\ \begin{bmatrix} dw_0 \\ 0_{nK\times 1} \end{bmatrix} \end{bmatrix}, \tag{30}$$

with the matrices above defined as follows:

$$\mathbb{A}_k = \begin{bmatrix} A_k & [G\ \ 0_{n\times nK}] \\ 0_{(nK+n)\times n} & \mathbb{A}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top\Pi_0 \end{bmatrix},$$
$$\mathbb{B}_k = \begin{bmatrix} B_k \\ 0_{(nK+n)\times m} \end{bmatrix} \quad \mathbb{M} = \begin{bmatrix} 0_{n\times 1} \\ \mathbb{M}_0 - \mathbb{B}_0 R_0^{-1}\mathbb{B}_0^\top s_0 \end{bmatrix},$$
$$\mathbb{D} = \begin{bmatrix} D & 0 \\ 0 & \mathbb{D}_0 \end{bmatrix}.$$

The individual cost for a minor agent $\mathcal{A}_i$, $1 \le i \le N$, is given by

$$J_i^\infty(u_i, u_{-i}) = \mathbb{E}\int_0^\infty e^{-\rho t}\left\{\left\|x_i - \Psi(\bar{x})\right\|_Q^2 + \|u_i\|_R^2\right\} dt, \tag{31}$$

$$\Psi(\cdot) = H_1 x_0 + H_2^\pi \bar{x} + \eta,$$

where

$$H_2^\pi = \pi \otimes H_2,$$

and the minor agent tracking problem solution is given by

$$\rho\Pi_k = \Pi_k\mathbb{A}_k + \mathbb{A}_k^\top\Pi_k - \Pi_k\mathbb{B}_k R^{-1}\mathbb{B}_k^\top\Pi_k + Q^\pi,$$
$$\rho s_k = \frac{ds_k}{dt} + (\mathbb{A}_k - \mathbb{B}_k R^{-1}\mathbb{B}_k^\top\Pi_k)^\top s_k + \Pi_k\mathbb{M} - \bar{\eta},$$
$$u_i^\circ = -R^{-1}\mathbb{B}_k^\top\left[\Pi_k(x_i^\top, x_0^\top, \bar{x}^\top)^\top + s_k\right],$$

where

$$\bar{\eta} = [I_{n\times n}, -H_1, -H_2^\pi]^\top Q\eta,$$
$$Q^\pi = [I_{n\times n}, -H_1, -H_2^\pi]^\top Q [I_{n\times n}, -H_1, -H_2^\pi].$$

Now define

$$\Pi_k = \begin{bmatrix} \Pi_{k,11} & \Pi_{k,12} & \Pi_{k,13} \\ \Pi_{k,21} & \Pi_{k,22} & \Pi_{k,23} \\ \Pi_{k,31} & \Pi_{k,32} & \Pi_{k,33} \end{bmatrix}, \quad 1 \le k \le K,$$

and $\mathbf{e}_k = [0_{n\times n}, \ldots, 0_{n\times n}, I_n, 0_{n\times n}, \ldots, 0_{n\times n}]$, where the $n\times n$ identity matrix I_n is at the kth block.

MM System ϵ-Nash Equilibria

This notation permits a compact description of the MM-MFG Equations determining $\bar{A}$, $\bar{G}$, $\bar{m}$ via the consistency requirements:

$$\begin{aligned}
\rho\Pi_0 &= \Pi_0\mathbb{A}_0 + \mathbb{A}_0^\top\Pi_0 - \Pi_0\mathbb{B}_0R_0^{-1}\mathbb{B}_0^\top\Pi_0 + Q_0^\pi,\\
\rho\Pi_k &= \Pi_k\mathbb{A}_k + \mathbb{A}_k^\top\Pi_k - \Pi_k\mathbb{B}_kR^{-1}\mathbb{B}_k^\top\Pi_k + Q^\pi, \ \ \forall k,\\
\bar{A}_k &= [A_k - B_kR^{-1}B_k^\top\Pi_{k,11}]\mathbf{e}_k - B_kR^{-1}B_k^\top\Pi_{k,13}, \ \ \forall k,\\
\bar{G}_k &= -B_kR^{-1}B_k^\top\Pi_{k,12}, \ \ \forall k,\\
\rho s_0 &= \frac{ds_0}{dt} + (\mathbb{A}_0 - \mathbb{B}_0R_0^{-1}\mathbb{B}_0^\top\Pi_0)^\top s_0 + \Pi_0\mathbb{M}_0 - \bar{\eta}_0,\\
\rho s_k &= \frac{ds_k}{dt} + (\mathbb{A}_k - \mathbb{B}_kR^{-1}\mathbb{B}_k^\top\Pi_k)^\top s_k + \Pi_k\mathbb{M} - \bar{\eta}, \ \ \forall k,\\
\bar{m}_k &= -B_kR^{-1}\mathbb{B}_k^\top s_k, \ \ \forall k.
\end{aligned} \tag{32}$$

Finally one defines:

$$M_1 = \begin{bmatrix} A_1 - B_1R^{-1}B_1^\top\Pi_{1,11} & & 0\\ & \ddots & \\ 0 & & A_K - B_KR^{-1}B_K^\top\Pi_{K,11}\end{bmatrix},$$

$$M_2 = \begin{bmatrix} B_1R^{-1}B_1^\top\Pi_{1,13}\\ \vdots \\ B_KR^{-1}B_K^\top\Pi_{K,13}\end{bmatrix},$$

$$M_3 = \begin{bmatrix} A_0 & 0 & 0\\ \bar{G} & \bar{A} & 0\\ \bar{G} & -M_2 & M_1\end{bmatrix},$$

$$L_{0,H} = Q_0^{1/2}[I, 0, -H_0^\pi].$$

The final set of hypotheses is as follows:

A2: The initial states are independent, $\mathbb{E}x_i(0) = \bar{x}_i(0)$ for each $i \geq 0$, with $\sup_{i\geq 0}\mathbb{E}|x_i(0)|^2 \leq c$.

A3: The pair $(L_{0,H}, M_3)$ is observable.

A4: The pair $(L_a, \mathbb{A}_0 - (\rho/2)I)$ is detectable, and for each $k = 1, \ldots, K$, the pair $(L_b, \mathbb{A}_k - (\rho/2)I)$ is detectable, where $L_a = Q_0^{1/2}[I, -H_0^\pi]$ and $L_b = Q^{1/2}[I, -H, -\hat{H}^\pi]$. The pair $(\mathbb{A}_0 - (\rho/2)I, \mathbb{B}_0)$ is stabilizable and $(\mathbb{A}_k - (\rho/2)I, \mathbb{B}_k)$ is stabilizable for each $k = 1, \ldots, K$.

A5: There exists a unique stabilizing solution $\Pi_0, s_0, \Pi_k, \bar{A}_k, \bar{G}_k, s_k, \bar{m}_k$ to the major-minor MF equations (32) in the sense that the matrices

$$A_0 - \mathbb{B}_0R_0^{-1}\mathbb{B}_0^\top\Pi_0 - \frac{\rho}{2}I,$$
$$\mathbb{A}_k - \mathbb{B}_kR^{-1}\mathbb{B}_k^\top\Pi_k - \frac{\rho}{2}I, \quad 1 \leq k \leq K,$$

are asymptotically stable, and

$$\sup_{t\geq 0, 1\leq k\leq K} e^{-\frac{\rho}{2}t}\left(|s_0(t)| + |s_k(t)| + |\bar{m}_k(t)|\right) < \infty.$$

Theorem 3.2: (After Huang, 2010) Nash Equilibria for Major-Minor Agent MF Systems

Subject to **A1**–**A5** the MF equations generate a set of stochastic control laws $\mathcal{U}_{MF}^N \triangleq \{u_i^\circ; 0 \leq i \leq N\}$, $1 \leq N < \infty$, such that

1) all agent systems $0 \leq i \leq N$, are $e^{-\frac{\rho}{2}t}$ discounted second order stable in the sense that
$$\sup_{t\geq 0, 0\leq i\leq N}\mathbb{E}\Big\{e^{-\frac{\rho}{2}t}\big(\|x_i(t)\|^2 + \|x^N(t)\|^2 + \|\bar{x}(t)\|^2\big)\Big\} < C,$$
with C independent of N;

2) $\{\mathcal{U}_{MF}^N; 1 \leq N < \infty\}$ yields a unique ϵ-Nash equilibrium within the set of linear controls $\mathcal{U}_L^N$ for all ϵ, i.e. for all $\epsilon > 0$, there exists $N(\epsilon)$ such that for all $N \geq N(\epsilon)$
$$J_i^N(u_i^\circ, u_{-i}^\circ) - \epsilon \leq \inf_{u_i\in\mathcal{U}_L^N} J_i^N(u_i, u_{-i}^\circ) \leq J_i^N(u_i^\circ, u_{-i}^\circ),$$
where $J_i^N(u_i^\circ, u_{-i}^\circ) \to J_i^\infty(u_i^\circ, u_{-i}^\circ)$, $0 \leq i \leq N$, as $N \to \infty$. ■

IV. Partially Observed Major-Minor Agent LQG Systems

The partial observations structure which is adopted for the MM-LQG-MF systems in this paper are specified first for a generic minor agent with respect to its extended state for the fully observed system (29), and second, similarly for the major agent system (24). The partially observed (PO) MM-LQG-MF control problem is then to establish the existence of an MF Nash equilbrium for the resulting partially observed system.

The observation process for any minor agent $\mathcal{A}_i$, $1 \leq i \leq N$, of type k is defined to be

$$dy_i(t) = \mathbb{L}_k x_i^{0,\bar{x}}dt + dv_i(t) \equiv \mathbb{L}_k \begin{bmatrix} x_i\\ x_0\\ \bar{x}\end{bmatrix} dt + dv_i(t), \tag{33}$$

where

$$\mathbb{L}_k = [L_1^k \ L_2^k \ 0],$$

while for the major agent $\mathcal{A}_0$ the observation process is defined to be the process of complete observations given by

$$dy_0(t) = dx_0(t). \tag{34}$$

We now introduce the family of partial observation information sets $\mathcal{F}_i^y$, $1 \leq i \leq N$, where each $\mathcal{F}_i^y$ denotes the increasing family of σ-fields $\{\mathcal{F}_{i,t}^y;\ 0 \leq t < \infty\}$ generated by agent $\mathcal{A}_i$'s partial observations $(y_i(\tau);\ 0 \leq \tau \leq t)$, on its own state and the major agent's state, as given in (33). Further, for the major agent $\mathcal{A}_0$, the information set is defined to be the family of increasing σ-fields $\{\mathcal{F}_{0,t};\ 0 \leq t\}$ generated by the process of complete observations (34). An immediate consequence of the definition above is that the major agent's filter equations for its state are simply the original major agent's dynamical equations (8a), which, when extended with the system's mean field, take the form (25).

A6: *Minor Agent σ-Fields and Linear Controls:* For each minor agent $\mathcal{A}_i$, $1 \leq i \leq N$, the set of control inputs $\mathcal{U}_{i,y}^{N,L}$ is defined to be the collection of linear feedback controls adapted to the increasing σ-fields of partial observations $\{\mathcal{F}_{i,t}^y;\ 0 \leq t\}$.

A7: *Major Agent σ-Fields and Linear Controls:* For the major agent $\mathcal{A}_0$ the set of control inputs $\mathcal{U}_0^L$ is defined to be the collection of linear feedback controls adapted to the increasing σ-fields of complete observations $\{\mathcal{F}_{0,t};\, 0 \leq t\}$.

By **A7** the minor agents are able to generate u_0° whenever the functional dependence of the major agent's control on its state is available to the minor players.

The motivation for adopting the complete observation hypothesis in **A7** is as follows. Consider the situation where the major agent's controls are adapted to an increasing family of σ-fields $\tilde{\mathcal{F}}_0$ generated by partial observations y_0 on the major agent $\mathcal{A}_0$'s state, where it is assumed these observations are only available to the major agent. The conditional expectation of the major agent $\mathcal{A}_0$'s control input, which is to be computed by an agent $\mathcal{A}_i$, would then be given by the conditional expectation with respect to $\mathcal{F}_i^y$ of the conditional expectation $\hat{x}_{0|\tilde{F}_0}$ (necessarily computed with respect to $\tilde{\mathcal{F}}_0$). Specifically, for linear control inputs the estimation process for agent $\mathcal{A}_i$ would involve the term

$$\begin{aligned}\mathbb{E}_{|\mathcal{F}_{i,t}^y} u_0(t) &= -\mathbb{E}_{|\mathcal{F}_{i,t}^y}(K_1 \mathbb{E}_{|\mathcal{F}_{0,t}} x_0^{\bar{x}}(t) + K_2) \\ &= -K_1 \mathbb{E}_{|\mathcal{F}_{i,t}^y}(\mathbb{E}_{|\mathcal{F}_{0,t}} x_0^{\bar{x}}(t)) - K_2, \quad (35)\end{aligned}$$

where the time subscript on the observation sigma algebras is displayed for explicitness.

The complete observations hypothesis **A7** rules out the recursive generation of iterated conditional expectations as found in (35). Consequently the analysis of such iterated estimates falls outside the scope of this paper, but it is the subject of current investigations (see e.g. [13]).

The standard assumption below implies the convergence of the solution to the filtering Riccati equation to a positive definite asymptotically stabilizing solution.

A8: The system parameter set $\Theta = \{1, \ldots, K\}$ is such that $[\mathbb{A}_k, Q_w^k]$ is controllable and $[\mathbb{L}_k, \mathbb{A}_k]$ is observable for all k, $1 \leq k \leq K$.

We begin the analysis of the recursive state estimation for the system's major and minor agents by noting that no non-trivial state estimation problem occurs for the major agent state in (25) since the x_0 component of the state $x_0^{\bar{x}}$ is completely observed by **A7** and the mean field component $\bar{x}$ is a function of x_0 and the deterministic process $\bar{m}$.

Concerning the state estimation problem for the minor agents, for any L^2 random variable z at any instant $t \geq 0$, let $\hat{z}_{|\mathcal{F}_i^y}$ denote the conditional expectation $\mathbb{E}_{|\mathcal{F}_{i,t}^y}$ of z with respect to the observation σ-field $\mathcal{F}_{i,t}^y$ of the agent $\mathcal{A}_i$ at the instant $t \geq 0$.

It is clear that the Riccati equation associated with the Kalman filtering equations for the extended state $x_i^{0,\bar{x}}$ of an agent type k is given by:

$$\dot{V}(t) = \mathbb{A}_k V(t) + V(t)\mathbb{A}_k^\top - K(t) R_v K^\top(t) + Q_w^k, \quad (36)$$

where we recall that

$$Q_w^k = \begin{bmatrix} \Sigma_k & 0 & 0 \\ 0 & \Sigma_0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

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

and

$$V(0) = \mathbb{E}\left[x_i^{0,\bar{x}}(0) - (\widehat{x_i^{0,\bar{x}}(0)})_{|\mathcal{F}_i^y}\right]\left[x_i^{0,\bar{x}}(0) - (\widehat{x_i^{0,\bar{x}}(0)})_{|\mathcal{F}_i^y}\right]^\top.$$

The innovations process for the estimation of $x_i^{0,\bar{x}}$ is evidently

$$d\nu_i = dy_i - \mathbb{L}_k \begin{bmatrix} \hat{x}_{i|\mathcal{F}_i^y} \\ \hat{x}_{0|\mathcal{F}_i^y} \\ \hat{\bar{x}}_{|\mathcal{F}_i^y} \end{bmatrix} dt, \quad (37)$$

and the Kalman filter gain is given by

$$K(t) = V(t)\mathbb{L}_k^\top R_v^{-1}.$$

Then the filter equations for a generic minor agent $\mathcal{A}_i$ with state equation (30) and observation equation (33) are given by the following equations

$$\begin{aligned}\begin{bmatrix} d\hat{x}_{i|\mathcal{F}_i^y} \\ d\hat{x}_{0|\mathcal{F}_i^y} \\ d\hat{\bar{x}}_{|\mathcal{F}_i^y} \end{bmatrix} &= \begin{bmatrix} A_k & [G \;\; 0_{n\times nK}] \\ \begin{bmatrix} 0_{n\times n} \\ 0_{nK\times n} \end{bmatrix} & \begin{bmatrix} A_0 & 0_{nK\times n} \\ \bar{G} & \bar{A} \end{bmatrix} \end{bmatrix} \\ &\times \begin{bmatrix} \hat{x}_{i|\mathcal{F}_i^y} \\ \hat{x}_{0|\mathcal{F}_i^y} \\ \hat{\bar{x}}_{|\mathcal{F}_i^y} \end{bmatrix} dt + \begin{bmatrix} B_k \\ 0_{n\times m} \\ 0_{nK\times m} \end{bmatrix} u_i dt \\ &+ \begin{bmatrix} 0_{n\times m} \\ B_0 \\ 0_{nK\times m} \end{bmatrix} \hat{u}_{0|\mathcal{F}_i^y} dt + \begin{bmatrix} 0_{n\times 1} \\ 0_{n\times 1} \\ \bar{m} \end{bmatrix} dt + K(t) d\nu_i(t), \quad (38)\end{aligned}$$

which may be written

$$\begin{aligned}\begin{bmatrix} d\hat{x}_{i|\mathcal{F}_i^y} \\ d\hat{x}_{0|\mathcal{F}_i^y} \\ d\hat{\bar{x}}_{|\mathcal{F}_i^y} \end{bmatrix} &= \begin{bmatrix} A_k & [G \;\; 0_{n\times nK}] \\ 0_{(nK+n)\times n} & \mathbb{A}_0 - \mathbb{B}_0 R_0^{-1} \mathbb{B}_0^\top \Pi_0 \end{bmatrix} \\ &\times \begin{bmatrix} \hat{x}_{i|\mathcal{F}_i^y} \\ \hat{x}_{0|\mathcal{F}_i^y} \\ \hat{\bar{x}}_{|\mathcal{F}_i^y} \end{bmatrix} dt + \begin{bmatrix} B_k \\ 0_{(nK+n)\times m} \end{bmatrix} u_i dt \\ &+ \begin{bmatrix} 0_{n\times 1} \\ \mathbb{M}_0 - \mathbb{B}_0 R_0^{-1} \mathbb{B}_0^\top s_0 \end{bmatrix} dt + K(t) d\nu_i(t), \quad (39)\end{aligned}$$

where, by **A7**,

$$\hat{u}_{0|\mathcal{F}_i^y} = \mathbb{E}_{|\mathcal{F}_{i,t}^y} u_0(t) = -K_1 \mathbb{E}_{|\mathcal{F}_{i,t}^y} x_0^{\bar{x}}(t) - K_2, \quad (40)$$

with the initial conditions

$$\begin{bmatrix} \hat{x}_{i|\mathcal{F}_i^y}(0) \\ \hat{x}_{0|\mathcal{F}_i^y}(0) \\ \hat{\bar{x}}_{|\mathcal{F}_i^y}(0) \end{bmatrix} = \mathbb{E}\begin{bmatrix} x_i(0) \\ x_0(0) \\ \hat{\bar{x}}(0) \end{bmatrix}.$$

Theorem 4.1: ϵ-Nash Equilibria for PO MM-MFG Systems Subject to **A1**–**A8**, the KF-MFG state estimation scheme (36)–(40) plus the MM-MFG equation scheme (32) generate the set of control laws $\hat{\mathcal{U}}_{MF}^N \triangleq \{\hat{u}_i^\circ; 0 \leq i \leq N\}$, $1 \leq N < \infty$, given by

$$u_0^\circ = -R_0^{-1}\mathbb{B}_0^\top \left[\Pi_0 x_0^{\bar{x}} + s_0\right],$$

$$\hat{u}_i^\circ = -R^{-1}\mathbb{B}_k^\top \left[\Pi_k (\hat{x}_{i|\mathcal{F}_i^y}^\top, \hat{x}_{0|\mathcal{F}_i^y}^\top, \hat{\bar{x}}_{|\mathcal{F}_i^y}^\top)^\top + s_k\right],$$

$\{1 \leq i \leq N\}$, such that

1) all agent systems $0 \le i \le N$, are $e^{-\frac{\rho}{2}t}$ discounted second order stable in the sense that

$$\sup_{t\ge 0, 0\le i\le N} e^{-\frac{\rho}{2}t}\mathbb{E}\left(\|\hat{x}_{i|\mathcal{F}_i^y}(t)\|^2 + \|\hat{\bar{x}}_{|\mathcal{F}_i^y}(t)\|^2\right) < C,$$

with C independent of N;

2) $\{\hat{\mathcal{U}}_{MF}^N; 0 \le i \le N < \infty\}$ yields a unique ϵ-Nash equilibrium within the class of linear controls $\mathcal{U}_{i,y}^{N,L}$ and $\mathcal{U}_0^L$ for all ϵ, i.e. for all $\epsilon > 0$, there exists $N(\epsilon)$ such that for all $N \ge N(\epsilon)$;

$$J_i^{s,N}(\hat{u}_i^\circ, \hat{u}_{-i}^\circ) - \epsilon \le \inf_{u_i \in \mathcal{U}_{i,y}^{N,L}} J_i^{s,N}(u_i, \hat{u}_{-i}^\circ) \le J_i^{s,N}(\hat{u}_i^\circ, \hat{u}_{-i}^\circ),$$

where the major agent's performance function $J_0^{s,N}(u_0, u_{-0})$, $u_i \in \mathcal{U}_{i,y}^{N,L}$, $0 \le i \le N$, takes the value $J_0^N(u_0, u_{-0})$ as in the completely observed case, and the generic minor agent type k's performance function $J_i^{s,N}(u_i, u_{-i})$, $u_i \in \mathcal{U}_{i,y}^{N,L}$, $0 \le i \le N$, is given by

$$J_i^N(u_i, u_{-i}) + \hat{E}_N,$$

where $J_i^N(u_i, u_{-i})$ is as in the completely observed case, $\hat{E}_N > 0$, and when $u_i = \hat{u}_i^\circ$ the following limits hold:

$$(\alpha)\ \lim_{N\to\infty} J_i^N(\hat{u}_i^\circ, \hat{u}_{-i}^\circ) = J_i^\infty(\hat{u}_i^\circ, \hat{u}_{-i}^\circ),$$

$$(\beta)\ \lim_{N\to\infty} \hat{E}_N = \int_0^\infty e^{-\rho t}\mathrm{tr}[Q^\pi V_k(t)]dt,$$

where V_k is the solution to (36).

Proof: Generalizing the standard procedure (see e.g. [14] and [15]), an initial decomposition shall be made of (i) the state processes and (ii) the performance functions of the major agent and the minor agents; this transformation takes them into their separated forms which depend explicitly on the estimation processes and the control independent estimation errors. In order to employ the NCE analysis [3], [7], this is done for the finite and infinite population cases for both the major agent and the minor agents.

State Decompositions

Finite Population Major Agent:

$$\begin{bmatrix} x_0 \\ x^N \end{bmatrix} = \begin{bmatrix} x_0 \\ \hat{x}_{|\mathcal{F}_0}^N \end{bmatrix} + \begin{bmatrix} 0 \\ x^N - \hat{x}_{|\mathcal{F}_0}^N \end{bmatrix}, \quad N \in \mathbb{Z}_1, \qquad (41)$$

where the singularity is due to the hypothesis of the theorem that the major agent observes its current state.

Infinite Population Major Agent:

$$\begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} = \begin{bmatrix} x_0 \\ \hat{\bar{x}}_{|\mathcal{F}_0} \end{bmatrix} + \begin{bmatrix} 0 \\ \bar{x} - \hat{\bar{x}}_{|\mathcal{F}_0} \end{bmatrix}, \qquad (42)$$

with $\bar{x} - \hat{\bar{x}}_{|\mathcal{F}_0} = 0$ as above.

Finite Population Minor Agent:

$$\begin{bmatrix} x_i \\ x_0 \\ x^N \end{bmatrix} = \begin{bmatrix} \hat{x}_{i|\mathcal{F}_i^y} \\ \hat{x}_{0|\mathcal{F}_i^y} \\ \hat{x}_{|\mathcal{F}_i^y}^N \end{bmatrix} + \begin{bmatrix} x_i - \hat{x}_{i|\mathcal{F}_i^y} \\ x_0 - \hat{x}_{0|\mathcal{F}_i^y} \\ x^N - \hat{x}_{|\mathcal{F}_i^y}^N \end{bmatrix}, \quad N \in \mathbb{Z}_1. \qquad (43)$$

Infinite Population Minor Agent:

$$\begin{bmatrix} x_i \\ x_0 \\ \bar{x} \end{bmatrix} = \begin{bmatrix} \hat{x}_{i|\mathcal{F}_i^y} \\ \hat{x}_{0|\mathcal{F}_i^y} \\ \hat{\bar{x}}_{|\mathcal{F}_i^y} \end{bmatrix} + \begin{bmatrix} x_i - \hat{x}_{i|\mathcal{F}_i^y} \\ x_0 - \hat{x}_{0|\mathcal{F}_i^y} \\ \bar{x} - \hat{\bar{x}}_{|\mathcal{F}_i^y} \end{bmatrix}. \qquad (44)$$

Performance Function Separation

We now apply the smoothing property of conditional expectations with respect to the increaing filtration families $\mathcal{F}_i^y$ and $\mathcal{F}_0^y$ to the major and minor agent performance functions (9) and (10) respectively to obtain the expressions (45), (46), (47), (48) below. The superscript 's' on the resulting performance functions indicates the separation into control dependent and control independent summands.

The crucial feature of the resulting decompositions is that the second terms in the separated performance functions are independent of the controls and the first terms depend upon the state estimation processes where the dynamics in each case are given by the filter equations (36)–(40).

Major Agent Finite Population J_0^N:

$$J_0^{s,N}(u_0, u_{-0}) = \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left\{\left\|x_0 - H_0^\pi \hat{x}_{|\mathcal{F}_0}^N - \eta_0\right\|_{Q_0}^2 + \|u_0\|_{R_0}^2\right\}dt\right] + \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left\|H_0^\pi(x^N - \hat{x}_{|\mathcal{F}_0}^N)\right\|_{Q_0}^2 dt\right]. \qquad (45)$$

Major Agent Infinite Population J_0^∞:

$$\begin{aligned} J_0^{s,\infty}(u_0, u_{-0}) &= \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left\{\left\|x_0 - H_0^\pi \hat{\bar{x}}_{|\mathcal{F}_0} - \eta_0\right\|_{Q_0}^2 + \|u_0\|_{R_0}^2\right\}dt\right] + \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left\|H_0^\pi(\bar{x} - \hat{\bar{x}}_{|\mathcal{F}_0})\right\|_{Q_0}^2 dt\right] \\ &= \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left\{\left\|x_0 - H_0\bar{x} - \eta_0\right\|_{Q_0}^2 + \|u_0\|_{R_0}^2\right\}dt\right], \end{aligned} \qquad (46)$$

since $\hat{\bar{x}}_{\mathcal{F}_0} = \bar{x}$ by (25).

Minor Agent Finite Population J_i^N:

$$\begin{aligned} J_i^{s,N}(u_i, u_{-i}) &= \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left\{\left\|\hat{x}_{i|\mathcal{F}_i^y} - H_1\hat{x}_{0|\mathcal{F}_i^y} - H_2^\pi \hat{x}_{|\mathcal{F}_i^y}^N - \eta\right\|_Q^2 + \|u_i\|_R^2\right\}dt\right] \\ &+ \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left\|\tilde{x}_{i|\mathcal{F}_i^y} - H_1\tilde{x}_{0|\mathcal{F}_i^y} - H_2^\pi \tilde{x}_{|\mathcal{F}_i^y}^N\right\|_Q^2 dt\right], \end{aligned} \qquad (47)$$

where $\tilde{x}_{i|\cdot}$ denotes the error $x_i - \hat{x}_{i|\cdot}$, and similarly for the other terms.

Minor Agent Infinite Population J_i^∞:

$$
\begin{aligned}
J_i^{s,\infty}(u_i,u_{-i}) = \mathbb{E}\bigg[\int_0^\infty e^{-\rho t}\Big\{\Big\|\hat{x}_{i|\mathcal{F}_i^y} - H_1\hat{x}_{0|\mathcal{F}_i^y} \\
- H_2^\pi \hat{\bar{x}}_{|\mathcal{F}_i^y} - \eta\Big\|_Q^2 + \|u_i\|_R^2\Big\}dt\bigg] \\
+\mathbb{E}\bigg[\int_0^\infty e^{-\rho t}\big\|(x_i - \hat{x}_{i|\mathcal{F}_i^y}) - H_1(x_0 - \hat{x}_{0|\mathcal{F}_i^y}) \\
- H_2^\pi(\bar{x} - \hat{\bar{x}}_{|\mathcal{F}_i^y})\big\|_Q^2 dt\bigg]. \qquad (48)
\end{aligned}
$$

We now solve the tracking LQG-MM problem in the infinite population case for the major and minor agents in their separated forms.

As shown in equation (47), in the case of the major agent the infinite population mean field tracking problem is that of a single agent LQG tracking problem with the function $H_0\bar{x}$ to be tracked generated by (25). While for a generic minor agent with extended estimated state process generated by (30) the infinite population tracking loss function in separated form is given by

$$
\begin{aligned}
J_i^{s,\infty}(u_i,u_{-i}) = \mathbb{E}\int_0^\infty e^{-\rho t}\{\|[I,\ -H_1,\ -H_2^\pi]\begin{bmatrix}\hat{x}_{i|\mathcal{F}_i^y}\\ \hat{x}_{0|\mathcal{F}_i^y}\\ \hat{\bar{x}}_{|\mathcal{F}_i^y}\end{bmatrix} \\
-\eta\|_Q^2 + \|u_i\|_R^2\}dt \\
+\mathbb{E}\int_0^\infty e^{-\rho t}\{\|[I,\ -H_1,\ -H_2^\pi]\begin{bmatrix}\tilde{x}_{i|\mathcal{F}_i^y}\\ \tilde{x}_{0|\mathcal{F}_i^y}\\ \tilde{\bar{x}}_{|\mathcal{F}_i^y}\end{bmatrix}\|_Q^2\}dt. \qquad (49)
\end{aligned}
$$

Now, (i) the controlled state processes appearing in (46) and (49) are respectively those of the completely observed major agent system (25) and the completely observed minor agent state estimation system (38), and (ii) the estimation error process $(\tilde{x})$ in (49) is independent of u_i. Hence by Theorem 3.2 the infinite population Nash Certainty Equivalence equilibrium controls are given by u_0° and $\hat{u}_i^\circ$ in the theorem statement.

Parallel to Theorem 3.2, this results in (i) $e^{-\frac{\rho}{2}t}$ second order system stability and (ii) the ϵ-Nash equilibrium property; but in the present case this is obtained first for the completely observed estimated systems with their performance functions J_i^N, J_i^∞, $0 \le i \le N < \infty$, with the final result for the minor agents obtained for the minor agent by adding the finite control independent loss, $\hat{E}_N$, corresponding to the second part of the loss functions in (47) and (48); these depend linearly on the solutions to the associated Riccati equations. In the major agent case the additional performance loss term evidently takes the value zero.

The form of the epsilon-Nash result in the theorem statement is finally obtained by adding the corresponding state estimation error loss terms. ∎

V. Simulation

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

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

and for the major agent we have

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

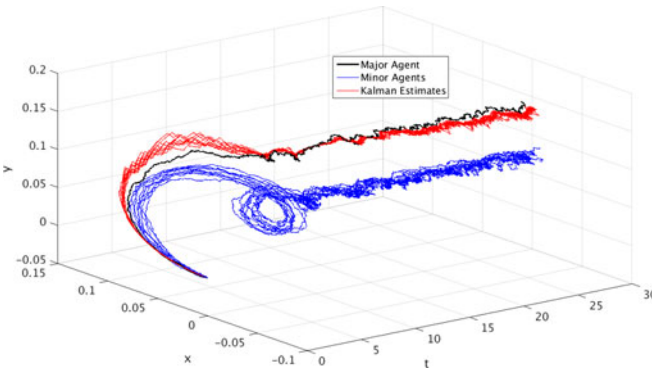

Fig. 1. State trajectories.

The parameters used in the simulation are: $t_{final} = 30s$, $\Delta t = 0.01s$, $\sigma_w = 0.005$, $\sigma_v = 0.0001$, $\rho = 0.01$, $\eta = [0.25, 0.25]^\top$, $\eta_0 = [0.25, 0.25]^\top$, $Q = 1 \times I_{2\times 2}$, $Q_0 = 1 \times I_{2\times 2}$, $R = 1$, $R_0 = 1$, $H = 0.6 \times I_{2\times 2}$, $H_0 = 0.6 \times I_{2\times 2}$, $\hat{H} = 0.6 \times I_{2\times 2}$, $G = 0_{2\times 2}$, and the mean field equation system is iterated 100 times. The state trajectories for a single realization can be displayed for the entire population of 101 agents together with their major agent state estimates, but in the figure only 10 minor agents are shown for the sake of clarity. The effect of the major agent's state on minor agents' states may be seen at the time horizon. In the case where the minor agents are not allowed to directly observe the major agent they apply Kalman filtering, and for this case their estimates of the major agent are also plotted in the figure. As expected, the estimates closely follow the true state values of the major over the entire time interval.

VI. Concluding Remarks

Based upon the MM-LQG-MFG theory for partially observed systems developed in this paper and in [13], and the nonlinear generalization of MM-MFG theory presented in [12], current work includes (i) the generation of a theory for nonlinear partially observed MM-MFG systems and initial results are announced in [16], and (ii) the generalization of the theory in the current paper to the case where the major agent has only partial observations on its own state, which, as stated earlier, has been initiated in [13]. Furthermore, building upon the work in [17], the application of partially observed MM-MFG methods to electrical power markets has been initiated in [10]; in that formulation, minor agents (customers and suppliers) receive intermittent observations on active major agents (such as utilities and international energy prices) and on passive major agents (such as wind and ocean behaviour).

Finally, in the non-classical information pattern framework of this paper, we have restricted attention to the class of controls which are linear in the system observation processes. In this connection, we observe that for the problem addressed in this paper sufficient conditions for unique MFG equilibria to exist, and for the corresponding control laws to be linear, within, say, a Lipschitz class, are currently unknown.

Acknowledgment

The authors gratefully acknowledge conversations with M. Huang, D. Firoozi, and S. Yüksel, and the comments of the anonymous reviewers.

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," *Proc. 42nd IEEE Conf. Decision Control*, Maui, HI, Dec. 2003, pp. 98–103.

[2] 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.

[3] M. Huang, P. E. Caines, and R. P. Malhamé, "Large population cost-coupled LQG problems with non-uniform agents: individual-mass behaviour and decentralized ϵ-Nash equilibria," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1560–1571, Sept. 2007.

[4] P. E. Caines, "Mean field games," in *Encyclopedia of Systems and Control*, T. Samad and J. Baillieul, Eds. Springer London, 2014, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1007/978-1-4471-5102-9_30-1.

[5] 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.

[6] 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.

[7] 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.

[8] S. L. Nguyen and M. Huang, "Linear-quadratic-gaussian mixed games with continuum-parametrized minor players," *SIAM J. Control Optim.*, vol. 50, no. 5, pp. 2907–2937, 2012.

[9] M. Huang, P. E. Caines, and R. P. Malhamé, "Distributed multi-agent decision-making with partial observations: Asymptotic Nash equilibria," *Proc. 17th Int. Symp. Math. Theory Netw. Syst.*, Kyoto, Japan, Jul. 2006, pp. 2725–2730.

[10] A. C. Kizilkale and P. E. Caines, "Recursive estimation of common partially observed disturbances in mfg systems with application to large scale power markets," in *Proc. 52nd IEEE Conf. Decision Control*, Dec. 2013, pp. 2505–2512.

[11] P. E. Caines and A. C. Kizilkale, "Mean field estimation for partially observed LQG systems with major and minor agents," *Proc. 19th World Congress Int. Fed. Autom. Control*, Aug. 2014, pp. 8705–8709.

[12] M. Nourian and P. E. Caines, "ϵ-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents," *SIAM J. Control Optim.*, vol. 51, no. 4, pp. 3302–3331, 2013.

[13] D. Firoozi and P. E. Caines, "ϵ-Nash equilibria for partially observed LQG mean field games with major player: Partial observation of all agents," *Proc. 54th IEEE Conf. Decision Control*, Dec. 2015, pp. 4430–4437.

[14] M. H. A. Davis, *Linear Estimation and Stochastic Control*. London, U.K., 1977.

[15] P. E. Caines, *Linear Stochastic Systems*, New York: Wiley, Apr. 1988.

[16] N. Sen and P. E. Caines, "Mean field game theory with a partially observed major agent", *SIAM J. Control Optim.*, vol. 54, no. 6, pp. 3174–3224, 2016. DOI: 10.1137/16M1063010.

[17] A. C. Kizilkale, S. Mannor, and P. E. Caines, "Large scale real-time bidding in the smart grid: A mean field framework," in *Proc. 51st IEEE Conf. Decision Control*, Dec. 2012, pp. 3680–3687.

Peter E. Caines (LF'11) received the B.A. degree in mathematics from Oxford University, Oxford, U.K., in 1967 and the Ph.D. in systems and control theory from Imperial College, University of London, London, U.K., in 1970.

He has been a Postdoctoral Researcher and Faculty Member at the University of Manchester Institute of Science and Technology, Manchester, U.K., Stanford University, Stanford, CA, University of California, Berkeley, University of Toronto, Toronto, ON Canada, and Harvard University, Cambridge, MA. 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 received the IEEE Control Systems Society Bode Lecture Prize in 2009. He is a Fellow of SIAM, IMA (UK), 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 Trans. Autom. Control, 1980, was recognized by the IEEE Control Systems Society as one of the 25 seminal control theory papers of the 20th century.

Arman C. Kizilkale (M'14) received the B.Sc. degree in telecommunications engineering from Sabanci University, Istanbul, Turkey, in 2004, the M.Sc. degree in electrical engineering from the University of Michigan, Ann Arbor, MI, in 2005, and the Ph.D. degree in electrical engineering from McGill University, Montreal, QC in 2013.

Until 2016, he was a Research Associate at École Polytechnique de Montréal and GERAD, Montreal. Since then, he has been a Research Scientist at Automat, Montreal.