This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

Collective Stochastic Discrete Choice Problems: A Min-LQG Dynamic Game Formulation

Rabih Salhab, Roland P. Malhamé and Jerome Le Ny

***Abstract*—We consider a class of dynamic collective choice models with social interactions, whereby a large number of non-uniform agents have to individually settle on one of multiple discrete alternative choices, with the relevance of their would-be choices continuously impacted by noise and the unfolding group behavior. This class of problems is modeled here as a so-called Min-LQG game, i.e., a linear quadratic Gaussian dynamic and non-cooperative game, with an additional combinatorial aspect in that it includes a final choice-related minimization in its terminal cost. The presence of this minimization term is *key to enforcing some specific discrete choice* by each individual agent. The theory of mean field games is invoked to generate a class of decentralized agent feedback control strategies, which are then shown to converge to an exact Nash equilibrium of the game as the number of players increases to infinity. A key building block in our approach is an explicit solution to the problem of computing the best response of a generic agent to some arbitrarily posited smooth mean field trajectory. Ultimately, an agent is shown to face a continuously revised discrete choice problem, where greedy choices dictated by current conditions must be constantly balanced against the risk of the future process noise upsetting the wisdom of such decisions. We show that any Nash equilibrium of the game is defined by an a priori computable probability matrix which describes the distribution of the players' choices over the alternatives. The results are illustrated through simulations.**

***Index Terms*—Mean Field Games, Stochastic Optimal Control, Discrete Choice Models.**

I. Introduction

Discrete choice problems arise in situations where an individual makes a choice among a discrete set of alternatives, such as modes of transportation [1], entry and withdrawal from the labor market, or residential locations [2]. In some situations, the individuals' choices are considerably influenced by the surrounding social behavior. For example, in schools, teenagers' decisions to smoke are affected by some personal factors, as well as by their peers' behavior [3].

In this paper, we study a discrete choice problem for a large population of agents in a dynamic setting, capturing in particular the impact of mean population behavior on individuals, the efforts involved by the latter in changing opinions, and possible dithering behavior until an ultimate choice is crystallized (See Figure 1 below, which illustrates the opinions' evolution of a group of agents choosing between two alternatives, -10 and 10). For example, as in [4]–[6], a group of robots exploring an unknown terrain might need to move within a finite time horizon from their initial positions towards one of multiple sites of interest. While moving, each robot optimizes its efforts to remain close to the group and to arrive at the end of the time horizon in the vicinity of one of the predefined sites. The group may split, but the size of the subgroups should remain large enough to perform some collective tasks, such as search and rescue. Our model could also be viewed as a mechanistic representation of opinion crystallization in elections [7], [8], where voters continuously update their opinions until forming a final decision regarding who they should vote for. Along the path to choose a candidate, changing one's decision requires an effort, while at the same time, persistent deviation from the current majority opinion involves a discomfort.

This work was supported by NSERC under Grants 6820-2011 and 435905-13. The authors are with the department of Electrical Engineering, Polytechnique Montreal and with GERAD, Montreal, QC H3T-1J4, Canada {rabih.salhab, roland.malhame, jerome.le-ny}@polymtl.ca.

This paper makes the following contributions:

1) We introduce a new class of stochastic dynamic games called "min-LQG" to model certain dynamic discrete choice problems involving population-influenced decisions.
2) We develop via the Mean Field Games (MFGs) methodology, briefly reviewed below, a set of decentralized closed-loop strategies which lead to a near Nash equilibrium for the finite population min-LQG game. The latter becomes exact asymptotically in the size of the population. To compute the associated feedback strategies, a generic agent must solve a novel class of optimal tracking problems for which we are able to obtain an explicit solution. Furthermore, provided the limiting MFG equilibrium exists, it is important to note that a generic agent need only know the other agents' initial states distribution, and observe its own state, to compute the Nash equilibrium inducing feedback policy. This means that the communication needed between agents, beyond what would be required in a consensus like algorithm for initial states distribution exchange (see [9] for example) is minimal.
3) According to the min-LQG control laws, the agents cannot commit to a choice from the beginning. Instead, they continuously revise their decisions to account for the risk of being driven by the noise process to one of the alternatives, while making a premature decision in favor of another alternative based on the current conditions. Accordingly, we show that one can interpret the solution to the min-LQG tracking problem at each instant as that of a modified static discrete choice model [10], where an agent's cost for choosing one of the alternatives includes an additional term penalizing myopic decisions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

4) We establish a one-to-one correspondence between the infinite population Nash equilibria and the fixed points of a finite dimensional map. These fixed points are the potential probability distributions of the agents' choices over the alternatives. This one-to-one correspondence reduces the infinite dimensional problem of finding a Nash equilibrium for the limiting game to finding a fixed point for the finite dimensional map. Furthermore, it simplifies the existence proofs and numerical schemes to compute the equilibria.
5) Our model lies somewhere between the LQG MFG models studied in [11], [12] and the fully nonlinear MFGs considered in [13]–[16]. On the one hand, in the LQG case the optimal strategies are linear feedback policies, whereas in our case they are nonlinear. On the other hand, an explicit solution in the general nonlinear case is not possible, while we provide in this paper an explicit form for the optimal strategies. Moreover, the existence and uniqueness of solutions of the mean field equations are shown in the general case under some strong regularity and boundedness conditions on the dynamics and cost coefficients [13]–[16]. In this paper, however, the quadratic running cost and linear dynamics make it possible to relax these assumptions.

The rest of the paper is organized as follows. Section II discusses related work on discrete choice models and MFGs. In Section III, we present the mathematical formulation of our problem. In Section IV, we develop a closed form solution to the min-LQG optimal tracking problem characterizing a generic agent's best response. In Section V we introduce the mean field equations and analyze their solutions. Finally, Section VI discusses some numerical simulations illustrating the results, while Section VII presents our conclusions.

II. Related Work and Mean Field Games

McFadden laid in [10] the theoretical foundations of static discrete choice models, where an agent has to choose among a finite set of alternatives the one that maximizes its utility. This utility depends on some potentially observable attributes that dictate a deterministic trend in the choice, and other attributes idiosyncratic to that agent, which are not known by the economist or social planner carrying out the macroscopic analysis, although they may influence the choice. As a result, utility is defined as the sum of a known deterministic term plus a random term. Later, Rust [17] introduced a dynamic discrete choice model involving a Markov decision process for each agent. While peer pressure effects are absent in Rust's and McFadden's models, Brock and Durlauf [18] discuss a discrete choice problem with *social interactions* modeled as a *static* non-cooperative game, where a large number of agents have to choose between two alternatives while being influenced by the average of the choices. The authors analyze the model using an approach similar to that of a static MFG and inspired by statistical mechanics.

The Brock-Durlauf model includes peer influence but is static, and in Rust's model, the agents are required to make a choice repeatedly at each discrete time period but under no social influence. In our model, the agents are instead continuously reassessing the adequacy of their would-be choices and current actions along their random state-space path, up until the end of the control horizon, at which point their ultimate choice of alternative becomes fully crystallized. Thus, our formulation helps in modeling situations where alternative choices are identified as potential destination points in a suitable state space (e.g., physical space in the robotics example, or opinion space in the election example), and implementation of a given choice involves movement towards a final destination state, requiring control effort and constrained by specific dynamics.

The MFG methodology that we follow in this paper was introduced in a series of papers by Huang *et al.* [11]–[13], and independently by Lions and Lasry [14]–[16]. It is a powerful technique to analyze dynamic games involving a large number of agents anonymously interacting through their *mean field*, i.e., the empirical distribution of the agents' states. The analysis starts by considering the limiting case of a continuum of agents. For agents evolving according to diffusion processes, the equilibrium of the game can be shown to be characterized by the solution of two coupled partial differential equations (PDE), a Hamilton-Jacobi-Bellman (HJB) equation propagating backwards and a Fokker-Planck (FP) equation propagating forwards. Indeed, the former characterizes the agents' best response to some posited candidate mean field trajectory, while the latter propagates the would-be resulting mean field when all agents implement the computed best responses. Consistency requires that sustainable mean field trajectories, if they exist, be replicated in the process. Limiting equilibria are thus required to satisfy a system of fixed point equations, herein given by the coupled HJB-FP equations. The corresponding best responses, when applied to the finite population, constitute an approximate Nash equilibrium (ϵ-Nash equilibrium) [12], [13], in the following sense.

Definition 1: Consider N agents, a set of strategy profiles $\mathcal{U} = \mathcal{U}_1 \times \ldots \times \mathcal{U}_N$, and for each agent $i \in \{1, \ldots, N\}$, a cost function $J_i(u_1, \ldots, u_N)$, for $(u_1, \ldots, u_N) \in \mathcal{U}$. A strategy profile $(u_1^*, \ldots, u_N^*) \in \mathcal{U}$ is called an ϵ-Nash equilibrium with respect to the costs $\{J_i, 1 \leq i \leq N\}$ if there exists an $\epsilon > 0$ such that, for any fixed $1 \leq i \leq N$, for all $u_i \in \mathcal{U}_i$, we have $J_i(u_i, u_{-i}^*) \geq J_i(u_i^*, u_{-i}^*) - \epsilon$.

Recently, we introduced a related dynamic collective choice model with social interactions in [19]–[21]. In these articles, we study using the MFG methodology a dynamic game involving a large number of players choosing between multiple destination points, when the agents' dynamics are *deterministic* with random initial conditions. We show that multiple $\epsilon-$Nash equilibria may exist. The strategies developed in these papers are *open loop* decentralized policies, in the sense that to make its choice of trajectory and destination, an agent needs to know only its initial state and the initial distribution of the population. In particular, in this formulation, each agent can commit from the start to its final choice, which implies that the model with deterministic dynamics is not sufficiently rich to fully capture opinion change phenomena. In contrast, we consider here the fully stochastic case, where the noise in the dynamics can model for example the unexpected events that influence a voter's opinion during electoral campaigns or the random forces that perturb a robot's trajectory while choosing a site to visit. The presence of this noise requires the agents to use bona fide feedback strategies and prevents them from committing to a choice before the final time. However, one can still anticipate asymptotically the manner in which an infinite population splits among the set of alternatives.

III. Mathematical Model

A. Notation

Given an Euclidean space X, $B(c,r)$ denotes the ball of radius r centered at c, x_k the k-th element of $x \in X$, and x_{-i} the vector $(x_1,\dots,x_{i-1},x_{i+1},\dots,x_N)$. The Kronecker product is denoted $\otimes$. The $n\times n$ identity matrix is I_n, M' denotes the transpose of a matrix M, Tr(M) its trace and $|M|$ its determinant. The notations $M\succ 0$ and $M\succeq 0$ stand respectively for M positive definite and positive semi-definite. Given an $n\times n$ matrix $Q\succeq 0$ and $x\in\mathbb{R}^n$, $\sqrt{\frac{1}{2}x'Qx}$ is denoted by $\|x\|_Q$. Let X and Y be two subsets of Euclidean spaces. The set of functions from X to Y is denoted by Y^X. For a subset $A\subset X$, we denote ∂A, Int(A) and $\bar{A}$ its boundary, interior and closure respectively. $C(X)$ refers to the set of $\mathbb{R}^n$-valued continuous functions on X with the standard supremum norm $\|.\|_\infty$, and $C^{i,j}(X\times Y)$ to the set of continuous real-valued functions $f(x,y)$ on $X\times Y$ such that the first i partial derivatives with respect to the x variable and the first j partial derivatives with respect to the y variable are continuous. The normal distribution of mean μ and covariance matrix Σ is denoted by $\mathcal{N}(\mu,\Sigma)$.

B. Problem Statement

We present in this section the dynamic collective choice model with social interactions. We consider a dynamic non-cooperative game involving a large number N of players with the following stochastic dynamics:

$$dx_i(t) = (A_i x_i(t) + B_i u_i(t))\,dt + \sigma_i dw_i(t), \qquad (1)$$

for $1\le i\le N$, where $A_i\in\mathbb{R}^{n\times n}$, $B_i\in\mathbb{R}^{n\times m}$, $\sigma_i\in\mathbb{R}^{n\times n}$, and $\{w_i, 1\le i\le N\}$ are N independent Brownian motions in $\mathbb{R}^n$ on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. In the remaining of the paper, $\mathbb{P}(A)$ denotes the probability of an event A and $\mathbb{E}(X)$ the expectation of a random variable X. We assume that the initial conditions $\{x_i(0), 1\le i\le N\}$ are independent and identically distributed (i.i.d.) and also independent of $\{w_i, 1\le i\le N\}$. Moreover, we assume that $\mathbb{E}\|x_i(0)\|^2<\infty$. The vector $x_i(t)\in\mathbb{R}^n$ is the state of player i at time t and $u_i(t)\in\mathbb{R}^m$ its control input. Let p_j, $1\le j\le l$, be l points in $\mathbb{R}^n$. Each player $i\in\{1,\dots,N\}$ is associated with the following individual cost functional:

$$J_i\left(x_i(0),u_i,u_{-i}\right) = \mathbb{E}\bigg[\int_0^T \left\{\|x_i-\bar{x}\|^2_{Q_i} + \|u_i\|^2_{R_i}\right\}dt + \min_{1\le j\le l}\|x_i(T)-p_j\|^2_{M_i}\ \bigg|\ x_i(0)\bigg], \qquad (2)$$

where $T>0$, $Q_i\succeq 0$, $R_i\succ 0$ and $M_i\succ 0$. Along the path, the running cost encourages the players to remain grouped around the average of the population $\bar{x}(t) := \sum_{i=1}^N x_i(t)/N$, which captures the social effect, and to develop as little effort as possible. The form of the final cost captures the discrete choice aspect, if we assume M_i large. That is, a player i at time T should be close to one of the destination points $p_j\in\mathbb{R}^n$, otherwise it is strongly penalized. Indeed, we show in [22] (for the scalar case) that the probability that an agent is close to one of the destination points increases with the final cost's coefficients M_i. Hence, the overall individual cost captures the problem faced by each agent of deciding between l alternatives, while trying to remain close to the mean population trajectory. It should be noted that the analysis of the game doesn't change if we allow the drift term in (1) to depend linearly on $\bar{x}$. The cost (2) has a LQG running cost and a final cost the minimum of l quadratic terms, hence the nomenclature "Min-LQG".

Let us denote the individual parameters $\theta_i := (A_i,B_i,\sigma_i,Q_i,R_i,M_i)$. We assume that there are k types of agents, that is, θ_i takes values in a finite set $\{\Theta_1,\dots,\Theta_k\}$, which does not depend on the size of the population N. As N tends to infinity, it is convenient to represent the limiting sequence of $(\theta_i)_{i=1,\dots,N}$ by realizations of a random vector θ, which takes values in the same finite set $\{\Theta_1,\dots,\Theta_k\}$. Let us denote the empirical measure of the sequence θ_i as $P_\theta^N(\Theta_s) = 1/N\sum_{i=1}^N \mathbf{1}_{\{\theta_i\}}(\Theta_s)$ for $s=1,\dots,k$. We assume that P_θ^N has a limit $P_\theta = (\alpha_1,\dots,\alpha_k)$ as $N\to\infty$, i.e. $\lim_{N\to\infty} P_\theta^N(\Theta_s) := \lim_{N\to\infty}\alpha_s^N = \alpha_s$ for $s=1,\dots,k$. For further discussions about this assumption, one can refer to [23].

Let $\mathcal{M}_i([0,T],\mathbb{R}^m)$, $1\le i\le N$ be the set of progressively measurable $\mathbb{R}^m$-valued functions with respect to the filtration generated by the initial condition of player i and its Brownian motion $\{\mathcal{F}(x_i(0),w_i(s),0\le s\le t)\}_{t\in[0,T]}$. We define the set of admissible control laws for an agent i as

$$\mathcal{U}_i = \left\{u_i\in\mathcal{M}_i([0,T],\mathbb{R}^m)\ \middle|\ \mathbb{E}\int_0^T\|u_i(s)\|^2 ds<\infty\right\} \qquad (3)$$

If $u_i\in\mathcal{U}_i$, then the stochastic differential equation (SDE) (1) has a unique strong solution [24, Section 5.2]. We define the set of admissible Markov policies

$$\mathcal{L} = \Big\{u\in(\mathbb{R}^m)^{[0,T]\times\mathbb{R}^n}\Big|\exists L_1>0, \forall(t,x)\in[0,T]\times\mathbb{R}^n,\ \|u(t,x)\|\le L_1(1+\|x\|),\ \text{and}\ \forall r>0, T'\in(0,T), \exists L_2>0,\ \forall\|(x,y)\|\le r, t\in[0,T'], \|u(t,x)-u(t,y)\|\le L_2\|x-y\|\Big\}. \qquad (4)$$

An agent i minimizes its cost with respect to its admissible control set $\mathcal{U}_i$. In addition, as shown in Theorem 2 below, it turns out that no loss of optimality is incurred if one further constrains the control strategies in $\mathcal{U}_i$ to be Markov (feedback policy) and satisfying the linear growth and Lipschitz properties in (4). These properties guaranty that the SDE (1) has a unique strong solution [24, Section 5.2].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

IV. The Min-LQG Optimal Tracking Problem, the Generic Agent's Best Response, and relation to the discrete choice models

A. The Min-LQG Optimal Tracking Problem

Following the MFG approach, we start by assuming a continuum of agents for which one can ascribe a *deterministic* macroscopic behavior $\bar{x}$ (posited mean population state trajectory), which is supposed known in this section. The problem of determining $\bar{x}$ is treated in Section V. In order to compute its best response to $\bar{x}$, a generic agent with parameters $\theta = (A, B, \sigma, Q, R, M) \in \{\Theta_1, \dots, \Theta_k\}$ solves the following optimal control problem, which we call the "Min-LQG" optimal tracking problem:

$$\begin{aligned}\inf_{u\in\mathcal{U}} J(x(0), u, \bar{x}) = \inf_{u\in\mathcal{U}} \mathbb{E}\Big[\int_0^T \{\|x-\bar{x}\|_Q^2 + \|u\|_R^2\}\, dt \\ + \min_{j=1,\dots,l} \|x(T) - p_j\|_M^2 \Big| x(0)\Big] \\ \text{s.t. } dx(t) = (Ax(t) + Bu(t))\, dt + \sigma dw(t),\end{aligned} \tag{5}$$

where w is a Brownian motion in $\mathbb{R}^n$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $x(0)$ is a random vector independent of w and distributed according to the known distribution of the agents' initial states. Herein, $\frac{\partial h(t,x)}{\partial x}$ and $\frac{\partial^2 h(t,x)}{\partial x^2}$ will denote respectively the gradient and Hessian matrix of the real function h with respect to $x \in \mathbb{R}^n$. The optimal cost-to-go function of (5) satisfies the following HJB equation [25]

$$\begin{aligned}-\frac{\partial V}{\partial t} &= x'A'\frac{\partial V}{\partial x} - \frac{1}{2}\left(\frac{\partial V}{\partial x}\right)' BR^{-1}B'\frac{\partial V}{\partial x} \\ &\quad + \frac{1}{2}\text{Tr}\left(\sigma'\frac{\partial^2 V}{\partial x^2}\sigma\right) + \|x-\bar{x}\|_Q^2 \\ V(T,x) &= \min_{1\le j\le l} \|x - p_j\|_M^2,\ \forall x \in \mathbb{R}^n.\end{aligned} \tag{6}$$

Equation (6) is similar to the LQG HJB equation except for the boundary condition, which is the minimum of l quadratic terms. In the following, we linearize equation (6) under appropriate conditions, using a generalized Hopf-Cole transformation [26, Section 4.4]. Moreover, we derive an explicit formula for the solution of (6) and the min-LQG optimal control law.

The following entities are used to define the solution of (6). Let $x^{(j)}$, $u^{(j)}$ and $V^{(j)}$ be respectively the optimal state trajectory, optimal control law and optimal cost-to-go of the LQG tracking problem that the generic agent must solve when p_j is the only available alternative, that is, (5) with $p_k = p_j$, for all $k \in \{1, \dots, l\}$. Recall that [27, Chapter 6]

$$V^{(j)}(t,x) = \frac{1}{2}x'\Pi(t)x + x'\beta^{(j)}(t) + \delta^{(j)}(t) \tag{7}$$

$$u^{(j)}(t,x) = -R^{-1}B'\left(\Pi(t)x + \beta^{(j)}(t)\right) \tag{8}$$

$$dx^{(j)}(t) = \left(Ax^{(j)}(t) + Bu^{(j)}\left(t, x^{(j)}(t)\right)\right) dt + \sigma dw^{(j)}(t), \tag{9}$$

where the positive definite matrix Π, the vector $\beta^{(j)}$ and scalar $\delta^{(j)}$ are the unique solutions of

$$\begin{aligned}\frac{d}{dt}\Pi(t) &= \Pi(t)BR^{-1}B'\Pi(t) - A'\Pi(t) - \Pi(t)A - Q, \\ \frac{d}{dt}\beta^{(j)}(t) &= -\left(A - BR^{-1}B'\Pi(t)\right)'\beta^{(j)}(t) + Q\bar{x}(t), \\ \frac{d}{dt}\delta^{(j)}(t) &= \frac{1}{2}\beta^{(j)}(t)'BR^{-1}B'\beta^{(j)}(t) - \frac{1}{2}\text{Tr}(\sigma'\Pi(t)\sigma) \\ &\quad - \|\bar{x}(t)\|_Q^2,\end{aligned} \tag{10}$$

with $\Pi(T) = M, \beta^{(j)}(T) = -Mp_j$ and $\delta^{(j)}(T) = \|p_j\|_M^2$. We denote by $\mathcal{W}^{(j)}$, for $j \in \{1, \dots, l\}$, the Voronoi cell associated with p_j, that is, $\mathcal{W}^{(j)} = \{x \in \mathbb{R}^n$ such that $\|x - p_j\|_M \le \|x - p_k\|_M$, for all $1 \le k \le l\}$. We define for all $j \in \{1, \dots, l\}$ the following notation for the conditional probability of an agent following the control law $u^{(j)}$ to be in the Voronoi cell j at time T given that its state at time t is x

$$g^{(j)}(t,x) \triangleq \mathbb{P}\left(x^{(j)}(T) \in \mathcal{W}^{(j)} \Big| x^{(j)}(t) = x\right). \tag{11}$$

Under Assumptions 1 and 2, the probability $g^{(j)}$ can be written as follows,

$$\begin{aligned}g^{(j)}(t,x) &= \mathbb{P}\left(x^{(j)}(T) \in \mathcal{W}^{(j)} \Big| x^{(j)}(t) = x\right) \\ &= \frac{1}{\sqrt{|2\pi\Sigma_t|}}\int_{\mathcal{W}^{(j)}} \exp\Bigg(-\Bigg\|y - \alpha(T,t)x \\ &\quad + \int_t^T \alpha(T,\tau)BR^{-1}B'\beta^{(j)}(\tau)d\tau\Bigg\|^2_{\Sigma_t^{-1}}\Bigg) dy \\ \text{for } \Sigma_t &= \int_t^T \alpha(T,\tau)\sigma\sigma'\alpha(T,\tau)'d\tau,\end{aligned} \tag{12}$$

where $x^{(j)}$, Π and $\beta^{(j)}$ are defined in (9) and (10), and the matrix-valued function $\alpha(t,s)$ is the unique solution of

$$\frac{d}{dt}\alpha(t,s) = \left(A - BR^{-1}B'\Pi(t)\right)\alpha(t,s), \tag{13}$$

with $\alpha(s,s) = I_n$. The matrix Σ_t, which is the Gramian [28] of $(A - BR^{-1}B'\Pi(t), \sigma) = (A - \eta\sigma\sigma'\Pi(t), \sigma)$, is invertible under Assumptions 1 and 2. The expression (12) follows from the fact that the solution of a linear SDE with deterministic initial condition has a normal distribution [24, Section 2.5].

To linearize and solve the HJB equation (6) using the Hopf-Cole transformation (see Appendix A), we make the following assumption.

Assumption 1: We assume that there exists a scalar $\eta > 0$ such that $BR^{-1}B' = \eta\sigma\sigma'$.

Remark 1: Note the following:

i. Assumption 1 always holds in the scalar case ($n = m = 1$).
ii. Assumption 1 is satisfied in particular if $B = R = \sigma = I_n$, a situation that has been studied previously in the context of other mean-field games (with $A = 0$) using the Hopf-Cole transformation, see [29, Chapter 2] and the references therein.
iii. Suppose that $n = m$, $A = Q = 0$, $B = \text{diag}(b_1, \dots, b_n)$, $R = \text{diag}(r_1, \dots, r_n)$ and $\sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$. σ_j is the intensity of the noise in the $j-$th direction. The control variable acts on the state in the $j-$th direction through the coefficient b_j, and the cost of this action in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

the same direction is evaluated through the coefficient r_j. Hence, the ratio b_j^2/r_j measures the efficiency of the control variable in the $j-$th direction. Following these interpretations, Assumption 1 requires that the ratio of the control efficiency to the noise intensity is identical in all the directions. In other words, it imposes a sort of isotropy on the ratio "control efficiency / noise intensity". Thus, the Hopf-Cole transformation is a linearizing transformation and success in this linearization relies on some global symmetries (isotropy "control efficiency / noise intensity"). In future work, it would be interesting to study the local Lie symmetries [30] of our HJB equation and develop an explicit solution under weaker assumptions.

iv. The coefficient R is a design parameter. It can be chosen to help satisfy Assumption 1.

Assumption 2: We assume that (A, σ) is controllable.

We now state the main result of this section, which is proved in Appendix A.

Theorem 1: Under Assumptions 1 and 2, the HJB equation (6) has a unique strong solution $(t,x) \mapsto V(t,x)$ in $C^\infty([0,T) \times \mathbb{R}^n) \cap C([0,T] \times \mathbb{R}^n)$, defined as

$$\begin{aligned} V(t,x) &= -\frac{1}{\eta} \log \left(\sum_{j=1}^{l} \exp\left(-\eta V^{(j)}(t,x)\right) g^{(j)}(t,x) \right), \\ &\qquad \forall (t,x) \in [0,T) \times \mathbb{R}^n \\ V(T,x) &= \min_{j=1,\dots,l} \|x - p_j\|_M^2, \ \forall x \in \mathbb{R}^n, \end{aligned} \tag{14}$$

where V_j and g_j are defined in (7) and (11), respectively.

Having solved the HJB equation related to the Min-LQG optimal control problem (5), we now prove the existence of a unique optimal control law. We define the following function:

$$\begin{aligned} u^*(t,x) &= -R^{-1} B' \frac{\partial V}{\partial x}, \qquad t \in [0,T) \\ u^*(T,x) &= 0. \end{aligned} \tag{15}$$

The proof of the following Theorem is given in Appendix A.

Theorem 2: Under Assumptions 1 and 2, the following statements hold:

i. The function u^* defined in (15) has the following form on $[0,T) \times \mathbb{R}^n$:

$$u^*(t,x) = \sum_{j=1}^{l} \frac{\exp\left(-\eta V^{(j)}(t,x)\right) g^{(j)}(t,x)}{\sum_{k=1}^{l} \exp\left(-\eta V^{(k)}(t,x)\right) g^{(k)}(t,x)} u^{(j)}(t,x), \tag{16}$$

with $V^{(j)}, u^{(j)}$ defined in (7) and (8), respectively.

ii. u^* is an admissible Markov policy.

iii. $u^*(t, x^*(t,w))$ is the unique optimal control law of (5), where $x^*(t,w)$ is the unique strong solution of the SDE in (5) with u equal to $u^*(t,x)$.

In the degenerate case with $\sigma = 0$, it is shown in [19]–[21] that the optimal strategy of an agent i in the Min-LQG problem is equal to $u^{(j)}$ (the optimal strategy in the presence of only one alternative p_j) if the Linear Quadratic Regulator (LQR) control problem associated with p_j is the least costly starting from $x_i(0)$. Therefore, a generic agent commits from the start to its final choice based on its initial state. When $\sigma \neq 0$, the generic agent can no longer be "decisive". Its optimal control law (16) is a convex combination of the optimal policies $u^{(j)}$, $j = 1, \dots, l$. The weights of $u^{(j)}$ form a spatio-temporal Gibbs distribution [31], which puts more mass on the less costly and risky destinations. A destination p_j is considered riskier in state x at time t if the Brownian motion has a higher chance of driving the state of an agent closer to a destination different from p_j at time T, when this agent implements $u^{(j)}$ from (x,t) onwards.

B. Relation to Discrete Choice Models

In the rest of this section, we discuss the relation between our solution to the Min-LQG optimal control problem in the scalar binary choice case and the solution of static discrete choice models. We start by recalling some facts about the static models. In the standard binary discrete choice models, a generic individual chooses between two alternatives 1 and 2. The cost paid by this individual when choosing an alternative j is defined by $v^{(j)} = k(j) + \nu$, where $k(j)$ is a deterministic function that depends on personal publicly observable attributes and on alternative j, while ν is a random variable accounting for personal idiosyncrasies unobservable by the social planner. When ν is distributed according to the extreme value distribution [10], then the probability that a cost-minimizing generic individual chooses an alternative j is equal to $P^{(j)} = \frac{\exp(-k(j))}{\exp(-k(1)) + \exp(-k(2))}$. Now, the Min-LQG optimal control law (16) can be written as follows:

$$\begin{aligned} u^*(t,x) &= \frac{\exp\left(-\eta \tilde{V}^{(1)}(t,x)\right)}{\exp\left(-\eta \tilde{V}^{(1)}(t,x)\right) + \exp\left(-\eta \tilde{V}^{(2)}(t,x)\right)} u^{(1)}(t,x) \\ &+ \frac{\exp\left(-\eta \tilde{V}^{(2)}(t,x)\right)}{\exp\left(-\eta \tilde{V}^{(1)}(t,x)\right) + \exp\left(-\eta \tilde{V}^{(2)}(t,x)\right)} u^{(2)}(t,x), \end{aligned} \tag{17}$$

where

$$\tilde{V}^{(j)}(t,x) = V^{(j)}(t,x) - \frac{1}{\eta} \log\left(g^{(j)}(t,x)\right), j = 1, 2. \tag{18}$$

Here $V^{(j)}(t,x)$ is the expected cost paid by a generic agent if p_j were the only available alternative, and $u^{(j)}(t,x)$ is the corresponding optimal policy. In the presence of two alternatives, the optimal policy at time t is given by (17), which can be interpreted as the mean of a randomized strategy which is a mix of two pure strategies $u^{(1)}(t,x)$ (picking alternative p_1) and $u^{(2)}(t,x)$ (picking alternative p_2). Within this framework, denoting by $-j$ the alternative other than j, a generic agent at time t chooses the alternative p_j with probability

$$Pr^{(j)} = \frac{\exp\left(-\eta \tilde{V}^{(j)}(t,x)\right)}{\exp\left(-\eta \tilde{V}^{(j)}(t,x)\right) + \exp\left(-\eta \tilde{V}^{(-j)}(t,x)\right)}.$$

Thus, the Min-LQG problem can be viewed at each time $t \in [0,T]$ as a static discrete choice problems, where the cost of choosing alternative p_j includes an additional term $-\frac{1}{\eta} \log\left(g^{(j)}(t,x)\right)$ accounting for the risk of landing nearer

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

to $-j$ at time T because of unexpected random fluctuations, while decisions were in fact optimized with a landing at p_j in mind.

V. The Mean Field Equations and The Fixed Point Problem

In Section IV, we assumed the mean trajectory $\bar{x}$ known and we computed the generic agent's best response to it, which is given by (16). In the following, a subscript s refers to an agent with parameters $\Theta_s \in \{\Theta_1, \ldots, \Theta_k\}$. We write $u_s^*(t, x, \bar{x})$ instead of $u_s^*(t, x)$ to emphasize the dependence on $\bar{x}$. We now seek a sustainable macroscopic behavior $\bar{x}$, in the sense that it is replicated by the mean of the generic agent's state under its best response to it. Thus, an admissible $\bar{x}$ satisfies the following mean field equations:

$$\bar{x}(t) = \sum_{s=1}^{k} \alpha_s \bar{x}_s(t), \text{ with } \bar{x}_s = \mathbb{E}[x_s^*],\ 1 \le s \le k, \quad (19)$$
$$dx_s^*(t) = \left(A_s x_*^s(t) + B_s u_s^*\left(t, x_s^*(t), \bar{x}\right)\right) dt + \sigma_s dw_s(t),$$

where $x_s^*(.)$, $x_s^*(0)$, respectively represent the trajectory of a generic agent of type s and its associated initial condition, $1 \le s \le k$. Initial conditions are independent and identically distributed according to assumed common probability law, i.e., the distribution of $x_i(0)$ irrespective of the type. α_s is the proportion of agents of type s as defined in Section III and $\{w_1, \ldots, w_k\}$ are k independent Brownian motions, assumed independent of $\{x_1^*(0), \ldots, x_k^*(0)\}$.

The mean field equations in (19) are a $n \times k$ nonlinear McKean-Vlasov equation [13], where the drift term depends on the joint probability law of the k types of agents through the mean trajectory $\bar{x}$. The solution of such equations corresponds to computing a fixed point on the space of joint probability laws of the k generic agent states over the time interval [0,T], given their evolution as law dependent SDE's. A direct solution can be quite challenging. Since $\bar{x}$ is really the trajectory of interest, we focus instead on a simpler characterization of that quantity. It is obtained via a consideration of the following stochastic maximum principle equations associated with the best responses of each of the k agents to a sustainable trajectory $\bar{x}$,

$$\begin{aligned} dx_s^*(t) &= \left(A_s x_s^*(t) - B_s R_s^{-1} B_s' q_s(t)\right) dt \\ &\quad + \sigma_s dw_s(t) \quad (20) \\ -dq_s(t) &= \left(A_s' q_s(t) + Q_s(x_s^*(t) - \bar{x}(t))\right) dt \\ &\quad - \frac{\partial^2 V_s}{\partial x^2}\left((t, x_s^*(t)\right) \sigma_s dw_s(t), \quad (21) \end{aligned}$$

with $q_s(t) = \frac{\partial V_s}{\partial x}(t, x_s^*(t))$ and $q_s(T) = M_s\left(x_s^*(T) - \sum_{j=1}^{l} 1_{\mathcal{W}^{(j)}}(x_s^*(T)) p_j\right)$. This maximum principle is derived in Lemma 7, Appendix B under the following assumption,

Assumption 3: We assume that σ_s is invertible, for $s = 1, \ldots, k$.

It should be noted that Assumption 3 implies Assumption 2. Taking expectations of these forward-backward SDEs over both agent initial conditions and type leads to linear forward backward deterministic coupled differential equations (See (23a)-(23b) below) with nonlinear coupling in the boundary condition characterizing sustainable $\bar{x}$ trajectories.

Define the vectors $X^* = (x_1^*, \ldots, x_k^*)$, $U^* = (u_1^*, \ldots, u_k^*)$, $\bar{X} = (\bar{x}_1, \ldots, \bar{x}_k)$ and $X^*(0) = (x_1^*(0), \ldots, x_k^*(0))$ of respectively, the optimal states, the optimal control laws, expected optimal states and initial states for k generic agents representing the k types, when they optimally respond to a sustainable $\bar{x}$. X^* and U^* satisfy

$$dX^*(t) = \left(AX^*(t) + BU^*\left(t, X^*(t), \bar{x}\right)\right) dt + \sigma dW(t), \quad (22)$$

which is the matrix form of (19), where $W = (w_1, \ldots, w_k)$, and A, B, σ are the block-diagonal matrices diag$(A_1, \ldots, A_k)$, diag$(B_1, \ldots, B_k)$ and diag$(\sigma_1, \ldots, \sigma_k)$, respectively.

The equivalent representation for $\bar{x}$ is given by the following equations

$$\frac{d}{dt}\bar{X}(t) = A\bar{X}(t) - BR^{-1}B'\bar{q}(t), \quad (23a)$$
$$\frac{d}{dt}\bar{q}(t) = -A'\bar{q}(t) + QL\bar{X}(t), \quad (23b)$$
$$\bar{x}(t) = P_1 \bar{X}(t), \quad (23c)$$

with $Q = \text{diag}(Q_1, \ldots, Q_k)$, $R = \text{diag}(R_1, \ldots, R_k)$, $M = \text{diag}(M_1, \ldots, M_k)$ and $L = I_{nk} - 1_k \otimes P_1$, where 1_k is a column of k ones and $P_1 = P_\theta' \otimes I_n$ with P_θ the vector of α_s as defined in Section III. The initial condition for (23a) is $\bar{X}(0) = \mathbb{E}X^*(0) = \mathbb{E}[x(0)] \otimes 1_k$, and the terminal condition for the backward equation (23b) is $\bar{q}(T) = M\left(\bar{X}(T) - (\Lambda \otimes I_n)p\right)$, where the $k \times l$ matrix Λ is defined as

$$\Lambda_{sj} = \mathbb{P}\left(x_s^*(T) \in \mathcal{W}^{(j)}\right), \qquad 1 \le s \le k,\ 1 \le j \le l, \quad (24)$$

and $p = (p_1, \ldots, p_l)$. Equations (23a)-(23b) are respectively the aggregates of the state and co-state equations (20)-(21) of the generic agents. They have a nonlinear coupling in the boundary condition $\bar{q}(T)$ through what we call a "Choice Distribution Matrix" (CDM) Λ. A CDM is a $k \times l$ row stochastic matrix with its (s, j) entry equal to the probability that the generic agent of type s is at time T closer (in the sense of the M-weighted 2-norm) to p_j than any of the other alternatives when it optimally responds to $\bar{x}$. The fact that (23a)-(23c) characterize $\bar{x}$ is proved in the first point of Theorem 3 below.

The next step is to exploit this new representation to characterize all candidate sustainable mean trajectories in (19). Indeed, the advantage of this new representation is that if one considers the CDM in the boundary condition $\bar{q}(T)$ as a parameter (say any $k \times l$ row stochastic matrix), then equations (23a)-(23b) become two coupled linear forward-backward differential equations. As a result, they have a unique solution under the following assumption.

Assumption 4: We assume the existence of a solution on $[0, T]$ to the following (nonsymmetric) Riccati equation:

$$\frac{d}{dt}\pi(t) = -A'\pi(t) - \pi(t)A + \pi(t)BR^{-1}B'\pi(t) + QL, \quad (25)$$

with $\pi(T) = M$.

Remark 2: If Assumption 4 is satisfied, the solution of (25) is unique as a consequence of the smoothness of the right-hand side of (25) with respect to π [32, Section 2.4, Lemma 1].

This assumption is satisfied, for example, when QL is positive semi-definite [33, Section 2.3]. This happens, for instance, in case of a uniform population ($k=1$), in which case $L=0$. For more details about Assumption 4, one can refer to [34].

The Riccati equation (25) arises in LQG dynamic games [35]. Just like the Riccati function in optimal control theory allows decoupling of the state and co-state equations, here the Riccati function π in (25), if it exists, permits decoupling of (23a)-(23b) (when Λ is considered as a parameter). As shown in the proof of the second point of Theorem 3, these equations have then the following explicit solution parametrized by Λ,

$$\bar{X}(t) = R_1(t,0)\bar{X}(0) + R_2(t)(\Lambda \otimes I_n)\, p, \tag{26}$$

where

$$\begin{aligned} \frac{d}{dt}R_1(t,s) &= \left(A - BR^{-1}B'\pi(t)\right) R_1(t,s), \\ \frac{d}{dt}R_2(t) &= \left(A - BR^{-1}B'\pi(t)\right) R_2(t) \\ &\quad + BR^{-1}B'R_1(T,t)'M, \end{aligned} \tag{27}$$

with $R_1(s,s) = I_{nk}$ and $R_2(0) = 0$. This implies that the solutions $\bar{x}$ of the mean field equations lie in the following family of paths parametrized by matrices Λ in the set S of $k \times l$ row stochastic matrices

$$\bar{x}_\Lambda(t) := P_1(R_1(t,0)\bar{X}(0) + R_2(t)(\Lambda \otimes I_n)\, p). \tag{28}$$

Conversely, for a path $\bar{x}_\Lambda$ parametrized by some candidate Λ to be a solution of (19), consistency requires that Λ be equal to the associated CDM as expressed in (24) when the generic agents optimally respond to $\bar{x}_\Lambda$. This is equivalent to requiring that Λ be a fixed point of the following finite dimensional map $F : S \mapsto S$,

$$F(\Lambda)_{sj} = \mathbb{P}\left(x^*_{s,\Lambda}(T) \in \mathcal{W}^{(j)}\right), \tag{29}$$

where $X^*_\Lambda = (x^*_{1,\Lambda}, \dots, x^*_{k,\Lambda})$ is the unique strong solution of the following SDE parameterized by Λ

$$dX^*_\Lambda(t) = \left(AX^*_\Lambda(t) + BU^*(t, X^*_\Lambda(t), \bar{x}_\Lambda)\right)dt + \sigma dW(t),$$
$$\text{with } X^*_\Lambda(0) = X^*(0). \tag{30}$$

Note that although the notation is partially shared with (19), (30) is no longer a McKean-Vlasov equation (with solutions joint probability laws) but rather an SDE for a vector of k generic agents, with each of the components of $X^*_\Lambda(0)$ independent, and each being a sample random variable from the common distribution assumed for agent initial conditions across all types. F maps a row stochastic matrix Λ to the CDM when the generic agents optimally respond to $\bar{x}_\Lambda$. The map F involves the probability distribution of the process X^*_Λ. Hence, to find the value of $F(\Lambda)$, one needs to solve the Fokker-Planck equation associated with (30).

In effect, we establish that there is a one to one map between the solutions $\bar{x}$ of (19) and the fixed points of the finite dimensional map F. Theorem 3 below summarizes the related results: point (ii) of the theorem states the existence of the one-to-one map in question, while point (iii) states that F has at least one fixed point. Thus, a Nash equilibrium CDM exists (equivalently a solution of (19)), which characterizes the way an infinite population splits between the destination points, when it has the same distribution of heterogeneous parameters as the original large but finite population.

The following results are proved in Appendix B.

Theorem 3: Under Assumptions 1, 3 and 4, the following statements hold:

(i) $\bar{x}$ satisfies (19) if and only if it satisfies (23a)-(23c).

(ii) $\bar{x}$ satisfies (19) if and only if

$$\bar{x} = \bar{x}_\Lambda \tag{31}$$

where $\bar{x}_\Lambda$ is defined in (28) and Λ is a fixed point of F defined in (29)-(30).

(iii) F is continuous and has at least one fixed point. Equivalently, (19) has at least one solution $\bar{x}$.

Remark 3: In [13], [36], which consider the general MFG theory, the authors show the existence and uniqueness of solutions for the McKean-Vlasov equation describing the mean field behavior via Banach's [13] or Schauder's fixed point theorem [36]. In [13], it is assumed that the optimal control law is regular enough (Lipschitz continuous with respect to the state and the distribution) in order to define a contraction, while in [36] the result is proved under the assumption of smooth and convex final cost. In our case, the control law (16) is not Lipschitz continuous with respect to $\bar{x}$. Moreover, the final cost is neither smooth nor convex. Hence, the McKean-Vlasov equation (19) might have multiple solutions. Indeed, (19) has a number of solutions equal to the number of fixed points of F.

Having solved the game for a continuum of players, we now return to the practical case of a finite population of players. Using arguments similar to those in [21, Theorem 8], one can show that the MFG-based decentralized strategies (16), when applied by a finite population, constitute an $\epsilon-$Nash equilibrium with respect to the sets of centralized strategies $\mathcal{U}^c_i = \{u_i \text{ adapted to } \mathcal{F}(x_k(s),\ 0 \leq s \leq t, 1 \leq k \leq N) \mid \mathbb{E}\int_0^T \|u_i(s)\|^2 ds < \infty\}$, $1 \leq i \leq N$. We give in Appendix B a sketch of the proof.

Theorem 4: The decentralized feedback strategies u^* defined in (16) for a fixed point path $\bar{x}$ of (19), when applied by N players with dynamics (1), constitute an ϵ_N-Nash equilibrium with respect to the costs (2) and control sets $\mathcal{U}^c_i$, $1 \leq i \leq N$, where $\epsilon_N = C(\frac{1}{\sqrt{N}} + \max\limits_{1\leq s \leq k} |\alpha_s - \alpha^N_s|)$ converges to zero as N increases to infinity. Here, C is a positive scalar independent of N, and α_s and α^N_s are defined in Section III-B.

The set of fixed points of the finite dimensional map F characterizes the game in terms of the number of distinct $\epsilon-$Nash equilibria and the distribution of the choices for each of them. In fact, Theorem 3 establishes a one to one map between the solutions of the mean field equations (19) and the set of fixed point CDMs. If Λ is a fixed point of F and the players optimally respond to the corresponding $\bar{x}_\Lambda$ given by (28), then Λ_{sj} is the fraction of agents of type s that go towards p_j, and $\sum_{s=1}^k \alpha_s \Lambda_{sj}$ is the total fraction of players choosing this alternative. Thus, to compute a path $\bar{x}$ satisfying (19), a player computes a fixed point Λ of F and then computes (28).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

The computation of the mean field term in the general MFG theory involves solving a forward Fokker-Planck equation coupled with a backward HJB equation [37]. In our case, however, Theorem 3 reduces this infinite dimensional problem to the computation of a fixed point for the finite dimensional map F. Thus, one can use the bisection method in the binary scalar case for a population with uniform dynamics and cost functions, i.e., $n = m = k = 1$, and $l = 2$, or a quasi Newton method such as Broyden's method [38] in the general case. In both cases, one has to propagate the Fokker-Planck equation associated to (30) to compute the values of F.

Let us briefly illustrate how to compute a fixed point of F in the first case. We consider a uniform population with initial probability density function f_0. In this case, a pair $(r, 1-r)$ for $r \in [0,1]$ is a fixed point of F if and only if r is a fixed point of G, where $G(r)$ is the first component of the 2-dimensional vector $F(r, 1-r)$. According to the third point of Theorem 3, F is continuous, and so is $G : [0,1] \to [0,1]$. Moreover, $(G(1) - 1)(G(0) - 0) \leq 0$. Thus, we can apply the bisection method to $G(r) - r$ to find a fixed point of G, if we can compute the value of G at any $r \in [0,1]$. But $G(r) = \int_{-\infty}^{c} f_r(T,x)dx$, where $c = (p_1 + p_2)/2$ and $f_r(t,x)$ is the probability density of $X^*_{(r,1-r)}(t)$ defined by (30), which can be computed numerically by solving the following Fokker-Planck equation associated to (30), e.g., via an implicit finite difference scheme [39]

$$\frac{\partial f_r(t,x)}{\partial t} = -\frac{\partial \left(\mu(t,x,r) f_r(t,x)\right)}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 f_r(t,x)}{\partial x^2}, \quad (32)$$

with $f_r(0,x) = f_0(x), \forall x \in \mathbb{R}$. Here $\mu(t,x,r) = Ax + Bu^*\left(t, x, \bar{x}_{(r,1-r)}\right)$ for $\bar{x}_{(r,1-r)}(t) = R_1(t,0)\bar{x}(0) + R_2(t)\left(rp_1 + (1-r)p_2\right)$, (see (28)). In the general case, one can apply Broyden's method to the function $F(\Lambda) - \Lambda$, where the values $F(\Lambda)$ are obtained by solving numerically the multi-dimensional version of (32).

VI. Simulation Results

To illustrate the dynamics of our collective decision mechanism, we consider a group of uniform agents with parameters $A = 0.1$, $B = 0.2$, $R = 5$, $M = 500$, $T = 2$ and $\sigma = 1.5$. The agents, initially drawn from the normal distribution $\mathcal{N}(0.3, 1)$, choose between the alternatives $p_1 = -10$ and $p_2 = 10$. Assumptions 1, 2 and 3 are satisfied because of the scalar dynamics, while Assumption 4 holds as a result of the uniform population assumption (See Remark 2). At first, we consider a weak social effect ($Q = 0.1$). Following the numerical scheme at the end of Section V, we find a fixed point $r = 0.39$. Accordingly, a player applying its decentralized MFG-based strategy is at time T closer to p_1 with probability 0.39. Equivalently, if we draw independently from the initial distribution a large number of players with independent Wiener processes, then the percentage of players that will be at time T closer to p_1 converges to 39% as the size of the population increases to infinity. Thus, the majority of the players choose p_2. Figure 1 shows the distribution at time $t = 0$, $t = 0.5T$ and $t = T$, the mean of a generic agent, the tracked path (admissible path (31)), and the sample paths of 10 players choosing between -10 and 10 under the weak social effect. As shown in this figure, the mean replicates the tracked path computed using the fixed point $r = 0.39$ and (31).

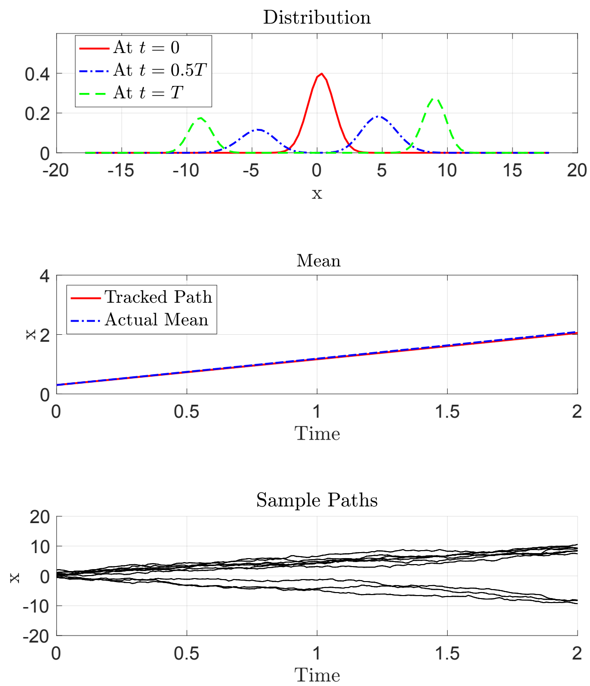


Fig. 1. Distribution evolution, mean trajectory, tracked path and sample paths for a weak social effect ($Q = 0.1$). Here $r = 0.39$.

Figure 2 shows that for sufficiently small values of Q ($Q < 21$) the fixed points of F are unique. When the social effect Q exceeds 21, F has three fixed points, where two of them correspond to consensus on one alternative. Indeed as Q increases arbitrarily, agents essentially forget temporarily about the final cost, and the problem becomes a classical rendez-vous MFG where they tend to merge towards each other rapidly. If this occurs around the middle of the destinations segment, then this is clearly an unstable situation where most of the time, they end up splitting classically according to initial conditions; however, some large deviations are possible whereby a significant fraction decides to choose one destination, thus pulling everyone else towards it, which may help explain the non uniqueness of outcomes. Although the frameworks are different, this behavior resembles to the pitchfork bifurcation diagrams studied in [40] to model the influence of the social effect on the behavior of a population of honeybees choosing between two nectar sites. Figure 3 illustrates the evolution of the distributions that correspond to the first fixed point with $Q = 10$ and $Q = 20$.

To illustrate the effect of the noise intensity on the behavior of the group, we fix $Q = 20$ and we increase σ from 1.5 to 5. For $\sigma = 1.5$, $r = 0.02$ (Figure 3), but for $\sigma = 3$, r increases to 0.28 (Figure 4) and for $\sigma = 5$, $r = 0.46$ (Figure 4). Thus, the higher the noise the more evenly distributed the players are between the alternatives.

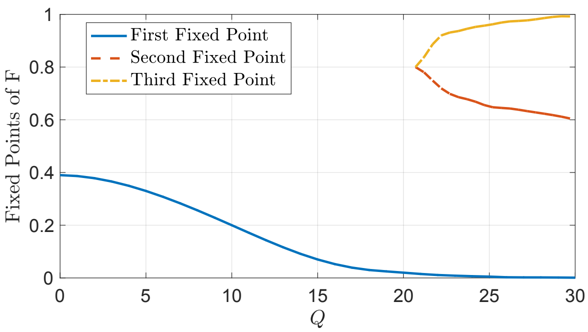


Fig. 2. Influence of the coefficient Q on the multiplicity of the fixed points

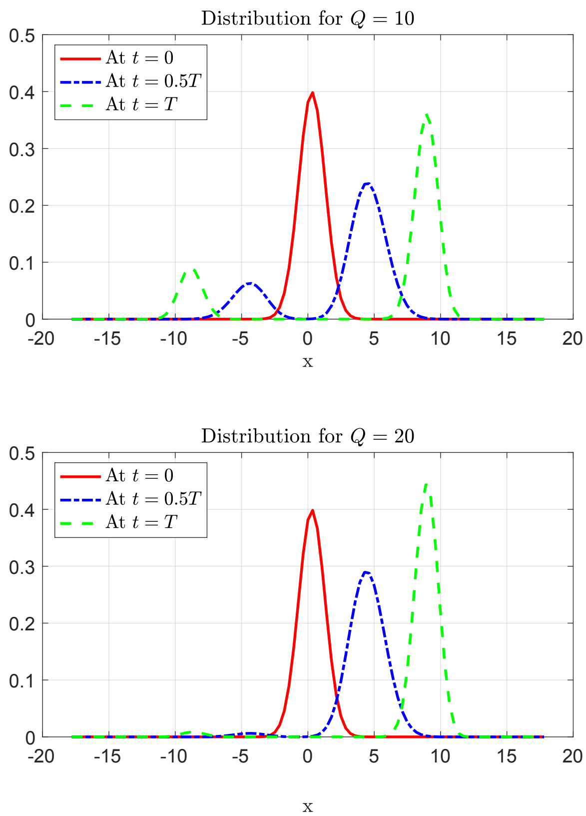


Fig. 3. Influence of the social effect on the distribution of the agents. Distribution evolution for medium ($Q = 10$) and strong ($Q = 20$) social effects. In the first case, $r = 0.2$, while in second case, $r = 0.02$.

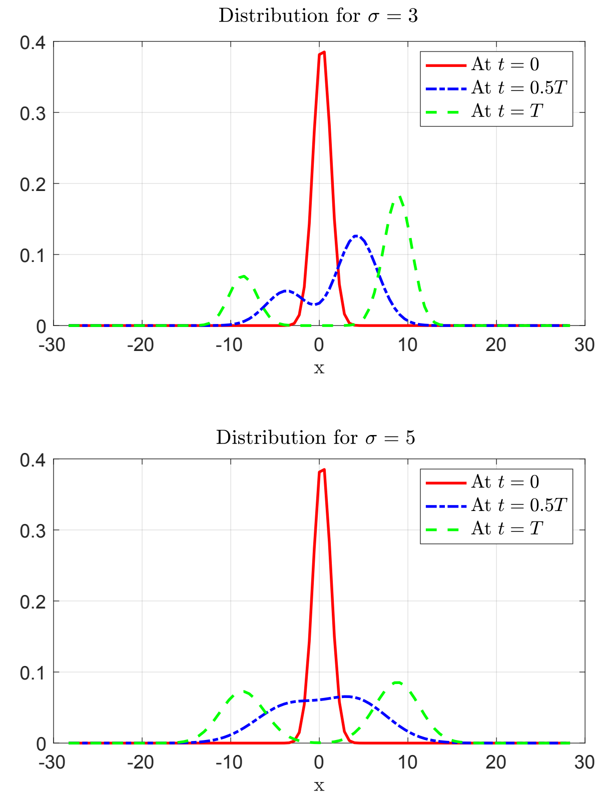


Fig. 4. Influence of the noise on the distribution of the agents. Distribution evolution for $\sigma = 3$ and $\sigma = 5$. In the first case, $r = 0.28$, while in the second case, $r = 0.46$.

VII. Conclusion

We have studied within the framework of MFG theory a dynamic collective choice model with social interactions. Towards that end, a novel class of optimal control problems called Min-LQG was formulated and an explicit form of a generic agent's best response was developed. The Min-LQG problem relates to the discrete choice literature, in that it can be interpreted at each time step as a static discrete choice model where the cost of choosing one of the alternatives has an additional term that increases with the risk of being driven by the process noise to the other alternatives. We have shown the existence of closed loop decentralized $\epsilon-$Nash strategies. Moreover, we have characterized these strategies by a so-called choice distribution matrix describing the way the population splits between the alternatives. It is shown to be a fixed point of a well defined finite dimensional map. The fixed points may in general be non unique. A numerical example of such behavior was given. In future work, it would be interesting to explore a situation of non uniqueness where agents would use pure or mixed strategies based on the possible anticipated distinct mean field equilibria. The policies would be refined via mutual observations of agents behavior as time unfolds. Furthermore, the nature of best responses under non isotropy along different dimensions should be further explored, in particular when some agent state components are not directly excited by noise. It is also of interest for future work to include a learning process in the game, where the players share periodically their states and parameters to learn the probability distributions of their initial conditions and parameters. Finally, we would like also to develop a maximum likelihood estimator that observes the sample paths of some agents and estimates the model parameters. Once the parameters are estimated, the model can be used for example to predict the evolution of opinions before elections and the distribution of voters' final choices between the candidates.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

Appendix A

In this appendix, we provide the proofs of lemmas and theorems related to a generic agent's best response.

Proof of Theorem 1

We start with a technical result on the mean-square convergence of random variables.

Lemma 5: Let I be a closed subset of $\mathbb{R}^n$. Let $X_k \in \mathbb{R}^n$ be a sequence of random variables with finite first and second moments. If $\mathbb{E}[X_k] =: \mu_k \to \mu$ for some vector μ not in I, and $\mathbb{E}[\|X_k - \mu_k\|^2] \to 0$, then $\lim_{k\to\infty} \mathbb{P}(X_k \in I) = 0$.

Proof: $I \subset \mathbb{R}^n$ is a closed set and $\mu \notin I$, so the distance d between μ and I is strictly positive. Since, μ_k converges to μ, there exists $k_0 > 0$ such that for all $k \geq 0$ and for all x in I we have $\|x - \mu_k\| \geq d/2$. Hence, using Chevyshev's inequality [41, Theorem 1.6.4],

$$\mathbb{P}(X_k \in I) \leq \mathbb{P}(\|X_k - \mu_k\| \geq d/2) \leq \frac{4}{d^2}\mathbb{E}[\|X_k - \mu_k\|^2],$$

for all $k \geq k_0$. The result follows since the right-hand side of the inequality is assumed to converge to 0. ∎

The following lemma is concerned with the regularity of the solution provided in Theorem 1.

Lemma 6: Under Assumptions 1 and 2, V defined in (14) is in $C^\infty([0,T) \times \mathbb{R}^n) \cap C([0,T] \times \mathbb{R}^n)$.

Proof: As discussed below equation (13), $\Sigma_t \succ 0$ for all $t \in [0,T)$. Hence, in view of (12) and (14), V is in $C^\infty([0,T) \times \mathbb{R}^n)$. It remains to show the continuity on $\{T\} \times \mathbb{R}^n$. We start by considering $x \in \mathbb{R}^n \setminus \cup_{j=1}^l \partial\mathcal{W}^{(j)}$, and $(t_k, x_k) \in [0,T) \times \mathbb{R}^n$ converging to (T,x). We have $x \in \mathrm{Int}(\mathcal{W}^{(j_0)})$ for some $j_0 \in \{1,\dots,l\}$, and $x \notin \mathcal{W}^{(j)}$ for $j \neq j_0$. In view of (12), $g^{(j)}(t_k,x_k)$ is the probability that a Gaussian vector of mean $\alpha(T,t_k)x_k - \int_{t_k}^T \alpha(T,\tau)BR^{-1}B'\beta^{(j)}(\tau)d\tau$ (which converges to x with k) and covariance Σ_{t_k} (which converges to 0 with k) is in the closed set $\mathcal{W}^{(j)}$. In this way, each j defines a distinct sequence of random variables associated with the (t_k,x_k)'s. Now, if one considers the closed set I of Lemma 5 to be any of the closed sets $\mathcal{W}^{(j)}$'s for $j \neq j_0$, one can conclude from this lemma that $g^{(j)}(t_k,x_k)$ must converge to 0 for $j \neq j_0$ and, as a consequence, to 1 for $j = j_0$ since the $\mathcal{W}^{(j)}$'s form a partition of the state space. Therefore, $V(t_k,x_k)$ converges to $V(T,x)$. Thus, V is continuous on $[0,T] \times \left(\mathbb{R}^n \setminus \cup_{j=1}^l \partial\mathcal{W}^{(j)}\right)$. Finally, consider a sequence $(t_k,x_k) \in [0,T) \times \mathbb{R}^n$ converging to (T,c), with $c \in \cup_{j=1}^l \partial\mathcal{W}^{(j)}$. We show that $V(t_k,x_k)$ converges to $V(T,c)$. Up to renumbering the Voronoi cells, we can assume without loss of generality that $c \in \partial\mathcal{W}^{(j)}$ for all j in $\{1,\dots,z\}$ and $c \notin \cup_{j=z+1}^l \mathcal{W}^{(j)}$, for some $1 \leq z \leq l$. We have,

$$\begin{aligned} I_0 &= \sum_{j=1}^{l} \exp\left(-\eta V^{(j)}(t_k,x_k)\right) g^{(j)}(t_k,x_k) \\ &= \sum_{j=1}^{z} \exp\left(-\eta V^{(j)}(t_k,x_k)\right) g^{(j)}(t_k,x_k) \\ &\quad + \sum_{j=z+1}^{l} \exp\left(-\eta V^{(j)}(t_k,x_k)\right) g^{(j)}(t_k,x_k). \end{aligned}$$

Since $c \notin \cup_{j=z+1}^l \mathcal{W}^{(j)}$, one can use an argument similar to the one above to show that the second term of the right-hand side of the second equality converges to 0.

Next, let $\epsilon > 0$ and fix $r > 0$ small enough so that the closed ball centered at c of radius r $\bar{B}(c,r) \subset \left(\cap_{j=z+1}^l \mathcal{W}^{(j)}\right)^C$. The value of r will be determined later. The first term can be written

$$\sum_{j=1}^{z} \exp\left(-\eta V^{(j)}(t_k,x_k)\right) g^{(j)}(t_k,x_k) = I_1 + I_2,$$

where

$$\begin{aligned} I_1 &= \sum_{j=1}^{z} \exp\left(-\eta V^{(j)}(t_k,x_k)\right) \\ &\qquad \times \mathbb{P}\left(x^{(j)}(T) \in \mathcal{W}^{(j)} \cap \bar{B}(c,r) \middle| x^{(j)}(t_k) = x_k\right) \\ I_2 &= \sum_{j=1}^{z} \exp\left(-\eta V^{(j)}(t_k,x_k)\right) \\ &\qquad \times \mathbb{P}\left(x^{(j)}(T) \in \mathcal{W}^{(j)} \setminus \bar{B}(c,r) \middle| x^{(j)}(t_k) = x_k\right). \end{aligned}$$

By Lemma 5, I_2 converges to zero. Next, by solving the linear differential equations in (10) and replacing the expressions of $\beta^{(j)}$ and $\delta^{(j)}$ in the expression (12) of $g^{(j)}$, one can show that under Assumptions 1 and 2

$$\begin{aligned} I_1 &= \exp\left(-\eta V^{(0)}(t_k,x_k)\right) \\ &\quad \times \sum_{j=1}^{z} \int_{\mathcal{W}^{(j)} \cap \bar{B}(c,r)} f_k(y) \exp\left(\eta(\|y\|_M^2 - \|y - p_j\|_M^2)\right) dy, \end{aligned}$$

where $f_k(y)$ is the probability density function of the Gaussian distribution of mean $\alpha(T,t_k)x_k - \int_{t_k}^T \alpha(T,\tau)BR^{-1}B'\beta^{(0)}(\tau)d\tau$ and variance Σ_{t_k}, and $V^{(0)}$ and $\beta^{(0)}$ are equal to $V^{(j)}$ and $\beta^{(j)}$ defined in (7)-(10) but for $p_j = 0$. By the definition of c, $\|c - p_1\|_M^2 = \dots = \|c - p_z\|_M^2$. Hence,

$$\begin{aligned} I_1 &= \exp\left(-\eta(V^{(0)}(t_k,x_k) - \|c\|_M^2 + \|c - p_j\|_M^2)\right) \\ &\quad \times \sum_{j=1}^{z} \int_{\mathcal{W}^{(j)} \cap \bar{B}(c,r)} f_k(y) dy \\ &\quad + \exp\left(-\eta V^{(0)}(t_k,x_k)\right)\sum_{j=1}^{z} \int_{\mathcal{W}^{(j)} \cap \bar{B}(c,r)} f_k(y) f(y) dy \\ &\triangleq I_3 + I_4 \end{aligned}$$

where $f(y) = \exp\left(\eta(\|y\|_M^2 - \|y - p_j\|_M^2)\right) - \exp\left(\eta(\|c\|_M^2 - \|c - p_j\|_M^2)\right)$. $V^{(0)}(t_k,x_k)$ converges to $V^{(0)}(T,c) = \|c\|_M^2$, f_k converges in distribution to a point mass at c and $\mathcal{W}^{(j)} \cap \bar{B}(0,c)$, $j = 1,\dots,z$, is a partition of $\bar{B}(0,c)$. Therefore, I_3 converges to $\exp(-\eta\|c - p_j\|_M^2) = \exp(-\eta V(T,c))$. f is continuous, and $f(c) = 0$. Hence, one can choose r small enough so that $|f(y)| < \epsilon$ for all $y \in \bar{B}(c,r)$. Thus, $|I_4| \leq \epsilon$, and $\limsup_k |I_0 - \exp(-\eta V(T,c))| \leq \epsilon$. Since ϵ is arbitrary, I_0 converges to $\exp(-\eta V(T,c))$. This proves the result. ∎

To finish the proof of Theorem 1, it remains to show that V satisfies the HJB equation (6). We define the transformations by a generalized Hopf-Cole transformation [26, Chapter 4-Section 4.4] of the LQG optimal cost-to-go $V^{(j)}(t,x)$, $\psi^{(j)}(t,x)=\exp\left(-\eta V^{(j)}(t,x)\right)$, for $j=1,\ldots,l$. Recall [27, Chapter 6] that the optimal cost-to-go $V^{(j)}$ satisfies the HJB equation,

$$\begin{aligned}-\frac{\partial V^{(j)}}{\partial t}&=x'A'\frac{\partial V^{(j)}}{\partial x}-\frac{1}{2}\left(\frac{\partial V^{(j)}}{\partial x}\right)'BR^{-1}B'\frac{\partial V^{(j)}}{\partial x}\\&\quad+\frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 V^{(j)}}{\partial x^2}\sigma\right)+\|x-\bar{x}\|_Q^2 \qquad (33)\\ V^{(j)}(T,x)&=\|x-p_j\|_M^2,\ \forall x\in\mathbb{R}^n.\end{aligned}$$

By multiplying the right-hand and left-hand sides of (33) by $-\eta\exp\left(-\eta V^{(j)}(t,x)\right)$, we obtain

$$\begin{aligned}&-\frac{\partial\psi^{(j)}}{\partial t}=x'A'\frac{\partial\psi^{(j)}}{\partial x}+\frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2\psi^{(j)}}{\partial x^2}\sigma\right)\\&-\eta\|x-\bar{x}\|_Q^2\psi^{(j)}+\\&\frac{\eta}{2}\exp\left(-\eta V^{(j)}(t,x)\right)\left(\frac{\partial V^{(j)}}{\partial x}\right)'\left(BR^{-1}B'-\eta\sigma\sigma'\right)\frac{\partial V^{(j)}}{\partial x}.\end{aligned}$$

Thus, under Assumption 1, we get

$$\begin{aligned}&-\frac{\partial\psi^{(j)}}{\partial t}=x'A'\frac{\partial\psi^{(j)}}{\partial x}+\frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2\psi^{(j)}}{\partial x^2}\sigma\right)-\eta\|x-\bar{x}\|_Q^2\psi^{(j)}\\&\psi^{(j)}(T,x)=\exp\left(-\eta\|x-p_j\|_M^2\right),\ \ \forall x\in\mathbb{R}^n. \qquad (34)\end{aligned}$$

Define $\psi(t,x)=\exp\left(-\eta V(t,x)\right)$ the transformation of $V(t,x)$ defined in (14). Hence, we have $\psi(t,x)=\sum_{j=1}^{l}\psi^{(j)}(t,x)g^{(j)}(t,x)$. Equation (34), Assumption 1 and the identity $\frac{\partial\psi^{(j)}}{\partial x}=-\eta\left(\Pi x+\beta^{(j)}\right)\psi^{(j)}$, where Π and $\beta^{(j)}$ are defined in (10), imply

$$\begin{aligned}&\frac{\partial\psi}{\partial t}+x'A'\frac{\partial\psi}{\partial x}+\frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2\psi}{\partial x^2}\sigma\right)-\eta\|x-\bar{x}\|_Q^2\psi\\&=\sum_{j=1}^{l}\left(\frac{\partial g^{(j)}}{\partial t}+\left(Ax-BR^{-1}B'\Pi x-BR^{-1}B'\beta^{(j)}\right)'\frac{\partial g^{(j)}}{\partial x}\right.\\&\left.+\frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 g^{(j)}}{\partial x^2}\sigma\right)\right)\psi^{(j)}.\end{aligned}$$

The process $x^{(j)}$ satisfies the SDE (9). Therefore, by Kolmogorov's backward equation [24, Section 5.B] applied to the conditional probability $g^{(j)}$,

$$\begin{aligned}&\frac{\partial g^{(j)}}{\partial t}+\left(Ax-BR^{-1}B'\Pi x-BR^{-1}B'\beta^{(j)}\right)'\frac{\partial g^{(j)}}{\partial x}\\&+\frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 g^{(j)}}{\partial x^2}\sigma\right)=0.\end{aligned}$$

with final condition $g^{(j)}(T,x)=1_{\mathcal{W}^{(j)}}(x)$. Hence ψ is the unique strong solution of

$$\frac{\partial\psi}{\partial t}+x'A'\frac{\partial\psi}{\partial x}+\frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2\psi}{\partial x^2}\sigma\right)-\eta\|x-\bar{x}\|_Q^2\psi=0. \quad (35)$$

By multiplying the right and left-hand sides of (35) by $\frac{1}{\eta}\exp(\eta V(t,x))$, $V(t,x)$ satisfies (6) in the strong sense. The uniqueness of the solution follows from the uniqueness of solutions to the uniform parabolic PDE (35) [24, Theorem 7.6].

Proof of Theorem 2

We have

$$\begin{aligned}u^*(t,x)&=-R^{-1}B'\frac{\partial V}{\partial x}\\&=\sum_{j=1}^{l}\frac{\psi^{(j)}(t,x)g^{(j)}(t,x)}{\sum_{k=1}^{l}\psi^{(k)}(t,x)g^{(k)}(t,x)}u^{(j)}(t,x)+\\&\frac{1}{\eta\sum_{k=1}^{l}\psi^{(k)}(t,x)g^{(k)}(t,x)}R^{-1}B'\sum_{j=1}^{l}\psi^{(j)}(t,x)\frac{\partial g^{(j)}}{\partial x}(t,x).\end{aligned}$$

In the following we show that the second summand is zero. By the change of variable $z=y-\alpha(T,t)x+\int_t^T\alpha(T,\tau)BR^{-1}B'\beta^{(j)}(\tau)d\tau$ in (12) and Leibniz integral rule, we have

$$\begin{aligned}&\frac{\partial g^{(j)}}{\partial x}(t,x)\\&=\frac{-\alpha(T,t)}{\sqrt{|2\pi\Sigma_t|}}\int_{\partial\mathcal{W}^{(j)}-\alpha(T,t)x+\int_t^T\alpha(T,\tau)BR^{-1}B'\beta^{(j)}(\tau)d\tau}\\&\exp\left(-\|z\|_{\Sigma_t^{-1}}^2\right)\vec{n}^{(j)}(z)ds(z)=\frac{-\alpha(T,t)}{\sqrt{|2\pi\Sigma_t|}}\int_{\partial\mathcal{W}^{(j)}}\exp\Bigg(\\&-\left\|y-\alpha(T,t)x+\int_t^T\alpha(T,\tau)BR^{-1}B'\beta^{(j)}(\tau)d\tau\right\|_{\Sigma_t^{-1}}^2\Bigg)\\&\times\vec{n}^{(j)}(y)ds(y),\end{aligned}$$

where α is defined in (13) and $\vec{n}^{(j)}(y)$ is the unit normal component of $\partial\mathcal{W}^{(j)}$ and its translation $\partial\mathcal{W}^{(j)}-\alpha(T,t)x+\int_t^T\alpha(T,\tau)BR^{-1}B'\beta^{(j)}(\tau)d\tau$. By solving for $\beta^{(j)}$ and $\delta^{(j)}$ in (10) and replacing the solutions in the expressions of the costs $V^{(j)}$ defined in (7) and in the derivatives $\frac{\partial g^{(j)}}{\partial x}$, one can show that under Assumptions 1 and 2,

$$\begin{aligned}&\sum_{j=1}^{l}\psi^{(j)}(t,x)\frac{\partial g^{(j)}}{\partial x}(t,x)=K_1(t,x)\sum_{j=1}^{l} \qquad (36)\\&\int_{\partial\mathcal{W}^{(j)}}\exp\left(K_2(t,x,y)+\eta\|y-p_j\|_M^2-\eta\|y\|_M^2\right)\vec{n}^{(j)}(y)ds(y).\end{aligned}$$

where K_1 and K_2 are functions that do not depend on p_j, $\forall j\in\{1,\ldots,l\}$. Note that $\partial\mathcal{W}^{(j)}=\cup_{i=1}^{k_j}O_i$, where the disjoint subsets (up to a subset of measure zero) $\{O_i\}_{i=1}^{k_j}$ are the common boundaries of $\mathcal{W}^{(j)}$ and the adjacent Voronoi cells. If O_i is the common boundary of $\mathcal{W}^{(j)}$ and some adjacent Voronoi Cell $\mathcal{W}^{(k)}$, then $\vec{n}^{(j)}(y)=-\vec{n}^{(k)}(y)$ for all $y\in O_i$. Moreover, by the definition of the Voronoi cells, $\|y-p_j\|_M=\|y-p_k\|_M$ for all $y\in O_i$. Therefore, the right-hand side of (36) is equal to zero. Thus, the optimal control u^* satisfies (16).

Next, we show that u^* is an admissible Markov policy, i.e. $u^*\in\mathcal{L}$ defined in (4). In view of (16), the function $\frac{\partial u^*}{\partial x}$ is continuous on $[0,T)\times\mathbb{R}^n$. Therefore, the local Lipschitz condition holds. Moreover, for all $(t,x) \in [0,T] \times \mathbb{R}^n$, we have

$$\|u^*(t,x)\| \leq \sum_{j=1}^{l} \left\|u^{(j)}(t,x)\right\| \leq \|R^{-1}B'\| \left(l\|\Pi\|_\infty\|x\| + \sum_{j=1}^{l}\|\beta^{(j)}\|_\infty \right). \quad (37)$$

Hence, the linear growth condition is satisfied and this proves the result. As a result, sufficient conditions are satisfied for the SDE defined in (5) and controlled by $u^*(t,x)$ to have unique strong solution denoted $x^*(.)$ [24, Section 5.2].

Finally, by the verification theorem [25, Theorem 4.3.1], we know that u^* is the unique optimal control law of (5) if it is the unique minimizer (up to a set of measure 0) of the Hamiltonian $H(x, \frac{\partial V}{\partial x}, u, t) = (Ax+Bu)'\frac{\partial V}{\partial x} + \|x-\bar{x}\|_Q^2 + \|u\|_R^2$, and if the cost-to-go $V(t,x)$ has a polynomial growth in x and satisfies (6). For the first condition, we have for Lebesgue $\times\,\mathbb{P}$-a.e $(t,\xi) \in [0,T]\times\Omega$ ($\mathbb{P}$ is the probability measure defined at the beginning of Section III),

$$u^*\left(t, x^*(t,\xi)\right) = -R^{-1}B'\frac{\partial V}{\partial x}\left(t, x^*(t,\xi)\right) = \operatorname*{argmin}_{u\in\mathbb{R}^n} H\left(x^*(t,\xi), \frac{\partial V}{\partial x}(t, x^*(t,\xi)), u, t\right).$$

In fact, the control law defined in (15) minimizes H except on the set $\{T\}\times\Omega$, which has a Lebesgue $\times\,\mathbb{P}$ measure zero. Next, in view of (37), we have for all $(t,x)\in[0,T)\times\mathbb{R}^n$

$$\|V(t,x)\| \leq \int_0^{\|x\|} \left\|\frac{\partial V}{\partial x}(t,y)\right\| dy \leq K_1(1+\|x\|^2),$$

for some $K_1>0$. Moreover, $\|V(T,x)\|\leq K_2(1+\|x\|^2)$, for some $K_2>0$. Hence, for all $(t,x)\in[0,T]\times\mathbb{R}^n$, $\|V(t,x)\|\leq K\left(1+\|x\|^2\right)$, for some $K>0$. Moreover, as established in Lemma 3, $V\in C^{1,2}\left([0,T)\times\mathbb{R}^n\right)\cap C([0,T]\times\mathbb{R}^n)$ satisfies the HJB equation (6). This proves the result.

APPENDIX B

This appendix includes the proofs of lemmas and theorems related to the existence of a solution to the mean field equations and the $\epsilon-$Nash property.

Proof of Theorem 3

First, we provide in the following lemma a stochastic maximum principle [27] for the "min-LQG" optimal control problem. Because of the non-smooth final cost, this result is derived using the relationship between dynamic programming and the stochastic maximum principle rather than the variational method used in [27].

Lemma 7: Under Assumptions 1 and 3, the processes $\left(q_s(t), \frac{\partial^2 V_s}{\partial x^2}(t, x_s^*(t)\right)$, $1\leq s\leq k$, with $q_s(t) = \frac{\partial V_s}{\partial x}(t, x_s^*(t))$, satisfy the following backward linear SDE (21).

Proof: As discussed below equation (13), $\Sigma_t \succ 0$ for all $t\in[0,T)$. Hence, in view of (12) and (14), the function V_s is in $C^\infty([0,T)\times\mathbb{R}^n)$, which means that $\frac{\partial V_s}{\partial x}(t,x)$ is smooth on $[0,T)\times\mathbb{R}^n$. By applying Itô's formula [24, Section 3.3.A] to $\frac{\partial V_s}{\partial x}(t, x_s^*(t))$, and by noting that V_s satisfies the HJB equation (6), we have

$$-dq_s(t) = \left(A_s'q_s(t) + Q_s(x_s^*(t)-\bar{x}(t))\right)dt - \frac{\partial^2 V_s}{\partial x^2}\left((t, x_s^*(t)\right)\sigma_s dw_s(t),$$

with $q_s(0) = \frac{\partial V_s}{\partial x}(0, x_s^*(0))$. It remains to show that $\mathbb{P}$-a.s

$$\lim_{t\to T}\frac{\partial V_s}{\partial x}(t, x_s^*(t)) = M_s\left(x_s^*(T) - \sum_{j=1}^{l} 1_{\mathcal{W}^{(j)}}(x_s^*(T))p_j\right). \quad (38)$$

By Theorem 2, we have on $[0,T)\times\mathbb{R}^n$

$$\frac{\partial V_s}{\partial x}(t,x) = \sum_{j=1}^{l}\frac{\exp\left(-\eta_s V_s^{(j)}(t,x)\right)g_s^{(j)}(t,x)}{\sum_{k=1}^{l}\exp\left(-\eta_s V_s^{(k)}(t,x)\right)g_s^{(k)}(t,x)}\frac{\partial V_s^{(j)}}{\partial x}.$$

Fix $j\in\{1,\dots,l\}$. By Lemma 5, we have on $\{x_s^*(T)\in \mathrm{Int}(\mathcal{W}^{(j)})\}$, $\lim_{t\to T} g_s^{(j)}(t, x_s^*(t)) = 1$ and $\lim_{t\to T} g_s^{(k)}(t, x_s^*(t)) = 0$, for all $k\neq j$. Hence, on $\{x_s^*(T)\in\mathrm{Int}(\mathcal{W}^{(j)})\}$, we have $\lim_{t\to T}\frac{\partial V_s}{\partial x}(t, x_s^*(t)) = \lim_{t\to T}\frac{\partial V_s^{(j)}}{\partial x}(t, x_s^*(t)) = M_s(x_s^*(T)-p_j)$. But, under Assumption 3, x_s^* is the solution of an SDE with non degenerate noise. Therefore, $\mathbb{P}\left(x_s^*(T)\in\partial\mathcal{W}^{(j)}\right) = 0$. Hence, (38) holds. ■

Remark 4: The backward SDE (21) is the adjoint equation [27] for the min-LQG optimal control problem.

We now prove point (i) of Theorem 3. By taking the expectations on the right and the left hand sides of (21) and the SDE in (19), and in view of $\sum_{s=1}^{k}\alpha_s\bar{x}_s(t) = \bar{x}$, we get the necessary condition. To prove the sufficient condition, we consider $(\bar{X}, \bar{x}, \bar{q})$ satisfying (23a)-(23c). We define $(\hat{x}_s, \hat{q}_s) = (\mathbb{E}x_s^*, \mathbb{E}q_s)$, where (x_s^*, q_s) are the s-type generic agent's optimal state and co-state when tracking $\bar{x}$. We define $e = (\hat{x}_1,\dots,\hat{x}_k) - \bar{X}$ and $\bar{q}_e = (\hat{q}_1,\dots,\hat{q}_k) - \bar{q}$. By taking expectations on the right and the left hand sides of (21) and the generic agent's dynamics, we obtain that

$$\begin{aligned} \frac{d}{dt}e(t) &= Ae(t) - BR^{-1}B'\bar{q}_e(t), & e(0) &= 0 \\ \frac{d}{dt}\bar{q}_e(t) &= -A'\bar{q}_e(t) + QLe(t), & \bar{q}_e(T) &= Me(T). \end{aligned} \quad (39)$$

Under Assumption 4, we define $q_e'(t) = \pi(t)e(t)$, where $\pi(t)$ is the unique solution of (25). We have $\frac{d(\bar{q}_e - q_e')}{dt} = -(A' - \pi(t)BR^{-1}B')(\bar{q}_e - q_e')$, with $(\bar{q}_e(T) - q_e'(T)) = 0$. Hence, $\bar{q}_e(t) = \pi(t)e(t)$. By replacing $\bar{q}_e(t) = \pi(t)e(t)$ in the forward equation in (39), we obtain that $e=0$. Hence, $\bar{x}$ satisfies (19).

To prove point (ii), we consider a path $\bar{x}$ satisfying (19), which by point (i) satisfies (23a)-(23c) and (22). Thus, a solution $(\bar{X}, \bar{q})$ of (23a) and (23b) exists. Under Assumption 4, using arguments similar to those used in the first point we obtain that this solution is unique. Moreover, one can check that $\bar{q} = \pi\bar{X} + \gamma$, where π is the unique solution of (25), and γ is the unique solution of $\dot{\gamma} = -(A - BR^{-1}B'\pi)'\gamma$ with $\gamma(T) = -M\Lambda\otimes I_n p$. By replacing, $\bar{q} = \pi\bar{X} + \gamma$ in (23a), we get that $\bar{x}$ is of the form (28). Next, by implementing this new form of $\bar{x}$ in the expression of (22) and by noting that Λ satisfies (24), Λ is a fixed point of F.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

Conversely, we consider Λ to be a fixed point of F, $\bar{X} = \big(R_1(t,0)\bar{X}(0) + R_2(t)\Lambda \otimes I_np\big)$ and $\bar{x} = P_1\bar{X}$. We define $\bar{q}(t) = -(BR^{-1}B')^{-1}(\frac{d}{dt}\bar{X}(t) - A\bar{X}(t))$. $(\bar{X},\bar{q})$ satisfies (23a)-(23b). We have $\Lambda_{sj} = F(\Lambda)_{sj} = \mathbb{P}(x^*_{s,\Lambda}(T) \in \mathcal{W}^{(j)})$, where $x^*_{s,\Lambda}$ is defined in (30). But $\bar{x}$ is of the form (31), hence $x^*_{s,\Lambda}$ is the unique strong solution of (22). Therefore, $\bar{x}$ satisfies (23a)-(23c), and by the first point, it satisfies (19). This proves the second point.

Next, to show the existence of a fixed point of F, it is sufficient to show that F is continuous, in which case Brouwer's fixed point theorem [42, Section V.9] ensures the existence of a fixed point. Equation (30) is a stochastic differential equation depending on the parameter Λ. By [43, Theorem 1] and the assumption that σ is invertible, the joint distribution of X^*_Λ and the Brownian motion W is weakly continuous in Λ. Consider a sequence of stochastic matrices $\{\Lambda_n\}_{n\geq 0}$ converging to the stochastic matrix Λ. The distribution of $X^*_{\Lambda_n}(T)$ converges weakly to the distribution of $X^*_\Lambda(T)$ Moreover, X^*_Λ is the solution of a non-degenerate SDE. Hence, $\mathcal{W}^{(j)}$, $j = 1,\dots,l$, is a continuity set of the distribution of X^*_Λ. Therefore, $\lim_n F(\Lambda_n)_{sj} = \lim_n \mathbb{P}(x^*_{s,\Lambda_n}(T) \in \mathcal{W}^{(j)}) = \mathbb{P}(x^*_{s,\Lambda}(T) \in \mathcal{W}^{(j)}) = F(\Lambda)_{sj}$, and so F is continuous.

Proof of Theorem 4

The idea of the proof is to show that the average state of the N agents when they optimally respond to a sustainable mean trajectory $\bar{x}$ converges to this trajectory in the L_2 norm (See Lemma 8 below). This result is subsequently used to prove that an individual cost in the finite population case converges to that in the infinite population case. This, combined with the fact that the MFG strategies form a Nash equilibrium in the infinite population case, implies the $\epsilon-$Nash property.

Lemma 8: We have

$$\mathbb{E}\int_0^T \|\bar{x}(t)-\frac{1}{N}\sum_{i=1}^N x^*_i(t)\|^2 dt = C\left(\frac{1}{N} + \max_{1\leq s\leq k}|\alpha_s - \alpha_s^N|^2\right),$$

where α and α_s are defined in Section III-B, $\bar{x}$ is a solution of the mean field equations (19), x^*_i is the optimal state of agent i when he optimally responds to $\bar{x}$, and $C > 0$ independent of N.

Proof: We have

$$\begin{aligned}
&\mathbb{E}\int_0^T \|\bar{x}(t) - \frac{1}{N}\sum_{i=1}^N x^*_i(t)\|^2 dt \\
&\leq 2\mathbb{E}\int_0^T \|\bar{x}(t) - \frac{1}{N}\sum_{i=1}^N \mathbb{E}x^*_i(t)\|^2 dt \\
&+ 2\mathbb{E}\int_0^T \|\frac{1}{N}\sum_{i=1}^N \mathbb{E}x^*_i(t) - \frac{1}{N}\sum_{i=1}^N x^*_i(t)\|^2 dt := I_1 + I_2.
\end{aligned}$$

Since the initial states are i.i.d and have finite second moment, the optimal states are independent and $\int_0^T \|x^*_i(t)\|^2 dt < \infty$. Hence, $I_2 \leq C_1/N$, where $C_1 > 0$ independent of N. We have

$$\begin{aligned}
I_1 &= \int_0^T \|\sum_{s=1}^k \alpha_s\bar{x}_s(t) - \sum_{s=1}^k \alpha_s^N\bar{x}_s(t)\|^2 dt \\
&\leq Tk^2 \sup_{1\leq s\leq k}\sup_{t\in[0,1]}\|\bar{x}_s(t)\|^2 \max_{1\leq s\leq k}|\alpha_s - \alpha_s^N|^2.
\end{aligned}$$

The term $\sup_{1\leq s\leq k}\sup_{t\in[0,1]}\|\bar{x}_s(t)\|^2$ is finite because $\bar{x}_s$, the $s-$th component of $\bar{X}$ in (26), is a continuous function of time. This proves the result. ■

Finally, we proceed with the proof of the $\epsilon-$Nash property. Fix i, and suppose that the strategy of agent i is $u_i \in \mathcal{U}_i^c$, while the strategies of the other agents are u^*_j, $j \neq i$, defined in (16) for a trjectory $\bar{x}$ solution of (19). Suppose that agent i profits by deviating from the mean field strategies, i.e. $J_i(u_i,u^*_{-i}) \leq J_i(u^*_i,u^*_{-i})$. This implies that u_i and the corresponding state x_i have finite L_2 norm $\mathbb{E}\int_0^T \|\|^2 dt$. We have,

$$\begin{aligned}
&J_i(u_i,u^*_{-i}) - J_i(u^*_i,u^*_{-i}) = J(u_i,\bar{x}) - J(u^*_i,\bar{x}) \\
&+ \mathbb{E}\int_0^T \|x_i(t) - \frac{1}{N}\sum_{j=1,j\neq i}^N x^*_j(t) - \frac{1}{N}x_i(t)\|^2_{Q_i} dt \\
&- \mathbb{E}\int_0^T \|x_i(t) - \bar{x}\|^2_{Q_i} dt + \mathbb{E}\int_0^T \|x^*_i(t) - \bar{x}\|^2_{Q_i} dt \\
&- \mathbb{E}\int_0^T \|x^*_i(t) - \frac{1}{N}\sum_{j=1}^N x^*_j(t)\|^2_{Q_i} dt \\
&:= I_1 - I_2 + I_3 - I_4 + I_5 - I_6.
\end{aligned}$$

Using the L_2-boundedness of x_i, x^*_i, Cauchy-Schwarz inequality [44] and Lemma 8, one can show that $I_3 - I_4$ and $I_5 - I_6$ are bounded by ϵ_N. Moreover, the optimality of u^*_i implies that $I_1 - I_2$ is positive. Hence, $J_i(u_i,u^*_{-i}) - J_i(u^*_i,u^*_{-i}) \geq -\epsilon_N$. This proves the result.

References

[1] F. Koppelman and V. Sathi, “Incorporating variance and covariance heterogeneity in the generalized nested logit model: an application to modeling long distance travel choice behavior,” *Transportation Research*, vol. 39, pp. 825–853, 2005.

[2] C. Bhat and J. Guo, “A mixed spatially correlated logit model: formulation and application to residential choice modeling,” *Transportation Research*, vol. 38, pp. 147–168, 2004.

[3] R. Nakajima, “Measuring peer effects on youth smoking behaviour,” *The Review of Economic Studies*, vol. 74, no. 3, pp. 897–935, 2007.

[4] D. Vail and M. Veloso, “Multi-robot dynamic role assignment and coordination through shared potential fields,” *Multi-robot systems*, pp. 87–98, 2003.

[5] A. Halász, M. A. Hsieh, S. Berman, and V. Kumar, “Dynamic redistribution of a swarm of robots among multiple sites,” in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2007, pp. 2320–2325.

[6] M. A. Hsieh, Á. Halász, S. Berman, and V. Kumar, “Biologically inspired redistribution of a swarm of robots among multiple sites,” *Swarm Intelligence*, vol. 2, no. 2-4, pp. 121–141, 2008.

[7] S. Merrill and B. Grofman, *A Unified Theory of Voting: Directional and Proximity Spatial Models*. Cambridge University Press, 1999.

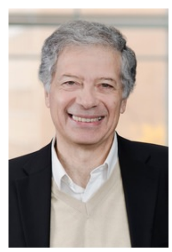
[8] G. B. Markus and P. E. Converse, “A dynamic simultaneous equation model of electoral choice,” *American Political Science Review*, vol. 73, no. 04, pp. 1055–1070, 1979.

[9] F. Fagnani and S. Zampieri, “Randomized consensus algorithms over large scale networks,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 634–649, 2008.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2941443, IEEE Transactions on Automatic Control

[10] D. McFadden, "Conditional logit analysis of qualitative choice behavior," in *Frontiers in Econometrics*, P. Zarembka, Ed. New York: New York: Academic Press, 1974, ch. 4, pp. 105–142.
[11] 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 *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2003, pp. 98–103.
[12] ——, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized epsilon-Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.
[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," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.
[14] J. M. Lasry and P. L. Lions, "Jeux à champ moyen. I–le cas stationnaire," *Comptes Rendus Mathématique*, vol. 343, no. 9, pp. 619–625, 2006.
[15] ——, "Jeux à champ moyen. II–horizon fini et contrôle optimal," *Comptes Rendus Mathématique*, vol. 343, no. 10, pp. 679–684, 2006.
[16] ——, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, pp. 229–260, 2007.
[17] J. Rust, "Structural estimation of markov decision processes," *Handbook of econometrics*, vol. 4, pp. 3081–3143, 1994.
[18] W. Brock and S. Durlauf, "Discrete choice with social interactions," *Review of Economic Studies*, pp. 147–168, 2001.
[19] R. Salhab, R. P. Malhamé, and J. Le Ny, "Consensus and disagreement in collective homing problems: A mean field games formulation," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, Dec 2014, pp. 916–921.
[20] ——, "A dynamic game model of collective choice in multi-agent systems," in *Proceedings of the 54th IEEE Conference on Decision and Control*, Dec 2015, pp. 4444–4449.
[21] R. Salhab, R. P. Malhamé, and J. Le Ny, "A dynamic game model of collective choice in multiagent systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 768–782, 2018.
[22] ——, "Collective stochastic discrete choice problems: A Min-LQG game formulation," *arXiv preprint arXiv:1981080*, 2017.
[23] M. Huang, P. E. Caines, and R. P. Malhamé, "Social optima in mean field LQG control: centralized and decentralized strategies," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1736–1751, 2012.
[24] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*. Springer Science & Business Media, 2012, vol. 113.
[25] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*. Springer Science & Business Media, 2006, vol. 25.
[26] L. Evans, *Partial Differential Equations*, ser. Graduate studies in mathematics. American Mathematical Society, 1998.
[27] J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*. Springer Science & Business Media, 1999, vol. 43.
[28] W. J. Rugh and W. J. Rugh, *Linear system theory*. prentice hall Upper Saddle River, NJ, 1996, vol. 2.
[29] D. Gomes, E. Pimentel, and V. Voskanyan, *Regularity Theory for Mean-Field Game Systems*, ser. SpringerBriefs in Mathematics. Springer, 2016.
[30] P. J. Olver, *Applications of Lie groups to differential equations*. Springer Science & Business Media, 2000, vol. 107.
[31] T. Liggett, *Interacting particle systems*. Springer Science & Business Media, 2012, vol. 276.
[32] L. Perko, *Differential equations and dynamical systems*. Springer Science & Business Media, 2013, vol. 7.
[33] B. D. Anderson and J. B. Moore, *Optimal control: linear quadratic methods*. Dover Publications, 2007.
[34] G. Freiling, "A survey of nonsymmetric Riccati equations," *Linear algebra and its applications*, vol. 351, pp. 243–270, 2002.
[35] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. Siam, 1999, vol. 23.
[36] R. Carmona and F. Delarue, "Probabilistic analysis of mean-field games," *SIAM Journal on Control and Optimization*, vol. 51, no. 4, pp. 2705–2734, 2013.
[37] Y. Achdou and I. Capuzzo-Dolcetta, "Mean field games: Numerical methods," *SIAM Journal on Numerical Analysis*, vol. 48, no. 3, pp. 1136–1162, 2010.
[38] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of computation*, pp. 577–593, 1965.
[39] L. Pichler, A. Masud, and L. Bergman, "Numerical solution of the Fokker–Planck equation by finite difference and finite element methods - A comparative study," in *Computational Methods in Stochastic Dynamics*. Springer, 2013, pp. 69–85.
[40] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, "Multi-agent decision-making dynamics inspired by honeybees," *IEEE Transactions on Control of Network Systems*, 2018.
[41] R. Durrett, *Probability: theory and examples*. Cambridge university press, 2010.
[42] J. B. Conway, *A Course in Functional Analysis*, ser. Graduate Texts in Mathematics. Springer-Verlag, 1985.
[43] A. V. Skorokhod, "Stochastic differential equations depending on a parameter," *Theory of Probability & Its Applications*, vol. 25, no. 4, pp. 659–666, 1981.
[44] W. Rudin, *Real and Complex Analysis*, 3rd ed. McGraw-Hill Inc., 1987.

Rabih Salhab received the Bachelors and Ph.D. degrees in Electrical Engineering from Ecole Supérieure d'Ingénieurs de Beyrouth (E.S.I.B), and the Ecole Polytechnique de Montreal, in 2008 and 2018 respectively. He is currently a Postdoctoral Associate at the MIT Sociotechnical Systems Research Center (SSRC). From 2018 to 2019, he was a Postdoctoral Fellow (IVADO) at HEC Montréal, Canada. From 2008 to 2013 he was an Electrical Engineer with Dar al Handasah Shair and Partners, Lebanon. His current research interests are in stochastic control, social learning, game theory and mean field games theory and its applications.



Roland Malhamé received the Bachelors, Masters and Ph.D. degrees in Electrical Engineering from the American University of Beirut, the University of Houston, and the Georgia Institute of Technology in 1976, 1978 and 1983 respectively.

After single year stays at University of Quebec , and CAE Electronics Ltd (Montreal), he joined in 1985 École Polytechnique de Montréal, where he is Professor of Electrical Engineering. In 1994, 2004, and 2012 he was on sabbatical leave respectively with LSS CNRS (France), École Centrale de Paris, and University of Rome Tor Vergata. His interest in statistical mechanics inspired approaches to the analysis and control of large scale systems has led him to contributions in the area of aggregate electric load modeling, and to the early developments of the theory of mean field games. His current research interests are in collective decentralized decision making schemes, and the development of mean field based control algorithms in the area of smart grids. From june 2005 to june 2011, he headed GERAD, the Group for Research on Decision Analysis. He is an Associate Editor of International Transactions on Operations Research.

Jerome Le Ny (S'05-M'09-SM'16) received the Engineering Degree from the École Polytechnique, France, in 2001, the M.Sc. degree in Electrical Engineering from the University of Michigan, Ann Arbor, in 2003, and the Ph.D. degree in Aeronautics and Astronautics from the Massachusetts Institute of Technology, Cambridge, in 2008. He is currently an Associate Professor with the Department of Electrical Engineering, Polytechnique Montreal, Canada. He is a member of GERAD, a multi-university research center on decision analysis. From 2008 to 2012 he was a Postdoctoral Researcher with the GRASP Laboratory at the University of Pennsylvania. His research interests include robust and stochastic control, mean-field control, privacy and security in large-scale networked control systems, dynamic resource allocation and active perception, with applications to autonomous multi-robot systems and intelligent infrastructures.