

<b>Titre:</b> Title:	A dynamic game model of collective choice in multi-agent systems
<b>Auteurs:</b> Authors:	Rabih Salhab, Roland P. Malhamé, Jérôme Le Ny
<b>Date:</b>	2018
<b>Type:</b>	Article de revue / Journal article
<b>Référence:</b> Citation:	Salhab, R., Malhame, R. P. & Le Ny, J. (2018). A dynamic game model of collective choice in multi-agent systems. <i>IEEE Transactions on Automatic Control</i> , 63(3), p. 768-782. doi:10.1109/tac.2017.2723956



### Document en libre accès dans PolyPublie

Open Access document in PolyPublie

<b>URL de PolyPublie:</b> PolyPublie URL:	<a href="https://publications.polymtl.ca/2865/">https://publications.polymtl.ca/2865/</a>
<b>Version:</b>	Version finale avant publication / Accepted version Révisé par les pairs / Refereed
<b>Conditions d'utilisation:</b> Terms of Use:	Tous droits réservés / All rights reserved



### Document publié chez l'éditeur officiel

Document issued by the official publisher

<b>Titre de la revue:</b> Journal Title:	IEEE Transactions on Automatic Control
<b>Maison d'édition:</b> Publisher:	IEEE
<b>URL officiel:</b> Official URL:	<a href="https://doi.org/10.1109/tac.2017.2723956">https://doi.org/10.1109/tac.2017.2723956</a>
<b>Mention légale:</b> Legal notice:	© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

**Ce fichier a été téléchargé à partir de PolyPublie,  
le dépôt institutionnel de Polytechnique Montréal**

This file has been downloaded from PolyPublie, the  
institutional repository of Polytechnique Montréal

<http://publications.polymtl.ca>

# A Dynamic Game Model of Collective Choice in Multi-Agent Systems

Rabih Salhab, Roland P. Malhamé and Jerome Le Ny

**Abstract**—Inspired by successful biological collective decision mechanisms such as honey bees searching for a new colony or the collective navigation of fish schools, we consider a scenario where a large number of agents engaged in a dynamic game have to make a choice among a finite set of different potential target destinations. Each individual both influences and is influenced by the group’s decision, as represented by the mean trajectory of all agents. The model can be interpreted as a stylized version of opinion crystallization in an election for example. In the most general formulation, agents’ biases are dictated by a combination of initial position, individual dynamics parameters and a priori individual preference. Agents are assumed linear and coupled through a modified form of quadratic cost, whereby the terminal cost captures the discrete choice component of the problem. Following the mean field games methodology, we identify sufficient conditions under which allocations of destination choices over agents lead to self replication of the overall mean trajectory under the best response by the agents. Importantly, we establish that when the number of agents increases sufficiently, (i) the best response strategies to the self replicating mean trajectories qualify as epsilon-Nash equilibria of the population game; (ii) these epsilon-Nash strategies can be computed solely based on the knowledge of the joint probability distribution of the initial conditions, dynamics parameters and destination preferences, now viewed as random variables. Our results are illustrated through numerical simulations.

**Index Terms**—Mean Field Games, Collective Choice, Discrete Choice Models, Multi-Agent Systems, Optimal Control.

## I. INTRODUCTION

Collective decision making is a common phenomenon in social structures ranging from animal populations [3], [4] to human societies [5]. Examples include honey bees searching for a new colony [6], [7], the navigation of fish schools [8], [9], or quorum sensing [10]. Collective decisions involve dynamic “microscopic-macroscopic” or “individual-social” interactions. On the one hand, individual choices are socially influenced, that is, influenced by the behavior of the group. On the other hand, the collective behavior itself results from aggregating individual choices.

In elections for example, an interplay between individual interests and collective opinion swings leads to the crystallization of final decisions [5], [11]. Our model may be an abstract representation of this process where: i) individual opinion dynamics are described in a state-space form [11];

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.malhamé, jerome.le-ny}@polymtl.ca.

Preliminary versions of this paper appeared at CDC 2014 and CDC 2015 [1], [2].

ii) changing one’s opinion requires an effort but deviation from the majority’s opinion involves a discomfort; and iii) a choice must be made before a finite deadline. The classical voter model [12] describes the evolution of opinions in an election. It considers a group of agents choosing between two alternatives. At each instant, the probability that an individual switches from one alternative to the other depends on its current choice, the others’ states, as well as the communication graph. In this paper, we consider a game theoretic approach to a limited randomness version of this problem, whereby agents, given enough deterministic or probabilistic information on the initial spatial distribution and preferences of other agents, crystallize at the outset what their final choices will be, through *anticipation* of the group behavior over the control horizon. Movement in opinion space towards a final choice requires costly efforts from voters. At the same time, they experience discomfort whenever their individual state differs from the mean population state.

“Homing” optimal control problems, first introduced by Whittle and Gait in [13] and studied later in [14]–[17] for example, are concerned with a single agent trying to reach one of multiple predefined final states. Here we consider a similar fundamental issue but in a multi-agent setting. A large number of agents initially spread out in  $\mathbb{R}^n$  need to move within a finite time horizon to one of multiple possible home or target destinations. They must do so while trying to remain tightly grouped and expending as little control effort as possible. Our goal is to model situations in which the choice made by each agent regarding which destination to reach both influences and depends on the behavior of the population. For example, when honey bees determine their next site to establish a colony they must make a choice between different alternatives based on the information provided by scouts, who are themselves part of the group. Even though certain colonies can be easier to reach and are more attractive for some bees, following the majority is still a priority to enhance the foraging ability [6], [7]. In animal collective navigation [18], [19], discrete choices must be made regarding the route to take, but at the same time, staying with the group offers better protection against predators [8].

Similarly, consider a situation as in [20]–[22] where a collection of robots is exploring an unknown terrain and should choose between multiple potential sites of interest to visit. The robots can possibly split, but each subgroup should remain sufficiently large to carry out collective tasks of interest [23]–[26]. Our framework allows modeling such a situation and provides parameters indicating for each robot the attractiveness of the sites and the cost of deviating from the

population's centroid. Moreover, we look for a coordination strategy that requires a limited amount of communication, thus increasing the robustness of our solution against intermittent loss of connectivity within the group for example. We do so by relying on the law of large numbers and the Mean-Field Games (MFG) methodology, where the agents only need to learn the initial distribution of the group [27] (for example, through a consensus-like algorithm), in order to compute an optimal decentralized control strategy. In practice in a finite group the agents would also have to communicate their states periodically in order to compensate for the prediction error due to the fact that the MFG methodology assumes an infinite population.

A related topic in economics is discrete choice models where an agent makes a choice between multiple alternatives such as mode of transportation [28], entry and withdrawal from the labor market, residential location [29], or a physician [30]. In many circumstances, these individual choices are influenced by the so called "Peer Effect", "Neighborhood Effect" or "Social Effect". In particular, Brock and Durlauf [31] use an approach similar to Mean Field Games (MFG) [32], [33] and inspired by statistical mechanics to study a static binary discrete choice model with a large number of agents, which takes into account the effect of the agents' interdependence on the individual choices. In their model, the individual choices are influenced by the mean of the other agents' choices, while for an infinite size population, the impact of an isolated individual choice on this mean is negligible. The authors show that in an infinite size rational population, each agent can predict this mean as the result of a fixed point calculation, and makes a decentralized choice based upon its prediction. Moreover, multiple anticipated means may exist. Our analysis leads to similar insights for a dynamic non-cooperative multiple choice game including situations where the agents have limited information about the dynamics of other agents.

## II. PROBLEM STATEMENT AND CONTRIBUTIONS

In this section, we formulate our problem, state our main contributions and provide an outline for the rest of the paper.

### A. Deterministic Initial Conditions

We consider a dynamic non-cooperative game involving  $N$  players with identical linear dynamics

$$\dot{x}_i = Ax_i + Bu_i \quad \forall i \in \{1, \dots, N\}, \quad (1)$$

where  $x_i \in \mathbb{R}^n$  is the state of agent  $i$  and  $u_i \in \mathbb{R}^m$  its control input. Player  $i$  is associated with an individual cost functional

$$J_i(u_i, \bar{x}, x_i^0) = \int_0^T \left\{ \frac{q}{2} \|x_i - \bar{x}\|^2 + \frac{r}{2} \|u_i\|^2 \right\} dt + \frac{M}{2} \min_{j \in \{1, \dots, l\}} \left( \|x_i(T) - p_j\|^2 \right), \quad (2)$$

where  $\bar{x}(t) \triangleq 1/N \sum_{i=1}^N x_i(t)$ ,  $p_j \in \mathbb{R}^n$  (for  $j \in \{1, \dots, l\}$ ) are the destination points,  $q, r$  are positive constants and  $M$  is a large positive number. The running cost requires the

agents to develop as little effort as possible while moving and to stay grouped around the mean of the population  $\bar{x}$ . Moreover, each agent should reach before final time  $T$  one of the destinations  $p_j, j \in \{1, \dots, l\}$ . Otherwise, it is strongly penalized by the terminal cost. Hence, the overall individual cost captures the problem faced by each agent of deciding between a finite set of alternatives, while trying to remain close to the mean population trajectory. It is sometimes convenient to write the costs in a game theoretic form, i.e.  $J_i(u_i, u_{-i})$ , where  $u_{-i} = (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$ . We seek  $\epsilon$ -Nash strategies, i.e. such that an agent can benefit at most  $\epsilon$  through unilateral deviant behavior, with  $\epsilon$  going to zero as  $N$  goes to infinity [34]. We assume that each agent can observe only its own state and the initial states of the other agents.

*Definition 1:* Consider  $N$  players, a set of strategy profiles  $S = S_1 \times \dots \times S_N$  and for each player  $i$ , a payoff function  $J_i(u_i, u_{-i}), \forall (u_i, u_{-i}) \in S$ . A strategy profile  $(u_i^*, u_{-i}^*) \in S$  is called an  $\epsilon$ -Nash equilibrium with respect to the costs  $J_i$  if there exists an  $\epsilon > 0$  such that for any fixed  $i \in \{1, \dots, N\}$  and for all  $u_i \in S_i$ , we have  $J_i(u_i, u_{-i}^*) \geq J_i(u_i^*, u_{-i}^*) - \epsilon$ .

Inspired by the framework of MFG theory [23], [32]–[36] discussed in Section II-C below, we develop in this paper a class of decentralized strategies satisfying a certain fixed point requirement. In particular, under the assumption of a continuum of agents, the problem of computing an agent's best response to a given macroscopic behavior of the population turns out to be an optimal control problem. Hence, the terms "best response" and "optimal control law / strategy" of the agents are used interchangeably in the paper. The fixed point requirement originates from the fact that collectively, the agents' best responses must reproduce the assumed macroscopic behavior. Identification of the strategies requires only that an agent knows its own state and the initial states of the other agents. As we later show, when the number of agents  $N$  is sufficiently large, these fixed point based strategies achieve their meaning as  $\epsilon$ -Nash equilibria.

### B. Random Initial Conditions

As  $N$  goes to infinity, it is also convenient to think of the initial states as realizations of random variables resulting from a common probability distribution function in a collection of independent experiments. Agent  $i$  is then associated with the following adequately modified cost:

$$J_i(u_i, \bar{x}, x_i^0) = \mathbb{E} \left( \int_0^T \left\{ \frac{q}{2} \|x_i - \bar{x}\|^2 + \frac{r}{2} \|u_i\|^2 \right\} dt + \frac{M}{2} \min_{j \in \{1, \dots, l\}} \left( \|x_i(T) - p_j\|^2 \right) \middle| x_i^0 \right). \quad (3)$$

We establish that an agent needs only to know its own state and the common probability distribution of initial states to construct one of the decentralized fixed point based strategies alluded to earlier. In this case, the only randomness lies in the agents' initial conditions, and the control strategies, while expressed as state feedback laws, correspond in effect to open loop policies [37].

### C. The MFG Approach and our Contributions

The MFG approach is concerned with a class of dynamic non-cooperative games involving a large number of players where the individual strategies are considerably affected by the mass behavior, while the influence of isolated individual actions on the group is negligible. Linear Quadratic Gaussian (LQG) MFG formulations were developed in [32], [34], [35], while the general nonlinear stochastic framework was considered in [33], [38]–[40]. To compute a Nash equilibrium, one has to produce in general  $N$  fixed point trajectories  $\frac{1}{N} \sum_{i \neq j} x_i$ ,  $j \in \{1, \dots, N\}$  (the average of the population without player  $j$ ). As the size  $N$  of the population increases to infinity, these trajectories become indistinguishable. The MFG approach posits at the outset an infinite population to which one can ascribe a deterministic although initially unknown macroscopic behavior. Hence, one starts by assuming that the mean field contributed term  $\bar{x}$  in the cost (2) or (3) is given, denoted  $\hat{x}$ . The cost functions being now decoupled, each agent optimally tracks  $\hat{x}$  (i.e. computes its best response to  $\hat{x}$ ). The resulting control laws (best responses) are decentralized. This analysis of the tracking problem is presented in Section III. With the agents implementing the resulting decentralized strategies, a new candidate average path is obtained by computing the corresponding mean population trajectory. Indeed, and it is a fundamental argument in MFG analysis, asymptotically as the population grows, the posited tracked path is an acceptable candidate only if it is reproduced as the mean of the agents when they optimally respond to it. Thus, we look for candidate trajectories that are fixed points of the tracked path to tracked path map defined above. In Section IV, these fixed points are studied for the deterministic initial conditions with a finite population, and an explicit expression is obtained by assuming that each agent knows the exact initial states of all other agents. The alternative probabilistic description of the agents' initial states is explored in Section V. In Section VI, we further generalize the problem formulation to include initial preferences towards the target destinations. Moreover, we consider that the agents have nonuniform dynamics and that each agent has limited information about the other agents dynamic parameters in the form of a statistical distribution over the matrices  $A$  and  $B$ . Section VII shows that the decentralized strategies developed when tracking the fixed point mean trajectories constitute  $\epsilon$ -Nash equilibria in all the cases considered above, with  $\epsilon$  going to zero as  $N$  goes to infinity. In Section VIII, we provide some numerical simulation results, while Section IX presents our conclusions.

The main contributions of the paper include the following:

- i. We introduce a novel class of linear quadratic non-convex games aimed at characterizing solutions of collective discrete choice problems in a variety of applications.
- ii. We show that an agent with knowledge about the dynamic parameters and initial preferences of the other agents can make its choice by observing only the initial conditions of the players (in the case of deterministic initial conditions), or by knowing their initial probability distribution (in the case of random initial conditions).
- iii. In the uniform dynamics case, we characterize the way the

- population splits between the destination points. In fact, we construct a finite dimensional map, which we call the “Choice Distribution Map” (CDM), such that the probability distribution of the choices between the alternatives is a fixed point of this map. In the probabilistic version of the problem, this corresponds to a fixed point vector equation of dimension  $l$  (total number of available destinations). Thus the computation of the corresponding strategies can be considerably simplified relative to that of the finite  $N$  case, which requires comparing the performance of  $l^N$  possible deployments of the  $N$  agents over  $l$  destinations.
- iv. We prove the existence of a decentralized  $\epsilon$ -Nash equilibrium, and in the uniform dynamics case, we develop a method to compute it. In essence, this indicates that our simplified and decentralized infinite population based control policies induce Nash equilibria asymptotically as  $N$  tends to infinity.

We further detail here some more technical aspects of our contribution. Although we rely on the MFG methodology in order to analyze the behavior of many agents choosing one of the available destinations, our model is not standard with respect to the LQG MFG literature. Specifically, our cost is non-convex and non-smooth (the final cost is a minimum of  $l$  quadratic functions), in order to capture the combinatorial aspect of the discrete decision-making problem. Hence, the existence proofs for a fixed point rely here on topological fixed point theorems rather than a contraction argument as in [32]. One of the main contributions of this paper is also to show that in case of a uniform population, the infinite dimensional MFG fixed point problem [33], [38] has a finite dimensional version that can be characterized via Brouwer's fixed point theorem [41]. For a nonuniform population, the existence of a fixed point mean trajectory relies on an abstract fixed point theorem, namely Schauder's fixed point theorem [41]. In both cases, to solve the MFG equation system, one needs to know the initial probability distribution of the players, whereas in the standard LQG MFG problems, it is sufficient to know the initial mean to anticipate the macroscopic behavior. Thus, in a nutshell, the theoretical tools needed to address this new formulation are thoroughly different. Further highlighting the differences between the two problems, the control laws when extending the current formulation to the stochastic dynamics case are entirely different from the LQG case [42].

Preliminary versions of our results appeared in the conference papers [1], [2]. Here we provide a unified discussion of our collective choice model for the deterministic and stochastic scenarios, as well as more extensive results. Many of the proofs were omitted from the conference papers due to space limitations and can be found here. The simulation section is also expanded with respect to [1], [2] and provides additional insight on the role of the different parameters in the model.

### D. Notation

The following notation is used throughout the paper. We denote by  $C(X, Y)$  the set of continuous functions from a normed vector space  $X$  to  $Y \subset \mathbb{R}^k$  with the standard supremum norm  $\|\cdot\|_\infty$ . We fix a generic probability space

$(\Omega, \mathcal{F}, \mathbb{P})$  and denote by  $\mathbb{P}(A)$  the probability of an event  $A$ , and by  $\mathbb{E}(X)$  the expectation of a random variable  $X$ . The indicator function of a subset  $X$  is denoted by  $1_X$  and its interior by  $\overset{\circ}{X}$ . We denote by  $|X|$  the size of a finite set  $X$ . The transpose of a matrix  $M$  is denoted by  $M^T$ . We denote by  $I_k$  the identity  $k \times k$  matrix. The subscript  $i$  is used to index entities related to the agents, while the subscripts  $j$  and  $k$  are used to index entities related to the home destinations. We denote by  $[x]_m$  the  $m$ -th component of a vector  $x$ .

### III. TRACKING PROBLEM AND BASINS OF ATTRACTION

In this section, we compute the agents' best responses (tracking problem) to the mean field contributed term  $\bar{x}$ . We establish the existence of an  $\bar{x}$  dependent partition of the state space into  $l$  basins of attraction, each associated with a distinct destination for all agents initially situated in it.

#### A. Tracking Problem

Following the MFG approach, we assume the trajectory  $\bar{x}(t)$  in (2) and (3) to be given for now and call it  $\hat{x}(t)$ . The cost functions (2) and (3) can be written as the minimum of  $l$  linear quadratic tracking cost functions, each corresponding to a destination point:

$$J_i(u_i, \hat{x}, x_i^0) = \min_{j \in \{1, \dots, l\}} J_{ij}(u_i, \hat{x}, x_i^0), \quad (4)$$

where

$$J_{ij}(u_i, \hat{x}, x_i^0) = \int_0^T \left\{ \frac{q}{2} \|x_i - \hat{x}\|^2 + \frac{r}{2} \|u_i\|^2 \right\} dt + \frac{M}{2} \|x_i(T) - p_j\|^2. \quad (5)$$

Moreover,  $\inf_{u_i(\cdot)} J_i(u_i, \hat{x}, x_i^0) = \min_{j \in \{1, \dots, l\}} \left( \inf_{u_i(\cdot)} J_{ij}(u_i, \hat{x}, x_i^0) \right)$ . Assuming a full (local) state feedback, the optimal control law for (4)  $u_i^*$  is the optimal control law of the less costly linear quadratic tracking problem, that is  $u_i^* = u_{ij}^*$  if  $J_{ij}(u_{ij}^*, \hat{x}, x_i^0) = \min_{k \in \{1, \dots, l\}} J_{ik}(u_{ik}^*, \hat{x}, x_i^0)$ , where  $u_{ik}^*$  is the optimal solution of the simple linear quadratic tracking problem with cost function  $J_{ik}$ .

In the following, we partition the space  $\mathbb{R}^n$  into  $l$  regions (basins of attraction), each corresponding to a distinct destination point, such that if an agent is initially in one of these basins, the linear tracking problem associated with the corresponding destination point is the least costly. We recall the optimal control laws [43],  $u_{ik}^*(t) = -\frac{1}{r} B^T (\Gamma(t)x_i + \beta_k(t))$ ,  $\forall k \in \{1, \dots, l\}$ , with the corresponding optimal costs

$$J_{ik}^*(\hat{x}, x_i^0) = \frac{1}{2} (x_i^0)^T \Gamma(0) x_i^0 + \beta_k(0)^T x_i^0 + \delta_k(0),$$

where  $\Gamma$ ,  $\beta_k$  and  $\delta_k$  are respectively matrix-, vector-, and real-valued functions satisfying the following backward propagating differential equations:

$$\dot{\Gamma} - \frac{1}{r} \Gamma B B^T \Gamma + \Gamma A + A^T \Gamma + q I_n = 0 \quad (6a)$$

$$\dot{\beta}_k = \left( \frac{1}{r} \Gamma B B^T - A^T \right) \beta_k + q \hat{x} \quad (6b)$$

$$\dot{\delta}_k = \frac{1}{2r} (\beta_k)^T B B^T \beta_k - \frac{1}{2} q \hat{x}^T \hat{x}, \quad (6c)$$

with the final conditions  $\Gamma(T) = M I_n$ ,  $\beta_k(T) = -M p_k$ ,  $\delta_k(T) = \frac{1}{2} M p_k^T p_k$ . We define the basins of attraction

$$D_j(\hat{x}) = \{x \in \mathbb{R}^n \mid J_{ij}^*(\hat{x}, x) \leq J_{ik}^*(\hat{x}, x), \forall k \in \{1, \dots, l\}\} \quad (7)$$

for  $j \in \{1, \dots, l\}$ . If an agent  $i$  is initially in  $D_j(\hat{x})$ , then the smallest optimal (simple) cost is  $J_{ij}^*$ , and player  $i$  goes towards the corresponding destination point  $p_j$ .

*Assumption 1:* Conventionally, we assume that if  $x_i^0 \in \bigcap_{m=1}^k D_{j_m}(\hat{x})$ , for some  $j_1 < \dots < j_k$ , then the player  $i$  goes towards  $p_{j_1}$ . Under Assumptions 2 and 5 (defined below in Sections V and VI), this convention does not affect the analysis in case of random initial conditions.

We summarize the above analysis in the following lemma.

*Lemma 1:* Under Assumption 1, the tracking problem (4) has a unique optimal control law

$$u_i^*(t) = -\frac{1}{r} B^T (\Gamma(t)x_i + \beta_j(t)) \quad \text{if } x_i^0 \in D_j(\hat{x}), \quad (8)$$

where  $\Gamma$ ,  $\beta_j$ ,  $\delta_j$  are the unique solutions of (6a)-(6c).

The optimal control laws (8) depend on the local state  $x_i$  and on the tracked path  $\hat{x}(t)$  via  $D_j$  and  $\beta_j$ . As mentioned above, each agent should reach one of the predefined destinations. We show in the next lemma that for any horizon length  $T$ ,  $M$  can be made large enough that each agent reaches an arbitrarily small neighborhood of some destination point by applying the control law (8). The result is proved for tracked paths  $\hat{x}(t)$  that are uniformly bounded with respect to  $M$ , a property that is shown to hold later in Lemma 9 for the desired tracked paths (fixed point tracked paths).

*Lemma 2:* Suppose that the pair  $(A, B)$  is controllable and for each  $M > 0$ , the agents are optimally tracking a path  $\hat{x}_M(t)$ . We suppose that the family  $\hat{x}_M(t)$  is uniformly bounded with respect to  $M$  for the norm  $\left( \int_0^T \|\cdot\|^2 dt \right)^{\frac{1}{2}}$ . Then, for any  $\epsilon > 0$ , there exists  $M_0 > 0$  such that for all  $M > M_0$ , each agent is at time  $T$  in a ball of radius  $\epsilon$  and centered at one of the  $p_j$ 's, for  $j \in \{1, \dots, l\}$ .

*Proof:* See Appendix A. ■

Given any continuous path  $\hat{x}(t)$ , there exist  $l$  basins of attraction where all the agents initially in  $D_j(\hat{x})$  prefer going towards  $p_j$ . Therefore, the mean of the population is highly dependent on the structure of  $D_j(\hat{x})$ . In the next paragraph, we provide an explicit form of these basins.

#### B. Basins of Attraction

We start by giving an explicit solution of (6b) and (6c). Let  $\Pi(t) = \frac{1}{r} \Gamma(t) B B^T - A^T$  and  $\Phi(\cdot, \eta)$ , for  $\eta \in \mathbb{R}$ , be the unique solution of

$$\frac{d\Phi(t, \eta)}{dt} = \Pi(t)\Phi(t, \eta) \quad \Phi(\eta, \eta) = I_n, \quad (9)$$

and  $\Psi(\eta_1, \eta_2, \eta_3, \eta_4) = \Phi(\eta_1, \eta_2)^T B B^T \Phi(\eta_3, \eta_4)$ . Two main properties of the state transition matrix  $\Phi$  are used in this paper, namely the matrix  $\Phi(\eta_1, \eta_2)$  has an inverse  $\Phi(\eta_2, \eta_1)$  and the state transition matrix  $\Phi(\eta_1, \eta_2)$  of  $-\Pi^T$  (i.e. solution

of (9), where  $\Pi$  in the right hand side of (9) is replaced by  $-\Pi^T$ ) is equal to  $\Phi(\eta_2, \eta_1)^T$ . For more details about the properties of the state transition matrix, one can refer to [44]. The unique solution of (6b)-(6c) is

$$\begin{aligned} \beta_k(t) &= -M\Phi(t, T)p_k + q \int_T^t \Phi(t, \sigma)\hat{x}(\sigma) d\sigma \\ \delta_k(t) &= \frac{1}{2}Mp_k^T p_k - \frac{q}{2} \int_T^t \hat{x}(\sigma)^T \hat{x}(\sigma) d\sigma \\ &+ \frac{M^2}{2r}p_k^T \int_T^t \Psi(\eta, T, \eta, T) d\eta p_k \\ &- \frac{Mq}{r}p_k^T \int_T^t \int_T^\eta \Psi(\eta, T, \eta, \sigma)\hat{x}(\sigma) d\sigma d\eta \\ &+ \frac{q^2}{2r} \int_T^t \int_T^\eta \int_T^\eta \hat{x}(\sigma)^T \Psi(\eta, \sigma, \eta, \tau)\hat{x}(\tau) d\tau d\sigma d\eta. \end{aligned} \quad (10)$$

By replacing (10) in the expression of  $D_j(\hat{x})$ , (7) can be written

$$D_j(\hat{x}) = \{x \in \mathbb{R}^n | \beta_{jk}^T x \leq \delta_{jk} + \alpha_{jk}(\hat{x}), \forall k \in \{1, \dots, l\}\}, \quad (11)$$

where

$$\begin{aligned} \beta_{jk} &= M\Phi(0, T)(p_k - p_j) \\ \delta_{jk} &= \frac{1}{2}Mp_k^T p_k - \frac{1}{2}Mp_j^T p_j \\ &+ \frac{M^2}{2r}p_k^T \int_T^0 \Psi(\eta, T, \eta, T) d\eta p_k \\ &- \frac{M^2}{2r}p_j^T \int_T^0 \Psi(\eta, T, \eta, T) d\eta p_j \\ \alpha_{jk}(\hat{x}) &= \frac{Mq}{r}(p_j - p_k)^T \int_T^0 \int_T^\eta \Psi(\eta, T, \eta, \sigma)\hat{x}(\sigma) d\sigma d\eta. \end{aligned} \quad (12)$$

Hence, we see from (11) that the regions  $D_j(\hat{x})$  for a given  $\hat{x}$  are separated by hyperplanes.

#### IV. FIXED POINT - DETERMINISTIC INITIAL CONDITIONS

Having computed the best responses to an arbitrary  $\hat{x}$ , we now seek a continuous path  $\hat{x}(t)$  that is sustainable, in the sense that it can be replicated by the average of the agents under their best responses to it. We start by analyzing a finite size population where the initial state of each agent is known to all agents. We exhibit a one to one map between the sustainable paths and the fixed points of a CDM (Choice distribution map), itself a *finite dimensional* operator, the fixed points of which characterize the set of fractions of agents going to each one of the possible destinations. An algorithm is proposed to compute these fixed points and the associated agent trajectories.

We start our search for the desired path  $\hat{x}(t)$  by computing the mean  $\bar{x}(t)$  when tracking any continuous path  $\hat{x}(t)$ . The dynamics of the mean when tracking  $\hat{x} \in C([0, T], \mathbb{R}^n)$  satisfies

$$\begin{aligned} \dot{\bar{x}} &= -\Pi^T \bar{x} - \frac{q}{r}BB^T \int_T^t \Phi(t, \sigma)\hat{x}(\sigma) d\sigma \\ &+ \frac{M}{r}BB^T \Phi(t, T)p_{\lambda(\hat{x})}, \end{aligned} \quad (13)$$

where  $\Phi$  is defined in (9),  $\bar{x}(0) = \frac{1}{N} \sum_{i=1}^N x_i^0 \triangleq \bar{x}_0$ ,  $p_{\lambda(\hat{x})} = \sum_{j=1}^l \frac{\lambda_j(\hat{x})}{N} p_j$  and  $\lambda_j(\hat{x})$  is the number of agents initially in  $D_j(\hat{x})$ , which therefore pick  $p_j$  as a destination. We obtain (13) by substituting (10) in (8) and the resulting control law in (1) to subsequently compute  $\bar{x} = 1/N \sum_{i=1}^N x_i$  and its derivative. Thus, the mean of the population  $\bar{x}$  when tracking any continuous path  $\hat{x}$  is the image of  $\hat{x}$  by a composite map  $G = G_2 \circ G_1$ , where

$$\begin{aligned} G_1 : C([0, T], \mathbb{R}^n) &\longrightarrow C([0, T], \mathbb{R}^n) \times \mathbb{N}^l \quad (14) \\ \hat{x} &\longrightarrow (\hat{x}, (\lambda_1(\hat{x}), \dots, \lambda_l(\hat{x}))) \end{aligned}$$

$$G_2 : C([0, T], \mathbb{R}^n) \times \mathbb{N}^l \longrightarrow C([0, T], \mathbb{R}^n) \quad (15)$$

such that  $\bar{x} = G_2(\hat{x}, (\lambda_1, \dots, \lambda_l))$  is the unique solution of (13) in which  $\lambda_j(\hat{x})$  is equal to an arbitrary  $\lambda_j$ ,  $j \in \{1, \dots, l\}$ .

The desired path describing the mean trajectory is a fixed point of  $G$ . In the following, we construct a one to one map between the fixed points of  $G$  and the fixed points of a finite dimensional operator  $F$  describing the way the population splits between the destination points. We define the following quantities:

$$R_1(t) = \Phi_P(t, 0) \quad (16)$$

$$R_2(t) = \frac{M}{r} \int_0^t \Phi_P(t, \sigma)BB^T \Phi_P(T, \sigma)^T d\sigma \quad (17)$$

$$\theta_{jk} = \frac{Mq}{r}(p_j^T - p_k^T) \int_T^0 \int_T^\eta \Psi(\eta, T, \eta, \sigma)R_1(\sigma) d\sigma d\eta \quad (18)$$

$$\xi_{jk} = \frac{Mq}{r}(p_j^T - p_k^T) \int_T^0 \int_T^\eta \Psi(\eta, T, \eta, \sigma)R_2(\sigma) d\sigma d\eta \quad (19)$$

where  $P$  and  $\Phi_P(t, \eta)$  are the unique solutions of

$$\dot{P} = -PA - A^T P + \frac{1}{r}PBB^T P, \quad P(T) = MI_n \quad (20)$$

$$\dot{\Phi}_P(t, \eta) = (A - \frac{1}{r}BB^T P)\Phi_P(t, \eta), \quad \Phi_P(\eta, \eta) = I_n.$$

We define the CDM  $F$  from  $\{0, \dots, N\}^l$  into itself, such that for all  $\lambda \in \{0, \dots, N\}^l$ ,

$$F(\lambda) = (|\{x_i^0 | x_i^0 \in H_1^\lambda\}|, \dots, |\{x_i^0 | x_i^0 \in H_l^\lambda\}|), \quad (21)$$

where  $H_j^\lambda = D_j(R_1 \bar{x}_0 + R_2 p_\lambda) = \{x \in \mathbb{R}^n | \beta_{jk}^T x \leq \delta_{jk} + \theta_{jk} \bar{x}_0 + \xi_{jk} p_\lambda \forall k \in \{1, \dots, l\}\}$ . The functions (16)-(17) are defined to compute the family of paths in which lies the set of fixed points. The regions  $H_j^\lambda$  are the basins of attraction associated with this family and the function  $F$  counts the number of initial conditions in each  $H_j^\lambda$ . We now state the main result of this section.

*Theorem 3:* The path  $\hat{x}$  is a fixed point of  $G$ , defined in (13), if and only if it has the following form

$$\hat{x}(t) = R_1(t)\bar{x}_0 + R_2(t)p_\lambda, \quad (22)$$

where  $\lambda$  is a fixed point of  $F$ .

*Proof:* See Appendix A. ■

Theorem 3 states that there exists a one to one map between the fixed points of  $G$  (sustainable macroscopic behaviors) and the fixed points of the CDM  $F$ . As a result, one can compute all the sustainable macroscopic behaviors and anticipate the

corresponding distributions of the choices between the alternatives. More precisely, for each  $\lambda = (\lambda_1, \dots, \lambda_l) \in \{0, \dots, N\}^l$  satisfying  $\sum_{i=1}^l \lambda_i = N$ , one can define the  $l$  regions  $H_j^\lambda$  and count the numbers  $\eta_j$  of initial positions inside each region. If  $\eta_j = \lambda_j$ , for all  $j \in \{1, \dots, l\}$ , then  $\lambda$  is a fixed point of  $F$ . The map  $F$  may have multiple fixed points. Hence, an a priori agreement on how to choose  $\lambda$  should exist. For example, although non-cooperative, the agents may anticipate that their majority will look for the most socially favorable Nash equilibrium if many exist and  $N$  is large. This  $\lambda$  corresponds to minimizing the total cost  $\frac{1}{N} \sum_{i=1}^N J_i(u_i^*(x_i, \hat{x}), \hat{x}, x_i^0) = \frac{1}{N} \sum_{i=1}^N \min_{k \in \{1, \dots, l\}} \left\{ \frac{1}{2} (x_i^0)^T \Gamma(0) x_i^0 + \beta_k(0)^T x_i^0 + \delta_k(0) \right\}$ , which is also computable by just knowing the exact initial conditions of all the agents. Once the agents agree on a  $\lambda$ , they start tracking the corresponding fixed point defined by (22). The fixed point vector  $\lambda$  describes the way the population splits between the destination points. In fact,  $\lambda_j$ ,  $j \in \{1, \dots, l\}$ , is the number of agents that go towards  $p_j$ . When  $N$  is large however, this algorithm is costly in terms of number of counting and verification operations. In the next section, we consider the limiting case of a large population with random initial conditions.

In the binary choice case, the quantities  $\beta_{12}^T x_i^0$ ,  $i \in \{1, \dots, N\}$ , can be used to compare the ‘‘attractiveness’’ of  $p_2$  w.r.t.  $p_1$  for the different players. In fact, if going to  $p_1$  has lower cost than going to  $p_2$  for player  $i$  and  $\beta_{12}^T x_j^0 \leq \beta_{12}^T x_i^0$ , then  $p_1$  is also more attractive than  $p_2$  for player  $j$ . Hence, the agents’ indices can be reordered in such a way that the agents that go towards  $p_1$  have indices lower than those going towards  $p_2$ , as follows

$$\beta_{12}^T x_1^0 \leq \beta_{12}^T x_2^0 \leq \dots \leq \beta_{12}^T x_N^0. \quad (23)$$

We then derive in the following theorem a necessary and sufficient condition for the existence of a fixed point path.

*Theorem 4:* For  $l = 2$ , the following statements hold:

- 1)  $\hat{x}$  is a fixed point of  $G$  if and only if there exists a separating  $\alpha \in \{0, \dots, N\}$  such that:

$$\begin{aligned} \beta_{12}^T x_\alpha^0 - \delta_{12} - \theta_{12} \bar{x}_0 - \xi_{12} p_2 &\leq \frac{\alpha}{N} \xi_{12} (p_1 - p_2) \\ &< \beta_{12}^T x_{\alpha+1}^0 - \delta_{12} - \theta_{12} \bar{x}_0 - \xi_{12} p_2. \end{aligned} \quad (24)$$

$$\text{For } \alpha = 0, \quad 0 < \beta_{12}^T x_1^0 - \delta_{12} - \theta_{12} \bar{x}_0 - \xi_{12} p_2, \quad (25)$$

$$\text{For } \alpha = N, \quad \beta_{12}^T x_N^0 - \delta_{12} - \theta_{12} \bar{x}_0 - \xi_{12} p_2 \leq 0, \quad (26)$$

where  $\beta_{12}$ ,  $\delta_{12}$ ,  $\theta_{12}$  and  $\xi_{12}$  are defined in (12), (18) and (19). In this case,  $\alpha$  is the number of agents that go towards  $p_1$ .

- 2) For  $\xi_{12}(p_1 - p_2) \geq 0$ , there exists  $\alpha \in \{0, \dots, N\}$  satisfying (24), (25) or (26).
- 3) For  $\xi_{12}(p_1 - p_2) < 0$ , there exists at most one  $\alpha \in \{0, \dots, N\}$  satisfying (24), (25) or (26). Moreover, there exist some initial distributions for which no such  $\alpha$  exists.

*Proof:* See Appendix A. ■

If the agents are indexed in ascending order of attractiveness of  $p_2$  with respect to  $p_1$  as in (23), then an  $\alpha$  satisfying (24)-(26) separates the population into two groups, the agents that

go towards  $p_1$  (with indices less or equal to  $\alpha$ ), and those that go towards  $p_2$  (with indices greater than  $\alpha$ ).

*Remark 1:* For the scalar case ( $n = 1$ ),  $\Phi$  and  $\Phi_P$  are real exponential functions, which implies  $\xi_{12}(p_1 - p_2) \geq 0$ .

## V. FIXED POINT - RANDOM INITIAL CONDITIONS

In this section, we assume that the agents’ initial conditions  $x_i^0$  are random and i.i.d. on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with distribution  $P_0$  on  $\mathbb{R}^n$ . We assume an infinite size population. As in the deterministic case, we construct a one to one map between the fixed point paths and the fixed points of a finite dimensional CDM. The latter characterizes the distribution of the choices between the alternatives. Moreover, we prove the existence of a fixed point and propose several methods to compute it.

We start our search for a fixed point path by considering  $\bar{x}$  in (3) given and call it  $\hat{x}$ . By Lemma 1, there exist  $l$  regions  $D_j(\hat{x})$  such that the agents initially in  $D_j(\hat{x})$  select the control law (8) when tracking  $\hat{x}$ . By substituting (10) in (8) and the resulting control law in (1), we show that the mean trajectory  $\mathbb{E}(x_i)$  of a generic agent satisfies:

$$\begin{aligned} \mathbb{E}(x_i) = G_s(\hat{x}) \triangleq &\Phi(0, t)^T \mu_0 + \frac{M}{r} \int_0^t \Psi(\sigma, t, \sigma, T) p_{\lambda(\hat{x})} d\sigma \\ &- \frac{q}{r} \int_0^t \int_T^\sigma \Psi(\sigma, t, \sigma, \tau) \hat{x}(\tau) d\tau d\sigma, \end{aligned} \quad (27)$$

where  $\Psi$  was defined below (9),  $\mathbb{E}x_i^0 \triangleq \mu_0$ ,  $p_{\lambda(\hat{x})} = \sum_{j=1}^l \lambda_j(\hat{x}) p_j$  and  $\lambda_j(\hat{x}) = P_0(D_j(\hat{x}))$ .  $G_s$  and its fixed points, if any, depend only on the initial statistical distribution of the agents.

We define the set  $\Delta_l = \{(\lambda_1, \dots, \lambda_l) \in [0, 1]^l \mid \sum_{j=1}^l \lambda_j = 1\}$  and the CDM  $F_s$  from  $\Delta_l$  into itself such that

$$F_s(\lambda_1, \dots, \lambda_l) = [P_0(\bar{H}_1^\lambda), \dots, P_0(\bar{H}_l^\lambda)], \quad (28)$$

with

$$\bar{H}_i^\lambda = \{x \in \mathbb{R}^n \mid \beta_{ij}^T x \leq \delta_{ij} + \theta_{ij} \mu_0 + \xi_{ij} p_\lambda, \forall j \in \{1, \dots, l\}\} \quad (29)$$

and  $p_\lambda = \sum_{k=1}^l \lambda_k p_k$ , where  $\beta_{kj}$ ,  $\delta_{kj}$ ,  $\theta_{kj}$  and  $\xi_{kj}$  are defined in (12), (18) and (19).

Theorem 5 below, proved in Appendix B, shows the existence of a fixed point of  $F_s$  using Brouwer’s fixed point theorem. The latter requires continuity of  $F_s$ , which is guaranteed by the following assumption.

*Assumption 2:* We assume that  $P_0$  is such that the  $P_0$ -measure of hyperplanes is zero.

Assumption 2 is satisfied when  $P_0$  is absolutely continuous with respect to the Lebesgue measure for example. We now state the main result of this section.

*Theorem 5:* Under Assumption 2, the following statements hold:

- (i)  $\hat{x}$  is a fixed point of  $G_s$  if and only if there exists  $\lambda = (\lambda_1, \dots, \lambda_l)$  in  $\Delta_l$  such that

$$F_s(\lambda) = \lambda, \quad (30)$$

for  $\hat{x}(t) = R_1(t) \mu_0 + R_2(t) p_\lambda$ , where  $R_1$  and  $R_2$  are defined in (16) and (17).

- (ii)  $F_s$  has at least one fixed point (equivalently  $G_s$  has at least one fixed point).
- (iii) For  $l = 2$ , if  $\xi_{12}(p_1 - p_2) \leq 0$ , then  $G_s$  has a unique fixed point.

In Theorem 5, (i) shows that computing the anticipated macroscopic behaviors is equivalent to computing all the vectors  $\lambda$  satisfying (30) under the corresponding constraint on  $\hat{x}$ . To compute a  $\lambda$  satisfying (30), each agent is assumed to know the initial statistical distribution of the agents. As in the deterministic case, multiple  $\lambda$ 's may exist. Hence, an a priori agreement on how to choose  $\lambda$  should exist. In that respect, the agents may implicitly assume that collectively they will opt for the  $\lambda$  (assuming it is unique!) that minimizes the total expected population cost  $\mathbb{E}J_i(u_i^*(x_i, \hat{x}), \hat{x}, x_i^0) = \mathbb{E} \min_{k \in \{1, \dots, l\}} \left\{ \frac{1}{2}(x_i^0)^T \Gamma(0)x_i^0 + \beta_k(0)^T x_i^0 + \delta_k(0) \right\}$  ( $\Gamma$ ,  $\beta_k$  and  $\delta_k$  are defined in (6a)-(6c)), which can be evaluated if the agents know the initial statistical distribution of the population.

While in the deterministic case a fixed point of the CDM  $F$  determines the number of players that go to each destination point, in the stochastic case a fixed point of  $F_s$  is the vector of probabilities that an agent chooses each of the alternatives. The CDM  $F$  and  $F_s$  defined respectively in the deterministic and stochastic cases have similar structures. In fact, in the deterministic case, if the sequence  $\{x_i^0\}_{i=1}^N$  of initial conditions is interpreted as a random variable on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with distribution  $P_0(\mathcal{A}) = 1/N \sum_{i=1}^N 1_{\{x_i^0 \in \mathcal{A}\}}$ , for all (Borel) measurable sets  $\mathcal{A}$ , then  $F(\lambda) = NF_s(\lambda/N)$ .

The fixed points of the CDM  $F_s$  characterize the game in terms of the number of what will be characterized in Section VII as  $\epsilon$ -Nash equilibria, whether there is a consensus or disagreement and the distribution of the choices between the alternatives. In Subsection V-A, we investigate the question of computing these fixed points, while Subsection V-B treats the problem of uniqueness and multiplicity of these fixed points in the Gaussian binary choice case.

#### A. Computation of The Fixed Points

The map  $F_s$  is not necessarily a contraction. Hence, it is sometimes impossible to compute its fixed points by the simple iterative method  $\lambda_{k+1} = F_s(\lambda_k)$ .

1) *Binary Choice Case:* We give two simple methods to compute a fixed point of  $F_s$  in the binary choice case. The first method is applicable if  $\xi_{12}(p_1 - p_2) > 0$ . We define in  $[0, 1]$  a sequence  $\alpha_k$  such that  $\alpha_0$  is an arbitrary number in  $[0, 1]$  and

$$\lambda_{k+1} = (\alpha_{k+1}, 1 - \alpha_{k+1}) = F_s(\alpha_k, 1 - \alpha_k) = F_s(\lambda_k). \quad (31)$$

The sequence  $\lambda_k$  converges to a fixed point of  $F_s$ . In fact, given that  $\xi_{12}(p_1 - p_2) > 0$ ,  $\left[ F_s(t, 1 - t) \right]_1$  increases with  $t$ . We show by induction that  $\alpha_k$  is monotone. But  $\alpha_k \in [0, 1]$ , therefore,  $\alpha_k$  converges to some limit  $\alpha$ . By the continuity of  $F_s$ ,  $(\alpha, 1 - \alpha)$  satisfies (30). Since in this case  $F_s$  may have multiple fixed points, the  $\lambda = (\alpha, 1 - \alpha)$  obtained using this approach depends on the initial value  $\lambda_0 = (\alpha_0, 1 - \alpha_0)$ . Moreover, the sequence  $\bar{x}_k = R_1(t)\mu_0 + R_2(t)p_{\lambda_k}$  converges to a fixed point of  $G_s$ . The second method is applicable if

$\xi_{12}(p_1 - p_2) \leq 0$ . In this case  $\left[ F_s(\alpha, 1 - \alpha) \right]_1 - \alpha$  decreases with  $\alpha$ . Hence, one can compute the unique zero of this function by the bisection method.

2) *General Case:* In general ( $l \geq 2$ ),  $F_s$  is a vector of probabilities of some regions delimited by hyperplanes. Although a fixed point could be computed using Newton's method, this is computationally expensive as it requires the values of the inverse of the Jacobian matrix at the root estimates. Alternatively, one can compute a fixed point of  $F_s$  using a quasi Newton method such as Broyden's method [45] (see Section VIII). Using this method, the inverse of the Jacobian can be estimated recursively provided that  $F_s$  is continuously differentiable; this will be the case if the initial probability distribution has a continuous probability density function.

#### B. Gaussian Binary Choice Case

We have shown in Theorem 5 that for the binary choice case ( $l = 2$ ), if  $\xi_{12}(p_1 - p_2) < 0$ , then  $G_s$  defined in (27) has a unique fixed point. We now prove that for the binary choice case and Gaussian initial distribution  $\mathcal{N}(\mu_0, \Sigma_0)$ , irrespective of the sign of  $\xi_{12}(p_1 - p_2)$ ,  $G_s$  has a unique fixed point provided that the initial spread of the agents is "sufficient". For any  $n \times n$  matrix  $\Sigma$  such that  $(\beta_{12}^T \Sigma \beta_{12}) < (\xi_{12}(p_1 - p_2))^2 / 2\pi$ , we define

$$\begin{aligned} a(\Sigma) &= \delta_{12} + \xi_{12}p_2 - c(\Sigma) \sqrt{2(\beta_{12}^T \Sigma \beta_{12})} \\ b(\Sigma) &= \delta_{12} + \xi_{12}p_1 + c(\Sigma) \sqrt{2(\beta_{12}^T \Sigma \beta_{12})} \\ c(\Sigma) &= \sqrt{\log \xi_{12}(p_1 - p_2) - \frac{1}{2} \log 2\pi(\beta_{12}^T \Sigma \beta_{12})} \\ S(\Sigma) &= \{ \mu \in \mathbb{R}^n, (\beta_{12}^T - \theta_{12})\mu \in (a(\Sigma), b(\Sigma)) \}, \end{aligned} \quad (32)$$

where  $\beta_{12}$ ,  $\delta_{12}$ ,  $\theta_{12}$  and  $\xi_{12}$  are defined in (12), (18) and (19).

*Theorem 6:*  $G_s$  has a unique fixed point if one of the following conditions is satisfied

- 1)  $\beta_{12}^T \Sigma_0 \beta_{12} \geq (\xi_{12}(p_1 - p_2))^2 / 2\pi$ .
- 2)  $\beta_{12}^T \Sigma_0 \beta_{12} < (\xi_{12}(p_1 - p_2))^2 / 2\pi$  and  $\mu_0 \notin S(\Sigma_0)$ .

*Proof:* See Appendix B. ■

Theorem 6 states that in the Gaussian binary choice case, if the initial distribution of the agents has enough spread, then the agents make their choices in a unique way. On the other hand, if the uncertainty in their initial positions is low enough and the mean of population is inside the region  $S(\Sigma_0)$  (a region delimited by two parallel hyperplanes), then the agents can anticipate the collective behavior in multiple ways.

## VI. NONUNIFORM POPULATION WITH INITIAL PREFERENCES

Hitherto, the agents' initial affinities towards different potential targets are dictated only by their initial positions in space. In this section, the model is further generalized by considering that in addition to their initial positions, the agents are affected by their a priori opinion. When modeling smoking decision in schools for example [46], this could represent a teenager's tendency towards "Smoking" or "Not



Smoking”, which is the result of some endogenous factors such as parental pressure, financial condition, health, etc. When modeling elections, this would reflect personal preferences that transcend party lines. Moreover, we assume in this section that the agents have nonuniform dynamics.

We consider  $N$  agents with nonuniform dynamics

$$\dot{x}_i = A_i x_i + B_i u_i \quad i \in \{1, \dots, N\}, \quad (33)$$

with random initial states as in Section V. Player  $i$  is associated with the following individual cost:

$$J_i(u_i, \bar{x}, x_i^0) = \mathbb{E} \left( \int_0^T \left\{ \frac{q}{2} \|x_i - \bar{x}\|^2 + \frac{r}{2} \|u_i\|^2 \right\} dt + \min_{j \in \{1, \dots, l\}} \left( \frac{M_{ij}}{2} \|x_i(T) - p_j\|^2 \right) \middle| x_i^0 \right). \quad (34)$$

In the costs (34), a small  $M_{ij}$  relative to  $M_{ik}$ ,  $k \neq j$ , reflects an a priori affinity of agent  $i$  towards the destination  $p_j$ .

As  $N$  tends to infinity, it is convenient to represent the limiting sequence of  $(\theta_i)_{i \in \{1, \dots, N\}} = ((A_i, B_i, M_{i1}, \dots, M_{il}))_{i \in \{1, \dots, N\}}$  by a random vector  $\theta$ . We assume that  $\theta$  is in a compact set  $\Theta$ . Let us denote the empirical measure of the sequence  $\theta_i$  as  $P_\theta^N(\mathcal{A}) = 1/N \sum_{i=1}^N 1_{\{\theta_i \in \mathcal{A}\}}$  for all (Borel) measurable sets  $\mathcal{A}$ . We assume that  $P_\theta^N$  has a weak limit  $P_\theta$ , that is, for all  $\phi$  continuous,  $\lim_{N \rightarrow \infty} \int_\Theta \phi(x) dP_\theta^N(x) = \int_\Theta \phi(x) dP_\theta(x)$ . For further discussions about this assumption, one can refer to [47]. We assume that the initial states  $x_i^0$  and  $\theta$  are independent, and that an agent  $i$  knows its initial position  $x_i^0$ , its parameters  $\theta_i$ , as well as the distributions  $P_0$  and  $P_\theta$ . We develop the following analysis for a generic agent with an initial position  $x^0$  and parameters  $\theta$ . Assuming an infinite size population, we start by tracking  $\hat{x}(t)$ , a posited deterministic although initially unknown continuous path. We can then show that, under the convention in Assumption 1, this tracking problem is associated with a unique optimal control law

$$u^*(t) = -\frac{1}{r} (B^\theta)^T (\Gamma_j^\theta(t)x + \beta_j^\theta(t)) \quad \text{if } x^0 \in D_j^\theta(\hat{x}), \quad (35)$$

where  $\Gamma_j^\theta$ ,  $\beta_j^\theta$ ,  $\delta_j^\theta$  are the unique solutions of

$$\dot{\Gamma}_j^\theta - \frac{1}{r} \Gamma_j^\theta B^\theta (B^\theta)^T \Gamma_j^\theta + \Gamma_j^\theta A^\theta + (A^\theta)^T \Gamma_j^\theta + qI_n = 0 \quad (36a)$$

$$\dot{\beta}_j^\theta = \left( \frac{1}{r} \Gamma_j^\theta B^\theta (B^\theta)^T - (A^\theta)^T \right) \beta_j^\theta + q\hat{x} \quad (36b)$$

$$\dot{\delta}_j^\theta = \frac{1}{2r} (\beta_j^\theta)^T B^\theta (B^\theta)^T \beta_j^\theta - \frac{1}{2} q \hat{x}^T \hat{x}, \quad (36c)$$

with the final conditions  $\Gamma_j^\theta(T) = M_j^\theta I_n$ ,  $\beta_j^\theta(T) = -M_j^\theta p_j$ ,  $\delta_j^\theta(T) = \frac{1}{2} M_j^\theta p_j^T p_j$ . The definition of the basins of attraction becomes

$$D_j^\theta(\hat{x}) = \left\{ x \in \mathbb{R}^n \text{ such that } \frac{1}{2} x^T \Gamma_{jk}^\theta x + x^T \beta_{jk}^\theta(\hat{x}) + \delta_{jk}^\theta(\hat{x}) \leq 0, \forall k \in \{1, \dots, l\} \right\}, \quad (37)$$

where  $\Gamma_{jk}^\theta = \Gamma_j^\theta(0) - \Gamma_k^\theta(0)$ ,  $\beta_{jk}^\theta(\hat{x}) = \beta_j^\theta(0) - \beta_k^\theta(0)$  and  $\delta_{jk}^\theta(\hat{x}) = \delta_j^\theta(0) - \delta_k^\theta(0)$ . In this case, the solutions

of the Riccati equations (36a) depend on both the initial preference vector  $M^\theta$  and the destination points. Hence, the basins of attraction are now regions delimited by quadric surfaces in  $\mathbb{R}^n$  instead of hyperplanes. This fact complicates the structure of the operator that maps the tracked path to the mean. The existence proof for a fixed point relies now on an abstract Banach space version of Brouwer’s fixed point theorem, namely Schauder’s fixed point theorem [41]. We define  $\Psi_j^\theta(\eta_1, \eta_2, \eta_3, \eta_4) = \Phi_j^\theta(\eta_1, \eta_2)^T B^\theta (B^\theta)^T \Phi_j^\theta(\eta_3, \eta_4)$ , where  $\Pi_j^\theta(t) = \frac{1}{r} \Gamma_j^\theta(t) B^\theta (B^\theta)^T - (A^\theta)^T$ , and  $\Phi_j^\theta$  is defined as in (9), where  $\Pi$  is replaced by  $\Pi_j^\theta$ . The state trajectory of the generic agent, on  $\{x^0 \in D_j^\theta(\hat{x})\}$ , is then equal to

$$x^{0\theta}(t) = \Phi_j^\theta(0, t)^T x^0 + \frac{M_j^\theta}{r} \int_0^t \Psi_j^\theta(\sigma, t, \sigma, T) p_j d\sigma - \frac{q}{r} \int_0^t \int_T^\sigma \Psi_j^\theta(\sigma, t, \sigma, \tau) \hat{x}(\tau) d\tau d\sigma, \quad (38)$$

*Assumption 3:* We assume that  $\mathbb{E} \|x^0\|^2 < \infty$ .

The functions defined by (36a), (36b) and (36c) are continuous with respect to  $\theta$ , which belongs to a compact set. Moreover,  $\theta$  and  $x^0$  are assumed to be independent. Thus, under Assumption 3, the mean of the infinite size population can be computed using Fubini-Tonelli’s theorem [48] as follows:

$$\mathbb{E}(x^{0\theta}(t)) = G_p(\hat{x}) \triangleq \sum_{j=1}^l \int_\Theta \int_{\mathbb{R}^n} 1_{D_j^\theta(\hat{x})}(x^0) \left\{ \Phi_j^\theta(0, t)^T x^0 + \frac{M_j^\theta}{r} \int_0^t \Psi_j^\theta(\sigma, t, \sigma, T) p_j d\sigma - \frac{q}{r} \int_0^t \int_T^\sigma \Psi_j^\theta(\sigma, t, \sigma, \tau) \hat{x}(\tau) d\tau d\sigma \right\} dP_0 dP_\theta. \quad (39)$$

In the next theorem, we show that  $G_p$  has a fixed point. We define

$$k_1 = \mathbb{E} \|x^0\| \times \left( \sum_{j=1}^l \max_{(\theta, t) \in \Theta \times [0, T]} \|\Phi_j^\theta(0, t)\| \right) \\ k_2 = \sum_{j=1}^l \max_{(\theta, t) \in \Theta \times [0, T]} \left\| \frac{M_j^\theta}{r} \int_0^t \Psi_j^\theta(\sigma, t, \sigma, T) p_j d\sigma \right\| \\ k_3 = \frac{q}{r} \sum_{j=1}^l \max_{(\theta, t, \sigma, \tau) \in \Theta \times [0, T]^3} \|\Psi_j^\theta(\sigma, t, \sigma, \tau)\|. \quad (40)$$

Since  $\Theta$  and  $[0, T]$  are compact and  $\Phi_j^\theta$  is continuous with respect to time and parameter  $\theta$ , then  $k_1$ ,  $k_2$  and  $k_3$  are well defined. Theorem 7 below, proved in Appendix B, establishes the existence of a fixed point of  $G_p$  using Schauder’s fixed point theorem. The latter requires boundedness of  $G_p$  on bounded subsets of its domain and continuity of  $G_p$ , which are guaranteed by the following two assumptions respectively.

*Assumption 4:* We assume that  $\sqrt{\max(k_1 + k_2, k_3)T} < \pi/2$ .

Noting that the left hand side of the inequality tends to zero as  $T$  goes to zero, Assumption 4 can be satisfied for a short time horizon  $T$  for example.

*Assumption 5:* We assume that  $P_0$  is such that the  $P_0$ -measure of quadric surfaces is zero.

Assumption 5, similar to Assumption 2, is satisfied when  $P_0$  is absolutely continuous with respect to the Lebesgue measure for example.

*Theorem 7:* Under Assumptions 3, 4 and 5,  $G_p$  has a fixed point.

Note that if  $T$  goes to zero, the costs become decoupled, and each agent will choose the “closest” destination in the minimum energy sense. It is then expected that a Nash equilibrium exists in this case. Assumption 4 gives an upper bound on the time horizon  $T$  under which we can prove that such an equilibrium continues to exist.

## VII. APPROXIMATE NASH EQUILIBRIUM

In the three cases above, deterministic, random initial conditions and random initial conditions with non-uniform dynamics and initial preferences, we defined three maps  $G$ ,  $G_s$  and  $G_p$  respectively (equations (13), (27) and (39)). Depending on the structure of the game, each player can anticipate the macroscopic behavior of the limiting population by computing a fixed point  $\hat{x}$  of  $G$ ,  $G_s$  or  $G_p$ , and compute its best response  $u_i^*(x_i, \hat{x})$  to  $\hat{x}$  as defined in (8), (35). When considering the finite population, the next theorem establishes the importance of such decentralized strategies in that they lead to an  $\epsilon$ -Nash equilibrium with respect to the costs (2), (3) and (34). This equilibrium makes the group’s behavior robust in the face of potential selfish behaviors as unilateral deviations from the associated control policies are guaranteed to yield negligible cost reductions when  $N$  increases sufficiently.

*Theorem 8:* Under Assumption 3, the decentralized strategies  $u_i^*$ ,  $i \in \{1, \dots, N\}$ , defined in (8) and (35) for a fixed point path  $\hat{x}$ , constitute an  $\epsilon_N$ -Nash equilibrium with respect to the costs  $J_i(u_i, u_{-i})$ , where  $\epsilon_N$  goes to zero as  $N$  increases to infinity.

*Proof:* See Appendix B. ■

## VIII. SIMULATION RESULTS

To illustrate the collective decision-making mechanism, we consider a group of agents (robots) moving on the real line according to a second order system,  $\ddot{x}_i = -3x_i - \dot{x}_i + u_i$ , where  $x_i$  is the position of robot  $i$ . They should move from their initial conditions towards the position 10, and arrive at the speed of 1, 4 or 10. Thus, in the state space (position, speed), the potential destination points are  $p_1 = (10, 10)$ ,  $p_2 = (10, 4)$  or  $p_3 = (10, 1)$ . We draw  $N = 300$  initial conditions from the Gaussian distribution  $P_0 \triangleq \mathcal{N}(0, I_2)$ . We simulate two cases. In the first one, each agent knows the exact initial states of the other agents and anticipates the mean of the population accordingly. Following the counting and verification operations described at the end of Section IV, we find that  $F$  defined in (21) has a fixed point  $\lambda = (91, 209, 0)$ . By implementing the control laws corresponding to this  $\lambda$ , 91 agents (30.33% of the agents) go towards  $p_1$  and 209 (69.67% of the agents) towards  $p_2$  (see Fig.1). Moreover, the actual average replicates the anticipated mean as shown in this figure.

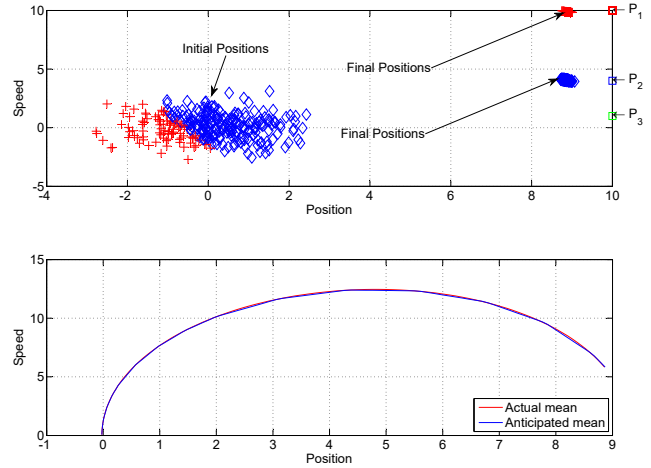


Fig. 1. Collective choice (deterministic initial conditions) with  $\lambda = (91, 209, 0)$ . Red Crosses = agents that go towards  $p_1$ . Blue diamonds = agents that go towards  $p_2$ .

In the second case, the agents know only the initial distribution  $P_0$  of the agents. Then, Broyden’s method converges to  $\lambda = (0.3066, 0.6921, 0.0013)$  satisfying (30). Accordingly, 30.66% of the agents go towards  $p_1$ , 69.21% towards  $p_2$  and the rest towards  $p_3$  (see Fig.2). The actual average and the anticipated mean are approximately the same.

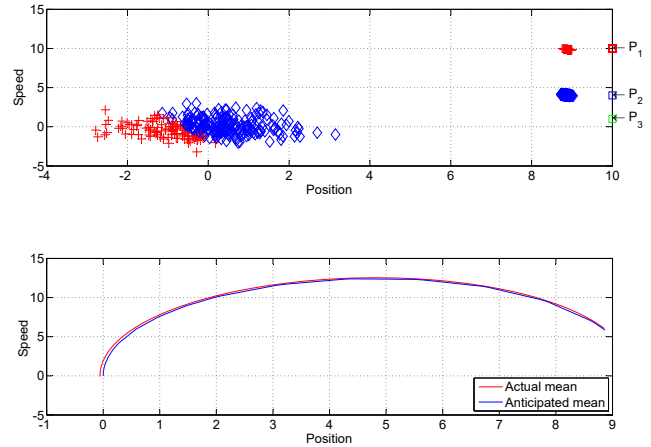


Fig. 2. Collective choice (random initial conditions) with  $\lambda = (0.3066, 0.6921, 0.0013)$ . Red Crosses = agents that go towards  $p_1$ . Blue diamonds = agents that go towards  $p_2$ .

To illustrate the social effect on the individual choices (see Fig. 3), we consider the same initial conditions. Without social effect ( $q = 0$ ),  $(0.4146, 0.5843, 0.0011)$  satisfies (30). In this case, the majority goes towards  $p_2$ . As the social effect increases to  $q = 5$ , some of the agents that went towards  $p_1$  in the absence of a social effect change their decisions and follow the majority towards  $p_2$  (see yellow balls in Fig. 3). In this case,  $(0.1775, 0.8216, 0.0009)$  satisfies (30). If the social impact increases more to  $q = 30$ , then a consensus to follow the majority occurs. Figure 4 shows that as the social effect increases, the basin of attraction of  $p_2$  (blue area) increases at the expenses of the two other basins.

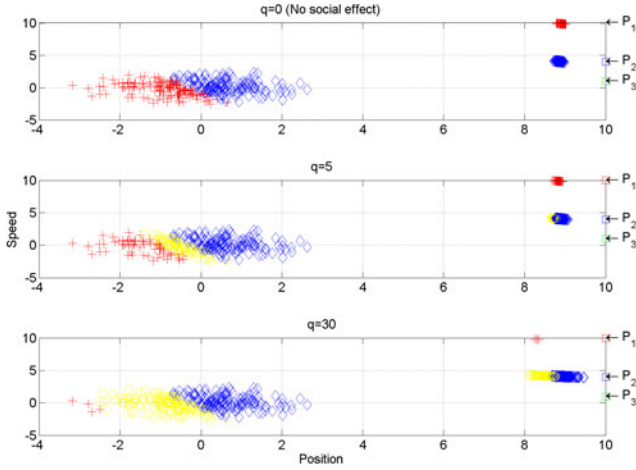


Fig. 3. Influence of the social effect  $q$ . Red Crosses = agents that go towards  $p_1$ . Blue diamonds = agents that go towards  $p_2$ . Yellow balls = agents that change their decisions when the social effect increases and follow the majority towards  $p_2$ .

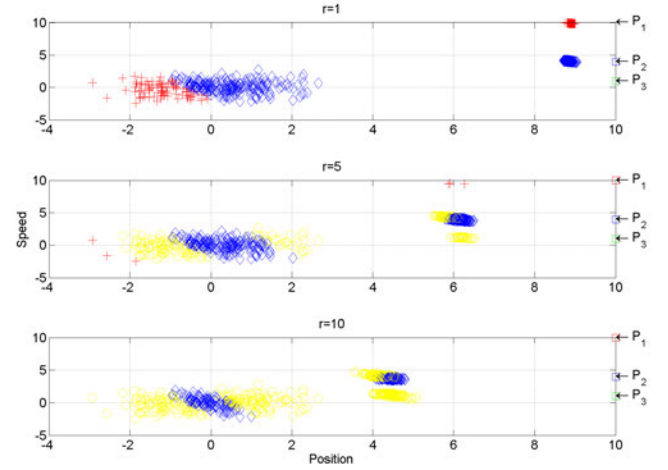


Fig. 5. Influence of  $r$ .

To illustrate the Gaussian Binary Choice Case, we consider a group of agents initially drawn from the normal distribution  $\mathcal{N}(\mu_0, I_2)$  and moving according to the same dynamics towards the destination points  $p_1 = (-10, 0)$  or  $p_2 = (10, 0)$ . Now the robots are moving on the real line towards  $-10$  or  $10$ , where they must arrive at a speed 0. For the covariance matrix  $\Sigma_0 = I_2$ ,  $S(\Sigma_0)$  defined in (32) is the region delimited by the lines  $y = -3.8203 - 0.9164x$  and  $y = 3.8203 - 0.9164x$ . If  $\mu_0 = (5, 0)$ , i.e., outside  $S(\Sigma_0)$ , only one  $\epsilon$ -Nash equilibrium corresponding to  $\lambda = (0, 1)$  exists. If  $\mu_0 = (0, 0)$ , i.e. inside  $S(\Sigma_0)$ , three  $\epsilon$ -Nash equilibria exist. The first corresponds to  $\lambda_1 = (0.06, 1 - 0.06)$ , the second to  $\lambda_2 = (0.5, 1 - 0.5)$ , and the third to  $\lambda_3 = (0.94, 1 - 0.94)$  (see Fig. 6).

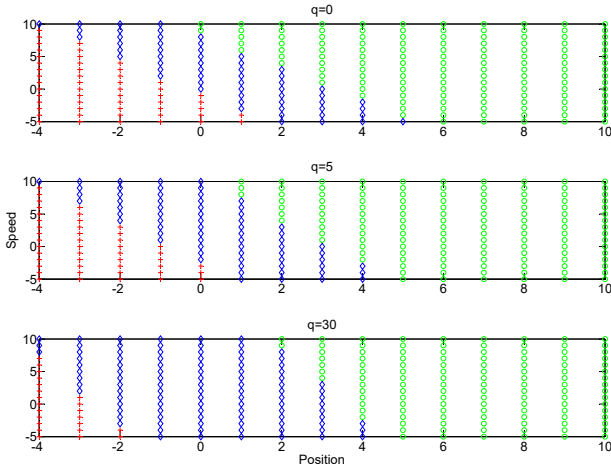


Fig. 4. Influence of the social effect  $q$  on the Basins of attraction. Region hatched with: red crosses = basin of attraction of  $p_1$ , blue diamonds = basin of attraction of  $p_2$  and green circles = basin of attraction of  $p_3$ .

To illustrate the impact of the individual efforts on the behavior of the population (see Fig. 5), we start with the case where the control effort is inexpensive ( $r = 1$ ). In this case,  $\lambda = (0.3066, 0.6921, 0.0013)$  is a fixed point of  $F_s$ , defined in (28). As the effort coefficient increases ( $r = 5$ ), the majority of the agents that went to  $p_1$  in the previous case ( $r = 1$ ) choose a closer alternative, namely  $p_2$  (yellow balls). Moreover, some of the agents that chose  $p_2$  move now towards a less expensive choice  $p_3$  (yellow balls). In this case  $\lambda = (0.0097, 0.9085, 0.0819)$ . As  $r$  increases to 10, more players change their choices from  $p_2$  to  $p_1$  (yellow balls), and  $\lambda$  is equal to  $(0.0004, 0.6664, 0.3332)$  in this case.

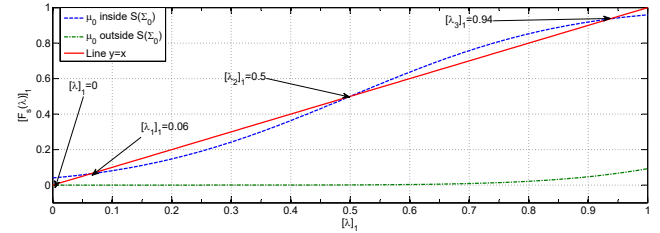


Fig. 6. Gaussian binary choice case. Fixed points of  $F_s$  defined in (28).

## IX. CONCLUSION

We consider in this paper a dynamic collective choice model where a large number of agents are choosing between multiple destination points while taking into account the social effect as represented by the mean of the population. The analysis is carried using the MFG methodology. We show that under this social effect, the population may split between the destination points in different ways. For a uniform population, we show that there exists a one to one map between the self-replicating mean trajectories and the fixed points of a function defined on  $\mathbb{R}^l$ . The latter describe the way the agents split between the  $l$  destination points. Finally, we prove that the decentralized strategies optimally tracking the self-replicating mean trajectories are approximate Nash equilibria. We suspect that in the uniform case the linear dynamics and the quadratic running costs are necessary to reduce the infinite dimensional

fixed point problem to a finite dimensional one. It is of interest for future work to extend our analysis to a model where players have stochastic dynamics as well. In that case, the optimal choices (feedback strategies) change along the path according to the occurring events (noise). This is in contrast to the current formulation where the agents can choose without loss of optimality their destination before they start moving. Also, we would like to extend the current formulation to certain nonlinear models, where the basins of attraction are delimited by more complex manifolds, and the fixed-point computations would require numerical methods for backward-forward systems of partial differential equations [49].

## APPENDIX A

This appendix includes the proofs of lemmas and theorems related to the tracking and fixed point problems in case of deterministic initial conditions.

### A. Proof of Lemma 2

In this proof, the subscript  $M$  indicates the dependence on the final cost's coefficient  $M$ . For any  $M > 0$ , the agents are optimally tracking a path  $\hat{x}_M$ . The agent  $i$  optimal state is denoted by  $x_{iM}^*(t)$ . We have

$$\frac{M}{2} \min_{j \in \{1, \dots, l\}} \left( \|x_{iM}^*(T) - p_j\|^2 \right) \leq J_{iM}(u_{iM}^*, \hat{x}_M, x_i^0),$$

where  $J_{iM}(u_{iM}^*, \hat{x}_M, x_i^0)$  is the cost defined by (2) with the final cost's coefficient equal  $M$ . It suffices to find an upper bound for  $J_{iM}(u_{iM}^*, \hat{x}_M, x_i^0)$  which is uniformly bounded with  $M$ . Since  $(A, B)$  is controllable, then there exists for each agent  $i$  a continuous control law  $u_{x_i^0, p_1}(t)$  on  $[0, T]$  which transfers this agent from the state  $x_i^0$  to  $p_1$  in a finite time  $T$ . By optimality, we have  $J_{iM}(u_{iM}^*, \hat{x}_M, x_i^0) \leq J_{iM}(u_{x_i^0, p_1}, \hat{x}_M, x_i^0)$ . But,

$$J_{iM}(u_{x_i^0, p_1}, \hat{x}_M, x_i^0) = \int_0^T \left\{ \frac{q}{2} \|x_i(u_{x_i^0, p_1}) - \hat{x}_M\|^2 + \frac{r}{2} \|u_{x_i^0, p_1}\|^2 \right\} dt,$$

which is uniformly bounded with  $M$ , since  $\hat{x}_M$  is uniformly bounded with  $M$ . Thus, for all  $\epsilon > 0$ , there exists an  $M_0 > 0$  such that for all  $M > M_0$ ,  $\min_{j \in \{1, \dots, l\}} \left( \|x_{iM}^*(T) - p_j\|^2 \right) < \epsilon$ .

### B. Fixed points of $G_2(\cdot, \lambda)$

For any  $\lambda = (\lambda_1, \dots, \lambda_l) \in \{0, \dots, N\}^l$ , we define the map  $T_\lambda$  from  $C([0, T], \mathbb{R}^n)$  to  $C([0, T], \mathbb{R}^n)$  by  $T_\lambda(\hat{x}) = G_2(\hat{x}, \lambda)$ , with  $G_2$  defined in (15).

**Lemma 9:** For any  $\lambda = (\lambda_1, \dots, \lambda_l) \in \{0, \dots, N\}^l$ ,  $T_\lambda$  has a unique fixed point equal to

$$y_\lambda = R_1(t)\bar{x}_0 + R_2(t)p_\lambda, \quad (41)$$

where  $R_1$  and  $R_2$  are defined in (16)-(17) and  $\bar{x}_0$  is the agents' initial average. Moreover, if  $(A, B)$  is controllable, then the paths  $y_\lambda$  are uniformly bounded with respect to  $(M, \lambda) \in [0, \infty) \times [0, N]^l$  for the norm  $\left( \int_0^T \|\cdot\|^2 dt \right)^{\frac{1}{2}}$ .

*Proof:* Consider  $y$  a fixed point of  $T_\lambda$ . We define

$$n(t) = \Gamma(t)y(t) + q \int_T^t \Phi(t, \sigma)y(\sigma)d\sigma - M\Phi(t, T)p_\lambda,$$

where  $\Gamma$  and  $\Phi$  are defined in (6a) and (9). One can easily check that  $(y, n)$  satisfies

$$\begin{aligned} \dot{y} &= Ay - \frac{1}{r}BB^T n & y(0) &= \bar{x}_0 \\ \dot{n} &= -A^T n & n(T) &= M(y(T) - p_\lambda). \end{aligned} \quad (42)$$

Therefore,  $y$  and  $n$  are respectively the optimal state and co-state of the following LQR problem:

$$\begin{aligned} \min_u \int_0^T \frac{r}{2} \|u\|^2 dt + \frac{M}{2} \|x(T) - p_\lambda\|^2 & \quad (43) \\ \text{Subject to } \dot{x} &= Ax + Bu & x(0) &= \bar{x}_0. \end{aligned}$$

Hence,  $n$  has the representation  $n(t) = P(t)y(t) + g(t)$ , where  $P$  is the unique solution of the Riccati equation (20) and  $g$  satisfies  $\dot{g} = -(A - \frac{1}{r}BB^T P(t))^T g$ , with  $g(T) = -Mp_\lambda$ . By solving  $g$  and implementing its expression in  $n = Py + g$ , and by implementing the new expression of  $n$  in the dynamics of  $y$ , one can show that  $y(t) = R_1(t)\bar{x}_0 + R_2(t)p_\lambda$ . Conversely, let  $(n, y)$  be the unique solution of (42). We define  $m(t) = \Gamma(t)y(t) + q \int_T^t \Phi(t, \sigma)y(\sigma)d\sigma - M\Phi(t, T)p_\lambda$ . One can easily check that  $d(m - n)/dt = (\frac{1}{r}\Gamma BB^T - A^T)(m - n)$ , with  $m(T) = n(T)$ . Therefore,  $n = m$  and from (42)  $y$  is a fixed point of  $T_\lambda$ .

We now prove the uniform boundedness of the fixed point paths  $y_\lambda$  with respect to  $(M, \lambda)$ . The paths  $y_\lambda$  are the optimal states of the control problem (43). Since  $(A, B)$  is controllable, the corresponding optimal control law  $u_\lambda$  satisfies

$$\int_0^T \frac{r}{2} \|u_\lambda\|^2 dt \leq \int_0^T \frac{r}{2} \|u_0^\lambda\|^2 dt,$$

where  $u_0^\lambda$  is a continuous control law that transfers the state  $y$  from  $y(0)$  to  $p_\lambda$ . But  $u_0^\lambda$  is independent of  $M$  and continuous with respect to  $\lambda$ . Hence,

$$\sup_{\lambda \in [0, N]^l} \int_0^T \frac{r}{2} \|u_\lambda\|^2 dt \leq \max_{\lambda \in [0, N]^l} \int_0^T \frac{r}{2} \|u_0^\lambda\|^2 dt.$$

We have

$$y_\lambda(t) = \exp(At)\bar{x}_0 + \int_0^t \exp(A(t - \sigma))Bu_\lambda(\sigma)d\sigma.$$

Therefore,

$$\int_0^T \|y_\lambda\|^2 dt \leq K_1 + K_2 \int_0^T \|u_\lambda\|^2 dt + K_3 \left( \int_0^T \|u_\lambda\|^2 dt \right)^{\frac{1}{2}},$$

for some positive constants  $K_1, K_2, K_3$  which are independent of  $(M, \lambda)$ . Hence,  $y_\lambda$  is uniformly bounded with  $(M, \lambda)$ . ■

### C. Proof of Theorem 3

If  $\hat{x}$  is a fixed point of  $G$ , then  $\hat{x}$  is a fixed point of the map  $T_\lambda(\hat{x})$ , with  $T_\lambda$  defined in the previous subsection. By Lemma 9,  $\hat{x}$  is of the form (22). To show that  $\lambda(\hat{x})$  is a fixed point of  $F$ , we replace the new expression of the fixed point  $\hat{x}$  (22)

in the expressions of the basins of attraction, which will then have the following form

$$H_j^{\lambda(\hat{x})} = D_j(R_1(t)\bar{x}_0 + R_2(t)p_{\lambda(\hat{x})}) = \{x \in \mathbb{R}^n | \beta_{jk}^T x \leq \delta_{jk} + \theta_{jk}\bar{x}_0 + \xi_{jk}p_{\lambda(\hat{x})} \quad \forall k \in \{1, \dots, l\}\}.$$

Therefore,

$$\begin{aligned} \lambda(\hat{x}) &= (|\{x_i^0 | x_i^0 \in D_1(\hat{x})\}|, \dots, |\{x_i^0 | x_i^0 \in D_l(\hat{x})\}|) \\ &= (|\{x_i^0 | x_i^0 \in H_1^{\lambda(\hat{x})}\}|, \dots, |\{x_i^0 | x_i^0 \in H_l^{\lambda(\hat{x})}\}|) \triangleq F(\lambda(\hat{x})). \end{aligned}$$

Thus, we proved that if  $\hat{x}$  is a fixed point of  $G$ , then  $\hat{x}$  is of the form (22), where  $\lambda = \lambda(\hat{x})$  is a fixed point of the finite dimensional operator  $F$  defined in (21).

To prove the converse, we consider a fixed point  $\lambda$  of  $F$  and the path  $\hat{x} = R_1(t)\bar{x}_0 + R_2(t)p_{\lambda}$ . We have

$$\begin{aligned} \lambda &= F(\lambda) = (|\{x_i^0 | x_i^0 \in H_1^{\lambda}\}|, \dots, |\{x_i^0 | x_i^0 \in H_l^{\lambda}\}|) \\ &= (|\{x_i^0 | x_i^0 \in D_1(\hat{x})\}|, \dots, |\{x_i^0 | x_i^0 \in D_l(\hat{x})\}|) = \lambda(\hat{x}), \end{aligned}$$

where the third equality is a consequence of the form of  $\hat{x}$ . By Lemma 9, the path  $\hat{x}$  is the unique fixed point of  $T_{\lambda}$ . But then  $\hat{x} = T_{\lambda}(\hat{x}) = T_{\lambda(\hat{x})}(\hat{x}) = G(\hat{x})$ . Therefore,  $\hat{x}$  is a fixed point of  $G$ .

#### D. Proof of Theorem 4

The first point follows from Theorem 3 and (23). For 2) and 3), we define

$$a_N(\alpha) = \frac{N}{\xi_{12}(p_1 - p_2)} (\beta_{12}^T x_{\alpha}^0 - \delta_{12} - \theta_{12}\bar{x}_0 - \xi_{12}p_2).$$

We start by proving 2). Suppose that there does not exist any  $\alpha$  in  $\{0, \dots, N\}$  satisfying (24), (25) or (26). Zero does not satisfy (25), hence  $a_N(1) \leq 0 < 1$ . One does not satisfy (24) and  $a_N(1) < 1$ , hence  $a_N(2) \leq 1$ . By induction, we have  $a_N(N) \leq N - 1$ . Therefore,  $N$  satisfies (26). Thus, by contradiction, there exists  $\alpha$  in  $\{0, \dots, N\}$  satisfying (24), (25) or (26). We now prove the third point. Suppose that there exist multiple  $\alpha$ 's satisfying (24), (25) or (26). Let  $\alpha_0$  be the least of these  $\alpha$ 's. If  $\alpha_0 < N$ , then in view of  $\xi_{12}(p_1 - p_2) < 0$ ,  $a_N(\alpha_0 + 1) < \alpha_0 \leq a_N(\alpha_0)$ .  $a_N(j)$  is decreasing. Hence, for all  $\alpha > \alpha_0$ ,  $a_N(\alpha) \leq a_N(\alpha_0 + 1) < \alpha_0 < \alpha$ . Therefore,  $\alpha_0$  is the unique  $\alpha$  satisfying (24), (25) or (26). Finally, if  $\xi_{12}(p_1 - p_2) < 0$ , then an initial distribution for which  $a_N(\alpha)$  is in  $(0, 1)$ , for all  $\alpha$  in  $\{0, \dots, N\}$ , does not have any  $\alpha$  in  $\{0, \dots, N\}$  satisfying (24), (25) or (26).

## APPENDIX B

In this appendix, we provide the proofs of theorems related to the fixed point problems in the random initial conditions and non-uniform population cases, as well as the proof of Theorem 8, which characterizes the decentralized control laws developed in Sections III to VI as  $\epsilon$ -Nash strategies.

#### A. Proof of Theorem 5

We start by proving (i). Let  $\hat{x}$  be a fixed point of  $G_s$  and  $\lambda_j = P_0(D_j(\hat{x}))$ . By replacing the probabilities in the expression of  $G_s$  by  $\lambda_j$ ,  $j \in \{1, \dots, l\}$ , we get  $\hat{x} = G_s(\hat{x}) = T_{\lambda}(\hat{x})$ , where  $\lambda = (\lambda_1, \dots, \lambda_l)$  and  $T_{\lambda}$  is defined above Lemma 9. Hence,  $\hat{x}$  is a fixed point of  $T_{\lambda}$ . By Lemma 9,  $\hat{x}(t) = R_1(t)\mu_0 + R_2(t)p_{\lambda}$ . By replacing this expression of  $\hat{x}$  in  $D_j(\hat{x})$  (11), we get  $\lambda = F_s(\lambda)$ . Conversely, consider  $\lambda = (\lambda_1, \dots, \lambda_l)$  in  $\Delta_l$  such that  $\lambda = F_s(\lambda)$  and let  $\hat{x}(t) = R_1(t)\mu_0 + R_2(t)p_{\lambda}$ . The path  $\hat{x}$  is the unique fixed point of  $T_{\lambda}$  and

$$\left( P_0(D_1(\hat{x})), \dots, P_0(D_l(\hat{x})) \right) = F_s(\lambda) = \lambda.$$

Hence,  $\hat{x} = T_{\lambda}(\hat{x}) = G_s(\hat{x})$ . We now prove the second point. Noting that the set  $\Delta_l$  is convex and compact in  $\mathbb{R}^l$ , we just need to show that  $F_s$  is continuous. Then, Brouwer's fixed point theorem [41] ensures the existence of a fixed point. Let  $\lambda_r$  be a sequence in  $\Delta_l$  converging to  $\lambda$ . We have

$$\begin{aligned} \left| [F_s(\lambda_r)]_k - [F_s(\lambda)]_k \right| &= \left| \int_{\mathbb{R}^n} (1_{\bar{H}_k^{\lambda_r}}(x) - 1_{\bar{H}_k^{\lambda}}(x)) dP_0(x) \right| \\ &\leq \int_{\mathbb{R}^n} \left| 1_{\bar{H}_k^{\lambda_r}}(x) - 1_{\bar{H}_k^{\lambda}}(x) \right| dP_0(x). \end{aligned}$$

But,  $\bar{H}_k^{\lambda_r}$  and  $\bar{H}_k^{\lambda}$ , defined in (29), are regions delimited by hyperplanes. Hence, under Assumption 2,

$$\begin{aligned} \int_{\mathbb{R}^n} \left| 1_{\bar{H}_k^{\lambda_r}}(x) - 1_{\bar{H}_k^{\lambda}}(x) \right| dP_0(x) &= \\ \int_{\mathbb{R}^n} \left| 1_{\bar{H}_k^{\lambda_r}}(x) - 1_{\bar{H}_k^{\lambda}}(x) \right| dP_0(x). \end{aligned}$$

But,  $\left| 1_{\bar{H}_k^{\lambda_r}}(x) - 1_{\bar{H}_k^{\lambda}}(x) \right| \leq 2$  and converges to zero for all  $x$  in  $\mathbb{R}^n$ . Thus, by Lebesgue dominated convergence theorem [48], the integral of this function converges to zero. This proves that  $F_s$  is continuous. Finally, we prove (iii). For  $l = 2$ , the fixed points of  $F_s$  are of the form  $(\alpha, 1 - \alpha)$ . The set of fixed points of  $F_s$  is compact. Thus, the set of the first components of these fixed points is compact. Let  $\alpha_0$  be the minimum of those first components. Consider  $\alpha > \alpha_0$ . Hence,

$$\begin{aligned} &\left\{ (\beta_{12})^T x_i^0 - \delta_{12} - \theta_{12}\mu_0 - \xi_{12}p_2 \leq \alpha \xi_{12}(p_1 - p_2) \right\} \\ &\subset \left\{ (\beta_{12})^T x_i^0 - \delta_{12} - \theta_{12}\mu_0 - \xi_{12}p_2 \leq \alpha_0 \xi_{12}(p_1 - p_2) \right\}, \end{aligned}$$

which implies  $\left[ F_s(\alpha, 1 - \alpha) \right]_1 \leq \left[ F_s(\alpha_0, 1 - \alpha_0) \right]_1 = \alpha_0 < \alpha$ . Thus,  $(\alpha_0, 1 - \alpha_0)$  is the unique fixed point of  $F_s$ , and  $\hat{x}(t) = R_1(t)\mu_0 + R_2(t)p_{(\alpha_0, 1 - \alpha_0)}$  is the unique fixed point of  $G_s$ .

#### B. Proof of Theorem 6

We show in Theorem 5 that the fixed points of  $G_s$  can be one to one mapped to the fixed points of  $F_s$ . The initial states  $x_i^0$  are distributed according to a Gaussian distribution  $\mathcal{N}(\mu_0, \Sigma_0)$ . Therefore,  $\beta_{12}^T x_i^0$  are distributed according to the normal distribution  $\mathcal{N}(\beta_{12}^T \mu_0, \beta_{12}^T \Sigma_0 \beta_{12})$ . Thus, one can analyze the dependence of  $\left[ F_s(\alpha, 1 - \alpha) \right]_1 - \alpha$  on  $\alpha$  to show

that this function has a unique zero in  $[0, 1]$  in case 1) or 2) holds. Indeed, if 1) or 2) holds, the sign of the derivative with respect to  $\alpha$  of  $[F_s(\alpha, 1 - \alpha)]_1 - \alpha$  does not change. Thus, this function is monotonic. This implies that  $F_s$  and  $G_s$  have unique fixed points.

### C. Proof of Theorem 7

We use Schauder's fixed point theorem [41] to prove the existence of a fixed point. We start by showing that  $G_p$  is a compact operator, i.e.,  $G_p$  is continuous and maps bounded sets to relatively compact sets. Let  $\hat{x}$  be in  $C([0, T], \mathbb{R}^n)$  and  $\{\hat{x}_k\}_{k \in \mathbb{N}}$  be a sequence converging to  $\hat{x}$  in  $(C([0, T], \mathbb{R}^n), \|\cdot\|_\infty)$ . Let

$$Q_j = \max_{(\theta, t) \in \Theta \times [0, T]^2} \|\Phi_j^\theta(t)\| + \max_{(\theta, t) \in \Theta \times [0, T]^4} \|\Psi_j^\theta(t)\| + \max_{\theta \in \Theta} \|M^\theta\|.$$

We have

$$\|G_p(\hat{x}_k) - G_p(\hat{x})\|_\infty \leq \sum_{j=1}^l Q_j \left\{ \frac{qT^2}{r} \|\hat{x}_k - \hat{x}\|_\infty + V_{1j} + \frac{Q_j \|p_j\|T + q\|\hat{x}\|_\infty T^2}{r} V_{2j} \right\},$$

where

$$V_{1j} = \int_{\Theta} \int_{\mathbb{R}^n} \left| 1_{D_j^\theta(\hat{x}_k)}(x^0) - 1_{D_j^\theta(\hat{x})}(x^0) \right| \|x^0\| dP_0 dP_\theta$$

$$V_{2j} = \int_{\Theta} \int_{\mathbb{R}^n} \left| 1_{D_j^\theta(\hat{x}_k)}(x^0) - 1_{D_j^\theta(\hat{x})}(x^0) \right| dP_0 dP_\theta.$$

Under Assumption 5,

$$V_{1j} = \int_{\Theta} \int_{\mathbb{R}^n} \left| 1_{\hat{D}_j^\theta(\hat{x}_k)}(x^0) - 1_{\hat{D}_j^\theta(\hat{x})}(x^0) \right| \|x^0\| dP_0 dP_\theta.$$

But,  $\left| 1_{\hat{D}_j^\theta(\hat{x}_k)}(x^0) - 1_{\hat{D}_j^\theta(\hat{x})}(x^0) \right| \|x^0\| \leq 2\|x^0\|$  and converges to zero for all  $(x^0, \theta)$  in  $\mathbb{R}^n \times \Theta$ . We have  $\mathbb{E}\|x^0\| < \infty$ . Therefore, by Lebesgue's dominated convergence theorem [48],  $V_{1j}$  converges to zero. By the same technique, we prove that  $V_{2j}$  converges to zero. Hence,  $G_p$  is continuous.

Let  $V$  be a bounded subset of  $C([0, T], \mathbb{R}^n)$ . We show in the following via Arzela-Ascoli Theorem that the closure of  $G_p(V)$  is compact. Let  $\{G_p(\hat{x}_k)\}_{k \in \mathbb{N}} \in G_p(V)$ . By the continuity of  $\Phi_j^\theta(\sigma, t)$  with respect to  $(\sigma, t, \theta)$ , of its derivative with respect to  $t$  and  $\sigma$ , and by the boundedness of  $\hat{x}_k$ , one can prove that for all  $(t, s)$  in  $[0, T]^2$ ,

$$\|G_p(\hat{x}_k)(t) - G_p(\hat{x}_k)(s)\| \leq (K_1 \mathbb{E}\|x^0\| + K_2) |t - s|,$$

where  $K_1$  and  $K_2$  are positive constants. This inequality implies the uniform boundedness and equicontinuity of  $\{G_p(\hat{x}_k)\}_{k \in \mathbb{N}}$ . By Arzela-Ascoli Theorem [41], there exists a convergent subsequence of  $\{G_p(\hat{x}_k)\}_{k \in \mathbb{N}}$ . Hence,  $G_p(V)$  and its closure are compact sets, and  $G_p$  is a compact operator.

Now we construct a nonempty, bounded, closed, convex subset  $U \subset C([0, T], \mathbb{R}^n)$  such that  $G_p(U) \subset U$ . Let  $Q = \max(k_1 + k_2, k_3)$ , where  $k_1$ ,  $k_2$  and  $k_3$  are defined in (40). We have  $\|G_p(x)(t)\| \leq Q + Q \int_0^t \int_\sigma^T \|\hat{x}(\tau)\| d\tau d\sigma$ . Consider the following set

$$U = \left\{ x \in C([0, T], \mathbb{R}^n) \mid \|x(t)\| \leq R(t), \forall t \in [0, T] \right\},$$

where  $R$  is a continuous positive function on  $[0, T]$  to be determined later.  $U$  is a nonempty, bounded, closed and convex subset of  $C([0, T], \mathbb{R}^n)$ . If we find an  $R$  positive such that  $R(t) = Q + Q \int_0^t \int_\sigma^T R(\tau) d\tau d\sigma$ , for all  $t \in [0, T]$ , then for all  $\hat{x} \in U$ ,

$$\|G_p(x)(t)\| \leq Q + Q \int_0^t \int_\sigma^T R(\tau) d\tau d\sigma = R(t). \quad (44)$$

Hence,  $G_p(U) \subset U$ . It remains to find such  $R$ . Note that the equality in (44) is equivalent to the second order differential equation  $\ddot{R} = -QR$ , with the boundary conditions,  $R(0) = Q$  and  $\dot{R}(T) = 0$ . Thus,  $R(t) = Q / \cos(\sqrt{Q}t)$ , which is positive under Assumption 4. Having found  $R$ ,  $U$  is well defined and by Schauder's Theorem,  $G_p$  has a fixed point in  $U$ .

### D. Proof of Theorem 8

We consider an arbitrary agent  $i \in \{1, \dots, N\}$  applying an arbitrary full state feedback control law  $u_i$ . Suppose that this agent  $i$  can profit by a unilateral deviation from the decentralized strategies. This means that

$$J_i(u_i, u_{-i}^*) \leq J_i(u_i^*, u_{-i}^*). \quad (45)$$

In the following, we prove that this potential cost improvement is bounded by some  $\epsilon_N$  that converges to zero as  $N$  increases to infinity. We denote respectively by  $x_i$  and  $x_j^*$  the states corresponding to  $u_i$  and  $u_j^*$ . In view of (34), the compactness of  $\Theta$ , the continuity of  $x_j^*$  with respect to  $\theta$  and  $\mathbb{E}\|x_i^0\|^2 < \infty$ , the right hand side of (45) is bounded by  $Q_1$  independently of  $N$ . For any  $X$  and  $Y$  in  $C([0, T], \mathbb{R}^n)$ , we define  $\langle X|Y \rangle = \mathbb{E}\left(\int_0^T X^T(t)Y(t) dt \mid x_i^0\right)$  and  $\|X\|_2 = \sqrt{\langle X|X \rangle}$ . We have

$$J_i(u_i, u_{-i}^*) = J_i(x_i, \hat{x}, x_i^0) + \frac{q}{2} \left\| \hat{x} - \frac{1}{N} \sum_{j=1}^N x_j^* \right\|_2^2 + \frac{q}{2N^2} \|x_i^* - x_i\|_2^2 + S_1 + S_2 + S_3,$$

where

$$S_1 = \frac{q}{N} \langle x_i^* - x_i \mid x_i - \hat{x} \rangle$$

$$S_2 = \frac{q}{N} \langle x_i^* - x_i \mid \hat{x} - \frac{1}{N} \sum_{j=1}^N x_j^* \rangle$$

$$S_3 = q \langle \hat{x} - \frac{1}{N} \sum_{j=1}^N x_j^* \mid x_i - \hat{x} \rangle,$$

with  $\hat{x}$  is a fixed point of  $G_p$ . By the Cauchy-Schwarz inequality,

$$|S_1| \leq \frac{q}{N} \|x_i^* - x_i\|_2 \|x_i - \hat{x}\|_2.$$

In view of (45) and the bound  $Q_1$ ,  $\|x_i^* - x_i\|_2$  and  $\|x_i - \hat{x}\|_2$  are bounded. Thus,  $|S_1| \leq \eta_1/N$ , where  $\eta_1 > 0$ . Similarly,  $|S_2| \leq \eta_2/N$ , where  $\eta_2 > 0$ . We define

$$\alpha_N = \left\| \hat{x} - \frac{1}{N} \sum_{j=1}^N \mathbb{E}x_j^* \right\|_2 = \left\| \int_{\Theta} \bar{x}^\theta dP_\theta - \int_{\Theta} \bar{x}^\theta dP_\theta^N \right\|_2,$$

where  $\bar{x}^\theta = \mathbb{E}(x^{0\theta}|\theta)$ , with  $x^{0\theta}$  defined in (38). We have

$$\left\| \hat{x} - \frac{1}{N} \sum_{j=1}^N x_j^* \right\|_2^2 \leq 2\alpha_N^2 + 2 \left\| \frac{1}{N} \sum_{j=1}^N (\mathbb{E}x_j^* - x_j^*) \right\|_2^2.$$

By the compactness of  $[0, T] \times \Theta$ , the family of functions  $\bar{x}^\theta(t)$  defined on  $\Theta$  and indexed by  $t$  is uniformly bounded and equicontinuous. By Corollary 1.1.5 of [50], we deduce

$$\lim_{N \rightarrow +\infty} \sup_{t \in [0, T]} \left\| \hat{x}(t) - \frac{1}{N} \sum_{j=1}^N \mathbb{E}x_j^*(t) \right\| = 0.$$

Thus,  $\alpha_N$  converges to 0 as  $N$  increases to infinity. By the independence of the initial conditions (and thus the independence of  $x_j^*$ ,  $j \in \{1, \dots, N\}$ ) and the assumption  $\mathbb{E}\|x_i^0\|^2 < \infty$ , we deduce that  $\left\| \frac{1}{N} \sum_{j=1}^N (\mathbb{E}x_j^* - x_j^*) \right\|_2^2 = O(1/N)$ . Thus,  $S_3$  and  $\left| J_i(x_i^*, \hat{x}, x_i^0) - J_i(u_i^*, u_{-i}^*) \right|$  converge to 0 as  $N$  increases to infinity. By optimality, we have  $J_i(x_i^*, \hat{x}, x_i^0) \leq J_i(x_i, \hat{x}, x_i^0)$ . Therefore,  $J_i(u_i, u_{-i}^*) \geq J_i(u_i^*, u_{-i}^*) + \epsilon_N$ , where  $\epsilon_N = J_i(x_i^*, \hat{x}, x_i^0) - J_i(u_i^*, u_{-i}^*) + S_1 + S_2 + S_3$  converges to 0 as  $N$  increases to infinity.

#### REFERENCES

- [1] 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.
- [2] —, "A dynamic game model of collective choice in multi-agent systems," in *Proceedings of the 54th IEEE Conference on Decision and Control*, dec 2015.
- [3] N. E. Leonard, T. Shen, B. Nabet, L. Scardovi, I. D. Couzin, and S. A. Levin, "Decision versus compromise for animal groups in motion," *Proceedings of the National Academy of Sciences*, vol. 109, no. 1, pp. 227–232, 2012.
- [4] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, "Effective leadership and decision-making in animal groups on the move," *Nature*, vol. 433, pp. 513–516, 2005.
- [5] S. Merrill and B. Grofman, *A Unified Theory of Voting: Directional and Proximity Spatial Models*. Cambridge University Press, 1999.
- [6] T. D. Seeley, S. Camazine, and J. Sneyd, "Collective decision-making in honey bees: how colonies choose among nectar sources," *Behavioral Ecology and Sociobiology*, vol. 28, pp. 277–290, 1991.
- [7] S. Camazine, P. K. Visscher, J. Finley, and R. S. Vetter, "House-hunting by honey bee swarms: collective decisions and individual behaviors," *Insectes Sociaux*, vol. 46, no. 4, pp. 348–360, November 1999.
- [8] J. H. Tien, S. A. Levin, and D. I. Rubenstein, "Dynamics of fish schools: identifying key decision rules," *Evolutionary Ecology Research*, vol. 6, pp. 555–565, 2004.
- [9] I. Aoki, "A simulation study on the schooling mechanism in fish," *Bulletin of the Japanese Society for the Science of Fish*, vol. 48, pp. 1081–1088, 1982.
- [10] S. C. Pratt and D. J. T. Sumpter, "A tunable algorithm for collective decision-making," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 15906–15910, 2006.
- [11] 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.
- [12] T. Liggett, *Interacting particle systems*. Springer Science & Business Media, 2012, vol. 276.
- [13] P. Whittle and P. Gait, "Reduction of a class of stochastic control problems," *IMA Journal of Applied Mathematics*, vol. 6, no. 2, pp. 131–140, 1970.
- [14] P. Whittle, *Optimization over time*. Wiley, 1982.
- [15] J. Kuhn, "The risk-sensitive homing problem," *Journal of Applied Probability*, vol. 22, pp. 796–803, 1985.
- [16] M. Lefebvre, "A homing problem for diffusion processes with control-dependent variance," *The Annals of Applied Probability*, vol. 14, pp. 786–795, 2004.
- [17] M. Lefebvre and F. Zitouni, "General LQG homing problems in one dimension," *International Journal of Stochastic Analysis*, vol. 2012, 2012.
- [18] A. Okubo, "Dynamical aspects of animal grouping: swarms, schools, flocks, and herds," *Advances in biophysics*, vol. 22, pp. 1–94, 1986.
- [19] C. Robinson and T. Pitcher, "Hunger motivation as a promoter of different behaviours within a shoal of herring: selection for homogeneity in fish shoal?" *Journal of Fish Biology*, vol. 35, no. 3, pp. 459–460, 1989.
- [20] 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.
- [21] 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.
- [22] D. Vail and M. Veloso, "Multi-robot dynamic role assignment and coordination through shared potential fields," *Multi-robot systems*, pp. 87–98, 2003.
- [23] N. Nourian, R. P. Malhamé, M. Huang, and P. E. Caines, "Mean-field NCE formulation of estimation-based leader-follower collective dynamics," *International Journal of Robotics and Automation*, vol. 26, no. 1, pp. 120–129, 2011.
- [24] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, First, Ed. Princeton University Press, 2010.
- [25] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009.
- [26] J. Le Ny and G. J. Pappas, "Adaptive deployment of mobile robotic networks," *IEEE Transactions on automatic control*, vol. 58, pp. 654–666, 2013.
- [27] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice hall Englewood Cliffs, NJ, 1989, vol. 23.
- [28] 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.
- [29] 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.
- [30] M. A. Burke, G. Fournier, and K. Prasad, *Physician social networks and geographical variation in medical care*. Center on Social and Economic Dynamics, 2003.
- [31] W. Brock and S. Durlauf, "Discrete choice with social interactions," *Review of Economic Studies*, pp. 147–168, 2001.
- [32] M. Huang, P. E. Caines, and R. P. Malhamé, "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.
- [33] J. M. Lasry and P. L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, pp. 229–260, 2007.
- [34] M. Huang, "Stochastic control for distributed systems with applications to wireless communications," Ph.D. dissertation, McGill University, 2003.
- [35] 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*, Maui, Hawaii, 2003, pp. 98–103.
- [36] M. Huang, R. P. Malhamé, and P. E. Caines, "Nash certainty equivalence in large population stochastic dynamic games: Connections with the physics of interacting particle systems," in *Proceedings of the 44th IEEE Conference on Decision and Control*, San Diego, CA, 2006, pp. 4921–4926.
- [37] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. Siam, 1999, vol. 23.
- [38] 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.
- [39] 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.
- [40] —, "Jeux à champ moyen. II—horizon fini et contrôle optimal," *Comptes Rendus Mathématique*, vol. 343, no. 10, pp. 679–684, 2006.
- [41] J. B. Conway, *A Course in Functional Analysis*, ser. Graduate Texts in Mathematics. Springer-Verlag, 1985.



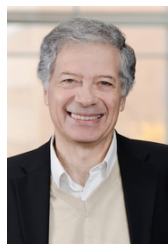
- [42] R. Salhab, R. P. Malhamé, and J. L. Ny, "A dynamic game model of collective choice: Stochastic dynamics and closed loop solutions," *CoRR*, vol. abs/1604.08136, 2016. [Online]. Available: <http://arxiv.org/abs/1604.08136>
- [43] B. D. Anderson and J. B. Moore, *Optimal control: linear quadratic methods*. Dover Publications, 2007.
- [44] W. Rugh, *Linear System Theory*, ser. Prentice-Hall information and systems sciences series. Prentice Hall, 1993.
- [45] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of computation*, pp. 577–593, 1965.
- [46] R. Nakajima, "Measuring peer effects on youth smoking behaviour," *The Review of Economic Studies*, vol. 74, no. 3, pp. 897–935, 2007.
- [47] M. Huang, P. E. Caines, and R. P. Malhamé, "Social optima in mean field LQG control: centralized and decentralized strategies," *Automatic Control, IEEE Transactions on*, vol. 57, no. 7, pp. 1736–1751, 2012.
- [48] W. Rudin, *Real and Complex Analysis*, 3rd ed. McGraw-Hill Inc., 1987.
- [49] Y. Achdou and I. Capuzzo-Dolcetta, "Mean field games: Numerical methods," *SIAM Journal on Numerical Analysis*, vol. 48, no. 3, pp. 1136–1162, 2010.
- [50] D. W. Stroock and S. S. Varadhan, *Multidimensional diffusion processes*. Springer Science & Business Media, 1979, vol. 233.



**Jerome Le Ny** (S'05-M'09) is an assistant professor in the Department of Electrical Engineering at the Ecole Polytechnique de Montréal since May 2012. He received the B.S. degree from the Ecole 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. 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, dynamic resource allocation problems, and privacy and security issues in cyber-physical systems, with applications to networked embedded control systems, autonomous multi-robot systems, and transportation systems.



**Rabih Salhab** is currently pursuing the Ph.D. degree in Electrical Engineering in the Department of Electrical Engineering, Ecole Polytechnique de Montréal, Canada. He received the B.S. degree in Electrical Engineering from Ecole Supérieure d'Ingénieurs de Beyrouth (E.S.I.B), Lebanon, in 2008. 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, 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*