Consensus and Disagreement in Collective Homing Problems: A Mean Field Games Formulation

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 mean field games (MFG) scenario producing decentralized homing decisions in large multi-agent systems. For our setup, we show that given an initial distribution of the agents, many strategies exist, with each one of them defining an $\epsilon-$Nash equilibrium. These strategies, on which the processes of consensus and disagreement within the group depend, collapse into one strategy as the number of agents goes to infinity.**

I. Introduction

Collective decision making in large groups occurs when individuals agree on a choice among several alternatives. This phenomenon exists in most social aggregations, e.g., in economic systems, biological populations [1], [2], or human societies. Important examples of successful biological collective decision making mechanisms include honey bees searching for a new colony [3], [4], the collective navigation of fish schools [5], [6], or quorum sensing [7].

Two important properties characterize such biological systems. The first is the aggregation property where, despite their inherent selfishness, the agents need to remain close, with the primary benefits being protection from predators (in the case of fish shoals) or enhanced foraging ability (in the case of honey bees). The second is the decentralized control of the agents' behaviors, which seems to be highly developed in advanced eusocial species (species hierarchically organized) of population on the order of 10^3 or larger, whereas centralized control is adopted in primitive eusocial species of population on the order of 10^2 or less [3]. The reason behind distributing the control among the individuals of a large population is to minimize communication and computational requirements and to maintain cohesiveness of the group in the face of potentially selfish behavior of some of its individuals.

In this paper, we consider a situation where a large number of agents, initially spread out in $\mathbb{R}^n$, need to move within a finite time horizon to one of two possible home or target destinations. They must do so while trying to remain tightly grouped, and expending as little control effort as possible. For example, in a navigation situation for a collection of micro robots exploring an unknown terrain, remaining grouped may be necessary for achieving coordinated collective tasks [8]–[11]. In animal collective navigation, staying within a large group offers better protection against predators. Finally, this model may be an abstract representation of opinion crystallization in an election where (i) relative distances measure current differences of opinions, (ii) individuals are sensitive to collective opinion swings, and (iii) a choice must be made before a finite deadline [12]–[14].

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

II. Mathematical Model

In this section, we present the mathematical formulation of our problem. Consider N agents, with identical and independent linear dynamics

$$\dot{x}_i = Ax_i + Bu_i, \qquad x_i^0 \in \mathbb{R}^n,\, u_i \in \mathbb{R}^m, \tag{1}$$

for $i = 1, \dots, N$. These agents must migrate in a finite time interval $[0, T]$, from their initial positions toward one of two predefined destinations $p_a, p_b \in \mathbb{R}^n$, while minimizing the cost

$$J_i(u_i, \bar{x}, x_i^0) = \int_0^T \left\{ \frac{q}{2}\|x_i - \bar{x}\|^2 + \frac{r}{2}\|u_i\|^2 \right\} \mathrm{d}t + \frac{M}{2} \min\left(\|x_i(T) - p_a\|^2, \|x_i(T) - p_b\|^2\right) \tag{2}$$

with $q, r, M > 0$. The running cost in (2) penalizes the control effort as well as individual deviations from the population mean

$$\bar{x} := \frac{1}{N}\sum_{i=1}^N x_i.$$

Moreover, the terminal cost strongly penalizes deviations from either p_a or p_b (M is large).

The agents minimizing (2) are cost coupled. The optimal control law u_i^* of each agent depends on the state of the population $(x_1, \dots, x_N)$. Hence, the exact optimal solution of (2) requires compilcated centralized data when N is large. Alternatively, by working within the framework of the Mean Field Games (MFG) theory [8], [15]–[19], we develop decentralized strategies where each agent only needs to know the initial distribution of the population and its own state. The resulting solutions are $\epsilon-$Nash equilibria instead of exact Nash solutions.

As described in [15], [16], [19], the MFG approach starts by approximating the mass behaviour $\bar{x}$ by an assumed known function x^*. This unknown trajectory x^* is calculated by requiring that it can be replicated by the mean of all agents when they optimally track it.

The rest of the paper is organized as follows. In Section III we develop the solution to the general tracking problem for any continuous path x^*. We establish that the choice of destinations p_a or p_b by a given agent is dictated by its initial condition and two basins of attraction exist. In Section IV we characterize these basins of attraction. In Section V we tackle the problem of identifying conditions for the existence of a trajectory x^* reproducing a mean $\bar{x} = x^*$. We also prove that the strategies developed when tracking these trajectories constitute an $\epsilon-$Nash equilibrium. In Section VI we provide some numerical simulation results, while Section VII presents our conclusions.

III. General Tracking Problem

Following the MFG approach, in this section we start by solving the following tracking problem. The N agents with dynamics (1) minimize the cost function

$$J_i(u_i, x^*, x_i^0) = \int_0^T \left\{\frac{q}{2}\|x_i - x^*\|^2 + \frac{r}{2}\|u_i\|^2\right\} \mathrm{d}t + \frac{M}{2}\min\Big(\|x_i(T) - p_a\|^2, \|x_i(T) - p_b\|^2\Big) \quad (3)$$

which corresponds to (2) with $\bar{x}$ replaced by some assumed known continuous path x^*. Note that the cost function J_i can be written

$$J_i(u_i, x^*, x_i^0) = \min\Big(J_i^a(u_i, x^*, x_i^0), J_i^b(u_i, x^*, x_i^0)\Big),$$

where

$$J_i^e(u_i, x^*, x_i^0) = \int_0^T \left\{\frac{q}{2}\|x_i - x^*\|^2 + \frac{r}{2}\|u_i\|^2\right\} \mathrm{d}t + \frac{M}{2}\|x_i(T) - p_e\|^2,$$

for $e = a$ or b. Moreover, we have

$$\inf_{u_i(.)} J_i(u_i, x^*, x_i^0) = \min\left(\inf_{u_i(.)} J_i^a(u_i, x^*, x_i^0), \inf_{u_i(.)} J_i^b(u_i, x^*, x_i^0)\right).$$

Assuming a full state feedback, the optimal control for (3) is

$$u_i^* = \begin{cases} u_i^a & \text{if } J_i^a(u_i^a, x^*, x_i^0) \le J_i^b(u_i^b, x^*, x_i^0) \\ u_i^b & \text{if } J_i^a(u_i^a, x^*, x_i^0) > J_i^b(u_i^b, x^*, x_i^0), \end{cases}$$

where u_i^a and u_i^b are the optimal solutions of the simple linear quadratic tracking problems with cost functions J_i^a and J_i^b, for which we recall [20] the optimal control laws

$$u_i^e(t) = -\frac{1}{r}B^T\Big(\alpha(t)x_i + \beta^e(t)\Big), \qquad \forall e \in \{a, b\},$$

and the corresponding optimal costs

$$J_i^{e,*}(x^*, x_i^0) = \frac{1}{2}(x_i^0)^T\alpha(0)x_i^0 + \beta^e(0)^T x_i^0 + \delta^e(0),$$

where $\alpha \in C([0,T], \mathbb{R}^{n\times n})$, $\beta^e \in C([0,T], \mathbb{R}^n)$ and $\delta^e \in C([0,T], \mathbb{R})$ are the unique solutions of

$$\dot{\alpha} - \frac{1}{r}\alpha BB^T\alpha + \alpha A + A^T\alpha + qI_n = 0 \quad (4a)$$

$$\dot{\beta^e} = (\frac{1}{r}\alpha BB^T - A^T)\beta^e + qx^* \quad (4b)$$

$$\dot{\delta^e} = \frac{1}{2r}(\beta^e)^T BB^T\beta^e - \frac{1}{2}q\, x^{*T}x^*, \quad (4c)$$

with the final conditions

$$\alpha(T) = MI_n, \;\; \beta^e(T) = -Mp_e, \;\; \delta^e(T) = \frac{1}{2}Mp_e^T p_e.$$

We summarize the above analysis in the following lemma.

Lemma 1: The tracking problem (3) has a unique optimal control function

$$u_i^*(t) = \begin{cases} -\frac{1}{r}B^T\Big(\alpha(t)x_i + \beta^a(t)\Big) & \text{if } x_i^0 \in D_a(x^*) \\ -\frac{1}{r}B^T\Big(\alpha(t)x_i + \beta^b(t)\Big) & \text{if } x_i^0 \notin D_a(x^*) \end{cases} \quad (5)$$

where α, β^e, δ^e are the unique solutions of (4a)-(4c) for $e = a, b$, and

$$D_a(x^*) = \Big\{x \in \mathbb{R}^n; \big(\beta^a(0) - \beta^b(0)\big)^T x \le \delta^b(0) - \delta^a(0)\Big\} \quad (6)$$

Hence, given any continuous path x^*, there exists a basin of attraction $D_a(x^*)$ where all the agents initially present in this region prefer going toward p_a whereas the others prefer going toward p_b. Therefore, the mean of the population is highly dependent on $D_a(x^*)$. In the next section we study the properties of this basin in more details.

Remark 1: We conventionally impose p_a as a destination for the agents initially present on the boundary of $D_a(x^*)$.

IV. Basin of Attraction

We start by giving an explicit solution of (4b) and (4c). Let $K(t) = \frac{1}{r}\alpha(t)BB^T - A^T$ and ϕ_K be the state-transition matrix of (4b). Thus,

$$\beta^e(t) = -M\phi_K(t,T)p_e + q\int_T^t \phi_K(t,\sigma)x^*(\sigma)\,\mathrm{d}\sigma$$

$$\begin{aligned}\delta^e(t) = {} & \frac{1}{2}Mp_e^T p_e - \frac{q}{2}\int_T^t (x^*)^T(\sigma)x^*(\sigma)\,\mathrm{d}\sigma \\ & + \frac{M^2}{2r}p_e^T\int_T^t \left\{\begin{matrix}\phi_K^T(\eta,T)B \\ B^T\phi_K(\eta,T)\end{matrix}\right\}\mathrm{d}\eta\, p_e \\ & - \frac{Mq}{r}p_e^T\int_T^t\int_T^\eta \left\{\begin{matrix}\phi_K^T(\eta,T)BB^T \\ \phi_K(\eta,\sigma)x^*(\sigma)\end{matrix}\right\}\mathrm{d}\sigma\mathrm{d}\eta \\ & + \frac{q^2}{2r}\int_T^t\int_T^\eta\int_T^\eta \left\{\begin{matrix}(x^*)^T(\sigma)\phi_K^T(\eta,\sigma)B \\ B^T\phi_K(\eta,\tau)x^*(\tau)\end{matrix}\right\}\mathrm{d}\tau\mathrm{d}\sigma\mathrm{d}\eta\end{aligned} \quad (7)$$

Finally, by replacing (7) in the expression of $D_a(x^*)$, (6) can be written

$$D_a(x^*) = \Big\{x \in \mathbb{R}^n; \beta_0^T x \le \delta_0 + \delta_1(x^*)\Big\} \quad (8)$$

where

$$\begin{aligned}\beta_0 &= M\phi_K(0,T)(p_b - p_a) \\ \delta_0 &= \frac{1}{2}Mp_b^T p_b - \frac{1}{2}Mp_a^T p_a \\ &\quad + \frac{M^2}{2r}p_b^T\int_T^0 \phi_K^T(\eta,T)BB^T\phi_K(\eta,T)\,\mathrm{d}\eta\, p_b \\ &\quad - \frac{M^2}{2r}p_a^T\int_T^0 \phi_K^T(\eta,T)BB^T\phi_K(\eta,T)\,\mathrm{d}\eta\, p_a \\ \delta_1(x^*) &= \frac{Mq}{r}(p_a^T - p_b^T)\int_T^0\int_T^\eta \left\{\begin{matrix}\phi_K^T(\eta,T)BB^T \\ \phi_K(\eta,\sigma)x^*(\sigma)\end{matrix}\right\}\mathrm{d}\sigma\mathrm{d}\eta\end{aligned} \quad (9)$$

V. Fixed Points and Nash Equilibria

Having solved the general tracking problem, we now seek a continuous path x^* that can be replicated by the mean of all agents when they optimally track it. We start our search by computing the dynamics of the mean $\bar{x}$ when tracking any continuous path x^*. We then show that this mean is the image of x^* by a map T_λ, where λ is the number of agents initially in $D_a(x^*)$ and T_λ is an element of a more general familly of maps $(T_k)_{k\in\{1,\dots,N\}}$. Based on this analysis, the sought path x^* is a fixed point of T_λ, where λ is the number of agents initially in $D_a(x^*)$. In order to identify such a path, we first prove in Lemma 2 the existence and uniqueness of a fixed point of any $T_k \in (T_k)_{k\in\{1,\dots,N\}}$ and we derive its explicit form. Then, we define in Theorem 3 a necessary and sufficient condition for the existence of the desired path. This condition is a direct consequence of the following observation: The mean is a fixed point of T_λ if and only if λ is the number of agents initially in $D_a(x^*)$. Finally, we prove in Theorem 4 that the control strategies developed while tracking this path constitute an $\epsilon-$Nash equilibrium.

We consider $x^* \in C([0,T],\mathbb{R}^n)$. By Lemma 1, there exists a region $D_a(x^*)$ such that, while tracking x^*, the agents initially present in this region select the control law $-\frac{1}{r}B^T(\alpha x+\beta^a)$, whereas the others select the control law $-\frac{1}{r}B^T(\alpha x+\beta^b)$. Suppose that initially λ agents are in $D_a(x^*)$. The dynamics of the mean

$$\dot{\bar{x}} = -K^T\bar{x} - \frac{q}{r}BB^T\int_T^t \phi_k(t,\sigma)x^*(\sigma)\,\mathrm{d}\sigma + \frac{M}{r}BB^T\phi_K(t,T)p_\lambda \quad (10)$$

with $\bar{x}(0)=\bar{x}_0, p_\lambda = \frac{\lambda}{N}p_a + \frac{N-\lambda}{N}p_b$, is obtained by substituting (7) in (5) and the resulting control law in (1) to subsequently compute $\bar{x}=\frac{1}{N}\sum_{i=1}^N x_i$ and its derivative. We define for any $\lambda\in\{0,\dots,N\}$ a map T_λ from $C([0,T],\mathbb{R}^n)$ to $C([0,T],\mathbb{R}^n)$, where $T_\lambda(x^*)$ is the unique solution of (10). Hence, $\bar{x}$ is the image of x^* by T_λ where λ is the number of agents initially in $D_a(x^*)$. The next lemma establishes that for any $\lambda\in\{0,\dots,N\}$, T_λ has a unique fixed point.

Lemma 2: Consider $\lambda\in\{0,\dots,N\}$. T_λ has a unique fixed point equal to

$$R_1(t)\bar{x}_0 + R_2(t)p_\lambda \quad (11)$$

where $\tilde{R}_1 = \phi_K^T(t,0)R_1(t)$ and $\tilde{R}_2 = \phi_K^T(t,0)R_2(t)$ are the unique solutions of

$$\begin{aligned}\dot{\tilde{R}}_1 &= -\frac{q}{r}\int_T^t \begin{Bmatrix}\phi_K^T(t,0)BB^T\phi_K(t,\sigma)\\ \phi_K^T(0,\sigma)\tilde{R}_1(\sigma)\end{Bmatrix}\mathrm{d}\sigma\\ \dot{\tilde{R}}_2 &= -\frac{q}{r}\int_T^t \begin{Bmatrix}\phi_K^T(t,0)BB^T\phi_K(t,\sigma)\\ \phi_K^T(0,\sigma)\tilde{R}_2(\sigma)\end{Bmatrix}\mathrm{d}\sigma \\ &\quad+\frac{M}{r}\phi_K^T(t,0)BB^T\phi_K(t,T)\end{aligned} \quad (12)$$

with initial conditions $\tilde{R}_1(0)=I_n,\ \tilde{R}_2(0)=0$.

Proof: Let $W_\lambda = L_2\circ T_\lambda\circ L_1$, where L_1 and L_2 are operators from $C([0,T],\mathbb{R}^n)$ to itself such that $\forall x\in C([0,T],\mathbb{R}^n)$, $L_1(x)(t) = \phi_K^T(0,t)x(t)$ and $L_2(x)(t) = \phi_K^T(t,0)x(t)$. T_λ has a unique fixed point if and only if W_λ has. Let $x\in C([0,T],\mathbb{R}^n)$,

$$W_\lambda(x)(t) = g(t)p_\lambda + \bar{x}_0 + \int_0^t\int_T^\tau f(\tau,\sigma)x(\sigma)\,\mathrm{d}\sigma\mathrm{d}\tau$$

where $f(t,\sigma) = -\frac{q}{r}\phi_K^T(t,0)BB^T\phi_K(t,\sigma)\phi_K^T(0,\sigma)$ and $g(t)=\frac{M}{r}\int_0^t \phi_K^T(\tau,0)BB^T\phi_K(\tau,T)\,\mathrm{d}\tau$. Let $\|.\|_\infty$ be the sup norm on $C([0,T],\mathbb{R}^{n\times n})$. We define on the Banach space $\big(C([0,T],\mathbb{R}^{n\times n}),\|.\|_\infty\big)$

$$\begin{aligned}R_{1k} &= I_n + \int_0^t\int_T^\tau f(\tau,\sigma)\,\mathrm{d}\sigma\mathrm{d}\tau + \\ &\int_0^t\int_T^\tau\int_0^\sigma\int_T^{\tau_1} f(\tau,\sigma)f(\tau_1,\sigma_1)\,\mathrm{d}\sigma_1\mathrm{d}\tau_1\mathrm{d}\sigma\mathrm{d}\tau + \\ &\vdots\\ &\int_0^t\int_T^{\tau_0}\cdots\int_0^{\sigma_{k-1}}\int_T^{\tau_k}\prod_{i=0}^k f(\tau_i,\sigma_i)\,\mathrm{d}\sigma_k\mathrm{d}\tau_k\ldots\mathrm{d}\sigma_0\mathrm{d}\tau_0\end{aligned}$$

$$\begin{aligned}R_{2k} &= g(t) + \int_0^t\int_T^\tau f(\tau,\sigma)g(\sigma)\,\mathrm{d}\sigma\mathrm{d}\tau + \\ &\int_0^t\int_T^\tau\int_0^\sigma\int_T^{\tau_1} f(\tau,\sigma)f(\tau_1,\sigma_1)g(\sigma_1)\,\mathrm{d}\sigma_1\mathrm{d}\tau_1\mathrm{d}\sigma\mathrm{d}\tau + \\ &\vdots\\ &\int_0^t\int_T^{\tau_0}\cdots\int_0^{\sigma_{k-1}}\int_T^{\tau_k}\prod_{i=0}^k f(\tau_i,\sigma_i)g(\sigma_k)\,\mathrm{d}\sigma_k\mathrm{d}\tau_k\ldots\mathrm{d}\sigma_0\mathrm{d}\tau_0\end{aligned}$$

Let $S = \max_{(t,\sigma)\in[0,T]^2}\|f(t,\sigma)\|$

$$\begin{aligned}&\|R_{1l}-R_{1k}\|_\infty \le\\ &\quad S^{k+2}\int_0^t\int_0^T\cdots\int_0^{\sigma_k}\int_0^T \mathrm{d}\sigma_{k+1}\mathrm{d}\tau_{k+1}\ldots\mathrm{d}\sigma\mathrm{d}\tau\\ &\quad\vdots\\ &\quad + S^{l+1}\int_0^t\int_0^T\cdots\int_0^{\sigma_{l-1}}\int_0^T \mathrm{d}\sigma_l\mathrm{d}\tau_l\ldots\mathrm{d}\sigma\mathrm{d}\tau\end{aligned}$$

$$\Rightarrow \|R_{1l}-R_{1k}\|_\infty \le \frac{S(STt)^{k+1}}{(k+1)!} + \cdots + \frac{S(STt)^l}{l!}$$

$$\Rightarrow \lim_{k,l\to+\infty}\|R_{1l}-R_{1k}\|_\infty = 0$$

Hence, R_{1k} is a Cauchy sequence in the Banach space $(C([0,T],\mathbb{R}^{n\times n}),\|.\|_\infty)$, therefore has a limit $\tilde{R}_1$. Using similar arguments, we prove that R_{2k} has a limit $\tilde{R}_2$. Let $y(t) = \tilde{R}_1(t)\bar{x}_0 + \tilde{R}_2(t)p_\lambda$. y is a fixed point of W_λ, and every fixed point of W_λ is equal to y. Hence, y is the unique fixed point of W_λ, and $x = L_1(y)$ is the unique fixed point of T_λ. By the same technique used to prove the existence and uniqueness of the fixed point of W_λ, we prove that $\tilde{R}_1$ and $\tilde{R}_2$ are the unique solutions of (12). ∎

Suppose now that the desired path x^* exists. This path is equal to the mean $\bar{x}$. Suppose that there exists initially λ

agents in $D_a(x^*)$, therefore $x^* = \bar{x}$ is the unique fixed point of T_λ. By Lemma 2, we know the form of x^*. By combining these facts, we elaborate in the next theorem a necessary and sufficient condition on the number of agents initially in $D_a(x^*)$ for the existence of the desired path. We define

$$\theta_1 = \frac{Mq}{r}(p_a^T - p_b^T)\times \int_T^0 \int_T^\eta \left\{ \begin{matrix} \phi_K^T(\eta,T)BB^T\phi_K(\eta,\sigma) \\ \Big(R_1(\sigma)\bar{x}_0 + R_2(\sigma)p_b\Big) \end{matrix} \right\} \mathrm{d}\sigma\mathrm{d}\eta$$

$$\theta_2 = \frac{Mq}{Nr}(p_a^T - p_b^T)\times \int_T^0 \int_T^\eta \left\{ \begin{matrix} \phi_K^T(\eta,T)BB^T\phi_K(\eta,\sigma) \\ R_2(\sigma)(p_a - p_b) \end{matrix} \right\} \mathrm{d}\sigma\mathrm{d}\eta$$

In order to facilitate the remaining analysis, we index the agents going toward p_a by numbers lower than those given for agents going toward p_b as follows.

$$\beta_0^T x_1^0 \le \beta_0^T x_2^0 \le \cdots \le \beta_0^T x_N^0 \tag{13}$$

Theorem 3: A path x^* that can be replicated by the mean of all agents optimally tracking it exists if and only if $\exists \lambda \in \{0, ..., N\}$ such that:

$$\beta_0^T x_\lambda^0 - \delta_0 - \theta_1 \le \lambda\theta_2 < \beta_0^T x_{\lambda+1}^0 - \delta_0 - \theta_1 \tag{14}$$

x^* is in this case the unique fixed point of T_λ.

Proof: Suppose that there exists a path x^* that can be replicated by the mean of all agents when optimally tracking it ($\bar{x} = x^*$). Hence, $\bar{x}$ is the image of x^* by T_λ where $\lambda \in \{0, ..., N\}$ is the number of agents in $D_a(x^*)$. But $\bar{x} = x^*$, therefore x^* is the unique fixed point of T_λ which is given by (11). By replacing (11) in the third equation in (9) we obtain $\delta_1(x^*) = \theta_1 + \lambda\theta_2$. By the indexing adopted in (13), the first λ agents are in D_a and the others outside. Hence, by the definition of $D_a(x^*)$ (8) we obtain:

$$\beta_0^T x_\lambda^0 \le \delta_0 + \delta_1(x^*)$$
$$\beta_0^T x_{\lambda+1}^0 > \delta_0 + \delta_1(x^*)$$
$$\Rightarrow \beta_0^T x_\lambda^0 - \delta_0 - \theta_1 \le \lambda\theta_2 < \beta_0^T x_{\lambda+1}^0 - \delta_0 - \theta_1$$

Suppose now that $\exists \lambda \in \{0, ..., N\}$ satisfying (14). We define x^* as the unique fixed point of T_λ which is given by (11).

$$\Rightarrow \delta_1(x^*) = \theta_1 + \lambda\theta_2$$

By (13) and (14)

$$\forall n \le \lambda,\ \beta_0^T x_n^0 - \delta_0 - \delta_1(x^*) \le \beta_0^T x_\lambda^0 - \delta_0 - \delta_1(x^*) \le 0$$

$$\forall n \ge \lambda+1,\ \beta_0^T x_n^0 - \delta_0 - \delta_1(x^*) \ge \beta_0^T x_{\lambda+1}^0 - \delta_0 - \delta_1(x^*) > 0$$

Hence, by (8) there exists initially λ agents in $D_a(x^*)$. Therefore $\bar{x} = T_\lambda(x^*)$. But x^* is the unique fixed point of T_λ, hence $\bar{x} = x^*$. ■

We developed in the above analysis decentralized control strategies by approximating the trajectory of the average of the population $\bar{x}$ by an assumed known function x^* and we found explicitly the form of this function in Lemma 2. To implement these strategies each agent needs to know, prior start moving, the anticipated mean trajectory and, while moving, its own state. How are these strategies related to the inital global cost (2)? In the next theorem we show that instead of defining an optimal solution for (2), the decentralized strategies contitute an $\epsilon-$Nash equilibrium with respect to (2). This type of equilibria makes the group's behaviour robust in the face of potential selfish behaviours. Indeed, in a decentralized mode, choosing strategies other than those defining an $\epsilon-$Nash-equilibrium is not profitable (for small ϵ) as the next definition shows [15].

Definition 1: Consider N players, a set of strategy profiles $S = S_1 \times \cdots \times S_N$ and for each player k a payoff function $J_k(u_1, \dots, u_N)$, $\forall (u_1, \dots, u_N) \in S$. A strategy $(u_1^*, \dots, u_N^*) \in S$ is called an $\epsilon-$Nash equilibrium with respect to the costs J_k, if there exists an $\epsilon > 0$ such that for any fixed $1 \le i \le N$, $\forall u_i \in S_i$, we have

$$J_i(u_1^*, \dots, u_{i-1}^*, u_i, , u_{i+1}^*, \dots, u_N^*) \ge J_i(u_1^*, \dots, u_N^*) - \epsilon.$$

We define the following technical hypothesis necessary for the next proof.

Hypothesis 1 (H-1): $\forall i \in \{1, ..., N\}$; $\|x_i^0\| \le Z$, where $Z > 0$ is independent of N.

Theorem 4: Suppose that H-1 holds. Suppose that $\exists \lambda \in \{0, ..., N\}$ satisfying (14). Let Σ be the set of decentralized controls that generates a fixed point of T_λ (i.e. solution of the optimal tracking problem (3), the tracked path being equal to the mean). We denote the elements of Σ by u_i^*, $i \in \{1, ..., N\}$. Then Σ is an $\epsilon-$Nash equilibrium with respect to the costs $J_i(u_i, \frac{1}{N}\sum_{j=1}^N x_j(u_j), x_i^0)$ where $\epsilon = o(\frac{1}{N})$.

Proof: The fixed point $x^* = \bar{x}$ of T_λ is given by (11). By H-1, $\|\bar{x}_0\| \le Z$. We also have $\|p_\lambda\| \le \|p_a\| + \|p_b\|$. By continuity of R_1 and R_2 on $[0, T]$, we have $\|R_1\|_\infty \le M_1$ and $\|R_2\|_\infty \le M_2$ with M_1 and M_2 independents of N. Hence, $\|\mu\|_\infty \le Q_1$, Q_1 independent of N. Depending on their initial positions, the state and control law of each agent are

$$x_i(u_i^*) = \phi_K^T(0,t)x_i^0 + \frac{M}{r}\int_0^t \phi_K^T(\sigma,t)BB^T\phi_K(\sigma,T)p_e\,\mathrm{d}\sigma - \frac{q}{r}\int_0^t \int_T^\sigma \phi_K^T(\sigma,t)BB^T\phi_K(\sigma,\tau)x^*(\tau)\,\mathrm{d}\tau\mathrm{d}\sigma$$

$$u_i^* = -\frac{1}{r}B^T\Big(\alpha x_i(u_i^*) - M\phi_K(t,T)p_e + q\int_T^t \phi_K(t,\sigma)x^*(\sigma)\,\mathrm{d}\sigma\Big)$$

where $e \in \{a, b\}$. The continuity on $[0, T]$ implies $\|x_i(u_i^*)\|_\infty \le Q_2$ and $\|u_i^*\|_\infty \le Q_3$ with Q_2 and Q_3 independents of N. The boundedness of u_i^*, $x_i(u_i^*)$ and x^* implies $J_i(u_i^*, \frac{1}{N}\sum_{j=1}^N x_j(u_j^*), x_i^0) \le Q_4$ with Q_4 independent of N.

Consider $i \in \{1,...N\}$, u_i an arbitrary complete state feedback control law for the agent i. Suppose that

$$J_i\left(u_i, \frac{1}{N}\sum_{j=1, j\neq i}^{N} x_j(u_j^*) + \frac{1}{N}x_i(u_i), x_i^0\right) \leq J_i\left(u_i^*, \frac{1}{N}\sum_{j=1}^{N} x_j(u_j^*), x_i^0\right)$$

$$\Rightarrow J_i\left(u_i, \frac{1}{N}\sum_{j=1, j\neq i}^{N} x_j(u_j^*) + \frac{1}{N}x_i(u_i), x_i^0\right) \leq Q_4$$

$$\Rightarrow \|u_i\| \leq Q_5$$

with Q_5 independent of N. By the boundedness of $\|u_i\|$ and initial positions, we have $\|x_i(u_i)\| \leq Q_6$ with Q_6 independent of N. The cost that the agent i pays when switching to alternative startegy to u_i is

$$\begin{aligned}
&J_i\left(u_i, \frac{1}{N}\sum_{j=1, j\neq i}^{N} x_j(u_j^*) + \frac{1}{N}x_i(u_i), x_i^0\right) \\
&= \frac{q}{2}\int_0^T \left\|x_i(u_i) - \frac{1}{N}\sum_{j=1, j\neq i}^{N} x_j(u_j^*) - \frac{1}{N}x_i(u_i)\right\|^2 \mathrm{d}t \\
&+\frac{r}{2}\int_0^T \|u_i\|^2 \,\mathrm{d}t + \frac{M}{2}\min\left(\|x_i(T)-p_a\|^2, \|x_i(T)-p_b\|^2\right) \\
&= J_i\left(u_i, \frac{1}{N}\sum_{j=1}^{N} x_j(u_j^*), x_i^0\right) + \frac{q}{2N^2}\int_0^T \|x_i(u_i) - x_i(u_i^*)\|^2 \,\mathrm{d}t \\
&+\frac{q}{N}\int_0^T \left(x_i(u_i^*) - x_i(u_i)\right)^T \left(x_i(u_i) - \frac{1}{N}\sum_{j=1}^{N} x_j(u_j^*)\right) \mathrm{d}t.
\end{aligned}$$

By optimality we have

$$J_i\left(u_i, \frac{1}{N}\sum_{j=1}^{N} x_j(u_j^*), x_i^0\right) \geq J_i\left(u_i^*, \frac{1}{N}\sum_{j=1}^{N} x_j(u_j^*), x_i^0\right).$$

By boundedness of $x_i(u_i)$, $x_i(u_i^*)$ and x^* we have

$$\left\|\frac{q}{N}\int_0^T (x_i(u_i^*) - x_i(u_i))^T (x_i(u_i) - \frac{1}{N}\sum_{j=1}^{N} x_j(u_j^*))\,\mathrm{d}t\right\| \leq \frac{Q_7}{N}$$

with Q_7 independent of N. Finally, we deduce

$$J_i\left(u_i, \frac{1}{N}\sum_{j=1, j\neq i}^{N} x_j(u_j^*) + \frac{1}{N}x_i(u_i), x_i^0\right) \geq J_i\left(u_i^*, \frac{1}{N}\sum_{j=1}^{N} x_j(u_j^*), x_i^0\right) - \frac{Q_7}{N}$$

Hence, Σ is an $\epsilon-$Nash equilibrium with respect to the costs $J_i(u_i, \frac{1}{N}\sum_{j=1}^{N} x_j(u_j), x_i^0)$, where $\epsilon = \frac{Q_7}{N} = o(\frac{1}{N})$. ■

Theorem 5: Suppose that there exists N_0 such that $\forall N \geq N_0$, $\max_{\lambda} \|x_{\lambda+1}^0 - x_\lambda^0\| \leq k\frac{1}{N}$, where $k = \frac{N|\theta_2|}{2\|\beta_0\|}$ is independent of N (i.e. the maximum inter-agent distance is bounded by $O(1/N)$). Then $\forall N \geq N_0$ at most one $\epsilon-$Nash equilibrium exists.

Proof: Let $a_N(\lambda) = \frac{1}{\theta_2}\left(\beta_0^T x_\lambda^0 - \delta_0 - \theta_1\right)$. $\forall N \geq N_0$, $|a_N(\lambda+1) - a_N(\lambda)| \leq 1/2, \forall \lambda \in \{1, \ldots, N-1\}$. Consider a population of size $N \geq N_0$. Suppose that there exists at least one $\epsilon-$Nash equilibrium. Let λ_0 be the smallest λ satisfying (14) (λ_0 is well defined because it was supposed that at least one $\epsilon-$Nash equilibrium exists). λ_0 satisfies $a_N(\lambda_0) \leq \lambda_0 < a_N(\lambda_0+1)$. Suppose that there exists λ_0+i satisfying (14) for some integer $i \geq 0$. $\lambda_0 + i$ satisfies $a_N(\lambda_0 + i) \leq \lambda_0 + i < a_N(\lambda_0 + i + 1)$. Therefore, $a_N(\lambda_0 + i + 1) > \lambda_0 + i \geq \lambda_0 \geq a_N(\lambda_0)$. Hence, $|a_N(\lambda_0 + i + 1) - a_N(\lambda_0)| > i$. But $|a_N(\lambda_0 + i + 1) - a_N(\lambda_0)| \leq \frac{i+1}{2}$. Hence, $i = 0$. Thus there exists at most one $\epsilon-$Nash equilibrium corresponding to λ satisfying (14). ■

Theorems 3 and 4 have three consequences. Firstly, inequality (14) implies that for any initial distribution of the agents, for any λ satisfying this inequality, there exists a unique $\epsilon-$Nash equilibrium in which λ agents decide to go toward p_a while the others toward p_b. Hence, given an initial distribution of the agents there might exist many $\epsilon-$Nash equilibria corresponding to different λs satisfying (14). Secondly, a consensus exists if $\lambda = 0$ or N satisfying (14). Finally, even though the developed control laws are decentralized, each agent needs to know the exact initial positions of other agents to compute the mean trajectory prior start moving. This fact can cause some problems when the number of agents tends to infinity or when just statistical data about the initial positions are available.

VI. Simulation Results

To illustrate the collective decision-making mechanisms, we consider a population of $N = 20$ agents with $A = B = I_2$, $p_a = -p_b = (-10, 0)$ and $T = 1$. We consider three cases where the population starts moving from three different initial distributions. In the first two cases, we set $q = r = 1$ and $M = 10000$. In the third case, we decrease M to 1000 and penalize more on the deviation from the mean by increasing q to 10. For the first case, two different control strategies generating two distinct $\epsilon-$Nash equilibria are possible, whereby for the first strategy 8 agents decide to go to p_a (Fig.1), and 12 agents choose the destination p_a in the second one (Fig.2). For the second case, only one strategy in which 5 agents decide to go to p_a can generate an $\epsilon-$Nash equilibrium (Fig.3). For the third case, a consensus occurs (Fig.4). Moreover, it should be noted that the mean perfectly replicates the tracked trajectory.

VII. Conclusion

We considered in this paper a large population dynamic game involving a binary choice, and showed that if the number of agents is finite then there might exist a priori

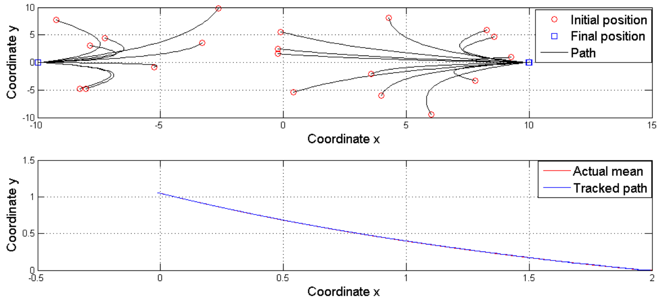

Fig. 1. First case: First $\epsilon-$Nash equilibrium

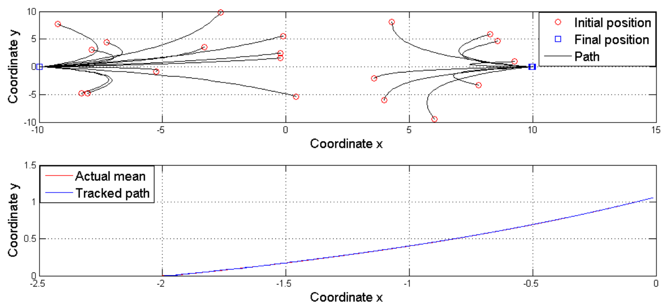

Fig. 2. First case: Second $\epsilon-$Nash equilibrium

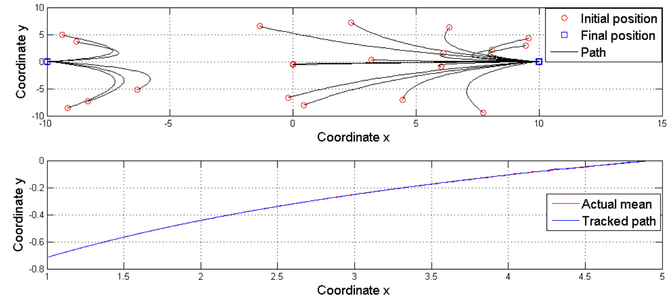

Fig. 3. Second case: The unique $\epsilon-$Nash equilibrium

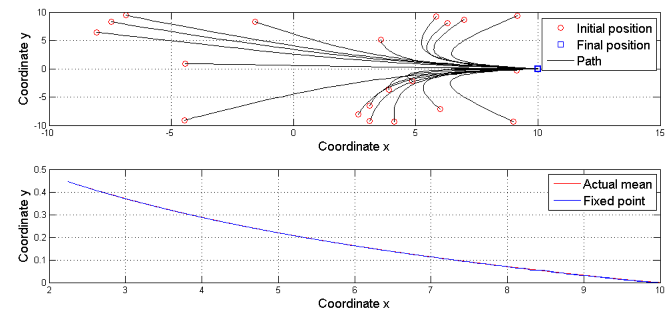

Fig. 4. Third case: Consensus

multiple $\epsilon-$Nash equlibria. Hence, in the absence of an a-priori convention on how to choose between the available potential equilibria, agents can choose freely between these and this prohibits the existence of an actual $\epsilon-$Nash equlibrium. When the number of agents goes to infinity, these equilibria collapse to one however. The "miscoordination" between the agents of low population groups, which can supress the existence of an actual $\epsilon-$Nash equilibrium, may be an additional cause for the absence of decentralized control strategies in primitively eusocial species of population on the order of 100 or less [3]. One approach for disambiguation is to assume that the majority will opt for the alternative corresponding to the lowest overall mean cost (i.e. the least damaging $\epsilon-$Nash equilibrium).

References

[1] 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.

[2] I. Couzin, J. Krause, N. Franks, and S. Levin, "Effective leadership and decision-making in animal groups on the move," *Nature*, vol. 433, pp. 513–516, 2005.

[3] T. 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.

[4] S. Camazine, P. Visscher, J. Finley, and R. Vetter, "House-hunting by honey bee swarms: collective decisions and individual behaviors," *Insectes Sociaux*, vol. 46, no. 4, pp. 348–360, November 1999.

[5] J. Tien, S. Levin, and D. Rubenstein, "Dynamics of fish schools: identifying key decision rules," *Evolutionary Ecology Research*, vol. 6, pp. 555–565, 2004.

[6] 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.

[7] 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. 15 906–15 910, 2006.

[8] M. Nourian, R. Malhamé, M. Huang, and P. 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.

[9] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, First, Ed. Princeton University Press, 2010.

[10] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009.

[11] J. Le Ny and G. Pappas, "Adaptive deployment of mobile robotic networks," *IEEE Transactions on automatic control*, vol. 58, pp. 654–666, 2013.

[12] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, vol. 1.1, pp. 3–49, 2010.

[13] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis, and simulation," *Journal of Artifical Societies and Social Simulation*, vol. 5, 2002.

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

[15] M. Huang, "Stochastic control for distributed systems with applications to wireless communications," Ph.D. dissertation, McGill University, 2003.

[16] M. Huang, P. Caines, and R. 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.

[17] M. Huang, R. Malhamé, and P. 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.

[18] J. Lasry and P. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, pp. 229–260, 2007.

[19] M. Huang, P. Caines, and R. 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.

[20] B. Anderson and J. Moore, *Linear Optimal Control*. Prentice-Hall, Inc., 1971.