Optimal Control of a Collective Migration Model

Benedetto Piccoli*, Nastassia Pouradier Duteil*, and Benjamin Scharf†

* Department of Mathematical Sciences, Rutgers University, Camden, NJ 08102, USA.
piccoli@camden.rutgers.edu, nastassia.pouradierduteil@rutgers.edu
† Technische Universität München, Fakultät Mathematik, Boltzmannstrasse 3 D-85748, Garching bei München, Germany. scharf@ma.tum.de

Abstract

Collective migration of animals in a cohesive group is rendered possible by a strategic distribution of tasks among members: some track the travel route, which is time and energy-consuming, while the others follow the group by interacting among themselves. In this paper, we study a social dynamics system modeling collective migration. We consider a group of agents able to align their velocities to a global *target velocity*, or to follow the group via interaction with the other agents. The balance between these two attractive forces is our control for each agent, as we aim to drive the group to consensus at the target velocity. We show that the optimal control strategies in the case of final and integral costs consist of controlling the agents whose velocities are the furthest from the target one: these agents sense only the target velocity and become *leaders*, while the uncontrolled ones sense only the group, and become *followers*. Moreover, in the case of final cost, we prove an "Inactivation" principle: there exist initial conditions such that the optimal control strategy consists of letting the system evolve freely for an initial period of time, before acting with full control on the agent furthest from the target velocity.

Introduction

A fascinating feature of large groups is their *self-organization* ability, i.e. the emergence from local interaction rules of certain global patterns. For instance, animal groups such as schools of fish, flocks of birds or herds of mammals exhibit strong coordination in their movements, see e.g. [1, 2, 5, 8, 9, 27, 28, 29, 34, 37, 38]. This collective behavior in animal groups also inspired applications to robotics (see [3]), in which the aim is to coordinate autonomous vehicles [7, 21, 25, 36] and flight formations [31, 35]. Other interests concern models in microbiology [19, 20, 22, 30, 32], pedestrian and crowd motions [10, 11] and financial markets [14, 23]. Such systems are usually referred to as social dynamics. Examples of self-organization include clustering of the agents, alignment of velocities, or other kinds of equilibria, see [5, 18, 26, 27, 28, 29, 37]. This raises the question of understanding the mechanisms behind the global pattern formation.

A well-known model was proposed by F. Cucker and S. Smale (see [12]) to describe the phenomenon of *consensus* in terms of alignment of velocities in a group on the move. The Cucker-Smale model in formula is written as:

$$\begin{cases} \dot{x_i} = v_i \\ \dot{v_i} = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} \dfrac{v_j - v_i}{(1+\|x_j - x_i\|^2)^{\beta}} \end{cases} \qquad \text{for } i \in \{1, ..., N\}, \tag{1}$$

where $\beta > 0$, and $x_i \in \mathbb{R}^d$ and $v_i \in \mathbb{R}^d$ are respectively the *state* and *velocity*. This model was originally designed to describe the formation and evolution of language, and the variables v_i can more generally represent opinions, preferences or invested capital. The system converges to consensus if $\beta \leq \frac{1}{2}$, which corresponds to a strong interaction even between distant agents, see [6]. On the other hand, if $\beta > \frac{1}{2}$, i.e. if the interaction is too weak, convergence to consensus only happens under certain conditions. More generally, the term $(1+\|x_j - x_i\|^2)^{-\beta}$ can be replaced by $a(\|x_j - x_i\|)$. Intuitively, it is natural to define a as a non-increasing function, since proximity often encourages interaction. On the other hand, it was proven that interactions modeled by non-decreasing functions a, called heterophilious, in fact enhance consensus (see [26]). When the system does not converge to a desired state, a natural question is to study the possibility of steering it via controls functions u_i, in which case the second equation of (1) becomes: $\dot{v_i} = \frac{1}{N}\sum_{j=1}^{N} a(\|x_j - x_i\|)(v_j - v_i) + u_i$ (see [6, 15]).

In the *collective migration* problem (see [24]), not only do agents interact with one another to travel as a group, but they also gather clues from the environment guiding them towards a global *target velocity*. In the case of migrating birds, for instance, this velocity can be sensed through a magnetic field, the direction of the sun, or environmental features. However, sensing the migration velocity is costly, both in used time and energy. A trade-off thus occurs between gathering this information, which ensures more precision, and following the group, which is less costly and saves time and energy for other tasks such as surveying for predators [13, 17]. This problem also applies to the field of robotics, in which gathering information from the environment is done at the expense of communicating with other robots (or planes, drones, etc.) or performing other tasks, and to the field of economics when one aims to influence decisions of a group based on limited information. This trade-off naturally separates the group into *leaders*, who gather information, and *followers*, who only interact with the other agents (see [17]).

We study a *Collective Migration Model*, where the agents' dynamics is determined by two forces: the attraction towards a target velocity V (which we assume can be sensed) and the consensus dynamics as in the Cucker-Smale model. More precisely, each agent's evolution is governed by a parameter $\alpha_i \in [0, 1]$ which provides the balance between the two forces. The system can be written as:

$$\begin{cases} \dot{x_i} = v_i \\ \dot{v_i} = \alpha_i(V - v_i) + (1 - \alpha_i)\dfrac{1}{N}\sum\limits_{j=1}^{N} a(\|x_j - x_i\|)(v_j - v_i) \end{cases} \qquad \text{for } i \in \{1, ..., N\}, \tag{2}$$

where: $x_i \in \mathbb{R}^d$ and $v_i \in \mathbb{R}^d$ are the state and velocity, $V \in \mathbb{R}^d$ is the target velocity, and $\alpha_i \in [0, 1]$ is the control, with the constraint $\sum_i \alpha_i \leq M$, $M > 0$. In this paper, we choose to set $a \equiv 1$, so that the strength of interaction does not depend on the agents' positions. This is a reasonable hypothesis for instance if we consider groups of planes or drones that can communicate just as easily from great distances.

While the Cucker-Smale model leads to alignment of all velocities to the average one (when there is consensus), the migration model tends to align all velocities to the preassigned *target velocity*. Our work focuses on finding optimal control strategies in order to achieve consensus to the target velocity, and in particular on selecting optimal *controlled leaders* among the agents when the control strength M is small with respect to the size of the group. In order to do that, we define the cost function $\tilde{\mathbb{V}} = \frac{1}{N}\sum_i \|v_i - V\|^2$, measuring the distance from consensus at the target velocity. We first show that, given any $M > 0$, the strategy to decrease $\tilde{\mathbb{V}}$ instantaneously, with the constraint $\sum_i \alpha_i \leq M$, consists of distributing the control among the agents with the largest positive projections of velocities along $\bar{v} - V$ (where $\bar{v}$ is the mean velocity). In particular, if $\langle v_i, \bar{v} - V\rangle < 0$, the agent i is not controlled ($\alpha_i = 0$).

We then study the optimal control strategy to minimize $\tilde{\mathbb{V}}$ at a fixed final time and first focus on the case of two agents, with control bounded by $M \in [0, 2]$. The optimal control strategies depend on M but, in all cases, we act with larger control on the agent with the largest projected velocity. Furthermore, if the final time is too short to bring the agents together, then initially the system must evolve with no control ($\alpha \equiv 0$). We call this phenomenon "Inactivation", in line with the "Inactivation Principle" proven in [16] in the case of muscle use. The latter claims that, during fast arm movements, it is optimal to simultaneously inactivate both agonistic and antagonistic muscles. We next generalize our results to any number of agents, but with the constraint $M \leq 1$. Then the optimal control strategy acts with full strength on a sub-group of agents to bring them together. Also in this case we observe "Inactivation", which occurs when the initial average velocity $\bar{v}$ is very close to the target velocity V. Indeed, to drive the system to V requires both achieving consensus and moving the average velocity towards V. If the average velocity is already close to V, then we are left with inducing consensus which happens naturally without control. However, simulations show that Inactivation is rare and its performance gain is hardly significant compared to a full-control strategy.

Then we move on to examine integral costs $\int_0^T \tilde{\mathbb{V}}(t)dt$ and show that the optimal control strategy never exhibits Inactivation. More precisely, we must use full control at all time splitting it evenly among the agents with the biggest projected velocity. Such a strategy is more restrictive than that with final cost, indeed the controls are completely determined by initial conditions, while previously we could use any strategy bringing agents together at final time.

The paper is organized as follows. In Section 1, we define the cost functional and make general observations. In Section 2, we determine the strategy to decrease it instantaneously in time. Then, in Section 3, we introduce the optimal control problem to minimize the cost function at a given final time. We first solve it for the particular case of two agents (Section 4) before generalizing to any number of agents with a control bounded by 1 (Section 5). Lastly, we find optimal control strategies to minimize the integral cost (Section 6).

1 Cost function and general observations

With no loss of generality, we set the target velocity V to zero. Having simplified the interaction function a, system (2) reduces to:

$$\begin{cases} \dot{x_i} = v_i \\ \dot{v_i} = -\alpha_i v_i + (1-\alpha_i)\dfrac{1}{N}\sum\limits_{j=1}^{N}(v_j - v_i) \end{cases} \qquad i \in \{1, ..., N\}. \tag{3}$$

Given $M > 0$, we define the set of controls $\mathcal{U}_M$ as:

$$\mathcal{U}_M = \Big\{ \alpha : [0,T] \to [0,1]^N \Big| \ \alpha \text{ measurable, s.t. for all } t, \ \sum_{i=1}^{N} \alpha_i(t) \leq M \Big\}. \tag{4}$$

1.1 Projection of the Dynamics

Note that the dynamics (3) can be written in the more compact way:

$$\begin{cases} \dot{x_i} = v_i \\ \dot{v_i} = -v_i + (1-\alpha_i)\,\bar{v} \end{cases} \tag{5}$$

where $\bar{v}$ represents the mean velocity $\bar{v} = \frac{1}{N}\sum_i v_i$. The evolution of $\bar{v}$ is given by $\dot{\bar{v}} = -\frac{1}{N}(\sum_i \alpha_i)\bar{v}$, so the direction of $\bar{v}$ is an invariant of the dynamics. Notice that if $\bar{v}(0) = 0$, then $\bar{v}(t) = 0$ for all $t \in [0,T]$. Hereafter, we shall assume that $\bar{v}(0) \neq 0$, since the case $\bar{v} \equiv 0$ would make the dynamics (5) trivial. We can then define the invariant unit vector $e = \frac{\bar{v}}{\|\bar{v}\|}$.

Let $w_i = v_i - \langle v_i, e \rangle\, e$ be the projection of v_i over $(\bar{v}^{\perp})$. Then

$$\dot{w_i} = -v_i + (1-\alpha_i)\,\bar{v} - \langle -v_i + (1-\alpha_i)\,\bar{v}, e\rangle\, e = -w_i.$$

Therefore the projection of v_i over $(\bar{v}^{\perp})$ decreases exponentially, independently of the controls α_i. Let us now define $\xi_i = \langle v_i, e\rangle$. Its evolution is given by: $\dot{\xi_i} = -\langle v_i, e\rangle + (1-\alpha_i)\langle \bar{v}, e\rangle = -\xi_i + (1-\alpha_i)\|\bar{v}\| = -\xi_i + (1-\alpha_i)\bar{\xi}$. In the following, we will only study the equations governing the evolution of the projected variables ξ_i:

$$\text{For all } i \in \{1, ..., N\}, \quad \dot{\xi_i} = -\xi_i + (1-\alpha_i)\bar{\xi} \tag{6}$$

where $\bar{\xi} = \frac{1}{N}\sum_j \xi_j$. This is a significant result: instead of studying a system evolving in $\mathbb{R}^{Nd}$, we consider a system in $\mathbb{R}$, thus greatly reducing the complexity of theoretical and numerical analyses. Hereafter we shall make the following hypothesis:

Hypothesis 1.
It holds $\bar{v}(0) \neq 0$ and $\xi_i(0) \geq \xi_{i+1}(0)$ for every $i \in \{1, ..., N-1\}$.

Proposition 1.1.
Having made Hyp. 1, it holds $\bar{v}(t) \neq 0$ and $\bar{\xi}(t) > 0$ for all $t \in [0,T]$.
Furthermore, let $\tau \in [0,T]$. If $\xi_i(\tau) \geq 0$, then $\xi_i(t) \geq 0$ for all $t \in [\tau, T]$. If $\xi_i(\tau) > 0$, then $\xi_i(t) > 0$ for all $t \in [\tau, T]$.

Proof. The proposition is mainly a consequence of Gronwall's inequality: It holds

$$\bar{\xi} = \frac{1}{N} \sum_j \langle v_j, \frac{\bar{v}}{\|\bar{v}\|} \rangle = \langle \bar{v}, \frac{\bar{v}}{\|\bar{v}\|} \rangle = \|\bar{v}\| \tag{7}$$

and

$$\dot{\bar{\xi}} = -\frac{1}{N} \left(\sum_{i=1} \alpha_i \right) \bar{\xi} \geq -\frac{M}{N} \bar{\xi}.$$

Hence, if $\bar{v}(0) \neq 0$ and therefore $\bar{\xi}(0) > 0$, then $\bar{\xi}(t) \geq e^{-Mt/N} \bar{\xi}(0) > 0$ and thus $\bar{v}(t) \neq 0$ for all $t \in [0, T]$. Now notice that from (6) we can compute for all $t \in [\tau, T]$: $\xi_i(t) = e^{-(t-\tau)}(\xi_i(\tau) + \int_\tau^t (1 - \alpha_i)(s) \bar{\xi}(s) e^{s-\tau} ds)$, so $\xi_i(t) \geq e^{-(t-\tau)} \xi_i(\tau)$, which proves the second part of the proposition. $\square$

1.2 Migration functional

We introduce the functional

$$\tilde{\mathbb{V}} = \frac{1}{N} \sum_{i=1}^{N} \|v_i - V\|^2, \tag{8}$$

which measures the distance from consensus at the desired velocity V. Since we set $V = 0$, $\tilde{\mathbb{V}}$ reduces to: $\tilde{\mathbb{V}} = \frac{1}{N} \sum_i \|v_i\|^2$. In the new projected coordinates ξ, the migration functional can be written as: $\tilde{\mathbb{V}} = \frac{1}{N} \sum_i (\|w_i\|^2 + \xi_i^2)$, where only the second term ξ_i^2 can be controlled. Hence, here onward we will only consider the controllable part of $\tilde{\mathbb{V}}$, which we denote $\mathbb{V}$:

$$\mathbb{V} = \frac{1}{N} \sum_{i=1}^{N} \xi_i^2. \tag{9}$$

Notice that $\mathbb{V}$ can be written as a sum of two terms:

$$\mathbb{V} = \bar{\xi}^2 + \frac{1}{N} \sum_{i=1}^{N} (\xi_i - \bar{\xi})^2, \tag{10}$$

which should be minimized simultaneously (where we remind that $\bar{\xi} = \frac{1}{N} \sum_i \xi_i$). Minimizing $\bar{\xi}^2$ (or $\bar{\xi}$, since according to Proposition 1.1, $\bar{\xi} > 0$) corresponds to steering the system as a whole to the desired velocity $V = 0$. On the other hand, minimizing $\frac{1}{N} \sum_i (\xi_i - \bar{\xi})^2$ corresponds to driving the system to consensus. However, the dynamics (6) of ξ_i show that if $\xi_i < 0$, decreasing $\bar{\xi}$ slows down the increase of ξ_i, resulting in a possible increase of $(\xi_i - \bar{\xi})^2$. Hence, minimizing $\mathbb{V}$ requires balancing the decrease of the two terms in (10).

1.3 Minimization problems

In the following sections, we will deal with the minimization of different quantities, in order to design a strategy for consensus at the migration velocity $V = 0$.

(i) The minimization of $\frac{d\mathbb{V}}{dt}$, i.e. the maximization of the instantaneous decrease of $\mathbb{V}$ (see Section 2).

(ii) The minimization of the cost at final time, $\mathbb{V}(T)$ (see Section 3).

(iii) The minimization of the integral cost $\int_0^T \mathbb{V}(t) dt$ (see Section 6).

In order to minimize *(ii)* $\mathbb{V}(T)$ and *(iii)* $\int_0^T \mathbb{V}(t) dt$, we will design an optimal control strategy using Pontryagin's maximum principle. The minimization of $\dot{\mathbb{V}}$, on the other hand, will not provide an optimal control.

2 Instantaneous Decrease

In this section we look for a control strategy maximizing the instantaneous decrease of $\mathbb{V}$.

$$\dot{\mathbb{V}} = \frac{2}{N}\sum_{i=1}^{N}\xi_i\dot{\xi}_i = \frac{2}{N}\left(\sum_{i=1}^{N}-\xi_i^2+\sum_{i=1}^{N}(1-\alpha_i)\bar{\xi}\xi_i\right) = -2\mathbb{V}+\frac{2}{N}\bar{\xi}\sum_{i=1}^{N}(1-\alpha_i)\xi_i.$$

Since $\bar{\xi}\geq 0$, minimizing $\dot{\mathbb{V}}$ amounts to the following problem:

$$\text{Find} \quad \min\sum_{i=1}^{N}(1-\alpha_i)\xi_i,$$

which can be done as follows (where $\lfloor M\rfloor$ and $\lceil M\rceil$ respectively denote the floor and the ceiling of M):

Proposition 2.1. *Assume Hyp. 1. Furthermore, suppose that $\xi_1(t)\geq ...\geq\xi_N(t)$ (or re-arrange the agents so that this is satisfied). Then the following strategy minimizes $\frac{d}{dt}\mathbb{V}$:*
Define $I^+(t)=\{i\in\{1,..,N\},\ \xi_i(t)>0\}$.
If $|I^+(t)|\leq M$, then set $\alpha_i(t)=1$ if $i\in I^+$ and $\alpha_i(t)=0$ otherwise.
If $|I^+(t)|>M$ and $\xi_{\lceil M-1\rceil}>\xi_{\lceil M\rceil}>\xi_{\lceil M+1\rceil}$ then set $\alpha_i(t)=1$ if $i\leq\lfloor M\rfloor$, $\alpha_{\lfloor M\rfloor+1}(t)=M-\lfloor M\rfloor$ and $\alpha_i(t)=0$ otherwise.
If $|I^+(t)|>M$ and $\xi_{\lceil M-1\rceil}=\xi_{\lceil M\rceil}$ or $\xi_{\lceil M\rceil}=\xi_{\lceil M+1\rceil}$, let $I_{\lceil M\rceil}=\{i\in\{1,...,N\},\ \xi_i(t)=\xi_{\lceil M\rceil}(t)\}$ and $I^*_{\lceil M\rceil}=\{1,...,\lceil M\rceil\}\setminus I_{\lceil M\rceil}$. Then set $\alpha_i(t)=1$ if $i\in I^*_{\lceil M\rceil}$, $\alpha_i(t)=\frac{M-|I^*_{\lceil M\rceil}|}{|I_{\lceil M\rceil}|}$ if $i\in I_{\lceil M\rceil}$ and $\alpha_i(t)=0$ otherwise.

3 Optimal control for final cost

In this section, we focus on problem (ii) (see Section 1.3), i.e. minimizing the migration functional $\mathbb{V}$ at final time T using Pontryagin's maximum principle.

Let us compute the hamiltonian H of the scalar system (6):

$$H=\sum_{i=1}^{N}\lambda_i\left(-\xi_i+(1-\alpha_i)\bar{\xi}\right)=-\bar{\xi}\sum_{i=1}^{N}\alpha_i\lambda_i+\sum_{i=1}^{N}\lambda_i\left(-\xi_i+\bar{\xi}\right). \tag{11}$$

By Pontryagin's maximum principle [33], if $\alpha\in\mathcal{U}_M$, associated with the trajectory ξ, is optimal on $[0,T]$, then there exists $\lambda:[0,T]\to\mathbb{R}^N$ such that $\dot{\xi}=\frac{\partial H}{\partial\lambda}$ and $\dot{\lambda}=-\frac{\partial H}{\partial\xi}$. Furthermore the following minimization condition holds for almost all $t\in[0,T]$:

$$H(t,\xi(t),\lambda(t),\alpha(t))=\min_{\beta\in\mathcal{U}_M}H(t,\xi(t),\lambda(t),\beta(t)). \tag{12}$$

Since $\bar{\xi}\geq 0$, minimizing H requires to set $\alpha_i=1$ on the biggest positive λ_i. The differential equation for the covectors λ_i gives:

$$\dot{\lambda}_i=-\frac{\partial H}{\partial\xi_i}=\frac{1}{N}\sum_{j=1}^{N}\alpha_j\lambda_j-\bar{\lambda}+\lambda_i,\quad i\in\{1,...,N\}, \tag{13}$$

From this we can also compute the evolution of $\bar{\lambda}=\frac{1}{N}\sum_i\lambda_i$:

$$\dot{\bar{\lambda}}=\frac{1}{N}\sum_{i=1}^{N}\alpha_j\lambda_j. \tag{14}$$

Since the final condition for ξ is not fixed, the final condition for λ at time T gives:

$$\lambda(T)=\nabla\mathbb{V}(\xi(T))=\left(\frac{2}{N}\xi_1(T),...,\frac{2}{N}\xi_N(T)\right) \tag{15}$$

Proposition 3.1. *If $\bar{t} > 0$, $i, j \in \{1, ..., N\}$, and $\lambda_i(\bar{t}) = \lambda_j(\bar{t})$, then $\lambda_i(t) = \lambda_j(t)$ for all t. In this case, for a given control α, any control $\tilde{\alpha}$ satisfying $\tilde{\alpha_i} + \tilde{\alpha_j} = \alpha_i + \alpha_j$ and $\tilde{\alpha_k} = \alpha_k$ for every $k \neq i, j$ gives the same evolution of λ. If the control α satisfies the Pontryagin Maximum Principle, then the control $\tilde{\alpha}$ also does.*

Proof. Assume that at time $\bar{t}$, $\lambda_i(\bar{t}) = \lambda_j(\bar{t})$. Let us define $z_{ij} = \lambda_i - \lambda_j$. The evolution of z_{ij} is given by: $\dot{z}_{ij} = \dot{\lambda}_i - \dot{\lambda}_j = \lambda_i - \lambda_j = z_{ij}$. Hence, $z_{ij}(t) = z_{ij}(\bar{t})e^{t-\bar{t}}$, and if $z_{ij}(\bar{t}) = 0$, then for all t, $z_{ij}(t) = 0$, i.e. $\lambda_i(t) = \lambda_j(t)$. From this it follows that if α minimizes the hamiltonian H, then any control $\tilde{\alpha}$ satisfying $\tilde{\alpha_i} + \tilde{\alpha_j} = \alpha_i + \alpha_j$ and $\tilde{\alpha_k} = \alpha_k$ also minimizes H, since one easily sees from (11) that $H^\alpha = H^{\tilde{\alpha}}$ (where we denote by H^α the hamiltonian obtained with the control function α). ☐

Still assuming that the projected velocities are initially ordered (Hypothesis 1), the following lemma will allow us to further assume that they are ordered at all time.

Lemma 3.2.
There exists an optimal strategy satisfying the following: For all $t \in [0, T]$,

$$\text{If } i < j, \text{ then } \xi_i(t) \geq \xi_j(t). \tag{16}$$

Proof. Consider an optimal control strategy $\alpha \in \mathcal{U}$.
Define $\tau = \sup\{t \in [0, T];\ \exists \beta \in \mathcal{U} \text{ s.t. } \mathbb{V}_\beta(T) = \mathbb{V}_\alpha(T) \text{ and } \xi^\beta \text{ satisfies (16) on } [0, t]\}$, where $\mathbb{V}_\beta$ and ξ^β denote respectively the migration functional and the dynamics driven by the control β. Let us prove by contradiction that $\tau = T$. Suppose that $\tau < T$. Then there exist $i, j \in \{1, ..., N\}$ with $i < j$ such that $\xi_i^\beta(\tau) = \xi_j^\beta(\tau)$ and $\xi_j^\beta(t) > \xi_i^\beta(t)$ on $]\tau, \tau + \delta]$ for some $\delta > 0$. Design a control strategy $\tilde{\beta}$ such that on $[\tau, T]$, $\tilde{\beta}_i = \beta_j$, $\tilde{\beta}_j = \beta_i$ and for every $k \in \{1, ..., N\} \setminus \{i, j\}$, $\tilde{\beta}_k = \beta_k$. Then for all $t \in [\tau, T]$, $\xi_i^{\tilde{\beta}}(t) = \xi_j^\beta(t)$, and $\xi_j^{\tilde{\beta}}(t) = \xi_i^\beta(t)$. So for all $t \in [\tau, \tau + \delta]$, $\xi_i^{\tilde{\beta}}(t) \geq \xi_j^{\tilde{\beta}}(t)$ and $\mathbb{V}^{\tilde{\beta}}(T) = \mathbb{V}^\beta(T)$. Proceeding likewise for every pair of indices (m, n) satisfying $m < n$ and $\xi_m^\beta(t) < \xi_n^\beta(t)$ on $]\tau, \tau + \delta]$ we are able to design a control strategy $\tilde{\beta}$ satisfying (16) on $[0, \tau + \delta]$ and $\mathbb{V}^{\tilde{\beta}}(T) = \mathbb{V}^\alpha(T)$, which contradicts the definition of τ. In conclusion, $\tau = T$, i.e. for all $t \in [0, T]$, for every $i, j \in \{1, ..., N\}$, if $i < j$ then $\xi_i(t) \geq \xi_j(t)$. ☐

Hence, from here onward we shall assume that the variables ξ_i are ordered at all time.

Hypothesis 2. *If $i < j$, then $\xi_i(t) \geq \xi_j(t)$ for all $t \in [0, T]$.*

From Hyp.2 and the transversality condition (15), we know that the covectors are ordered at final time, i.e. $\lambda_1(T) \geq ... \geq \lambda_N(T)$. Prop. 3.1 allows us to generalize this for any time t:

$$\lambda_1(t) \geq ... \geq \lambda_N(t) \quad \text{ for all } t \in [0, T]. \tag{17}$$

The Pontryagin Maximum Principle allows us to state the following:

Proposition 3.3. *The optimal strategy requires controlling the agents with the biggest positive covectors. Let $\alpha \in \mathcal{U}_M$ be an optimal strategy and λ_i, $i \in \{1, ..., N\}$ the corresponding covectors. Define:*

$$I_\lambda(t) := \Big\{ i \in \{1, ..., N\} \ \Big| \ \lambda_i(t) \geq 0 \Big\} \quad \text{ and } \quad I_\lambda^+(t) := \Big\{ i \in \{1, ..., N\} \ \Big| \ \lambda_i(t) > 0 \Big\}. \tag{18}$$

If the set $I_\lambda(t)$ is empty, then there is no control on any agent: $\alpha_i(t) = 0$ for every i.
If the set $I_\lambda^+(t)$ is not empty, then there exists $i \in I_\lambda^+(t)$ such that $\alpha_i(t) > 0$. Furthermore, $\sum_j \alpha_j \geq \min(|I_\lambda^+(t)|, M)$.

Proof. According to Pontryagin's maximum principle (12), if the control α is optimal, then it minimizes the hamiltonian H (11) for almost all $t \in [0, T]$. The only controllable part of H is $\tilde{H} = -\bar{\xi} \sum_i \alpha_i \lambda_i$. Minimizing H requires to control the largest positive λ_i with the maximum strength allowed, while setting $\alpha_i = 0$ if $\lambda_i < 0$. If $\lambda_i = 0$, Pontryagin's maximum principle gives no information on α_i. ☐

This leads to a trichotomy of cases.

- The biggest positive λ_i's are always controlled with maximum control: $\sum_{i \in I_\lambda^+} \alpha_i = \min(|I_\lambda^+|, M)$ (where $|\cdot|$ denotes the cardinality of a set).

- If for i, j, λ_i and λ_j coincide (at a certain time, which implies at all time) then α_i and α_j are underdetermined. The PMP only requests that $\alpha_i + \alpha_j = c$ where c is given by the strength of the control to be used on the two agents.
- The negative λ_i's are never controlled: if $\lambda_i < 0$, then $\alpha_i = 0$.

Remark 3.4. *The existence of an optimal control for the problem described above is ensured by the convexity of the sets* $F(t,\xi) = \{\left(\xi_i + (1-\alpha_i)\bar{\xi}\right)_{i=1\dots N},\ \alpha \in [0,1]^N,\ \sum_i \alpha_i \le M\}$ *(see [4]).*

4 Final cost with two agents

For a clearer understanding of the mechanisms taking place, we consider the simple case of two agents in $\mathbb{R}^d$. We consider the sets of controls $\mathcal{U}_M$, where $0 < M \le 2$. Thus, system (6) becomes:

$$\begin{cases} \dot{\xi}_1 = -\xi_1 + (1-\alpha_1)\,\bar{\xi} \\ \dot{\xi}_2 = -\xi_2 + (1-\alpha_2)\,\bar{\xi}. \end{cases} \tag{19}$$

Computing the difference of the two projected variables will also prove useful:

$$\dot{\xi}_1 - \dot{\xi}_2 = -(\xi_1 - \xi_2) - (\alpha_1 - \alpha_2)\bar{\xi} \tag{20}$$

Three different situations may arise, depending on the value of the constraint on the control. Indeed, two constraints are set: $\alpha_1 + \alpha_2 \le M$, and $0 \le \alpha_i \le 1$ for $i = 1, 2$. We differentiate the cases (a) $0 < M \le 1$, (b) $1 < M < 2$ and (c) $M = 2$.

4.1 Pontryagin's Maximum Principal

Notice that the migration functional can be written as:

$$\mathbb{V} = \frac{1}{2}(\xi_1^2 + \xi_2^2) = \frac{1}{4}\left((\xi_1+\xi_2)^2 + (\xi_1-\xi_2)^2\right) = \bar{\xi}^2 + \left(\frac{\xi_1 - \xi_2}{2}\right)^2, \tag{21}$$

once again emphasizing the necessary trade-off between two terms: the mean velocity $\bar{\xi}$ and the distance between the agents $|\xi_1 - \xi_2|$. Computing the hamiltonian of the system gives:

$$H(t,\xi,\lambda,\alpha) = -\bar{\xi}\,(\alpha_1\lambda_1 + \alpha_2\lambda_2) + \frac{\xi_2 - \xi_1}{2}(\lambda_1 - \lambda_2), \tag{22}$$

In line with Hyp. 1, two cases are possible: $\xi_1(0) = \xi_2(0)$ or $\xi_1(0) > \xi_2(0)$. If $\xi_1(0) = \xi_2(0)$, one easily sees in (21) that the optimal strategy consists in keeping $\xi_1(t) = \xi_2(t)$ for all time, while maximizing the decay of $\bar{\xi}$. This is done by choosing $\alpha_1(t) = \alpha_2(t) = \frac{M}{2}$ for all t. Hence, in the following we will assume:

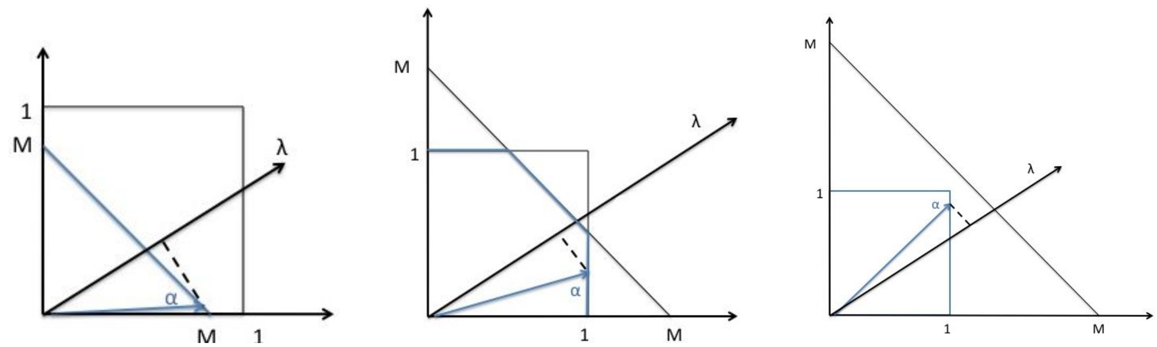
Hypothesis 3. $\xi_1(0) > \xi_2(0)$.

Before studying each case in detail, we give general considerations on the relation between the control α and λ:

(a) If $M \le 1$, minimizing H (i.e. maximizing $\langle\lambda,\alpha\rangle$) gives (see Fig.1a): $(\alpha_1,\alpha_2) = (M,0)$ if $0 < \lambda_2 < \lambda_1$; $(\alpha_1,\alpha_2) = (M/2, M/2)$ if $0 < \lambda_2 = \lambda_1$; $(\alpha_1,\alpha_2) = (M,0)$ if $\lambda_2 < 0 < \lambda_1$; $(\alpha_1,\alpha_2) = (0,0)$ if $\lambda_2 < 0$ and $\lambda_1 < 0$.

(b) If $1 < M < 2$, minimizing H gives (see Fig.1b): $(\alpha_1,\alpha_2) = (1, M-1)$ if $0 < \lambda_2 < \lambda_1$; $(\alpha_1,\alpha_2) = (M/2, M/2)$ if $0 < \lambda_2 = \lambda_1$; $(\alpha_1,\alpha_2) = (1,0)$ if $\lambda_2 < 0 < \lambda_1$; $(\alpha_1,\alpha_2) = (0,0)$ if $\lambda_2 < 0$ and $\lambda_1 < 0$.

(c) If $M \ge 2$, minimizing H gives (see Fig.1c): $(\alpha_1,\alpha_2) = (1,1)$ if $0 < \lambda_2 \le \lambda_1$; $(\alpha_1,\alpha_2) = (1,0)$ if $\lambda_2 < 0 < \lambda_1$; $(\alpha_1,\alpha_2) = (0,0)$ if $\lambda_2 < 0$ and $\lambda_1 < 0$.



(a) Case $M \leq 1$ (b) Case $1 < M < 2$ (c) Case $M \geq 2$

Figure 1: Minimizing $-\langle \lambda, \alpha \rangle$

Notice that in all three cases, if $\lambda_1 = \lambda_2$, then the Pontryagin maximum principle does not give sufficient information since any combination of α_1 and α_2 such that $\alpha_1 + \alpha_2 = M$ minimizes the scalar product $-\langle \lambda, \alpha \rangle$ (see Figure 1).

The dynamics for λ are given by $\dot{\lambda} = -\nabla H = \begin{pmatrix} \frac{1+\alpha_1}{2}\lambda_1 - \frac{1-\alpha_2}{2}\lambda_2 \\ \frac{1+\alpha_2}{2}\lambda_2 - \frac{1-\alpha_1}{2}\lambda_1 \end{pmatrix}$, which allows us to compute the evolution of the difference $\lambda_1 - \lambda_2$:

$$\frac{d}{dt}(\lambda_1 - \lambda_2) = \lambda_1 - \lambda_2. \tag{23}$$

The transversality conditions give: $\lambda(T) = \nabla \mathbb{V}(T) = (\xi_1(T), \xi_2(T))^T$. Hence, if the final configuration is such that $\xi_1(T) \neq \xi_2(T)$, i.e. $\lambda_1(T) \neq \lambda_2(T)$, the difference $\lambda_1 - \lambda_2$ increases with time. On the other hand, if $\lambda_1(T) = \lambda_2(T)$, then $\forall t \leq T,\ \lambda_1(t) = \lambda_2(t)$. If the dynamics allow us to drive ξ_1 and ξ_2 together before time T, then $\lambda_1(t) = \lambda_2(t)$ for all t, and the Pontryagin maximum principle does not give sufficient information, as seen above.

4.2 Global Strategy

According to equation (21), the functional $\mathbb{V}$ can be written as:

$$\mathbb{V} = \bar{\xi}^2 + \frac{(\xi_1 - \xi_2)^2}{4}. \tag{24}$$

Minimizing $\mathbb{V}$ requires minimizing $\bar{\xi}$ and $(\xi_1 - \xi_2)^2$ simultaneously. The evolution of $\bar{\xi}$ is given by:

$$\dot{\bar{\xi}} = -\frac{1}{2}(\alpha_1 + \alpha_2)\,\bar{\xi}, \tag{25}$$

while that of $(\xi_1 - \xi_2)^2$ is:

$$\frac{d}{dt}\left((\xi_1 - \xi_2)^2\right) = -2(\xi_1 - \xi_2)^2 - 2(\xi_1 - \xi_2)\bar{\xi}(\alpha_1 - \alpha_2). \tag{26}$$

Thus, minimizing $\bar{\xi}^2$ (both instantaneously and globally) requires using full control, i.e. setting $\alpha_1 + \alpha_2 = M$. On the other hand, the strategy to minimize $(\xi_1 - \xi_2)^2$ is less clear. It would require both maximizing $\bar{\xi}$ and maximizing the difference $\alpha_1 - \alpha_2$ (assuming that $\xi_1 - \xi_2 \geq 0$), and these conditions might not be compatible.

4.3 Case $M = 1$

Theorem 4.1.
Let $T > 0$ and let $M = 1$. Furthermore, let $\alpha = (\alpha_1, \alpha_2) \in \mathcal{U}_1$ (see (4)) be an optimal control and ξ be the corresponding trajectory of system (19). Define $t_0 = 2\ln(\xi_1(0)/\bar{\xi}(0))$. Then

(i) $T \geq t_0$ *if and only if* $\xi_1(T) = \xi_2(T)$*. In such a case, the the control satisfies:* $\alpha_1 + \alpha_2 \equiv 1$ *(so* $\bar{\xi}(t) = \bar{\xi}(0)e^{-t/2}$*). For instance, the strategy* $(\alpha_1, \alpha_2)(t) = (1, 0)$ *for all* $t \in [0, t_0[$ *and* $(\alpha_1, \alpha_2)(t) = (1/2, 1/2)$ *for all* $t \in [t_0, T]$ *is optimal.*

(ii) If $T < t_0$*, then* $\alpha(t) = (0, 0)$ *for all* $t \in [0, t^*[$ *and* $\alpha(t) = (1, 0)$ *for all* $t \in [t^*, T]$*, where* $t^* = 2\ln(\bar{X})$ *and* $\bar{X} \in [1, e^{T/2}[$ *is defined as follows:*

$$\bar{X} = \arg \min_{X \in [1, e^{T/2}]} \left[\left(\xi_1(0) + \bar{\xi}(0)(X^2 - 1)\right)^2 + \left(\xi_2(0) + \bar{\xi}(0)(X^2 - 1) + 2\bar{\xi}(0)X(e^{T/2} - X)\right)^2 \right].$$

Proof. Let ξ be an optimal trajectory achieved with optimal control α.
To prove *(i)*, we shall show that the three statements (a) $T \geq t_0$, (b) there exists $t \in [0, T]$ such that $\xi_1(t) = \xi_2(t)$ and (c) $\xi_1(T) = \xi_2(T)$ are equivalent.
Suppose (b), i.e. that there exists $\tau \in [0, T]$ such that $\xi_1(\tau) = \xi_2(\tau)$. Then necessarily $\xi_1(T) = \xi_2(T)$. Indeed, suppose that $\xi_1(T) \neq \xi_2(T)$. Then any strategy $\tilde{\alpha}$ such that on $[0, \tau]$, $\tilde{\alpha} = \alpha$ and on $]\tau, T]$, $(\tilde{\alpha}_1, \tilde{\alpha}_2) = (\frac{\alpha_1+\alpha_2}{2}, \frac{\alpha_1+\alpha_2}{2})$ achieves: $\tilde{\bar{\xi}}(T) = \bar{\xi}(T)$ and $(\tilde{\xi}_1 - \tilde{\xi}_2)^2(T) = 0 < (\xi_1 - \xi_2)(T)$ (where $\tilde{\xi}$, $\tilde{V}$ denote the trajectory and cost corresponding to $\tilde{\alpha}$), so according to equation (24), $\tilde{\mathbb{V}}(T) < \mathbb{V}(T)$ and control strategy α cannot be optimal. Hence, $\xi_1(T) = \xi_2(T)$.
Now suppose (c) $\xi_1(T) = \xi_2(T)$. The transversality condition (15) gives $\lambda_1(T) = \lambda_2(T)$ and from Proposition 3.1 we get: $\lambda_1(t) = \lambda_2(t)$ for all $t \in [0, T]$. Then, $\dot{\bar{\lambda}} = \sum \alpha_i \lambda_i = (\sum \alpha_i)\bar{\lambda}$. Since $\bar{\xi}(T) > 0$, the transversality condition (15) gives: $\bar{\lambda}(T) > 0$, and $\bar{\lambda}(t) = \lambda_1(t) = \lambda_2(t) > 0$ for all $t \in [0, T]$. Therefore, the set I_λ, see (18), is not empty, so according to Proposition 3.3, the optimal control strategy requires using maximal control strength: $\alpha_1 + \alpha_2 \equiv 1$. According to equation (25), this suffices to fully determine $\bar{\xi}(t) = \bar{\xi}(0)\, e^{-t/2}$. Then $\xi_1(t) - \xi_2(t) = e^{-t}\left((\xi_1 - \xi_2)(0) - \bar{\xi}(0)\int_0^t (\alpha_1 - \alpha_2)e^{s/2}ds\right)$, and $\xi_1(t) - \xi_2(t) = 0$ if, and only if, $\int_0^t e^{s/2}(\alpha_1 - \alpha_2)(s)ds = (\xi_1(0) - \xi_2(0))/\bar{\xi}(0)$. Notice that $\min_{(\alpha_1,\alpha_2)\in\mathcal{U}_1}\{t \mid (\xi_1 - \xi_2)(t) = 0\}$ is obtained when $\alpha_1 - \alpha_2$ is maximal, i.e. for $(\alpha_1, \alpha_2) \equiv (1, 0)$. With this strategy, $\min_{(\alpha_1,\alpha_2)\in\mathcal{U}_1}\{t \mid (\xi_1 - \xi_2)(t) = 0\} := t_0 = 2\ln(\xi_1(0)/\bar{\xi}(0))$. Hence, we must have: $T \geq t_0$.
Lastly, suppose (a) $T \geq t_0$. Design a strategy $\tilde{\alpha}$ so that for all $t < t_0$, $(\tilde{\alpha}_1, \tilde{\alpha}_2) = (1, 0)$ and for all $t \geq t_0$, $(\tilde{\alpha}_1, \tilde{\alpha}_2) = (1/2, 1/2)$. This strategy is optimal since it maximizes the decrease of $\tilde{\bar{\xi}}$, see (25), and achieves $(\tilde{\xi}_1 - \tilde{\xi}_2)(T) = 0$, see (26). Hence, our optimal strategy α must also satisfy: $\xi_1(T) = \xi_2(T)$ and $\alpha_1 + \alpha_2 \equiv 1$. This proves (b).
We showed that (a), (b) and (c) are equivalent. We thus proved the first part of the proposition: $T \geq t_0$ if and only if $\xi_1(T) = \xi_2(T)$. In this case, it also holds: $\alpha_1 + \alpha_2 \equiv 1$.
If on the other hand, *(ii)* $T < t_0$, then $\xi_1(t) > \xi_2(t)$ for all $t \in [0, T]$ (since (b) implies (a)). According to condition (15) and to Prop. 3.1, $\lambda_1(t) > \lambda_2(t)$ for all $t \in [0, T]$ and $\lambda_1(T) > 0$. The evolution of λ_1 is given by: $\dot{\lambda}_1 = \frac{1}{2}(\alpha_1\lambda_1 + \alpha_2\lambda_2) + \lambda_1 - \bar{\lambda} > 0$ since $\lambda_1 > \bar{\lambda}$. Hence, two cases must be distinguished: either $\lambda_1 > 0$ at all time, so the set I_λ^+ is non-empty and full control will be used at all time, or there exists $t^* \in]0, T[$ such that $\lambda_1 < 0$ on $[0, t^*[$, $\lambda_1(t^*) = 0$ and $\lambda_1 > 0$ on $]t^*, T]$, in which case $\alpha = (0, 0)$ on $[0, t^*[$ and $\alpha = (1, 0)$ on $]t^*, T]$. Knowing this, it is easy to express ξ_1, ξ_2 and $\mathbb{V}$ as functions of t^*:

$$\forall t \in [t^*, T], \begin{cases} \xi_1(t) = e^{-t}(\xi_1(0) + \bar{\xi}(0)(e^{t^*} - 1)) \\ \xi_2(t) = e^{-t}(\xi_2(0) + \bar{\xi}(0)(e^{t^*} - 1) + 2\bar{\xi}(0)e^{t^*/2}(e^{t/2} - e^{t^*/2})) \\ \mathbb{V}(t) = \xi_1^2(t) + \xi_2^2(t) \end{cases} \tag{27}$$

Denoting $X = e^{t^*/2}$, $\mathbb{V}(T)$ can be written as a biquadratic polynomial in X:

$$\mathbb{V}(T)(X) = e^{-2T}\left[\left(\xi_1(0) + \bar{\xi}(0)(X^2 - 1)\right)^2 + \left(\xi_2(0) + \bar{\xi}(0)(X^2 - 1) + 2\bar{\xi}(0)X(e^{T/2} - X)\right)^2 \right].$$

We look for $\bar{X}$ minimizing $\mathbb{V}(T)(X)$ in the interval $[1, e^{T/2}]$ (so that $t^* \in [0, T]$). Notice that the leading term is $2e^{-2T}\bar{\xi}(0)^2 \cdot X^4$. Hence, there are at most two local minima in the interval $[1, e^{T/2}]$. Furthermore, $\mathbb{V}(T)(1) = e^{-2T}\left[\xi_1(0)^2 + (\xi_2(0) + 2\bar{\xi}(0)(e^{T/2} - 1))^2\right]$ and $\mathbb{V}(T)(e^{T/2}) = e^{-2T}[(\xi_1(0) + \bar{\xi}(0)(e^{T/2} - 1))^2 + (\xi_2(0) + \bar{\xi}(0)(e^T - 1))^2]$, so $\mathbb{V}(T)(1) < \mathbb{V}(T)(e^{T/2})$, which means that $\bar{X} < e^{T/2}$. If $\bar{X} = 1$, then $t^* = 0$ so it is optimal to act with control $(1, 0)$ on the full interval $[0, T]$. If $1 < \bar{X} < e^{T/2}$, then $0 < t^* < T$. The optimal control strategy will require leaving the system to evolve without control on $[0, t^*[$, and acting with control $\alpha = (1, 0)$ on $[t^*, T]$. □

Remark 4.2. *The existence of an initial "Inactivation" period can be proven also with any number of agents (see Theorem 5.3). Numerical simulations with any number of agents (see Section 5.2) show that in some cases it is indeed optimal to let the system evolve without control on an initial time interval $[0, t^*]$, where $t^* > 0$.*

4.4 Case $M < 1$

Generalizing to the case of any $M < 1$, we conduct the same analysis and the optimal control strategy is similar.

Theorem 4.3. *Let T>0 and M<1. Let $\alpha = (\alpha_1, \alpha_2) \in \mathcal{U}_M$ (see (4)) be an optimal control and ξ be the corresponding trajectory of system (19). Define $t_0 = \frac{2}{2-M} \ln\left(\frac{2-M}{2M}(\xi_1(0) - \xi_2(0))/\bar{\xi}(0) + 1\right)$, Then*

(i) $T \geq t_0$ if and only of $\xi_1(T) = \xi_2(T)$. In such a case, the control satisfies: $\alpha_1 + \alpha_2 \equiv M$ (so $\bar{\xi}(t) = \bar{\xi}(0)e^{-Mt/2}$).

(ii) If $T < t_0$, then there exists $t^* \in [0, T[$ such that $\alpha(t) = (0, 0)$ for all $t \in [0, t^*[$ and $\alpha(t) = (1, 0)$ for all $t \in [t^*, T]$.

Remark 4.4. *To compute $t^* \in [0, T[$ in the case $T < t_0$, one can compute $\mathbb{V}(T)(e^{t^*/2})$ depending on $t^* \in [0, T]$ similarly to the case $M = 1$.*

Proof. Let ξ be an optimal trajectory achieved with optimal control $\alpha \in \mathcal{U}_M$. We argue as in the case $M = 1$. To prove *(i)*, first suppose that there exists $\tau \in [0, T]$ such that $\xi_1(\tau) = \xi_2(\tau)$. Then, as in the case $M = 1$, necessarily it holds $\xi_1(T) = \xi_2(T)$ and any strategy achieving $\xi_1(T) = \xi_2(T)$ while using maximum control $\alpha_1 + \alpha_2 \equiv M$ is optimal. Then $\xi_1(t) - \xi_2(t) = 0 \Leftrightarrow \int_0^t e^{\frac{2-M}{2}s}(\alpha_1 - \alpha_2)(s)ds = (\xi_1(0) - \xi_2(0))/\bar{\xi}(0)$. Hence, $\min_{\alpha \in \mathcal{U}_M}\{t \mid (\xi_1 - \xi_2)(t) = 0\}$ is obtained when $\alpha_1 - \alpha_2$ is maximal, i.e. for $(\alpha_1, \alpha_2) \equiv (M, 0)$. With this strategy, $\min_{\alpha \in \mathcal{U}_M}\{t \mid (\xi_1 - \xi_2)(t) = 0\} = t_0$ as defined above. Hence, if there exists $\tau \leq T$ such that $\xi_1(\tau) = \xi_2(\tau)$, then $T \geq t_0$. This proves the first implication of the proposition: if $\xi_1(T) = \xi_2(T)$, then $T \geq t_0$.

Conversely, if $T \geq t_0$, then the strategy $(\tilde{\alpha}_1, \tilde{\alpha}_2) = (M, 0)$ on $[0, t_0[$ and $(\tilde{\alpha}_1, \tilde{\alpha}_2) = (\frac{M}{2}, \frac{M}{2})$ on $[t_0, T]$ is optimal since it minimizes $\bar{\tilde{\xi}}(T)$ and achieves $\tilde{\xi}_1(T) = \tilde{\xi}_2(T)$. Hence, if α is optimal, it must also satisfy $\alpha_1 + \alpha_2 \equiv M$ and $\xi_1(T) = \xi_2(T)$, which proves the second implication.

Now assume *(ii)* $T < t_0$. From *(i)* we get: $\xi_1(t) > \xi_2(t)$ for all $t \in [0, T]$. One can then argue as in the case $M = 1$. According to Pontryagin's Maximum Principle, $\alpha_2 \equiv 0$ and two cases have to be distinguished: either $\lambda_1 > 0$ at all time, so the set I_λ^+ (see (18)) is non-empty and full control will be used at all time, or there exists $t^* \in]0, T[$ such that $\lambda_1 < 0$ on $[0, t^*[$, $\lambda_1(t^*) = 0$ and $\lambda_1 > 0$ on $]t^*, T]$, in which case $\alpha = (0, 0)$ on $[0, t^*[$ and $\alpha = (M, 0)$ on $]t^*, T]$.

□

Remark 4.5. *Notice that in the limit case $M \to 1$ of Theorem 4.3, one finds the same expression for t_0 as in Theorem 4.1.*

4.5 Case $M = 2$

In order to determine the optimal strategy, let us first study the evolution of the covectors λ. From $\xi_1(T) \geq \xi(T) > 0$ (see Prop. 1.1 and Hyp. 2) and the transversality condition (15), we get $\lambda_1(T) > 0$.

Proposition 4.6. *Let $M = 2$ and λ_1 and λ_2 be the covectors corresponding to an optimal control strategy for the system (19). Then they satisfy the following properties:*

(i) If $\lambda_2(T) > 0$, then $\lambda_1(t) > 0$ and $\lambda_2(t) > 0$ for all $t \in [0, T]$.

(ii) If $\lambda_2(T) = 0$, then $\lambda_1(t) > 0$ and $\lambda_2(t) = 0$ for all $t \in [0, T]$.

(iii) If $\lambda_2(T) < 0$, then $\lambda_2(t) < 0$ for all $t \in [0, T]$.

Proof.
(i) Let $\lambda_2(T) > 0$. Suppose that there exists $\tau \in [0, T[$ such that $\lambda_2(\tau) = 0$ and $\lambda_2(t) > 0$ for all $t \in]\tau, T]$. Then since $\lambda_1 \geq \lambda_2 > 0$ on $]\tau, T]$, according to Pontryagin's maximum principle (see Section 4.1), $(\alpha_1, \alpha_2) \equiv (1, 1)$ on $]\tau, T]$, which gives the following evolutions: $\dot{\lambda}_1 = \lambda_1$ and $\dot{\lambda}_2 = \lambda_2$. Hence, $\lambda_2(\tau) = \lambda_2(T)e^{\tau - T} > 0$, which contradicts the definition of τ. Therefore, $\lambda_2(t) > 0$ for all $t \in [0, T]$, and by (17), $\lambda_1(t) > 0$.
(ii) Let $\lambda_2(T) = 0$. Let $\tau := \inf_{[0,T]}\{\bar{t} \in [0, T]$ s.t. $\lambda_2(t) = 0$ for all $t > \bar{t}\}$ and suppose that $\tau > 0$. By definition of τ, $\lambda_2(\tau) = 0$. Since $\lambda_1(t) > \lambda_2(t)$ for all t (see Prop. 3.1), there exists an interval $[\tau - \delta, \tau[$ on which $\lambda_1 > 0$ and either $\lambda_2 > 0$ or $\lambda_2 < 0$. If $\lambda_2(t) > 0$ for all $t \in [\tau - \delta, \tau[$, then according to Pontryagin's maximum principle (Section 4.1), the control satisfies $\alpha_1(t) = \alpha_2(t) = 1$, which gives: $\dot{\lambda}_2(t) = \lambda_2(t) > 0$. So $\lambda_2(\tau) > 0$, which contradicts the definition of τ. If on the other hand $\lambda_2(t) < 0$ for all $t \in [\tau - \delta, \tau[$, then $\alpha_2(t) = 0$ and $\dot{\lambda}_2(t) = \frac{1}{2}\lambda_2(t) < 0$, which is impossible since it implies $\lambda_2(\tau) < 0$. Hence, $\tau = 0$. Furthermore, since $\lambda_1(T) > 0$ and $\lambda_2 \equiv 0$, then $\dot{\lambda}_1 = \lambda_1$ in a neighborhood of T, which ensures that $\lambda_1(t) > 0$ for all $t \in [0, T]$ (by the same reasoning as in (i)).
(iii) Let $\lambda_2(T) < 0$. Define $\tau := \inf_{[0,T]}\{\bar{t} \in [0, T]$ s.t. $\lambda_1(t) > 0$ and $\lambda_2(t) < 0$ for all $t > \bar{t}\}$. Then on $]\tau, T]$, as seen in Section 4.1, $\alpha_1 \equiv 1$ and $\alpha_2 \equiv 0$, which gives: $\lambda_2(t) = \lambda_2(\tau)e^{T-\tau}$. Since $\lambda_2(T) < 0$, it follows that $\lambda_2(\tau) < 0$. Hence, either $\tau = 0$ or $\lambda_1(\tau) = 0$. Notice that since $\lambda_1(t) > \lambda_2(t)$ for all t, λ_1 is strictly increasing (see (13)). Then the former case implies that $\lambda_2(t) < 0$ for all $t \in [0, T]$. In the latter case, we get that $\lambda_2(t) < 0$ for all $t \leq \tau$. □

This information about the covectors allows us to solve the optimization problem based on the initial conditions and the final time. Recall from Proposition 1.1 that $\xi_1(0) > 0$.

Theorem 4.7. *Let M=2. Let $(\alpha_1, \alpha_2) \in \mathcal{U}_2$ be an optimal control strategy and ξ be the corresponding trajectory for system* (19). *Define $t_0 = 2\ln\left(\xi_1(0)/(2\bar{\xi}(0))\right)$.*

- *(i) If $\xi_2(0) > 0$, then $(\alpha_1, \alpha_2) \equiv (1, 1)$.*
- *(ii) If $\xi_2(0) \leq 0$ and $T \geq t_0$, then $\xi_2(T) = 0$ and $\alpha_1 \equiv 1$. For instance the strategy $(\alpha_1, \alpha_2) = (1, 0)$ for all $t \in [0, t_0[$ and $(\alpha_1, \alpha_2) = (1, 1)$ for all $t \in [t_0, T]$ is optimal. Furthermore, if there exists $\bar{t} \in [0, T[$ such that $\xi_2(\bar{t}) = 0$, then $\xi_2(t) = 0$ for all $t \in [\bar{t}, T]$.*
- *(iii) If $\xi_2(0) \leq 0$ and $T < t_0$, then there exists $t^* \in [0, T[$ such that $\alpha(t) = (0, 0)$ for all $t \in [0, t^*[$ and $\alpha(t) = (1, 0)$ for all $t \in [t^*, T]$.*

Proof. Let (α_1, α_2) be an optimal control strategy and ξ be the corresponding trajectory.
(i) Let $\xi_2(0) > 0$. According to Prop. 1.1, for all $t \in [0, T]$ it holds $\xi_1(t) > 0$ and $\xi_2(t) > 0$. Then $\lambda_1(T) > 0$ and $\lambda_2(T) > 0$. From Prop. 4.6 it follows that $\lambda_1(t) > 0$ and $\lambda_2(t) > 0$ for all $t \in [0, T]$. According to the PMP (see Section 4.1), maximal control has to be used at all time, i.e. $(\alpha_1, \alpha_2)(t) = (1, 1)$ for all $t \in [0, T]$.

For cases *(ii)* and *(iii)*, let $\xi_2(0) \leq 0$. By Prop. 1.1 it holds $\xi_1(t) > 0$ for all $t \in [0, T]$. Suppose that $\xi_2(T) > 0$. Then from Prop. 4.6 we get $\lambda_1(t) \geq \lambda_2(t) > 0$ for all $t \in [0, T]$, so $(\alpha_1, \alpha_2) \equiv (1, 1)$. But with this strategy $\dot{\xi}_2 = -\xi_2$, so $\xi_2(t) = \xi_2(0)e^{-t} \leq 0$ for all $t \in [0, T]$, which contradicts $\xi_2(T) > 0$. Hence $\xi_2(T) \leq 0$.
(ii) First assume that $T \geq t_0$. Let us show that $\xi_2(T) = 0$ and $\alpha_1 \equiv 1$. Such a strategy exists, since for instance the control $(\beta_1, \beta_2)(t) = (1, 0)$ for $t \in [0, t_0[$ and $(\beta_1, \beta_2)(t) = (1, 1)$ for $t \in [t_0, T]$ achieves $\xi_2^\beta(t) = 0$ for all $t \in [t_0, T]$ (where ξ^β denotes the trajectory corresponding to the control strategy β) – by direct computation of (19). Suppose that $\xi_2(T) < 0$. Then α cannot be optimal since the control strategy β achieves the minimum of $\xi_1^\beta(T)^2$, see (19), and of $\xi_2^\beta(T)^2$ and therefore the minimum of $\mathbb{V}(T) = \xi_1^\beta(T)^2 + \xi_2^\beta(T)^2$. Hence α must satisfy $\alpha_1 \equiv 1$ and $\xi_2(T) = 0$ in order to perform as well as β. Obviously, all strategies that achieve $\xi_2(T) = 0$ with $\alpha_1 \equiv 1$ achieve the same final positions (see (19)) and thus have the same $\mathbb{V}(T)$. Furthermore, if there exists a $\hat{t} < T$ such that $\xi_2(\hat{t}) = 0$, then $\xi_2(t) = 0$ for all $t \in [\hat{t}, T]$: if $\xi_2(\bar{t}) = 0$, then $\dot{\xi}_2(\bar{t}) = (1 - \alpha_2)\bar{\xi}(\bar{t}) \geq 0$ and therefore ξ_2 cannot become negative, once it reaches 0. On the other hand, if $\xi_2(t) > 0$, then $\xi_2(T) > 0$ by Prop. 1.1.
(iii) Assume now that $T < t_0$. Firstly, we show that an optimal strategy (by PMP) always achieves $\xi_2(T) < 0$. We argue by contradiction: Assume that $\xi_2(T) = 0$. Then $\lambda_1(T) > 0$ and $\lambda_2(T) = 0$ and, according to Proposition 4.6, it follows that $\lambda_1(t) > \lambda_2(t) = 0$ for all $t \in [0, T]$. According to the PMP,

$\alpha_1 \equiv 1$. Then the growth of ξ_2 is maximal if, and only if, $\alpha_2 \equiv 0$ since in this case $\bar{\xi}$ is maximal. But with this strategy ξ_2 cannot reach 0 before t_0 – by direct computation of (19). Therefore $\xi_2(T) < 0$, so $\lambda_2(T) < 0$ and $\lambda_2(t) < 0$ for all $t \in [0,T]$ by Prop. 4.6. Hence we are in the same situation as in the case $M = 1$ and $M < 1$. Two cases are possible: either $\lambda_1 > 0$ at all time, so the set I_λ^+ is non-empty and full control on ξ_1 is used at all time, or there exists $t^* \in]0,T[$ such that $\lambda_1 < 0$ on $[0,t^*[$, $\lambda_1(t^*) = 0$ and $\lambda_1 > 0$ on $]t^*,T]$, in which case $\alpha = (0,0)$ on $[0,t^*[$ and $\alpha = (1,0)$ on $]t^*,T]$. □

Remark 4.8. *To compute $t^* \in [0,T[$ in the case $\xi_1(0) > -\xi_2(0) > 0$ and $T < t_0$, one can compute $\mathbb{V}(T)(X)$ depending on $t^* \in [0,T[$ similarly to the case $M = 1$.*

4.6 Case $1 < M < 2$

As in the case $M = 2$, we state the following properties concerning the covectors λ.

Proposition 4.9. *Let $M \in]1,2[$ and λ_1 and λ_2 be the covectors corresponding to an optimal control strategy for the system* (19). *They satisfy the following properties:*

(i) If $\lambda_2(T) > 0$, then $\lambda_1(t) > 0$ and $\lambda_2(t) > 0$ for all $t \in [0,T]$.

(ii) If $\lambda_2(T) = 0$, then $\lambda_1(t) > 0$ and $\lambda_2(t) = 0$ for all $t \in [0,T]$.

(iii) If $\lambda_2(T) < 0$, then $\lambda_2(t) < 0$ for all $t \in [0,T]$.

Proof. The proof is very similar to that of Prop 4.6.
(i) Let $\lambda_2(T) > 0$. Suppose that there exists $\tau \in [0,T[$ such that $\lambda_2(\tau) = 0$ and $\lambda_2(t) > 0$ for all $t \in]\tau,T]$. Then if $\lambda_1 > \lambda_2 > 0$ on $]\tau,T]$, according to Pontryagin's maximum principle (see Section 4.1), $(\alpha_1,\alpha_2) \equiv (1, M-1)$ on $]\tau,T]$, which gives $\dot{\lambda}_2 = \frac{M}{2}\lambda_2$. If $\lambda_1 = \lambda_2 > 0$ on $]\tau,T]$, then $\alpha_1 + \alpha_2 \equiv M$ (see Figure 1b), which also gives $\dot{\lambda}_2 = \frac{M}{2}\lambda_2$. Hence, $\lambda_2(\tau) = \lambda_2(T)e^{\frac{M}{2}(\tau - T)} > 0$, which contradicts the definition of τ.
For (ii) and (iii) we reason the same way as in the proof of Proposition 4.6. □

As in the previous sections, this allows us to solve the optimal control problem by distinguishing cases based on the initial conditions and the final time. The case $\xi_2(0) < 0$ is illustrated in Figure 2.

Theorem 4.10.
Let $M \in]1,2[$. Let $\alpha \in \mathcal{U}_M$ be an optimal control strategy and ξ be the corresponding trajectory.
Define $t_0 \le t_1 \le t_2$ as: $t_0 = 2\ln\left(\frac{\xi_1(0)}{2\bar{\xi}(0)}\right)$, $t_1 = \frac{2}{2-M}\ln\left(\frac{\xi_1(0)}{2\bar{\xi}(0)}\right)$ and $t_2 = \frac{2}{2-M}\ln(\frac{\xi_1(0)}{\bar{\xi}(0)})$.
If $\xi_2(0) > 0$, two subcases are to be distinguished:

- *If $T < t_2$, then $(\alpha_1,\alpha_2) \equiv (1, M-1)$ and $0 < \xi_2(T) < \xi_1(T)$.*
- *If $T \ge t_2$, $\xi_1(T) = \xi_2(T)$ and $\alpha_1 + \alpha_2 = M$.*

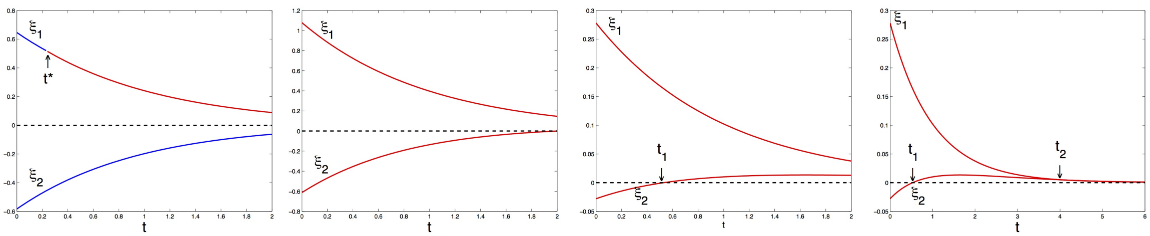
In the case $\xi_2(0) < 0$, four subcases appear:

- *If $T < t_0$, then $\xi_2(t) < 0$ and there exists $t^* \in [0,T[$ such that $(\alpha_1,\alpha_2)(t) = (0,0)$ for all $t \in [0,t^*]$ and $(\alpha_1,\alpha_2)(t) = (1,0)$ for all $t \in]t^*,T]$.*
- *If $t_0 \le T \le t_1$, then $\alpha_1 \equiv 1$ and $\xi_2(T) = 0$.*
- *If $t_1 < T < t_2$, then $(\alpha_1,\alpha_2) \equiv (1, M-1)$ and $0 < \xi_2(T) < \xi_1(T)$.*
- *If $t_2 \le T$, then $\alpha_1 + \alpha_2 \equiv M$ and $\xi_1(T) = \xi_2(T)$.*

Remark 4.11. *Notice that if $\xi_1(0) = \xi_2(0)$, then $t_2 = 0$.*

Remark 4.12. *In the limit case $M \to 1$, the times t_0 and t_1 are equal, which is in line with Theorem 4.1. In the limit case $M \to 2$, t_1 and t_2 are undefined, in line with Theorem 4.7.*

Proof. See appendix. □



(a) $T < t_0$ (b) $t_0 \leq T \leq t_1$ (c) $t_1 < T < t_2$ (d) $t_2 \leq T$

Figure 2: Control strategies in the case $\xi_2(0) < 0$ (controlled agents in red, uncontrolled ones in blue)

5 Final cost with any number of agents and control bounded by M=1

5.1 Theroretical Analysis

In this section, we address the optimal control problem of minimizing $\mathbb{V}(T)$ with any number of agents, setting the upper bound $M = 1$ on the strength of the control, i.e. $\sum_{i=1}^{N} \alpha_i \leq 1$. We define the set of such controls:

$$\mathcal{U} = \left\{ \alpha : [0,T] \to [0,1]^N \,\middle|\, \alpha \text{ measurable, s.t. for all } t \in [0,T] \ \sum_{i=1}^{N} \alpha_i(t) \leq 1 \right\}. \tag{28}$$

We remind the equations governing the evolution of ξ_i and $\bar{\xi}$ for $i \in \{1, ..., N\}$:

$$\dot{\xi}_i = -\xi_i + (1 - \alpha_i)\bar{\xi} \qquad \text{and} \qquad \dot{\bar{\xi}} = -(\sum_i \alpha_i)\ \bar{\xi}. \tag{29}$$

As before, we aim to minimize the migration functional $\mathbb{V} = \frac{1}{N} \sum_{i=1}^{N} \xi_i^2$ over the space $\mathcal{U}$ at final time:

Problem 1. *Find* $\arg\min_{\alpha \in \mathcal{U}} \mathbb{V}(T)$.

Let us consider the restricted set of full-strength controls $\mathcal{U}_{FS} \subset \mathcal{U}$:

$$\mathcal{U}_{FS} = \left\{ \alpha : [0,T] \to [0,1]^N \,\middle|\, \alpha \text{ measurable, s.t. for all } t, \ \sum_{i=1}^{N} \alpha_i(t) = 1 \right\}. \tag{30}$$

We also introduce the set of optimal controls $\mathcal{U}_{\text{opt}}$:

$$\mathcal{U}_{\text{opt}} = \left\{ \alpha \in \mathcal{U} \ \text{ s.t. } \mathbb{V}_\alpha = \min_{\beta \in \mathcal{U}} \mathbb{V}_\beta \right\}. \tag{31}$$

A question then arises naturally: are there optimal controls among full-strength controls? In other words, we study the intersection $\mathcal{U}_{\text{FS}} \cap \mathcal{U}_{\text{opt}}$. To answer this, we first look for an optimal control strategy among the restricted set of controls $\mathcal{U}_{\text{FS}}$, i.e. we consider the problem:

Problem 2. *Find* $\arg\min_{\alpha \in \mathcal{U}_{FS}} \mathbb{V}(T)$.

Introducing the partial mean $\bar{\xi}_{1,l} = \frac{1}{l} \sum_{i=1}^{l} \xi_i$, we design the following optimal control strategy to solve Problem 2.

Theorem 5.1 (Full-control strategy)**.**
Let $T > 0$. The strategy designed in Prop 2.1 to decrease $\dot{\mathbb{V}}$ instantaneously is an optimal control strategy for Problem 2. It can be explicitly described as follows:
Define $t_1 = 0$ and for $l \in \{2, ..., N\}$, $t_l = \frac{N}{N-1} \ln \left((l-1) \frac{N-1}{N} \frac{\bar{\xi}_{1,l-1}(0) - \xi_l(0)}{\bar{\xi}(0)} + 1 \right)$.
If there exists $l \in \{1, ..., N-1\}$ such that $T \in [t_l, t_{l+1}[$, then any strategy satisfying: $\xi_i(T) = \bar{\xi}_{1,l}(T)$ for every $i \in \{1, ..., l\}$, $\sum_{i=1}^{l} \alpha_i \equiv 1$ and $\alpha_i \equiv 0$ for every $i \in \{l+1, ..., N\}$ is optimal.
If $T \geq t_N$, then any strategy satisfying $\xi_i(T) = \bar{\xi}(T)$ for all $i \in \{1, ..., N\}$ and $\sum_{i=1}^{N} \alpha_i \equiv 1$ is optimal.
For instance, if $T \in [t_l, t_{l+1}[$, one optimal strategy would consist in defining the following piecewise constant control:

$$\forall k \leq l, \ \forall t \in [t_k, t_{k+1}[, \ \begin{cases} \alpha_i(t) = \frac{1}{k} \text{ if } i \leq k \\ \alpha_i(t) = 0 \text{ if } i > k. \end{cases} \tag{32}$$

Proof. Let us first show that if $T \geq t_l$, then the optimal control strategy for Problem 2 must achieve $\xi_i(T) = \bar{\xi}_{1,l}(T)$ for all $i \in \{1, ..., l\}$, reasoning by contradiction.
Suppose that there exists $k \in \{1, ..., l\}$ such that $\xi_k(T) \neq \bar{\xi}_{1,l}(T)$. Using Hyp. 2, we can suppose that there exists $m < l$ such that for every $i \in \{1, ..., m\}$, $\xi_i(T) = \bar{\xi}_{1,m}(T)$, and for every $i \in \{1, ..., m\}$ and $j \in \{m+1, ..., N\}$, $\xi_j(T) < \xi_i(T)$.

Let $j \in \{m+1, ..., l\}$. The transversality condition (15) gives: for every $i \in \{1, ..., m\}$, $\lambda_j(T) < \lambda_i(T)$. According to Proposition 3.1, for all $t \in [0, T]$, for every $i \in \{1, ..., m\}$, $\lambda_j(t) < \lambda_i(t)$. According to the PMP, as seen in Section 3, only the biggest covectors are controlled, and since $\alpha \in \mathcal{U}_{\text{FS}}$, with maximum control. So $\sum_{i=1}^{m} \alpha_i \equiv 1$ and $\alpha_j \equiv 0$. The evolutions of ξ_j and $\bar{\xi}_{1,m}$ are then given by:

$$\begin{cases} \dot{\xi}_j = -\xi_j + \bar{\xi} \\ \dot{\bar{\xi}}_{1,m} = -\bar{\xi}_{1,m} + \frac{m-1}{m} \bar{\xi}. \end{cases} \tag{33}$$

Since $\sum_{i=1}^{N} \alpha_i \equiv 1$, the evolution of the mean is given by $\dot{\bar{\xi}} = -\frac{1}{N} \bar{\xi}$, and we can compute $\bar{\xi} = \bar{\xi}(0) e^{-t/N}$, which in turn allows us to solve:

$$\forall t \in [0, T], \begin{cases} \xi_j(t) = e^{-t} \left(\xi_j(0) + \frac{N}{N-1} \bar{\xi}(0) (e^{\frac{N-1}{N} t} - 1) \right) \\ \bar{\xi}_{1,m}(t) = e^{-t} \left(\bar{\xi}_{1,m}(0) + \frac{m-1}{m} \frac{N}{N-1} \bar{\xi}(0) (e^{\frac{N-1}{N} t} - 1) \right) \end{cases} \tag{34}$$

We get:

$$(\bar{\xi}_{1,m} - \xi_j)(T) = e^{-T} \left(\bar{\xi}_{1,m}(0) - \xi_j(0) - \frac{1}{m} \frac{N}{N-1} \bar{\xi}(0) (e^{\frac{N-1}{N} T} - 1) \right). \tag{35}$$

We made the hypothesis that $T \geq t_l = \frac{N}{N-1} \ln \left((l-1) \frac{N-1}{N} \frac{\bar{\xi}_{1,l-1}(0) - \xi_l(0)}{\bar{\xi}(0)} + 1 \right)$. Hence,

$$\begin{aligned} (\bar{\xi}_{1,m} - \xi_j)(T) &\leq e^{-T} \left(\bar{\xi}_{1,m}(0) - \xi_j(0) - \frac{1}{m} (l-1) (\bar{\xi}_{1,l-1}(0) - \xi_l(0)) \right) \\ &= \frac{1}{m} e^{-T} \left[m \bar{\xi}_{1,m}(0) - m \xi_j(0) - (l-1) \bar{\xi}_{1,l-1}(0) + (l-1) \xi_l(0) \right] \\ &\overset{(*)}{\leq} \frac{1}{m} e^{-T} \left[\sum_{i=1}^{m} \xi_i(0) - \sum_{i=1}^{l-1} \xi_i(0) + (l-1-m) \xi_l(0) \right] \\ &= \frac{1}{m} e^{-T} \left[- \sum_{i=m+1}^{l-1} \xi_i(0) + (l-1-m) \xi_l(0) \right] \\ &\overset{(*)}{\leq} \frac{1}{m} e^{-T} \left[-(l-1-m) \xi_l(0) + (l-1-m) \xi_l(0) \right] \\ &= 0, \end{aligned} \tag{36}$$

where inequalities $(*)$ derive from Hypothesis 1: since $j \leq l$, $\xi_j(0) \geq \xi_l(0)$. However, $(\bar{\xi}_{1,m} - \xi_j)(T) \leq 0$ contradicts that $\xi_j(T) < \xi_i(T)$ for every $i \in \{1, ..., m\}$. From this we conclude that if $T \geq t_l$, then for every $i \in \{1, ..., l\}$, $\xi_i(T) = \bar{\xi}_{1,l}(T)$ for an optimal control strategy fulfilling Hypothesis 2.

Let us now show that if $T < t_{l+1}$, then for every $k \in \{l+1, ..., N\}$, $\alpha_i \equiv 0$ and $\xi_k(T) < \bar{\xi}_{1,l}(T)$.

$$
\begin{aligned}
\bar{\xi}_{1,l}(T) - \xi_k(T) &\overset{(1)}{=} e^{-T}\left(\bar{\xi}_{1,l}(0) - \xi_k(0) - \int_0^T e^{\frac{N-1}{N}s}(\frac{1}{l}\sum_{j=1}^{l}\alpha_j - \alpha_k)(s)\bar{\xi}(0)ds\right) \\
&\overset{(2)}{\geq} e^{-T}\left(\bar{\xi}_{1,l}(0) - \xi_k(0) - \int_0^T e^{\frac{N-1}{N}s}\frac{1}{l}\bar{\xi}(0)ds\right) \\
&= e^{-T}\left(\bar{\xi}_{1,l}(0) - \xi_k(0) - \frac{N}{N-1}\frac{1}{l}\bar{\xi}(0)(e^{\frac{N-1}{N}T} - 1)\right) \\
&\overset{(3)}{>} e^{-T}\left(\bar{\xi}_{1,l}(0) - \xi_k(0) - (\bar{\xi}_{1,l}(0) - \xi_{l+1}(0))\right) \\
&= e^{-T}\left(\xi_{l+1}(0) - \xi_k(0)\right) \\
&\overset{(4)}{\geq} 0,
\end{aligned}
\tag{37}
$$

where
(1) was computed using the evolutions of ξ_k and $\bar{\xi}_{1,l}$: $\dot{\xi}_k = -\xi_k + (1-\alpha_k)\bar{\xi}$ and $\dot{\bar{\xi}}_{1,l} = -\bar{\xi}_{1,l} + (1 - \frac{1}{l}\sum_{i=1}^{l}\alpha_i)\bar{\xi}$,
(2) was obtained from inequalities $\sum_{j=1}^{l}\alpha_j(t) \leq 1$ and $\alpha_k(t) \geq 0$ for all t,
(3) comes from the inequality: $T < t_{l+1} = \frac{N}{N-1}\ln(\frac{N-1}{N}l\frac{\bar{\xi}_{1,l}(0)-\xi_{l+1}(0)}{\bar{\xi}(0)} + 1)$,
(4) derives from Hypothesis 1 since $k \geq l+1$.
Hence, for every $k \in \{l+1, ..., N\}$, $\xi_k(T) \geq \bar{\xi}_{1,l}(T)$. Furthermore, the transversality condition (15) along with Proposition 3.1 imply that for all $t \in [0,T]$ for every $i \in \{1, ..., l\}$, $\lambda_k(t) < \lambda_i(t)$ and the Pontryagin Maximum Principle as seen in Section 3 states that $\alpha_k \equiv 0$. So $\xi_k(T) \geq \bar{\xi}_{1,l}(T)$.

We proved that if $T \in [t_l, t_{l+1}[$, then for every $i \in \{l+1, ..., N\}$, $\alpha_i \equiv 0$. Since $\bar{\xi}$ is fully determined as $\alpha \in \mathcal{U}_{\mathrm{FS}}$, this means that for all $i \in \{l+1, ..., N\}$, $\xi_i(T)$ is also fully determined (satisfying the equation $\dot{\xi}_i = -\xi_i + \bar{\xi}$). On the other hand, we proved that for all $i \in \{1, ..., l\}$, $\xi_i(T) = \bar{\xi}_{1,l}(T)$ and that $\sum_{i=1}^{l}\alpha_i \equiv 1$, so $\bar{\xi}_{1,l}$ is also fully determined (satisfying the equation $\dot{\bar{\xi}}_{1,l} = -\bar{\xi}_{1,l} + \frac{l-1}{l}\bar{\xi}$). Hence, any strategy such that for all $i \in \{1, ..., l\}$, $\xi_i(T) = \xi_{1,l}(T)$ with $\sum_{i=1}^{l}\alpha_i \equiv 1$ and for all $i \in \{l+1, ..., N\}$, $\alpha_i \equiv 0$ is optimal for Problem 2. □

Notice that this optimal control strategy is not sparse, as control is split among more and more agents as time goes. However, it is not unique and one could very well act on one agent at a time until all reach the known final velocities. Going back to the general Problem 1, we prove that under certain conditions, the optimal control strategy uses full strength at all time, i.e. $\alpha^{\mathrm{opt}} \in \mathcal{U}_{FS}$.

Theorem 5.2 (Sufficient condition for full control)**.**
Define the time $t_N = \frac{N}{N-1}\ln\left(\frac{(N-1)^2}{N}\frac{\bar{\xi}_{1,N-1}(0)-\xi_N(0)}{\bar{\xi}(0)} + 1\right)$ *as in Theorem 5.1.*
If $T \geq t_N$*, then the optimal strategies* α^{opt} *to Problem 1 belong to* $\mathcal{U}_{FS}$ *and for these controls* $\xi_i(T) = \bar{\xi}(T)$ *for every* $i \in \{1, ..., N\}$*.*

Proof. If $T \geq t_N$, then the instantaneous decrease strategy designed in Theorem 5.1 is optimal. Indeed, we noticed that the migration functional can be written as the sum of two terms (10): $\mathbb{V} = \bar{\xi}^2 + \frac{1}{N}\sum(\xi_i - \bar{\xi})^2$. The strategy designed in Theorem 5.1 minimizes $\bar{\xi}(T)$ by using full control at all time, hence minimizing $\bar{\xi}(T)^2$ since $\bar{\xi} > 0$. Furthermore, it achieves $\xi_i(T) = \bar{\xi}(T)$ for all $i \in \{1, ..., N\}$, thus minimizing the second term $\frac{1}{N}\sum(\xi_i - \bar{\xi})^2$. Hence any optimal control strategy has to use full control at all time and achieve $\xi_i(T) = \bar{\xi}(T)$ for every $i \in \{1, ..., N\}$ in order to perform as well. □

We finally address the general case stated in Problem 1: minimize $\mathbb{V}(T)$ over the set of controls $\mathcal{U}$, for any time T. In the following theorem, we show the existence of an initial "inativation" time interval: the optimal strategy can require to let the system evolve freely (i.e. without control) at initial time, before acting on it with full strength.

Theorem 5.3 (Inactivation Principle)**.**
If $T < t_N$, then one of the two holds: any control strategy α^{opt} either belongs to $\mathcal{U}_{FS}$ and the strategy designed in Theorem 5.1 is optimal, or there exists some $\delta < T$ such that $\alpha^{opt} \equiv 0$ on $[0, \delta]$, and $\sum \alpha_i^{opt} \equiv 1$ on $[\delta, T]$.

Proof. According to Hypothesis 2, we can assume that $\xi_1(T) \geq \xi_i(T)$ for every $i \in \{1, ..., N\}$. Furthermore, $\bar{\xi}(T) > 0$, so $\xi_1(T) > 0$. From the transversality condition (15) we deduce: $\lambda_1(T) \geq \lambda_i(T)$ for every $i \in \{1, ..., N\}$ and $\lambda_1(T) > 0$. From Prop. 3.1, we know that for all $t \in [0, T]$, $\lambda_1(t) \geq \lambda_i(t)$. According to Prop. 3.3, full control is used at time t if $\lambda_1(t) > 0$ and no control is used if $\lambda_1(t) < 0$. Let us study the evolution of λ_1: $\dot{\lambda}_1 = \frac{1}{N} \sum \alpha_j \lambda_j - \bar{\lambda} + \lambda_1$. By the Pontryagin maximum principle, we always have $\sum \alpha_j \lambda_j \geq 0$. Furthermore, $\lambda_1 - \bar{\lambda} \geq 0$. So $\dot{\lambda}_1(t) \geq 0$ for all $t \in [0, T]$. We show that $\lambda_1 = 0$ at most at one point. Indeed, suppose that $\lambda_1(\tau) = 0$ for some $\tau \in [0, T]$ and that $\dot{\lambda}_1(\tau) = 0$. Then $\dot{\lambda}_1(\tau) = -\bar{\lambda}(\tau)$ so $\bar{\lambda}(\tau) = \lambda_1(\tau) = 0$, and since the λ_i's are ordered, $\lambda_i(\tau) = \bar{\lambda}(\tau)$ for every $i \in \{1, ..., N\}$. According to Proposition 3.1, $\lambda_i(t) = \bar{\lambda}(t)$ for all time t and every i. Since $\lambda_1(T) > 0$, there exists a time interval $[\tau^*, T]$ such that $\lambda_1(t) > 0$ for all $t \in [\tau^*, T]$. On this interval, $\dot{\lambda}_1 = \frac{1}{N} \lambda_1 \sum_j \alpha_j = \frac{1}{N} \lambda_1$, which gives: $\lambda_1(T) = \lambda_1(\tau^*) e^{\frac{1}{N}(T - \tau^*)}$. This contradicts the existence of a time τ at which $\lambda_1(\tau) = 0$. In conclusion, if $\lambda_1(\tau) = 0$, then $\dot{\lambda}_1(\tau) > 0$ so $\lambda_1 = 0$ at most at one point.
Hence, there is a dichotomy of cases:
Either $\lambda_1(t) \geq 0$ for all time, so $I(t) \neq \emptyset$ for all t, which implies that $\alpha^{\text{opt}} \in \mathcal{U}_{\text{FS}}$ according to Prop. 3.3. In this case, $\arg\max_{\alpha \in \mathcal{U}} \mathbb{V} = \arg\max_{\alpha \in \mathcal{U}_{\text{FS}}} \mathbb{V}$ and the control strategy designed in Theorem 5.1 for Problem 2 is optimal also for Problem 1.
Or there exists $\delta \in [0, T]$ such that $\lambda_1(t) < 0$ on $[0, \delta[$ and $\lambda_1(t) \geq 0$ on $[\delta, T]$, which implies that $\alpha(t) \equiv 0$ on $[0, \delta]$ and $\sum \alpha_i(t) \equiv 1$ on $]\delta, T]$. Practically, an optimal control strategy would consist in letting the system evolve without control on $[0, \delta[$. Then the full-control strategy from Theorem 5.1 can be applied on $[\delta, T]$ with the new initial positions $\xi(\delta)$. ☐

Remark 5.4. *Although this result may seem counter-intuitive, in certain cases it makes sense to let the system evolve freely, at least initially. Indeed, without control the system naturally regroups in order to reach consensus, minimizing $\sum_{i=1}^{N} (\xi_i - \bar{\xi})$ in (10), but keeping $\bar{\xi}$ constant. Actual examples of such cases are shown in the next section.*

Remark 5.5. *Note that a constraint $M < 1$ would not change the nature of the results. It would only mean acting with lesser strength on the controlled agents, therefore changing the values of the times t_l defined in Theorem 5.1, but the optimal control strategy would be unchanged. With a constraint $M > 1$, we can expect results similar to those of Section 4, with two kinds of Inactivation periods, consisting either in letting the system evolve freely, or in controlling it with a non-maximal total strength $0 < \sum_i \alpha_i < M$ (see Theorem 4.7 (ii) and (iii)).*

5.2 Practical Approach

We proved in the previous section that the optimal strategy can either be to act with full control as in Theorem 5.1, or to let the system evolve without control on some time interval $[0, \delta]$, before acting with full control on $]\delta, T]$. The migration functional $\mathbb{V}_\delta$ can be computed explicitly as a function of δ. We then look for the value of δ that minimizes $\mathbb{V}_\delta(T)$.

Let us denote by ξ^δ the solution to system (29) when no control is applied on $[0, \delta]$ and full control is used on $]\delta, T]$. Equation (29) gives:

$$\begin{cases} \dot{\xi}_i^\delta = -\xi_i^\delta + \bar{\xi}^\delta \\ \dot{\bar{\xi}}^\delta = 0, \end{cases} \qquad \text{on } [0, \delta], \tag{38}$$

which allows us to solve: $\xi_i^\delta(\delta) = e^{-\delta} \left(\xi_i^\delta(0) + \bar{\xi}^\delta(0)(e^\delta - 1) \right)$.

We then apply the strategy designed in Theorem 5.1 with the new initial conditions $\xi^\delta(\delta)$ and the new final time $T - \delta$. Define the times $t_1^\delta = 0$ and for $l \in \{2, ..., N\}$, $t_l^\delta = \frac{N}{N-1} \ln \left((l-1) \frac{N-1}{N} \frac{\bar{\xi}_{1,l-1}^\delta(\delta) - \xi_l^\delta(\delta)}{\bar{\xi}^\delta(\delta)} + 1 \right)$.
Find $l \in \{1, ..., N-1\}$, such that $T - \delta \in [t_l^\delta, t_{l+1}^\delta[$. Then any strategy satisfying $\xi_i^\delta(T) = \bar{\xi}_{1,l}^\delta(T)$ for every

$i \in \{1,...,l\}$, $\sum_{i=1}^{l} \alpha_i(t) = 1$ for all $t \in [\delta, T]$, and $\alpha_i \equiv 0$ for every $i \in \{l+1,...,N\}$ is optimal. From equation (29) we get:

$$\begin{cases} \dot{\bar{\xi}}^{\delta}_{1,l} = -\bar{\xi}^{\delta}_{1,l} + \frac{l-1}{l}\bar{\xi}^{\delta} \\ \dot{\xi}^{\delta}_{i} = -\xi^{\delta}_{i} + \bar{\xi}^{\delta} \quad \text{for } i \in \{l+1,...,N\} \qquad \text{on } [\delta, T], \\ \dot{\bar{\xi}}^{\delta} = -\frac{1}{N}\bar{\xi}^{\delta}, \end{cases} \tag{39}$$

from which we can solve:

$$\begin{cases} \xi^{\delta}_i(T) = \bar{\xi}^{\delta}_{1,l}(T) = e^{-(T-\delta)}\left(\bar{\xi}^{\delta}_{1,l}(\delta) + \frac{l-1}{l}\frac{N}{N-1}\bar{\xi}^{\delta}(0)(e^{\frac{N-1}{N}(T-\delta)} - 1)\right) \quad \text{for all } i \in \{1,...,l\}, \\ \xi^{\delta}_i(T) = e^{-(T-\delta)}\left(\xi^{\delta}_i(\delta) + \frac{N}{N-1}\bar{\xi}^{\delta}(0)(e^{\frac{N-1}{N}(T-\delta)} - 1)\right) \quad \text{for all } i \in \{l+1,...,N\}. \end{cases} \tag{40}$$

We then compute $\mathbb{V}^{\delta}(T) = \frac{1}{N}\sum_{i=1}^{N} \xi^{\delta}_i(T)^2$ and look for $\min_{\delta \in [0,T]} \mathbb{V}^{\delta}(T)$ (see Figure 3).

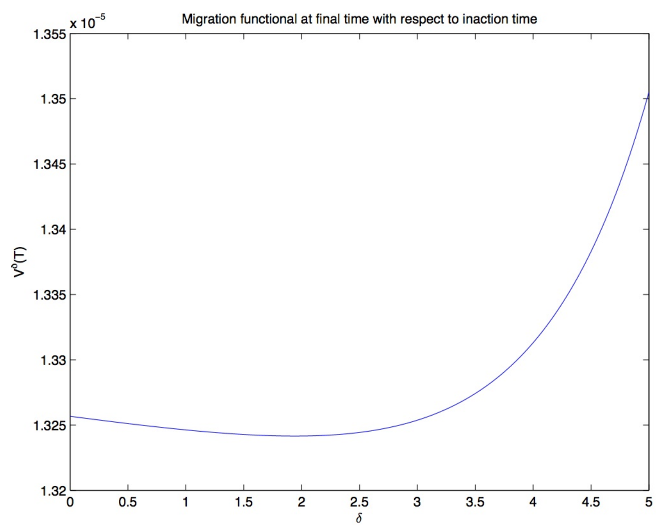


Figure 3: $V^{\delta}(T)$ with respect to Inactivation time δ. Here the optimal Inactivation time is $\delta = 1.94$.

Series of simulations were then run to look for cases in which $\delta > 0$. Tables 1 and 2 list the number of such cases found over 1000 simulations, for different values of the number of agents and of the final time. For the first series of simulations (Table 1), initial projected variables $\xi_i(0)$ were chosen randomly in the interval $[-1, 1]$ and such that the mean $\bar{\xi}$ is strictly positive. In Table 2, initial projected variables were chosen randomly in the interval $[-0.5, 1.5]$. As expected (and proven in Theorem 5.2), for larger values of T, it is always optimal to act with full control at all time (in other words $\delta = 0$). One also notices that as the number of agents increases, "Inactivation" cases become less and less frequent. They are also less frequent when $\bar{\xi}(0)$ is larger (Table 2). This can be explained again by looking at the two terms in the migration functional $\mathbb{V}$ (10). When $\bar{\xi}$ is small (as is the case in the first set of simulations), one can concentrate on minimizing the second term $\frac{1}{N}\sum(\xi_i - \bar{\xi})^2$, which does not necessarily require full control since the system naturally evolves to minimize this term. On the other hand, when $\bar{\xi}$ is large, full control is required to minimize it, which explains the smaller number of "Inactivation" cases in the second set of simulations.

Table 3 shows the average of the relative difference $\frac{\mathbb{V}_{\text{fc}} - \mathbb{V}^{\delta}}{\mathbb{V}_{\text{fc}}}$, where $\mathbb{V}^{\delta}$ was obtained by using optimal control and $\mathbb{V}_{\text{fc}}$ by using full control at all time (as designed in Theorem 5.1), in the case of the first set of simulations, i.e. $\xi_i(0) \in [-1, 1]$. The gain in performance when using the optimal strategy is minor (significantly less than 1% in most cases), and decreases as the number of agents increases.

Hence, $\mathcal{U}_{\text{opt}} \cap \mathcal{U}_{\text{FS}} = \emptyset$ occurs in very few cases, and when it does, the gain in performance compared to the full control strategy is hardly significant. For reasons of computational speed and complexity, it is very reasonable to neglect those cases and to apply the full control strategy at all time.

Number of agents	5	10	20	50
T=3	16 ‰	9 ‰	0	0
T=4	18 ‰	7 ‰	3 ‰	0
T=5	10 ‰	2 ‰	2 ‰	0
T=6	2 ‰	1 ‰	0	1 ‰
T=7	0	0	0	0

Table 1: Number of cases in which $\delta > 0$ out of 1000 simulations. $\xi_i(0)$ chosen randomly in $[-1, 1]$.

Number of agents	5	10	20	50
T=3	4 ‰	1 ‰	0	0
T=4	2 ‰	0	0	0
T=5	1 ‰	0	0	0
T=6	0	0	0	0
T=7	0	0	0	0

Table 2: Number of cases in which $\delta > 0$ out of 1000 simulations. $\xi_i(0)$ chosen randomly in $[-0.5, 1.5]$.

Number of agents	5	10	20	50
T=3	0.073%	0.001%	-	-
T=4	0.27%	0.018%	0.001%	-
T=5	0.91%	0.056%	0.0069%	-
T=6	1.53%	0.2%	-	0.00003 %

Table 3: Relative improvement of $\mathbb{V}^\delta$ w.r.t. $\mathbb{V}_{\text{fc}}$

Figure 4 shows the evolution of the projected velocities ξ_i, $i \in \{1, ..., 10\}$ with respect to time, in a case where the optimal strategy requires full control at all time. The control function is the one designed in Theorem 5.1 and acts first on ξ_1, then on ξ_1 and ξ_2, and so on until all have reached consensus, at which point it acts with equal strength on all agents to drive $\bar{\xi}$ down to 0.

6 Optimal control for integral cost

In this section we focus on minimizing the integral of the migration functional, with the constraint on the controls $M = 1$. As done in Section 5, we define two problems (where $\mathcal{U}$ (28) and $\mathcal{U}_{\text{FS}}$ (30) are defined as before).

Problem 3. *Find* $\arg\min\limits_{\alpha \in \mathcal{U}} \int_0^T \mathbb{V}(t)dt$.

Problem 4. *Find* $\arg\min\limits_{\alpha \in \mathcal{U}_{FS}} \int_0^T \mathbb{V}(t)dt$.

6.1 Pontryagin's Maximum Principle

We first prove general results, with the aim of solving Problem 3. In order to use Pontryagin's maximum principle, we introduce the new hamiltonian $H = \langle \lambda, f \rangle + \lambda^0 \mathbb{V}$ and the equations governing the covectors' evolution $\dot{\lambda}_i = -\frac{\partial H}{\partial \xi_i}$. Considering normal trajectories, we set $\lambda^0 = 1$ and obtain:

$$\begin{cases} H = \sum\limits_{i=1}^{N} (-\lambda_i \xi_i) + \bar{\xi} \sum\limits_{i=1}^{N} (1 - \alpha_i)\lambda_i + \sum\limits_{i=1}^{N} \xi_i^2 \\ \dot{\lambda}_i = \lambda_i - \frac{1}{N} \sum_j (1 - \alpha_j)\lambda_j - 2\xi_i \end{cases} \tag{41}$$

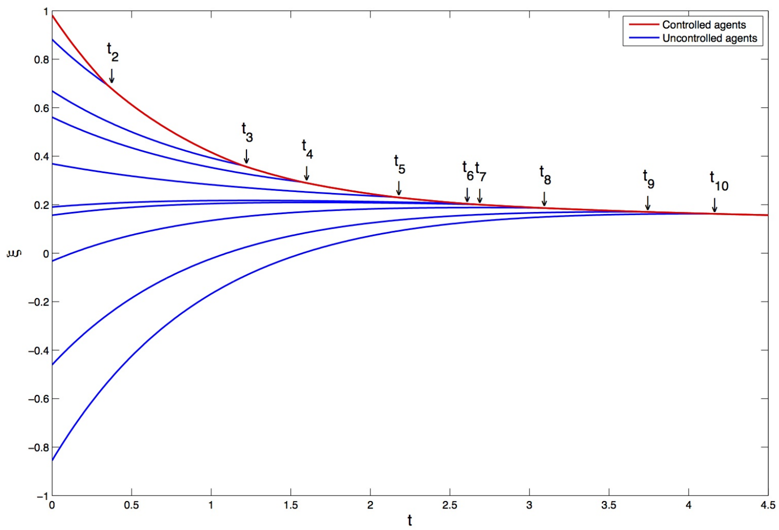


Figure 4: Evolution of the projected velocities ξ_i with the full strength optimal control for a system of 10 agents. In this example $\bar{\xi}(0) = 0.25$ so full control at all time is needed to drive $\bar{\xi}$ to the desired velocity $V = 0$ (i.e. $\delta = 0$). At final time $T = 4.5$ the system has reached consensus, but not yet at the desired velocity.

Since the final condition is not fixed, we have the following transversality condition for the covectors:

$$\lambda(T) = 0. \tag{42}$$

As in the minimization of the migration functional at final time (Section 3), we define I_λ and I_λ^+ (see (18)). Then minimizing $H = \sum_{i=1}^{N} -\alpha_i \lambda_i + \tilde{H}$ (where $\tilde{H}$ contains only uncontrolled terms) requires the following : if $k \notin I_\lambda$, $\alpha_k = 0$; furthermore, if $I_\lambda^+ \neq \emptyset$, then $\sum_{i \in I_\lambda^+} \alpha_i = 1$.

As in Section 5, we make Hypothesis 1. Given the initial order on the agents' projected velocities ξ_i, we prove the following:

Lemma 6.1. *There exists an optimal control strategy satisfying:*

$$\forall t \in [0, T],\ \forall i, j \in \{1, ..., N\},\ i < j \Rightarrow \xi_i(t) \geq \xi_j(t) \tag{43}$$

Proof. The proof is very similar to that of Lemma 3.2. Consider an optimal control strategy $\alpha \in \mathcal{U}$. Define $\tau = \sup\{t \mid \exists \beta \in \mathcal{U} \text{ s.t. } \int_0^T \mathbb{V}_\beta(s)ds = \int_0^T \mathbb{V}_\alpha(s)ds \text{ and } \xi^\beta \text{ satisfies (43) on } [0,t]\}$. Let us prove by contradiction that $\tau = T$. Suppose that $\tau < T$. Then there exist $i, j \in \{1, ..., N\}$ with $i < j$ such that $\xi_i^\beta(\tau) = \xi_j^\beta(\tau)$ and $\xi_j^\beta(t) > \xi_i^\beta(t)$ on $]\tau, \tau + \delta]$ for some $\delta > 0$. Design a control strategy $\tilde{\beta}$ such that on $[\tau, T]$, $\tilde{\beta}_i = \beta_j$, $\tilde{\beta}_j = \beta_i$ and for every $k \in \{1, ..., N\} \setminus \{i, j\}$, $\tilde{\beta}_k = \beta_k$. Then for all $t \in [\tau, T]$, $\xi_i^{\tilde{\beta}}(t) = \xi_j^\beta(t)$, and $\xi_j^{\tilde{\beta}}(t) = \xi_i^\beta(t)$. So for all $t \in [\tau, \tau + \delta]$, $\xi_i^{\tilde{\beta}}(t) \geq \xi_j^{\tilde{\beta}}(t)$ and for all $t \in [0, T]$, $\mathbb{V}^{\tilde{\beta}}(t) = \mathbb{V}^\beta(t)$. Proceeding likewise for every pair of indices (m, n) satisfying $m < n$ and $\xi_n^\beta(t) > \xi_n^\beta(t)$ on $]\tau, \tau+\delta]$ we are able to design a control strategy $\tilde{\beta}$ satisfying (43) on $[0, \tau + \delta]$ and $\int_0^T \mathbb{V}_\beta(t)dt = \int_0^T \mathbb{V}_\alpha(t)dt$, which contradicts the definition of τ. In conclusion, $\tau = T$, i.e. for all $t \in [0, T]$, for every $i, j \in \{1, ..., N\}$, $i < j \Rightarrow \xi_i(t) \geq \xi_j(t)$. $\square$

Hence, as in Section 5, we can assume Hypothesis 2: for all $t \in [0, T]$, if $i < j$, then $\xi_i(t) \geq \xi_j(t)$. By the following proposition, we shall prove that the same order is observed among the covectors λ_i.

Proposition 6.2.

$$\forall t \in [0, T],\ i < j \Rightarrow \lambda_i(t) \geq \lambda_j(t). \tag{44}$$

Proof. Let us reason by contradiction. Suppose that there exists $\tau \in [0, T]$ such that for some i<j, $(\lambda_i - \lambda_j)(\tau) < 0$. From the evolution of the covectors (41) we derive for all $t \geq \tau$: $(\lambda_i - \lambda_j)(t) = e^{t-\tau}\left((\lambda_i - \lambda_j)(\tau) - 2\int_\tau^t e^{-(s-\tau)}(\xi_i - \xi_j)(s)ds\right)$. Since $(\lambda_i - \lambda_j)(\tau) < 0$ and for all $s \in [0, T]$, $(\xi_i - \xi_j)(s) \geq 0$, we deduce that for all $t \in [\tau, T]$, $(\lambda_i - \lambda_j)(t) < 0$, which contradicts the final condition (42). ☐

Proposition 6.3. *Let $\tau \in [0, T]$ and $i, j \in \{1, ..., N\}$, such that $(\lambda_i - \lambda_j)(\tau) = 0$. Then for all $t \geq \tau$, $(\lambda_i - \lambda_j)(t) = 0$ and $(\xi_i - \xi_j)(t) = 0$.*

Proof. Let $\tau \in [0, T]$ and $i, j \in \{1, ..., N\}$, such that $(\lambda_i - \lambda_j)(\tau) = 0$. Then for all $t \geq \tau$,

$$(\lambda_i - \lambda_j)(t) = -2e^{t-\tau}\int_\tau^t e^{-(s-\tau)}(\xi_i - \xi_j)(s)ds. \tag{45}$$

Suppose for instance that $i < j$. According to Proposition 6.2, for all $t \in [0, T]$, $(\lambda_i - \lambda_j)(t) \geq 0$. Since we made Hypothesis 2, the right-hand side of equation (45) is nonpositive. This is only possible if both sides are equally zero. Hence, for all $t \geq \tau$, $(\lambda_i - \lambda_j)(t) = 0$ and $(\xi_i - \xi_j)(t) = 0$. ☐

The following proposition states that if at a certain point in time, two agents have the same projected velocities, then these should stay identical until final time.

Proposition 6.4. *Suppose that there exists $\tau \in [0, T]$ and $i, j \in \{1, ..., N\}$ such that $\xi_i(\tau) = \xi_j(\tau)$. Then*

$$\text{for all } t \geq \tau,\ \xi_i(t) = \xi_j(t). \tag{46}$$

As a consequence, for almost all $t \geq \tau$, $\alpha_i(t) = \alpha_j(t)$.

Proof. Let $\tau \in [0, T]$ and $i, j \in \{1, ..., N\}$. Define $\tilde{\tau} = \sup\{t \geq \tau \mid \xi_i(t) = \xi_j(t) \text{ for all } t \in [\tau, \tilde{\tau}]\}$. Notice from (29) that this implies that $\alpha_i(t) = \alpha_j(t)$ for almost every $t \in [\tau, \tilde{\tau}]$. Let us prove that $\tilde{\tau} = T$.
Suppose that $\tilde{\tau} < T$. Then there exists $\delta > 0$ such that for all $t \in]\tilde{\tau}, \tilde{\tau} + \delta]$, $\xi_i(t) \neq \xi_j(t)$. Define β such that $\beta = \alpha$ on $[0, \tilde{\tau}]$ and

$$\begin{cases} \beta_i = \beta_j = \frac{1}{2}(\alpha_i + \alpha_j) \\ \beta_k = \alpha_k \text{ for } k \neq i,\ k \neq j \end{cases} \qquad \text{on }]\tilde{\tau}, T],$$

and denote by ξ^β the corresponding trajectory. Notice that $\sum_k \alpha_k \equiv \sum_k \beta_k$, so according to (29), $\bar{\xi} \equiv \bar{\xi}^\beta$. This implies that $\xi_k = \xi_k^\beta$ for all $k \neq i, j$. Moreover, $\alpha_i + \alpha_j \equiv \beta_i + \beta_j$ so for all $t \in [\tilde{\tau}, T]$, $(\xi_i + \xi_j)(t) = (\xi_i^\beta + \xi_j^\beta)(t)$. Furthermore, ξ_i^β and ξ_j^β satisfy the same differential equation on $[\tilde{\tau}, T]$ and $\xi_i^\beta(\tau) = \xi_j^\beta(\tau)$, so for all $t \in [\tilde{\tau}, T]$, $\xi_i^\beta(t) = \xi_j^\beta(t) = \frac{1}{2}(\xi_i + \xi_j)(t)$. Define $\mathbb{V}_\alpha$ and $\mathbb{V}_\beta$ as the cost functions associated respectively with the controls α and β. Then $\mathbb{V}_\beta = \mathbb{V}_\alpha$ on $[0, \tilde{\tau}]$. On $]\tilde{\tau}, T]$,

$$\begin{aligned} \mathbb{V}_\alpha - \mathbb{V}_\beta &= \sum_k (\xi_k)^2 - \sum_k (\xi_k^\beta)^2 = (\xi_i)^2 + (\xi_j)^2 - (\xi_i^\beta)^2 - (\xi_j^\beta)^2 \\ &= (\xi_i)^2 + (\xi_j)^2 - 2(\frac{1}{2}(\xi_i + \xi_j))^2 = (\xi_i - \xi_j)^2. \end{aligned}$$

Hence, for all $t \in]\tilde{\tau}, \tilde{\tau} + \delta]$, $\mathbb{V}_\alpha(t) > \mathbb{V}_\beta(t)$, and for all $t \in [\tilde{\tau} + \delta, T]$, $\mathbb{V}_\alpha(t) \geq \mathbb{V}_\beta(t)$. We get $\int_0^T \mathbb{V}_\beta < \int_0^T \mathbb{V}_\alpha$, which contradicts that α is an optimal control. In conclusion, $\tau = T$, which proves the proposition. ☐

6.2 Optimal full-strength control

We design an optimal control strategy for Problem 4:

Theorem 6.5. *Let $J(t) = \{i \in \{1, ..., N\} \mid \xi_i(t) = \max_j \xi_j(t)\}$. The following control α is optimal for Problem 4:*

$$\begin{cases} \forall i \in J(t),\ \alpha_i(t) = \frac{1}{|J(t)|} \\ \forall i \notin J(t),\ \alpha_i(t) = 0. \end{cases} \tag{47}$$

Proof. According to Pontryagin's maximum principle and the expression of the Hamiltonian (41), the optimal control strategy solving Problem 4 requires to set $\sum_{i\in I(t)} \alpha_i(t) = 1$ and $\alpha_k(t) = 0$ for $k \notin I(t)$, where $I(t) := \{i \mid \lambda_i(t) = \max_j \lambda_j(t)\}$. Furthermore, according to Proposition 6.3, if $\lambda_i(\bar{t}) = \lambda_j(\bar{t})$, then $\xi_i(t) = \xi_j(t)$ for all $t \geq \bar{t}$, and according to Proposition 6.4, $\alpha_i(t) = \alpha_j(t)$ for almost every $t \geq \bar{t}$. Hence, the optimal strategy in fact requires to set, for almost every $t \in [0, T]$,

$$\begin{cases} \forall i \in I(t),\ \alpha_i(t) = \frac{1}{|I(t)|} \\ \forall i \notin I(t),\ \alpha_i(t) = 0, \end{cases} \tag{48}$$

where $|\cdot|$ denotes the cardinality of a set. Let us prove that $I(t) = J(t)$ for almost every t. Assume that $i \in I(t)$ and (48) holds true. According to Proposition 6.2, the covectors are ordered, so $\lambda_1(t) = \cdots = \lambda_i(t)$. From Proposition 6.3 and Hypothesis 2, this implies $\xi_1(t) = \cdots = \xi_i(t)$, so $i \in J(t)$. Conversely, assume that $i \in J(t)$. Then from Hypothesis 2, $\xi_1(t) = \cdots = \xi_i(t)$. According to Proposition 6.4, $\alpha_1(t) = \cdots = \alpha_i(t)$. Since $\alpha(t)$ verifies (48), we deduce that $i \in I(t)$. Therefore, $I(t) = J(t)$ for almost every $t \in [0, T]$ and the optimal strategies (48) and (47) are equivalent.

□

Notice that the control strategy in the case of integral cost minimization with full control (Problem 4) is more restrictive than the optimal strategy to minimize the final value of the migration functional with full control (Problem 2) seen in Section 5. Indeed, this control strategy cannot be sparse. In order to minimize $\int_0^T \mathbb{V}(t)dt$, one has to split the control among more and more agents. However, any optimal control solving Problem 4 is also optimal for Problem 2.

6.3 Optimal control in the general case

After designing the optimal strategy for Problem 4, we show that Problems 3 and 4 are actually equivalent, i.e. that the optimal control solving Problem 3 belongs to $\mathcal{U}_{\text{FS}}$.

Theorem 6.6. *The optimal control strategy for Problem 3 requires using full-strength control, i.e.* $\alpha \in \mathcal{U}_{FS}$.

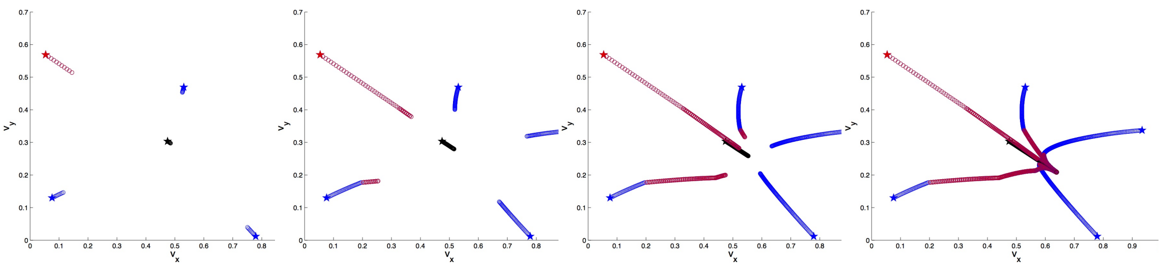
Proof. According to the Pontryagin Maximum Principle (see Section 6.1), if $\lambda_1(t) > 0$ for all t, then full control must be used at all time. Combining the final condition (42) and the evolution (41), we get $\lambda_1(T) = 0$ and $\dot{\lambda}_1(T) = -2\xi_1(T) < 0$. Hence there exists an interval $]t, T[$ on which $\lambda_1 > 0$. Let $\tau = \inf\{t \in [0, T] \text{ s.t. } \lambda_1(s) > 0 \text{ for all } s \in]t, T[\}$. Suppose that $\tau > 0$. Then $\lambda_1(\tau) = 0$. Furthermore, $\dot{\lambda}_1(\tau) = (\lambda_1 - \bar{\lambda})(\tau) - 2\xi_1(\tau)$. We compute: $\dot{\lambda}_1 - \dot{\bar{\lambda}} = \lambda_1 - \bar{\lambda} - 2(\xi_1 - \bar{\xi})$. Denoting $\Lambda = \lambda_1 - \bar{\lambda}$, we get the following evolution backwards in time: $\dot{\Lambda} = -\Lambda + 2(\xi_1 - \bar{\xi})$. Recall that backwards in time, we also have: $\dot{\xi}_1 = \xi_1 - (1 - \alpha_1)\bar{\xi}$. If $\Lambda = \xi_1$, then $\dot{\Lambda} = \xi_1 - 2\bar{\xi} = \dot{\xi}_1 + (1 - \alpha_1)\bar{\xi} - 2\bar{\xi} = \dot{\xi}_1 - (1 + \alpha_1)\bar{\xi} < \dot{\xi}_1$. Since $\Lambda(T) = 0 < \xi_1(T)$, this implies that $\Lambda(t) < \xi_1(t)$ for all $t \in [\tau, T]$. Hence, $\dot{\lambda}_1(\tau) = \Lambda(\tau) - 2\xi_1(\tau) < 0$, which contradicts the definition of τ. We conclude that $\lambda_1(t) > 0$ for all $t \in]0, T[$, and that $\sum_i \alpha_i \equiv 1$. □

Hence, the control strategy designed in Theorem 6.5 is an optimal strategy for the minimization of integral cost, i.e. Problem 3. Unlike in the minimization of the final cost (Problem 1), there is no initial Inactivation period.

Figure 5 illustrates the control strategy designed in Theorem 6.5. In this example, 5 agents are to be controlled optimally to reach consensus at the target velocity $V = (1, 0)$. Initially (Figure 5a), only one agent is controlled, the agent with the biggest projected velocity over $\bar{v} - V$. The set $J(t) = \arg\max_{i\in\{1,\ldots,N\}} \langle v_i, \frac{\bar{v}-V}{\|\bar{v}-V\|} \rangle$ contains more and more agents as time goes (5b, 5c) and eventually, control is split evenly among all agents (see Figure 5d).

Acknowledgment

The authors acknowledge the support of the NSF project KI-Net, DMS Grant # 1107444. Benjamin Scharf acknowledges the support of the ERC-Starting Grant "High-Dimensional Sparse Optimal Control".



(a) $t = 0.1$ (b) $t = 0.48$ (c) $t = 1.35$ (d) $t = 6.0$

Figure 5: Control of 5 agents to reach the target velocity $V = (1, 0)$. Agents are represented in the velocity space, controlled ones in red, uncontrolled ones in blue, and the mean velocity in black. Initial positions are marked by stars.

7 Appendix

Proof. (Theorem 4.10)
First, let $\xi_1(0) \geq \xi_2(0) > 0$. According to Prop. 1.1, $\xi_1(t) > 0$ and $\xi_2(t) > 0$ for all $t \in [0, T]$. The transversality condition gives $\lambda_1(T) > 0$ and $\lambda_2(T) > 0$, and according to Prop. 4.9, $\lambda_1(t) > 0$ and $\lambda_2(t) > 0$ for all $t \leq T$. According to Pontryagin's maximum principle (see Section 4.1), the global strategy requires setting $\alpha_1 + \alpha_2 \equiv M$. In this case, $\bar{\xi}(t) = \bar{\xi}(0)\exp(-\frac{M}{2}t)$ does not depend on the choice of α_1 and α_2. Minimizing $\mathbb{V}$ (24) therefore amounts to minimizing $(\xi_1 - \xi_2)^2$.

(i) If $T \geq t_2$, we will show that in addition to satisfying $\alpha_1 + \alpha_2 \equiv M$, the optimal control α must achieve $\xi_1(T) = \xi_2(T)$. Such a control strategy exists, since for instance (as one can see by direct computation of (19)) the control $(\beta_1, \beta_2)(t) = (1, M-1)$ for all $t \in [0, t_2[$ and $(\beta_1, \beta_2)(t) = (M/2, M/2)$ for all $t \in [t_2, T]$ achieves $\xi_1^\beta(t) = \xi_2^\beta(t)$ for all $t \in [t_2, T]$, where ξ^β denotes the corresponding trajectory. Notice that β minimizes $\bar{\xi}(T)$ by using the full strength M of the control at all time (see (25)), and minimizes $(\xi_1 - \xi_2)^2(T)$, so it minimizes $\mathbb{V}(T)$ (see (21)). Hence, in order to be optimal, α must satisfy $\xi_1(T) = \xi_2(T)$ as well as $\alpha_1 + \alpha_2 \equiv M$.

(ii) If $T < t_2$, we will show that $(\alpha_1, \alpha_2) \equiv (1, M-1)$ and that ξ_1 and ξ_2 cannot be brought together (i.e. $\xi_1(T) > \xi_2(T)$). Indeed, knowing that $\alpha_1 + \alpha_2 \equiv M$, one can use (20) to compute: $(\xi_1 - \xi_2)(t) = e^{-t}\left((\xi_1 - \xi_2)(0) - \int_0^t (\alpha_1 - \alpha_2)(s)\bar{\xi}(s)e^s ds\right)$. Since $\bar{\xi}$ is fully determined, $t_{\min} := \min_{\alpha \in \mathcal{U}_M, \alpha_1 + \alpha_2 \equiv M}\{t \in [0, T]$ s.t. $(\xi_1 - \xi_2)(t) = 0\}$ is achieved by maximizing $(\alpha_1 - \alpha_2)$, which gives: $(\alpha_1, \alpha_2) \equiv (1, M-1)$. As seen previously, by direct computation of (19), $t_{\min} = t_2$ as defined above. Hence, if $T < t_2$, necessarily $\xi_1(T) > \xi_2(T)$. Then $\lambda_1(T) > \lambda_2(T)$ and according to Prop. 4.9, and to Prop. 3.1, $\lambda_1(t) > \lambda_2(t) > 0$ for all t. According to the PMP (see 4.1), the optimal strategy is $(\alpha_1, \alpha_2) \equiv (1, M-1)$.

Now let $\xi_1(0) > 0$, $\xi_2(0) < 0$ and $\bar{\xi}(0) > 0$. We then distinguish four subcases.
Firstly, let us prove that if $\xi_2(T) > 0$, then necessarily $T > t_1$. Indeed, if $0 < \xi_2(T) \leq \xi_1(T)$, then $0 < \lambda_2(T) \leq \lambda_1(T)$, and according to Proposition 4.9, $0 < \lambda_2(t) \leq \lambda_1(t)$ for all $t \in [0, T]$. According to the PMP (see Section 4.1), $\alpha_1 + \alpha_2 \equiv M$. Hence $\bar{\xi}(t) = \bar{\xi}(0)e^{-Mt/2}$, and $\xi_2(t) = e^{-t}(\xi_2(0) + \bar{\xi}(0)\int_0^t (1 - \alpha_2)e^{\frac{2-M}{2}s} ds)$. The minimum time $t_{\min}$ needed to achieve $\xi_2(t_{\min}) > 0$ is achieved for $(\alpha_1, \alpha_2) \equiv (1, M-1)$, which, after computation, gives $t_{\min} = t_1$ as defined above. Hence, if $\xi_2(T) \geq 0$, then $T > t_1$.

(iii) Let $T \geq t_2$. Let us prove that $\alpha_1 + \alpha_2 \equiv M$ and $\xi_1(T) = \xi_2(T)$. Such a control strategy exists. Indeed, take for example $(\beta_1, \beta_2)(t) = (1, M-1)$ on $[0, t_2]$ and $(\beta_1, \beta_2)(t) = (M/2, M/2)$ on $]t_2, T]$. Then, by direct computation of (19), $\xi_1^\beta(t) = \xi_2^\beta(t)$ for all $t \in [t_2, T]$ (where ξ^β denotes the trajectory corresponding to the control β). Furthermore, β is optimal since it minimizes $\bar{\xi}^\beta$ by using full control at all time and achieves $(\xi_1^\beta - \xi_2^\beta)^2(T) = 0$ (see (21)). In order to perform optimally, the control α must also satisfy $\alpha_1 + \alpha_2 \equiv M$ and $\xi_1(T) = \xi_2(T)$.

(iv) Let $T < t_0$. Since $t_0 < t_1$, then as proved above, $\xi_2(T) \leq 0$. Suppose that $\xi_2(T) = 0$. Then $\lambda_1(T) > \lambda_2(T) = 0$ and according to Proposition 4.9, $\lambda_1(t) > \lambda_2(t) = 0$ for all time t. Hence, $\alpha_1 \equiv 1$ (see Section 4.1). Then $\min_{\alpha_2}\{t \in [0, T]$ s.t. $\xi_2(t) = 0\} = t_0$ as defined above (obtained for

$\alpha_2 \equiv 0$). This contradicts the condition on T. Hence, if $T < t_0$, then $\xi_2(T) < 0$ and according to Proposition 4.9 and Section 4.1, $\lambda_2 < 0$ so $\alpha_2 \equiv 0$. However, there is no information on λ_1 other than $\dot{\lambda}_1 = \alpha_1/2\lambda_1 + \bar{\lambda} - \lambda_1 \geq 0$ and $\lambda_1(\tau) = 0$ implies $\dot{\lambda}_1(\tau) > 0$. Hence, as in the previous sections, there exists $t^* \in [0, T[$ such that $\lambda_1 < 0$ on $[0, t^*[$, $\lambda_1(t^*) = 0$ and $\lambda_1 > 0$ on $]t^*, T]$. This implies that $(\alpha_1, \alpha_2) = (0, 0)$ on $[0, t^*[$ and $(\alpha_1, \alpha_2) = (1, 0)$ on $[t^*, T[$.

(v) Let $t_0 \leq T \leq t_1$. We shall prove that $\xi_2(T) = 0$ and that $\alpha_1 \equiv 1$. As seen previously, if $T \leq t_1$, then $\xi_2(T) \leq 0$. Suppose that $\xi_2(T) < 0$. Then $\lambda_1(T) > 0$ and $\lambda_2(T) < 0$ which according to Proposition 4.9 gives $\lambda_2(t) < 0$ for all t, and according to the PMP (see Section 4.1), $\alpha_2 \equiv 0$. Then $\xi_2(t) = e^{-t}(\xi_2(0) + \bar{\xi}(0)\int_0^T e^{-\int_0^s \frac{1}{2}\alpha_1(r)dr} e^s ds)$. Thus $t_{\text{sup}} := \sup_{\alpha_1}\{\tau \in [0, T] \text{ s.t. } \xi_2(t) < 0 \text{ for all } t \in [0, \tau[\}$ is obtained for $\alpha_1 \equiv 1$ and by direct computation, $t_{\text{sup}} = t_0$. Since $T \geq t_0$, there exists $\tau \leq T$ such that $\xi_2(\tau) = 0$. However, by Proposition 1.1, once $\xi_2 = 0$ it cannot become negative again, which contradicts $\xi(T) < 0$. Therefore, $\xi_2(T) = 0$, and according to Proposition 4.9 and the PMP (Section 4.1), $\lambda_1(t) > 0$ for all $t \in [0, T]$ so $\alpha_1 \equiv 1$. Furthermore, if $\xi_2(\tau) = 0$, then $\dot{\xi}_2(\tau) = (1 - \alpha_2(\tau))\bar{\xi}(\tau) > 0$ since $\alpha_2 = M - \alpha_1 = M - 1 < 1$. According to Proposition 1.1, once ξ_2 becomes positive it cannot become zero again. Hence we must have $\xi_2(t) < 0$ for all $t < T$ and $\xi_2(T) = 0$.

(vi) Let $t_1 < T < t_2$. We shall prove that $0 < \xi_2(T) < \xi_1(T)$ and that $(\alpha_1, \alpha_2) \equiv (1, M - 1)$. As in the previous case, since $T \geq t_0$, one must have: $\xi_2(T) \geq 0$. Suppose that $\xi_2(T) = 0$. Then according to Proposition 4.9 and the PMP, $\alpha_1 \equiv 1$ and $\xi_2(t) = e^{-t}(\xi_2(0) + \bar{\xi}(0)\int_0^T (1 - \alpha_2)(s) e^{-\int_0^s \frac{1}{2}(1+\alpha_2)(r)dr} e^s ds)$. Then the minimum of $\xi_2(T)$ is obtained for $\alpha_2 \equiv M - 1$, so

$$\xi_2(T) \geq e^{-T}(\xi_2(0) + \bar{\xi}(0)\int_0^T (2 - M)e^{-\frac{1}{2}Ms} e^s ds) > e^{-T}(\xi_2(0) + \bar{\xi}(0)(e^{\frac{2-M}{2}t_1} - 1)) > 0$$

by definition of t_1. This contradicts $\xi_2(T) = 0$, so necessarily $\xi_2(T) > 0$. Then $\lambda_1(t) > 0$ and $\lambda_2(t) > 0$ for all t, which implies that $\alpha_1 + \alpha_2 \equiv M$. In this case we prove as in case *(ii)* that $\xi_1(T) > \xi_2(T)$, which implies $(\alpha_1, \alpha_2) \equiv (1, M - 1)$.

$\square$

References

[1] N. Bellomo, A. Bellouquid, *On the modeling of crowd dynamics: Looking at the beautiful shapes of swarms*, Netw. Heter. Media (6), 383–399, 2011.

[2] N. Bellomo, C. Bianca, M. Delitala *Complexity analysis and mathematical tools towards the modelling of living systems*, Physics of Life Reviews (6) 144–175, 2009.

[3] S. Berman, Q. Lindsey, M. S. Sakar, V. Kumar, and S. C. Pratt. *Experimental study and modeling of group retrieval in ants as an approach to collective transport in swarm robotic systems*, Proceedings of the IEEE, 99 (9):1470–1481, 2011.

[4] Alberto Bressan and Benedetto Piccoli, *Introduction to the Mathematical Theory of Control*, AIMS on Applied Math, Vol.2, 2007.

[5] S. Camazine, J. Deneubourg, N. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau. *Self organization in biological systems*, Princeton University Press, 2003.

[6] Marco Caponigro, Massimo Fornasier, Benedetto Piccoli, Emmanuel Trélat. *Sparse stabilization and optimal control of the Cucker-Smale model*, Math. Cont. Related Fields 3, no. 4, 447–466, 2013.

[7] Y. Chuang, Y. Huang, M. D'Orsogna, and A. Bertozzi. *Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials*, IEEE International Conference on Robotics and Automation, pages 2292–2299, 2007.

[8] I. Couzin and N. Franks. *Self-organized lane formation and optimized traffic flow in army ants*, Proc. R. Soc. Lond., B 270:139–146, 2002.

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

[10] E. Cristiani, B. Piccoli, and A. Tosin. *Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints,* In G. Naldi, L. Pareschi, G. Toscani, and N. Bellomo, editors, Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences, Modeling and Simulation in Science, Engineering and Technology. Birkhĺauser Boston, 2010.

[11] E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale modeling of granular flows with application to crowd dynamics*, Multiscale Model. Simul., 9(1):155–182, 2011.

[12] F. Cucker and S. Smale. *Emergent behavior in flocks*, IEEE Trans. Automat. Control, 52(5):852–862, 2007.

[13] S. R. X. Dall, L.-A. Giraldeau, O. Olsson, J. M. McNamara, and D. W. Stephens. *Information and its use by animals in evolutionary ecology*, Trends in Ecology & Evolution, 20(4):187–193, 2005.

[14] B. Düring, D. Matthes, and G. Toscani. *Kinetic equations modelling wealth redistribution: A comparison of approaches*, Phys. Rev. E, 78:056103, 2008.

[15] Massimo Fornasier, Benedetto Piccoli, and Francesco Rossi. *Mean-field sparse optimal control*, Phil. Trans. R. Soc. A., 2014.

[16] J.P. Gauthier, *The Inactivation Principle: Mathematical Solutions minimizing the Absolute Work and Biological Implications for the Planning of Arm Movements*, PLoS Comput. Biol. 4, no. 10, 2008.

[17] V. Guttal and I. D. Couzin. *Social interactions, information use, and the evolution of collective migration*, Proceedings of the National Academy of Sciences, 107 (37):16172–16177, 2010.

[18] S.-Y. Ha and E. Tadmor, *From particle to kinetic and hydrodynamic descriptions of flocking*, Kinet. Relat. Models , 1 , pp. 415–435, 2008.

[19] D. Horstmann. *From 1970 until present: The Keller-Segel model in chemotaxis and its consequences*, I. Jahresber. Dtsch. Math.-Ver., 105(3):103–165, 2003.

[20] D. Horstmann. *From 1970 until present: the Keller-Segel model in chemotaxis and its consequences*, II. Jahresber. Dtsch. Math.-Ver., 106:51–69, 2004.

[21] A. Jadbabaie, J. Lin, and A. S. Morse. Correction to: "Coordination of groups of mobile autonomous agents using nearest neighbor rules" [IEEE Trans. Automat. Control 48, no. 6, 2003. 988–1001; MR 1986266]. IEEE Trans. Automat. Control, 48(9):1675, 2003.

[22] E. F. Keller and L. A. Segel. *Initiation of slime mold aggregation viewed as an instability*, J. Theor. Biol., 26(3):399–415, 1970.

[23] J.-M. Lasry and P.-L. Lions. *Mean field games*, Jpn. J. Math. (3), 2(1):229–260, 2007.

[24] N. Leonard, *Multi-Agent System Dynamics: Bifurcation and Behavior of Animal Groups*, Proc. 9th IFAC Symposium on Nonlinear Control Systems.

[25] N. Leonard and E. Fiorelli. *Virtual leaders, artificial potentials and coordinated control of groups*, Proc. 40th IEEE Conf. Decision Contr., pages 2968–2973, 2001.

[26] S. Motsch, E. Tadmor, *How heterophilious dynamics enhance consensus*, SIAM review, Vol. 56, No. 4, pp. 577–621, 2014.

[27] H. Niwa. *Self-organizing dynamic model of fish schooling*, J. Theor. Biol., 171:123–136, 1994.

[28] J. Parrish and L. Edelstein-Keshet. *Complexity, pattern, and evolutionary trade-offs in animal aggregation*, Science, 294:99–101, 1999.

[29] J. Parrish, S. Viscido, and D. Gruenbaum. *Self-organized fish schools: An examination of emergent properties*, Biol. Bull., 202:296–305, 2002.

[30] C. S. Patlak. *Random walk with persistence and external bias*, Bull. Math. Biophys., 15:311–338, 1953.

[31] L. Perea, G. Gťomez, and P. Elosegui. *Extension of the Cucker-Smale control law to space flight formations*, AIAA Journal of Guidance, Control, and Dynamics, 32:527–537, 2009.

[32] B. Perthame. *Transport Equations in Biology*, Basel: Birkhauser, 2007.

[33] L.S. Pontryagin and V.G.Boltyanskii, R.V. Gamkrelidze and E.F. Mishenko, *The Mathematical Theory of Optimal Processes*, John Wiley, New York, 1962.

[34] W. Romey. *Individual differences make a difference in the trajectories of simulated schools offish*, Ecol. Model., 92:65–77, 1996.

[35] R. Sepulchre, D. Paley, and N. E. Leonard. *Stabilization of planar collective motion with all-to-all communication*, IEEE Transactions on Automatic Control, 52(5):811–824, May 2007.

[36] K. Sugawara and M. Sano. *Cooperative acceleration of task performance: Foraging behavior of interacting multi-robots system*, Physica D, 100:343–354, 1997.

[37] J. Toner and Y. Tu. *Long-range order in a two-dimensional dynamical xy model: How birds fly together*, Phys. Rev. Lett., 75:4326–4329, 1995.

[38] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet. *Novel type of phase transition in a system of self-driven particles*, Phys. Rev. Lett., 75:1226–1229, 1995.