MathematicS

MathS In A.

In Action

Felipe Cucker & Cristián Huepe

Flocking with informed agents

Volume 1 (2008), p. 1-25.

<http://msia.cedram.org/item?id=MSIA_2008__1_1_1_0>

MathematicS In Action
Vol. 1, 1-25 (2008)

Flocking with informed agents

Felipe Cucker *
Cristián Huepe **

* Department of Mathematics City University of Hong Kong HONG KONG
E-mail address: macucker@cityu.edu.hk
** 614 N Paulina Street, Chicago, IL 60622-6062 U.S.A.
E-mail address: cristian@northwestern.edu.

Abstract

Two similar Laplacian-based models for swarms with informed agents are proposed and analyzed analytically and numerically. In these models, each individual adjusts its velocity to match that of its neighbors and some individuals are given a preferred heading direction towards which they accelerate if there is no local velocity consensus. The convergence to a collective group swarming state with constant velocity is analytically proven for a range of parameters and initial conditions. Using numerical computations, the ability of a small group of informed individuals to accurately guide a swarm of uninformed agents is investigated. The results obtained in one of our two models are analogous to those found for more realistic and complex algorithms for describing biological swarms, namely, that the fraction of informed individuals required to guide the whole group is small, and that it becomes smaller for swarms with more individuals. This observation in our simple system provides insight into the possibly robust dynamics that contribute to biologically effective collective leadership and decision-making processes. In contrast with the more sophisticated models mentioned above, we can describe conditions under which convergence to consensus is ensured.

1. Introduction

Swarming models describe the dynamics of groups of interacting self-propelled agents, and provide a new set of systems to study complex non-equilibrium dynamics. They are developed to mimic the collective motion of groups of living individuals such as such as bird flocks, fish schools, herds of quadrupeds or bacteria colonies [1, 12, 14, 21, 27]. More recently, these models have been the focus of renewed interest as basic coordination and consensus algorithms for systems of mobile autonomous agents, such as groups of robotic vehicles, mobile sensors, or even satellites [2, 11, 13, 17, 18, 20, 22, 24]. In simple agent-based swarming models, point-particles are set to move in space at non-zero speeds while matching the headings of neighboring particles [8, 9, 26]. Different models consider different specific dynamics (e.g. with or without inertia [25], discrete or continuous [10, 19], with imposed or emerging non-zero agent speed [12, 26], etc.), but all of them follow some variation of these same basic rules. As models are made more realistic, the effect of attraction and repulsion forces (to form cohesive groups while avoiding agent collisions) are often included [4]. Noise can also be added either to the agent dynamics or to their communication. In spite of these variations, a similar behavior is observed in most simple swarming models. Namely, agents self-organize for certain parameter values into coherent collective states where they group in space and/or move together by converging, respectively, in the position and/or velocity spaces. The conditions required for such convergence, however, have not been rigorously established even for simple models, with the notable exception of the work in [17]. This is partly due to the fact that model interaction rules have been designed for

The work of Cristián Huepe was supported by the US National Science Foundation under Grant No. DMS-0507745.
Keywords: Particle systems, flocking.
Math. classification: 93C15.

their numerical implementation rather than for their mathematical analysis. It is also a consequence of the nontrivial nature of biological interactions, which often include nonlinearities even in the simplest cases. Very few solid mathematical results exist for swarming dynamics. As a consequence, a number the basic properties of swarming systems remain unknown, such as the conditions required to ensure group cohesion, the mechanisms of group size selection, or the fraction of biased agents needed to guide the rest in a given direction. While many of these questions could depend on the specificities of the model, the robust generic behavior often found in other statistical systems lead us to expect that some model-independent mechanisms may also appear in swarms. In this paper we address this issue by studying a family of simple Laplacian-based swarming models that are well suited for mathematical analysis. We use analytical and numerical tools to characterize the group cohesion and its collective decision making capabilities when few informed individuals are present. Our results mimic qualitatively what is obtained for more complicated and biologically realistic models. They could therefore help unveil the generic mechanisms behind such behaviors.

Laplacian-based models (such as those in [6, 26]) postulate the following behavior: every agent adjusts its velocity by adding to it a weighted average of the differences of its velocity with those of the other agents. That is, at time t, and for agent i,

$$\boldsymbol{v}_i(t+h) = \boldsymbol{v}_i(t) + h\sum_{j=1}^{N} a_{ij}(\boldsymbol{v}_j(t) - \boldsymbol{v}_i(t)) \tag{1.1}$$

where the weights $\{a_{ij}\}$ quantify the way the agents influence each other, and $h > 0$ is the time step. It is commonly assumed that these weights are a non-increasing function of the distance between agents. This is the case in [6, 26] and the major difference between these models is in the choice of the a_{ij}. In this paper we will follow [6] and take the *adjacency matrix* A_x to have entries

$$a_{ij} = \frac{H}{(1 + \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2)^\beta} \tag{1.2}$$

for some fixed $H > 0$ and $\beta \geq 0$. The equations in (1.1) describe a model with discrete time. A corresponding model for continuous time is obtained by letting h tend to zero. We then obtain

$$\boldsymbol{v}_i'(t) = \sum_{j=1}^{N} a_{ij}(\boldsymbol{v}_j(t) - \boldsymbol{v}_i(t)). \tag{1.3}$$

The role played by the agents in the model given by (1.1) (or (1.3)) is the same for all of them. Differentiated roles due to a hierarchy among the agents is studied in [23]. Another role differentiation, which will be the focus of this paper, is that arising by a possible preferred direction for some of the agents. Such a possibility is studied in [3] for a model which, in addition to the Laplacian-based averaging, features attracting and repelling forces between agents. In this paper we will do so for a simple extension of (1.3) which, together with its natural equation for position changes, has the form

$$\begin{aligned} \boldsymbol{x}_i'(t) &= \boldsymbol{v}_i(t) \\ \boldsymbol{v}_i'(t) &= \sum_{j=1}^{N} a_{ij}(\boldsymbol{v}_j(t) - \boldsymbol{v}_i(t)) + C_i(t)\boldsymbol{d}_i(t) \end{aligned} \tag{1.4}$$

where $C_i(t) \geq 0$ is some local measure of the alignment consensus at time t (with absolute consensus for $C_i(t) = 0$) and $\boldsymbol{d}_i(t) \in \mathbb{E}$ is the preferred direction of agent i. Here $\mathbb{E}$ denotes Euclidean space (either 2-dimensional or 3-dimensional). Individuals with a non-trivial preferred direction $\boldsymbol{d}_i(t)$ are called *informed*.

The first two terms in the right-hand side of (1.4) (which are the right-hand side of (1.3)) capture the "social forces" that stir the individuals towards alignment. The last term, in contrast, stirs individuals towards their preferred direction. A balance must be achieved between these

two terms: if informed individuals are stubborn (i.e., their last term is always significant) there is no hope to reach consensus as long as we have two informed individuals with different preferred directions. The role of the terms $C_i(t)$ is to allow for such balance. Informed individuals may start with a determined steering towards their preferred direction but as a consensus is built this determination gives place to an acceptance of the socially preferred direction. They attempt to drive the flock to their own direction but prefer to give up the latter than to get disentangled from the flock.

In this paper we consider two possibilities for $C_i(t)$ (Models A and B described in §2.1 below). In addition we take $\|\boldsymbol{d}_i(t)\| = \nu$ for some $\nu \geq 0$ and all $i \leq k$. Under these assumptions the ratio $\frac{H}{\nu}$ determines a fixed (independent of time and of the actual consensus) balance between the roles of social cohesion and information.

The numerical experiments described in [3] show that, for the model considered in that paper, the existence of a small group of informed agents is enough, under certain conditions, to guide the whole group towards a preferred direction.

The goal of this paper is twofold. Firstly, we extend the model in [6] to include informed agents as described above. This extension allows us to generalize the proofs of convergence in [6] to flocks with informed agents. This is in contrast with the prevalent character of papers in flocking whose conclusions rely only on simulations, and puts our paper within the stream of mathematical papers [15, 16, 23] dealing with various aspects and extensions of the model proposed in [6]. Secondly, we perform some simulations as well. These simulations are useful for measuring how sharp the worst-case bounds in our formal result are. They are also useful, as usual, to observe different features of the flock's behavior.

The rest of the paper is structured as follows. In Section 2 we provide the mathematical preliminaries necessary to describe both our consensus measures $C_i(t)$ and our main formal result. The latter is formally stated and proved in Section 4. Before doing so, however, in Section 3, we describe the results of the numerical simulations (which are performed, needless to say, with discrete time for a small enough h).

2. A Convergence Result

2.1. Measuring consensus

The inner product on $\mathbb{E}$ naturally induces an inner product on $\mathbb{E}^N$. Let Δ be the diagonal of $\mathbb{E}^N$, i.e.,

$$\Delta = \{(u, u, \ldots, u) \mid u \in \mathbb{E}\}$$

and $\Delta^\perp$ be the orthogonal complement of Δ in $\mathbb{E}^N$. Then, every point $\boldsymbol{v} \in \mathbb{E}^N$ decomposes in a unique way as $\boldsymbol{v} = \boldsymbol{v}_\Delta + \boldsymbol{v}_\perp$ with $\boldsymbol{v}_\Delta \in \Delta$ and $\boldsymbol{v}_\perp \in \Delta^\perp$. This decomposition has a simple explicit form. Denote by

$$\boldsymbol{m} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{v}_i$$

the mean of the $\boldsymbol{v}_i$. Then $\boldsymbol{v}_\Delta = (\boldsymbol{m}, \ldots, \boldsymbol{m})$ and $\boldsymbol{v}_\perp = (\boldsymbol{v}_1 - \boldsymbol{m}, \ldots, \boldsymbol{v}_N - \boldsymbol{m})$. This follows immediately from the equality

$$\begin{aligned}\langle \boldsymbol{v}_\Delta, \boldsymbol{v}_\perp \rangle &= \sum_{i=1}^{N} \langle \boldsymbol{m}, (\boldsymbol{v}_i - \boldsymbol{m}) \rangle = \left\langle \boldsymbol{m}, \sum_{i=1}^{N} \boldsymbol{v}_i - \boldsymbol{m} \right\rangle \\ &= \left\langle \boldsymbol{m}, \left(\sum_{i=1}^{N} \boldsymbol{v}_i \right) - N\boldsymbol{m} \right\rangle = \langle \boldsymbol{m}, 0 \rangle = 0.\end{aligned}$$

We can look at the evolution of the velocities $\boldsymbol{v}_i(t)$ decomposing into the evolution of their mean $\boldsymbol{m}(t)$ and that of the distances to that mean $\boldsymbol{v}_\perp = (\boldsymbol{v}_1 - \boldsymbol{m}, \ldots, \boldsymbol{v}_N - \boldsymbol{m})$ and a key observation at

this stage is the fact that convergence to a common velocity is a feature of the second evolution only. More precisely, the condition "the velocities $\boldsymbol{v}_i(t)$ tend to alignment is equivalent to the condition "$\boldsymbol{v}_\perp(t) \to 0$".

It is natural now to take the norm $\|\boldsymbol{x}_\perp\|$ of the projection $\boldsymbol{x}_\perp$ as the *dissimilarity* of $\boldsymbol{x}$ and similarly for $\|\boldsymbol{v}_\perp\|$. In the case of $\boldsymbol{x}$ we may call this measure the *dispersion* of the flock. It relates with its "diameter" as shown in the next lemma.

Lemma 2.1. *For all* $\boldsymbol{x} \in \mathbb{E}^N$, $\max_{i\neq j} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 \leq 2\|\boldsymbol{x}_\perp\|^2$.

Proof. Write $\boldsymbol{x} = \boldsymbol{x}_\Delta + \boldsymbol{x}_\perp = (\widetilde{u}, \dots, \widetilde{u}) + ((\boldsymbol{x}_\perp)_1, \dots, (\boldsymbol{x}_\perp)_N)$. Then, for all $i \neq j$, $\boldsymbol{x}_i - \boldsymbol{x}_j = (\boldsymbol{x}_\perp)_i - (\boldsymbol{x}_\perp)_j$ and

$$\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_{\mathbb{E}} = \|(\boldsymbol{x}_\perp)_i - (\boldsymbol{x}_\perp)_j\|_{\mathbb{E}} \leq \|(\boldsymbol{x}_\perp)_i\|_{\mathbb{E}} + \|(\boldsymbol{x}_\perp)_j\|_{\mathbb{E}} \leq \sqrt{2}\|\boldsymbol{x}_\perp\|_{\mathbb{E}^N}. \qquad \square$$

In the case of velocities, the dissimilarity $\|\boldsymbol{v}_\perp\|$ relates to the consensus in a straight-forward manner: the smaller is $\|\boldsymbol{v}_\perp\|$ the larger this consensus (with complete consensus when, and only when, $\|\boldsymbol{v}_\perp\| = 0$). Note that we are using an absolute measure of consensus (as oposed to a measure that would take into account the magnitude $\|\boldsymbol{v}_\Delta\|$ as well). This is enough for our purposes. We will actually use $\boldsymbol{v}_\perp$ to describe the consensus measure $C_i(t)$. We propose two ways of doing so.

Model A The simplest measure of consensus is $\|v(t)\|$ itself. It is this quantity what is actually shown to approach zero in proofs of convergence to consensus (cf. [6]). Therefore, in our first model we take $C_i(t) = \|\boldsymbol{v}_\perp(t)\|$ for $i = 1, \dots, N$.

Model B This is a more local (and slightly easier to compute) version of Model A. It is simply the mean of the velocities of other agents relative to agent i velocity. That is,

$$C_i(t) = \frac{1}{N-1} \sum_{j\neq i} \|\boldsymbol{v}_j(t) - \boldsymbol{v}_i(t)\|.$$

2.2. **Statement of the Main Result**

In Section 4 we provide detailed proofs of convergence, under certain conditions, towards states where all agents move with the same velocity and in one group. We consider a general setting that includes various cases (Model A, Model B, and different configurations of informed agents). We next describe this convergence result.

Consider the system described by (1.4) with an adjacency matrix $A = (a_{ij})$ as given in (1.2), we can define a *Main Convergence Condition* as:

$$U_0^\beta D \leq \frac{1}{3} NH. \tag{MCC}$$

Here, U_0 is a constant that depends mainly on initial conditions of the swarm given by $\|\boldsymbol{x}_\perp(0)\|^2$ and $\|\boldsymbol{v}_\perp(0)\|^2$. It can also depend on the system parameters β, H, and N. Its specific definition changes for different values of β and will be detailed below in each case. The parameter D only depends on the characteristics of the informed individuals. In cases with q arbitrarily informed individuals with the same bias strength, which corresponds to having q non-zero vectors $\boldsymbol{d}_i$ in (1.4) with any direction but the same magnitude ν, we find that D is given by:

$$D = \begin{cases} \nu\sqrt{q} & \text{for Model A} \\ \nu\sqrt{2q} & \text{for Model B.} \end{cases}$$

We can sharpen inequality (MCC) in some special cases. As shown in Proposition 4.6, in case of Model A and if there are only two groups of informed individuals, one with q_1 and the other with q_2 members, where both groups have the same bias strength ν and all agents in each group

have the same preferred direction, then (MCC) will be satisfied for:

$$D = \nu\sqrt{(q_1+q_2) - \frac{1}{N}(q_1^2+q_2^2+2q_1q_2\omega)}.$$

Here ω denotes the cosine of the angle between the two group's directions. It follows that if $\boldsymbol{d}_i$ in (1.4) contains q non-zero entries and all are identical vectors with magnitude ν, a stronger MCC inequality is obtained where:

$$D = \nu\sqrt{q\left(1-\frac{q}{N}\right)}.$$

Using these quantities together with the following ones (depending on the system parameters and the initial state of the flock only),

$$\boldsymbol{a} = \frac{16}{(NH)^2}\|\boldsymbol{v}_\perp(0)\|^2$$

and

$$\boldsymbol{b} = 1 + 2\|\boldsymbol{x}_\perp(0)\|^2$$

we can state our main result.

Main Theorem. *Assume that one of the following three conditions hold:*

Case i:: $\beta < 1/2$ *and* (MCC) *is satisfied with*

$$U_0 = \max\left\{(2\boldsymbol{a})^{\frac{1}{1-2\beta}}, 2\boldsymbol{b}\right\}$$

Case ii:: $\beta = 1/2$ *and* (MCC) *is satisfied with*

$$U_0 = \frac{\boldsymbol{b}}{1-\boldsymbol{a}}$$

and $\boldsymbol{a} < 1$,

Case iii:: $\beta > 1/2$ *and* (MCC) *is satisfied with*

$$U_0 = \frac{2\beta}{2\beta-1}\boldsymbol{b}$$

and

$$\left(\frac{1}{2\beta\boldsymbol{a}}\right)^{\frac{1}{2\beta-1}} > \frac{2\beta}{2\beta-1}\boldsymbol{b}.$$

Then, $\boldsymbol{v}_\perp(t)$ *converges exponentially fast to zero,*

$$\|\boldsymbol{v}_\perp(t)\| \le \|\boldsymbol{v}_\perp(0)\|\exp\left(-\frac{2NH}{3U_0^\beta}t\right),$$

and the maximum size reached by the group is bounded by the relation

$$\|\boldsymbol{x}_\perp(t)\|^2 \le \frac{U_0-1}{2},$$

for all $t \ge 0$.

3. Numerical Simulations

In this section we will study numerically the dynamics of the model. The equations in (1.4) were directly implemented as the time-step algorithm for numerical integration, with all new positions and velocities computed and updated synchronously. A typical run consisted of from a few hundred to several thousand time-steps between the initial condition and the end of the run. The runs were terminated when the system reached one of the following two conditions. Either $\|\boldsymbol{v}_\perp(t)\| < 0.001$, which almost certainly implies that the system converged to its final swarming configuration, or $\|\boldsymbol{v}_\perp(t)\| > 1000$, which is a near sure indication that the agents are diverging and will never reach consensus to swarm together. As detailed below, several realizations were carried out for each set of parameters.We are interested in the case with small h, since it is closer to the continuous-time dynamics studied analytically in Sections 2 and 4. In all the computations below, we used a time-step of $h = 0.001$ which was validated by verifying that a smaller time-step $h = 0.0001$ produced no significant difference. To reduce the parameter space, we concentrate in the most physically interesting case of $\beta = 1$ and $H = 1$.

3.1. Analysis of Convergence Theorems

We first study numerically the conditions under which the system converges to a flocking dynamics and analyze the resulting states. We compare here this numerical convergence to the theoretical bounds derived above. We begin by noting that our theoretical convergence results only depend on the initial variance of the agents' positions and velocities, but not on their specific values. Because of this, we will consider two different distributions for the initial conditions, which we detail as follows. The first initial condition (IC1) is chosen as a random uniform distribution of positions and velocities within a disc of radius $R^1_{\boldsymbol{x}(0)}$ and $R^1_{\boldsymbol{v}(0)}$ in the x_1–x_2 and v_1–v_2 planes, respectively. The second initial condition (IC2) is chosen to maximize the chances of a group to break for a given $\boldsymbol{x}_\perp(0)$ and $\boldsymbol{v}_\perp(0)$, and corresponds to having all agents placed on a circle of radius $R^2_{\boldsymbol{x}(0)}$, heading outwards with speed $R^2_{\boldsymbol{v}(0)}$. Figures 1 and 2 show convergence results for the simulations of Models A and B (respectively), using $N = 10$ agents of which one is informed ($q = 1$), and with $\nu = 1$.

Simulations were carried out using initial conditions given by IC1 (left panels) or IC2 (right panels), with $R^1_{x(0)}$, $R^1_{v(0)}$, $R^2_{x(0)}$, and $R^2_{v(0)}$ spanning from 0.05 to 2.5. Each point indicates the initial $\|\boldsymbol{x}_\perp(0)\|$ and $\|\boldsymbol{v}_\perp(0)\|$ values at which the numerical simulation did not converge to a single swarm moving as a group, but instead split into two or more groups heading in different directions with different speeds. Two hundred and fifty thousand runs spanning values of $\|\boldsymbol{x}_\perp(0)\|$ and $\|\boldsymbol{v}_\perp(0)\|$ ranging from approximately 0.05 to 7 where tested for convergence, so regions with no points imply that all simulations converged there.

The dashed line displays the $\|\boldsymbol{x}_\perp(0)\|$ bound for convergence given by the MCC and the solid curve shows the bound imposed by the additional convergence condition in Case (iii). Both conditions must be satisfied for the convergence to be assured, so the region that contains the origin where $\|\boldsymbol{x}_\perp(0)\|$ and $\|\boldsymbol{v}_\perp(0)\|$ are smaller than these curves is where we have proven that the system converges.

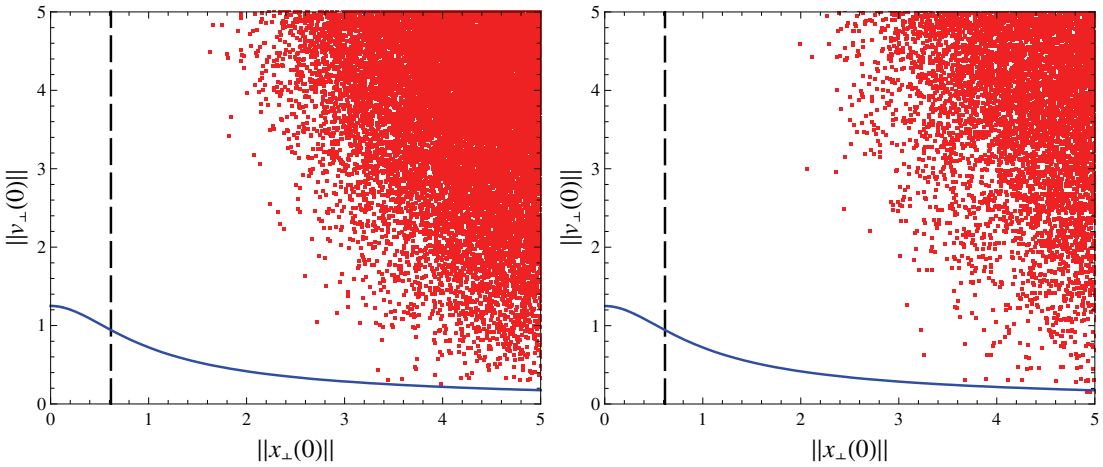

FIGURE 1. Values of the initial position and velocity dispersions $\|\boldsymbol{x}_\perp(0)\|$ and $\|\boldsymbol{v}_\perp(0)\|$ (red points) for which a group of $N = 10$ agents with one leader ($q = 1$ and $\nu = 1$) do not converge under the dynamics of Model A to a collective swarm where all have the same velocity. The dashed line and the solid curve correspond to the analytical bounds specified by the Main Convergence Condition and by the additional condition in Case (iii) (see text). The Main Theorem that we prove in the text guarantees that no dynamics will diverge in the region next to the origin and bounded by both curves, as confirmed here numerically. The left panel corresponds to an initial condition (IC1) where agents' positions and velocities are randomly distributed within a region. In the right panel, agents' initial positions are randomly distributed on a circle pointing outwards (IC2).

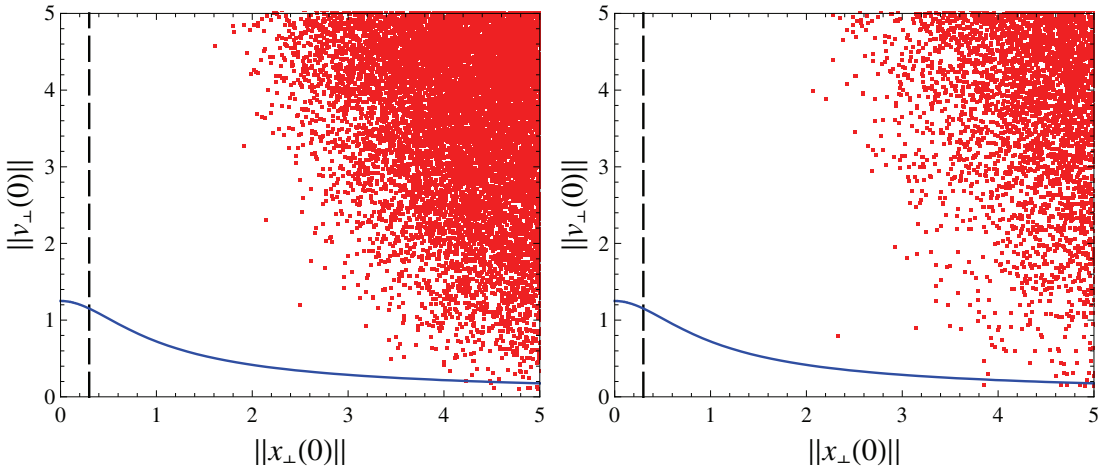

FIGURE 2. Values of the initial position and velocity dispersions $\|\boldsymbol{x}_\perp(0)\|$ and $\|\boldsymbol{v}_\perp(0)\|$ (red points) for which a group of $N = 10$ agents with one leader ($q = 1$ and $\nu = 1$) do not converge under the dynamics of Model B to a collective swarm. As on Figure 1, the solid and dashed lines bound a region next to the origin, within which no dynamics can diverge. The left panel corresponds to an IC1 initial condition and the right one to an IC2 (see text and Figure 1).

The figures confirm that for a set of parameters within Case (iii) of the Main Theorem, convergence is indeed observed numerically.

3.2. **Swarming Dynamics with Informed Individuals**

We next focus not only on whether the system will converge, or whether instead the flock will break, but also on how effective it is in heading as a group towards the preferred direction of the informed individuals. In [3] it was pointed out that a biologically relevant feature of a swarming algorithm is to help a small fraction of informed individuals guide the whole group reliably in their preferred direction. Using a detailed model that included various specific biological features, they showed that a great degree of accuracy could be achieved with a small fraction of informed individuals and that the fraction required to reach a certain accuracy became smaller for larger groups.

The main conclusion resulting from our analysis is that, under certain conditions, both Model A and Model B display an ability to guide the group accurately in the preferred direction of a small fraction of informed individuals similar to the one observed in [3], but with a much simpler and mathematically manageable algorithm.

In the following figures we display the Accuracy that the group achieves in heading towards the preferred direction of the informed individuals, which we impose to be the same for all of them and denote by the unit vector $\widehat{I}$. This quantity is defined as the variance of the final direction angle with respect to the angle of $\widehat{I}$, and is given by:

$$\mathsf{Accuracy} = \frac{1}{N_c}\sum_{j=1}^{N_c}\left[\mathrm{Angle}(v_j^{\mathrm{end}},\widehat{I})\right]^2,$$

where N_c corresponds to the total number of realizations in which $\boldsymbol{v}_\perp$ converges to zero and therefore the group does not split. This implies that the Accuracy cannot be computed in cases where the groups split in all realizations. The figures also display the fraction of realizations where $\boldsymbol{v}_\perp$ eventually diverges and the group splits. The results are presented as a function of the fraction of informed agents q/N for groups of various sizes ranging from $N = 10$ to $N = 200$ agents. Each point represents the mean Accuracy obtained after averaging over 400 realizations.

In Figures 3 to 9, the positions of all agents are picked at random with equal probability from anywhere within a disc of radius $\sqrt{N/(2\pi\rho)}$ (which is equal to $R^1_{\boldsymbol{x}(0)}$ and $R^1_{\boldsymbol{v}(0)}$ radii in the IC1 initial condition detailed in §3.1) and centered at the origin on the x_1–x_2 and v_1–v_2 planes, respectively. This makes the initial density equal to $\rho = 1$ regardless of the number of agents N in the group.

Model A is generally much better in achieving high accuracy for a small fraction of informed individuals in larger groups. Two different mechanisms contribute to this effect in our simulations. The first corresponds to simply orienting all informed individuals in their preferred direction as initial conditions. We implemented this mechanism in some of the simulations, which we will refer to as having Oriented Initial Conditions (OIC). The second is achieved by increasing the acceleration ν of informed agents in their preferred direction. The accuracy plots for Model A with $\nu = 0$ and OIC (Figure 3) and those with $\nu > 0$ and non-OIC (e.g. Figure 4) show that either mechanism can achieve independently high accuracy for small fractions of informed agents in large groups.

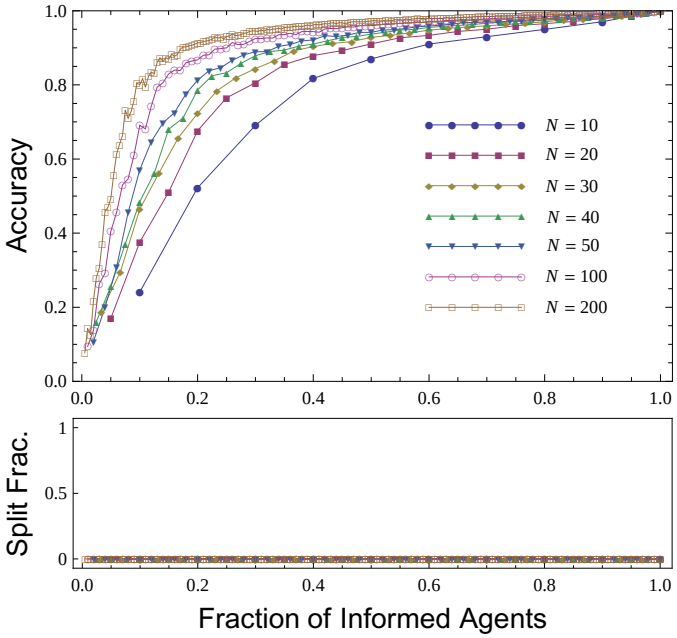

FIGURE 3. Model A, Oriented, $\nu = 0$

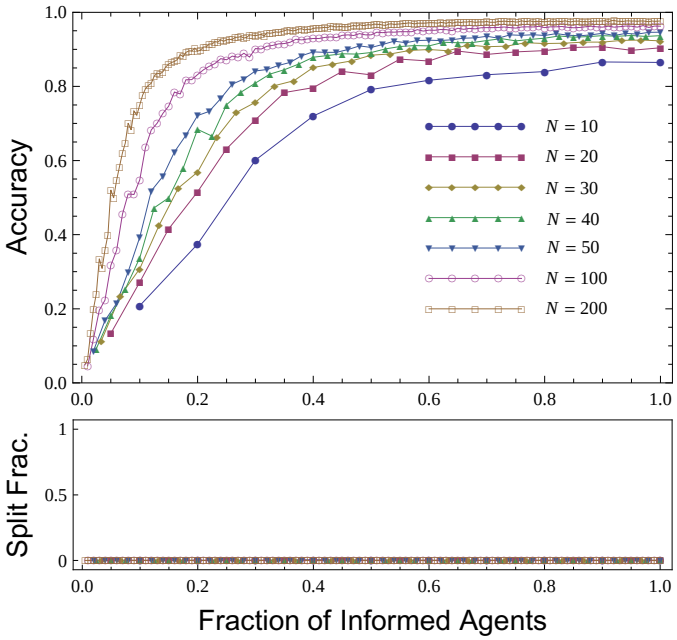

FIGURE 4. Model A, Non Oriented, $\nu = 2$

As ν is increased, the OIC and non-OIC cases become more similar (cf. Figures 5 and 6) since the initial non-OIC transient dynamics will be faster and informed individuals will rapidly assume their preferred oriented direction as in the OIC case.

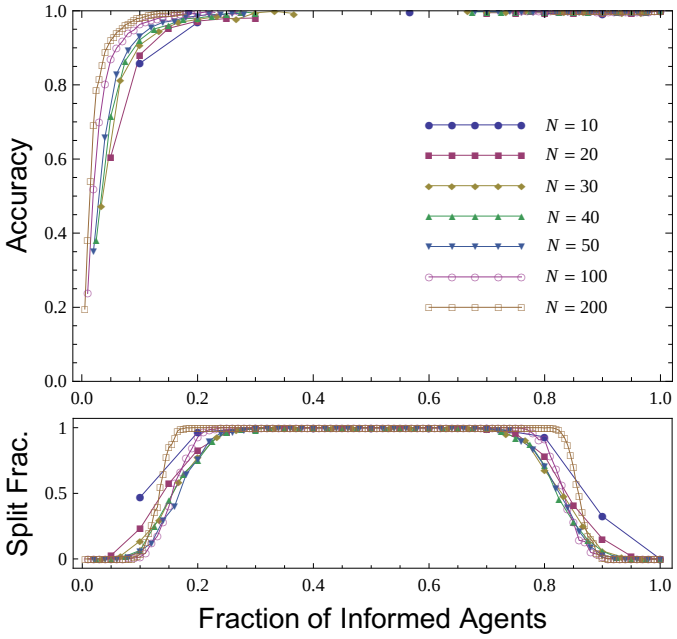

Figure 5. Model A, Oriented, $\nu = 5$

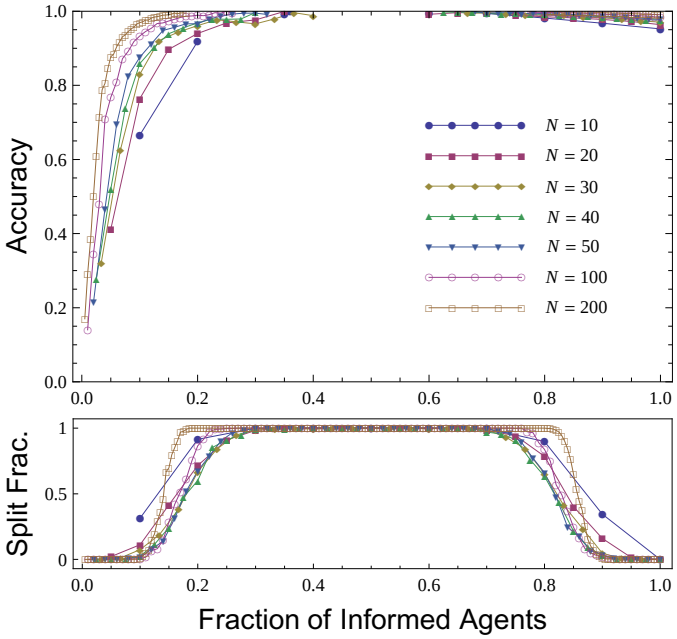

Figure 6. Model A, Non Oriented, $\nu = 5$

Our results also show that as ν increases, there is a greater chance that the group will split, thus failing to achieve a consensus in their final velocity. In both the OIC and non-OIC cases, if ν is large enough the groups will split for intermediate fractions of informed agents. This is because the group cohesion is harder to maintain if it contains a similar fraction of informed and non-informed individuals. In this case the lack of a strong majority will produce a slower

convergence to consensus, thus increasing the chances that the group splits before all of its agents are "convinced" to move in the same way. *Groups must therefore achieve a compromise between higher values of ν that will increase accuracy (for a given fraction of informed individuals) and lower ones that will avoid group splitting.*

Model B behaves quite differently than Model A. The only conditions under which Model B also displays higher accuracy for lower fractions of informed individuals as the group size N grows, is for the OIC case with $\nu = 0$ or small. This is because both models will become identical for $\nu = 0$, since the term containing $C_i(t)$ in equation (3) will then be equal to zero. For larger values of ν, however, the behavior is quite different in Model B, where the accuracy values vs. fraction of informed agents seem to collapse to a unique curve for different group sizes as, for example, on Figure 7.

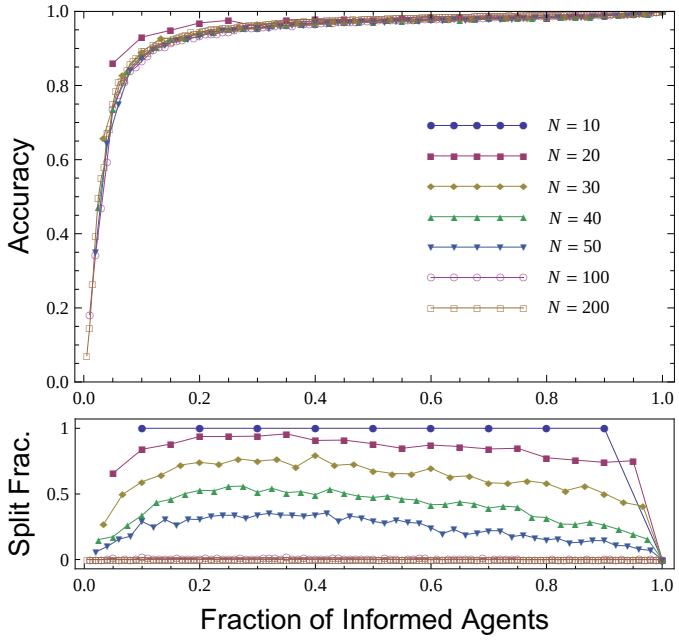

FIGURE 7. Model B, Oriented, $\nu = 10$

As ν is increased in Model B, the fraction of groups that split also increases but in a different way than in Model A. Here, the probability of splitting grows almost homogeneously for all fractions of informed agents, in contrast to Model A where situations with similar fractions of informed and non-informed individuals have a much higher chance of splitting. When non-OIC are used in Model B, however, the situation is different. In this case, the system displays for small ν values (but not for $\nu = 0$, which is a trivial case where no preferred direction is imposed and therefore the accuracy is zero) accuracy curves that do not depend on the group sizes as in Figure 8.

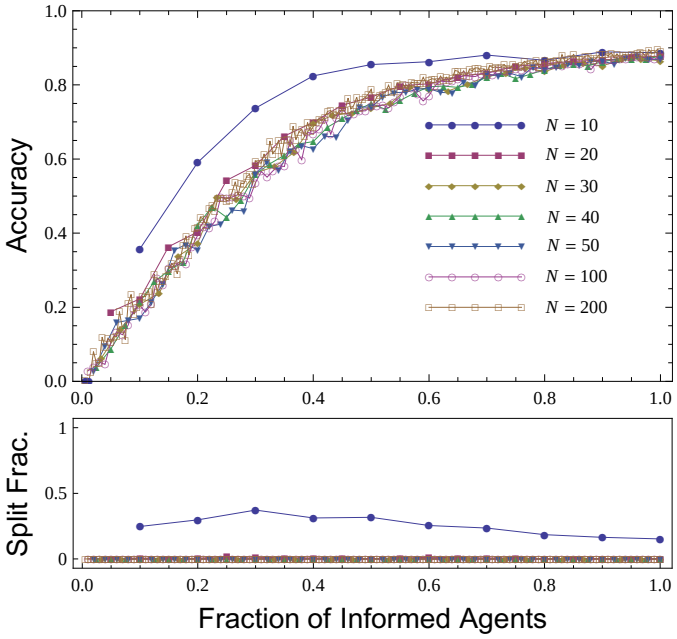

Figure 8. Model B, Non Oriented, $\nu = 5$

As ν is increased, the accuracy decreases for larger group sizes as shown on Figure 9, which is opposite to the typical behavior observed in Model A.

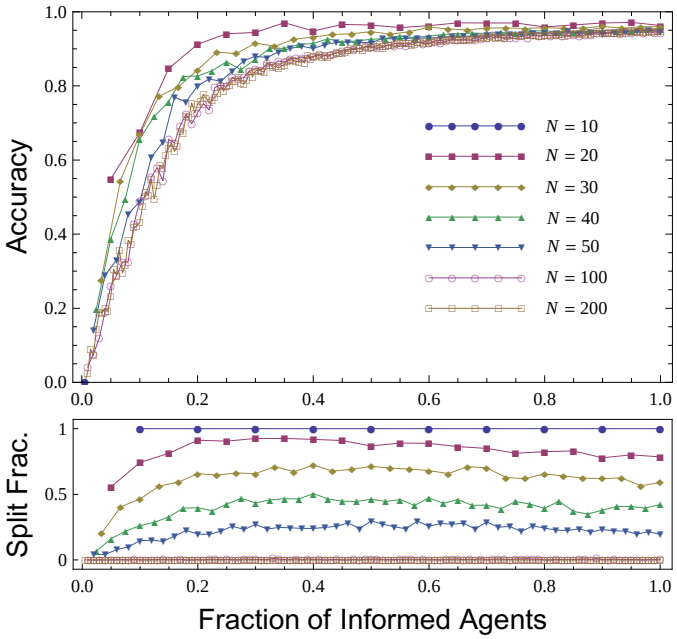

Figure 9. Model B, Non Oriented, $\nu = 10$

Note that the higher average accuracy measured for smaller groups in this case may be an effect of having more small groups that split, leaving only the ones with better convergence (and perhaps better accuracy) to be considered in this average calculations.

The similarities and differences between Models A and B may help us understand the basic underlying mechanism that produces an accuracy that increases for larger groups at fixed fraction of informed agents, as first observed in [3]. The behavior of both models can be partially explained as follows. Two factors contribute to the group final direction of motion: The initial orientation of the informed agents and their acceleration ν. The former is a robust way to achieve greater accuracy in larger systems with the same fraction of informed agents that does not depend on the term containing $C_i(t)$ in equation (3). In contrast, the latter depends on the specific model interactions, generating higher accuracies for larger groups in Model A but lower or equal accuracies in Model B. This scenario is supported by computations carried out at different initial density values, which also achieve higher accuracies for both models at low ν values and in a way that does not seem to depend strongly on this initial density (see, e.g., Figures 10 or 11 carried out for initial constant volume and therefore with increasing initial density as N grows).

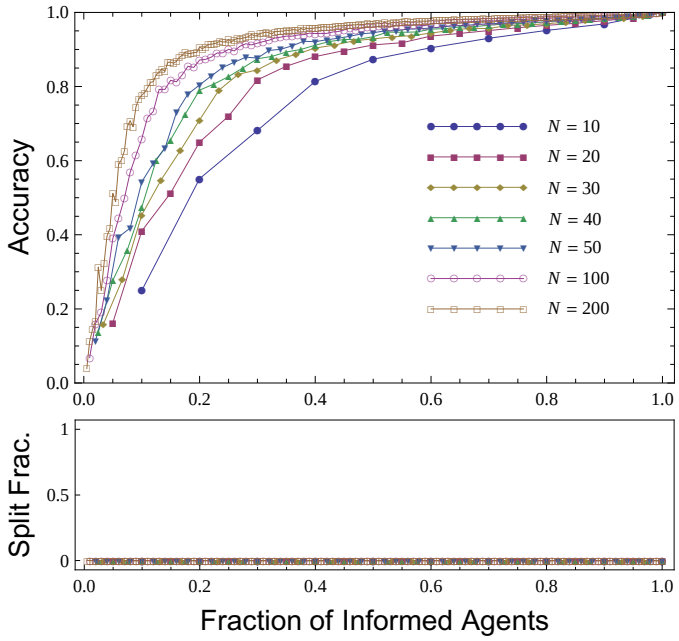

Figure 10. Model A, Constant Initial Volume, Oriented, $\nu = 0$

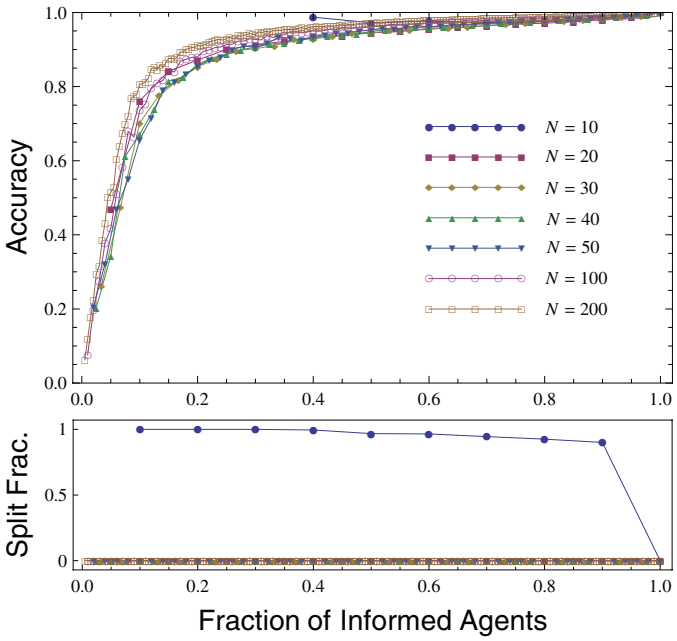

Figure 11. Model B, Constant Initial Volume, Oriented, $\nu = 10$

The accuracy curves behave very differently when the initial density is changed for non-OIC. In this case, the acceleration mechanism is seen to depend strongly on this density and on the specific model (as shown on the results for non-OIC and constant initial volume; Figures 12 and 13).

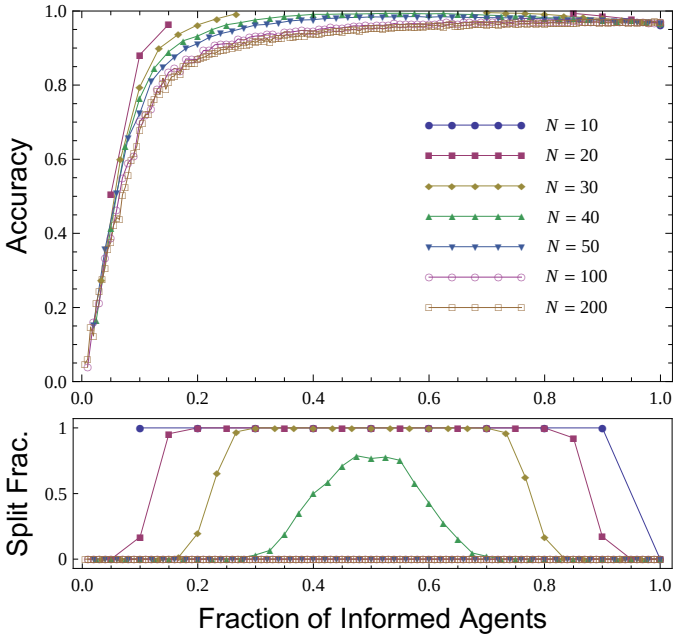

Figure 12. Model A, Constant Initial Volume, Non Oriented, $\nu = 10$

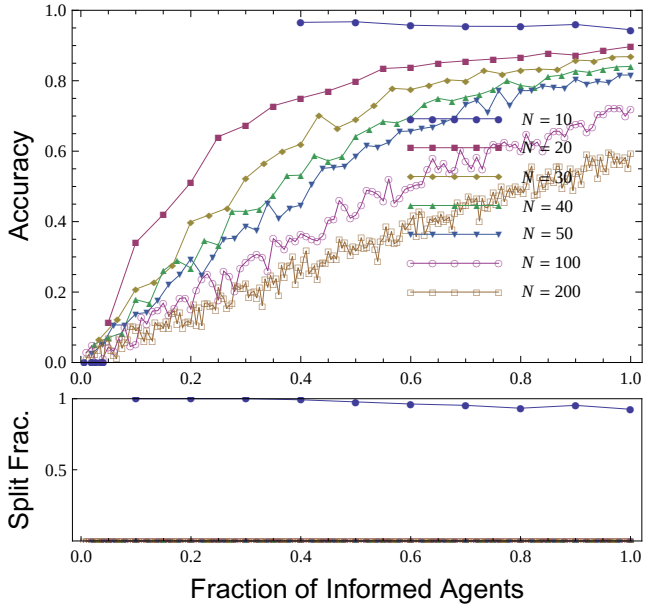

FIGURE 13. Model B, Constant Initial Volume, Non Oriented, $\nu = 10$

In spite of the similarities between the accuracy results obtained with our models and the ones in [3], other aspects of the dynamics are very different. If we defined the group elongation as the ratio of the span of the group in the direction of motion to that perpendicular to the direction of motion, the results in [3] displayed significant group elongations that increased with the group size. In contrast, in all computations carried with our models the elongation value achieved when the velocities converged was approximately equal to one, which implies no significant deformation. Furthermore, for larger groups the final elongation becomes even closer to one. The only consistent increase in the group elongation is observed in Model A as a small increase above one in regions close to the transition to group splitting.

Another significant difference in the dynamics of our models is found in the final positions of the informed agents within the group. While it was stated in [3] that the informed agents tend to move towards the front of the group in their simulations, in all the regimes tested with our algorithms the informed individuals tend to be homogeneously distributed about the center of the group with respect to the direction of motion. This difference may be due to the lack of noise in our system or because the long-range nature of the aligning interactions of our model are too stringent to allow for a significant redistribution of the agents positions in the transient dynamics towards convergence.

3.3. **Accuracy and splitting fraction against information strength ν**

There is an interesting contrast between Figures 5 and 6 (or Figures 7 and 9 for Model B) and similar runs with $\nu = 1$. In the first two there is not much difference between Oriented and non-Oriented initial conditions. But it is clear that as soon as ν is small enough this can make a big difference. This motivates looking at the way the accuracy and splitting fraction vary with respect to the information strength ν.

We do so in the next four plots (Figures 14, 15, 16, and 17). We take $N = 50$ in all runs. As in §3.2, each point is obtained by averaging over 400 realizations.

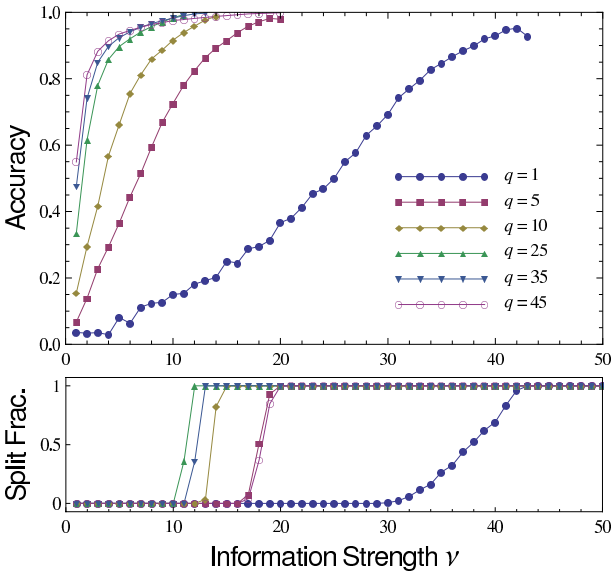

Figure 14. Model A, Non Oriented

We observe that when ν increases so does the accuracy. At a certain value of ν, however, the splitting fraction surges and quickly reaches 1. We see again the trade-off between accuracy and convergence observed in §3.2. For every number of informed agents q we reach a certain level of accuracy by increasing ν but at some moment we can not further increase the accuracy without (rapidly) increasing the frequency of dispersion. Also, curiously, the role of q is monotonous for the accuracy (the larger q the larger the accuracy) but it is not so for the splitting fraction.

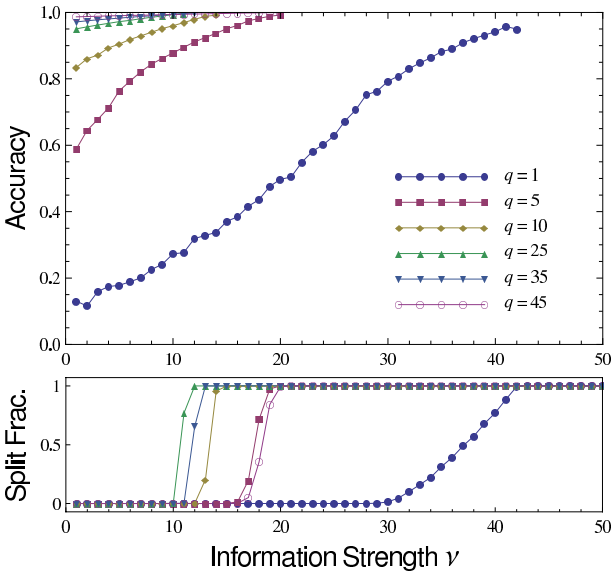

Figure 15. Model A, Oriented

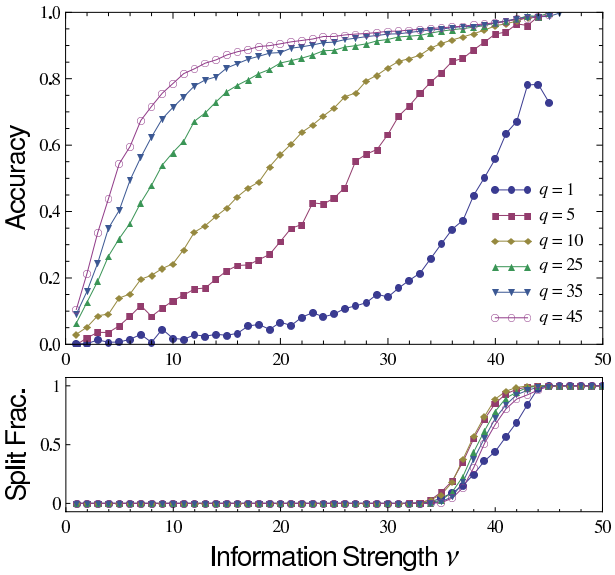

FIGURE 16. Model B, Non Oriented

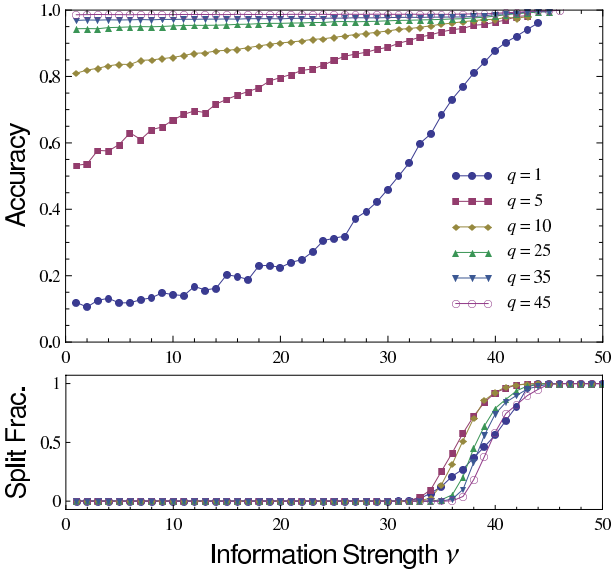

FIGURE 17. Model B, Oriented

4. **Proof of the Main Results**

We can write the set of equalities (1.3) in a more concise form. Let D_x be the $N \times N$ diagonal matrix whose ith diagonal entry is $\sum_{j \leq N} a_{ij}$ and $L_x = D_x - A_x$. Then (cf. [6])

$$\boldsymbol{v}(t+h) - \boldsymbol{v}(t) = -hL_x\boldsymbol{v}(t).$$

Note that the matrix notation $L_x\boldsymbol{v}(t)$ does not have the usual meaning of a $N \times N$ matrix acting on $\mathbb{R}^N$. Instead, the matrix L_x is acting on $\mathbb{E}^N$ by mapping $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N)$ to $((L_x)_{i1}\boldsymbol{v}_1 + \ldots + (L_x)_{iN}\boldsymbol{v}_N)_{i \leq N}$.

We may therefore rewrite (1.4) as

$$\begin{aligned} \boldsymbol{x}' &= \boldsymbol{v} \\ \boldsymbol{v}' &= -L_x\boldsymbol{v} + C\boldsymbol{d}(t). \end{aligned} \tag{C}$$

4.1. **Laplacians**

Given a nonnegative, symmetric, $N \times N$ matrix A the *Laplacian* L of A is defined to be

$$L = D - A$$

where $D = \mathsf{diag}(d_1, \ldots, d_N)$ and $d_\ell = \sum_{j=1}^N a_{\ell j}$. Some features of L are immediate. It is symmetric and it does not depend on the diagonal entries of A.

The matrix L_x in (C) is thus the Laplacian of A_x. It satisfies that for all $u \in \mathbb{E}$, $L_x(u, \ldots, u) = 0$. In addition, it is positive semidefinite.

The smallest eigenvalue of L_x is zero. Its second eigenvalue is called the *Fiedler number* of A_x. We will denote it by ϕ_x.

Proposition 4.1. **(**[7, Proposition 1]**)** *Let A be a $N \times N$ nonnegative, symmetric matrix, $L = D - A$ its Laplacian, ϕ its Fiedler number, and $\mu = \min_{i \neq j} a_{ij}$. Then $\phi \geq N\mu$.* □

4.2. **Projecting the solutions**

Consider the projection $(\boldsymbol{x}_\perp(t), \boldsymbol{v}_\perp(t))$ over $\Delta^\perp \times \Delta^\perp$ of the solutions $(\boldsymbol{x}(t), \boldsymbol{v}(t))$ of the system (C). It is easy to show (see [6]) that these projections are the solutions of the restriction of (C) to $\Delta^\perp \times \Delta^\perp$.

More precisely, they are the solutions of

$$\begin{aligned} \boldsymbol{x}(t+h)_\perp &= \boldsymbol{x}(t)_\perp + h\boldsymbol{v}(t)_\perp \\ \boldsymbol{v}(t+h)_\perp &= (\mathrm{Id} - hL_{\boldsymbol{x}_\perp})\,\boldsymbol{v}(t)_\perp + hC\boldsymbol{d}(t). \end{aligned}$$

Hence, in what follows, we will consider positions in

$$X := \mathbb{E}^N/\Delta \simeq \Delta^\perp$$

and velocities in

$$V := \mathbb{E}^N/\Delta \simeq \Delta^\perp.$$

For $\boldsymbol{x}, \boldsymbol{v} \in \mathbb{E}^N$ we will denote $x = \boldsymbol{x}_\perp$ and $v = \boldsymbol{v}_\perp$. Finally, we will denote $Cd = (C\boldsymbol{d})_\perp$. With these notations, our focus is now on the system

$$\begin{aligned} x' &= v \\ v' &= -L_x v + Cd(t). \end{aligned} \tag{C}$$

Note that we still refer to this system as (C).

4.3. **The proofs**

In what follows we fix a solution (x, v) of (C). At time t, $x(t)$ and $v(t)$ are elements in X and V, respectively. In particular, $x(t)$ determines an adjacency matrix $A_{x(t)}$. For notational simplicity we will denote this matrix by A_t and its Laplacian and Fiedler number by L_t and ϕ_t, respectively.

Proposition 4.2. *For $t \geq 0$ we have $\|Cd(t)\| \leq D\|v(t)\|$.*

Proof. In case of Model A we have

$$\|Cd(t)\| \leq \|C\boldsymbol{d}(t)\| = \|v(t)\|\|\boldsymbol{d}(t)\| = \|v(t)\|\sqrt{\sum_{j \in \mathscr{I}} \|\boldsymbol{d}_j(t)\|^2} = \|v(t)\|\nu\sqrt{q}.$$

For Model B,

$$\begin{aligned}
\|Cd(t)\| &\leq \|C\boldsymbol{d}(t)\| = \sqrt{\sum_{j\in\mathscr{I}} \|C_j(t)\boldsymbol{d}_j(t)\|^2} = \sqrt{\sum_{j\in\mathscr{I}} C_j(t)^2\|\boldsymbol{d}_j(t)\|^2} \\
&= \nu\sqrt{\sum_{j\in\mathscr{I}} C_j(t)^2} = \nu\sqrt{\sum_{j\in\mathscr{I}} \left(\frac{1}{N-1}\sum_{j\neq i}\|\boldsymbol{v}_j(t)-\boldsymbol{v}_i(t)\|\right)^2} \\
&= \frac{\nu}{N-1}\sqrt{\sum_{j\in\mathscr{I}} \left(\sum_{j\neq i}\|\boldsymbol{v}_j(t)-\boldsymbol{v}_i(t)\|\right)^2} \\
&\leq \frac{\nu}{N-1}\sqrt{\sum_{j\in\mathscr{I}} \Big((N-1)\sqrt{2}\|v(t)\|\Big)^2} \qquad \text{(by Lemma 2.1)} \\
&= \nu\sqrt{2q}\|v(t)\|.
\end{aligned}$$

□

For $x \in X$ we denote $\Gamma(x) = \|x\|^2$ and for $v \in V$ we denote $\Lambda(v) = \|v\|^2$. Again, we will write $\Lambda(t)$ and $\Gamma(t)$ for the values of Λ and Γ, respectively, at $(v(t), x(t))$. Finally, we will write Γ_0 for $\Gamma(0)$ and similarly for Λ_0.

Denote $\Phi_t = \min_{\tau\in[0,t]} \phi_\tau$.

Proposition 4.3. *For all $0 \leq t < T$,*

$$\|v(t)\| \leq \|v(0)\|e^{-t(\Phi_t - D)}.$$

Proof. Let $\tau \in [0,t]$. Then

$$\begin{aligned}
\Lambda'(\tau) &= \frac{d}{d\tau}\langle v(\tau), v(\tau)\rangle \\
&= 2\langle v'(\tau), v(\tau)\rangle \\
&= -2\langle L_\tau v(\tau) + Cd(\tau), v(\tau)\rangle \\
&= -2\langle L_\tau v(\tau), v(\tau)\rangle - 2\langle Cd(\tau), v(\tau)\rangle \\
&\leq -2\phi_{x(\tau)}\Lambda(\tau) + 2\|Cd(\tau)\|\|v(\tau)\| \\
&\leq -2\phi_{x(\tau)}\Lambda(\tau) + 2D\|v(\tau)\|^2 \\
&= -2\Lambda(\tau)(\phi_{x(\tau)} - D).
\end{aligned}$$

Here we used that L_τ is symmetric positive semidefinite on V. Using this inequality,

$$\ln(\Lambda(\tau))\Big|_0^t = \int_0^t \frac{\Lambda'(\tau)}{\Lambda(\tau)}d\tau \leq \int_0^t -2(\phi_\tau - D)d\tau \leq -2t(\Phi_t - D)$$

i.e.,

$$\ln(\Lambda(t)) - \ln(\Lambda_0) \leq -2t(\Phi_t - D)$$

from which the statement follows. □

Proposition 4.4. *Assume that $\Phi_t > D$ for all $0 \leq t < T$. Then, for all $0 \leq t < T$,*

$$\|x(t)\|^2 \leq 2\|x(0)\|^2 + 2\frac{\|v(0)\|^2}{(\Phi_t - D)^2}.$$

Proof. For $\tau \leq t$ we have $|\Gamma'(\tau)| = |2\langle v(\tau), x(\tau)\rangle| \leq 2\|v(\tau)\|\|x(\tau)\|$. But $\|x(\tau)\| = \Gamma(\tau)^{1/2}$ and $\|v(\tau)\|^2 = \Lambda(\tau) \leq \Lambda_0 e^{-2\tau(\Phi_\tau - D)}$, by Proposition 4.3. Therefore,

$$\Gamma'(\tau) \leq |\Gamma'(\tau)| \leq 2\left(\Lambda_0 e^{-2\tau(\Phi_\tau - D)}\right)^{1/2}\Gamma(\tau)^{1/2} \tag{4.1}$$

and, using that $\tau \mapsto \Phi_\tau$ is non-increasing and that $\Phi_\tau - D > 0$ for all $\tau \le t$,

$$\begin{aligned}\int_0^t \frac{\Gamma'(\tau)}{\Gamma(\tau)^{1/2}} d\tau &\le 2\int_0^t \left(\Lambda_0 e^{-2\tau(\Phi_\tau - D)}\right)^{1/2} d\tau \\ &\le 2\int_0^t \Lambda_0^{1/2} e^{-\tau(\Phi_t - D)} d\tau \\ &= 2\Lambda_0^{1/2}\left(-\frac{1}{\Phi_t - D}\right) e^{-\tau(\Phi_t - D)}\Big|_0^t \le \frac{2\Lambda_0^{1/2}}{\Phi_t - D}\end{aligned}$$

the last inequality because $\Phi_t > D$. This implies

$$\Gamma(\tau)^{1/2}\Big|_0^t = \frac{1}{2}\int_0^t \frac{\Gamma'(\tau)}{\Gamma(\tau)^{1/2}} d\tau \le \frac{\Lambda_0^{1/2}}{\Phi_t - D}$$

from which it follows that

$$\Gamma(t)^{1/2} \le \Gamma_0^{1/2} + \frac{\Lambda_0^{1/2}}{\Phi_t - D}.$$

The result now follows from the elementary inequality $(p+q)^2 \le 2p^2 + 2q^2$. □

A proof of the following lemma is in [5, Lemma 7].

Lemma 4.5. *Let $c_1, c_2 > 0$ and $s > q > 0$. Then the equation*

$$F(z) = z^s - c_1 z^q - c_2 = 0$$

has a unique positive zero z_*. In addition

$$z_* \le \max\left\{(2c_1)^{\frac{1}{s-q}}, (2c_2)^{\frac{1}{s}}\right\}$$

and $F(z) \le 0$ for $0 \le z \le z^*$.

Proof of Main Theorem. Let

$$\Upsilon = \left\{t \ge 0 \mid (1 + 2\|x(t)\|^2)^\beta \le \frac{NH}{2D}\right\}.$$

Note that in all three cases ((i), (ii), and (iii)), we have $\boldsymbol{b} \le U_0$. This implies that $0 \in \Upsilon$ since

$$(1 + 2\|x(0)\|^2)^\beta = \boldsymbol{b}^\beta \le U_0^\beta \le \frac{NH}{3D} < \frac{NH}{2D}.$$

Hence, $\Upsilon \ne \emptyset$. Assume that $\Upsilon \ne [0, +\infty)$ and let $\widehat{t} = \inf\{[0, +\infty) \setminus \Upsilon\}$. Clearly, $(1+2\|x(\widehat{t})\|^2)^\beta = \frac{NH}{2D}$.

By Proposition 4.1 and Lemma 2.1, for all $x \in X$,

$$\phi_x \ge \frac{NH}{(1 + \max_{i\ne j}\|x_i - x_j\|^2)^\beta} \ge \frac{NH}{(1 + 2\|x\|^2)^\beta}.$$

Let $t < \widehat{t}$ and $t^* \in [0, t]$ be the point maximizing $\|x\|$ in $[0, t]$. Then

$$\Phi_t = \min_{\tau\in[0,t]} \phi_\tau \ge \min_{\tau\in[0,t]} \frac{NH}{(1 + 2\|x(\tau)\|^2)^\beta} \ge \frac{NH}{(1 + 2\|x(t^*)\|^2)^\beta}.$$

Moreover, since $t^* \le t < \widehat{t}$, $t^* \in \Upsilon$ and we have

$$\Phi_t - D \ge \frac{NH}{(1 + 2\|x(t^*)\|^2)^\beta} - D \ge \frac{NH}{2(1 + 2\|x(t^*)\|^2)^\beta} > 0. \tag{4.2}$$

Hence, we may apply Proposition 4.4 to obtain

$$\begin{aligned}\|x(t)\|^2 &\le 2\|x(0)\|^2 + 2\|v(0)\|^2 \frac{1}{(\Phi_t - D)^2} \\ &\le 2\|x(0)\|^2 + \frac{8\|v(0)\|^2(1 + 2\|x(t^*)\|^2)^{2\beta}}{(NH)^2}.\end{aligned} \tag{4.3}$$

Since t^* maximizes Γ in $[0,t]$ it also does so in $[0,t^*]$. Thus, for $t=t^*$, (4.3) takes the form

$$\left(1+2\|x(t^*)\|^2\right) - 16\|v(0)\|^2\frac{(1+2\|x(t^*)\|^2)^{2\beta}}{(NH)^2} - \left(1+2\|x(0)\|^2\right) \leq 0. \tag{4.4}$$

Let $z = 1+2\|x(t^*)\|^2$. Then (4.4) can be rewritten as $F(z)\leq 0$ with $F(z) = z - \boldsymbol{a}z^{2\beta} - \boldsymbol{b}$.

We next reason by cases.

(i) Assume that $\beta < 1/2$. By Lemma 4.5, $F(z)\leq 0$ implies $(1+2\|x(t^*)\|^2)\leq U_0$. Since U_0 is independent of t we deduce that, for all $t<\widehat{t}$,

$$\|x(t)\|^2 \leq \frac{U_0-1}{2}.$$

Therefore, for all $t<\widehat{t}$,

$$(1+2\|x(t)\|^2)^\beta \leq (1+2\|x(t^*)\|^2)^\beta \leq U_0^\beta \leq \frac{NH}{3D}$$

It follows that

$$(1+2\|x(\widehat{t})\|^2)^\beta \leq \frac{NH}{3D}$$

which contradicts the main property of $\widehat{t}$. This shows that no such $\widehat{t}$ exists, i.e., that $\Upsilon = [0,+\infty)$. Hence, for all $t\geq 0$, $1+2\|x(t)\|^2\leq U_0$. This implies that

$$\phi_t \geq \frac{NH}{U_0^\beta}$$

for all $t\geq 0$ and therefore, the same bound holds for Φ_t. It follows that

$$\Phi_t - D \geq \frac{NH}{U_0^\beta} - \frac{NH}{3U_0^\beta} = \frac{2NH}{3U_0^\beta}$$

and by Proposition 4.3 we conclude that

$$\|v(t)\| \leq \|v(0)\| e^{-t\frac{2NH}{3U_0^\beta}}. \tag{4.5}$$

(ii) Assume that $\beta = 1/2$. Then (4.4) takes the form

$$(1+2\|x(t^*)\|^2)\,(1-\boldsymbol{a}) - \boldsymbol{b} \leq 0$$

and our hypothesis ensure that $1-\boldsymbol{a}>0$. This implies that

$$1+2\|x(t^*)\|^2 \leq \frac{\boldsymbol{b}}{1-\boldsymbol{a}} = U_0.$$

We now proceed as in case (i).

(iii) Assume finally $\beta > 1/2$ and let $\alpha = 2\beta$ so that $F(z) = z - \boldsymbol{a}z^\alpha - \boldsymbol{b}$. The derivative $F'(z) = 1-\alpha\boldsymbol{a}z^{\alpha-1}$ has a unique zero at $z_* = \left(\frac{1}{\alpha\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}}$ and

$$\begin{aligned}
F(z_*) &= \left(\frac{1}{\alpha\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} - \boldsymbol{a}\left(\frac{1}{\alpha\boldsymbol{a}}\right)^{\frac{\alpha}{\alpha-1}} - \boldsymbol{b} \\
&= \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}}\left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} - \boldsymbol{b} \\
&= \left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}}\left[\left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\right] - \boldsymbol{b} \\
&\geq 0
\end{aligned}$$

the last by our hypothesis. Since $F(0) = -\boldsymbol{b}<0$, $F''(z) = \alpha(\alpha-1)\boldsymbol{a}z^{\alpha-2}>0$ for all $z>0$, and $F(z)\to-\infty$ when $z\to\infty$ we deduce that the shape of F is as follows:

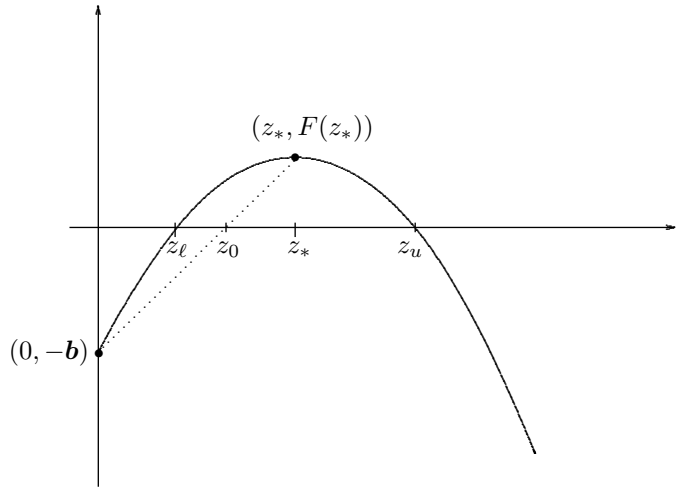

Figure 18: General shape of F

For $t \geq 0$ let $z(t) = 1 + 2\|x(t^*)\|^2$. Even though t^* is not continuous as a function of t, the mapping $t \mapsto z(t)$ is continuous and therefore, so is the mapping $t \mapsto F(z(t))$. This fact, together with (4.4), shows that, for all $t \geq 0$, $F(z(t)) \leq 0$. In addition, when $t = 0$ we have $t^* = 0$ as well and

$$z(0) = 1 + 2\|x(0)\|^2 = \boldsymbol{b} < \left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} = z_*.$$

This implies that $z(0) < z_\ell$ and hence, that for all $t < \widehat{t}$, $z(t) < z_\ell$.

Let z_0 be the intersection of the z axis with the line segment joining $(0, -\boldsymbol{b})$ and $(z_*, F(z_*))$ (see Figure 18). The line where this segment lies has equation

$$y + \boldsymbol{b} = z\frac{z_* - \boldsymbol{a}z_*^\alpha}{z_*}$$

from which it follows that

$$z_0 = \frac{\boldsymbol{b}}{1 - \boldsymbol{a}z_*^{\alpha-1}} = \boldsymbol{b}\frac{\alpha}{\alpha - 1}.$$

It follows that for all $t < \widehat{t}$

$$z(t) = 1 + 2\|x(t^*)\|^2 \leq \boldsymbol{b}\frac{\alpha}{\alpha - 1} = U_0.$$

We now proceed as in case (i).

4.4. **On shared information**

A case deserving particular attention is when some of the q informed agents share the same preferred direction. Assume, for instance, that the set of informed agents is partitioned in two subsets with all agents in each subset sharing the same information. In this case we can sharpen our results. Indeed, let

$$\overline{D} = \nu\sqrt{(q_1 + q_2) - \frac{1}{N}(q_1^2 + q_2^2 + 2q_1q_2\omega)}.$$

Then Proposition 4.2 can be sharpened as follows.

Proposition 4.6. *Assume that, in Model A, the set $\mathscr{I}$ is partitioned into $\mathscr{I} = \mathscr{I}_1 \cup \mathscr{I}_2$ and let $q_i = \#\mathscr{I}_i$, $i = 1, 2$. Assume further that the vectors $\boldsymbol{d}_i(t)$ are the same for the q_1 agents in $\mathscr{I}_1$ and that the same happens for $\mathscr{I}_2$. Denote by $\boldsymbol{d}_1^*$ and $\boldsymbol{d}_2^*$ these common vectors and let $\omega = \cos(\boldsymbol{d}_1^*, \boldsymbol{d}_2^*)$, $\omega \in [0, 1]$. Then, for $t \geq 0$, we have $\|Cd(t)\| = \overline{D}\|v(t)\|$.*

Proof. Assume, without loss of generality, that $\boldsymbol{d}_1^* = (\nu, 0)$ and $\boldsymbol{d}_2^* = (\nu\omega, \nu\overline{\omega})$ where $\overline{\omega} = \sin(\boldsymbol{d}_1^*, \boldsymbol{d}_2^*)$. Then, at all time t, the mean $\boldsymbol{m}$ for the information is given by

$$\boldsymbol{m} = \frac{\nu}{N}(q_1 + q_2\omega, q_2\overline{\omega})$$

and we have

$$\frac{N}{\nu}(\boldsymbol{m} - \boldsymbol{d}_1^*) = (q_1 + q_2\omega - N, q_2\overline{\omega})$$

as well as

$$\frac{N}{\nu}(\boldsymbol{m} - \boldsymbol{d}_2^*) = (q_1 + q_2\omega - N\omega, q_2\overline{\omega} - N\overline{\omega}).$$

Therefore,

$$\frac{N^2}{\nu^2}\|\boldsymbol{m}\|^2 = q_1^2 + q_2^2 + 2q_1q_2\omega,$$

$$\frac{N^2}{\nu^2}\|(\boldsymbol{m} - \boldsymbol{d}_1^*)\|^2 = q_1^2 + q_2^2 + 2q_1q_2\omega + N^2 - 2Nq_1 - 2Nq_2\omega$$

as well as

$$\frac{N^2}{\nu^2}\|(\boldsymbol{m} - \boldsymbol{d}_2^*)\|^2 = q_1^2 + q_2^2 + 2q_1q_2\omega + N^2 - 2Nq_1\omega - 2Nq_2.$$

It follows that

$$\begin{aligned}
\frac{N^2}{\nu^2}\|\boldsymbol{d}_\perp\|^2 &= \frac{N^2}{\nu^2}(N - q_1 - q_2)\|\boldsymbol{m}\|^2 + \frac{N^2}{\nu^2}q_1\|\boldsymbol{d}_1^* - \boldsymbol{m}\|^2 + \frac{N^2}{\nu^2}q_2\|\boldsymbol{d}_2^* - \boldsymbol{m}\|^2 \\
&= (N - q_1 - q_2)(q_1^2 + q_2^2 + 2q_1q_2\omega) \\
&\quad + q_1(q_1^2 + q_2^2 + 2q_1q_2\omega + N^2 - 2Nq_1 - 2Nq_2\omega) \\
&\quad + q_2(q_1^2 + q_2^2 + 2q_1q_2\omega + N^2 - 2Nq_1\omega - 2Nq_2) \\
&= N(q_1^2 + q_2^2 + 2q_1q_2\omega) + q_1(N^2 - 2Nq_1 - 2Nq_2\omega) + q_2(N^2 - 2Nq_1\omega - 2Nq_2) \\
&= N^2(q_1 + q_2) - N(q_1^2 + q_2^2 + 2q_1q_2\omega)
\end{aligned}$$

from where the statement follows. □

In case all the q informed agents share the same direction we obtain the following sharpenning of Proposition 4.2.

Proposition 4.7. *Assume that, in Model A, the vectors $\boldsymbol{d}_i(t)$ are the same for the q informed agents. Then, for $t \geq 0$, we have $\|Cd(t)\| = \overline{D}\|v(t)\|$ with*

$$\overline{D} = \nu\sqrt{q\left(1 - \frac{q}{N}\right)}.$$

Proof. Take $q = q_1 + q_2$ and $d_1^* = d_2^*$ (which implies $\omega = 1$) in Proposition 4.6. □

5. Conclusions

We have studied two simple Laplacian-based models for swarms with informed agents. Using the linear properties of their time-steps, we were able to prove analytically that there is a range of parameters and initial conditions for which the convergence to a swarming state is guaranteed. Despite their apparent simplicity, we showed numerically that both models can display complex consensus dynamics similar to those observed in more realistic algorithms. In particular, we found that for a wide range of initial configurations of Model A, a few informed individuals can accurately guide a large swarm in a target direction, which is a biologically desirable feature. This suggests that the underlying mechanism of such behavior could be found in simple dynamical interactions that are being captured by this model.

References

[1] C.M. Breder. Equations descriptive of fish schools and other animal aggregations. *Ecology*, 35:361–370, 1954.

[2] J. Cortes, S. Martinez, and F. Bullo. Spatially-distributed coverage optimization and control with limited-range interactions. *ESAIM Control Optim. Calc. Var.*, 11:691–719, 2005.

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

[4] I.D. Couzin, J. Krause, R. James, G.D. Ruxton, and N.R. Franks. Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218:1–11, 2002.

[5] F. Cucker and S. Smale. Best choices for regularization parameters in learning theory. *Found. Comput. Math.*, 2:413–428, 2002.

[6] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Trans. on Autom. Control*, 52:852–862, 2007.

[7] F. Cucker and S. Smale. On the mathematics of emergence. *Japan J. Math.*, 2:197–227, 2007.

[8] A. Czirok, H.E. Stanley, and T. Vicsek. Spontaneous ordered motion of self-propelled particles. *J. Phys. A: Math. Gen.*, 30:1375–1385, 1997.

[9] A. Czirok and T. Vicsek. Collective behavior of interacting self-propelled particles. *Physica A*, 281:17–29, 2000.

[10] U. Erdmann, W. Ebeling, and A. S. Mikhailov. Noise-induced transition from translational to rotational motion of swarms. *Phys. Rev. E*, 71:051904, 2005.

[11] J.A. Fax and R.M. Murray. Information flow and cooperative control of vehicle formation. *IEEE Trans. Aut. Contr.*, 49:1465–1476, 2004.

[12] G. Flierl, D. Grünbaum, S. Levin, and D. Olson. From individuals to aggregations: the interplay between behavior and physics. *J. Theor. Biol.*, 196:397–454, 1999.

[13] V. Gazi and K.M. Passino. Stability analysis of swarms. *IEEE Trans. Aut. Contr.*, 48:692–697, 2003.

[14] D. Grunbaum and A. Okubo. *Modeling social animal aggregations*, volume 100 of *Lecture Notes in Biomathematics*, pages 296–325. Springer-Verlag, 1994.

[15] S.-Y. Ha and J.-G. Liu. A simple proof of the Cucker-Smale flocking dynamics and mean-field limit. Preprint, 2008.

[16] S.-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinetic and Related Models*, 1:415–435, 2008.

[17] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. on Autom. Control*, 48:988–1001, 2003.

[18] E.W. Justh and P.S. Krishnaprasad. Equilibria and steering laws for planar formations. *Syst. and Contr. Lett.*, 52:25–38, 2004.

[19] A. Kolpas, J. Moehlis, and I.G. Kevrekidis. Coarse-grained analysis of stochasticity-induced switching between collective motion states. *Proc. Nat. Acad. Sc.*, 104:5931–5935, 2007.

[20] P. Ogren, E. Fiorelli, and N. E. Leonard. Cooperative control of mobile sensor networks: adaptive gradient climbing in a distributed environment. *IEEE Trans. Aut. Contr.*, 49:1292–1302, 2004.

[21] A. Okubo. Dynamical aspects of animal grouping: Swarms, schools, flocks, and herds. *Adv. Biophys*, 22:1–94, 1986.

[22] L. Perea, P. Elosegui, and G. Gómez. Extension of the Cucker-Smale control law to space flight formations. Preprint, 2008.

[23] J. Shen. Cucker-Smale flocking under hierarchical leadership. *SIAM J. Appl. Math*, 68:694–719, 2007.

[24] I. Suzuki and M. Yamashita. Distributed anonymous mobile robots: Formation of geometric patterns. *SIAM Journal on Computing*, 28:1347–1363, 1999.

[25] C.M. Topaz and A.L. Bertozzi. Swarming patterns in a two-dimensional kinematic model for biological groups. *SIAM J. Appl. Math.*, 65:152–174, 2004.

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

[27] K. Warburton and J. Lazarus. Tendency-distance models of social cohesion in animal groups. *J. Theoret. Biol.*, 150:473–488, 1991.