SOME FLOCKING PROPERTIES FOR A MODEL OF COLLECTIVE DYNAMICS WITH TOPOLOGICAL INTERACTIONS.

PATRICK CATTIAUX ♠ , **FANNY DELEBECQUE** ♠ ,
AND **OLYMPIO HACQUARD** ♣

♠ Université de Toulouse

♣ ENS Paris Saclay

Abstract. The mechanism of self-organization resulting in coordinated collective motion has received wide attention, in particular because collective behavior emerging out of self-organization is one of the most striking property of an animal group. In almost all the associated models, it is assumed that each individual in an animal group tends to align its direction of motion and velocity with those of its neighbors. The rules of alignment are often assumed to be of metric type, i.e. depending on the distances between individuals. The observation of large starlings flocks suggested in [BCC$^+$08] to rather consider nearest neighbor interactions, and several attempts were done to provide us with such models.

In the present note, we are interested in some properties of such a recent model presented in [BD16] and [BD17]. This model is a discrete model on velocities of N agents that interact according to their rank (and not their pairwise distance). In this paper, we present a detailed analysis of its flocking properties (flocking of the whole population as well as flocking of some subgroups of the population) and bounds for the flocking time. In the final section we discuss some consequences and interpretations of these results compared with existing numerical results in the literature.

Key words : Rank-based interactions, Flocking, Collective behavior, Birth and Death models.

MSC 2010 : 92D50, 91C20, 60K35, 60K25.

1. Introduction

The mechanism of self-organization resulting in coordinated collective motion has received wide attention, in particular because collective behavior emerging out of self-organization is one of the most striking property of an animal group. In almost all the associated models, it is assumed that each individual in an animal group tends to align its direction of motion and velocity with those of its neighbors. The rules of alignment are often assumed to be of metric type, i.e. depending on the distances between individuals. If there are numerous numerical studies of various models, only few have received a rigorous mathematical treatment.

Among the most known models of this type, let us mention the various Vicsek models (see [VZ12]) or the Cucker-Smale models ([CS07a, CS07b]) introduced in 2007. These models are thought as possible predictive models for "flocking".

Date: February 17, 2023.

The Cucker-Smale Model.

More precisely, the Cucker-Smale model considers N individuals determined by their respective positions and velocities $(x(t), v(t)) = (x_i(t), v_i(t))_{i=1,\dots,N}$ in $\mathbb{R}^d \otimes \mathbb{R}^d$ evolving in time via the dynamics

$$\begin{aligned} dx_i(t) &= v_i(t)\, dt \\ dv_i(t) &= -\frac{1}{N}\sum_{j=1}^{N} \Psi(|x_i(t) - x_j(t)|)\,(v_i(t) - v_j(t))\, dt\,; \end{aligned} \tag{1.1}$$

where Ψ is a non-negative and non increasing function. The minus sign in front of (1.1) facilitates a collective behavior, here the alignment of all velocities with the mean velocity as time goes to infinity. If in addition all positions stay at a bounded distance of the mean position (swarming property), this property is called "flocking".

The classical *communication rate* is given by $\Psi(u) = \lambda/(1+u^2)^r$. It was proved in these papers that for strong interaction ($r \le 1/2$) the system is flocking, while for $r > 1/2$ flocking may or not happen, depending on the initial configuration. The choice of the communication rate is of course artificial and no statistical method of calibration has been proposed. A lot of modified models have then been studied in the deterministic context, including delays, no collisions and many other features (such as leadership, cone of vision, pattern formation...) Some of them are presented in the review on Cucker-Smale type models [CHL17]. Among them, it was proposed in [MT11] to self-normalize the communication rates (instead of dividing by N). This model is particularely well-adapted to describe the dynamics of distant groups of various importance in terms of number of agents. In that prospect, it indeed may lead to the emergence of several flocking subgroups (*subflocks*).

Many of these models are deterministic, while randomness certainly appears at several levels: environment, individual degree of freedom or perturbation of the communication rate itself. Some noisy versions of Cucker-Smale like models have been considered, adding independent noises (Brownian motions) in [BCnC11, BCnC12, Péd18] in the dynamics of velocities, so that flocking becomes impossible, or correlated noises in various ways. These "noisy Cucker-Smale" models allowing flocking have been discussed in a few papers and we refer to [CDP18] for a first description of the situation and rigorous statements of how the flocking property may be adapted to the stochastic situation. Some stochatsic Cucker-Smale type models have also been derived and studied such as delayed equations in [HM18], or stochastic singular model for collision-avoiding in [HZ22] .

Another main question about these models is the micro-macro transition *i.e.* the behaviour of a single, or a finite number of individuals, as the size N of the population grows to infinity. Since (1.1) is of mean field type, it is expected that, provided the initial configuration is favorable, each individual will perform the same asymptotic averaged dynamics (McKean-Vlasov limit) and that any finite number of individuals will become independent in the limit (chaoticity). This is shown in various works for deterministic or stochastic models with independent Brownian noises, see e.g. [BCnC11, BCnC12, Péd18]. But of course such studies for general mean-field stochastic models have a long history, starting with M. Kac and H.P. McKean Jr. One may refer to [GM99] for a comprehensive survey of stochastic particles microscopic description of a Boltzmann's equation and to the very recent general approach in connection with biological models in [Car18]. Another nice survey of various models and

tools on the propagation of chaos in microscopic models can be found in [CD22] among the impressive amount of papers on the subject.

As a model for collective dynamics these models suffer some defaults. One of the main default is the following: if an attractive behavior depending on the distance between individuals is plausible for bacterias or cells (though aggregation type descriptions are presumably more convincing), it is much less clear for animals that have the option of free will. A human in a crowd of unknowns will try to join the closest familiar face, whatever the distance. This has been clearly observed for some birds in [BCC$^+$08] from which we extract the following:

Numerical models indicate that collective animal behavior may emerge from simple local rules of interaction among the individuals. However, very little is known about the nature of such interaction, so that models and theories mostly rely on aprioristic assumptions. By reconstructing the three-dimensional positions of individual birds in airborne flocks of a few thousand members, we show that the interaction does not depend on the metric distance, as most current models and theories assume, but rather on the topological distance. In fact, we discovered that each bird interacts on average with a fixed number of neighbors (six to seven), rather than with all neighbors within a fixed metric distance. We argue that a topological interaction is indispensable to maintain a flock's cohesion against the large density changes caused by external perturbations, typically predation. We support this hypothesis by numerical simulations, showing that a topological interaction grants significantly higher cohesion of the aggregation compared with a standard metric one.

If one can discuss the word "indispensable", since classical Cucker-Smale models may exhibit flocking situations, and if one can discuss the fact that "simulation is proof", these observations have pointed out the topological nature of interactions for some populations. Once again some models where interactions between individuals is restricted to nearest neighbors have been proposed and studied numerically (see e.g. [BFW11, SB14, KB13, CD16] among others). One can also mention [MNG17] where another type of interaction based on anticipation is proposed.

In the recent [KD21], the authors explore numerically the compared effects of metric versus topological local interactions on the dynamics of flocking in the following simple model :

$$\begin{aligned} dx_i(t) &= v_i(t)\,dt \\ dv_i(t) &= -\frac{1}{N_{int}}\sum_{j=1}^{N_{int}}(v_i(t)-v_j(t))dt. \end{aligned}$$

Here, the set $\{1,\dots,N_{int}\}$ represents, in the case of metric interactions, the particles lying in a ball of fixed radius R centered in x_i and, in the case of topological interactions, the N_{int} nearest neighbors. Varying the number of interacting particles (*i.e* the radius R for metric interactions and the number of interacting neighbors for topological interactions), the authors evidence that topological interactions seem to be more effective in bringing order in the agents group.

It is thus interesting to replace interactions depending on distances between individuals by interactions only depending on their closeness, or their ranks. Replacing metric interactions my topological ones can be done in different ways. In taht prospect, in [Has13], the authors turns the Cucker-Smale deterministic model into a topological one by replacing the interaction

term $\Psi(|x_i(t) - x_j(t)|)$ in (1.1) by,

$$\Psi\left(\sum_l \mathbf{1}_{|x_l(t)-x_i(t)|<|x_j(t)-x_i(t)|}\right) ,$$

hence particles are interacting through their ranks and not in a metric way, see the discussion below. In the spirit of our previous paper [CDP18] one may introduce different noise in such a model. Some stochastic models of this kind have been introduced for other purposes, for instance finance in dimension 1 (see e.g [Shk12, JM08, Rey15]).
Another way to make classical metric interactions asymetric, to introduce leadership or more complicated interactions is to impose a graph of interactions. The drawback of these models lies in the proofs of flocking. They indeed usually require invariant in time connectivity properties in the communication network along the dynamics. In [Mar14] and later in [CD16], the authors manage to determine practical conditions, depending only on initial data, that ensure sufficient robustness in the communication network to lead to flocking, or to show unconditionnal flocking if the number of interacting particles stays sufficient along time. We refer the reader to [DHJK20] and references therin for more such models.

Recently another stochastic "toy" model was proposed by A. Blanchet and P. Degond ([BD16, BD17]). This model also adresses the question of topological vs. metric interactions. Let us describe this model, which will be called the B-D model in the sequel.

The Degond-Blanchet Model.

(1) The dynamics is a succession of free flights and jumps (collisions in Blanchet-Degond terminology), that is, between two jump times, particles follow straight paths with constant velocity, *i.e.* $dx_i(t) = v_i(t)dt$ and $dv_i(t) = 0$.
(2) Jumps occur according to a Poisson process $P(t)$ with intensity (rate) $\lambda(N)$, i.e. times between jumps are independent, identically distributed according to an exponential distribution with parameter $\lambda(N)$.
When the clock of the Poisson process rings (*i.e.* at each jump time), a particle i is chosen at random (*i.e.* uniformly with probability $1/N$).
(3) When it "jumps", particle i chooses its partner $j \neq i$ according to a probability $\pi^N_{i,j}$ to be described below, and then (x_i, v_i) is changed into (x_i, v_j) *i.e.* i immediately aligns its velocity on v_j.

Remark 1.1. Note that the choice of one Poisson process $P(t)$ with intensity rate $\lambda(N)$ is equivalent to the choice of N Poisson process $P_i(t)$ with intensity $\lambda(N)/N$, all these processes being independent and $P(t) = \sum_{i=1}^N P_i(t)$. In that case, the different clocks almost surely ring at only one at the same time. ♢

Now let us precisely define the probabilities $\pi^N_{i,j}$. As they depend on the rank of j with respect to i, let us start with the definition of this rank.

Definition 1.2. *For each i, consider the increasing ordered set of distances $\{|x_i - x_j|,\ j \neq i\}$. Define $R^N(i,j)$ the rank of particle j for this order if $j \neq i$ and $R^N(i,i) = 0$. If distances are the same for two different indices or more, one chooses one of the corresponding indices randomly with uniform probability. We then introduce the normalized rank*

$$r^N(i,j) = \frac{R^N(i,j)}{N-1} .$$

Remark 1.3. First note that, $R^N(i,i) = 0$ and $R^N(i,j) = 1, ..., N-1$.
Also note that, in the case where there are no tied particles (*i.e* the distances between any pair $(i,j),\ i \neq j$ are different), one can write

$$r^N(i,j) = \frac{1}{N-1} \sum_{l \neq i, 1 \leq l \leq N-1} \mathbf{1}_{B(x_i, |x_i - x_j|)}(x_l) \ = \ \frac{1}{N-1} \sum_{l \neq i, 1 \leq l \leq N-1} \mathbf{1}_{|x_i - x_l| < |x_i - x_j|} \, .$$

♢

We now describe the transition process

Definition 1.4. *Let $k^N : [0,1] \to \mathbb{R}^+$ be continuous.*
Define, for all $i = 1, \ldots, N$, $\pi^N_{i,i} = 0$ and for $j \neq i$,

$$\pi^N_{i,j} = \frac{k^N(r^N(i,j))}{\sum_{l=1}^{N-1} k^N(l/N-1)} \, ,$$

Note that, because $r^N_{i,i} = 0$, $\sum_{j=1, j \neq i}^{N} \pi^N_{i,j} = 1$, which is consistent with the choice of $\pi^N_{i,i} = 0$, and that we may also assume, or not, that $\int_0^1 k(u)du = 1$.

The model, as simple as it may seem, has the following several interesting properties.

(1) It corresponds to the expected notion of topological interactions : if in particular, for some $m \leq N-1$,
$$k^N(s) = \mathbf{1}(s \in [0, m/(N-1)]) \, , \tag{1.2}$$
then each particle only interacts with its m-nearest neighbors. When N grows one can also take $m = \alpha\,(N-1)$ so that each particle interacts with a given fraction (α) of the whole population.
(2) It is a crucial point to remark that permutation invariance (exchangeability) is preserved by the dynamics, since the rank is permutation invariant.
(3) From the mathematical point of view, the use of the rank creates new difficulties. The first one is the break of symmetry in the interaction. Indeed, in the usual Cucker-Smale system for instance, it is crucial that the communication term $\Psi_{i,j} := \Psi(|x_i - x_j|)$ is symmetric : $\Psi_{i,j} = \Psi_{j,i}$. But, in our case, note that the rank is not symmetric : $R^N_{i,j} \neq R^N_{j,i}$. (think of a case where agent j might be the closest neigbour of agent i, whereas i is not the closest neighbour of agent i.) Therefore, the transitions $\pi^N_{i,j}$ won't be symmetric except if k^N is constant.
(4) Similarly, the usual "Cucker-Smale" interaction is a pair-wise interaction : the interaction term $\Psi_{i,j}(v_j - v_i)$ only depends on the positions and velocties of agents i and j, whereas, in this model, the rank $R^N_{i,j}$ (and therefore $\pi^N_{i,j}$) does not depend only on both particles but on the whole population of particles so that the "mean field" nature of the model is not completely clear.

Remarks (3) and (4) have an immediate consequence: the first one for the study of flocking, the second one for the large population limit.

In this paper, we study the flocking property, *i.e.* the long time behaviour of the solution for a fixed N. This problem is not discussed (except via some simulations) in [BD16]. It is worth noticing that the set of all velocities is always a subset of the initial configuration of

velocities. Indeed, at each jump time, a particle aligns its velocity on another one : no new velocity is created, contrary to the usual Cucker-Smale type models. Hence alignment will mean that all velocities are the same after some almost surely finite stopping time. After this time the distances between all positions will no more change. It follows that alignment of velocities automatically implies "swarming" (distances between particles remain bounded), and is thus enough for flocking. Our main results are the following:

(1) If interaction is allowed with a sufficiently large number of neighbors (more than one half of the population), in particular with all of them, then the system is almost surely flocking, whatever the initial configuration is. This is called "unconditional" flocking. In addition we shall provide theoretical bounds for the mean flocking time (i.e the first time when all the agents share the same velocity).
(2) In any case, after some almost surely finite stopping time, the population is divided into several flocking sub-populations that do no more interact one with another. The size of these subgroups of individuals is random, and we may have situations where flocking holds with a positive probability while non flocking also holds with a positive probability. This is somehow different from what is usually called "conditional" flocking in the literature (*i.e* flocking under conditions on the initial positions and velocities) : here same initial conditions may lead to flocking (or not) with positive probability.

The large population limit is studied in [BD16] for $k^N = k$ independent of N, and in [BD17] for k^N as in (1.2). In both papers the authors show that, for an appropriate choice of the intensity $\lambda(N)$, the limiting behaviour of the distribution of each particle, as $N \to \infty$, is given by some kinetic non linear P.D.E. In [DP18], a rigorous proof of convergence is given when $k^N = k$ belongs to some subclass of the analytic functions (including polynomials). The result is obtained by solving the whole underlying BBGKY hierarchy. Notice that in [Has13] the large population limit is studied by showing in particular that the associated "topological" Cucker-Smale model satisfies mean-field like properties.

The paper is organized as follows :

Section 2 is devoted to the case where all particle interact with the whole population, *i.e* the intercation kernel k is positive. We show, that in that case, the system amost-surely flocks.

In section 3, we focuse on the special case where k is constant, which means that all the agets have the same probability to be chosen for velocity alignement. We first draw a parrallel with the study of Moran birth and death process, deduce an upper-bound for the flocking time and present numerical simulations.

In section 4, back to the more general case where k is positive, we use the results on birth and death models to build a slowed process and get a better bound for the flocking time.

Finally, section 5 is devoted to the case where k may cancel and almost sure flocking does not hold anymore.

We end up with a discussion on the relevance of the model.

2. Almost sure flocking, case $k > 0$.

We recall that, for teh rest of the paper, the size N of the population is fixed. We will therefore suppress the superscripts N in all the notations. In addition we do not indicate the initial configuration when an estimate is true for all initial configurations.
We start with a few notations that will be usefull in the rest of the paper.

Notations 2.1. Consider the subsequence of jump times $(T_n)_{n\in\mathbb{N}}$, we denote by $\bar{V}(n) = \{v_i(T_n)\,;\, 1 \le i \le N\}$ the set of all velocities at time T_n.
We also define the skeleton Markov Chain $(X(n), V(n)) = (x(T_n), v(T_n))$, whose transition probabilities are given by the $\pi_{i,j}$'s.

As already said, no new velocity is created along the dynamics of the B-D model. It is therefore clear that $\bar{V}(n) \subseteq \bar{V}(m)$ for $n \ge m$ so that the limiting set $\bar{V}(\infty)$ exists. Let us now recall the definition of flocking in our setting.

Definition 2.2. *We shall say that the system of particles is (almost surely) flocking if $\bar{V}(\infty)$ is reduced to a single (random) vector $v(\infty)$. If $S_\infty = \inf\{n\,;\, \#\bar{V}(n) = 1\}$, the system flocks if and only if $S_\infty < +\infty$ almost surely. We will call $T_\infty = T_{S_\infty}$ the flocking time for the system in the sequel.*

It is easy to build situations where flocking does not hold

Example 2.3. For $d = 1$, $N = 4$ and $k(s) = 3\,\mathbf{1}_{\{s\le 1/3\}}$, *i.e.* an interaction with the closest neighbor (i.e $k(s) = \mathbf{1}(s \in [0, 1/(N-1)])$), choose $x(0) = (-3, -2, 2, 3)$ and $v(0) = (-2, -1, 1, 2)$. It is immediately seen that for all t the clusters of particles $1, 2$ and $3, 4$ remain unchanged, so that $\bar{V}(\infty)$ contains two elements. ♢

In the Example 2.3, the lack of flocking is due to the fact that two initial clusters of agents never interact. In the following theorem, we claim that, if any pair of particles has a positive probability to interact (i.e $\pi_{i,j} > 0$ as soon as $j \ne i$) then flocking almost surely happens.

Theorem 2.4. *Assume that $k(j/N - 1) > 0$ for all $j = 1, ..., N-1$. Then for any initial condition, the system is almost surely flocking.*

Proof. Note that the assumption on k is equivalent to the positivity of the $\pi_{i,j}$ for any pair (i, j). With Notations 2.1, denote by $\mathcal{F}$ the following event:

$\mathcal{F}$:= "*at each time T_i, $i = 1, ..., N-1$, agent i is selected and selects agent N to align with : $V_i(i)$ jumps to $V_N(i)$.*"

It is clear that, on $\mathcal{F}$, the system is flocking as all the agents have velocity $V_N(0)$ at T_{N-1}. Moreover,

$$\mathbb{P}(\mathcal{F}) \;=\; \frac{\prod_{j=1}^{N-1} \pi_{j,N}}{N^{N-1}}\,.$$

Define

$$\beta_N = \min_{1\le j\le N-1} \pi_{j,N} = \min_{1\le j\le N-1} \frac{k(j/N-1)}{\sum_{l=1}^{N-1} k(l/N-1)} \quad \text{(see Definition 1.4.)} \qquad (2.1)$$

It follows from the assumption of Theorem 2.4 that $\beta_N > 0$ and thus

$$\mathbb{P}(\mathcal{F}) \geq \left(\frac{\beta_N}{N}\right)^{N-1} := p > 0\,.$$

Note that p does not depend on the initial condition. Denote by S_∞ the flocking time for the skeleton chain. It follows that $\mathbb{P}(S_\infty = +\infty) \leq (1-p)$ ans this lower bound does not depend on the initial configuration. But using the Markov property we have

$$\mathbb{P}_{X(0),V(0)}(S_\infty = +\infty) \leq \mathbb{P}_{X(0),V(0)}(\mathbf{1}_{\mathcal{F}^c}\,\mathbb{P}_{X(N),V(N)}(S_\infty = \infty)) \leq (1-p)^2$$

so that by induction $\mathbb{P}(S_\infty = +\infty) = 0$. Since the T_n are almost surely growing to infinity, it follows that $\mathbb{P}(T_\infty = +\infty) = 0$. □

The previous proof also gives a first upper bound for the expectation

$$\mathbb{E}(S_\infty) = O\left(\left(\frac{N}{\beta_N}\right)^{N-1}\right)$$

which is presumably disastrous. In order to get an estimate for the expecttaion of the system's flocking time T_∞, remember that $T_\infty = T_{S_\infty}$.
Using Wald identity in the identity

$$T_\infty = \sum_{j=1}^{S_\infty} (T_j - T_{j-1})\,,$$

since S_∞ and the increments $T_{j+1} - T_j$ are independent, and since the increments $T_{j+1} - T_j$ follow an exponential law of parameter $\lambda(N)$, we deduce

$$\mathbb{E}(T_\infty) = \frac{1}{\lambda(N)}\, O\left(\left(\frac{N}{\beta_N}\right)^{N-1}\right).$$

We may refine the argument in order to improve this upper bound and state the following proposition.

Proposition 2.5. *Under the assumption of Theorem 2.4,*

$$\mathbb{E}(T_\infty) = \frac{1}{\lambda(N)\,((N-1)!)^2)} \underset{N\to\infty}{O}\left(\frac{N}{\beta_N}\right)^{N-1}.$$

And thus, using Stirling formula,

$$\mathbb{E}(T_\infty) = \underset{N\to\infty}{O}\left(\frac{e^{2N}}{N\,\lambda(N)\,(N\beta_N)^{N-1}}\right).$$

Proof. To refine the previous upper bound, let us study the evolution along time of the cardinal of a cluster of velocities.
Consider a jump time T_n and divide the N agents into clusters of different velocities. We recall that, at each jump time, a given cluster of velocity can either integrate a new agent (that has chosen an agent of the cluster to align on), keep the same number of agents or loose an agent (that has chosen to change its velocity to an other one). Its cardinal at time T_{n+1} therefore is

$$c(n+1) = c(n) + \varepsilon \quad \text{where} \quad \varepsilon \in \{-1, 0, +1\}.$$

Let us state the following lemma describing the evolution along time of the cardial of a given cluster of velocity.

Lemma 2.6. *Consider at time T_n a given cluster of velocity $\bar{v} \in \bar{V}(n)$, denote by $c(n)$ its cardinal : $c(n) = \#\{i = 1 \dots N,\ \ V_i(n) = \bar{v}\}$ and $p(n, r, r+\varepsilon) = \mathbb{P}\left(c(n+1) = r + \varepsilon \mid c(n) = r\right)$. Without loss of generality, we may number from 1 up to $c(n)$ the agents of the considered cluster. Then, if $1 \leq r \leq N$,*

$$\left\{ \begin{array}{c} p(n,r,r+1) = \frac{1}{N} \sum_{j=r+1}^{N} \sum_{i=1}^{r} \pi_{j,i} \\ p(n,r,r-1) = \frac{1}{N} \sum_{j=r+1}^{N} \sum_{i=1}^{r} \pi_{i,j} \\ p(n,r,r) = 1 - p(n,r,r+1) - p(n,r,r-1) \end{array} \right.$$

Proof of the Lemma : These formula are straight forward. Indeed, remember that $p(n, r, r+1)$ is the probability that the cluster of r agent may integrate a new one : it means that at time T_n, an agent j is chosen among those not in the cluster ($i.e\, j \in \{r+1, \dots, N\}$) with probability $\frac{1}{N}$, an agent i is chosen in the cluster ($i.e\ i \in \{1, \dots, r\}$) that j aligns on, with probability $\pi_{j,i}$.

Let us now finish the proof of Proposition 2.5. We use Lemma 2.6 choosing $C_{max}(n)$ to be one of the largest clusters at time T_n. Denoting $c_{max}(n) = \#\left(C_{max}(n)\right)$, and with the definition (2.1) of β_N, it holds :

$$\begin{aligned} p(n, c_{max}(n), c_{max}(n)+1) &= \sum_{j > c_{max}(n)} \frac{\pi_{j,1} + \dots + \pi_{j,c_{max}(n)}}{N} \\ &\geq \frac{\beta_N\, c_{max}(n)\left(N - c_{max}(n)\right)}{N}. \end{aligned}$$

Now, introduce the following event $\mathcal{E}$: during $N-1$ successive steps c_{max} increases of 1 or is fixed equal to N. What we have done before shows that, for $1 \leq \ell \leq N-1$,

$$\mathbb{P}(\mathcal{E} | c_{max}(0) = \ell) \geq \prod_{n=\ell}^{N-1} \left(\frac{\beta_N\, n\,(N-n)}{N} \right) \geq \prod_{n=1}^{N-1} \left(\frac{\beta_N\, n\,(N-n)}{N} \right).$$

Therefore, for any initial configuration :

$$\mathbb{P}(\mathcal{E}) \geq \left(\frac{\beta_N}{N} \right)^{N-1} \left((N-1)!\right)^2.$$

We may then argue as we did with the event $\mathcal{F}$ and deduce the bounds given in Proposition 2.5. □

In the next subsection we shall study more carefully the expected flocking time in the particular cas where $k = 1$.

3. The special case k=1

3.1. Expected flocking time, connection with the neutral Moran model.

We first assume that $k = 1$. From the modelling point of view it is certainly the less

interesting situation, but it will help to better understand the general case. It indeed corresponds to a case where any random particle changes its velocity with any other one, regardless to its rank, position or velocity, *i.e* for all $i \neq j$, $\pi_{i,j} = \frac{1}{N-1}$. Notice that in this case, because the dynamics of the velocities do not depend on the positions, $V(n)_{n\in\mathbb{N}}$ is a Markov chain.
It turns out that this new model is (more or less) well known in Evolutionary Dynamics: it is close to the multi-allele Moran model for an haploid population of size N.

Connection to the Moran Model : The Moran Model for 2 alleles is a birth and death process that describes the dynamics of a population of constant size N in which two alleles A and B of a gene are in competition.
At each time-step, a pair of agents (i, j) is chosen with uniform probability : one dies and one reproduces. Note that the pair (i, i) might be chosen, and in that case, the number of A-allele bearers and B-allele bearers does not change at this particular time-step. In the simplest case of Neutral drift, no fitness advantage or disandvantage is given to any of the alleles.
In that case, counting the number of A-bearers among N agents at each time-step leads to a Markov chain with the following transition probabilities :

$$Q(i, i-1) \;=\; Q(i, i+1) \;=\; \frac{N-i}{N}\frac{i}{N}, \qquad Q(i,i) = 1 - 2\,Q(i, i-1). \tag{3.1}$$

In the Moran model, no mutation is taken into account, so, as soon as A bearers have dispearred (or achieved the whole population), they can't be any new A-bearer : $P_{0,0} = P_{N,N} = 1$. We say that 0 and N are absorbing states for the Markov chain, and the time to reach N or 0 is called the "fixation time". In the case of the Moran model, the expression of the mean fixation time can be computed.

Back to our model : The parrallel with our model is easily drawn, the main difference is that, in our model, an agent cannot "align its velociy with itself" : $\pi_{i,i} = 0$, whereas it is possible in the Moran Model that an agent is chosen to die and reproduce at the same time step : $P_{i,i} > 0$.
We assume in addition (in order to stick to the Moran model) that $\#\bar{V}(0) = 2$, *i.e.* all initial velocities are equal to either $\bar{v}_1$ or $\bar{v}_2$. Then, the number $c_1(n)$ of agents with velocities $\bar{v}_1$ at time n is a (discrete-time) birth and death Markov chain, whose transition probabilities for l in $\{1, ..., N-1\}$ are given by

$$Q(l, l+1) \;=\; Q(l, l-1) \;=\; \frac{(N-l)l}{N(N-1)}, \qquad Q(l,l) \;=\; 1 - 2\,Q(l, l-1). \tag{3.2}$$

and which is absorbed when it hits 0 or N. It can therefore be adapted from Moran's computations, that

$$\mathbb{E}(S_\infty | c_1(0) = r) \;=\; (N-1)\left(\sum_{y=1}^{r}\frac{N-r}{N-y} \;+\; \sum_{y=r+1}^{N-1}\frac{r}{y}\right). \tag{3.3}$$

Thanks to this equality, we are now able to state the following result in the case where $k = 1$:

Theorem 3.1. *Assume that $k = 1$ (and $N \geq 6$). There exists C such that for all initial configuration (x, v) satisfying $\#\bar{V}(0) = m$,*

$$\frac{(N-1)(N-6)}{4\lambda(N)} \leq \mathbb{E}_{(x,v)}(T_\infty) \leq \frac{\log_2(m)\, N^2}{\lambda(N)}\,.$$

Proof. Due to (3.3), it is easily seen that, for any $1 \leq r \leq N$, $\mathbb{E}(S_\infty | c_1(0) = r) \leq N^2$. A general lower bound can also be obtained writing

$$\begin{aligned}
\mathbb{E}(S_\infty | c_1(0) = r) &= (N-1)\left(\sum_{y=1}^{r} \frac{N-r}{N-y} + \sum_{y=r+1}^{N-1} \frac{r}{y}\right) \\
&= (N-1)(N-r) \sum_{y=N-r}^{N-1} \frac{1}{y} + (N-1)r \sum_{y=r+1}^{N-1} \frac{1}{y} \\
&\geq \frac{(N-1)^2}{2} \sum_{y=[N/2]+1}^{N-1} \frac{1}{y}
\end{aligned}$$

where $[y]$ denotes the integer part of y. We thus deduce that for $N \geq 6$, and all $r = 1, ..., N-1$

$$\mathbb{E}(S_\infty | c_1(0) = r) \geq \frac{(N-1)(N-6)}{4}\,. \tag{3.4}$$

More precisely, if $r = pN$ for some $0 < p < 1$, as N goes to infinity, using (3.3) and the asymptotics of the harmonic series, we get

$$\mathbb{E}(S_\infty | c_1(0) = r) \sim -N^2\,(p \ln p + (1-p)\ln(1-p)).$$

We did not find in the literature an analogue for more than 2 different alleles, nor succeeded in deriving such a formula. We thus will only give somewhat rough upper and lower bounds. Indeed if the initial set $\bar{V}(0)$ contains $2^a \leq m < 2^{a+1}$ elements, we may divide it into two blocks of size $2^{a-1} \leq m' < 2^a$ and consider all velocities in each block as the same one. This new chain behaves like the previous one with two initial velocities ($k = 1$ is of course crucial here). The fixation time of this new chain has an expectation less than CN^2. Iterating the procedure and using the Markov property, the fixation time of our "m-allele" model has an expectation less than $\log_2(m)\, N^2$.

Also remark that, if S_∞^j denotes the first time the initial velocity $\bar{v}_j$ disappears or becomes the unique remaining velocity, then, it appears that

$$S_\infty = \max_{1 \leq j \leq m} S_\infty^j\,.$$

Now, using the fact that $k = 1$ shows that each of the S_∞^j have the same distribution as the fixation time for the "2-allele" model, starting for r_j the number of particles sharing the initial velocity $\bar{v}_j$. We may thus bound from below the expectation of S_∞ by the one of the corresponding 2-allele model, use (3.4) and Wald identity to get the anounced estimates.

□

Remark 3.2. An explicit (but tedious) computation of the mean flocking time can be done for small N (≤ 4) and leads to the exact value $(N-1)^2$. ♢

3.2. **Numerical simulations, case $k = 1$.**

In this section, we present a few numerical experiments that we have carried out in order to illustrate Theorem 3.1. Note that in the present case where $k = 1$, the positions of the particles do not interfere in the dynamics. We build the following experiment (in one dimension) :

- Consider N initial positions in $[-5, 5]$, choose a set of initial velocities among $\{-1, -0.5, 0.5, 1\}$ (we carry out experiments with 2, 3 or 4 possible initial velocities)
- Simulating a Poisson process with rate $\lambda(N) = N$ we define the sequence of jump times T_n.
- Between each jump time, the particles follow straight paths with constant velocity.
- At each jump time, one chooses a particle i uniformly with probability $\frac{1}{N}$ and another one, j uniformly with probability $\frac{1}{N-1}$ and one changes v_i into v_j.
- At each time step, if the particles have not flocked (*i.e* all the velocities are the same), they follow their free flight until the following jump time.

We carry out 100 experiments with the same initial data (especially the same distribution of velocities) and compute the mean flocking time (*i.e* the first time when all the particles flock). We also wonder how the initial proportions of each velocity impact the flocking time and the "finally chosen velocity".

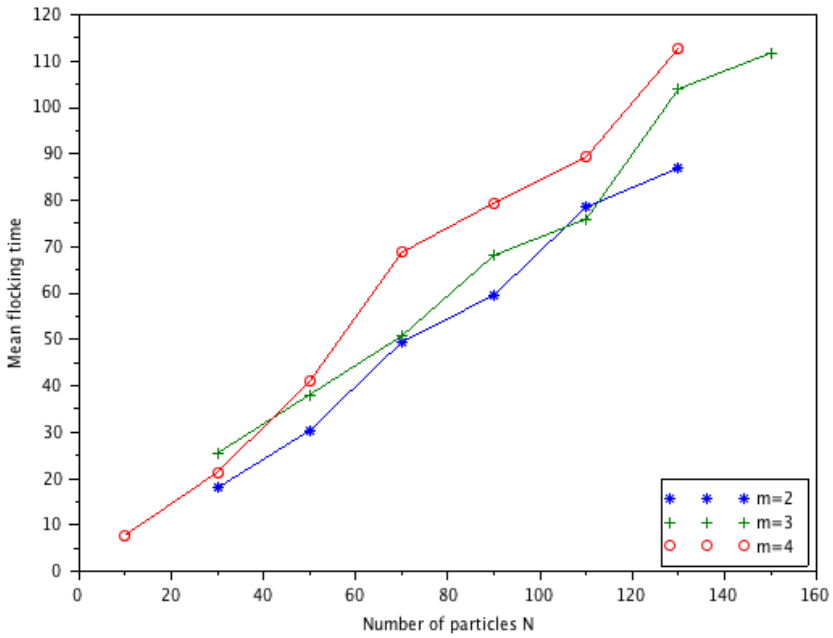


FIGURE 1. Mean flocking time as a function of N, case $k = 1$

	$m=2$		$m=3$			$m=4$			
	v_1	v_2	v_1	v_2	v_3	v_1	v_2	v_3	v_4
$N=50$ initial prop.	77	23	18	24	58	26	44	10	20
$N=50$ final prop.	81	29	17.6	30.4	52	24.5	42.1	10.1	23.3
Mean Flocking time	24.4	41.4	39.6	42.7	34	36.3	37.7	47.2	44.2
$N=70$ initial prop.	21	79	18.6	60	21.4	24.3	34.3	28.5	12.9
$N=70$ final prop.	19	81	18.8	59.5	21.7	24.1	34.6	28.5	12.8
Mean Flocking time	55.8	29.8	60.6	40.4	60	56	58	57.1	64
$N=100$ initial prop.	37	63	15	46	39	30	35	19	16
$N=100$ final prop.	38.6	61.4	14.4	45.5	40	27.2,	36.2	20.2	16.4
Mean Flocking time	81	54.8	75.7	96.1	73.8	79.2,	85.3	92.6	91.9

TABLE 1. Proportion of experiments flocking on each velocity w.r.t N and m

As we can see on Figure 1, regardless to the cardinal m of the initial set of velocities (here m=2,3,4), the mean flocking time increases linearly with respect to N, as shown in Theorem 3.1 in the case where $\lambda(N) = N$.

In Table 1 we present, for $N = 50, 70, 100$ particles and in the cases of 2, 3 or 4 different initial velocities the initial percentage of each velocity, the percentage of experiments (we carried out 1000 experiments with the same initial condition) flocking on each velocity and the mean flocking time for the experiments flocking on each different velocity. In the case of the Moran model, it is well-known that the probability of flocking on a velocity v_i equals the proportion of v_i in the initial data. It appears clearly in Table 1 that the proportion of a velocity in the initial data is basically the same as the probability to flock on that velocity. It is not so easy to draw a conclusion about the mean time needed to flock on a given velocity. It doesn't really seem to be linked to the initial proportion, even if, of course, it takes a bit longer to flock on v_i if there are far less particles with velocity v_i at initial time.

4. The general $k > 0$ case : a bound for the expected flocking time.

Using the same idea as before to deduce the general case $\#\bar{V}(0) = m$ from the case $\#\bar{V}(0) = 2$, we only need to look at the case $\#\bar{V}(0) = 2$ using Notations 2.1.

In the general case where $k > 0$ is not constant, $V(n)$ is no more a Markov chain, but $c_1(n)$ (the number of particles with velocity $\bar{v}_1$) still satisfies $c_1(n+1) = c_1(n) + \varepsilon(n)$ where $\varepsilon(n)$ takes its values in $\{-1, 0, 1\}$. It is thus natural to compare the evolution of $c_1(.)$ with a birth and death process for which we can get an explicit formula for the mean fixation time.

As a first step, we build a "slowed birth and death process" that simulates the worst case of our dynamics at each time step using lower and upper bounds for the transition probabilities. We also show that it's fixation time is greater than the flocking time of our dynamics, and will thus provide us with an upper bound for T_∞. As a second step, we compute estimates for the expectation fixatuion time of the constructed birth and death model.

4.1. **Construction of a "slowed birth and death process" to simulate the worst case.**

Let us assume, for simplicity, that k is positive and non-increasing. In this section, we aim at constructing a coupling of two different processes in order to approach the worst possible case for the dynamics and study the expected flocking time for this coupling.
Let us first, for any fixed $j \in \{1, ..., N-1\}$, define

$$\sigma'_j = \frac{\sum_{l=1}^{j} k(l/N-1)}{\sum_{l=1}^{N-1} k(l/N-1)}, \quad \text{and} \quad \sigma_j = \frac{\sum_{l=N-j}^{N-1} k(l/N-1)}{\sum_{l=1}^{N-1} k(l/N-1)}. \tag{4.1}$$

Basically, σ'_j (resp. σ_j) represents the probability for a particle to align with one of its j nearest neighbors, ($i.e$ with normalized rank $\frac{1}{N-1}...\frac{j}{N-1}$) (resp. with one of its j most distant neighbors, $i.e$ neighbors with normalized ranks $\frac{j+1}{N-1}, \ldots, \frac{N-1}{N-1}$). Now, as k is assumed non increasing, they also provide us with upper bounds (resp lower bounds) for any sum of j probabilities $\pi_{k,\ell}$.
Let us now consider the number $c_1(n)$ of particles with velocity $\bar{v}_1$ at time T_n. With the same notations used in Lemma 2.6, combining the equations for $p(n, r, r+1)$ and $p(n, r, r-1)$ given in Lemma 2.6 and the above remark on σ'_j and σ_j, it holds

$$\begin{cases} \frac{N-r}{N}\,\sigma_r \;\le\; p(n,r,r+1) \;\le\; \frac{N-r}{N}\sigma'_r, \\ \\ \frac{r}{N}\,\sigma_{N-r} \;\le\; p(n,r,r-1) \;\le\; \frac{r}{N}\,\sigma'_{N-r} \\ \\ \frac{1}{N}\left((N-r)\sigma_{N-r} + r\sigma_r\right) \;\le\; p(n,r,r) \;\le\; \frac{1}{N}\left((N-r)\sigma'_{N-r} + r\sigma'_r\right). \end{cases} \tag{4.2}$$

Let us now construct two coupled processes $R'(n)$ and $R(n)$ as follow : $R'(n)$ is a non-homogeneous Markov chain with transitions $p(n, r, r+1)$, $p(n, r, r-1)$ and $p(n, r, r)$. Of course the law of the process $R'(.)$ is by definition the same as the one of $c_1(.)$.
To define $R(n)$, we keep in mind that our goal is to create a birth and death process $\{0, ..., N\}$ that is a somehow "slowed version" of $R'(n)$: at each time-step, we minimize the probability to go towards the closest absorbing state (adding or suppressing an agent with velocity $\bar{v_1}$) and maximize the probability to go towards the other absorbing state. Thus, at each time-step, we choose the smallest probability to gain (resp loose) a new agent in the cluster of velocity $\bar{v}_1$ thanks to inequalities (4.2).
We thus define $R(n)$ as a Markov chain with transition matrix $Q(i, j)$ such that, at each time step :

- if $1 \le r < \frac{N}{2}$, the probability chosen for $Q(r, r-1)$ is the smallest possible for the dynamics. We thus choose the lower bound for $p(n, r, r-1)$ computed in equation (4.2) : $Q(r, r-1) = \frac{r}{N}\sigma_{N-r}$. On the contrary, the probability chosen for $Q(r, r+1)$ is the biggest possible given by equation (4.2) : $Q(r, r+1) = \frac{N-r}{N}\sigma'_r$. To sum up,

$$Q(r, r-1) = \frac{r}{N}\,\sigma_{N-r}, \quad Q(r, r+1) = \frac{N-r}{N}\,\sigma'_r\,, \tag{4.3}$$

- Similarly, if $N-1 \ge r > \frac{N}{2}$, we define

$$Q(r, r-1) = \frac{r}{N}\,\sigma'_{N-r}, \quad Q(r, r+1) = \frac{N-r}{N}\,\sigma_r\,, \tag{4.4}$$

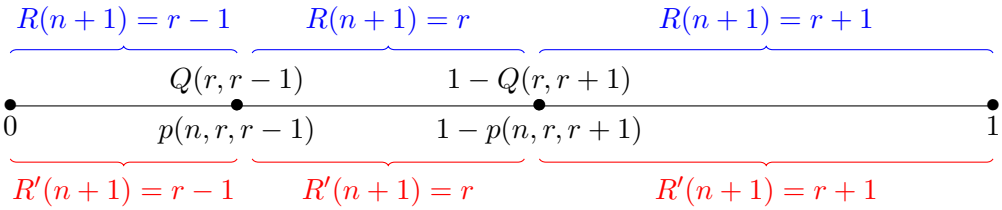
- Finally, if $r = \frac{N}{2}$ (in the case where N is even), we decide to choose the lowest possible probabilities to gain or loose an agent :

$$Q(r, r-1) = \frac{r}{N}\,\sigma_{N-r}\ ,\ Q(r, r+1) = \frac{N-r}{N}\,\sigma_r\ , \tag{4.5}$$

In any case, $Q(r,r) = 1 - Q(r,r-1) - Q(r,r+1)$, and $R(n)$ is stopped as soon as it reaches 0 or N. It is worth noticing that here again, the symmetry (4.9) is satisfied.

Technically, the coupling is performed as follows. Let U_n be a sequence of i.i.d. random variables uniformly distributed on $[0,1]$.

- For $N > r \geq (N+1)/2$, conditionally to $R(n) = r$, $R(n+1) = r+1$ (resp. $r-1$, resp. r) if $U_n \geq 1 - Q(r, r+1)$ (resp. $Q(r,r-1) \geq U_n$, resp. $1 - Q(r,r+1) \geq U_n \geq Q(r,r-1)$); conditionally to $R'(n) = r$, $R'(n+1) = r+1$ if $U_n \geq 1 - p(n,r,r+1)$, $R'(n) = r$ if $p(n,r,r-1) \leq U_n \leq 1 - p(n,r,r+1)$ and $R'(n+1) = r-1$ if $U_n \leq p(n,r,r-1)$ (see figure below).



- For $1 \leq r < N/2$, conditionally to $R(n) = r$, $R(n+1) = r-1$ (resp. r, resp. $r+1$) if $U_n \geq 1 - Q(r,r-1)$ (resp. $Q(r,r+1) \leq U_n \leq 1 - Q(r,r-1)$, resp. $U_n \leq Q(r,r+1)$); conditionally to $R'(n) = r$, $R'(n+1) = r-1$ if $U_n \geq 1 - p(n,r,r-1)$, $R'(n) = r$ if $p(n,r,r+1) \leq U_n \leq 1 - p(n,r,r-1)$ and $R'(n+1) = r+1$ if $U_n \leq p(n,r,r+1)$.
- For $r = N/2$, conditionally to $R(n) = r$, $R(n+1) = r-1$ (resp. r, resp. $r+1$) if $U_n \geq 1 - Q(r,r-1)$ (resp. $Q(r,r+1) \leq U_n \leq 1 - Q(r,r-1)$, resp. $U_n \leq Q(r,r+1)$); conditionally to $R'(n) = r$, $R'(n+1) = r-1$ if $U_n \geq 1 - p(n,r,r-1)$, $R'(n+1) = r$ if $p(n,r,r+1) \leq U_n \leq 1 - p(n,r,r-1)$ and $R'(n+1) = r+1$ if $U_n \leq p(n,r,r+1)$.

We now need to show that $R(n)$ indeed simulates a "worst case" of our dynamics in the sense that it leads to a fixation time that is bigger than the flocking time for $R'(n)$ and then to estimate this fixation time. In that prospect, let us study the distance of $R(n)$ and $R'(n)$ to the absorbing states 0 and N. We thus get interested in the evolution of their distance to $\{0, N\}$: defining $d_{0,N}(r) = r(N-r)$ for any $r \in \{0, \dots N\}$, we consider the quantities $D(n) = d(R(n)) = R(n)(N - R(n))$ and $D'(n) = d(R'(n)) = R'(n)(N - R'(n))$.

Lemma 4.1. *If, for a given $n \geq 0$,the inequality $0 < D'(n) \leq D(n)$ holds, then, almost surely, $D'(n+1) \leq D(n+1)$.*

Proof. Using the symmetry $d_{0,N}(N-r) = d_{0,N}(r)$, we may assume that $R'(n) = r$ with $N - 1 \geq r > N/2$ or $r = N/2$ with N even.

Case $r = N/2$: This case is simple as $r = \frac{N}{2}$ realizes the unique maximum of $d_{0,N}$ over $[0, N]$. Therefore, as by assumption $D'(n) \leq D(n) = N^2/4$, thus $D(n) = N^2/4$ and $R(n) = N/2$. The possible evolutions at time $n+1$ are :

- $R'(n+1) = r = N/2$. In that case, knowing that $R'(n) = r$, the construction gives

$$Q(r, r+1) = \frac{N-r}{N}\sigma_r \leq p(n, r, r+1) \leq U_n \leq 1 - p(n, r, r-1) \leq 1 - \frac{r}{N}\sigma_{N-r} = 1 - Q(r, r-1)$$

  and thus $R(n) = r$, which leads to $D(n+1) = D'(n+1) = N^2/4$.
- $R'(n+1) = r \pm 1$: in both cases, $D'(n+1) = N^2/4 - 1$. In the same way, whatever $R(n+1)$ is, $D(n+1) \in \{N^2/4,\ N^2/4 - 1\}$ which gives $D'(n+1) \leq D(n+1)$.

Case $r > N/2$:

- if $R'(n+1) = r+1$, $D'(n+1) < D'(n)$ as $d_{0,N}$ is decreasing over $[N/2, N]$. Moreover, as $D'(n) \leq D(n)$, then $R(n) \in [N-r, r]$ and so $R(n+1) \in [N-r-1, r+1]$, thus $D(n+1) \geq D'(n+1)$.
- if $R'(n+1) = r$, either $D(n) > D'(n)$ or $D(n) = D'(n)$.
  - If $D(n) > D'(n)$, then $R(n) \in]N-r, r[$ and so $R(n+1) \in [N-r, r]$ and thus, $D(n+1) \geq D'(n) = D'(n+1)$.
  - If $D(n) = D'(n)$, then either $R(n) = r$ or $R(n) = N - r$. Let us investigate the case $R(n) = r$, the other case can be treated the same way thanks to the constructon and its symmetries. As $R'(n) = r = R'(n+1)$ by assumption, the construction of the coupling imposes

$$p(n, r, r-1) \ \leq \ u_n \ \leq \ 1 - p(n, r, r+1)$$

    and inequalities (4.2) together with definition (4.4) yields

$$p(n, r, r-1) \leq u_n \leq 1 - p(n, r, r+1) \leq 1 - \frac{N-r}{N}\sigma_r = 1 - Q(r, r+1),$$

    and so, as $R(n) = r$, $R(n+1)$ might be r or $r-1$. In both cases, $D(n+1) \geq D(n) = D'(n) = D'(n+1)$.
- If $R'(n+1) = r-1$, once again, as $D(n) \geq D'(n)$ and $R'(n) = r$, then, $R(n) \in [N-r, r]$. Let us investigate different cases : either $R(n) \in]N-r+1, r-1[$, $D(n) = D'(n) = r(N-r)$ or $D'(n) = (r-1)(N-r+1)$.
  - If $R(n) \in]N-r+1, r-1[$, then, whatever $R(n+1)$ is, $R(n+1) \in [N-r+1, r-1]$ and so, thanks to $d_{0,N}$ behavior, $D(n+1) \geq d_{0,N}(r-1) = D'(n+1)$.
  - If $D(n) = r(N-r)$, *i.e* $R(n) = r$ or $R(n) = N - r$. As previously seen, both cases can be treated with the same arguments thanks to the symmetries at play. In the case where $R(n) = r$, then, by construction of the coupling, as $R'(n) = r$ and $R'(n+1) = r-1$, and inequalities (4.2) together with definition (4.4)

$$u_n \leq p(n, r, r-1) \leq \frac{r}{N}\sigma'_{N-r} = Q(r, r-1).$$

    Therefore, $R(n+1) = r-1$, and thus $D(n+1) = D'(n+1)$.
  - If $D(n) = (r-1)(N-r+1)$, (*i.e* $R(n) = r-1$ or $R(n) = N-r+1$), again, both cases can be treated with the same arguments thanks to the symmetry. Let us assume that $R(n) = r-1$. It might be possible that $r - 1 = N/2$ when N is even, which changes a little bit the following inequalities, but the arguments are the same than if $r - 1 > N/2$, which we assume now. Then, the possible values for $R(n)$ are $R(n) = r-2,\ r-1$ or r. As $r - 2 \geq N/2, r - 1 > N/2$ and $d_{0,N}$ is decreasing over $[N/2, N]$, if $R(n+1) \in \{r-2,\ r-1\}$, then $D(n+1) \geq D'(n+1) = d_{0,N}(r-1)$. The only case where $D(n+1)$ could be greater ou equal

to $D'(n+1)$ is the case when $R(n+1)=r$. Let us now prove that it is almost surely impossible.
Once again, the construction of the coupling gives $u_n \leq p(n,r,r-1)$. Moreover, given $R(n)=r-1$, $R(n+1)=r$ if and only if $u_n \geq 1-Q(r-1,r)$. The latter is of course impossible if

$$p(n,r,r-1) \leq 1-Q(r-1,r)\,. \tag{4.6}$$

Moreover, $Q(r-1,r)=\frac{N-r+1}{N}\sigma_{r-1}$ and $p(n,r,r-1)\leq \frac{r}{N}\sigma'_{N-r}$, thus, condition (4.6) is fulfilled provided

$$r\sigma'_{N-r} \leq N-(N-r+1)\sigma_{r-1}$$

which is straight-forward as $r \in]0,N[$ and $\sigma'_{N-r}+\sigma_{r-1}=1$.

This ends the proof of the lemma. □

We now need to be able to estimate the fixation time of the birth and death process R in order to get an upper bound of the flocking time of our process R'. In the next subsection, we provide a computation of the fixation time of any birth and death process, and, in the last one, we apply it to R and finally state an upper bound for the flocking time of R'.

4.2. **Getting an upper-bound for the expected fixation time for Birth and Death processes.**

In this section, let us consider the birth and death process $R'(.)$ with transition $Q(i,i+\varepsilon)$ with $\varepsilon=-1,0,1$ described in the previous section. The fixation time (denoted by F_∞) is the first time when R' equals 0 or N. The expectation of F_∞, starting from r :

$$m(r) := \mathbb{E}(F_\infty | R'(0)=r)\,,$$

can in that prospect be explicitly computed as it solves the followig system of equations for $0<r<N$:

$$m(r) = 1+Q(r,r+1)\,m(r+1)+Q(r,r)\,m(r)+Q(r,r-1)\,m(r-1)\,, \tag{4.7}$$

with boundary conditions $m(0)=m(N)=0$.
Using the fact that $Q(r,r)=1-Q(r,r-1)-Q(r,r+1)$, it follows that $m(r)$ satisfies the following recursion, for $r=1\dots N-1$:

$$m(r+1)-m(r) = \frac{Q(r,r-1)}{Q(r,r+1)}\,(m(r)-m(r-1)) - \frac{1}{Q(r,r+1)}. \tag{4.8}$$

In our case, according to equations (3.1), the transitions are symmetric around $N/2$:

$$Q(r,r') = Q(N-r,N-r')\,, \tag{4.9}$$

which implies that $m(r)=m(N-r)$ by a simple recursion and equation (4.8). It is therefore more natural to compute $m(r)$ starting with $m(N/2)$ or $m\left(\frac{N+1}{2}\right)$.

Case where N is even : We have, using recursion (4.8)

$$m((N/2)+1)-m(N/2) = \frac{Q(N/2,(N/2)-1)}{Q(N/2,(N/2)+1)}\big(m(N/2)-m((N/2)-1)\big) - \frac{1}{Q(N/2,(N/2)+1)}$$

so that using, $m((N/2)+1) = m((N/2)-1)$ and $Q(N/2,(N/2)+1) = Q(N/2,(N/2)-1)$, we get

$$m((N/2)+1) - m(N/2) = \frac{-1}{2\,Q(N/2,(N/2)+1)} < 0. \tag{4.10}$$

Similarly,

$$m((N/2)+2) - m((N/2)+1) = \frac{Q((N/2)+1,N/2)}{Q((N/2)+1,(N/2)+2)}\big(m((N/2)+1) - m(N/2)\big) - \frac{1}{Q((N/2)+1,(N/2)+2)} < 0,$$

and by induction it is easily seen that m is decreasing on $\{N/2,(N/2)+1,...,N\}$. By symmetry, m is therefore maximal at $N/2$ and $m(N/2)$ furnishes a bound for m over $\{1,\dots N\}$.

Case where N is odd : $N = 2p+1$ Equation (4.8) becomes :

$$m(p+2) - m(p+1) \;=\; \frac{Q(p+1,p)}{Q(p+1,p+2)}\,(m(p+1) - m(p)) - \frac{1}{Q(p+1,p+2)},$$

which, using symmetry $m(p+1) = m(p) = m((N+1/2)-1)$ leads to

$$m\left(\frac{N+1}{2}+1\right) - m\left(\frac{N+1}{2}\right) \;=\; m(p+2) - m(p+1) \;=\; \frac{-1}{Q(p+1,p+2)} < 0. \tag{4.11}$$

As before it easily follows that $m(r)$ is decreasing on $\left\{\frac{N+1}{2}, \frac{N+1}{2}+1, ..., N\right\}$ too, which again furnishes the maximality of $m\left(\frac{N+1}{2}\right)$ over $\{1,\dots N\}$ in that case.

Thus, let us now get some upper bound for $m(N/2)$ or $m\left(\frac{N+1}{2}\right)$. Since the expressions are similar we only give the formulas for N even. After the dust settles one gets

$$\begin{aligned} m(N/2) \;=\; & \frac{1}{2\,Q(N/2,(N/2)+1)}\left(1 + \sum_{l=1}^{(N/2)-1}\prod_{j=1}^{l}\frac{Q((N/2)+j,(N/2)+j-1)}{Q((N/2)+j,(N/2)+j+1)}\right) \\ & + \sum_{i=1}^{(N/2)-1}\frac{1}{Q((N/2)+i,(N/2)+i+1)}\times \\ & \left(1 + \sum_{l=1}^{(N/2)-1-i}\prod_{j=1}^{l}\frac{Q((N/2)+i+j,(N/2)+i+j-1)}{Q((N/2)+i+j,(N/2)+i+j+1)}\right). \end{aligned} \tag{4.12}$$

4.3. **Expected flocking time in the worst case.**

It follows from the lemma that, if $N > R(0) = R'(0) > 0$, the flocking time S_∞ (equal to the hitting time of 0 by $D'(n)$) is bounded by the hitting time F_∞ of $\{0,N\}$ by $R(.)$ (fixation time for the birth and death process). Thanks to Equation (4.12), we can state an upper bound for T_∞.

Theorem 4.2. *Assume that k is non increasing, strictly positive and $\#\bar{V}(0) = m$. Then there exists a constant C such that for any $N \geq 4$,*

$$\mathbb{E}(T_\infty) \leq \frac{CN^2 \log_2(m)}{\lambda(N)} \frac{\sum_{l=1}^{N-1} k(l/N-1)}{\sum_{l=[N+1/2]}^{N-1} k(l/N-1)} \ (1+B_N)$$

where

$$B_N = \prod_{r=[N/2]}^{N-1} \frac{r \sum_{z=1}^{N-r} k(z/N-1)}{(N-r) \sum_{z=N-r}^{N-1} k(z/N-1)} .$$

Proof. Since k is non-increasing, suppose $r > N/2$, then, by definition,

$$\begin{cases} Q(r,r-1) = \frac{r}{N} \frac{\sum_{\ell=1}^{N-r} k(\ell/N-1)}{\sum_{\ell=1}^{N-1} k(\ell/N-1)} \geq r(N-r)k\left(\frac{N-r}{N-1}\right) \\ \\ Q(r,r+1) = \frac{N-r}{N} \frac{\sum_{\ell=N-r}^{N-1} k(\ell/N-1)}{\sum_{\ell=1}^{N-1} k(\ell/N-1)} \leq r(N-r)k\left(\frac{N-r}{N-1}\right). \end{cases}$$

and thus

$$\frac{Q(r,r-1)}{Q(r,r+1)} \geq 1 . \tag{4.13}$$

Note that the equality holds for $r = N/2$ and for every $1 \leq r \leq N-1$ when k is constant. It follows that there exists a constant C such that for $N \geq 4$,

$$m(N/2) \leq CN\, A_N \frac{\sum_{j=1}^{N-1} k(j/N-1)}{\sum_{j=N/2}^{N-1} k(j/N-1)}$$

with

$$A_N = 1 + \sum_{l=1}^{N/2-1} \prod_{j=1}^{l} \frac{(j+(N/2)) \sum_{z=1}^{N/2-j} k(z/N-1)}{(N/2-j) \sum_{z=N/2-j}^{N-1} k(z/N-1)} .$$

Note that for k constant we get the good order N^2 for $m(N/2)$. Using a rough bound for A_N we have thus obtained the desired bound. □

Of course

$$\frac{\sum_{z=1}^{N-r} k(z/N-1)}{\sum_{z=N-r}^{N-1} k(z/N-1)} \leq \frac{1}{r\,\beta_N} ,$$

and in the worse case i.e when $\frac{k(z/N-1)}{\sum_{z=1}^{N-1} k(z/N-1)} = \beta_N$ for all $z \geq 2$ and $N\beta_N \ll 1$, this rough bound is the best possible. We thus have

Corollary 4.3. *Assume that k is non increasing and that β_N given by (2.1) is positive. Let $\#\bar{V}(0) = m$. Then for large N,*

$$\mathbb{E}(T_\infty) = O\left(\log_2(m) \frac{\int_0^1 k(u)du}{\int_{1/2}^1 k(u)du} \frac{N^2}{\lambda(N)} \left(\frac{e}{2N\beta_N}\right)^{N/2} \right) .$$

Notice that the previous result is better (for large N) than the one obtained in Proposition 2.5. Indeed $\beta_N N \leq 1$, and up to a power of N we replace $(e/\beta_N N)^N$ by its square root. However we still have an exponential in N bound.

We can modify the previous bound noticing that for $r \geq N/2$,

$$\frac{\sum_{z=1}^{N-r} k(z/N-1)}{\sum_{z=N-r}^{N-1} k(z/N-1)} \leq \frac{\sum_{z=1}^{N/2} k(z/N-1)}{\sum_{z=N/2}^{N-1} k(z/N-1)} .$$

Using Stirling formula again, we obtain for $k > 0$ and then any k by taking limits

Corollary 4.4. *Assume that k is non increasing and that $\int_{1/2}^1 k(u)du > 0$. Let $\#\bar{V}(0) = m$. Then for large N,*

$$\mathbb{E}(T_\infty) = O\left(\log_2(m) \frac{\int_0^1 k(u)du}{\int_{1/2}^1 k(u)du} \frac{N^{3/2}}{\lambda(N)} \left(4 \frac{\int_0^{1/2} k(u)du}{\int_{1/2}^1 k(u)du} \right)^{N/2} \right) .$$

In particular if k is non increasing and $\int_{1/2}^1 k(u)du > 0$, the flocking time is almost surely finite.

We have thus obtained a more general sufficient condition for flocking. We shall get another proof for this condition as a consequence of the general study of the behaviour of the system in the next subsection 5.1.

Remark 4.5. In [DM09] another explicit expression for the expectation $\mathbb{E}(T'_\infty)$ corresponding to the continuous time birth and death process, is obtained. It is equal to the sum of the inverse of the elements of the spectrum, hence it is bounded by N times the inverse of the spectral gap. Of course it remains to estimate this spectral gap. ♢

4.4. **Numerical simulations.**

In that section we aim at illustrating the bounds proved in Corollaries 4.3 and 4.4. As they both only give an upper bound for the mean flocking time, our goal is to test if these bounds can be "optimal" or not. We therefore have carried out the following experiments, in one or two dimensions :

One dimensional case : Consider N particles with positions randomly chosen in $[-\varepsilon, \varepsilon]$ (with uniform probability), for $N = 50,\ 100,\ 200$. Their velocities are uniformly chosen in $\{-1, 1\}$.

- Simulating a Poisson process with rate $\lambda(N) = N$ we define a sequence of jump times T_n.
- Between two jump times the particles follow straight paths with constant velocity.
- At each jump time, one chooses a particle i uniformly with probability $\frac{1}{N}$, compute the ranks (by ordering $\|x_i - x_j\|$ for all $j \neq i$) and therefore the probabilities $\pi_{i,j}$ for all $j \neq i$. One chooses a particle $j \neq i$ with probability $\pi_{i,j}$ and changes v_i into v_j.
  In the following simulations, we have tested different kernels k . Note that, for these functions, we don't have $\int_0^1 k(x)dx = 1$ which is not really necessary in this work

(as said in the introduction) the important point being the normalization of the probabilities π_{ij}.
  * $k(r) = 1 - r$ (note that in this case $k(1) = 0$)
  * $k(r) = 1.01 - r$ (a strictly positive approximation of the preceeding case)
  * $k(r) = \frac{8}{(2+2r)^3}$
  * $k(r) = N$ if $r \in [0, \frac{1}{N-1}[$ and $k(r) = \frac{1}{N}$ if $r \in [\frac{1}{N-1}, 1]$
- If the particles have not flocked (*i.e* all the velocities are not identical yet), one computes a free-flight until the following jump time.

For a given initial data, we carry out 500 experiments and compute the mean flocking time in each case. The results are given in Table 2

	$k(r) = 1.01 - r$	$k(r) = 1 - r$	$k(r) = \frac{8}{(2+2r)^3}$	$k(r) \in \{N, \frac{1}{N}\}$
$N = 50$	20.5	18.04	27.32	623.63
$N = 100$	38.52	36.27	50.35	2240
$N = 200$	82.5	79.7	85.41	9721.8

TABLE 2. Mean flocking time for different kernels k depending on N.

What can be deduced from Table 2 is first that, for some k the flocking time seems to be basically linear in N, as previously when we took $k = 1$ in section 3.2, (or even better ?) for the first functions. An interesting point is that, in the case where

$$k(r) = N \text{ if } r \in [0, \frac{1}{N-1}[\text{ and } k(r) = \frac{1}{N} \text{ if } r \in [\frac{1}{N-1}, 1]$$

the flocking time grows very fast with N and requires a huge number of iteration of the process before flocking. It was so long for $N = 100$ that we decided to carry out only 120 experiments and only 25 in the case $N = 200$. This can be understood using Corollary 4.4 which states that in the general case, $\mathbb{E}(T_\infty)$ may grow up to exponentially with N.

Two dimensional case : Consider $N = 12p$ particles, p being a positive integer. We carry out two different kind of simulations :

- half of them (*i.e* $6p$) are initially equally distributed at the vertices of a regular hexagon (p particles are located at each vertex) and the other $6p$ are located at the center of the hexagon
- all the particles are equally distributed at the vertices of the hexagone (none of them is in the center of the hexagone anymore).

The particles located at the vertices of the hexagon have their velocities directed by the line joining the vertex and the center of the hexagon and with norm 1. The particles located at the center of the hexagon have $\vec{0}$ velocity (see top left of Figure 5)

With this initial data, we test the exact same functions k as for the one dimensional case, and an additional one :

$$k(r) = N \text{ if } r \in [0, \frac{1}{N-1}[, \ k(r) = \frac{1}{N} \text{ if } r \in [\frac{1}{N-1}, \frac{N-10}{N-1}[\text{ and 0 else.}$$

Note that in the case where $N = 12$, we only cancel the four last rank to guarantee almost sure flocking. Of course, in this particular example k is compactly supported (and therefore

does not fall into this section's scope ($k > 0$), but it will give us clues for the next section where we investigate cases where k vanishes for some ranks). We carry out 100 experiments (25 in the two last case that are much more costly in time) with these initial data (agents in the center or no agents in the center) and then compute the mean flocking time for these experiments. The results are presented in Table 3 and Table 4. In that case, as in the previous one the mean flocking time seems to grow linearly with N for the linear or cubic decreasing k showing that our bound is presumably disastrous. The cases where k is strong for the nearest neighbor and weak for the others takes much more time to flock. It is also interesting to see that the case where the agents are located at the vertices and the initial data equally distributed between them takes much more time to flock than the situation where, from the start, half of the agents already have the same velocity (which is, therefore much more likely to be chosen to align on).

	$k(r) = 1 - r$	$k(r) = \frac{8}{(2+2r)^3}$	$k(r) \in \{N, \frac{1}{N}\}$	k compact. supp.
$N = 12$	4.6	6.54	13.82	16.43
$N = 24$	6.80	11.93	120.12	129.93
$N = 36$	10.26	16.48	364.6	227.67
$N = 72$	14.45	22.74	1505.47	936.08
$N = 108$	15.04	26	2846.52	2010.76

TABLE 3. Mean flocking time for different kernels k depending on N in the 2D hexagonal case with agents at the center of the hexagone.

	$k(r) = 1 - r$	$k(r) = \frac{8}{(2+2r)^3}$	$k(r) \in \{N, \frac{1}{N}\}$	k compact. supp.
$N = 12$	8.08	10.19	57.24	115.07
$N = 24$	16.19	20.32	182.56	454.73
$N = 36$	24.62	26.93	397.67	557.97
$N = 72$	53.45	54.34	1929.10	1750.27
$N = 108$	71.89	90.15	3768.14	3414.1

TABLE 4. Mean flocking time for different kernels k depending on N in the 2D hexagonal case without agents at the center.

What can also be said on these experiments is the behavior of the probability to flock on the different velocities. In the case where half of the agents are initially located in the center with velocity $\vec{0}$, we only care about the probability to flock on velocity $\vec{0}$ (which means that the particles stop). Again, things are very different in the case where k is strong for the nearest neighbors and weak for the other ones. Indeed, in the cases where $k(r) = 1 - r$ or $k(r) = \frac{8}{(2+2r)^3}$, the probability to flock on $\vec{0}$ is more than 90%, whatever N is, whereas in the last case, this probability varies between 54% and 69%. In the case where no agent is in the center, the probability to flock on a velocity is basically the same for all the velocities which is rather reasonable as all the velocities play the same role in that example.

On the following figures, we represent the evolution along time of the trajectories of each 12 agents. In that prospect, at each jump time the position (in the plane) of each particle is drawn with a circle until flocking time, for different kernels k.

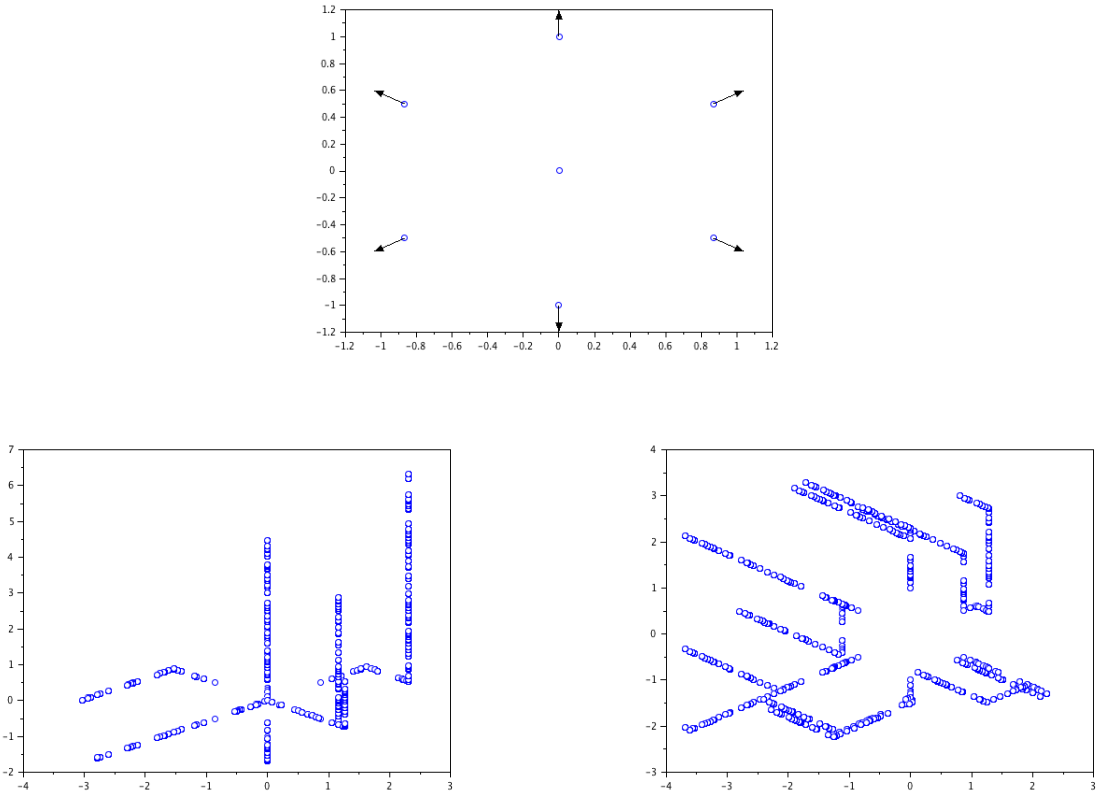


FIGURE 2. Different trajectories in the case $N = 12$. Top : initial data, Bottom $k(r) = 1 - r$ left : agents at the center, $V_f = (0, 0)$, $T_f = 8.15$, right no agent at the center : $T_f = 3.26$, $V_f = (-\cos(\pi/6), 1/2)$

5. The general case: sub-flocking.

5.1. **What can be proved...**

In this subsection we do no more assume that k is everywhere positive. For simplicity and because it is very natural (the closest particles are, the strongest their interaction) we shall nevertheless assume that k is non-increasing. Hence we assume that $k(s) > 0$ for $s \in [0, a[$ and $k(s) = 0$ on $[a, 1]$ and state the following result

Theorem 5.1. *Assume that k is non increasing and that $k(s) > 0$ for $s \in [0, a[$ and $k(s) = 0$ on $[a, 1]$. Then for almost all ω there exists an $n(\omega)$ such that for $t \geq T_{n(\omega)}$,*

$$v(t) = v(T_{n(\omega)}) = v(\infty)$$

i.e. the process $v(t)$ is constant after time $T_{n(\omega)}$.
Furthermore if $\frac{j}{N-1} < a \leq \frac{j+1}{N-1}$, for each $i = 1, ..., N$, $\#\{l\ ;\ v_l(\infty) = v_i(\infty)\} \geq j + 1$.

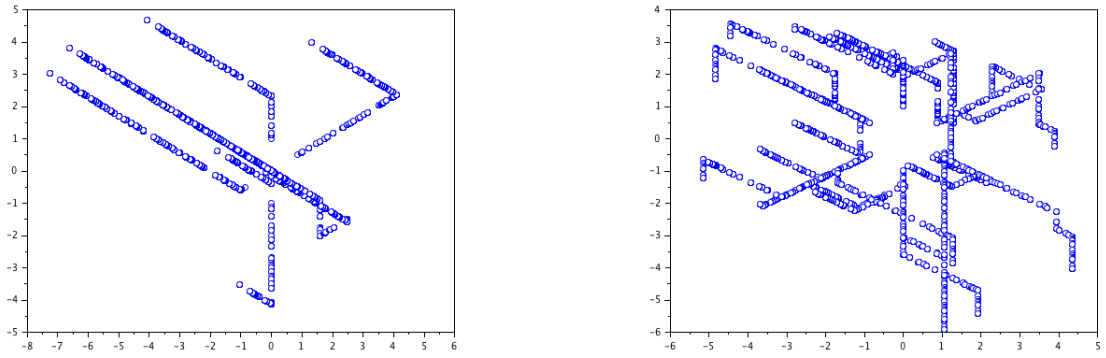


FIGURE 3. Different trajectories in the case $N = 12,\ k$ cubic, left : agents at the center, $T_f = 9.8,\ V_f = (-\cos(\pi/6), 1/2)$, right : $T_f = 5.53,\ V_f = (0, -1)$

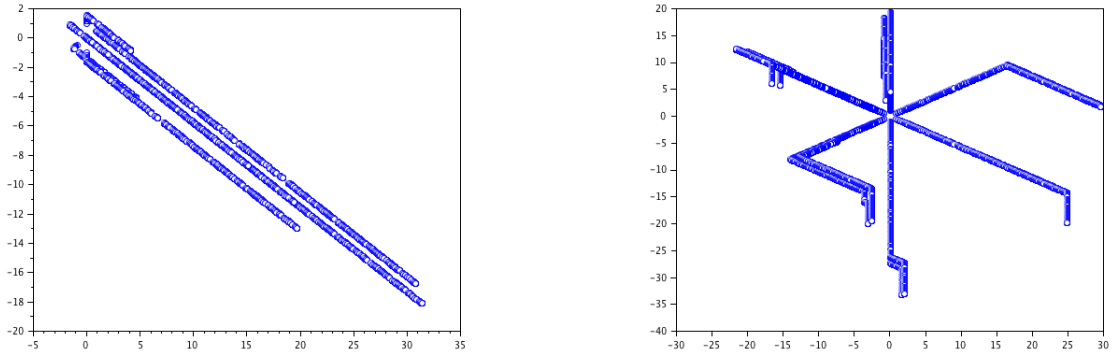


FIGURE 4. Different trajectories in the case $N = 12,\ k \in \{1, N\}$, left : agents at the center, $T_f = 35.25,\ V_f = (\cos(\pi/6), -1/2)$, right : $T_f = 33.28,\ V_f = (0, -1)$

Remark 5.2. It is worth noticing that the final set of velocities $\overline{V}(\infty)$ depends on ω. The theorem states that, almost surely, velocities are all kept fixed after some random time. Moreover, as soon as the final distribution of speeds is reached, any cluster of velocity in $\bar{V}(\infty)$ contains at least $j + 1$ agents. ♢

Proof. We choose j such that $k(j/(N-1)) > k(j + 1/(N-1)) = 0$. Hence a given particle chosen at a jump time aligns on the velocity of one of its j nearest neighbors. At each time t we can divide the set of velocities $\overline{V}(t)$ into $\bar{c}(t) = \#\overline{V}(t)$ clusters, each cluster gathering all particles with the same velocity. Along the dynamics, $\bar{c}(t)$ is by construction a non increasing sequence of integers, hence it is stationary : there exists $t_\infty > 0$ such that for $t \geq t_\infty,\ \bar{c}(t) = \bar{c}(\infty)$. We first claim

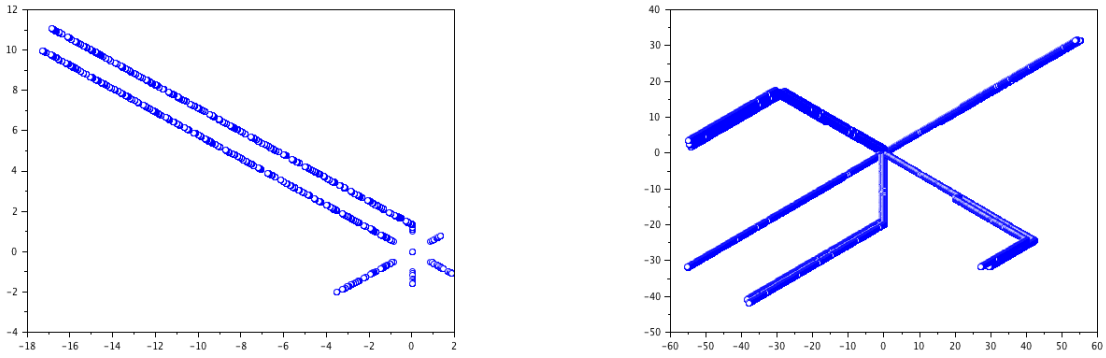


FIGURE 5. Different trajectories in the case $N = 12, k$ compactly supported, right : $T_f = 19.8,\ V_f = (0, 0)$, left : $T_f = 62.63,\ V_f = (-\cos(\pi/6), -1/2)$

Lemma 5.3. *Let k be non-increasing and such that $k(j/N - 1) > k(j + 1/N - 1) = 0$. Then there exists a time t'_∞ such that for $t \geq t'_\infty$ all clusters have at least $j + 1$ elements.*

Proof. of the lemma. Let $t > t_\infty$, $C(t)$ be one cluster and assume that $\#C(t) = \ell \leq j$. Then any particle in C has its neighbor(s) with rank j that does not belong to C. Let us show that, in that case, it is possible to "empty" $C(t)$ by choosing at each jump time an agent of $C(t)$ and force him to align with its neighbor with rank j (thus outside of C) and thus "leave" $C(t)$.
Introduce event $\mathcal{G} = \{$ At each of the ℓ first jumping times T_k following $t > t_\infty$, one chooses an agent of C to align with its neighbor with rank j $\}$. We have

$$\mathbb{P}(\mathcal{G}) \;=\; \frac{(k(j/N-1))^\ell}{N^\ell\,(\sum_{i=1}^{N-1} k(i/N-1))^\ell} \;=\; p > 0\,.$$

Hence, using the Markov property again, if there exists an infinite number of random times $t > t_\infty$ such that $\#C(t) = l \leq j$, $\mathcal{G}$ will be satisfied almost surely after time t_∞. But when $\mathcal{G}$ is satisfied, the cluster C disappears, which is in contradiction with the fact that after time t_∞ the number of clusters is fixed. Since the possible number of ℓ's and of clusters is finite, we get the lemma. □

From now on we assume that $t > t'_\infty$ that is for all cluster $C(t)$, $\#C(t) \geq j + 1$. Consider $c_{min}(t) = \#C_{min}(t)$ where $C_{min}(t)$ is the smallest cluster at time t. Then, $(c_{min}(T_n \vee t'_\infty))_n$ is a bounded sequence taking its values in $\mathbb{N}*$. Let $c* = \liminf_n\, c_{min}(T_n \vee t'_\infty)$.
First of all, according to Lemma 5.3, $c* \geq j+1$. Second, there exists at most a finite number of n such that $c_{min}(T_n \vee t'_\infty) < c*$. We shall look only at t larger than the maximum of these T_n say $t"_\infty$, so that we may assume that $t > t"_\infty$, *i.e.* $c_{min}(t) \geq c*$.
But since $c*$ is a limit point there exists an n_0 such that $c_{min}(T_{n_0}) = c*$. At time T_{n_0+1} the size of the cluster is either $c*$ or $c* + 1$. If the latter never occurs, the cluster C does no more change, hence it is stationary, and we may repeat the argument for the new system of particles where we have erased all particles of this cluster.

Hence it remains to study the case where $c_{min}(T_{n_0}) = c*$ and $c_{min}(T_{n_0+1}) = c* + 1$. For simplicity we may assume that the indices of these particles are $1, ..., c* + 1$. Take some $l > c* + 1$. As long as these particles have no jump, $x_i(t) = tv_i + x_i(T_{n_0+1})$ so that for $1 \leq i, m \leq c* + 1$ the distance $|x_i(t) - x_m(t)|$ is unchanged equal to $|x_i(T_{n_0+1}) - x_m(T_{n_0+1})|$, while $|x_i(t) - x_l(t)|$ grows like $t|v_i - v_l|$ for large t. It follows that, as long as these particles have no jump, for $t \geq A$, A being large enough, $|x_i(t) - x_l(t)| > |x_i(t) - x_m(t)|$ for all $i = 1, ..., c* + 1$ and all $l > c* + 1$. Since $c*$ is the lim inf, the first time T, $T > T_{n_0+1}$ such that $x(T) = c*$ is almost surely finite. T can be chosen as a jump time of the Poisson process, and so $\mathbb{P}(T - T_{n_0+1} > A) > 0$. According to what precedes, on the set $\{T - T_{n_0+1} > A\}$ the size of the cluster cannot decrease, yielding a contradiction.
The proof is completed. □

We immediately deduce from the second part of the Theorem, the following flocking result

Corollary 5.4. *If each particle only interacts with its j nearest neighbors, but $j \leq N \leq 2j+1$, then the system is almost surely flocking.*

We thus recover the second part of Corollary 4.4.

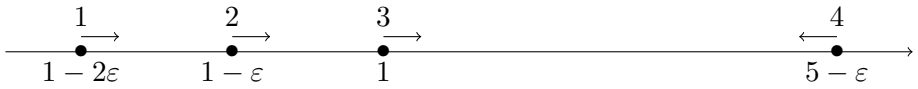
5.2. **Illuminating examples and simulations in one dimension, nearest neighbor case.**

In the spirit of example 2.3, let us describe some situations. For simplicity we assume that $d = 1$. Consider the case where each particle may only interact with its nearest neighbor (resp. j nearest neighbors). Then if for each i, $\#\{l, v_l(0) = v_i(0)\} \geq 2$ (resp. $j + 1$), the probability for these initial velocities to never change is positive. Indeed if $t > T$ for some T large enough, whatever the initial positions, the potential positions $x_i(t) = x_i(0) + t\, v_i(0)$ will be organized in such a way that nearest neighbors always have the same velocity. Hence, since the probability for always choosing a partner with the same velocity before time T is positive, we get the result. Of course if the initial velocities are not all the same, this means that there is no flocking.
At the same time, it is also possible to flock with a positive probability *i.e* $0 < \mathbb{P}(\text{flock}) < 1$.

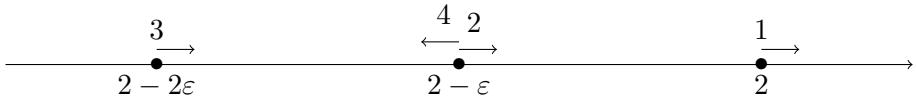
Example 5.5. In this example, we take

$$X_1(0) = (1 - 2\varepsilon,\ 1 - \varepsilon,\ 1,\ 5 - \varepsilon), \quad \text{and} \quad V_1(0) = (1, 1, 1, -1).$$

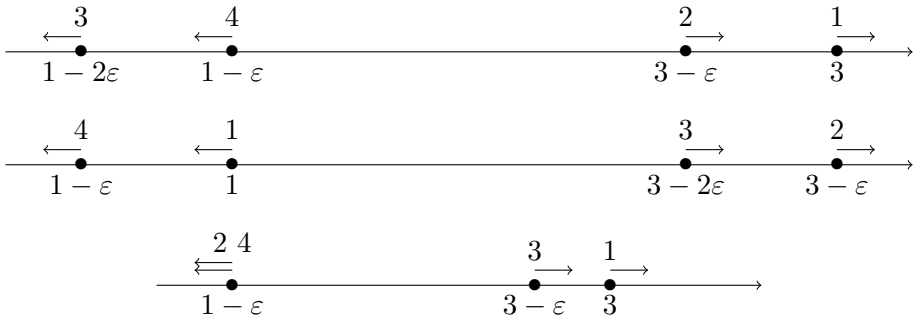
and discuss the difference between deterministic and random jump times following a Poisson process of parameter $\lambda(N) = N$ when the particles only interact with their nearest neighbor. In the case where time steps are deterministic, it can be shown that there is flocking with probability $\frac{5}{8}$.



Indeed, as soon as particle 4 is chosen, it aligns with a particle of velocity 1 and there is flocking. If particle 4 is not chosen until time T_3, at time T_3, the situation is the following :



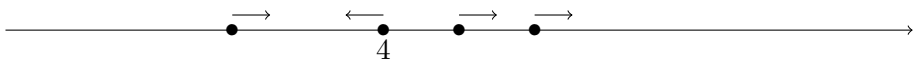
Now, if particle 4 is chosen, it has to align with 2 and the system flocks. If particle 3 is chosen (and the same occurs with particle 1), it can stay with velocity 1 or align with 4 and get velocity −1. If particle 2 is chosen, it aligns with 4 and takes velocity −1. The different possibilities at time T_4 are therefore the following :



which all lead to sub-flocking of to groups of 2 particles moving away from each other (and therefore, thanks to the "nearest neighbor" interaction, not interacting with each other anymore).

What is highlighted here is the importance of the choice of deterministic/stochastic time steps. Indeed, if the time steps become random, then we cannot foresee where particle 4 in our example will arrive and therefore, we cannot draw the different possibilities anymore. If particle 4 is not chosen and crosses particles 1, 2 and 3 without arriving between them, i.e, if, at one time step 4 has a coordinate that is greater than the others and, at the following step, its coordinate has become much smaller than all the others, then, as soon as 4 is chosen, it has to align with the others. If the particle with the smallest coordinate between 1, 2 and 3 is chosen (say 1 for example), its nearest neighbor will be 2 or 3 and the situation won't change until 4 is chosen. Such a case necessarily leads to flocking. Let's make it simple with two drawings :

will necessarily lead to flocking whereas a situation like



may lead, or not, to flocking.

Numerical experiments on example 5.5 : We carry out the same experiments as in the previous example and get the following results :

- If the time steps are deterministic ($T_n = n$), then the estimated flocking probability is 0.622 (theoretically $\frac{5}{8} = 0.625$).
- If the time steps are random with $\lambda(N) = N$, then we get the following results

	$\varepsilon = 0.01$	$\varepsilon = 0.1$	$\varepsilon = 0.3$	$\varepsilon = 0.6$
flocking probability	0.996	0.974	0.956	0.951

Example 5.6. It is also possible to slightly modify the initial data of example 5.5 in order to have almost surely flocking with deterministic time steps and different possibilities with random time steps :

$$X_0(0) = (-2\varepsilon,\ -\varepsilon,\ 0,\ 4+5\varepsilon), \quad \text{and } V_0(0) = (1,1,1,-1).$$

In the case of deterministic jump times (say for example that $T_n = n$), it is possible to show that :

- if $\varepsilon < \frac{1}{4}$, there is flocking on velocity 1 with probability 1
- if $\varepsilon \geq \frac{1}{4}$, there is flocking with positive probability (if $\varepsilon = \frac{1}{4}$ the flocking probability can be explicitly computed and is $\frac{119}{128} \simeq 0.93$)

Numerical experiments on example 5.6 : In order to estimate the flocking probability, starting with the same initial data, we run the algorithm and decide if there is flocking or not. Then, we count the number of experiments needed until the first one flocks and we do this 100 times and compute the mean number of experiments needed to flock. Doing so, we estimate the expectation the first success in a Bernoulli scheme ($X = 0$ if flock, 1 otherwise), which is $\frac{1}{p}$ if p denotes the probability of non-flocking.

Carrying these numerical experiments leads to the following results

- If the time steps are deterministic ($T_n = n$) and $\varepsilon < \frac{1}{4}$, then the system flocks with probability 1.
- If the time steps are deterministic ($T_n = n$) then the system flocks with the following probability depending on ε :

	$\varepsilon = 0.25$	$\varepsilon = 0.251$	$\varepsilon = 0.28$	$\varepsilon = 0.29$	$\varepsilon = 0.3$	$\varepsilon = 0.6$
flocking probability	0.9309	0.862	0.856	0.714	0.663	0.626

- If the time steps are random (and $\lambda(N) = N$), we get the following results, and we see that the flocking probability does not really change with ε for random time steps :

	$\varepsilon = 0.01$	$\varepsilon = 0.1$	$\varepsilon = 0.3$	$\varepsilon = 0.6$
flocking probability	0.996	0.98	0.983	0.994

6. Biological relevance of the model. Discussion and conclusion.

The B-D model can be viewed as a discrete version of the model of trail formation proposed in [BDM13]. In the latter work the change of direction in the motion of ants is continuous, the new direction being chosen according to some probability distribution on the circle. The goal of [BDM13] was to explain the mechanism of trail formation, due to pheromone depository, through an individual-based model.

As we said, if it looks like a toy-model, due to the finite set of allowed velocities, the B-D model nevertheless furnishes a very interesting attempt to understand collective motions in the situation of nearest neighbors interactions for at least three reasons.

First the intermediate free flights appear as a convincing issue integrating the fact that an individual cannot continuously adapt its own motion.

Second, it integrates natural randomness both for the times between jumps and for the choice of a new velocity. For some statistical issues of other type of randomness see [NM18]. As we have seen in examples, deterministic or random time jumps can yield very different behavior depending on the initial conditions.

The final reason is not detailed in the present work, but is easy to explain: the model is not too difficult to calibrate. Indeed starting with a set of observations, one can easily perform successively the following three statistical steps:

- Estimate the intensity λ of the Poisson process of jumps via the inverse empirical mean of the number of observed jumps,
- Estimate the number q of interacting neighbors, by looking at the maximal rank of interactions, according to the following procedure. Each time a jump occurs, take the chosen particle i and consider its new velocity v. Among all particles with velocity v, choose the nearest from i and compute its rank. A natural estimator for q is then the maximum of all observed such ranks. Of course it has to be proved that this estimator is consistent.
- eventually, perform a parametric estimation of the interactions strengths, i.e. of k, using the maximum likelihood estimator.

Of course such a program can only be performed with real data and not using comparison with simulated data obtained via different models. Except the seminal [BCC$^+$08], we did not find such a statistical approach in many papers on the subject.

What kind of informations can we deduce from the rigorous study of the B-D model ? In our previous work on Cucker-Smale models with noise ([CDP18]), we observed (at the theoretical level) that noise can help to flock (in some specific cases). Here we observed that stabilization (towards a finite number of subflocks) always occurs and that flocking occurs as soon as each individual interacts with more than one half of the whole population. This result seems to be a little bit disappointing in view of the claims in [BCC$^+$08] or the model in [SB14, KB13], we shall discuss later. However, contrary to the usual metric interactions, it is not needed that an individual interacts with *all* individuals in the population.

This latter assertion has to be tempered. What should be an analogue *metric* model ? Keeping the same procedure for jumps and choice of the jumping particle, we could turn the transition probability $\pi_{i,j}$ into a metric one as follows: choosing this time a non-increasing function k taking values in $\mathbb{R}^+$, define the transition probability

$$\pi_{i,j} = \frac{k(|x_i - x_j|)}{\sum_l k(|x_i - x_l|)}\,.$$

Due to the self normalization, this model (referred to as M-T model below) is not a *purely metric* model, but an intermediate model between topological and metric interactions: the strength of the interaction depends on the *relative distance* between individuals. In the usual context of deterministic dynamical systems *à la Cucker-Smale*, this self normalization

has been introduced and studied in [MT11]. It is particularly adapted to the study of the emergence of subflocks. Up to our knowledge, there is no noisy version of this model.

A quick exploration of the proofs in subsections 4 and 5.1 show that they cannot be adapted to the self normalized metric interaction introduced above. For instance (4.1) and (4.2) do no more furnish universal upper and lower bounds allowing us to compare to the worst case, because even if particles i and j are close (with a rank larger than 2), the transition probability from i to j can be as small as we want. This will easily furnish situations where the system does not flock. For instance, starting with the following initial configuration

we know that the B-D 2 nearest neighbors dynamics flocks, while it is easy to see that the M-T model does not flock as soon as the function k is too rapidly decaying at infinity (as in the standard Cucker-Smale case).

Of course as for the Cucker-Smale model, if we assume that the system is swarming (i.e. all positions stay in a bounded domain, as assumed in many papers), we may mimic the proofs we have done and get the flocking property.

Hence, at the level of these toy models, topological interaction improves on (self normalized) metric interactions and furnishes a more stable flocking property.

In a sense, though very different, the closest model (numerically) studied in the literature is the one in [BFW11]. A similar random selection is done and alignment holds, at deterministic times, with individuals in a given subarea of the plane (a vision cone). The alignment probability is inversely proportional to the distance between particles (introducing thus a metric aspect). Numerical results are compared with the observations in [BCC$^+$08]. Other references are based on continuous mathematical models. In [BCC$^+$08, CCG$^+$12], where the authors are trying to get an idea on how many neighbors have to interact in order to get stable flocks (w.r.t. external perturbations like predator attacks for instance), the underlying mathematical model is (a slight modification of) Vicsek's model continuous, i.e. the set of allowed velocities is the circle.

In all these works, it appears that the number of interacting individuals can be reduced to a *fixed one* between 6 and 10 (also see [SB14, KB13] for this number but for a different dynamics directly expressed on the graph of connectivities).

What the examples we have detailed show is that there is no theoretical evidence that communication with only 6 to 10 nearest neighbors may lead to flock. It does not mean that animals do not respect such a rule, but this rule does not necessarily imply flocking.

In particular, sensitivity to the initial configuration is of key importance and hardly ever studied. Nevertheless for continuous models starting from a not too chaotic initial configuration (in our examples initial velocities have an opposite sign on the line, hence are in a sense too different), one can expect to get such a type of *conditional almost flocking* for large times. This will be the aim of future works.

References

[BCC+08] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proc. Natl. Acad. Sci. U.S.A.*, 105:1232–1237, 2008.

[BCnC11] F. Bolley, J. A. Cañizo, and J. A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 11:2179–2210, 2011.

[BCnC12] F. Bolley, J. A. Cañizo, and J. A. Carrillo. Mean-field limit for the stochastic Vicsek model. *Appl. Math. Letters*, 25(3):339–343, 2012.

[BD16] A. Blanchet and P. Degond. Topological interactions in a Boltzmann-type framework. *J. Stat. Phys.*, 163:41–60, 2016.

[BD17] A. Blanchet and P. Degond. Kinetic models for topological nearest-neighbor interactions. *J. Stat. Phys.*, 169:929–950, 2017.

[BDM13] E. Boissard, P. Degond, and S. Motsch. Trail formation based on directed pheromone deposition. *J. Math. Biol.*, 66(6):1267–1301, 2013.

[BFW11] W. F. Bode, D. W. Franks, and J. Woods. Limited interactions in flocks: relating model simulations to empirical data. *J. R. Soc. Interface*, 8:301–304, 2011.

[Car18] P. Carmona. Some deterministic structured population models which are limit of stochastic individual based models.. Available on arXiv:1804.04913, math.pr, 2018.

[CCG+12] M. Camperi, A. Cavagna, I. Giardina, G. Parisi, and E. Silvestri. Spatially balanced topological interaction grants optimal cohesion in flocking models. *Interface Focus*, 2:715–725, 2012.

[CD16] F. Cucker and J-G. Dong. On flocks influenced by closest neighbors. *Mathematical Models and Methods in Applied Sciences*, 26, 11 2016.

[CD22] L-P. Chaintron and A. Diez. Propagation of chaos: A review of models, methods and applications. i. models and methods. *Kinetic and Related Models*, 30(5), 2022.

[CDP18] P. Cattiaux, F. Delebecque, and L. Pédèches. Stochastic Cucker-Smale models: Old and new. *Ann. Appl. Probab.*, 28(5):3239–3286, 2018.

[CHL17] Y P. Choi, S Y. Ha, and Z. Li. *Emergent dynamics of the Cucker-Smale flocking model and its variants*, pages 299–331. Number 9783319499949 in Modeling and Simulation in Science, Engineering and Technology. Springer Basel, 2017.

[CS07a] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Trans. Automat. Control*, 52:852–862, 2007.

[CS07b] F. Cucker and S. Smale. On the mathematics of emergence. *Jpn. J. Math.*, 2:197–227, 2007.

[DHJK20] J-G. Dong, S Y. Ha, J. Jung, and D. Kim. On the stochastic flocking of the cucker–smale flock with randomly switching topologies. *SIAM Journal on Control and Optimization*, 58:2332–2353, 01 2020.

[DM09] P. Diaconis and L. Miclo. On times to quasi-stationnarity for birth and death processes. *J. Theoret. Probab.*, 22(3):558–586, 2009.

[DP18] P. Degond and M. Pulvirenti. Propagation of chaos for topological interactions. Available on arXiv:1803.01922, math.ph, to appear in Ann. Appl. Probab., 2018.

[GM99] C. Graham and S. Méléard. Probabilistic tools and Monte Carlo approximations for some Boltzmann's equations. *ESAIM Proc.*, 10:77–126, 1999.

[Has13] J. Haskovec. Flocking dynamics and mean-field limit in the Cucker-Smale type models with topological interactions. *Phys. D*, 261:42–51, 2013.

[HM18] J. Haskovec and I. Markou. Delayed Cucker-Smale model with and without noise revisited. Available on arXiv:1810.01084, 2018.

[HZ22] Q. Huang and X. Zhang. On the stochastic singular cucker–smale model: Well-posedness, collision-avoidance and flocking. *Mathematical Models and Methods in Applied Sciences*, 32(01):43–99, 2022.

[JM08] B. Jourdain and F. Malrieu. Propagation of chaos and Poincaré inequalities of particles interacting through their CDF. *Ann. Appl. Probab.*, 18(5):1706–1736, 2008.

[KB13] M. Komareji and R. Bouffanais. Resilience and controllability of dynamic collective behaviors. *PLoS ONE*, 8(12), 2013.

[KD21] V. Kumar and R. De. Efficient flocking: metric versus topological interactions. *Royal Society Open Science*, 8, 09 2021.

[Mar14] S. Martin. Multi-agent flocking under topological interactions. *Systems and Control Letters*, 69:53–61, 2014.

[MNG17] H. Murakami, T. Niizato, and Y. P. Gunji. Emergence of a coherent and cohesive swarm based on mutual anticipation. *Sci. Rep.*, 7(46447), 2017.

[MT11] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behaviour. *J. Stat. Phys.*, 144:923–947, 2011.

[NM18] T. Niizato and H. Murakami. Entangled time in flocking: Multi-time-scale interaction reveals emergence of inherent noise. *PLoS ONE*, 13(4), 2018.

[Péd18] L. Pédèches. Asymptotic properties of various stochastic Cucker-Smale dynamics. *Discrete Contin. Dyn. Syst.*, 38(6):2731–2762, 2018.

[Rey15] J. Reygner. Chaoticity of the stationary distribution of rank-based interacting diffusions. *Electron. Commun. Probab.*, 20(60):1–20, 2015.

[SB14] Y.L. Shang and R. Bouffanais. Influence of the number of topologically interacting neighbors on swarm dynamics. *Sci. Rep.*, 4(4184), 2014.

[Shk12] M. Shkolnikov. Large systems of diffusions interacting through their ranks. *Stochastic Process. Appl.*, 122(4):1730–1747, 2012.

[VZ12] T. Vicsek and A. Zafeiris. Collective motion. *Physics Reports*, 517(3):71–140, 2012.

Patrick CATTIAUX,, Institut de Mathématiques de Toulouse. CNRS UMR 5219., Université Paul Sabatier,, 118 route de Narbonne, F-31062 Toulouse cedex 09.

E-mail address: patrick.cattiaux@math.univ-toulouse.fr

Fanny DELEBECQUE,, Institut de Mathématiques de Toulouse. CNRS UMR 5219., Université Paul Sabatier,, 118 route de Narbonne, F-31062 Toulouse cedex 09.

E-mail address: fanny.delebecque@math.univ-toulouse.fr

Olympio HACQUARD,, Ecole Normale Supérieure Paris Saclay,, .

E-mail address: olympio.hacquard@ens-paris-saclay.fr