TOPOLOGICAL INTERACTIONS FOR COLLECTIVE DYNAMICS: A STUDY OF A (NOT SO) SIMPLE MODEL.

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 (not so) toy-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 in a first part 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. The second part deals with partial results on the limiting behavior for large populations (N goes to infinity). We indeed show that the convergence of the first marginals is equivalent to the propagation of chaos. The propagation of chaos in full generality stays an open problem that has been dealt with in [DP18] for a subclass of the communication function. 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

Date: July 16, 2019.

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". 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 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. Among them, it was proposed in [MT11] to self-normalize the communication rates (instead of dividing by N). This model is also adapted to the emergence of several subgroups (*subflocks*).

A weakness of these models is that they 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 these 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 recent description, almost up to date, of the situation and rigorous statements of what flocking means in the stochastic situation (one can also look some extensions to delayed equations in [HM18]).

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], 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] among others). One can also mention [MNG17] where another type of interaction based on anticipation is proposed.

It is thus interesting to replace interactions depending on distances between individuals by interactions only depending on their closeness, or their ranks. This was done for the Cucker-Smale deterministic model in [Has13]. In the latter the author replaces $\Psi(|x_i(t) - x_j(t)|)$ 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]) or networks (see e.g. [Mar14]).

Recently another stochastic (not so) baby model was proposed by A. Blanchet and P. Degond ([BD16, BD17]). This model has some similarities with the discussion in [BFW11]. Let us describe this model, which will sometimes be called the B-D model in the sequel.

The 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 $N(t)$ 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 is chosen at random (*i.e.* uniformly with probability $1/N$). Notice that it is equivalent to say that to each particle i is attached a Poisson process $N_i(t)$ with intensity $\lambda(N)/N$, all these processes being independent and $N(t) = \sum_{i=1}^N N_i(t)$. When one of the clock rings (almost surely only one at the same time), the corresponding particle decides to "jump".
(3) When it "jumps", particle i chooses its partner $j \neq i$ according to a probability $\pi_{i,j}^N$ 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 j.

This immediate alignment of velocities seems to be a too simple alignment mechanism. We shall discuss this point in the final discussion section. Now let us precisely define the notion of rank.

Definition 1.1. *For each i one introduces the increasing ordering of distances $|x_i - x_j|$ between x_i and the other particles (here x_j $j \neq i$). Define $R^N(i,j)$ the rank of particle j for this order. If distances are the same for two different indices or more, one chooses one randomly with uniform probability. Thus $R^N(i,i) = 0$ and $R^N(i,j) = 1, ..., N-1$. We then introduce the normalized rank*

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

Remark 1.2. Note that, in the case where there are no tied particles (*i.e* the distances between any pair (i,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 1.3. *Let $k^N : [0,1] \to \mathbb{R}^+$ be continuous. Define*

$$K^N(r) = \frac{k^N(r)}{\sum_{l=1}^{N-1} k^N(l/N-1)}\,,$$

so that $\sum_{j=1, j\neq i}^N K^N(r^N(i,j)) = 1$. We thus define $\pi_{i,i}^N = 0$ and for $j \neq i$,

$$\pi_{i,j}^N = K^N(r^N(i,j))\,.$$

Note that we may assume or not that $\int_0^1 k(u)du = 1$. The model has several interesting properties. It corresponds to the required notion of topological interactions. In particular if for some $m \leq N-1$,

$$k^N(s) = \mathbf{1}(s \in [0, m/(N-1)])\,, \tag{1.2}$$

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.

Another interesting feature is that permutation invariance (exchangeability) is preserved by the dynamics, since the rank is permutation invariant.

From the mathematical point of view, we have to face some particular difficulties. Except if k^N is constant the transition probabilities $\pi_{i,j}^N$ are not symmetric: if j is the nearest neighbor of i, the converse is not necessarily true. Similarly the interaction between i and 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. Both previous remarks have an immediate consequence: the first one for the study of flocking, the second one for the large population limit.

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., we shall discuss in section 3. Actually the proofs in both papers are formal and lie in particular on an assumption of "propagation of chaos" that cannot be satisfied. These papers are nevertheless a very interesting first modelling attempt. More recently 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.

For the B-D model, we shall look at this problem in section 3, still with $k^N = k$. We rigorously prove that propagation of chaos is equivalent to convergence. The main point is thus propagation of chaos which remains an open problem in full generality.

Before to do this we shall first study the flocking property, *i.e.* the long time behaviour for a fixed N in section 2. 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. The main results of this section are the following:

(1) If interaction is allowed with a 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.
(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.

In the final discussion section, we shall examine the B-D model and the rigorous results we obtained from a more biological point of view. In particular, beyond the accuracy of the model, we shall see whether it predicts or not results in accordance with [BCC$^+$08].

2. Some flocking properties

In the sequel, the size N of the population being fixed, we will 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.
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, and $c(n) = \#\bar{V}(T_n)$ the cardinal of $\bar{V}(n)$. It is 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 2.1. *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 $T_\infty = \inf\{n\,;\, \#\bar{V}(n) = 1\}$, the system flocks if and only if $T_\infty < +\infty$ almost surely. We will call T_∞ the flocking time for the system.*

It is easy to build situations where flocking does not hold

Example 2.2. For $d = 1$, $N = 4$ and $k(s) = 3\,\mathbf{1}_{\{s\le 1/3\}}$, *i.e.* an interaction with the closest neighbor, 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 $3, 4$ are unchanged, so that $\bar{V}(\infty)$ contains two elements. ♢

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

In the previous example the lack of flocking follows from the fact that two initial clusters of velocities never meet. If any pair of particles has a positive probability to meet, one can think that flocking will happen. This is shown in the next theorem

Theorem 2.3. *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. Denote by $(T_n)_{n\ge1}$ the jump times and introduce 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. Note that the assumption on k is equivalent to the positivity of the $\pi_{i,j}$.
Denote by $\mathcal{F}$ the following event: at each time $i = 1, ..., N - 1$ the velocity $V_i(i)$ jumps to $V_N(i)$. It is clear that on $\mathcal{F}$ the system is flocking. Then

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

Define

$$\beta_N = \min_j \pi_{j,N} = \min_j \frac{k(j/N-1)}{\sum_{l=1}^{N-1} k(l/N-1)} \qquad \text{(see Definition 1.3.)} \tag{2.1}$$

It follows

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

The previous lower bound does not depend on the initial condition. Denote by S_∞ the flocking time for the skeleton chain. It follows that $\mathbb{P}(S_\infty = +\infty) \leq (1-p)$. 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. 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, 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. Though we still do not obtain our best bound, the line of reasoning will be interesting in the general non positive case.

At each time T_n we may divide the N particles into $c(n)$ clusters of particles with the same velocities. Once again we look at the skeleton chain $(X(n), V(n))$, and at each time n for this chain we look at the cardinal $c_{max}(n)$ of the largest cluster(s). Choose $C_{max}(n)$ one of the largest clusters at time n. Without loss of generality we may number from 1 up to $c_{max}(n)$ the elements of $C_{max}(n)$. With the definitions of β_N introduced before, it holds

$$\begin{aligned}\mathbb{P}(\#C_{max}(n) \text{ increases of 1 at time } n+1) &= \sum_{j > c_{max}(n)} \frac{\pi_{j,1} + ... + \pi_{j,c_{max}(n)}}{N} \\ &\geq \frac{\beta_N\, c_{max}(n)\,(N - c_{max}(n))}{N}\,.\end{aligned}$$

Of course, if $\#C_{max}(n)$ increases of 1 at time $n+1$, then $c_{max}(n+1) = c_{max}(n) + 1$.
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 $N - 1 \geq l \geq 1$,

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

This product is bounded from below by the product $\Pi_{n=1}^{N-1}$, so that for any initial configuration

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

We may then argue as we did with the event $\mathcal{F}$ and deduce

Proposition 2.4. *If β_N defined in* (2.1) *is positive,*

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

Using Stirling formula we get, for N large,

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

If much better than the previous one, does this estimate look good ? In the next subsection we shall study more carefully the expected flocking time.

2.2. **Case k=1.**

2.2.1. *Expected flocking time, connection with the neutral Moran model.*

Except when explicitly stated, in this subsection we focus on the skeleton chain introduced in the previous subsection. 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, for which, contrary to the B.D. model ,a particle j can decide to interact with itself (and thus stay as it is, which means that for all i, j, $\pi_{i,j} = \frac{1}{N}$). The flocking time in this context is called the fixation time.

Back to our model, assume in addition that $\#\bar{V}(0) = 2$, *i.e.* all initial velocities are equal to either $\bar{v}_1$ or $\bar{v}_2$. Estimates for the expected fixation time are then known, due to the fact that the number $c_1(n)$ of velocities equal to $\bar{v}_1$ at time n is a (discrete-time) birth and death Markov chain, whose transition probabilities for l in $\{1, ..., N-1\}$ are given by

$$\begin{aligned} Q(l,l+1) &= \frac{(N-l)l}{N(N-1)} \\ Q(l,l-1) &= Q(l,l+1) \\ Q(l,l) &= 1 - Q(l,l-1) - Q(l,l+1)\,, \end{aligned} \tag{2.2}$$

and which is absorbed when it hits 0 or N. It is thus known that

$$\mathbb{E}(S_\infty | c_1(0) = r) \;=\; (N-1)\left(\sum_{y=1}^{r}\frac{N-r}{N-y} + \sum_{y=r+1}^{N-1}\frac{r}{y}\right). \tag{2.3}$$

It is elementary to see that there exists C such that for any $r \leq N$,

$$\mathbb{E}(S_\infty | c_1(0) = r) \leq C\,N^2\,.$$

More precisely if $r = pN$ for some $0 < p < 1$,

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

as N goes to infinity.
A general lower bound can also be obtained writing

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

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

$$\mathbb{E}(S_\infty|c_1(0)=r) \geq \frac{(N-1)(N-6)}{4}. \tag{2.4}$$

We did not find in the literature an analogue for more than 2 different alleles, nor succeeded in deriving such a formula (one should perhaps use the results in [DMMPR10]). 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 $C \log_2(m) N^2$.
Also remark that, if S_∞^j denotes the first time the initial velocity $\bar{v}_j$ disappears or becomes the unique remaining velocity, then, a second of reflexion shows that

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

Another second of reflexion 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 (2.4) and Wald identity to get

Theorem 2.5. *Assume that $k = 1$ (and $N \geq 6$). There exists C such that for all initial configuration (x, v) satisfying $\#\bar{V}(0) = m$,*

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

Remark 2.6. An explicit (but tedious) computation of the mean flocking time can be done for small N (≤ 4) and leads to an exact value $(N-1)^2$. ♢

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

In this section, we present a few numerical experiments that we have carried out in order to illustrate Theorem 2.5. 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".

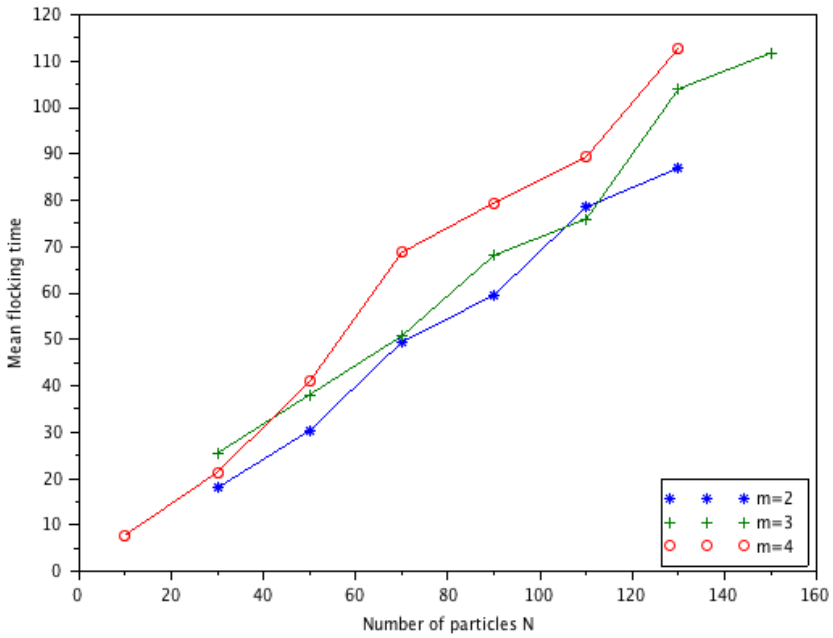


FIGURE 1. Mean flocking time as a function of N, case $k = 1$

	$m=2$		$m=3$			$m=4$			
	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 2.5 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.

2.3. **The general case $k > 0$: 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 have to look at the case $\#\bar{V}(0) = 2$ using the notations of subsubsection 2.2.1.

$V(n)$ 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 have an explicit formula for the mean fixation time (see below).

2.3.1. *A reminder of Birth and Death processes.*

Recall that for a birth and death process $R'(.)$ with transition $Q(i, i+\varepsilon)$ with $\varepsilon = -1, 0, 1$, the fixation time (denoted by F_∞) is the first time when R' equals 0 or N. The expectation of F_∞, starting from r :

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

solves the system of equations, for $0 < r < N$

$$m(r) = 1 + Q(r, r+1)\, m(r+1) + Q(r,r)\, m(r) + Q(r, r-1)\, m(r-1)\,, \tag{2.5}$$

with the boundary conditions $m(0) = m(N) = 0$.

Elementary calculations yield, for $2 \le r \le N$,

$$m(r) - m(r-1) \;=\; \prod_{j=1}^{r-1} \frac{Q(j,j-1)}{Q(j,j+1)} m(1) \;-\; \sum_{j=1}^{r-1} \frac{\prod_{l=j+1}^{r-1} Q(l,l-1)}{\prod_{l=j}^{r-1} Q(l,l+1)} \tag{2.6}$$

where as usual an empty sum equals 0 while an empty product equals 1. Using $m(0) = 0$ we deduce,

$$m(r) - m(1) = \left(\sum_{j=1}^{r} \prod_{l=1}^{j} \frac{Q(l,l-1)}{Q(l,l+1)} \right) m(1) \;-\; \sum_{j=1}^{r} \sum_{l=1}^{j} \frac{\prod_{i=l+1}^{j} Q(i,i-1)}{\prod_{i=l}^{j} Q(i,i+1)}\,. \tag{2.7}$$

Using $m(N) = 0$ we may obtain $m(1)$ and finally for $1 \le r \le N$,

$$m(r) \;=\; \frac{\left(\sum_{j=1}^{r} \prod_{l=1}^{j} \frac{Q(l,l-1)}{Q(l,l+1)} \right) \left(\sum_{j=1}^{N-1} \sum_{l=1}^{j} \frac{\prod_{i=l+1}^{j} Q(i,i-1)}{\prod_{i=l}^{j} Q(i,i+1)} \right)}{1 \;+\; \left(\sum_{j=1}^{N-1} \prod_{l=1}^{j} \frac{Q(l,l-1)}{Q(l,l+1)} \right)} \;-\; \sum_{j=1}^{r} \sum_{l=1}^{j} \frac{\prod_{i=l+1}^{j} Q(i,i-1)}{\prod_{i=l}^{j} Q(i,i+1)}\,. \tag{2.8}$$

This scary formula is hard to use in order to get a bound for the mean fixation time.

If in addition (which will be true in our case)

$$Q(r,r') = Q(N-r, N-r')\,, \tag{2.9}$$

it is more natural to compute $m(r)$ starting with $m(N/2)$ or $m(N+1/2)$.

Assume first that N is even. We thus have, since $m(r) = m(N-r)$,

$$\begin{aligned} m(N/2) = 1 + Q(N/2, (N/2)+1)\, m((N/2)+1) + Q(N/2, N/2)\, m(N/2) \\ + Q(N/2, (N/2)-1)\, m((N/2)-1) \end{aligned}$$

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

$$m((N/2)+1) - m(N/2) = \frac{-1}{2\, Q(N/2, (N/2)+1)} \;=\; \frac{-1}{\sigma_{N/2}}\,. \tag{2.10}$$

Similarly

$$m((N/2)+2) - m((N/2)+1) = \frac{-1}{\sigma_{(N/2)+1}} \left(1 \;+\; \frac{(Q((N/2)+1, N/2)}{2\, Q(N/2, (N/2)+1)} \right)$$

and by induction it is easily seen that $m(r)$ is decreasing on $\{N/2, (N/2)+1, ..., N\}$. It is thus enough to get a bound for $m(N/2)$.

When N is odd we have

$$\begin{gathered} m(N+1/2) = 1 + \\ + Q(N+1/2, (N+1/2)+1)\, m((N+1/2)+1) + Q(N+1/2, N+1/2)\, m(N+1/2) + \\ + Q(N+1/2, (N+1/2)-1)\, m((N+1/2)-1) \end{gathered}$$

and using symmetry $m(N+1/2) = m(N-1/2) = m((N+1/2)-1)$. It follows

$$m((N+1/2)+1) - m(N+1/2) = \frac{-1}{Q(N+1/2,(N+1/2)+1)}\,. \tag{2.11}$$

As before it easily follows that $m(r)$ is decreasing on $\{N/2, (N/2)+1, ..., N\}$ too.

We will thus get some upper bound for $m(N/2)$ or $m(N+1/2)$. Since the expressions are similar we only give the formulas for N even. After the dust settles one gets

$$\begin{aligned} m(N/2) \;=\; & \frac{1}{2\,Q(N/2,(N/2)+1)}\left(1+\sum_{l=1}^{(N/2)-1}\frac{\prod_{j=1}^{l} Q((N/2)+j,(N/2)+j-1)}{\prod_{j=1}^{l} Q((N/2)+j,(N/2)+j+1)}\right) + \\ & + \sum_{i=1}^{(N/2)-1}\frac{1}{Q((N/2)+i,(N/2)+i+1)} \\ & \left(1+\sum_{l=1}^{(N/2)-1-i}\frac{\prod_{j=1}^{l} Q((N/2)+i+j,(N/2)+i+j-1)}{\prod_{j=1}^{l} Q((N/2)+i+j,(N/2)+i+j+1)}\right). \end{aligned} \tag{2.12}$$

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

Coming back to the B.D. model, for simplicity we will assume that k is positive and non-increasing. In this section, we aim at constructing a coupling of two different processes in order to "simulate" the worst possible case and study the expected flocking time for this coupling. For any fixed $j \in \{1, ..., N-1\}$, define

$$\sigma'_j = \frac{\sum_{l=1}^{j} k(l/N-1)}{\sum_{l=1}^{N-1} k(l/N-1)}\,, \quad \text{and} \quad \sigma_j = \frac{\sum_{l=N-j}^{N-1} k(l/N-1)}{\sum_{l=1}^{N-1} k(l/N-1)}. \tag{2.13}$$

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). 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 the number $c_1(n)$ of particles with velocity $\bar{v}_1$ at time n. We recall that $c_1(n+1) = c_1(n) + \varepsilon$, $\varepsilon \in \{-1, 0, 1\}$. Here, $c_1(n+1) = c_1(n) + 1$ provided one chooses a particle with velocity $\bar{v}_2$ to align with a particle of velocity $\bar{v}_1$. Denote,

$$p(n, r, r+\varepsilon) \;=\; \mathbb{P}(c_1(n+1) = r+\varepsilon | c_1(n) = r)\,. \tag{2.14}$$

One may expect that the worst case (*i.e* leading to the maximal flocking time) is the case where the transitions in the direction of 0 or N are the smallest ones (see inequalities (2.15)). We will therefore prove that this guess is correct and get explicit bounds in the worst case.

It clearly holds

$$\frac{N-r}{N}\,\sigma'_r \geq p(n, r, r+1) \geq \frac{N-r}{N}\,\sigma_r\,. \tag{2.15}$$

Similarly,

$$\frac{r}{N}\,\sigma'_{N-r} \geq p(n, r, r-1) \geq \frac{r}{N}\,\sigma_{N-r}$$

and

$$\frac{1}{N}\left(r\sigma'_{r-1}+(N-r)\sigma'_{N-r-1}\right)\geq p(n,r,r)\geq\frac{1}{N}\left(r\sigma_{r-1}+(N-r)\sigma_{N-r-1}\right).$$

We introduce two coupled new processes $R(n)$ and $R'(n)$. $R(n)$ is a Markov chain with transition matrix $Q(i,j)$ which is a birth and death process on $\{0,...,N\}$, and such that

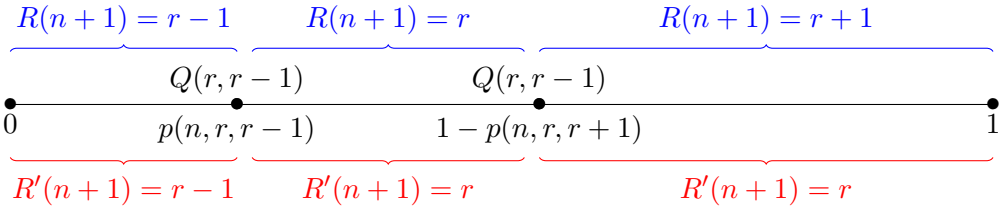
- for $N-1\geq r\geq(N+1)/2$,
$$Q(r,r+1)=\frac{N-r}{N}\,\sigma_r\ ,\ Q(r,r-1)=\frac{r}{N}\,\sigma'_{N-r}\ , \tag{2.16}$$
and $Q(r,r)=1-Q(r,r+1)-Q(r,r-1)$;
- for $1\leq r<N/2$,
$$Q(r,r-1)=\frac{r}{N}\,\sigma_{N-r}\ ,\ Q(r,r+1)=\frac{N-r}{N}\,\sigma'_r\ , \tag{2.17}$$
and $Q(r,r)=1-Q(r,r-1)-Q(r,r+1)$;
- and finally if $r=N/2$ (when N is even),
$$Q(r,r-1)=\frac{r}{N}\,\sigma_r\ ,\ Q(r,r+1)=\frac{N-r}{N}\,\sigma_r\ , \tag{2.18}$$
and $Q(r,r)=1-Q(r,r-1)-Q(r,r+1)$.

We do not define the transition matrix for $r=0$ and $r=N$ since we will study $R(n)$ stopped at the first time it hits 0 or N. It is worth noticing that (2.9) is satisfied.

$R'(n)$ 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 the same as the one of $c_1(.)$.
$R(n)$ is in a sense a "slowed down" version of $R'(n)$ since the probability to approach the boundary $\{0,N\}$ is at each step less than the corresponding one for $R'(n)$.

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 $U_n \geq 1 - p(n, r, r-1)$, $R'(n) = r$ if $p(n, r, r+1) \leq U_n \leq 1 - p(n, r, r-1)$ and $R'(n+1) = r+1$ if $U_n \leq p(n, r, r+1)$.

This construction furnishes a coupling of the desired processes (*i.e.* $R(.)$ and $R'(.)$ have the desired distribution).

We are now interested in the quantities $D(n) = R(n)(N - R(n))$ and $D'(n) = R'(n)(N - R'(n))$, describing the proximity of $\{0, N\}$.

Lemma 2.7. *If $0 < D'(n) \leq D(n)$, then almost surely $D'(n+1) \leq D(n+1)$.*

Proof. Using symmetry we may assume that $R'(n) = r$ with $N - 1 \geq r > N/2$ or $r = N/2$ with N even.
The second case $r = \frac{N}{2}$ is easy since $D(n) \leq N^2/4 = D'(n)$ so that, since $D(n) \geq D'(n)$ we have $D(n) = N^2/4$ and $R(n) = N/2$. If $D'(n+1) = D'(n)$, then

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

so that $D(n+1) = D(n) = D'(n+1)$. In all other cases $D'(n+1)$ decreases and $D(n+1) \geq D'(n+1)$.
If $r > N/2$, then

- if $R'(n+1) = r+1$, $D'(n+1) < D'(n)$ and whatever $R(n+1)$ is, $D(n+1) \geq D'(n+1)$.
- If $R'(n+1) = r$ (when $p(n, r, r-1) \leq U_n \leq 1 - p(n, r, r+1) \leq 1 - Q(r, r+1)$), $D'(n+1) = D'(n) \leq D(n+1)$ if $D(n) > D'(n)$. The case $D(n) = D'(n)$ splits in two cases: either $R(n) = R'(n)$, but in this case $R(n+1) = R(n)$ or $R(n+1) = r-1$, due to the position of U_n since $Q(r, r-1) \geq p(n, r, r-1)$, or $R(n) = N - R'(n)$ and the same holds thanks to our construction and the symmetry (2.9).
- When $R'(n+1) = r-1$, we have to look at two cases (in all other cases $D(n+1) \geq D'(n+1)$ is immediate): either $D(n) = D'(n) = r(N-r)$ or $D(n) = (r-1)(N - r + 1) > D'(n)$. In the first case it is easy to see that when $R(n) = R'(n)$ necessarily $R(n+1) = r-1$ looking at the position of U_n. The case $R(n) = N - R'(n)$ follows from (2.9). The remaining case is thus $R(n) = r - 1 = R'(n) - 1$ (or the symmetric $R(n) = N - r + 1$), $R'(n+1) = r-1$. The only thing to show is that, in this situation the case $R(n+1) = r$ is almost surely impossible.

  The latter is of course impossible if

$$p(n, r+1, r) \leq 1 - Q(r, r+1)\,. \tag{2.19}$$

  Since $Q(r+1, r) \geq p(n, r+1, r)$, condition (2.19) is fulfilled provided

$$(N-r)\,\sigma'_{N-r-1} \leq N \,-\, (N-r)\,\sigma_r\,,$$

  which is immediate since $\sigma'_{N-r-1} + \sigma_r = 1$.

This ends the proof of the lemma. □

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 less than the hitting time F_∞ of $\{0, N\}$ by $R(.)$ (fixation time for the birth and death process).

It is not difficult to check that, since k is non-increasing, $Q(i,i+1) \leq Q(i,i-1)$ for $i \geq N/2$, with equality when k is constant, more precisely, for $r \geq N/2$

$$\frac{Q(r,r-1)}{Q(r,r+1)} = \frac{r \sum_{l=1}^{N-r} k(l/N-1)}{(N-r) \sum_{l=N-r}^{N-1} k(l/N-1)} \geq 1\,. \tag{2.20}$$

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

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

Notice 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

Theorem 2.8. *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

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

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 2.9. *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 2.4. 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 2.10. *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 2.4.1.

Remark 2.11. 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. ♢

2.3.3. *Numerical simulations.*

In that section we aim at illustrating the bounds proved in Corollaries 2.9 and 2.10. 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 preceding 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.

	$k(r) = 1.01 - r$	$k(r) = 1 - r$	$k(r) = \frac{8}{(2+2r)^3}$	$k(r) \in \{N, \frac{1}{N}\}$
$N = 50$	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.

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

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 2.2.2, (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 2.10 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.

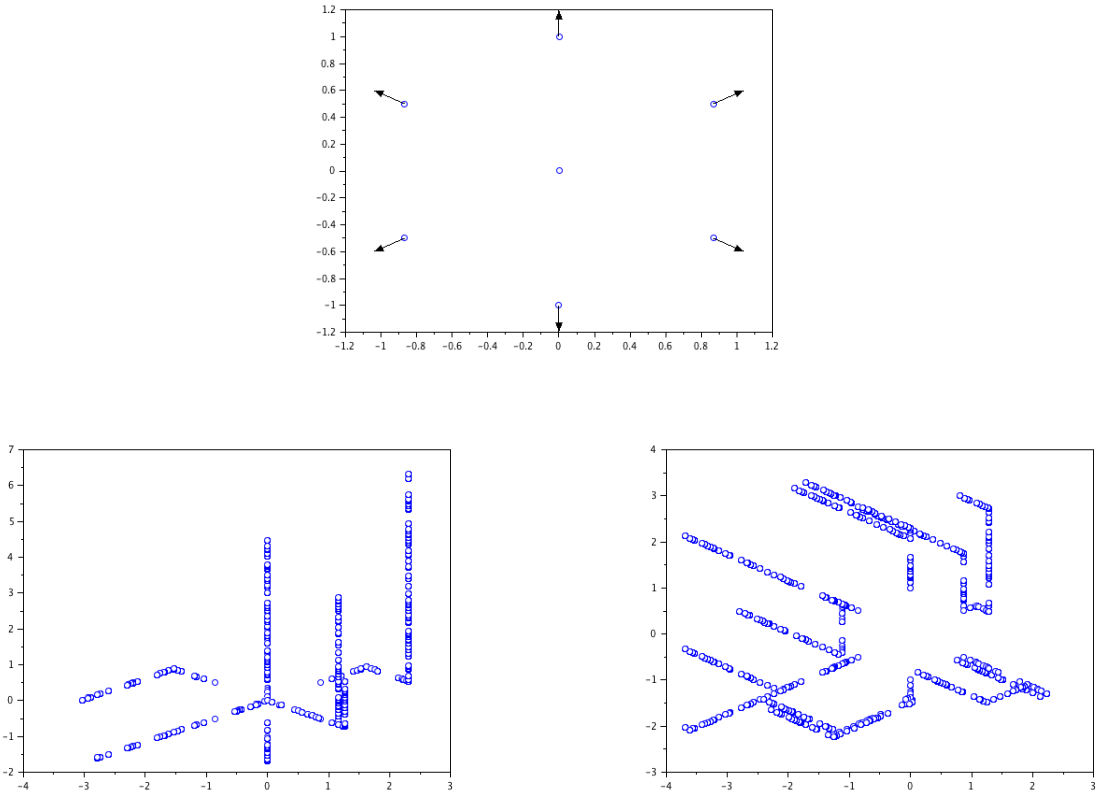


FIGURE 2. Different trajectories in the case $N = 12$. Top : initial data, Bottom $k(r) = 1 - r$ left : agents at the center, $V_f = (0, 0)$, $T_f = 8.15$, right no agent at the center : $T_f = 3.26$, $V_f = (-\cos(\pi/6), 1/2)$

2.4. **The general case: sub-flocking.**

2.4.1. *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 have that $k(s) > 0$ for $s \in [0, a[$ and $k(s) = 0$ on $[a, 1]$. We then have the following result

Theorem 2.12. *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 $i = 1, ..., N$, $\#\{l\ ;\ v_l(\infty) = v_i(\infty)\} \geq j + 1$.

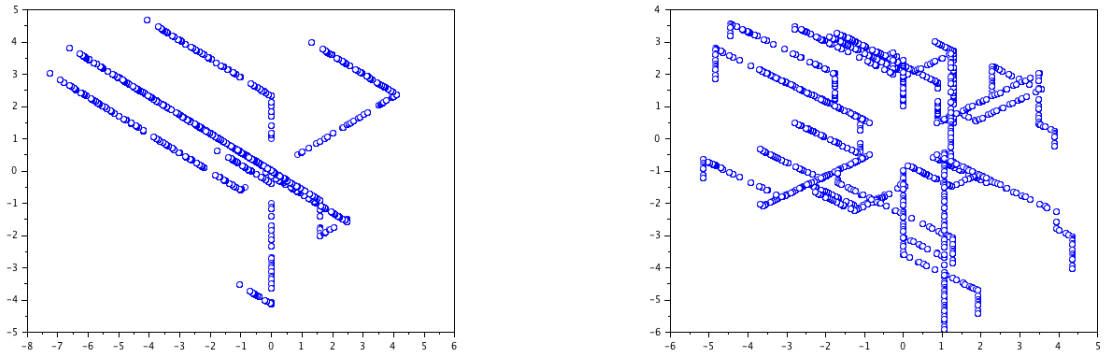


FIGURE 3. Different trajectories in the case $N = 12$, k cubic, left : agents at the center, $T_f = 9.8$, $V_f = (-\cos(\pi/6), 1/2)$, right : $T_f = 5.53$, $V_f = (0, -1)$

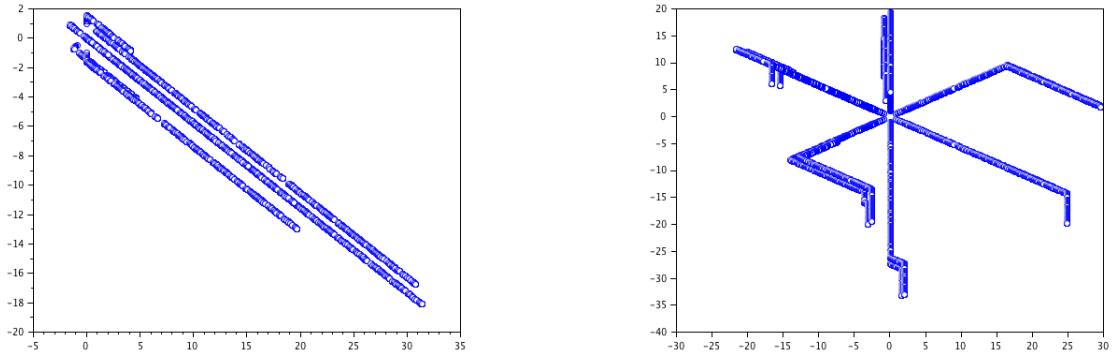


FIGURE 4. Different trajectories in the case $N = 12$, $k \in \{1, N\}$, left : agents at the center, $T_f = 35.25$, $V_f = (\cos(\pi/6), -1/2)$, right : $T_f = 33.28$, $V_f = (0, -1)$

It is worth noticing that the final $v(\infty)$ depends on ω. The theorem means that, in any case, almost surely, velocities are all kept fixed after some random time. The second part say that, as soon as one initial velocity is preserved in $v(\infty)$ it will appear at least $j + 1$ times.

Proof. We choose j such that $k(j/N - 1) > k(j + 1/N - 1) = 0$. Hence a given particle, when it jumps, has to choose the velocity of one of its j nearest neighbors. At each time t we can divide $v_1(t), ..., v_N(t)$ into $c(t) = \#\bar{V}(t)$ clusters, each cluster gathering all particles with the same velocity. As we already said $c(t)$ is a non increasing sequence, hence it is stationary, equal to $c(\infty)$ for $t \geq t_\infty$ for some finite t_∞. We first claim

Lemma 2.13. *Let k be as before, i.e. 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.*

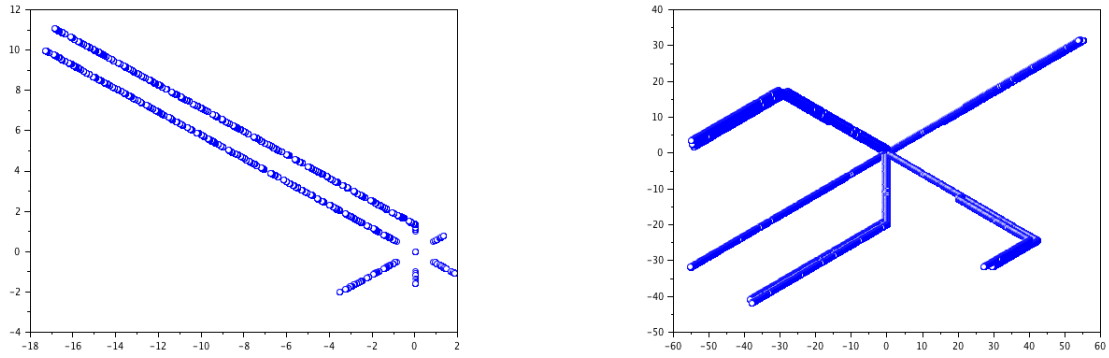


FIGURE 5. Different trajectories in the case $N = 12, k$ compactly supported, right : $T_f = 19.8,\ V_f = (0, 0)$, left : $T_f = 62.63,\ V_f = (-\cos(\pi/6), -1/2)$

Proof. of the lemma. Let $t > t_\infty$, $C(t)$ be one cluster and assume that $\#C(t) = l \leq j$. Then any particle in C has its neighbor with rank j that does not belong to C.
Introduce the event $\mathcal{G} = \{$For the first jump, one chooses one element of C and this element chooses as partner its neighbor with rank j, second, one chooses a second different element of C and this element chooses as partner its neighbor with rank j, and so on up to the lth element in $C\}$. We have

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

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. $(c_{min}(T_n \vee t'_\infty))_n$ is a bounded sequence taking its values in $\mathbb{N}*$. Let $c* = \liminf_n c_{min}(T_n \vee t'_\infty)$. First of all $c* \geq j + 1$. 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 $|x_i(t) - x_l(t)|$ growths 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 2.14. *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 2.10.

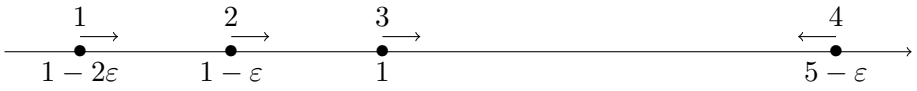
2.4.2. *Illuminating examples and simulations in one dimension, nearest neighbor case.*

In the spirit of example 2.2, 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* $0 < \mathbb{P}(\text{flock}) < 1$.

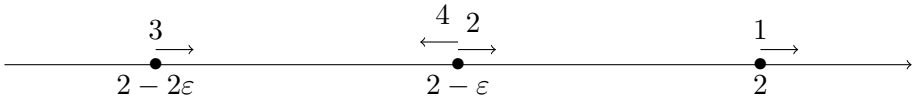
Example 2.15. In this example, we take

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

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 $\frac{5}{8}$.

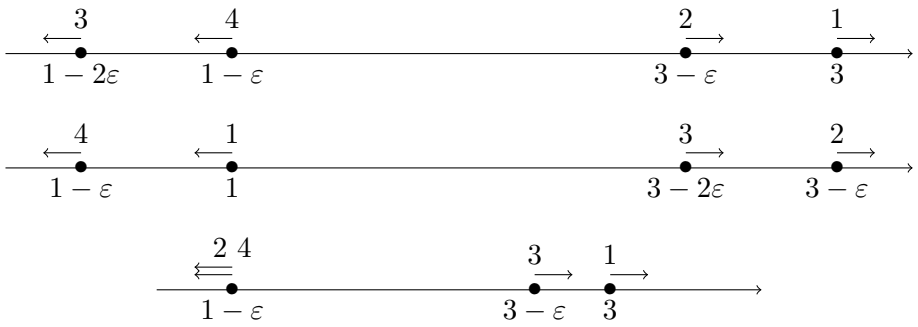


Indeed, as soon as particle 4 is chosen, it aligns with a particle of velocity 1 and there is flocking. If particle 4 is not chosen until time T_3, at time T_3, the situation is the following :



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

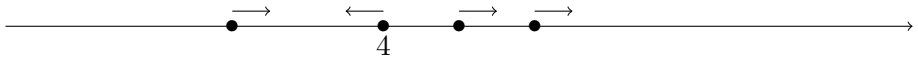
get velocity -1. If particle 2 is chosen, it aligns with 4 and takes velocity -1. The different possibilities at time T_4 are therefore the following :



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 2.15 : 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 2.16. It is also possible to slightly modify the initial data of example 2.15 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 2.16 : 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

3. Large population asymptotics: the case $k^N = k$.

In this section we shall study the behavior of the distribution of one particle, or a finite number of particles, as N growths to infinity. In order to get some non trivial limit, this will impose a value for $\lambda(N)$. As already said, a formal derivation of the limit is obtained in [BD16, BD17].

Assumption 3.1. *We will denote $z(t) = (x(t), v(t))$ and assume that the initial $(z_i^N(0))_{i=1,\dots,N}$ are i.i.d. with common distribution $p(x,v)dxdv$ that does not depend on N. For technical reasons we also assume that p is bounded by M.*
It is thus easy to see that for all t, $(z_i^N(t))_{i=1,\dots,N}$ is exchangeable, i.e. its distribution is invariant by permutation. In addition the law of $x^N(t)$ still has a density $\rho^N(x,t)$ w.r.t.

Lebesque's measure for all t because the jump times are independent of the dynamics. This simplification avoid ties, since equality of two positions at a given time doe not occur almost surely.
Finally it is not difficult to see that, for any fixed j, the joint density of $(x_1^N(t), ..., x_j^N(t))$ denoted by $\rho^N_{1,...,j}(x,t)$ is still bounded by M^j which is a bound at time 0 using independence.

Notation 3.2. Even if the distribution of $z^N(0)$ has a density w.r.t. Lebesgue on $\mathbb{R}^{2Nd}$, the distribution of $z(t)$ is singular, since the probability for at least two velocities to be equal, is positive. Nevertheless, by abuse of notation, and in order to make easier the comparison with [BD16, BD17, DP18] where this notation is used, we will write $f^N(z,t)\,dz$ for the distribution of $z^N(t)$.
As for the assumption we will denote by $f^N_{1,...,j}$ the marginal distribution of the first j particles under f^N. ♢

The distribution of $z^N(t)$ satisfies, in a weak sense, the so called "master equation": for all test function ϕ on $\mathbb{R}^{2Nd}$,

$$\begin{aligned}\partial_t \int f^N(z,t)\,\phi(z)\,dz \;-\; \sum_{i=1}^N \int (v_i.\nabla_{x_i}\phi(z))\, f^N(z,t)\,dz \\ = \;-\lambda(N)\int \phi(z)\, f^N(z,t)\,dz \\ +\frac{\lambda(N)}{N}\sum_{i,j=1,...,N;i\neq j}\int \pi^N_{i,j}(x)\,\phi(\tilde z(i,j))\, f^N(z,t)\,dz\,, \end{aligned} \tag{3.1}$$

where,

$$\tilde z(i,j) = (z_1,, z_{i-1}, (x_i, v_j), z_{i+1}, ..., z_N)$$

i.e. the velocity of particle i has moved to v_j.
Using (3.1) it is not difficult to see that Assumption 3.1 propagates, *i.e.* the distribution of $z(t)$ is exchangeable for all t.
Using this exchangeability, we may particularize the master equation to ϕ depending on a given (fixed) number of particles (say $1, ..., j$), and get an equation satisfied by their distribution. We shall only write the one satisfied by one single particle (say 1), whose law is denoted by $f_1^N(z_1,t)\,dz_1$ (the 1 marginal law). The following equation is derived in [BD16, DP18] in the particular case $\lambda(N) = N$.

$$\begin{aligned}\partial_t \int f_1^N(z_1,t)\,\phi(z_1)\,dz_1 \;-\; \int (v_1.\nabla_{x_1}\phi(z_1))\, f_1^N(z_1,t)\,dz_1 \\ = \;\frac{N-1}{N}\,\lambda(N)\int \pi^N_{1,2}(x)\,[\phi(x_1,v_2)-\phi(x_1,v_1)]\, f^N(z,t)\,dz \end{aligned} \tag{3.2}$$

which is still equal to

$$= \;\frac{N-1}{N}\,\lambda(N)\int \pi^N_{1,2}(x)\,\phi(x_1,v_2)\, f^N_{1,2}(z,t)\,dz \;-\; \frac{\lambda(N)}{N}\int \phi(x_1,v_1)\, f_1^N(z,t)\,dz\,. \tag{3.3}$$

Here we have explicitly written the dependence of $\pi^N_{1,2}$ in the whole system of positions, so that the right hand side cannot be simplified. Writing the explicit expression

$$\lambda(N)\,\pi^N_{1,2}(x) = \frac{\lambda(N)}{\sum_{j=1}^{N-1} k^N(j/N-1)}\;\; k^N\left(\frac{1}{N-1}\sum_{l\neq 1, 1\leq l\leq N} \mathbf{1}_{B(x_1,|x_1-x_2|)}(x_l)\right),$$

we see what kind of difficulty we have to handle with to pass to the limit.

If we assume that $k^N = k$ does not depend on N we may split the problem into two parts: choose $\lambda(N)$ such that we get some limit for $k'_N = \frac{\lambda(N)}{\sum_{j=1}^{N-1} k(j/N-1)}$ and show some "law of large numbers" type for $U_N = \frac{1}{N-1}\sum_{l\neq 1,1\leq l\leq N} \mathbf{1}_{B(x_1,|x_1-x_2|)}(x_l)$.

For the first problem, if $\lambda(N) \ll N$ the limit is 0 and the limiting behavior of the right hand side of (3.2) is also 0, so that the asymptotic behavior of each particle is governed by the simple transport equation. The interesting case is thus

$$\lambda(N) = N$$

for which k'_N converges to the inverse of the integral of k we have chosen equal to 1.

The study of U_N is more delicate. However, conditionally (w.r.t. $f^N(z,t)dz$) to (x_1, x_2), the $W_l = \mathbf{1}_{B(x_1,|x_1-x_2|)}(x_l)$, $l \neq 1, 2$ are exchangeable Bernoulli variables with parameter $M_{\rho^N}(x_1, |x_1 - x_2|)$ defined as follows. Recall that

$$\rho^N(.,t) = \int f^N(.,v,t)\,dv\,,$$

is the distribution at time t of $x^N(t)$. Under assumption 3.1 this distribution is absolutely continuous and $\rho^N(y,t)$ is then its density at point y defined on $\mathbb{R}^{dN}$. We can also consider

$$\rho^N_{1,2,3}(x_1, x_2, x_3, t)$$

the density of the distribution of $(x^N_1(t), x^N_2(t), x^N_3(t))$. Finally

$$\begin{aligned} M_{\rho^N}(x_1, |x_1 - x_2|) &= \int_{\mathbb{R}^d} \mathbf{1}_{|u-x_1|\leq|x_1-x_2|}\,\rho^N_{1,2,3}(x_1, x_2, u, t)\,du \\ &= \mathbb{P}\left[(|x^N_3 - x_1| < |x_2 - x_1|)\big|x^N_1 = x_1, x^N_2 = x_2\right]. \end{aligned} \tag{3.4}$$

We may thus state a first lemma

Lemma 3.3. *Denote by $\delta(\varepsilon)$ the modulus of (uniform) continuity of k, i.e. $|k(s) - k(t)| \leq \varepsilon$ as soon as $|s - t| \leq \delta(\varepsilon)$. Then as $N \to +\infty$,*

$$\begin{aligned} &\rho^N\left(\left|k\left(\frac{1}{N-1}\sum_{l\neq 1,1\leq l\leq N} \mathbf{1}_{B(x_1,|x_1-x_2|)}(x_l)\right) - k(M_{\rho^N}(x_1, |x_1 - x_2|))\right| > \varepsilon\right) \\ &\qquad = \frac{1}{\delta(\varepsilon)^2}\;\; O\left(\max(\frac{1}{N}\,,\; cov^N)\right), \end{aligned}$$

where

$$cov^N := \int_{\mathbb{R}^{2d}} Cov\left(\mathbf{1}_{B(x_1,|x_1-x_2|)}(x^N_3(t))\,,\, \mathbf{1}_{B(x_1,|x_1-x_2|)}(x^N_4(t))\right)\,\rho^N_{1,2}(x_1, x_2, t)\,dx_1\,dx_2\,.$$

Notice that, thanks to exchangeability again

$$\begin{aligned} cov^N &= \mathbb{E}\left(\mathbf{1}_{B(x_1,|x_1-x_2|)}(x_3^N(t))\,,\,\mathbf{1}_{B(x_1,|x_1-x_2|)}(x_4^N(t))\right) \\ &- \mathbb{E}\left(\mathbb{E}^2\left(\mathbf{1}_{B(x_1,|x_1-x_2|)}(x_3^N(t))\Big|x_1^N(t),x_2^N(t)\right)\right)\,. \end{aligned} \tag{3.5}$$

Proof. If $|k(a)-k(b)|\geq\varepsilon$ then $|a-b|\geq\delta(\varepsilon)$, so it is enough to bound the probability for

$$\left|\frac{1}{N-1}\sum_{l\neq 1,1\leq l\leq N}\mathbf{1}_{B(x_1,|x_1-x_2|)}(x_l)-M_{\rho^N}(x_1,|x_1-x_2|)\right|$$

to be larger than $\delta(\varepsilon)$.
To this end we use Tchebychev inequality, and thus we calculate the expectation of the square of the expression into $|.|$. We first condition with respect to the first two variables, then use exchangeability, apply (3.4) and finally integrate w.r.t. the law of the first two variables. After the dust settle, this expectation is less than

$$\frac{1}{(N-1)}+\frac{(N-2)(N-3)}{(N-1)^2}\,cov^N\,.$$

□

Convergence to 0 of this quantity is thus linked to the chaoticity of the system. Let us recall some definitions:

Definition 3.4. *We say that a triangular array of random vectors $(Z_1^N,...,Z_N^N)$ defined on E^N is ν chaotic if there exists some probability measure ν on E such that for all fixed j, the distribution of the (finite) random vectors $(Z_1^N,...,Z_j^N)$ weakly converges towards $\nu^{\otimes j}$ as N goes to infinity. In other words, for any fixed j, the (finite) vectors $(Z_1^N,...,Z_j^N)$ are asymptotically i.i.d. with common distribution ν.*
For a random dynamical system we say that propagation of chaos holds when chatocity at time 0 implies chaoticity for all times.
Recall that for an exchangeable array i.e. if for all N, $(Z_1^N,...,Z_N^N)$ is an exchangeable vector, ν chaoticity is equivalent to the weak convergence of (Z_1^N,Z_2^N) to $\nu^{\otimes 2}$.

With all these considerations in mind we may state a first result

Theorem 3.5. *Assume that*

1. *For all N assumption 3.1 is satisfied,*
2. *$k^N=k$ is fixed, $\lambda(N)=N$,*
3. *up to a subsequence, $f_{1,2}^N(z_1,z_2,t)dz_1dz_2$ weakly converges as $N\to\infty$ to some limiting measure $f_{1,2}(z_1,z_2,t)dz_1dz_2$, for all t. Here weak convergence is supposed to hold for the weak* convergence of measures (convergence in law).*

Assume in addition that for all t,

$$cov^N\ \to\ 0 \tag{3.6}$$

as N goes to infinity, where cov^N is defined by (3.5).
Then, provided it is absolutely continuous (w.r.t. Lebesgue), the limiting $f_1(z,t)dz$ solves the

following P.D.E, in weak form

$$\partial_t \int f_1(z_1,t)\,\phi(z_1)\,dz_1 - \int (v_1.\nabla_{x_1}\phi(z_1))\,f_1(z_1,t)\,dz_1$$
$$= \int [\phi(x_1,v_1') - \phi(x_1,v_1)]\,k(M_{\rho_1}(x_1,|x_1-x_1'|))\,f_1(z_1,t)\,f_1(z_1',t)\,dz_1\,dz_1'\,, \quad (3.7)$$

with

$$M_{\rho_1}(x_1,|x_1-x_1'|) = \int \mathbf{1}_{|y-x_1|\leq|x_1-x_1'|}\,\rho_1(y,t)\,dy \quad ; \quad \rho_1(y,t) = \int_{\mathbb{R}^d} f_1(y,v,t)\,dv\,.$$

We may complete the previous statement

Proposition 3.6. *If assumption 3.1 is satisfied, then*

(1) all $z_i^N(0)$ have the same distribution $f_1^N(z,0)dz$ for $i=1,...,N$. If all $f_1^N(z,0)dz$, for $N\geq 1$, have compact support included into a given fixed compact set, then condition (3) in theorem 3.5 is satisfied.

(2) If for all t the system $(x^N(t))_{N\geq 4}$ is chaotic, then (3.6) is satisfied.

Proof. of Proposition 3.6. For (1), it is enough to remark that the distribution of the $v_i^N(t)$ has compact support included into the one of $v_1^N(0)$ since the set of possible velocities does not change. $x_i(t)$ thus has a support included into $supp(x_1(0)) + t\max_i|v_i(0)|$ which is still compact so that the support of the distribution at time t is supported by some compact subset for all N. The result is thus a direct consequence of Prohorov tightness theorem.

The proof of (2) is immediate using the definition of chaoticity for 4 variables, thanks to (3.4). Indeed if $U_1,...,U_4$ are i.i.d with an absolutely continuous distribution ρ,

$$\mathbb{E}[h(U_3-U_1,U_2-U_1)h(U_4-U_1,U_2-U_1)] = \int \rho(u_1)\rho(u_2)$$
$$\left(\int h(u_3-u_1,u_2-u_1)h(u_4-u_1,u_2-u_1)\,\rho(u_3)\rho(u_4)du_3du_4\right)du_1du_2$$
$$= \int \rho(u_1)\rho(u_2)\left(\int h(u_3-u_1,u_2-u_1)\rho(u_3)du_3\right)\left(\int h(u_4-u_1,u_2-u_1)\rho(u_4)du_4\right)du_1du_2$$

□

Proof. of Theorem3.5. If (3) is satisfied, the left hand side of (3.2) converges to the left hand side of (3.7), at least after integrating with respect to the time t in an interval.

To get the almost sure convergence of $M_{\rho_1^N}$ to M_{ρ_1}, we use the assumption that the limiting measure $f_1(z,t)dz$ is absolutely continuous, the fact that the discontinuity set of the function $y\mapsto \mathbf{1}_{|y-x_1|\leq|x_1-x_1'|}$ is Lebesgue measurable and the portmanteau theorem.

It remains to prove the convergence of

$$\frac{N-1}{N}\,\lambda(N)\int \pi_{1,2}^N(x)\,[\phi(x_1,v_2)-\phi(x_1,v_1)]\,f^N(z,t)\,dz\,,$$

and since $\lambda(N)=N$ the one of

$$\int [\phi(x_1,v_2)-\phi(x_1,v_1)]\,k\left(\frac{1}{N-1}\sum_{l\neq 1,1\leq l\leq N}\mathbf{1}_{B(x_1,|x_1-x_2|)}(x_l)\right)f^N(z,t)\,dz$$

to

$$\int \left[\phi(x_1, v_1') - \phi(x_1, v_1)\right] \, k(M_{\rho_1}(x_1, |x_1 - x_1'|)) \, f_1(z_1) \, f_1(z_1') \, dz_1 \, dz_1' \, .$$

We introduce

$$\int \left[\phi(x_1, v_2) - \phi(x_1, v_1)\right] k(M_{\rho_1}(x_1, |x_1 - x_2|)) \, (f^N(z,t) \, - \, f_1(z_1,t) f_1(z_2,t)) \, dz \, .$$

This term goes to 0 thanks to the continuity (and boundedness) of k and the weak convergence of $f^N_{1,2}$. For the remaining

$$\int \left[\phi(x_1, v_2) - \phi(x_1, v_1)\right] \left(k\left(\frac{1}{N-1} \sum_{l \neq 1, 1 \leq l \leq N} \mathbf{1}_{B(x_1, |x_1 - x_2|)}(x_l) \right) - k(M_{\rho_1}(x_1, |x_1 - x_2|)) \right) f^N(z,t) \, dz \, ,$$

and since ϕ is bounded, it is enough to look at

$$\left| k\left(\frac{1}{N-1} \sum_{l \neq 1, 1 \leq l \leq N} \mathbf{1}_{B(x_1, |x_1 - x_2|)}(x_l) \right) \, - \, k(M_{\rho_1}(x_1, |x_1 - x_1'|)) \right| \, f^N(z,t) \, .$$

Recall that we are making an abuse of notation by writing $f^N(z,t)$ but it is easily seen below that this is not relevant.
We decompose the latter in the sum of two terms $A_1 + A_2$, A_1 given by,

$$\left| k\left(\frac{1}{N-1} \sum_{2 \leq l \leq N} \mathbf{1}_{B(x_1, |x_1 - x_2|)}(x_l) \right) \, - \, k(M_{\rho_1^N}(x_1, |x_1 - x_2|)) \right| \, f^N(z,t)$$

A_2 given by

$$\left| k(M_{\rho_1^N}(x_1, |x_1 - x_2|)) \, - \, k(M_{\rho_1}(x_1, |x_1 - x_2|)) \right| \, f^N(z,t) \, .$$

According to lemma 3.3 we have

$$A_1 \, \leq \, \varepsilon \, + \, \frac{\| \, k \, \|_\infty}{\delta(\varepsilon)^2} \, O\left(\max(\frac{1}{N} \, ; \, cov^N) \right) \, ,$$

so that we may choose a sequence ε_N going to 0 such that A_1 goes to 0 thanks to (3.6).
For A_2 we may first integrate with respect to all except x_1 and x_2 variables, *i.e.* replace $f^N(z,t)$ by $\rho^N_{1,2}(x,t)$. But $\rho^N_{1,2}$ is bounded by M^2 thanks to assumption 3.1, and $k(M_{\rho_1^N}(x_1, |x_1 - x_2|)) - k(M_{\rho_1}(x_1, |x_1 - x_2|))$ is bounded by $2 \, \| \, k \, \|_\infty$ and goes to 0 thanks to what precedes and the continuity of k. We may thus apply Lebesgue's bounded convergence theorem to conclude that A_2 goes to 0 too. The proof is completed. □

Remark 3.7. Using (3.3) instead of (3.2), the above proof can be mimic in order to show that the limit $f_1(z_1, t)$ also satisfies

$$\begin{aligned} \partial_t \int f_1(z_1,t) \, \phi(z_1) \, dz_1 \, - \, & \int (v_1 . \nabla_{x_1} \phi(z_1)) \, f_1(z_1,t) \, dz_1 \\ = & \int \phi(x_1, v_1') \, k(M_{\rho_1}(x_1, |x_1 - x_1'|)) \, f_1(z_1,t) \, f_1(z_1',t) \, dz_1 \, dz_1' \, - \\ - & \int \phi(x_1, v_1) \, f_1(z_1,t) \, dz_1 \, . \end{aligned} \tag{3.8}$$

Actually both (3.7) and (3.8) are the same once one has remarked that for a probability density q on $\mathbb{R}^d$,

$$\int k\left(\int \mathbf{1}_{|z|\leq|y|}\, q(z)\, dz\right) q(y)\, dy \,=\, 1\,,$$

since $\int k(u)du = 1$. ♢

In strong form, equation (3.8) is written as

$$\partial_t f_1(z_1,t) \,+\, v_1.\nabla_{x_1} f_1(z_1,t) \,=\, Q(f_1,f_1)(z_1,t) \,-\, f_1(z_1,t)\,, \tag{3.9}$$

with

$$Q(f_1,f_1)(z_1,t) = \rho_1(x_1,t)\int k(M_{\rho_1}(x_1,|x_1-y|))\, f_1(y,v_1,t)\, dy\,. \tag{3.10}$$

This equation looks like a Boltzmann's equation. Actually "collisions" hold without collisions, since particles can change their velocities at any time.
It is worth noticing that this formulation is also the strong formulation of (3.7) provided one adds the constraint $\int f_1(z_1,t)\, dz_1 \,=\, 1$ for all t, which is implicit in the analysis of [BD16, DP18]. This conservation of mass condition is clearly explicit in (3.7).
We turn now to the study of (3.9). We have

Proposition 3.8. *Let $f(z,0)$ be a density of probability on $\mathbb{R}^{2d}$. Assume that k is Lipschitz continuous. Then, there exists a unique flow $t \mapsto f(z,t)$ of probability densities satisfying* (3.8) *with initial condition $f(z,0)$.*

Proof. The proof is more or less standard. We only give the main steps.
For the existence one uses an iteration scheme by solving inductively, for v fixed, $f^0(x,v,t) = f(x,v,0)$ and

$$\partial_t f^n(z,t) + v.\nabla_x f^n(z,t) + f^n(z,t) = Q(f^{n-1},f^{n-1})(z,t)\,.$$

An explicit solution is given by

$$f^n(z,t) = e^{-t}\, f(x-tv,v,0) \,+\, \int_0^t Q(f^{n-1},f^{n-1})(x-(t-s)v,v,s)\, e^{-(t-s)}\, ds\,. \tag{3.11}$$

One checks inductively that positivity and total mass equal to 1 is preserved, using the same argument as in remark 3.7.
Since k is Lipschitz, the existence of a limit in $\mathbb{L}^1(dz)$ follows for t small enough as for Banach (or Picard) fixed point theorem, thanks to contraction.
Indeed, if k is L-Lipschitz, f and f' are two Probability densities for which one define M_ρ and $M_{\rho'}$ it holds first

$$\begin{aligned} |M_\rho - M_{\rho'}|(x,|x-x'|) \;&\leq\; \int \mathbf{1}_{|y-x|<|x-x'|}\, |f(y,u,t) - f'(y,u,t)|\, dydu \\ &\leq\; \|\, f(.,t) - f'(.,t)\,\|_{\mathbb{L}^1(dz)}\ . \end{aligned}$$

Next

$$\begin{aligned} \| Q(f,f) - Q(f',f') \|_{\mathbb{L}^1(dz)} \quad \leq \quad & \int k(M_\rho(x,|x-x'|))\,|f(z,t)-f'(z,t)|\,f(z',t)\,dzdz' \\ & + \int k(M_\rho(x,|x-x'|))\,|f(z',t)-f'(z,t)|\,f'(z,t)\,dzdz' \\ & + \int \left|(k(M_\rho)-k(M_{\rho'}))(x,|x-x'|)\right|\, f'(z,t)f'(z',t)\,dzdz'. \end{aligned}$$

The first two terms are bounded by

$$\| k \|_\infty \| f(.,t) - f'(.,t) \|_{\mathbb{L}^1(dz)}$$

and the third one by

$$L \ \| f(.,t) - f'(.,t) \|_{\mathbb{L}^1(dz)}$$

since mass is preserved. This proves the contraction property in (3.11) for t small enough. Uniqueness is proved similarly by using Lipschitz contraction. One can then extends the solution to any time, this time by time induction since mass is preserved. □

Remark 3.9. If the initial $f(.,0)$ is compactly supported by $B(0,R)\times B(0,S)$, the previous proof shows that $f(.,t)$ is compactly supported by $B(0,R+tS)\times B(0,S)$. ♢

We may thus deduce from all what precedes

Corollary 3.10. *Assume that assumption 3.1 is satisfied, $k^N = k$ is Lipschitz continuous, $\lambda(N) = N$ and that the initial density p in assumption 3.1 is compactly supported. If for all t the system $(x^N(t))$ is chaotic, then for any $j \in \mathbb{N}^*$, $f^N_{1,\dots,j}$ weakly converges as $N \to +\infty$ to $f_1^{\otimes j}$ where f_1 is the unique solution of (3.7) with initial condition p. Conversely if the latter convergence is true for all j, then $(z^N(t))$ is chaotic for all t.*

What now remains to do is prove that the propagation of chaos holds.

Since the model looks like a particle system approach of some Boltzmann's equation, one should call upon the existing results in the literature, for instance [GM99] or the results for the "Paveri-Fontana" traffic model obtained in [MMJD16]. The latter model is close to the B-D model, but the transition probabilities $\pi_{i,j}$ are obtained via pairwise interactions. In our case a direct proof of propagation of chaos induces a difficult combinatorial study.
There are only few cases where propagation of chaos is known to hold :

- the case where $k = 1$ is standard in mean field theory,
- when k is linear it can be easily proved by using De Finetti's theorem (unpublished).
- In [DP18], Degond and Pulvirenti directly prove the convergence for a typically polynomial k using the whole BBGKY hierarchy. Propagation of chaos thus follows.

4. 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 2.3 and 2.4.1 show that they cannot be adapted to the self normalized metric interaction introduced above. For instance (2.13) and (2.15) 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.

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

[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.

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

[DMMPR10] P. Del Moral, L. Miclo, F. Patras, and S. Rubenthaler. The convergence to equilibrium of neutral genetic models. *Stoch. Anal. Appl.*, 28(1):123–143, 2010.

[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.

[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.

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

[MMJD16] J. Mint-Moustapha, B. Jourdain, and D. Daucher. A probabilistic particle approximation of the "Paveri-Fontana" kinetic model of traffic flow. *SMAI J. Comput. Math.*, 2:229–253, 2016.

[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.

Email 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.

Email address: fanny.delebecque@math.univ-toulouse.fr

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

Email address: olympio.hacquard@ens-paris-saclay.fr