Control Theory Tech, Vol. 15, No. 4, pp. 245–257, November 2017

Control Theory and Technology

http://link.springer.com/journal/11768

On the minimum number of neighbors needed for consensus of flocks

Chen CHEN [1], Ge CHEN [2], Lei GUO [2†]

1.*Noah's Ark Laboratory, 2012 Lab, Huawei Technologies Co. Ltd, Beijing 100085, China;*

2.*LSC & NCMIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China*

Received 10 September 2017; revised 9 October 2017; accepted 9 October 2017

Abstract

This paper investigates consensus of flocks consisting of n autonomous agents in the plane, where each agent has the same constant moving speed v_n and updates its heading by the average value of the k_n nearest agents from it, with v_n and k_n being two prescribed parameters depending on n. Such a topological interaction rule is referred to as k_n-nearest-neighbors rule, which has been validated for a class of birds by biologists and verified to be robust with respect to disturbances. A theoretical analysis will be presented for this flocking model under a random framework with large population, but without imposing any *a priori* connectivity assumptions. We will show that the minimum number of k_n needed for consensus is of the order $O(\log n)$ in a certain sense. To be precise, there exist two constants $C_1 > C_2 > 0$ such that, if $k_n > C_1 \log n$, then the flocking model will achieve consensus for any initial headings with high probability, provided that the speed v_n is suitably small. On the other hand, if $k_n < C_2 \log n$, then for large n, with probability 1, there exist some initial headings such that consensus cannot be achieved, regardless of the value of v_n.

Keywords: k-nearest-neighbor, consensus, topological interaction, random geometric graph

DOI https://doi.org/10.1007/s11768-017-7097-7

1 Introduction

Collective behavior, which is widely observed in physical, chemical, social, and biological systems, does not seem to have global information transfer among the components of the system, but the overall can display some highly ordered behavior. From a scientific point of view, how locally interacting rules lead to ordered phenomena is a basic and challenging issue to be understood. In recent years, much attempt has been made to observe, describe and model the collective behavior ranging from molecules to groups of animals, trying to find the mechanism behind these phenomena [1–24]

†Corresponding author.

E-mail: lguo@amss.ac.cn. Tel.: 86-10-62620724; fax: 86-10-62626870.

This paper is dedicated to Professor T. J. Tarn on the occasion of his 80th birthday. This work was supported by the National Natural Science Foundation of China (No. 91427304, 61673373, 11688101), the National Key Basic Research Program of China (973 program) (No. 2014CB845301/2/3), and the Leading Research Projects of Chinese Academy of Sciences (No. QYZDJ-SSW-JSC003).

etc. To mimic the flock of birds, Reynolds proposed a Boid model which employs three simple local interaction rules: flocking cohesive, collision avoidance and velocity alignment [1]. These rules have been realized by (discrete or continuous) dynamical systems [2, 3]. To carry out a theoretical study on Boid model, the authors of [3,4] constructed some collective potential functions to characterize the attraction and repulsion among agents, adopted consensus algorithm to achieve velocity consensus, and provided the corresponding stability analysis. Note that in many practical systems, the velocities of the neighboring individuals tend to become parallel to each other, and such motion seems to be safe, stable and collision-free [5]. Consequently, the velocity alignment or consensus problem has drown wide attention from researchers in recent years. In particular, Vicsek et al. proposed a simplified Boid model [6], which consist of n autonomous agents moving in the plane with the same constant speed and with the heading of each being updated by the average of its geometric neighbors'. This model has also been generalized to other forms together with numerical simulations, see, e.g., [7] and [8]. To analyze the so-called Vicsek's model introduced in [6], Jadbabaie et al. [9] further simplified the Vicsek model, and initiated a theoretical study by resorting to some connectivity assumptions on the system dynamics, followed by many researchers, see, e.g., [10] and [11]. Another typical flocking model is the so-called Cucker-Smale model [12], in which the interaction between two agents is a monotonously decreasing function with respect to their distance. Some variants of this model can be found in e.g., [13], and the convergence time of flocks has been studied in-depth in [14].

We would like to point out that, in most of the local interaction-based flocking models studied in the existing literature, the "neighbor" is often defined via the geometrical distance, that is, each agent's neighbors are defined as the ones within a prescribed geometric distance from it, as can be seen from the models mentioned above. However, the geometric distance cannot cover all the interesting situations either practically or theoretically.

Take a group of animals for instance, as pointed out by [25], under the geometric interaction rule, once the inter-individual distance became larger than the prescribed geometric distance, there would be no interaction and stragglers would "evaporate" from the aggregation, and so, the cohesion in the case of strong perturbations or predators invasion cannot be kept. Hence, whether the practical interaction is indeed determined by a geometric distance remains to be a question.

In fact, a group of scientists has carried out an experimental observation for starlings within flocks, with a significant finding that the starlings in huge flocks interact with a fixed number (6 or 7) of nearest individuals (i.e., "topological distance") [25], instead of those within a given geometric distance. Moreover, they have also made comparisons with the geometric distance based rules via numerical simulations, and revealed that the topological interaction significantly outperforms the geometrical interaction towards maintaining the connectivity of the flocks. Based on this, they claimed that "topological interaction is the only mechanism granting the robust cohesion with higher biological fitness" [25]. Such interaction have also been valided in [26] through establishing a maximum entropy model to empirical data. Furthermore, Ballerini et al. [25] also discussed why the neighbor's number is 6 or 7, which may be explained as follows: on one hand, birds cannot distinguish and track too many individuals due to the limited visual capacity and this has been validated in trained pigeons [27] and shoaling fish [28]. On the other hand, the number of 6–7 is the result of some optimization. In [29], the authors showed that the flock interaction networks with 6–7 neighbors optimizes the trade-off between group cohesion and individual effort.

It is worth mentioning that the topological interaction has also been studied in wireless networks, where nodes are located randomly on the plane according to a Poisson point process and each node is connected to a fixed number of nearest ones. In order to conserve energy and reduce disturbance from communication noise, it is meaningful to find the minimum number of neighbors that each node should link to, so that the overall network becomes connected. To address this issue, [30] pioneered the investigation of the connectivity of a random topological graph denoted by $G(n, m_n)$ with n nodes and m_n neighbors, and successfully proved that there exist two constants $0 < c_1 < c_2$ such that $G(n, c_1 \log n)$ is disconnected and $G(n, c_2 \log n)$ is connected with high probability. This result has later been refined in [31].

Hereinafter, the topological interaction rule or the "k-nearest-neighbor rule" will be used exclusively in the paper. We will investigate the consensus property of flocks in the following scenario: n autonomous agents move in the plane with the same constant speed v_n and with heading of each agent updated according to the averaged direction of its k_n nearest neighbors. This model is obviously related to but different from those with geometric distance based rules, including the above-mentioned Vicsek's model and its variations.

We remark that, to the best of our knowledge, there are few theoretical results on the flocking model with k-nearest-neighbor rule, although there is a vast literature on the related geometric distance based flocking models, see, e.g., [3, 4], [9–11]and [18–21]. The theoretical difficulties in the current paper lie at least in the following two aspects: one is that some kind of connectivity is required in the theoretical investigation of consensus which is also adopted as a basic assumption in [3, 4] and [9–11]. This is a well-known "bottleneck" problem, because the topology of the flocks is time-varying and state-dependent, and thus, how to sidestep or verify the connectivity conditions turns out to be a difficult and challenging mathematical problem. Another difficulty is that the underlying topological graphs are directed due to the k-nearest-neighbor rule, and therefore a lot of nice properties with beauty of symmetry for undirected graphs are lost, which brings a big difference from the undirected case as in Vicsek's model. Therefore, the results and methods used in [19–21] for flocks with undirected graphs cannot be directly applied here, and new methods in analyzing nonlinearly coupled dynamical flocks with directed position graphs should be developed. This constitutes one new contribution of the paper, with parts of the results presented in [22].

Next, since the neighbors number k_n can be treated as a parameter of the system, then how k_n affects the ordering phenomenon on earth? It is obvious that if $k_n = 0$, the system cannot achieve consensus in general whatever the speed is, but if $k_n = n$, then consensus would be achieved immediately. Thus, it is nontrivial to ask "dose a critical neighbor number k_n, or at least a critical order exist for the emergence of consensus?" From an engineering viewpoint, the critical number of neighbors also plays an important role in designing distributed cooperative control or communication networks. Recall that for a static random k-nearest-neighbor graph with n nodes to be asymptotically connected, the order $\Theta(\log n)$ neighbors are kind of necessary and sufficient [30]. We would then naturally expect and will actually prove that for the current nonlinear dynamical system, similar consensus results can also be established, under some conditions on the speed and initial settings. This will constitutes another original contribution of the work, with parts of the results presented in [23] without full proof.

The rest of this paper is organized as follows: some notations used in the paper are defined in Section 2. In Section 3, we will present the formulation of the problem and the main results, with their detailed proofs given in Section 4 and the Appendix. A simulation example will be showed in Section 5, followed by the concluding remarks in Section 6.

2 Some basic definitions

Graph theory plays an important role in the research of dynamical network and some basic notations and concepts deserve to point out first. A directed graph (digraph) $G = \{V, E\}$ is composed of a vertex (or node) set $V = \{1, 2, \dots, n\}$ and an edge set $E = \{(i, j) \subseteq V \times V\}$, where $(i, j) \in E$ is an directed edge from i to j, and also means that j is a neighbor of i. If vice versa, then G is undirected. For any vertex $i \in V$, if $(i, i) \in E$, then it is called a loop of G. A path of length l in G that joins vertexes i and j means that there is a sequence of vertexes $i_1, i_2, \dots, i_l$ such that $(i_m, i_{m+1}) \in E$, $0 \leqslant m \leqslant l$ with $i_0 = i$, $i_{l+1} = j$. A digraph is called strongly connected if for any two different vertexes i and j, there always exists a path from i to j. If a strongly connected graph is undirected, then it is called connected. A digraph is said to have a spanning tree if and only if there exists a vertex $i \in V$, called root, such that there is a path from i to any other vertex. The adjacency matrix $M = (m_{ij})_{n\times n}$ of graph G is a $0-1$ matrix, where $m_{ij} = 1$ if and only if $(i, j) \in E$.

In this article, we use the following standard notations. The symbol := denotes definition. The set of real numbers is denoted by $\mathbb{R}$ and the set of non-negative integers is denoted by $\mathbb{Z}_+$. For a set U, $|U|$ denotes the cardinality of U. Given $t \in \mathbb{R}$, we write $\lfloor t\rfloor$ for the value of t rounded down to the nearest integer, and $\lceil t\rceil$ for the value of t rounded up to the nearest integer. For integers $n \geqslant m \geqslant 1$, $C_n^m := \dfrac{n!}{m!(n-m)!}$. Hereinafter, all logarithms are base e.

Suppose $\{a_n\}_{n\geqslant 1}$ and $\{b_n\}_{n\geqslant 1}$ are sequences of positive real numbers with $b_n > 0$ for all n. We write $a_n = \mathrm{O}(b_n)$ if $\limsup\limits_{n\to\infty}\dfrac{a_n}{b_n} < \infty$, and write $a_n = o(b_n)$ if $\limsup\limits_{n\to\infty}\dfrac{a_n}{b_n} = 0$. We write $a_n = \Theta(b_n)$ if both $a_n = \mathrm{O}(b_n)$ and $b_n = \mathrm{O}(a_n)$.

For all $x \in \mathbb{R}^d$ with $x := (x_1, \dots, x_d)$, the so-called l_p norms of x, $\|\cdot\|_p$, are defined for $1 \leqslant p < \infty$ by

$$\|(x_1, \dots, x_d)\|_p := \Big(\sum_{i=1}^{d} |x_i|^p\Big)^{1/p}. \tag{1}$$

The l_2 norm is also denoted by the Euclidean norm. Let $B(x,r) := \{y \in \mathbb{R}^2 : \|x-y\|_2 \leqslant r\}$ denotes the ball centered at x with radius r. The following notation is quite important in our paper, we highlight it.

Definition 1 We say that a sequence of events E_n, $n \geqslant 1$ occur with high probability (w.h.p.) if $\lim\limits_{n\to\infty} \Pr[E_n] = 1$. Moreover, we say E_n occur with probability 1 for large n if almost surely $E_n^{\rm c}$, $n \geqslant 1$ only happen finite times in terms of n.

3 Model and main results

3.1 Model

Let us assume that n autonomous agents move in the plane with the same speed v_n ($v_n > 0$) but with different headings. At any time $t \in \mathbb{Z}_+$, the position and heading of agent i are denoted by $X_i(t)$ ($\in \mathbb{R}^2$) and $\theta_i(t)$ ($\in (-\pi,\pi]$) respectively. The distance between agents i and j is denoted by $d_{ij}(t) := \|X_i(t) - X_j(t)\|_2$. For any agent i ($1 \leqslant i \leqslant n$), the neighbors of i are defined as the nearest k_n individuals from it, where k_n is a prescribed value depending on n, and the neighbor set of i at t is denoted by $\mathcal{N}_i(t)$. If at time t, there is more than one agents who are eligible to be the k_n-th nearest one from agent i, then i chooses one randomly among them. In particular, we define that each agent is a neighbor of itself. For arbitrary $t \in \mathbb{Z}_+$ and $1 \leqslant i \leqslant n$, the position's updating rule for i is as follows:

$$X_i(t) = X_i(t-1) + v_n(\cos\theta_i(t), \sin\theta_i(t)) \tag{2}$$

with $\theta_i(t)$ updated by

$$\theta_i(t) = \frac{1}{|\mathcal{N}_i(t-1)|} \sum_{j\in\mathcal{N}_i(t-1)} \theta_j(t-1). \tag{3}$$

Since $|\mathcal{N}_i(t)| \equiv k_n$, then we can rewrite (3) as follows:

$$\theta_i(t) = \frac{1}{k_n} \sum_{j\in\mathcal{N}_i(t-1)} \theta_j(t-1). \tag{4}$$

This paper will mainly investigate the consensus property of the model (2)–(4). Following Tang and Guo [19], we give the definition of "consensus".

Definition 2 If there exists a constant $\bar{\theta} \in (-\pi,\pi]$ such that $\lim\limits_{t\to\infty} \theta_i(t) = \bar{\theta}$, $\forall 1 \leqslant i \leqslant n$, then we say the model (2)–(4) achieve consensus.

3.2 Main results

For $t \in \mathbb{Z}_+$, let

$$\mathcal{X}_n(t) := \{X_1(t), X_2(t), \dots, X_n(t)\}$$

be the set including the positions of n agents at t. To analyze the consensus behavior, we will use the following graph sequence $\{\mathcal{G}(t), t = 0,1,\dots\}$ to describe the relationship among neighbors. For $t \in \mathbb{Z}_+$, define

$$\mathcal{G}(t) = \mathcal{G}_{n,k_n}(t) := \mathcal{G}(\mathcal{X}_n(t), \mathcal{E}(t))$$

to be the position graph of the model at t, where $(i,j) \in \mathcal{E}(t)$ if and only if $j \in \mathcal{N}_i(t)$. Note that $(i,i) \in \mathcal{E}(t)$ for all $1 \leqslant i \leqslant n$, since self-loop is contained. It worth mentioning that the graphs formed in this way are directed. Denote $P(t) = P_{n,k_n}(t)$ as the average matrix of the graph $\mathcal{G}(t)$, i.e.,

$$(P(t))_{ij} = \begin{cases} 1/k_n, & \text{if } (i,j) \in \mathcal{E}(t), \\ 0, & \text{else}, \end{cases} \quad \forall i,j = 1,2,\dots,n.$$

It can be seen immediately that $P(t)$ is a stochastic matrix. Set $\theta(t) := (\theta_1(t), \theta_2(t), \dots, \theta_n(t))^\tau$, then the iteration rule of the model based on (2) and (4) can be rewritten as the following compact matrix form:

$$\begin{cases} \theta(t) = P(t-1)\theta(t-1), \\ X_i(t) = X_i(t-1) + v_n(\cos\theta_i(t), \sin\theta_i(t)). \end{cases} \tag{5}$$

We will proceed our analysis under the assumption that the initial positions $\{X_i(0) \in \mathbb{R}^2, 1 \leqslant i \leqslant n\}$ are independently and uniformly distributed in the unit square $[0,1]^2$ with the initial headings $\{\theta_i(0) \in \mathbb{R}^2, 1 \leqslant i \leqslant n\}$ arbitrarily distributed in $(-\pi,\pi]$. Then the position graph at the initial time instant $\mathcal{G}(0)$ is the random k_n-nearest neighbor graph, which has been investigated in [30] and [31], etc.

Theorem 1 For the flocking model (2)–(4), suppose that the initial positions of n agents are uniformly and independently distributed in the unit square $[0,1]^2$. Then there exist some constants $0 < C_2 < C_1$ and $C \in (0,1)$, such that the following two assertions are true:

i) If $k_n > C_1 \log n$ and $v_n = \mathrm{O}(C^{\sqrt{\frac{n}{k_n}}})$, then the system will achieve consensus w.h.p. for arbitrary initial headings.

ii) If $k_n \leqslant C_2 \log n$, then for large n with probability 1, there exist some initial headings' configurations such that the flocking model cannot achieve consensus for any speed $v_n \geqslant 0$.

Remark 1 The precise value of the constants C_1, C_2, C can be found from the proofs of Theorem 3.1 in the next section. Here, we just mention that some calculations can give rough estimates for C_1 and C_2 as 50 and 0.1360, respectively.

4 Proof of Theorem 1

Theorem 1 consists of two parts whose proof will be proceeded in Section 4.1 and 4.2, respectively.

4.1 Analysis of Theorem 1 i)

Throughout the proof, all analysis will be carried out under the assumption that the positions of all the agents are independently and uniformly distributed in $[0,1]^2$, then the initial random k_n-nearest-neighbor graph will have some nice properties.

Let $[0,1]^2$ be divided equally into $M_{n,K} := \left\lceil \sqrt{\frac{n}{Kk_n}} \right\rceil^2$ small squares whose side length is defined as $a_{n,K} := 1/\left\lceil \sqrt{\frac{n}{Kk_n}} \right\rceil$ and set $L_{n,K} := \left\lceil \frac{1}{a_{n,K}} \right\rceil = \left\lceil \sqrt{\frac{n}{Kk_n}} \right\rceil$, as depicted in Fig. 1, where $K > 0$ is a tunable parameter. We label the small squares as $S^i_{n,K}, i = 1, 2, \ldots, M_{n,K}$, from left to right, and from bottom to top. Denote by $N^i_{n,K}$, the number among the n agents, that fall into the square $S^i_{n,K}$. Then the following estimation for $N^i_{n,K}$ holds.

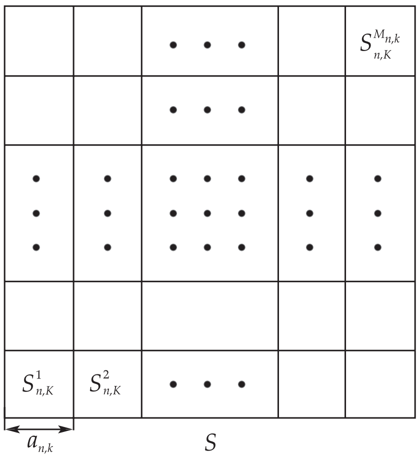

Fig. 1 The unit square $[0,1]^2$ is equally divided to $M\left\lceil \sqrt{\frac{n}{Kk_n}} \right\rceil^2$ small squares which are labeled as $S^i_{n,K}, i = 1, 2, \ldots, M_{n,K}$, from left to right, and from bottom to top.

Lemma 1 Assume that $Kk_n > \frac{\log n}{\log(4/e)}$, and let $\mu^0 \in (0,1)$ be the sole root of the equation

$$-\mu + (1+\mu)\log(1+\mu) = \frac{\log n}{Kk_n} \tag{6}$$

with respect with μ. Then for any $\mu > \mu^0$:

$$\max_i |N^i_{n,K} - Kk_n| \leqslant \mu Kk_n \quad \text{w.h.p.} \tag{7}$$

Proof This result can be obtained by the method of the proof of Lemma 3.1 in [30] with slight adjustment. □

Before proceeding further, define the large deviations rate function $H : [0,\infty) \to \mathbb{R}$ by $H(0) = 1$ and

$$H(a) = 1 - a + a\log a, \quad a > 0. \tag{8}$$

Note that $H(1) = 0$ and the unique turning point of H is the minimum at 1. Also $H(a)$ is increasing on $(1,\infty)$.

For any fixed agent i, the following lemma estimates the number of agents falling into a ball centered at i, see Fig. 2.

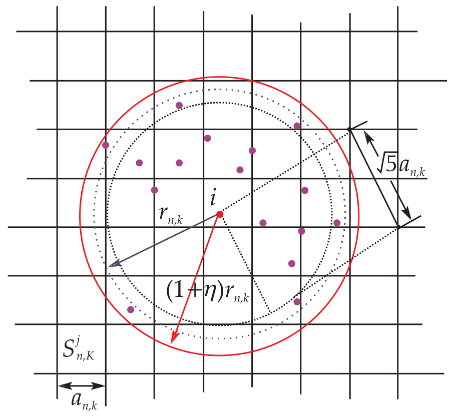

Fig. 2 The ball with red boundary represents $B(X_i(0), (1+\eta)r_{n,K})$ and $r_{K,\eta} = \frac{1}{1-\eta}\sqrt{5}a_{n,K}$.

Lemma 2 Suppose r_n ($n \geqslant 1$) is a positive real number sequence satisfying $\pi n r_n^2 \geqslant 2\log n$. Then with probability 1,

$$\max_i |\mathcal{X}_n(0) \cap B(X_i(0), r_n)| \leqslant a_n n\pi r_n^2(1+o(1)) \tag{9}$$

for large n, where a_n is the solution to the following equation:

$$H(a) = \frac{\log n}{\pi n r_n^2}, \quad a > 1.$$

Proof This result can be deduced directly from Theorem 6.14 of [32]. □

Lemma 3 Pick arbitrary $0 < \eta < 1$ and define $r_{K,\eta} = r_{n,K,\eta} := \sqrt{\frac{5Kk_n}{n}}/(1-\eta)$. If $k_n > C_1 \log n$ with $C_1 = (5\pi \times 1.23)/\log(4/e)$, then we can find some K and η such that $Kk_n > \frac{\log n}{\log(4/e)}$ and the following assertion is true:

$$\max_i |\mathcal{X}_n(0) \cap B(X_i(0), (1+\eta)r_{K,\eta})| \leqslant k_n \quad \text{w.h.p.} \tag{10}$$

Proof The relationship among $a_{n,K}$, $r_{K,\eta}$ and $(1+\eta)r_{K,\eta}$ can be seen in Fig. 2.

Let $k_n \geqslant C_1 \log n$ and set $K = 1/(5\pi \times 1.23)$, then by the value of C_1, we have $KC_1 > \frac{1}{\log(4/e)}$, which is followed by $Kk_n > \frac{\log n}{\log(4/e)}$.

By computing, we have

$$n\pi[(1+\eta)r_{K,\eta}]^2 = 5\pi K\Big(\frac{1+\eta}{1-\eta}\Big)^2 k_n > 2\log n,$$

then the condition of Lemma 2 is satisfied. Applying Lemma 2 directly, we can obtain that w.h.p.

$$\max_i |\mathcal{X}_n(0) \cap B(X_i(0), r_{K,\eta})| \leqslant 5\pi K\Big(\frac{1+\eta}{1-\eta}\Big)^2 a_n k_n, \tag{11}$$

where a_n is a root of $H(a_n) = \log n / (5\pi K\big(\frac{1+\eta}{1-\eta}\big)^2 k_n)$. Again by the value of K and C_1, we have $H(a_n) < \frac{\log(4/e)}{5\pi}$, and we can also verify that $H(1.23) > \frac{\log(4/e)}{5\pi}$. Since $H(a)$ is monotonously increasing on $a \in (1,\infty)$, we can get $1 < a_n < 1.23$, then there always exists some $0 < \eta < 1$ such that $5\pi K\big(\frac{1+\eta}{1-\eta}\big)^2 a_n = 1$. Combing this with (11), we have

$$\max_i |\mathcal{X}_n(0) \cap B(X_i(0), r_{K,\eta})| \leqslant k_n, \quad \text{w.h.p.} \qquad \square$$

From now on, when we refer to $r_{K,\eta}$, it means the same as that in Lemma 3. Next, we define a new graph based on the agents' initial positions.

Definition 3

$$\mathcal{G}_K = \mathcal{G}_{n,k_n,K} := \mathcal{G}(\mathcal{X}_n(0), \mathcal{E}_K),$$

where

$$\mathcal{E}_K = \mathcal{E}_{n,k_n,K} := \{(i,j) : X_j(0) \in \mathcal{X}_n(0) \cap B(X_i(0), (1-\eta)r_{K,\eta})\}.$$

Remark 2 Evidently, $\mathcal{G}_K$ is undirected. By the construction of $r_{K,\eta}$, it can be seen that $(1-\eta)r_{K,\eta} = \sqrt{\frac{5Kk_n}{n}} = \sqrt{5}a_{n,K}$, which is equal to the diagonal line length of two adjacent small squares as depicted in Fig. 2. We also provide an example of $\mathcal{G}_K$ with $n = 21, k_n = 5$ in Fig. 3.

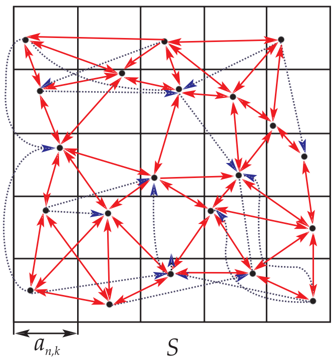

Fig. 3 Here is an example of $\mathcal{G}(0)$ and $\mathcal{G}_K$ with $n = 21$ and $k_n = 5$. We use arrows (both red and blue dotted arrows)to represent edges in $\mathcal{E}(0)$, which are defined according to the 5-nearest-neighbor rule, where double arrows represent the mutual neighbor relationship and one-way arrows represent the unidirectional neighbor relationship. When two agents' distance is smaller than $\sqrt{5}a_{n,K}$, then there is a red double arrow between them which belong to $\mathcal{E}_K$.

Throughout the sequel, let $k_n > C_1 \log n$. Fix $K^* = 1/(5\pi \cdot 1.23)$ and η^* such that Lemma 3 holds. For this K^*, we can also find some $\mu^* \in (0,1)$ such that Lemma 1 holds. And the variables L_{n,K^*}, r_{K^*,η^*} are as defined in the above. The following Lemmas 4-8 are all based on this premise.

Lemma 4 Suppose that $k_n > C_1 \log n$, then w.h.p. $\mathcal{G}_{K^*} \subset \mathcal{G}(0)$ and $\mathcal{G}_{K^*}$ is connected.

Proof According to Lemma 3, we obtain that w.h.p.

$$\forall i, \ \ |\mathcal{X}_n(0) \cap B(X_i(0), (1+\eta^*)r_{K^*,\eta^*})| \leqslant k_n.$$

Then by k_n-nearest-neighbor rule, we have w.h.p.

$$\forall i, \ \ \mathcal{X}_n(0) \cap B(X_i(0), (1+\eta^*)r_{K^*,\eta^*}) \subset \mathcal{N}_i(0). \tag{12}$$

Pick arbitrary $(i,j) \in \mathcal{E}_{K^*}$, by the construction of $\mathcal{E}_{K^*}$, it can be seen that

$$\begin{aligned} d_{ij}(0) = \|X_i(0) - X_j(0)\|_2 &\leqslant \sqrt{5}a_{n,K^*} \\ &= (1-\eta^*)r_{K^*,\eta^*} < (1+\eta^*)r_{K^*,\eta^*}. \end{aligned} \tag{13}$$

Combining (13) with (12), we can obtain that w.h.p. for arbitrary i, j,

$$j \in \mathcal{N}_i(0),$$

which means

$$(i, j) \in \mathcal{E}(0).$$

Then we have $\mathcal{E}_{K^*} \subset \mathcal{E}(0)$ w.h.p., which is followed by $\mathcal{G}_{K^*} \subset \mathcal{G}(0)$ w.h.p.

Now we prove the connectivity of $\mathcal{G}_{K^*}$. Notice that $\mathcal{G}_{K^*}$ is actually a standard random geometric graph with radius $\sqrt{5}a_{n,K^*} = \sqrt{\frac{5K^*k_n}{n}} = \sqrt{\frac{5K^*\pi k_n - \log n + \log n}{\pi n}}$. And it has been proved in [33] that the random geometric graph with radius $\sqrt{\frac{c_n + \log n}{\pi n}}$ will be connected w.h.p., if and only if $c_n \to \infty$. Hence, $\mathcal{G}_{K^*}$ is connected w.h.p., by the fact that $K^*k_n > 1/\log(4/e)$. □

Lemma 5 Assume that there exists a virtual vertex v^i_{n,K^*} in the center of each square S^i_{n,K^*}, $i = 1, 2, \dots, M_{n,K^*}$, and a virtual edge $(v^i_{n,K^*}, v^j_{n,K^*})$ if either $i = j$ or $\|v^i_{n,K^*} - v^j_{n,K^*}\|_2 = a_{n,K^*}$. Then for arbitrary i, j, the number of virtual undirected paths with length $2(L_{n,K^*}-1)$ joining v^i_{n,K^*} and v^j_{n,K^*} is not smaller than $C^{L_{n,K^*}}_{2(L_{n,K^*}-1)}$.

The proof of Lemma 5 is given in the appendix.

Let M_{K^*} be the adjacency matrix of the graph $\mathcal{G}_{K^*}$, we have:

Lemma 6 Suppose that $k_n \geqslant C_1 \log n$, then

$$\begin{aligned}&(M_{K^*})^{2(L_{n,K^*}-1)}\\ &\geqslant C^{L_{n,K^*}}_{2(L_{n,K^*}-1)}((1-\mu^*)K^*k_n)^{2(L_{n,K^*}-1)-1}\mathbf{1}\mathbf{1}^\tau, \text{ w.h.p.}\end{aligned}$$

where $\mathbf{1}$ is all 1's vector.

Proof For any i, j, $(M_{K^*})^{2(L_{n,K^*}-1)}_{ij}$ represents the total number of paths from agent i to agent j in $\mathcal{G}_{K^*}$ with length $2(L_{n,K^*} - 1)$. Assume that i and j locate in $S^{l_i}_{n,K^*}$ and $S^{l_j}_{n,K^*}$, respectively. From Lemma 5, the total number of virtual undirected path from $v^{l_i}_{n,K^*}$ to $v^{l_j}_{n,K^*}$ with length $2(L_{n,K^*} - 1)$ is at least $C^{L_{n,K^*}}_{2(L_{n,K^*}-1)}$. And by Lemma 1, each virtual directed path can be substituted by $((1-\mu^*)K^*k_n)^{2(L_{n,K^*}-1)-1}$ real paths in $\mathcal{G}_{K^*}$ w.h.p., which derives Lemma 6 immediately. □

Lemma 7 Suppose that $k_n > C_1 \log n$. If $\mathcal{G}_{K^*} \subset \mathcal{G}(s)$ on $s \in \mathbb{Z}_+ \cap [t+1, t+2(L_{n,K^*}-1)]$, then w.h.p.,

$$\begin{aligned}&P(t+2(L_{n,K^*}-1)) \times \dots \times P(t+1)\\ &\geqslant C^{L_{n,K^*}}_{2(L_{n,K^*}-1)} \times \frac{1}{k_n}((1-\mu^*)K^*)^{2(L_{n,K^*}-1)-1}\mathbf{1}\mathbf{1}^\tau.\end{aligned}$$

Proof By the assumption that $\mathcal{G}_{K^*} \subset \mathcal{G}(s)$, we have $M(t) \geqslant M_{K^*}$ on $s \in \mathbb{Z}_+ \cap [t+1, t+2(L_{n,K^*}-1)]$. Then Lemma 7 can be derived immediately noting that $P(t) = \frac{1}{k_n}M(t)$. □

Corollary 1 Under the same condition of Lemma 7, the following inequality holds w.h.p.:

$$P(t+2(L_{n,K^*}-1)g) \times \dots \times P(t+1) \geqslant \frac{1}{k_n}C^{\sqrt{\frac{n}{k_n}}},$$

where $C \in (0, 1)$ is a computable constant only depending on K^*, η^*, μ^*.

Proof As $n \to \infty$, $L_{n,K^*} = \Theta\Big(\sqrt{\frac{n}{\log n}}\Big) \to \infty$, therefore $C^{L_{n,K^*}}_{2(L_{n,K^*}-1)} \geqslant 2^{(L_{n,K^*}-1)}$. Plus the fact that $L_{n,K^*} = \sqrt{\frac{n}{K^*k_n}}$, the inequality can be obtained immediately. □

Lemma 8 Suppose that $k_n > C_1 \log n$. If for any $i \neq j$ and $t \in [0, T] \cap \mathbb{Z}_+$, the following inequality holds:

$$| d_{ij}(t) - d_{ij}(0) | < \eta^* r_{K^*,\eta^*}, \tag{14}$$

then w.h.p. $\mathcal{G}_{K^*} \subset \mathcal{G}(t)$ for all $t \in [0, T] \cap \mathbb{Z}_+$.

The proof of Lemma 8 is similar to that of Lemma 3.3 in [22], so we omit it due to space limitation.

Now we are ready to prove Theorem 1 i).

Proof of Theorem 1 i) Set $\varDelta_t := \max_{i,j}\{\theta_i(t) - \theta_j(t)\}$. By the heading iteration (4), it can be seen immediately that $\varDelta_t$ is monotonously decreasing with respect to t. Now we prove that w.h.p. $\varDelta_t$ is exponentially decreasing, that is,

$$\forall s,\ \ \varDelta_{s\cdot 2(L_{n,K^*}-1)} \leqslant \Big(1 - \frac{n}{k_n}C^{\sqrt{\frac{n}{k_n}}}\Big)^s \varDelta_0, \text{ w.h.p.,} \tag{15}$$

where $C \in (0, 1)$ is defined in Corollary 1.

The main idea to prove (15) is that once v_n is moderately small, $\mathcal{G}_{K^*}$, as a subgraph of $\mathcal{G}(0)$, can be maintained as the associated dynamical position graphs $\mathcal{G}(t)$ evolve, therefore a generic "convergence factor" of the corresponding stochastic matrices can be estimated only with respect to $\mathcal{G}_{K^*}$, then the convergence speed of $\varDelta_t$

can be computed. To this end, we need not only to verify the connectivity of position graphs but also to prove the headings' consensus at the same time on a bounded period of time and then repeat the process again and again. Similar proof line has been presented in [22], and we omit the details for saving space. □

4.2 Analysis of Theorem 1 ii)

In this part we will give the proof of Theorem 1 ii). To achieve this, we still focus on investigating the initial position graph $\mathcal{G}(0)$ and try to find moderately small k_n such that $\mathcal{G}(0)$ is disconnected.

Some new notations are introduced first. In subsequent paper, let $\mathcal{L}(A)$ denote the area for the set $A \subset \mathbb{R}^2$. For a point $x \in \mathbb{R}^2$ and a set $S \subset \mathbb{R}^2$, the distance and the biggest distance between x and S are denoted by $d(x, S) := \inf_{y\in S} \|x - y\|_2$ and $\mathrm{dia}(x, S) := \sup_{y\in S} \|x - y\|_2$, respectively. Pick $\lambda > 0$ arbitrarily, we write $\mathrm{Po}(\lambda)$ for any Poisson random variable with parameter λ. Define a Poisson point process $\mathcal{P}_\lambda$ by $\mathcal{P}_\lambda := \{Y_1, Y_2, \ldots, Y_{\mathrm{Po}(\lambda)}\}$, where $\{Y_1, Y_2, \ldots\}$ is the set of points independently and uniformly distributed in $[0, 1]^2$ and $\mathcal{P}_\lambda$ is independently of $\{Y_1, Y_2, \ldots\}$. For a set $A \subset [0, 1]^2$, $|\mathcal{P}_\lambda \cap A|$, the number of points lying in A is a Poisson random variable with parameter $\lambda\mathcal{L}(A)$. For any two sets $A_1, A_2 \subseteq [0, 1]^2$, if $\mathcal{L}(A_1 \cap A_2) = 0$, then the random variables $|\mathcal{P}_\lambda \cap A_1|$ and $|\mathcal{P}_\lambda \cap A_2|$ are mutually independent. This property is called *spatial independence* of a Poisson point process.

The following lemma will be useful.

Lemma 9 [31] Let $A_1, \ldots, A_r$ be disjoint regions of $\mathbb{R}^2$ and $\rho_1, \ldots, \rho_r \geqslant 0$ be real numbers such that $\rho_i\mathcal{L}(A_i)\lambda \in \mathbb{Z}_+$, where $\lambda > 0$. Then the probability that a Poisson process with intensity λ has precisely $\rho_i\mathcal{L}(A_i)\lambda$ points in each region A_i is

$$\exp\{\sum_{i=1}^{r}(\rho_i - 1 - \rho_i \log \rho_i)\lambda\mathcal{L}(A_i) + O(r \log_+ \sum \lambda\rho_i\mathcal{L}(A_i))\},$$

with the convention that $0 \log 0 = 0$, and $\log_+ x = \max(\log x, 1)$.

We redivide $[0, 1]^2$ as followes: let $[0, 1]^2$ be divided equally into $M_{n,K} := \lceil \sqrt{\dfrac{n}{K \log n}} \rceil^2$ small squares with side length

$$a_{n,K} = \sqrt{\frac{K \log n}{n}} \text{ and } L_{n,K} = \lceil \frac{1}{a_{n,K}} \rceil = \lceil \sqrt{\frac{n}{K \log n}} \rceil,$$

where $K > 0$ is a tunable parameter. The small squares are labeled as $S^i_{n,K}$, $i = 1, 2, \ldots, M_{n,K}$, from left to right, and from bottom to top. Now we construct some special position configurations, whose uses will be showed later.

Definition 4 ($\mathrm{Trap}^{\varepsilon}_{r_0}$) We call the configuration in Fig. 4 a $\mathrm{Trap}^{\varepsilon}_{r_0}$. It is a semi-disk D with center on x-axis and radius $5r_0$, which is located in one of the bottom squares, i.e., some $S^i_{n,K}$, where $i = 1, \ldots, L_{n,K}$, r_0 is pending. Inside D, A_1 is a concentric semi-disk with radius r_0, and A_2 is a concentric semi-annulus with radii r_0 and $3r_0$. The remaining region of D is denoted by A, which is divided into $N - 2$ small regions, i.e., $A = \bigcup_{3\leqslant i\leqslant N} A_i$, with each A_i of diameter at most εr_0.

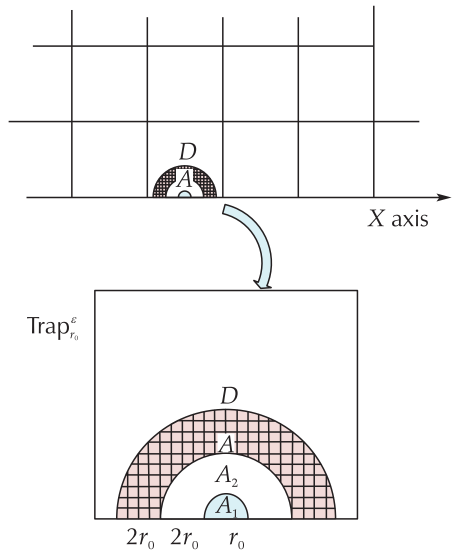

Fig. 4 A $\mathrm{Trap}^{\varepsilon}_{r_0}$ is a semi-disk D with center on x-axis and radius $5r_0$. Inside D, A_1 is a concentric semi-disk with radius r_0, and A_2 is a concentric semi-annulus with radii r_0 and $3r_0$. The remaining region of D is denoted by A, which is divided into $N - 2$ small regions, i.e., $A = \bigcup_{3\leqslant i\leqslant N} A_i$, with each A_i of diameter at most εr_0.

Definition 5 (the smallest cover in $\mathrm{Trap}^{\varepsilon}_{r_0}$) For any region $D' \subset A$, define

$$\mathcal{A}_{D'} = \bigcup_{3\leqslant i\leqslant N,\, A_i\cap D'\neq\emptyset} A_i \tag{16}$$

as D''s smallest cover in $\mathrm{Trap}^{\varepsilon}_{r_0}$, see Fig. 5.

Remark 3 It can be deduced immediately that $\mathcal{L}(\mathcal{A}_{D'}) = \sum_{A_i\subset\mathcal{A}_{D'}} \mathcal{L}(A_i)$ and $\mathcal{L}(D') \leqslant \mathcal{L}(\mathcal{A}_{D'})$. See Fig. 5.

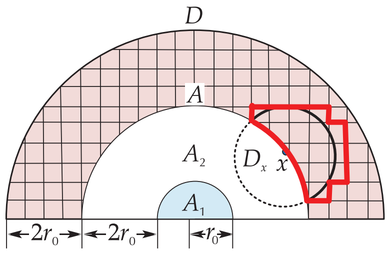

Fig. 5 $\mathcal{A}_{D_x\cap A}$ is the region with the red boundary, which is composed of all the A_i intersecting with $D_x \cap A$.

Definition 6 (k-filling event) We say a k-filling event occurs in $\mathrm{Trap}^{\varepsilon}_{r_0}$ if i) $|A_1 \cap \mathcal{X}_n(0)| \geqslant k$, ii) $|A_2 \cap \mathcal{X}_n(0)| = 0$ and iii) for arbitrary point $x \in A$, $|\mathcal{A}_{D_x\cap A} \cap \mathcal{X}_n(0)| \geqslant k$, where $D_x := B(x, r - (1+\varepsilon)r_0)$ and r is the distance between x and the center of D. See Fig. 6.

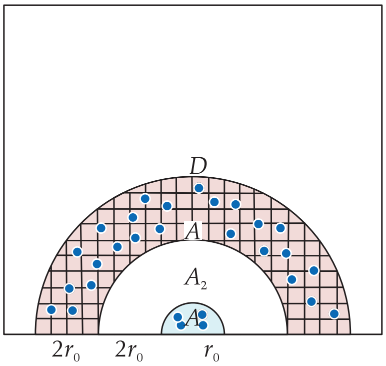

Fig. 6 A k-filling event occurs in $\mathrm{Trap}^{\varepsilon}_{r_0}$ if (i) $|A_1 \cap \mathcal{X}_n(0)| \geqslant k$, (ii) $|A_2 \cap \mathcal{X}_n(0)| = 0$ and (iii) for arbitrary point $x \in A$, $|\mathcal{A}_{D_x\cap A} \cap \mathcal{X}_n(0)| \geqslant k$, where $D_x := B(x, r - (1+\varepsilon)r_0)$ and r is the distance between x and the center of D.

Remark 4 Intuitively, (iii) guarantees that the points in A are relatively uniform with the "average density" in each ball D_x being larger than k.

Lemma 10 For some ε and r_0, if a k_n-filling event occurs in a $\mathrm{Trap}^{\varepsilon}_{r_0}$, then under k_n-nearest-neighbor rule, $\mathcal{G}(0)$ is disconnected.

Now for the bottom row squares of $[0,1]^2$: $\{S^i_{n,K}, i = 1,2,\dots,L_{n,K}\}$, where $K > \dfrac{1}{\log(4/e)}$, define the event

$$\begin{aligned} E^i_n :=\{&S^i_{n,K} \text{ contains a } \mathrm{Trap}^{\varepsilon}_{r_0}, \text{ in which} \\ &a\ k_n\text{-filling event occurs}\} \end{aligned}$$

and another event

$$\begin{aligned} \tilde{E}^i_n :=\{&S^i_{n,K} \text{ contains a } \mathrm{Trap}^{\varepsilon}_{r_0}, \text{ in which} \\ &|A_1 \cap \mathcal{X}_n(0)| = 2\rho\mathcal{L}(A_1) \geqslant k_n,\ |A_2 \cap \mathcal{X}_n(0)| = 0, \\ &|A_i \cap \mathcal{X}_n(0)| = \rho\mathcal{L}(A_i),\ 3 \leqslant i \leqslant N\}. \end{aligned}$$

Intuitively, 2ρ and ρ represent a kind of "densities" in A_1 and A, respectively, and the value of ρ can be chosen arbitrarily. Then we have

Lemma 11 For small enough ε, the following assertion holds:

$$\tilde{E}^i_n \subset E^i_n.$$

Set $\lambda^n_1 := n - n^{\frac{3}{4}}$ and $\lambda^n_2 := n + n^{\frac{3}{4}}$. Let $\mathcal{P}_{\lambda^n_1}$ and $\mathcal{P}_{\lambda^n_2}$ denote Poisson point processes in $[0,1]^2$ with parameters λ^n_1, λ^n_2, respectively. Define the event

$$\begin{aligned} F^i_n :=\{&S^i_{n,K} \text{ contains a } \mathrm{Trap}^{\varepsilon}_{r_0}, \text{ in which} \\ &|A_1 \cap \mathcal{P}_{\lambda^n_1}| = 2\rho\mathcal{L}(A_1) \geqslant k_n, |A_2 \cap \mathcal{P}_{\lambda^n_2}| = 0, \\ &|A_i \cap \mathcal{P}_{\lambda^n_1}| = \rho\mathcal{L}(A_i), 3 \leqslant i \leqslant N\}. \end{aligned}$$

Then we can get the following two lemmas:

Lemma 12

$$\Pr\Big(\bigcup_{i=1}^{L_{n,K}} \tilde{E}^i_n\Big) \geqslant \Pr\Big(\bigcup_{i=1}^{L_{n,K}} F^i_n\Big) - 2\mathrm{e}^{-n^{1/4}}. \tag{17}$$

Lemma 13 If $k_n < C_2 \log n$ with $C_2 = (2(\log\dfrac{25}{9} + 8\log\dfrac{25}{18}))^{-1}$, then for n large enough, $\bigcup\limits_{1\leqslant i\leqslant L_{n,K}} E^i_n$ happens with probability 1.

The proofs of Lemma 10-Lemma 13 are given in the appendix.

Proof of Theorem 1 ii) In Lemma 13, we have proved that under the condition $k_n < C_2 \log n$ and n large enough, at least one of the bottom row squares contains a $\mathrm{Trap}^{\varepsilon}_{r_0}$, in which a k_n-filling event occurs. Hence, we set the initial headings of the agents in A_1 to be $-\dfrac{\pi}{2}$, and the others to be $\dfrac{\pi}{2}$. For such case, the system cannot achieve consensus regardless of the value of v_n, which completes the proof. □

Remark 5 The idea stems from [21] but has a key difference rooted in the different interaction rules, and a much more complicated way is needed in our case to construct the disconnected component. The design of the k_n-filling event is partially inspired by [31], however we demand the configuration occur along the border of the $[0,1]^2$ due to the headings' specific configuration, while [31] does not, which makes construction design, connectivity analysis and probability computation quite different.

5 Simulations

In this section, we provide a simulation example. Here, we take the population as $n = 5000$, and set the neighbors number $k_n = 80 \log n$. The initial positions and headings of the n agents are mutually independent, with positions and headings uniformly and independently distributed in $[0, 1]^2$ and $(-\pi, \pi]$, respectively. Fig. 7 shows how the probability of consensus changes with moving speed. From this simulation, we see that if the speed is small, the system can achieve consensus with high probability, and the probability of consensus will tend to small as the speed increases.

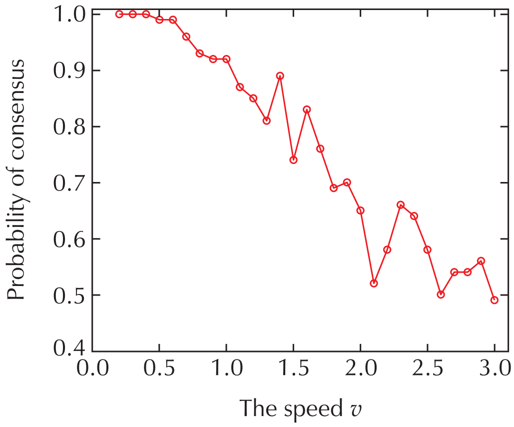

Fig. 7 Simulation example.

6 Conclusions

Most of the existing literature on flocks is concerned with interaction rules that are based on geometric distance in nature. In this paper, we have investigated a rather different class of flocks with k-nearest-neighbor rule. Such a topological distance-based interaction rule has been validated by biologists and verified to be robust with respect to disturbances.By overcoming the mathematical difficulties concerning with connectivity of the underlying nonlinear flocking dynamical systems, and with non-symmetry of the underlying dynamical topology resulted from the used directed k-nearest-neighbor rule, we are able to establish that the minimum number of neighbors k_n needed for consensus is of the order $O(\log n)$ for large population with size n. It goes without saying that this nice result may have meaningful implications in many related fields including biology, communication and social networks. For further investigation, it is desirable to shrink the “gap” between the two constants in our main theorem, and to extend the results to more complicated anisotropic updating rules.

References

[1] C. W. Reynolds. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, 1987, 21(4): 25 – 34.

[2] I. D. Couzin, J. Krause, R. James, et al. Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 2002, 218(1): 1 – 11.

[3] R. Olfati-Saber. Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on Automatic Control*, 2006, 51(3): 401 – 420.

[4] H. G. Tanner, A. Jadbabaie, G. J. Pappas. Flocking in fixed and switching networks. *IEEE Transactions on Automatic Control*, 2007, 52(5): 863 – 868.

[5] C. Viragh, G. Vasarhelyi, N. Tarcai, et al. Flocking algorithm for autonomous flying robots. *Bioinspiration and Biomimetics*, 2014, 9(2): DOI 10.1088/1748-3182/9/2/025012.

[6] T. Vicsek, A. Czirók, E. Ben-Jacob, et al. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 1995, 75(6): 1226 – 1229.

[7] P. Szabó, M. Nagy, T. Vicsek. Transitions in a self-propelled-particles model with coupling of accelerations. *Physical Review E*, 2009, 79(2): DOI 10.1103/PhysRevE.79.021908.

[8] I. D. Couzin, J. Krause, N. R. Franks, et al. Effective leadership and decision-making in animal groups on the move. *Nature*, 2005, 433(7025): 513 – 516.

[9] A. Jadbabaie, J. Lin, A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 2003, 48(6): 988 – 1001.

[10] L. Moreau. Stability of multi-agent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 2005, 50(2): 169 – 182.

[11] W. Ren, R. W. Beard, E. M. Atkins. A survey of consensus problems in multi-agent coordination. *Proceedings of the American Control Conference*, Portland: IEEE, 2005: 1859 – 1864.

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

[13] J. A. Carrillo, M. Fornasier, J. Rosado, et al. Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM Journal on Mathematical Analysis*, 2010, 42(1): 218 – 236.

[14] B. Chazelle. The geometry of flocking. *Proceedings Of the 26th Annual Symposium on Computational Geometry*, New York: ACM, 2010: 19 – 28.

[15] A. Okubo. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. *Advances in Biophysics*, 1986, 22: 1 – 94.

[16] D. Grünbaum, A. Okubo. Modeling social animal aggregations. *Frontiers in Mathematical Biology*. Berlin: Springer, 1994: 296 – 325.

[17] A. Mogilner, L. Edelstein-Keshet, L. Bent, et al. Mutual interactions, potentials, and individual distance in a social aggregation. *Journal of Mathematical Biology*, 2003, 47(4): 353 – 389.

[18] Z. Liu, J. Han, X. M. Hu. The proportion of leaders needed for the expected consensus. *Automatica*, 2009, 47(12): 2697 – 2703.

[19] G. Tang, L. Guo. Convergence of a class of multi-agent systems in probabilistic framework. *Journal of Systems Science and Complexity*, 2007, 20(2): 173 – 197.

[20] Z. Liu, L. Guo. Synchronization of multi-agent systems without connectivity assumptions. *Automatica*, 2009, 45(12): 2744 – 2753.

[21] G. Chen, Z. Liu, L. Guo. The smallest possible interaction radius for synchronization of self-propelled paricles. *SIAM Review*, 2014, 56(3): 499 – 521.

[22] C. Chen, G. Chen, L. Guo. Synchronization of mobile autonomous agents with M-nearest-neighbor rule. *Journal of Systems Science and Complexity*, 2015, 28(1): 1 – 15.

[23] C. Chen, G. Chen, L. Guo. The number of neighbors needed for consensus of flocks. *Proceedings of the 32nd Chinese Control Conference*, Xi'an: IEEE, 2013: 7060 – 7065.

[24] L. Wang, X. Wang, X. Hu. Synchronization of multi-agent systems with topological interaction. *Proceedings of the 18th IFAC World Congress*, Milano, Italy, 2011.

[25] M. Ballerini, N. Calbibbo, R. Candeleir, et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences*, 2008, 105(4): 1232 – 1237.

[26] W. Bialek, A. Cavagna, I. Giardina, et al. Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Sciences*, 2012, 109(13): 4786 – 4791.

[27] J. Emmerton, J. D. Delius. Beyond sensation: Visual cognition in pigeons. *Vision, Brain, and Behavior in Birds*. Cambridge: MIT Press, 1993: 377 – 390.

[28] R. W. Tegeder, J. Krause. Density dependence and numerosity in fright stimulated aggregation behaviour of shoaling fish. *Philosophical Transactions of the Royal Society of London*, 1995, 350(1334): 381 – 390.

[29] G. F. Young, L. Scardovi, A. Cavagna, et al. Starling flock networks manage uncertainty in consensus at low cost. *PLOS Computational Biology*, 2013, 9(1): DOI 10.1371/journal.pcbi.1002894.

[30] F. Xue, P. R. Kumar. The number of neighbors needed for connectivity of wireless networks. *Wireless Networks*, 2004, 10(2): 169 – 181.

[31] P. Balister, B. Bollobás, A. Sarkar, et al. Connectivity of random *k*-nearest-neighbour graphs. *Advances in Applied Probability*, 2005, 37(1): 1 – 24.

[32] M. Penrose. *Random Geometric Graphs*. Oxford: Oxford University Press, 2003.

[33] P. Gupta, P. R. Kumar. Critical power for asymptotic connectivity in wireless networks. *Stochastic Analysis, Control, Optimization and Applications*. Boston: Birkhäuser, 1999: 547¨C- 566.

Appendix

Proof of Lemma 5 In the following proofs, we will get rid of the subscripts from all the variables v^i_{n,K^*}, S^i_{n,K^*}, L_{n,K^*} for the sake of convenience.

For arbitrary two virtual vertices v^j and v^k, we can construct an undirected path with length $2(L-1)$ joining v^j and v^k. Without loss of generality, assume that v^j lies on the left of v^k as illustrated in Fig. a1. We begin from v^j and select virtual edges on the straight line from left to right until we arrive at right above or below v^k, then we select virtual edges on the straight line from top to bottom or bottom to top. By such method, the length of the virtual path from v^j to v^k is not larger than $2(L-1)$. If the length is strictly smaller than $2(L-1)$, then we add a number of loops (v^k, v^k) to lengthen it. It worth mentioning that the loops play an important role in the proof to be seen later. Now we prove the number of such undirected paths is not smaller than $C^L_{2(L-1)}$ in two situations:

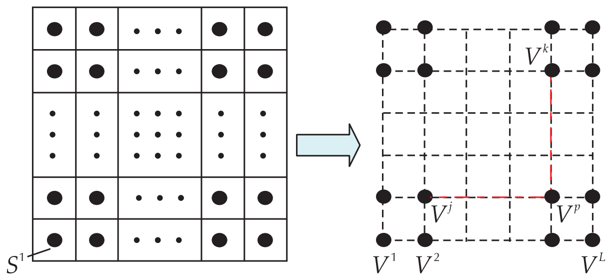

Fig. a1 The construction of virtual vertexes and edges.

i) v^j and v^k lie in the opposite corners of $[0,1]^2$ respectively, for example, v^j lies in the bottom-left square and v^k lies in the top-right square. We use a walk sequence from v^j to v^k to represent a path. In order to arrive at v^k exactly at the $2(L-1)$-th walk, each walk has to be either from left to right or from bottom to top, denoted as "$\rightarrow$" and "$\uparrow$", respectively, and a path is determined only by the order of $\rightarrow$ and $\uparrow$ (For example, "$\rightarrow\rightarrow\cdots\uparrow\uparrow$" and "$\rightarrow\uparrow\cdots\rightarrow\uparrow$" represent different paths). For the walk sequence in demand, the number of $\rightarrow$ should be $(L-1)$, and so is the number of $\uparrow$. As a result, we can choose $(L-1)$ walks as $\rightarrow$ among the total walks and the others as $\uparrow$, with the combinatorial number $C^{L-1}_{2(L-1)}$.

ii) v^j and v^k do not lie in the opposite corners of $[0,1]^2$. Assume that their coordinates are $(j_1a-\frac{1}{2}a, j_2a-\frac{1}{2}a)$, $(k_1a-\frac{1}{2}a, k_2a-\frac{1}{2}a)$ respectively, where a is the side length of the square S^i. Then from the construction of virtual vertices, we can deduce that $|k_1-j_1|, |k_2-j_2|$ are both integers and $\min\{|k_1-j_1|, |k_2-j_2|\} < L-1$. In such case, in order to arrive at v^k at exactly the $2(L-1)$-th walk, each walk may have more choices to move, not only to right and top but also to left, bottom and itself, denoted as "$\leftarrow$", "$\downarrow$" and "$\circlearrowleft$" respectively. For convenience, we denot $\rightarrow, \leftarrow$ as $\leftrightarrow$, and $\uparrow, \downarrow$ as $\updownarrow$. Now assume that $L-1$ is even without loss of generality:

. If j_1-k_1 and j_2-k_2 are both even, then let the walk sequence from v^j to v^k contain exactly $(L-1)$ $\leftrightarrow$ and $(L-1)$ $\updownarrow$, therefore the combinatorial number is $C^{L_n-1}_{2(L_n-1)}$. Now the problem is converted into a new one that whether v^j can arrive at v^k through exactly $L-1$ $\leftrightarrow$ and $L-1$ $\updownarrow$. Since $L-1$ is also even, such a walk sequence can be constructed easily: it first takes k_1-j_1 $\rightarrow$ from v^j to v^p, and then $L-1-(k_1-j_1)$ walks from v^p to v^p with $\rightarrow$ and $\leftarrow$ alternating, next takes k_2-j_2 $\uparrow$ from v^p to v^k, and finally $L-1-(k_2-j_2)$ walks from v^k to v^k with $\uparrow$

and $\downarrow$ alternating, See Fig. a1.

. If $j_1 - k_1$ is even and $j_2 - k_2$ is odd, let the walk sequence contain exactly $(L-1) \leftrightarrow$ and $(L-1) \updownarrow$ and $\circlearrowleft$, then the combinatorial number is $C^{L-1}_{2(L-1)}$ as expected. Such path can be constructed similarly to the design above: it takes $L-1 \leftrightarrow$ from v^j to v^p just the same as above, and then take $L-2 \updownarrow$ from v^p to v^k because $L-1$ is odd, finally we can add a $\circlearrowleft$ from v^k to v^k.

. If $j_1 - k_1$ is odd while $j_2 - k_2$ is even, then the combinatorial number of the walk sequences is $C^{L-1}_{2(L-1)}$ using the same analysis as above.

. If $j_1 - k_1$ and $j_2 - k_2$ are both odd, let the walk sequence contain exactly $L \leftrightarrow$ and $(L-2) \updownarrow$, then the combinatorial number is $C^{L}_{2(L-1)}$. $\square$

Proof of Lemma 10 For any $x \in A$ with distance r from D's center, $d(x, A_1) = r - r_0$. Now we claim that $\mathrm{dia}(x, \mathcal{A}_{D_x\cap A}) < r - r_0$. Pick arbitrary $A_i \subset \mathcal{A}_{D_x\cap A}$, if $A_i \subset D_x$, then from the construction of D_x, we have $\mathrm{dia}(x, A_i) \leqslant r - (1+\varepsilon)r_0 < r - r_0$ immediately; If $A_i \not\subset D_x$, then a portion of A_i is outside D_x, since the diameter of $A_i (3 \leqslant i \leqslant N)$ is at most εr_0, then $\mathrm{dia}(x, A_i) < r - (1+\varepsilon)r_0 + \varepsilon r_0 < r - r_0$. Hence, the points in $\mathcal{A}_{D_x\cap A}$ are closer to x than the points in A_1. Since $|\mathcal{A}_{D_x\cap A} \cap \mathcal{X}_n(0)| \geqslant k_n$ by (iii), then the neighbors of x all lie in $\mathcal{A}_{D_x\cap A} \subset A$. Notice that $|A_1 \cap \mathcal{X}_n(0)| \geqslant k_n$, then for any point in A_1, its neighbors all lie in A_1 itself, which makes no edge between A_1 and A, therefore $\mathcal{G}(0)$ is disconnected. $\square$

Proof of Lemma 11 By the definition of $\tilde{E}^i_n$, it is obvious that the conditions i) and ii) of Definition 6 are satisfied, and it remains to check condition iii) of Definition 6.

Pick any x with $r = 3r_0$ and ε small enough, then $\mathcal{L}(D_x \cap A) \geqslant (1/2+\delta)\mathcal{L}(D_x)$ for some $\delta > 0$, independent of ε. Hence for sufficiently small ε,

$$\mathcal{L}(D_x \cap A) \geqslant \Big(\frac{1}{2}+\delta\Big)\cdot\frac{1}{2}\pi[(2-\varepsilon)r_0]^2 > \frac{1}{2}\pi r_0^2 \geqslant 2\mathcal{L}(A_1). \quad \text{(a1)}$$

If we move the point x radially outwards from the center of D, the regions D_x form a nested family. Thus $\mathcal{L}(D_x \cap A) \geqslant 2\mathcal{L}(A_1)$ for all x.

From $\mathcal{A}_{D_x\cap A} \supset D_x \cap A$ and $|A_i \cap \mathcal{X}(0)| \geqslant \rho\mathcal{L}(A_i), 3 \leqslant i \leqslant N$, we have

$$\begin{aligned}|\mathcal{A}_{D_x\cap A} \cap \mathcal{X}(0)| &= \sum_{A_i \subset \mathcal{A}_{D_x\cap A}} |A_i \cap \mathcal{X}(0)| \geqslant \rho\mathcal{L}(D_x \cap A)\\ &\geqslant \rho\cdot 2\mathcal{L}(A_1) \geqslant k_n,\end{aligned}$$

which satisfies condition iii) of Definition 6. $\square$

Proof of Lemma 12 Using Lemma 1.4 in [32], for large n we can get

$$\Pr(\mathcal{X}_n(0) \subseteq \mathcal{P}_{\lambda^n_2}) \geqslant \Pr(\mathrm{Po}(\lambda^n_2) \geqslant n) = 1 - \mathrm{e}^{-n^{1/4}},$$
$$\Pr(\mathcal{P}_{\lambda^n_1} \subseteq \mathcal{X}_n(0)) \geqslant \Pr(\mathrm{Po}(\lambda^n_1) \leqslant n) = 1 - \mathrm{e}^{-n^{1/4}}.$$

Then by these two inequalities,

$$\begin{aligned}&\Pr\Big(\bigcup_{i=1}^{L_{n,K}} \tilde{E}^i_n\Big)\\ &\geqslant \Pr\Big(\bigcup_{i=1}^{L_{n,K}} F^i_n, \mathcal{P}_{\lambda^n_1} \subset \mathcal{X}_n(0), \mathcal{X}_n(0) \subset \mathcal{P}_{\lambda^n_2}\Big)\\ &\geqslant \Pr\Big(\bigcup_{i=1}^{L_{n,K}} F^i_n\Big) + \Pr(\mathcal{P}_{\lambda^n_1} \subset \mathcal{X}_n(0))\\ &\quad + \Pr(\mathcal{X}_n(0) \subset \mathcal{P}_{\lambda^n_2}) - 2\\ &> \Pr\Big(\bigcup_{i=1}^{L_{n,K}} F^i_n\Big) - 2\mathrm{e}^{-n^{1/4}}.\end{aligned}$$

$\square$

Proof of Lemma 13 Now we fix the value of ρ as $\rho_n\lambda^n_1$ with $\rho_n = \frac{25}{18} + \frac{8}{9}\frac{n^{3/4}}{n - n^{3/4}}$, then the number of points in D of a $\mathrm{Trap}^{\varepsilon}_{r_0}$ is as expected, i.e.,

$$\begin{aligned}&2\rho_n\lambda^n_1\mathcal{L}(A_1) + 0\cdot\mathcal{L}(A_2) + \sum_{3\leqslant i\leqslant N}\rho_n\lambda^n_1\mathcal{L}(A_i)\\ &= \lambda^n_1\sum_{i\neq 2}\mathcal{L}(A_i) + \lambda^n_2\mathcal{L}(A_2).\end{aligned}$$

Suppose that for $3 \leqslant i \leqslant N$, $\rho_n\lambda^n_1\mathcal{L}(A_i) \in \mathbb{Z}$ (for large enough n and suitable ε, r_0, this can be realized) and exactly $\rho_n\lambda^n_1\mathcal{L}(A_i)$ points lie in each A_i for $i \neq 2$, then from Lemma 9 and the spatial independence of the Poisson point process,

$$\begin{aligned}&\Pr(F^i_n)\\ &\geqslant Pr\{|\mathcal{P}_{\lambda^n_1} \cap A_1| = 2\rho_n\lambda^n_1\mathcal{L}(A_1) = k_n, |\mathcal{P}_{\lambda^n_2} \cap A_2| = 0,\\ &\quad |\mathcal{P}_{\lambda^n_1} \cap A_i| = \rho_n\lambda^n_1\mathcal{L}(A_i), 3 \leqslant i \leqslant N\}\\ &= \exp\{(-\log 2\rho_n - 8\log\rho_n)2\rho_n\lambda^n_1\mathcal{L}(A_1)\\ &\quad + \mathrm{O}(N\log(2\rho_n\lambda^n_1\mathcal{L}(A_1)))\}\\ &= \exp\{\{-\log\Big(\frac{50}{18} + \frac{16}{9}\frac{n^{3/4}}{n-n^{3/4}}\Big)\\ &\quad -8\log\Big(\frac{25}{18} + \frac{8}{9}\frac{n^{3/4}}{n-n^{3/4}}\Big)\}k_n\\ &\quad +7\mathrm{O}(N\log(2\rho_n\lambda^n_1\mathcal{L}(A_1)))\}\\ &:= \exp\{-c_nk_n + \mathrm{O}(N\log k_n)\},\end{aligned} \quad \text{(a2)}$$

where c_n monotonously decreases and satisfies $\lim_{n\to\infty} c_n := c = \log\frac{25}{9} + 8\log\frac{25}{18}$.

Note that under the condition $k_n < \frac{1}{2c}\log n = C_2\log n$, there always exists some $M > 0$ such that for $n > M$, $k_n < \frac{1}{2c_M}\log n$ and $c_n < c_M$. Again using the spatial independence of the Poisson point process, we have for $n > M$,

$$\begin{aligned}&\Pr\{(\bigcup_{1\leqslant i\leqslant L_{n,K}} F^i_n)^c\}\\ &= \Pr(\bigcap_{1\leqslant i\leqslant L_{n,K}} (F^i_n)^c) = \prod_{1\leqslant i\leqslant L_{n,K}} [1 - \Pr(F^i_n)]\\ &= \exp(\sum_{1\leqslant i\leqslant L_{n,K}} \log[1 - \Pr(F^i_n)])\\ &\leqslant \exp(-\sum_{1\leqslant i\leqslant L_{n,K}} \Pr(F^i_n))\end{aligned}$$

$$\leqslant \exp(-\sqrt{\frac{n}{K\log n}}\exp\{-\frac{c_n}{2c_M}\log n\})$$
$$\leqslant \exp(-\frac{n^{\frac{1}{2}-\frac{c_n}{2c_M}}}{\sqrt{K\log n}}),$$

where we have used the inequality $\log(1-x) \leqslant -x$ for $x \in (0,1)$ and (a2). Combing this with Lemmas 11 and 12,

$$\Pr\{(\bigcup_{1\leqslant i\leqslant L_{n,K}} E_n^i)^{\mathrm{c}}\} \leqslant \Pr\{(\bigcup_{1\leqslant i\leqslant L_{n,K}} \tilde{E}_n^i)^{\mathrm{c}}\}$$
$$\leqslant \Pr\{(\bigcup_{1\leqslant i\leqslant L_{n,K}} F_n^i)^{\mathrm{c}}\} + 2\exp(-n^{1/4})$$
$$\leqslant \exp(-\frac{n^{\frac{1}{2}-\frac{c_n}{2c_M}}}{\sqrt{K\log n}}) + 2\exp(-n^{1/4}),$$

which is followed by

$$\sum_{n=M}^{\infty} \Pr\{(\bigcup_{1\leqslant i\leqslant L_{n,K}} E_n^i)^{\mathrm{c}}\} < \infty. \qquad \text{(a3)}$$

From Borel-Cantelli theorem, (a3) means that

$$\Pr\{(\bigcup_{1\leqslant i\leqslant L_{n,K}} E_n^i)^{\mathrm{c}} \text{ happens infinitely often}\} = 0,$$

which completes the proof of Lemma 13.

Chen CHEN was born in Shannxi, China, in 1987. She received the B.Sc. degree in Mathematics from Beihang University in 2009, and the Ph.D. degree in Control Theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, in 2014. She is currently a researcher at Huawei Technologies Co. Ltd. Her research interests include complex systems, distributed filters, reinforcement learning and deep learning. E-mail: chenchen9@huawei.com.

Ge CHEN received the B.Sc. degree in Mathematics from the University of Science and Technology of China in 2004, and the Ph.D. degree in Mathematics from the University of Chinese Academy of Sciences, China, in 2009. He jointed the National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences in 2011, and is currently an Associate Professor. His current research interest is the collective behavior of multi-agent systems. Dr. Chen received the First Prize of the Application Award from the Operations Research Society of China (2010). One of his papers was selected as a SIGEST paper by the SIAM Review (2014). He was also a finalist for the OR in Development prize from the International Federation of Operations Research Societies (2011), and for the best theoretical paper award at the 10th World Congress on Intelligent Control and Automation (2012). E-mail: chenge@amss.ac.cn.

Lei GUO was born in China in 1961. He received the B.Sc. degree in Mathematics from Shandong University in 1982, and the Ph.D. degree in Control Theory from the Chinese Academy of Sciences (CAS) in 1987. He is currently a professor at the Academy of Mathematics and Systems Science, CAS, and Director of National Center for Mathematics and Interdisciplinary Sciences (NCMIS), CAS. He is a Fellow of both IEEE and IFAC, Member of CAS, Foreign Member of the Royal Swedish Academy of Engineering Sciences. His current research interests include feedback capability, nonlinear and adaptive control, multi-agent systems, distributed adaptive filtering, and game-based control systems, among others. E-mail: lguo@amss.ac.cn.