Jrl Syst Sci & Complexity (2007) 20: 173–197

CONVERGENCE OF A CLASS OF MULTI-AGENT SYSTEMS IN PROBABILISTIC FRAMEWORK*

Gongguo TANG · Lei GUO

Received: 12 March 2007

Abstract Multi-agent systems arise from diverse fields in natural and artificial systems, and a basic problem is to understand how locally interacting agents lead to collective behaviors (e.g., synchronization) of the overall system. In this paper, we will consider a basic class of multi-agent systems that are described by a simplification of the well-known Vicsek model. This model looks simple, but the rigorous theoretical analysis is quite complicated, because there are strong nonlinear interactions among the agents in the model. In fact, most of the existing results on synchronization need to impose a certain connectivity condition on the global behaviors of the agents' trajectories (or on the closed-loop dynamic neighborhood graphs), which are quite hard to verify in general. In this paper, by introducing a probabilistic framework to this problem, we will provide a complete and rigorous proof for the fact that the overall multi-agent system will synchronize with large probability as long as the number of agents is large enough. The proof is based on a detailed analysis of both the dynamical properties of the nonlinear system evolution and the asymptotic properties of the spectrum of random geometric graphs.

Key words Connectivity, large deviation, local interaction rules, multi-agent systems, random geometric graph, spectral graph theory, synchronization, Vicsek model.

1 Introduction

In recent years, the collective behaviors of multi-agent systems have drawn much attention from researchers [1–11]. The most salient characteristic of these systems is that the interactions among agents are based on local rules, that is, each agent interacts with those agents neighboring to it in some sense. Amazingly, without central control and global information exchange, the system as a whole can spontaneously generate various kinds of "macro" behaviors, such as synchronization, whirlpool, etc., merely based on local interactions.

Typical examples of multi-agent systems include animal aggregations such as flocks, schools and herds. Biologists have given detailed descriptions and discussions on the mechanisms of flying, swimming, and migrating in these aggregations. In many cases, agents have the tendency to move as other agents do in their neighborhood. Inspired more or less by this, Vicsek et al. proposed a model to simulate and explain the clustering, transportation and phase transition in nonequilibrium systems[5,6]. The model consists of finite agents (particles, animals, robots,

Gongguo TANG · Lei GUO
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, *China.*
Email: Lguo@amss.ac.cn.
*The research is supported by National Natural Science Foundation of China under the Grants No. 60221301 and No. 60334040.

etc.) on the plane, each of which moves with the same constant speed. At each time step, a given agent assumes the average direction of agents' motions in its neighborhood of radius r. Through simulation, Vicsek et al. explained the kinetic phase transition exhibited in the model by the spontaneous symmetry breaking of the rotational symmetry. The Vicsek model can also be viewed as a special case of the well-known Boid model introduced by Reynolds in 1987[7], where the purpose was to simulate the behaviors in flocks of flying birds and schools of fishes. Agents in the Boid model obey three rules in their movement: Collision Avoidance, Velocity Matching and Flock Centering. All of these three rules are local ones, which means that each agent adjusts its behavior based on the behaviors of the agents in its neighborhood.

Inspired by both the nature phenomena and the computer simulations, scientists have kept trying to give rigorous theoretical foundations and explanations. One of the most notable attempts is made recently by Jadbabaie et al.[8]. In this paper, the motion direction iteration rule in the Vicsek model is linearilized, and it was shown that the motion directions of all agents will converge to a common one provided that the closed-loop neighborhood graphs of the system are jointly connected with sufficient frequency. It was later found that some related results have been given in an earlier paper[12] in a somewhat different context. These results provide preliminary theoretical explanations of the phenomena observed in simulations. However, from a rigorous theoretical perspective, the existing results are far from complete. The main reason is that all the conditions in the existing theoretical analysis are imposed on the "closed-loop" graphs, which are resulted from the iteration of the system dynamics, and should be determined by both the initial states and model parameters. These results do not give any clue to how the neighboring graphs evolve, and how to verify the connectivity conditions. It is worth mentioning that if the local rules are modified to be weighted but global ones along the way, for example, suggested in [11], a complete theoretical analysis can be given with the convergence conditions imposed on system initial states and model parameters only[11]. The first complete result which guarantees the synchronization of the Vicsek model by imposing conditions only on the system initial states and model parameters seems to have been given in [13], but these conditions are still not satisfactory in the sense that they may not be valid for large population. Nevertheless, the results in [8,9,11,13] all suggest that the connectivity of the closed-loop graphs resulted from the system iteration is crucial to synchronization.

In this paper, we will take a somewhat different perspective to introduce a probabilistic framework for investigating the convergence of a class of multi-agent systems described by the linearized Vicsek model. We will first give a detailed analysis of both the dynamical properties of the nonlinear system evolution and the asymptotic properties of the spectrum of random geometric graphs, and then demonstrate that the linearized Vicsek model will synchronize with large probability for any given interaction radius r and motion speed v, whenever the population size is large enough.

The paper is organized as follows: In the next section, we will describe the multi-agent systems by a simplified Vicsek model and will present the main result (Theorem 1); The analysis of the system dynamics and the estimation of the characteristics of random geometric graphs will be given in Sections 3 and 4 respectively, and the proof of the main theorem will be given in Section 5; Section 6 will give some concluding remarks. The paper also contains two appendices giving the basic concepts and results in graph theory that are used in the paper.

2 The Main Results

We first introduce the model to be studied in the paper, with the related basic concepts in Graph Theory given in Appendix A.

The original Vicsek Model[5] consists of n agents on the plane, which are labelled by $1, 2, \cdots, n$. At any time t, each agent moves with a constant absolute velocity v, and assumes the average direction of motion of agents in its neighborhood with radius r at time $t-1$. Thus at time t, the neighborhood set of any agent k is defined as

$$N_k(t) = \{j : \|x_j(t) - x_k(t)\| < r\}, \tag{1}$$

where $x_k(t) \in \mathbb{R}^2$ denote the location of the kth agent at time t, and r is the interaction radius. Obviously, if we denote $x(t) = (x_1(t), x_2(t), \cdots, x_n(t))^\tau$, then the graph induced by the neighborhood relationship is a geometric graph $G(x(t), r)$, which can be abbreviated as $G(t)$. Quantities related to $G(t)$ may also be functions of time t.

As far as the original Vicsek Model is concerned, the moving direction of the kth agent is represented by its angle $\theta_k \in (-\pi, \pi]$ with the moving direction iteration rule of any agent k $(1 \le k \le n)$ given by

$$\theta_k(t) = \arctan \frac{\sum\limits_{j \in N_k(t-1)} \sin\theta_j(t-1)}{\sum\limits_{j \in N_k(t-1)} \cos\theta_j(t-1)}. \tag{2}$$

Suppose the moving speed of each agent is denoted by v,then the position iteration rule of any agent k $(1 \le k \le n)$ is

$$x_k(t) = x_k(t-1) + v\tilde{s}(\theta_k(t)), \tag{3}$$

where $\tilde{s}(\theta) \triangleq (\cos\theta, \sin\theta)^\tau$.

Vicsek et al.[5] attribute the kinetic phase transition exhibited in the model by the spontaneous symmetry breaking of the rotational symmetry. However, the intrinsic nonlinearity in the moving direction iteration rule makes the theoretical analysis quite complicated. The paper [8] proposed the following linearized Vicsek model through linearizing the equation (2) $(1 \le k \le n)$,

$$\theta_k(t) = \frac{1}{n_k(t-1)} \sum_{j \in N_k(t-1)} \theta_j(t-1). \tag{4}$$

Although obtained for the sake of mathematical analysis, the linearized Vicsek model has its own interest because the moving direction iteration rule (4) can be viewed as the solution to the following optimization problem

$$\min_{\theta} \sum_{j \in N_k(t-1)} \big(\theta - \theta_j(t-1)\big)^2. \tag{5}$$

Now, let $\theta(t) = (\theta_1(t), \theta_2(t), \cdots, \theta_n(t))^\tau$, $s(\theta(t)) = (\tilde{s}(\theta_1(t)), \tilde{s}(\theta_2(t)), \cdots, \tilde{s}(\theta_n(t)))^\tau$, then the iteration rules (3) and (4) of the linearized Vicsek model can be rewritten as

$$\begin{cases} \theta(t) = P(t-1)\theta(t-1), \\ x(t) = x(t-1) + vs(\theta(t)), \end{cases} \tag{6}$$

where $P(t-1)$ is the average matrix of the graph $G(t-1)$.

It is easy to see that for any fixed model parameters v and r, the graph sequence $\{G(t), t \ge 0\}$ is totally determined by the initial states $\theta(0)$ and $x(0)$. Our main question is: Under what conditions the angles $\{\theta_k(t), 1 \le k \le n\}$ will converge to a common one $\bar{\theta}$, i.e., $\theta(t) \to \bar{\theta}1_n$ $(t \to \infty)$ with $1_n = (1, 1, \cdots, 1)^\tau$. When this happens, we call that system (6) converges, or synchronizes. However, the "entanglement" between the moving direction iteration and position iteration makes the analysis of this nonlinear system quite difficult. Recently, Jadbabaie[8] investigated the first equation of (6) only by viewing it as a switched system to explore what conditions on the graph sequence $\{G(t), t \geq 0\}$ will guarantee $\{\theta_k(t), 1 \leq k \leq n\}$ converge to a common value. It was shown that if the closed-loop graphs resulted from the iteration of system (6) satisfy some connectivity property, then the system will converge.

In this paper, we will consider the above model in the following probabilistic framework: Let $(\Omega, \mathcal{F}, \mathcal{P})$ be the underlying probability space, and assume that in the system (6), the initial positions $\{x_j(0),\ 1 \leq j \leq n\}$ are i.i.d. random vectors uniformly distributed on the unit square $\mathcal{S}$, that the initial angles $\{\theta_j(0),\ 1 \leq j \leq n\}$ are i.i.d. random variables uniformly distributed on the interval $(-\pi, \pi]$, and that the initial positions and initial angels are mutually independent. Under these hypotheses, the initial graph $G(x(0), r)$ is a random geometric graph. The main result of this paper is as follows:

Theorem 1 *Consider the above probabilistic framework for the multi-agent system described by* (6)*. Then for any given speed $v > 0$, any radius $r > 0$ and all large population size n, the system will synchronize on a set with probability not less than $1 - O(\frac{1}{n^{g_n}})$, where $g_n = \frac{n}{\log^6 n}$.*

The main task of the following three sections is to provide a complete proof of this theorem.

3 Analysis of System Dynamics

In order to prove the main result of this paper, the analysis of the dynamics of system (6) is necessary. In this section, two lemmas concerning the estimation of convergence rate will be provided, one applicable to the case with constant closed-loop graphs, and the other holding true when the closed-loop graphs undergo small changes.

Denote

$$\delta(\theta) = \max_{1 \leqslant j \leqslant n} \theta_j - \min_{1 \leqslant j \leqslant n} \theta_j, \tag{7}$$

then we have $\delta(\theta) \leqslant \sqrt{2}\|\theta\|$, and by (4), $\delta(\theta(t))$ is nonincreasing with t.

Lemma 1 *For any $\theta_0 \in \mathbb{R}^n$ and for any undirected graph G, let P and T be its average matrix and degree matrix respectively, and let $\{\phi_0, \phi_1, \cdots, \phi_{n-1}\}$ be a system of orthogonal basis in R^n composed of the unit eigenvectors of the normalized Laplacian of G with $\phi_0 = \frac{1}{\sqrt{\mathrm{Vol}(G)}} T^{\frac{1}{2}} 1_n$. Moreover, let $T^{\frac{1}{2}}\theta_0$ be expanded as $\sum_{j=0}^{n-1} a_j \phi_j$, then*

$$\left\| P^t \theta_0 - \frac{a_0}{\sqrt{\mathrm{Vol}(G)}} 1_n \right\| \leq \kappa \bar{\lambda}^t \|\theta_0\|, \qquad \forall\, t \geq 0, \tag{8}$$

where $\bar{\lambda}$ is the spectral gap of G, and κ denotes the square root of the ratio of the maximum degree to the minimum degree of G, i.e., $\sqrt{\frac{d_{\max}}{d_{\min}}}$.

Proof First we note that by the definition of the Laplacian matrix L, it is known that $L1_n = 0$, and so

$$\mathcal{L}\phi_0 = T^{-\frac{1}{2}} L T^{-\frac{1}{2}} \left(\frac{1}{\sqrt{\mathrm{Vol}(G)}} T^{\frac{1}{2}} 1_n \right) = 0,$$

i.e., ϕ_0 is indeed the unit eigenvector of the normalized Laplacian matrix $\mathcal{L}$ corresponding to the eigenvalue $\lambda_0 = 0$. By $T^{\frac{1}{2}}\theta_0 = \sum_{j=0}^{n-1} a_j \phi_j$ it follows that

$$\theta_0 = \sum_{j=0}^{n-1} a_j T^{-\frac{1}{2}} \phi_j = \sum_{j=1}^{n-1} a_j T^{-\frac{1}{2}} \phi_j + \frac{a_0}{\sqrt{\mathrm{Vol}(G)}} 1_n,$$

and hence

$$
\begin{aligned}
P^t\theta_0 &= T^{-\frac{1}{2}}(I-\mathcal{L})^t T^{\frac{1}{2}}\theta_0 = T^{-\frac{1}{2}}(I-\mathcal{L})^t\left(\sum_{j=0}^{n-1} a_j\phi_j\right)\\
&= \sum_{j=1}^{n-1}(1-\lambda_j)^t a_j T^{-\frac{1}{2}}\phi_j + \frac{a_0}{\sqrt{\mathrm{Vol}(G)}}1_n.
\end{aligned}
$$

From this and the fact that $\{\phi_0, \phi_1, \cdots, \phi_{n-1}\}$ is a unit orthogonal basis, we have

$$
\begin{aligned}
&\left\|P^t\theta_0 - \frac{a_0}{\sqrt{\mathrm{Vol}(G)}}1_n\right\| = \left\|\sum_{j=1}^{n-1}(1-\lambda_j)^t a_j T^{-\frac{1}{2}}\phi_j\right\|\\
\le\ & \|T^{-\frac{1}{2}}\|\left(\sum_{j=1}^{n-1}(1-\lambda_j)^{2t}a_j^2\right)^{\frac{1}{2}} \le \|T^{-\frac{1}{2}}\|\bar{\lambda}^t\left(\sum_{j=1}^{n-1}a_j^2\right)^{\frac{1}{2}}\\
\le\ & \bar{\lambda}^t\|T^{-\frac{1}{2}}\|\|T^{\frac{1}{2}}\theta_0\| \le \kappa\bar{\lambda}^t\|\theta_0\|.
\end{aligned}
$$

∎

The following lemma plays a key role in the paper whose proof is inspired by the stability analysis of time-varying linear systems[14].

Lemma 2 *Let $\{G(t), t \geqslant t_0\}$ be a sequence of time-varying undirected graphs, with the corresponding characteristic quantities $\{\mathcal{L}(t), P(t), d_{\min}(t), d_{\max}(t), \bar{\lambda}(t), t \geqslant t_0\}$(see Appendix* A)*, and let $\{\theta(t), t \geqslant t_0\}$ be recursively defined by*

$$\theta(t+1) = P(t)\theta(t).$$

If there exists an undirected graph G with the corresponding $\{\mathcal{L}, P, d_{\min}, d_{\max}, \bar{\lambda}\}$, such that $\|P(t) - P\| \le \varepsilon$, for some $\varepsilon > 0$, then

$$\delta(\theta(t)) \le \sqrt{2}\kappa\big(\bar{\lambda} + \kappa\varepsilon\big)^{t-t_0}\|\theta(t_0)\|, \quad t \ge t_0.$$

Proof Let us denote $\Delta P(t) = P(t) - P$. Then

$$
\begin{aligned}
\theta(t+1) &= P(t)\theta(t) = P\theta(t) + \Delta P(t)\theta(t)\\
&= P^{t+1-t_0}\theta(t_0) + \sum_{s=t_0}^{t} P^{t-s}\Delta P(s)\theta(s).
\end{aligned}
$$

Similar to Lemma 1, let $0 = \lambda_0 \le \lambda_1 \le \cdots \le \lambda_{n-1}$ be the eigenvalues of the normalized Laplacian $\mathcal{L}$ of G, with the corresponding unit orthogonal eigenvectors $\{\phi_j, 0 \le j \le n-1\}$. If we denote

$$T^{\frac{1}{2}}\theta(t_0) = \sum_{j=0}^{n-1} a_j\phi_j\ , \quad T^{\frac{1}{2}}\Delta P(s)\theta(s) = \sum_{j=0}^{n-1} a_j(s)\phi_j,\ s \geqslant t_0,$$

and

$$\bar{\theta}(t+1) \triangleq \frac{a_0}{\sqrt{\mathrm{Vol}(G)}}1_n + \sum_{s=t_0}^{t}\frac{a_0(s)}{\sqrt{\mathrm{Vol}(G)}}1_n,$$

we then have

$$
\begin{aligned}
&\theta(t+1) - \bar{\theta}(t+1)\\
=&\left(P^{t+1-t_0}\theta(t_0) - \frac{a_0}{\sqrt{\mathrm{Vol}(G)}}1_n\right) + \sum_{s=t_0}^{t}\left(P^{t-s}\Delta P(s)\theta(s) - \frac{a_0(s)}{\sqrt{\mathrm{Vol}(G)}}1_n\right).
\end{aligned}
$$

By Lemma 1 and the fact that $\Delta P(s)1_n = 0$, we have

$$\begin{aligned}
&\|\theta(t+1)-\bar{\theta}(t+1)\| \\
\le\ & \kappa\bar{\lambda}^{t+1-t_0}\|\theta(t_0)\| + \sum_{s=t_0}^{t}\kappa\bar{\lambda}^{t-s}\|\Delta P(s)\theta(s)\| \\
=\ & \kappa\bar{\lambda}^{t+1-t_0}\|\theta(t_0)\| + \kappa\sum_{s=t_0}^{t}\bar{\lambda}^{t-s}\big\|\Delta P(s)\big(\theta(s)-\bar{\theta}(s)\big)\big\| \\
\le\ & \kappa\bar{\lambda}^{t+1-t_0}\|\theta(t_0)\| + \kappa\varepsilon\sum_{s=t_0}^{t}\bar{\lambda}^{t-s}\big\|\theta(s)-\bar{\theta}(s)\big\|.
\end{aligned}$$

Now, denote $\xi(t) = \|\theta(t)-\bar{\theta}(t)\|$, we have

$$\xi(t+1) \le \kappa\bar{\lambda}^{t+1-t_0}\|\theta(t_0)\| + \kappa\varepsilon\sum_{s=t_0}^{t}\bar{\lambda}^{t-s}\xi(s) \triangleq z(t+1).$$

It is easy to see that

$$z(t+1) \le (\bar{\lambda}+\varepsilon\kappa)z(t), \qquad z(t_0) = \kappa\|\theta(t_0)\|,$$

so

$$\xi(t+1) \le z(t+1) \le \kappa\|\theta(t_0)\|(\bar{\lambda}+\kappa\varepsilon)^{t+1-t_0}.$$

Finally, we get the desired result

$$\delta(\theta(t)) = \delta\big(\theta(t)-\bar{\theta}(t)\big) \le \sqrt{2}\|\xi(t)\| \le \sqrt{2}\kappa\Big(\bar{\lambda}+\kappa\varepsilon\Big)^{t-t_0}\|\theta(t_0)\|. \qquad ∎$$

Lemma 3 *Let $\mathcal{L}$ be the normalized Laplacian matrix of a geometric graph $G(x,r)$and $\widehat{G}$ be another graph formed by changing the neighborhood of $G(x,r)$. If the number of points changed in the neighborhood of the k-th ($1 \le k \le n$) node satisfies $R_k \le R_{\max} < d_{\min}$, then the corresponding normalized Laplacian matrix $\widehat{\mathcal{L}}$ satisfies*

$$\|\mathcal{L}-\widehat{\mathcal{L}}\| \le 2\frac{R_{\max}}{d_{\min}}\left(1+\frac{d_{\min}(d_{\max}+R_{\max})}{(d_{\min}-R_{\max})^2}\right). \tag{9}$$

Similarly, for the average matrix P and $\widehat{P}$, we have

$$\|P-\widehat{P}\| \le \frac{R_{\max}}{d_{\min}}\left(\frac{d_{\max}+d_{\min}}{d_{\min}-R_{\max}}\right). \tag{10}$$

Proof 1) By the definition of the normalized Laplacian, we have

$$\begin{aligned}
\mathcal{L}-\widehat{\mathcal{L}} &= T^{-\frac{1}{2}}LT^{-\frac{1}{2}} - \widehat{T}^{-\frac{1}{2}}\widehat{L}\widehat{T}^{-\frac{1}{2}} \\
&= T^{-\frac{1}{2}}(L-\widehat{L})T^{-\frac{1}{2}} + (T^{-\frac{1}{2}}-\widehat{T}^{-\frac{1}{2}})\widehat{L}T^{-\frac{1}{2}} + \widehat{T}^{-\frac{1}{2}}\widehat{L}(T^{-\frac{1}{2}}-\widehat{T}^{-\frac{1}{2}}) \\
&\triangleq I + II + III.
\end{aligned}$$

We first estimate the term I. Since the diagonal elements of the matrix $L-\widehat{L}$ are bounded by $R_{\max}$, other elements belong to $\{-1,1,0\}$, and the nonzero elements of each row cannot exceeds $R_{\max}$, by the Disk Theorem[14,15] it is known that any eigenvalue of $L-\widehat{L}$ is bounded by $|\lambda(L-\widehat{L})| \le 2R_{\max}$. Furthermore, by the symmetry of $L-\widehat{L}$, we know that $\|L-\widehat{L}\| = \max\{|\lambda(L-\widehat{L})|\} \le 2R_{\max}$. Hence

$$\|I\| = \|T^{-\frac{1}{2}}(L-\widehat{L})T^{-\frac{1}{2}}\| \le \frac{2R_{\max}}{d_{\min}}.$$

Next, we estimate the second term II. First note that $\widehat{d}_j \ge d_{\min} - R_{\max}$, and so

$$\begin{aligned}\|T^{-\frac{1}{2}} - \widehat{T}^{-\frac{1}{2}}\| &= \max_j \left| \frac{1}{\sqrt{d_j}} - \frac{1}{\sqrt{\widehat{d}_j}} \right| \\ &= \max_j \frac{|d_j - \widehat{d}_j|}{\sqrt{d_j\widehat{d}_j}\left(\sqrt{d_j}+\sqrt{\widehat{d}_j}\right)} \le \frac{R_{\max}}{2\sqrt{d_{\min}}(d_{\min}-R_{\max})}.\end{aligned}$$

By using the Disk Theorem again,

$$\|\widehat{L}\| = \max\{|\lambda(\widehat{L})|\} \le 2(d_{\max}+R_{\max}).$$

Hence

$$\begin{aligned}\|II\| &\le \|T^{-\frac{1}{2}} - \widehat{T}^{-\frac{1}{2}}\|\|\widehat{L}\|\|T^{-\frac{1}{2}}\| \\ &\le \frac{R_{\max}}{2(d_{\min}-R_{\max})^{\frac{3}{2}}} \cdot 2(d_{\max}+R_{\max}) \cdot \frac{1}{\sqrt{d_{\min}}} \\ &\le \frac{R_{\max}(d_{\max}+R_{\max})}{(d_{\min}-R_{\max})^2}.\end{aligned}$$

Finally, we estimate the last term III.

$$\begin{aligned}\|III\| &\le \|\widehat{T}^{-\frac{1}{2}}\|\|\widehat{L}\|\|T^{-\frac{1}{2}} - \widehat{T}^{-\frac{1}{2}}\| \\ &\le \frac{1}{\sqrt{d_{\min}-R_{\max}}} \cdot 2(R_{\max}+d_{\max}) \cdot \frac{R_{\max}}{2(d_{\min}-R_{\max})^{\frac{3}{2}}} \\ &\le \frac{R_{\max}(d_{\max}+R_{\max})}{(d_{\min}-R_{\max})^2}.\end{aligned}$$

Therefore, combining all the above analysis, we get

$$\|\mathcal{L}-\widehat{\mathcal{L}}\| \le 2\frac{R_{\max}}{d_{\min}}\left(1+\frac{d_{\min}(d_{\max}+R_{\max})}{(d_{\min}-R_{\max})^2}\right).$$

2) To prove the second assertion, we first note that

$$\begin{aligned} & P-\widehat{P} \\ &= T^{-1}A - \widehat{T}^{-1}\widehat{A} \\ &= (T^{-1}-\widehat{T}^{-1})A + \widehat{T}^{-1}(A-\widehat{A}) \qquad (11) \\ &\triangleq I + II. \qquad (12)\end{aligned}$$

For the first term I, by the Disk Theorem[14,15] we have $\|A\| \le |\lambda_{\max}(A)| \le d_{\max}$. Furthermore,

$$\|T^{-1}-\widehat{T}^{-1}\| \le \max_j \left|\frac{1}{d_j} - \frac{1}{\widehat{d}_j}\right| = \max_j \frac{|d_j-\widehat{d}_j|}{d_j\widehat{d}_j} \le \frac{R_{\max}}{d_{\min}(d_{\min}-R_{\max})}.$$

Hence, we get

$$\|I\| \leq \|T^{-1} - \widehat{T}^{-1}\|\|A\| \leq \frac{R_{\max} d_{\max}}{d_{\min}(d_{\min} - R_{\max})}.$$

Now, we estimate the second term II. By the Disk Theorem again, we have $\|A - \widehat{A}\| \leq \lambda_{\max}(A - \widehat{A}) \leq R_{\max}$ and so

$$\|II\| \leq \|\widehat{T}^{-1}\|\|A - \widehat{A}\| \leq \frac{R_{\max}}{d_{\min} - R_{\max}}.$$

Finally, combining the above estimations we obtain

$$\|P - \widehat{P}\| \leq \frac{R_{\max}}{d_{\min}}\left(\frac{d_{\max} + d_{\min}}{d_{\min} - R_{\max}}\right). \qquad \blacksquare$$

In the following analysis, the number n which denotes the number of vertexes of a graph or the number of agents in the model is taken as a variable, and we will analyze the asymptotic properties of the Laplacian for large n .

Corollary 1 *Assume that there exist two positive constants $\alpha \in (0, 1)$ and $\beta \geq 1$ such that for large n, $R_{\max} \leq \alpha d_{\min}(1 + o(1))$ and $d_{\max} \leq \beta d_{\min}(1 + o(1))$, then*

$$\|\mathcal{L} - \widehat{\mathcal{L}}\| \leq 2\left(1 + \frac{\beta + \alpha}{(1 - \alpha)^2}\right)\frac{R_{\max}}{d_{\min}}(1 + o(1)),$$
$$\|P - \widehat{P}\| \leq \frac{1 + \beta}{1 - \alpha} \cdot \frac{R_{\max}}{d_{\min}}(1 + o(1)).$$

Combining the above lemmas we can obtain a sufficient condition guaranteeing the synchronization of the system (6). For simplicity of notations, we will omit the subscript 0 in all the variables corresponding to the initial graph $G(x(0), r)$. For any node j, we introduce the following ring,

$$\mathcal{R}_j \triangleq \{x : (1 - \eta)r \leq \|x - x_j(0)\| \leq (1 + \eta)r\}, \tag{13}$$

where $0 < \eta < 1$ is any given positive number.

Proposition 1 *For the linearized Vicsek model* (6), *if the number of agents is sufficiently large and the following three conditions are satisfied, the the systems will synchronize:*

i) *For any node j, the number of nodes within the ring $\mathcal{R}_j$ has an upper bound $R_{\max}$, which satisfies*

$$R_{\max} \leq \alpha d_{\min}(1 + o(1)), \quad d_{\max} \leq \beta d_{\min}(1 + o(1)), \tag{14}$$

where $0 < \alpha < 1$ and $\beta \geq 1$ are constants.

ii) *The spectral gap $\bar{\lambda}$ of the initial graph $G(x(0), r)$ satisfies*

$$\bar{\lambda} + \varepsilon < 1, \tag{15}$$

where $\varepsilon \triangleq 2(1 + \frac{\beta+\alpha}{(1-\alpha)^2})\sqrt{\beta}\frac{R_{\max}}{d_{\min}}$.

iii) *The speed v of each agent satisfies the following inequality*

$$\frac{v\delta(\theta(1))}{1 - (\bar{\lambda} + \varepsilon)}\left(2 + \log\frac{2\sqrt{\beta}\|\theta(1)\|}{\delta(\theta(1))}\right) \leq \eta r, \tag{16}$$

where $\delta(\cdot)$ is defined by (7).

Proof We only need to prove the following claim: at any time t, for any agent j in the graph $G(t)$, the number of neighbors of which are different from those in $G(0)$ does not exceed $R_{\max}$. Because if this is true, according to Corollary 1, we get for n large enough

$$\begin{aligned}\lambda_1\big(\mathcal{L}(t)\big) &\geq \lambda_1\big(\mathcal{L}(0)\big) - \big\|\mathcal{L}(t) - \mathcal{L}(0)\big\| \\ &\geq \lambda_1\big(\mathcal{L}(0)\big) - 2\left(1 + \frac{\beta+\alpha}{(1-\alpha)^2}\right)\frac{R_{\max}}{d_{\min}}(1+o(1)) > 0.\end{aligned}$$

Therefore graph $G(t)$ is connected, and Theorem 1 in [8] guarantees the convergence of the linearized Vicsek's Model (6).

Next we prove the above claim by induction. At $t = 0$, the claim is obviously true.

Suppose the claim is valid for $s < t$. As a result of Corollary 1, we get $\|P(s) - P(0)\| \leq \frac{\varepsilon}{\sqrt{\beta}}$, $\forall\ s < t$. Hence, by Lemma 2, when n is large enough it is true that for arbitrary $s \leq t$, $\delta(\theta(s)) \leq 2\sqrt{\beta}(\bar{\lambda}+\varepsilon)^{s-1}\|\theta(1)\|$. By this and Condition ii), we can calculate the maximal distance between any two agents in motion as follows.

First of all, for arbitrary $1 \leq j \neq k \leq n$,

$$\begin{aligned}&\|x_j(t) - x_k(t)\| \\ &\leq \|x_j(t-1) - x_k(t-1)\| + v\left|\ 2\sin\left(\frac{\theta_j(t) - \theta_k(t)}{2}\right)\right| \\ &\leq \|x_j(t-1) - x_k(t-1)\| + v\delta(\theta(t)) \\ &\leq \|x_j(0) - x_k(0)\| + v\sum_{s=1}^{t}\delta(\theta(s)). \end{aligned}\tag{17}$$

Similarly, we can get

$$\|x_j(0) - x_k(0)\| \leq \|x_j(t) - x_k(t)\| + v\sum_{s=1}^{t}\delta(\theta(s)). \tag{18}$$

Now, let us denote $s_0 = \min\{s : 2\sqrt{\beta}\ (\bar{\lambda}+\varepsilon)^{s-1}\|\theta(1)\| \leq \delta(\theta(1))\}$, then

$$s_0 = \left\lceil \frac{\log\frac{\delta(\theta(1))}{2\sqrt{\beta}\|\theta(1)\|}}{\log(\bar{\lambda}+\varepsilon)} + 1\right\rceil \leq \frac{\log\frac{\delta(\theta(1))}{2\sqrt{\beta}\|\theta(1)\|}}{\log(\bar{\lambda}+\varepsilon)} + 2.$$

Hence, we have

$$\begin{aligned}&v\sum_{s=1}^{t}\delta(\theta(s)) \\ &= v\left(\sum_{s=1}^{s_0-1}\delta(\theta(s)) + \sum_{s=s_0}^{t}\delta(\theta(s))\right) \\ &< v(s_0-1)\delta(\theta(1)) + 2v\sqrt{\beta}(\bar{\lambda}+\varepsilon)^{s_0-1}\|\theta(1)\|\sum_{s=s_0}^{t}(\bar{\lambda}+\varepsilon)^{s-s_0} \\ &\leq v\delta(\theta(1))\left(\frac{\log\frac{\delta(\theta(1))}{2\sqrt{\beta}\|\theta(1)\|}}{\log(\bar{\lambda}+\varepsilon)} + 1 + \frac{1}{1-(\bar{\lambda}+\varepsilon)}\right) \\ &\leq \frac{v\delta(\theta(1))}{1-(\bar{\lambda}+\varepsilon)}\left(2 + \log\frac{2\sqrt{\beta}\|\theta(1)\|}{\delta(\theta(1))}\right) \\ &\leq \eta r,\end{aligned}$$

where for the last but one inequality we have used the following simple facts: $\log \frac{\delta(\theta(1))}{2\sqrt{\beta}\|\theta(1)\|} < 0$ and $\log x \leq x - 1,\ \forall\ 0 < x < 1$.

According to this and the inequality (17), we conclude that if $\|x_j(0) - x_k(0)\| \leq (1-\eta)r$, then $\|x_j(t) - x_k(t)\| < r$; Otherwise if $\|x_j(0) - x_k(0)\| \geq (1+\eta)r$, then by (18), $\|x_j(t) - x_k(t)\| \geq r$. Hence at time t the variation of the neighbors for any agent j cannot exceed the number of agents in the ring $\mathcal{R}_j = \{x : (1-\eta)r \leq \|x - x_j(0)\| \leq (1+\eta)r\}$ at time 0, hence cannot exceed $R_{\max}$. This completes the induction arguments, and the proof of the proposition is complete. ▮

It is worth noting that all the conditions in the above proposition are imposed on the model parameters and the initial conditions. The task of the next section is to show how these conditions can be satisfied by analyzing random geometric graphes.

4 Estimation for the Characteristics of Random Geometric Graph

Throughout the sequel, we denote $\{a_n, g_n, n \in \mathbb{N}\}$ as positive sequences satisfying

$$\sqrt{\frac{\log n}{n}} \ll a_n \ll 1 \ll g_n \ll \frac{n{a_n}^2}{\log n},$$

where by definition $a_n \ll b_n$ means that $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = 0$ for any positive sequences $\{a_n, b_n, n \in \mathbb{N}\}$.

Let us partition the unit square $\mathcal{S}$ into $M_n = \lceil \frac{1}{a_n} \rceil^2$ equal size small squares with the length of each sides equal to $a_n(1 + o(1))$, where $\lceil x \rceil$ is the smallest integer not less than x. Furthermore, we label these small squares as $S_j,\ j = 1, 2, \cdots, M_n$, from left to right, and from top to bottom. This idea of tessellation and the following lemma is inspired by [16].

Now, we place n agents independently on $\mathcal{S}$ according to the uniform distribution with their positions denoted by $X = (X_1, X_2, \cdots, X_n)^\tau$. Denote by N_j the number of agents that fall into the small square S_j. The following lemma gives a uniform estimation for N_j.

Lemma 4

$$Pr\big\{N_j = na_n^2(1 + o(1)), 1 \leq j \leq M_n\big\} = 1 - O\Big(\frac{1}{n^{g_n}}\Big). \tag{19}$$

Proof First consider the small square S_1. Denote X_j as the indicator function of the event where the jth agent falls into S_1. Then $\{X_j, 1 \leq j \leq n\}$ are i.i.d. Bernoulli random variables with success probability $p = a_n^2(1+o(1))$ and $N_1 = \sum_{j=1}^n X_j$. According to Chernoff Bound[17], for arbitrarily given $\varepsilon \in (0, 1)$, it is true that

$$Pr\{|N_1 - np| > \varepsilon np\} \leq 2\exp\Big(-\frac{\varepsilon^2 np}{3}\Big).$$

Obviously $\{N_j, 1 \leq j \leq M_n\}$ are identically distributed (but not independent) random variables, hence

$$\begin{aligned}
&Pr\Big\{\max_{1\leq j\leq M_n} |N_j - np| \leq \varepsilon np\Big\} \\
&\geq 1 - \sum_{j=1}^{M_n} Pr\{|N_j - np| > \varepsilon np\} \\
&\geq 1 - \frac{2}{a_n^2}(1 + o(1))\exp\Big(-\frac{\varepsilon^2 np}{3}\Big).
\end{aligned}$$

Take

$$\varepsilon = \varepsilon_n = \sqrt{\frac{3(g_n \log n - \log a_n^2)}{na_n^2}} = o(1),$$

then when n is sufficiently large

$$\begin{aligned}&Pr\Big\{\max_{1\le j\le M_n}|N_j-np|\le\varepsilon_n np\Big\}\\&\ge 1-2\exp\left\{-\frac{1}{3}\left(\sqrt{\frac{3(g_n\log n-\log a_n^2)}{na_n^2}}\right)^2 na_n^2(1+o(1))-\log a_n^2\right\}(1+o(1))\\&=1-O\Big(\frac{1}{n^{g_n}}\Big).\end{aligned}$$

∎

Denote the set

$$B(a_n)=\big\{\omega\in\Omega: N_j=na_n^2(1+o(1)),1\le j\le M_n\big\}.$$

The following analysis is carried out on this set. By Lemma 4, it is easy to prove the following lemma.

Lemma 5 *For random geometric graph $G(X,r)$ in $\mathbb{R}^2$, given $\omega\in B(a_n)$, suppose that one of the following three figures intersects with the unit square $\mathcal{S}$ with an area A of the intersecting part and a length L of the arc in $\mathcal{S}$:*

i) *Rectangle $\{x=(x^1,x^2)\in\mathbb{R}^2:|x^1-x_0^1|<a,|x^2-x^0|<b\}$;*

ii) *Disk $\{x\in\mathbb{R}^2:\parallel x-X_j\|<r\}$;*

iii) *Ring $\{x\in\mathbb{R}^2:(1-\eta)r\le\parallel x-X_j\|\le(1+\eta)r\}$,*

where $x_0=(x_0^1,x_0^2)$ is a fixed point on the plane, and a,b and $0<\eta<1$ are positive constants, j is an arbitrary vertex in $G(X,r)$ with X_j as its position (random vector), then the number of vertexes in the intersection part is $M_d=nA(1+o(1))$.

Proof Given $\omega\in B(a_n)$, denote by N_s^- and N_s^+ the number of small squares lying in the interior of the intersection part and intersecting with the intersection part respectively, then

$$N_s^-\ge\frac{A-\sqrt{2}La_n(1+o(1))}{a_n^2(1+o(1))}=\frac{A}{a_n^2}(1+o(1))-\frac{\sqrt{2}L}{a_n}(1+o(1)).$$

On the other hand

$$N_s^+\le\frac{A+\sqrt{2}La_n(1+o(1))}{a_n^2(1+o(1))}=\frac{A}{a_n^2}(1+o(1))+\frac{\sqrt{2}L}{a_n}(1+o(1)).$$

Hence, by

$$na_n^2N_s^-(1+o(1))\le M_d\le na_n^2N_s^+(1+o(1)),$$

we get

$$|M_d-nA(1+o(1))|\le\sqrt{2}Lna_n(1+o(1)).$$

Notice that $\frac{La_n}{A}=o(1)$, therefore

$$M_d=nA(1+o(1)).$$

∎

Theorem 2 *For random geometric graph $G(X,r)$, on the set $B(a_n)$, we have for n sufficiently large,*

i) *for $0<r<\frac{1}{2}$*

$$d_{\min}=\frac{n\pi r^2}{4}(1+o(1)),\quad d_{\max}=n\pi r^2(1+o(1));\tag{20}$$

ii) *for* $r \geq \frac{1}{2}$

$$\frac{\pi}{64}n(1+o(1)) \leq d_{\min} \leq d_{\max} \leq n; \tag{21}$$

iii) *denote by* R_j *the number of vertexes in the intersection part of the ring* $\mathcal{R}_j = \{x : (1-\eta)r \leq \|x - X_j\| \leq (1+\eta)r\}$ *with the unit square* $\mathcal{S}$*, where* $R_{\max} = \max\limits_j R_j$*, then*

$$R_{\max} \leq 4n\pi\eta r^2(1+o(1)). \tag{22}$$

Proof i) Given $\omega \in B(a_n)$, when n is sufficiently large, it is always possible to find a vertex j in the small square nearest to the center of $\mathcal{S}$, whose neighborhood disk is entirely contained in the interior of $\mathcal{S}$, making the area $A = \pi r^2$ in Lemma 5. Hence

$$d_j = n\pi r^2(1+o(1)),$$

which in conjunction with the obvious fact that $d_{\max} \leq n\pi r^2(1+o(1))$ gives

$$d_{\max} = n\pi r^2(1+o(1)).$$

As far as vertexes in the margin of $\mathcal{S}$ are concerned, their neighborhood disks intersect with $\mathcal{S}$ with a minimal area $\frac{\pi r^2}{4}$ (that of vertexes in the four corners is most close to this value). When n is sufficiently large, it is always possible to find vertexes in the corners, thus

$$d_{\min} = \frac{n\pi r^2}{4}(1+o(1)), \quad d_{\max} = n\pi r^2(1+o(1))\,.$$

ii) When $r \geq \frac{1}{2}$, the area of the intersection part between the neighborhood disk and $\mathcal{S}$ satisfies $A \geq \frac{1}{4}\pi(\frac{1}{4})^2$, hence

$$d_{\min} \geq \frac{\pi}{64}n(1+o(1))\,.$$

iii) Given $\omega \in B(a_n)$, since the maximal area of the intersection part between $\mathcal{R}_j$ and $\mathcal{S}$ is not greater than $4\pi\eta r^2(1+o(1))$, hence

$$R_{\max} \leq 4n\pi\eta r^2(1+o(1)).$$

∎

Remark 1 i) In Proposition 1, when $0 < r < \frac{1}{2}$, β can take the value 4, and when $r > \frac{1}{2}$, β can take the value $\frac{64}{\pi}$.

ii) When $0 < r < \frac{1}{2}$, $\frac{R_{\max}}{d_{\min}} = 16\eta(1+o(1))$; when $r \geq \frac{1}{2}$, $\frac{R_{\max}}{d_{\min}} \leq 256r^2\eta(1+o(1))$. No matter in which case, we can always pick η so small that $\alpha = \frac{3}{4}$ in Proposition 1, hence

$$\begin{aligned}
\varepsilon &= \begin{cases} 308\dfrac{R_{\max}}{d_{\min}}, & 0 < r < \frac{1}{2}; \\ \left(\dfrac{208}{\sqrt{\pi}} + \dfrac{2^{14}}{\pi^{\frac{3}{2}}}\right)\dfrac{R_{\max}}{d_{\min}}, & r \geq \dfrac{1}{2}. \end{cases} \\
&\leq \begin{cases} 308 \times 16\eta(1+o(1)), & 0 < r < \dfrac{1}{2}, \\ \left(\dfrac{208}{\sqrt{\pi}} + \dfrac{2^{14}}{\pi^{\frac{3}{2}}}\right) \times 256r^2\eta(1+o(1)), & r \geq \dfrac{1}{2}. \end{cases}
\end{aligned}$$

Due to the importance of $\bar{\lambda}$ in Proposition 1, we will give an estimation for it, part of which is to calculate λ_{n-1} which in turn depends on the following lemma whose proof is given by Prof. Feng TIAN and Dr. Mei LU, see Appendix B.

Lemma 6 *Let triangles be extracted from a complete graph K^n in such a way that every time one triangle is extracted with its three edges deleted while the three vertexes remain. Then there exists an algorithm such that the number of residual edges at each vertex is no more than three.*

Proposition 2 *For random geometric graph $G(X,r)$, on the set $B(a_n)$, we have for n sufficiently large*

$$\lambda_{n-1} \leq 2\left(1 - \frac{1}{4(1+2\sqrt{3})^2}(1+o(1))\right). \tag{23}$$

Proof Given $\omega \in B(a_n)$, firstly we split $\mathcal{S}$ equally into $M(\frac{r}{\sqrt{3}}) = \lceil\frac{\sqrt{3}}{r}\rceil^2$ small squares labelled by $S_k (1 \leq k \leq M(\frac{r}{\sqrt{3}}))$ with side length b satisfying

$$\frac{r}{r+\sqrt{3}} \leq b = \sqrt{\frac{1}{M(\frac{r}{\sqrt{3}})}} \leq \frac{\sqrt{3}}{3}r < \frac{\sqrt{2}}{2}r$$

(when n is sufficiently large), thus any two vertexes in each small square have a distance less than r, making them linked by an edge, which implies that all the vertexes in the small square and edges among them form a clique. Since the area of each small square satisfies $\frac{r^2}{(\sqrt{3}+r)^2} \leq A = b^2 \leq \frac{r^2}{3}$, according to Lemma 5 we know the number of vertexes M_d in the small square satisfying

$$\frac{nr^2}{(\sqrt{3}+r)^2}(1+o(1)) \leq M_d = nb^2(1+o(1)) \leq \frac{nr^2}{3}(1+o(1)).$$

Suppose the triangles extracted from the clique in S_k according to the algorithm in Lemma 6 form a set Δ_k, the elements of which take the form $G_{\Delta_k} = \{(x,y),(y,z),(z,x)\}$, where x,y,z lie in S_j. Let

$$\Delta = \bigcup_{k=1}^{M(\frac{r}{\sqrt{3}})} \Delta_k, \quad \Delta_e = \{(x,y) \in G_\Delta : G_\Delta \in \Delta\}, \quad \text{and} \quad \Delta_e^c = E\{G(X,r)\} - \Delta_e.$$

For each vertex j which lies in S_k and thus the neighborhood disk entirely contains S_k, there are $d_j - 1$ edges linking to it except the self-loop one, hence at least M_d edges among them belong to Δ_e. Therefore for vertex j, the ratio of the number of edges in Δ_e to the total number of edges linking to j except the self-loop one satisfies

$$\begin{aligned}
\frac{M_d}{d_j - 1} &\geq \frac{nb^2}{d_{\max}}(1+o(1)) \\
&\geq \begin{cases} \dfrac{1}{\pi(r+\sqrt{3})^2}(1+o(1)), & 0 < r < \dfrac{1}{2} \\ \dfrac{r^2}{(r+\sqrt{3})^2}(1+o(1)), & r \geq \dfrac{1}{2} \end{cases} \\
&\geq \begin{cases} \dfrac{1}{\pi(\frac{1}{2}+\sqrt{3})^2}(1+o(1)), & 0 < r < \dfrac{1}{2} \\ \dfrac{1}{(1+2\sqrt{3})^2}(1+o(1)), & r \geq \dfrac{1}{2}. \end{cases}
\end{aligned}$$

Hence for any vector $z \in \mathbb{R}^n$, when n is sufficiently large, we have

$$\begin{aligned}
&\sum_{j\sim k}(z_j - z_k)^2\\
&= \sum_{\{(j,\ k),\ (k,\ l),\ (l,\ j)\}\in\Delta} [(z_j - z_k)^2 + (z_k - z_l)^2 + (z_l - z_j)^2] + \sum_{(j,\ k)\in\Delta_e^c} (z_j - z_k)^2\\
&\leq \sum_{\{(j,\ k),\ (k,\ l),\ (l,\ j)\}\in\Delta} 3(z_j^2 + z_k^2 + z_l^2) + \sum_{(j,\ k)\in\Delta_e^c} 2(z_j^2 + z_k^2)\\
&= \sum_j \left(\sum_{k:\ (j,\ k)\in\Delta_e} \frac{3}{2} z_j^2 \right) + \sum_j \left(\sum_{k:\ (j,\ k)\in\Delta_e^c} 2z_j^2 \right)\\
&= \sum_j \left(M_d \frac{3}{2} z_j^2 \right) + \sum_j ((d_j - 1 - M_d) 2z_j^2)\\
&= \sum_j (d_j - 1)\left(\frac{M_d}{d_j - 1} \frac{3}{2} z_j^2 \right) + \sum_j (d_j - 1)\left(\left(1 - \frac{M_d}{d_j - 1} \right) 2z_j^2 \right)\\
&\leq \sum_j d_j \left(\left(1 - \frac{M_d}{4(d_j - 1)} \right) 2z_j^2 \right)\\
&\leq 2\left(1 - \frac{nb^2}{4d_{\max}}(1 + o(1)) \right) \sum_j d_j z_j^2,
\end{aligned}$$

where we have employed the elementary inequality:

$$(a - b)^2 + (b - c)^2 + (c - a)^2 \leqslant 3(a^2 + b^2 + c^2).$$

Therefore, according to equation (31), we get

$$\begin{aligned}
\lambda_{n-1} &= \sup_z \frac{\sum_{j\sim k}(z_j - z_k)^2}{\sum_j z_j^2 d_j}\\
&\leq 2\left(1 - \frac{nb^2}{4d_{\max}}(1 + o(1)) \right)\\
&\leq \begin{cases} 2\left(1 - \dfrac{1}{\pi(1 + 2\sqrt{3})^2}(1 + o(1)) \right), & 0 < r < \dfrac{1}{2}\\ 2\left(1 - \dfrac{1}{4(1 + 2\sqrt{3})^2}(1 + o(1)) \right), & r \geq \dfrac{1}{2} \end{cases}\\
&\leq 2\left(1 - \frac{1}{4(1 + 2\sqrt{3})^2}(1 + o(1)) \right).
\end{aligned}$$

∎

In the following, we will estimate λ_1. First of all, we need the following lemma which is a slight improvement of a Lemma in [18].

Lemma 7[18] *Let $G = (V, E)$ be an undirected graph with n vertexes and suppose that there exists a set $\mathcal{P}$ of $\binom{n}{2}$ pathes joining all pairs of vertexes such that each path in $\mathcal{P}$ has a length at most l and each edge of G is contained in at most m paths in $\mathcal{P}$. Then the eigenvalue λ_1 satisfies*

$$\lambda_1 \geq \frac{nd_{\min}}{d_{\max}^2 ml}.$$

Proposition 3 *For random geometric graph $G(X, r)$, on the set $B(a_n)$, we have for n sufficiently large*

$$\lambda_1 \geq \frac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)). \tag{24}$$

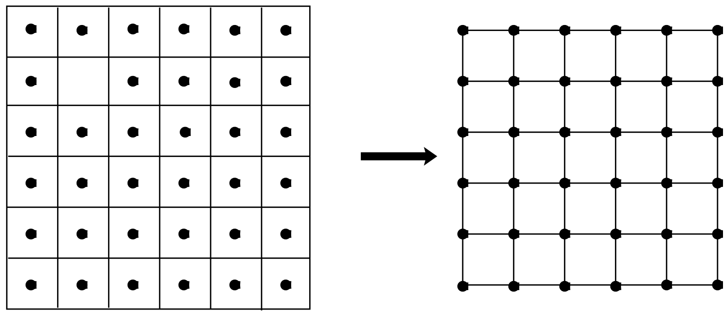
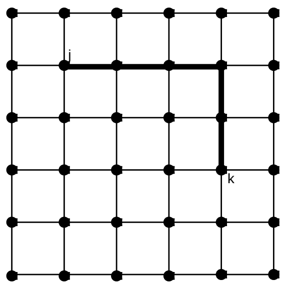


Figure 1 Virtual graph G'



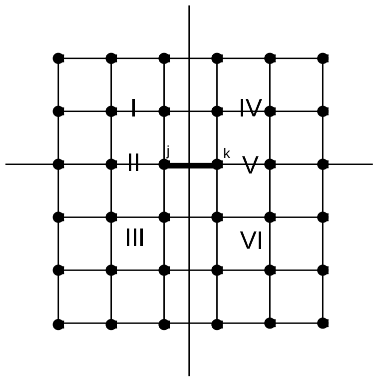


Figure 2 Construction of $\mathcal{P}_S$ **Figure 3** Path usage frequency

Proof Given $\omega \in B(a_n)$, we first split $\mathcal{S}$ equally into $M(\frac{r}{\sqrt{6}}) = \lceil \frac{\sqrt{6}}{r} \rceil^2$ small squares labelled by $S_j (1 \leq j \leq M(\frac{r}{\sqrt{6}}))$ with side length b satisfying $\frac{r}{r+\sqrt{6}} \leq b \leq \frac{\sqrt{6}}{6} r < \frac{\sqrt{5}}{5} r$ (when n is sufficiently large), thus for any small square, any vertex in it has a distance less than r with any vertex in those small squares vertically or horizontally adjacent to it, making the two vertices linked by an edge. Similar to Proposition 2, we could calculate the number of vertices in each small square to be $M_d = nb^2(1+o(1))$.

As illustrated in Figure 1, suppose that there is a virtual vertex in the center of each small square and every virtual vertex is jointed by virtual edges with at most 4 other virtual vertices surrounding it. These virtual vertices together with virtual edges form a grid graph G' with $M(\frac{r}{\sqrt{6}})$ vertices.

We will construct a set $\mathcal{P}_S$ of $\binom{M(\frac{r}{\sqrt{6}})}{2}$ virtual pathes joining all pairs of virtual vertices in G' as follows: for any two virtual vertices j and k as illustrated in Figure 2, without lose of generality, we assume that j lies on the left of k. First we begin from j and select virtual edges on the straight line from left to right until we arrive the right above or below the virtual vertex

k, then we select virtual edges on the straight line from top to bottom or bottom to top. With such a method we have constructed a virtual path from j to k, the length of which is not larger than $l_v = \frac{2}{b}$.

We can now compute how many times at most a virtual edge is used in pathes of $\mathcal{P}_S$. Without lose of generality we pick a virtual edge, for example edge (j,k) from left to right. Divide $\mathcal{S}$ into six parts as illustrated in Figure 3, then one virtual path in $\mathcal{P}_S$ uses edge (j,k) if and only if the starting virtual vertex lies in II and the ending virtual vertex lies in IV, V, or VI. According to this, we can compute that each virtual edge of G' is contained in at most $m_v = \frac{1}{4b^3}$ virtual pathes.

Next, let us construct the path set $\mathcal{P}$ in Lemma 7 for graph $G(X,r)$: for any vertices j,k in $G(X,r)$, if they lie in the same small square, then select the edge joining them into $\mathcal{P}$; Otherwise, the virtual vertices in the two small squares, say S_μ, S_ν, in which the vertices j,k lie, must have a virtual path in $\mathcal{P}_S$ joining them. Now the problem has been reduced to how to replace these virtual edges in the virtual path by edges in $G(X,r)$ to form a path in $\mathcal{P}$. On the one hand, both S_μ and S_ν have M_d vertices in them, thus a virtual path joining the virtual vertices centered in them will be used for $M_d \times M_d$ times; On the other hand, for each virtual edge in the virtual path, we have $M_d \times M_d$ real edges of $G(X,r)$ to substitute it. A careful allocation of these real edges will make them contained in at most 4 paths in $\mathcal{P}$ joining the real vertices in S_μ and S_ν. For $\mathcal{P}$ constructed with this method, any edge in $G(X,r)$ is contained in at most $m = 4 \times m_v = \frac{1}{b^3}$ pathes in $\mathcal{P}$, and the length of each path is not larger than $l = l_v = \frac{2}{b}$. Hence according to Lemma 7 with n sufficiently large, we get for $0 < r < \frac{1}{2}$

$$\begin{aligned}\lambda_1 &\geq \frac{nd_{\min}}{d_{\max}^2 ml} \geq \frac{n\frac{n\pi r^2}{4}(1+o(1))}{(n\pi r^2)^2 \times \frac{1}{b^3} \times \frac{2}{b}(1+o(1))} \\ &= \frac{b^4}{8\pi r^2}(1+o(1)) \geq \frac{r^2}{8\pi(r+\sqrt{6})^4}(1+o(1)) \\ &\geq \frac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)).\end{aligned}$$

Similarly, for $r \geq \frac{1}{2}$,

$$\begin{aligned}\lambda_1 &\geq \frac{nd_{\min}}{d_{\max}^2 ml} \geq \frac{n\frac{\pi n}{64}(1+o(1))}{n^2 \times \frac{1}{b^3} \times \frac{2}{b}(1+o(1))} \\ &= \frac{\pi b^4}{128}(1+o(1)) \geq \frac{\pi r^4}{128(r+\sqrt{6})^4}(1+o(1)) \\ &\geq \frac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)).\end{aligned}$$

Hence, we get

$$\lambda_1 \geq \frac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)). \qquad \blacksquare$$

Combining the above propositions, we get an estimation for the spectral gap $\bar{\lambda}$:

Theorem 3 *For the random geometric graph $G(X,r)$ with n sufficiently large, we have on the set $B(a_n)$,*

$$\bar{\lambda} \leq 1 - \frac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)). \tag{25}$$

Proof We have

$$n\lambda_{n-1} \geq \sum_{j=0}^{n-1} \lambda_j = \mathrm{trace}\{\mathcal{L}\} = \sum_{k=1}^{n} \Big(1 - \frac{1}{d_k}\Big) \geq n\Big(1 - \frac{1}{d_{\min}}\Big)$$

thus $\lambda_{n-1} \geq 1 - \frac{1}{d_{\min}}$. Combining Theorem 5, Propositions 2 and 3, we get

$$2\bigg(1 - \frac{1}{4(1+2\sqrt{3})^2}(1+o(1))\bigg) \geq \lambda_{n-1} \geq 1 - \frac{1}{d_{\min}}, \tag{26}$$

$$\lambda_1 \geq \frac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)). \tag{27}$$

When n is sufficiently large,

$$\begin{aligned} &|1-\lambda_{n-1}| \\ &\leq \max\left\{\frac{1}{d_{\min}}, 1 - \frac{1}{2(1+2\sqrt{3})^2}(1+o(1))\right\} \\ &\leq 1 - \frac{1}{2(1+2\sqrt{3})^2}(1+o(1)) \end{aligned}$$

and

$$\begin{aligned} |1-\lambda_1| &= \begin{cases} 1-\lambda_1, & \lambda_1 \leq 1 \\ \lambda_1 - 1, & \lambda_1 > 1 \end{cases} \\ &\leq \begin{cases} 1 - \dfrac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)), & \lambda_1 \leq 1 \\ \lambda_{n-1} - 1, & \lambda_1 > 1 \end{cases} \\ &\leq \begin{cases} 1 - \dfrac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)), & \lambda_1 \leq 1 \\ 1 - \dfrac{1}{2(1+2\sqrt{3})^2}(1+o(1)), & \lambda_1 > 1. \end{cases} \end{aligned}$$

Therefore

$$\begin{aligned} \bar{\lambda} &= \max_{1\leq j\leq n-1}\{|1-\lambda_j|\} \\ &= \max\{|1-\lambda_1|, |1-\lambda_{n-1}|\} \\ &\leq 1 - \frac{\pi r^2}{512(r+\sqrt{6})^4}(1+o(1)). \end{aligned}$$ ∎

Remark 2 According to Theorem 3 and Remark 1, we can always take η so small that

$$\bar{\lambda} + \varepsilon \leq 1 - \frac{\pi r^2}{1024(r+\sqrt{6})^4}(1+o(1)) = 1 - C_r,$$

where

$$C_r = \frac{\pi r^2}{1024(r+\sqrt{6})^4}(1+o(1)). \tag{28}$$

Next, we will deal with $\|\theta(1)\|$ and $\delta(\theta(1))$. For that we need the following lemma. Throughout the sequel, we denote $h_n = \frac{na_n^2}{g_n logn}$, which satisfies $\lim\limits_{n\to\infty} h_n = \infty$ by the choice of a_n and g_n.

Lemma 8 *Let* $\widetilde{\boldsymbol{S}}_j = \sum_{k\in S_j} \theta_k(0)$, $j = 1, 2, \cdots, M_n(a_n)$, *then*

$$Pr\left\{\max_{1\le j\le M_n(a_n)} |\widetilde{\boldsymbol{S}}_j| \le na_n^2\pi\sqrt{\frac{2}{h_n}}(1+o(1))\middle| B(a_n)\right\} = 1 - O\Big(\frac{1}{n^{g_n}}\Big). \tag{29}$$

Proof Let $B_n = B(a_n), M_n = M_n(a_n)$. First decompose B_n as a finite union, i.e., $B_n = \bigcup_\alpha B_{n_\alpha}$, where $B_{n_\alpha} \triangleq \{\text{Fixed } N_j \text{ out of } n \text{ agents lie in } S_j,\ j = 1, 2, \cdots, M_n\} \bigcap B_n$. Hence under the condition of B_{n_α}, $\{\widetilde{\boldsymbol{S}}_j, j = 1, 2, \cdots, M_n(a_n)\}$ are independent random variables. Let

$$\varepsilon_n = \sqrt{\frac{2\pi^2(g_n\log n - \log a_n^2)}{na_n^2}} = \sqrt{\frac{2\pi^2 g_n\log n}{na_n^2}(1+o(1))} = \pi\sqrt{\frac{2}{h_n}}(1+o(1)),$$

then

$$\begin{aligned}
&Pr\left\{\bigcap_{1\le j\le M_n} |\widetilde{\boldsymbol{S}}_j| \le \varepsilon_n N_j \middle| B_{n_\alpha}\right\}\\
&= \prod_{j=1}^{M_n} Pr\Big\{|\widetilde{\boldsymbol{S}}_j| \le \varepsilon_n N_j \Big| B_{n_\alpha}\Big\} = \prod_{j=1}^{M_n}\Big(1 - Pr\Big\{|\widetilde{\boldsymbol{S}}_j| > \varepsilon_n N_j\Big|B_{n_\alpha}\Big\}\Big)\\
&= \prod_{j=1}^{M_n}\exp\left\{\log\Big(1 - Pr\Big\{|\widetilde{\boldsymbol{S}}_j| > \varepsilon_n N_j\Big|B_{n_\alpha}\Big\}\Big)\right\}.
\end{aligned}$$

For arbitrarily given j,

$$Pr\Big\{|\widetilde{\boldsymbol{S}}_j| > \varepsilon_n N_j\Big|B_{n_\alpha}\Big\} \le \frac{Var(\widetilde{\boldsymbol{S}}_j|B_{n_\alpha})}{(\varepsilon_n N_j)^2} = \frac{\pi^2}{3\varepsilon_n^2 N_j} = O\left(\frac{h_n}{na_n^2}\right) \to 0.$$

Thus, by the Hoeffding inequality[17],

$$\begin{aligned}
&Pr\left\{\bigcap_{1\le j\le M_n} |\widetilde{\boldsymbol{S}}_j| \le \varepsilon_n N_j \middle| B_{n_\alpha}\right\}\\
&= \prod_{j=1}^{M_n}\exp\Big\{-Pr\Big\{|\widetilde{\boldsymbol{S}}_j| > \varepsilon_n N_j\Big|B_{n_\alpha}\Big\}(1+o(1))\Big\}\\
&\ge \prod_{j=1}^{M_n}\exp\left\{-2\exp\left(-\frac{\varepsilon_n^2 na_n^2}{2\pi^2}(1+o(1))\right)\right\}\\
&= \exp\left\{-2\frac{1}{a_n^2}\exp\left(-\frac{\varepsilon_n^2 na_n^2}{2\pi^2}\right)(1+o(1))\right\}\\
&\ge 1 - \frac{2}{a_n^2}\exp\left(-\frac{\varepsilon_n^2 na_n^2}{2\pi^2}\right)(1+o(1))\\
&= 1 - O\Big(\frac{1}{n^{g_n}}\Big).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&Pr\bigg\{\bigcap_{1\le j\le M_n}|\widetilde{\boldsymbol{S}}_j|\le\varepsilon_n N_j\bigg|B_n\bigg\}\\
&=\frac{Pr\bigg\{\bigcap_{1\le j\le M_n}|\widetilde{\boldsymbol{S}}_j|\le\varepsilon_n N_j,B_n\bigg\}}{Pr\{B_n\}}\\
&=\frac{\sum_\alpha Pr\bigg\{\bigcap_{1\le j\le M_n}|\widetilde{\boldsymbol{S}}_j|\le\varepsilon_n N_j,B_{n_\alpha}\bigg\}}{Pr\{B_n\}}\\
&=\frac{\sum_\alpha Pr\bigg\{\bigcap_{1\le j\le M_n}|\widetilde{\boldsymbol{S}}_j|\le\varepsilon_n N_j\bigg|B_{n_\alpha}\bigg\}Pr\{B_{n_\alpha}\}}{Pr\{B_n\}}\\
&=\frac{\sum_\alpha(1-O(\frac{1}{n^{g_n}})Pr\{B_{n_\alpha}\}}{Pr\{B_n\}}\\
&=1-O\bigg(\frac{1}{n^{g_n}}\bigg).
\end{aligned}$$

Finally, by noticing $\varepsilon_n N_j=na_n^2\pi\sqrt{\frac{2}{h_n}}(1+o(1))$, we get

$$Pr\bigg\{\max_{1\le j\le M_n(a_n)}|\widetilde{\boldsymbol{S}}_j|\le na_n^2\pi\sqrt{\frac{2}{h_n}}(1+o(1))\bigg|B(a_n)\bigg\}=1-O\bigg(\frac{1}{n^{g_n}}\bigg). \qquad \blacksquare$$

Now, let us denote

$$D(a_n,g_n,h_n)=\bigg\{\omega:\max_{1\le j\le M_n(a_n)}|\widetilde{\boldsymbol{S}}_j|\le na_n^2\pi\sqrt{\frac{2}{h_n}}(1+o(1))\bigg\}.$$

Then, it is obvious that

$$Pr\{D(a_n,g_n,h_n)\cap B(a_n)\}=Pr\{D(a_n,g_n,h_n)|B(a_n)\}Pr\{B(a_n)\}=1-O\bigg(\frac{1}{n^{g_n}}\bigg).$$

Theorem 4 *On the set $B(a_n)\bigcap D(a_n,g_n,h_n)$, we have for n sufficiently large*

$$\|\theta(1)\|\le\sqrt{n}\pi\sqrt{\frac{2}{h_n}}(1+o(1)),\quad \delta(\theta(1))\le 2\pi\sqrt{\frac{2}{h_n}}(1+o(1)).$$

Proof Given $\omega\in B(a_n)\bigcap D(a_n,g_n,h_n)$, for any agent k, denote by J_k the index set of those small squares intersecting with the neighborhood disk of agent k, and A the area of the intersection part between the neighborhood disk and the unit square $\mathcal{S}$, then

$$\begin{aligned}
\bigg|\sum_{j\in\mathcal{N}_k(0)}\theta_j(0)\bigg|&\le\sum_{j\in J_k}|\widetilde{\boldsymbol{S}}_j|\le\sum_{j\in J_k}na_n^2\pi\sqrt{\frac{2}{h_n}}(1+o(1))\\
&=\frac{A}{a_n^2}(1+o(1))\times na_n^2\pi\sqrt{\frac{2}{h_n}}(1+o(1))\\
&=\pi nA\sqrt{\frac{2}{h_n}}(1+o(1)),
\end{aligned}$$

hence

$$| \theta_k(1)| = \frac{1}{n_k(0)}\left|\sum_{j\in\mathcal{N}_k(0)}\theta_j(0)\right| \le \frac{1}{n_k(0)}\times \pi n A\sqrt{\frac{2}{h_n}}(1+o(1)) = \pi\sqrt{\frac{2}{h_n}}(1+o(1)),$$

therefore

$$\|\theta(1)\| \le \sqrt{\sum_{j=1}^{n}\theta_j^2(1)} \le \sqrt{n}\pi\sqrt{\frac{2}{h_n}}(1+o(1))$$
$$\delta(\theta(1)) \le 2\pi\sqrt{\frac{2}{h_n}}(1+o(1)).$$

∎

5 The Proof of Theorem 1

Let us take the positive sequences $\{a_n, g_n, h_n, n\in\mathbb{N}\}$ as

$$a_n = \frac{1}{\log n}, \quad h_n = \log^3 n, \quad g_n = \frac{n}{\log^6 n}.$$

Then, on the set $B(a_n)\bigcap D(a_n, g_n, h_n)$ with n sufficiently large, we have by Theorem 4 and Remark 2,

$$\begin{aligned}&\frac{v\delta(\theta(1))}{1-(\bar{\lambda}+\varepsilon)}\left(2+\log\frac{2\sqrt{\beta}\parallel\theta(1)\parallel}{\delta\theta(1)}\right)\\ \le\ &\frac{2v\pi\sqrt{\frac{2}{h_n}}}{C_r}\left(2+\log\frac{2\sqrt{2\beta}\pi\sqrt{\frac{n}{h_n}}}{2\pi\sqrt{\frac{2}{h_n}}}\right) = \frac{v\pi}{C_r}\sqrt{\frac{2}{h_n}}\log n(1+o(1)).\end{aligned}$$

Thus in order to satisfy condition iii) in Proposition 1, it is sufficient to take n large enough so that

$$v\pi\sqrt{\frac{2}{h_n}}\frac{\log n}{C_r}(1+o(1)) \leqslant \eta r,$$

which is obviously true by the choice of the sequences $\{a_n, g_n, h_n, n\in\mathbb{N}\}$ above. Thus when n is sufficiently large, the probability for convergence will be greater than or equal to $Pr\{B(a_n)D(a_n,g_n,h_n)\} = 1-O(\frac{1}{n^{g_n}})$. This completes the proof.

6 Conclusions

By working in a probabilistic framework, we are able to show that the multi-agent systems described by a simplified Vicsek model will synchronize with large probability for large population and for any model parameters. In another paper [19], we also considered the case where both the radius r and the speed v may depend on the population size n. If we denote them by r_n and v_n respectively, then under the parameter conditions that $\frac{v_n}{r_n^5} = O(\frac{1}{\log n})$, and $\left(\frac{\log n}{n}\right)^{\frac{1}{6}} = o(r_n)$, $r_n = o(1)$, similar synchronization result can also be established[19]. To the best of our knowledge, this kind of synchronization results for multi-agent systems are established for the first time. Of course, many interesting problems still remain open, for example, the robustness to noises, the phase transition, and the analysis of the more complicated Boid model, etc., and all of these belong to future investigation.

Appendix A: Some Preliminaries in Graph Theory

The problem formulation and analysis in this paper relies on some basic concepts in graph theory, algebraic graph theory, spectral graph theory and random graph theory which are collected in the following and may be found in [18, 20–27].

A graph (undirected) is an ordered pair $G = (V, E)$ that consists of a set of vertexes $V = V(G) = \{1, 2, \cdots, n\}$ and a set of edges $E = E(G) \subseteq \big\{(i,j) : (i,j)\,\text{is an unordered pair of vertexes}\big\}$, where self-loop is allowed. Two vertexes are called adjacent, or neighboring with each other, if $(i,j) \in E$, denoted by $i \sim j$. If all vertexes of a graph$G$ are adjacent, the graph G is called complete; a complete graph with n vertexes is denoted by K^n. The neighborhood set of all the vertexes that are adjacent to vertex i in graph G is denoted by $N_i = N_i(G) = \{j \in V : (i,j) \in E\}$; denote $n_i = |N_i|$. A vertex with empty neighborhood set N_i is called isolate.

The intersection $G \cap G'$ of two graphs $G = (V, E)$ and $G' = (V', E')$ is also a graph $(V \cap V', E \cap E')$, and the union $G \cup G'$ of them is $(V \cup V', E \cup E')$. If $V' \subseteq V$ and $E' \subseteq E$, then we call G' a subgraph of G; a complete subgraph is called a clique.

A path of a graph G is a subgraph $P = (W, H)$, where $W = \{i_1, i_2, \cdots, i_k\} \subseteq V(G)$, $H = \{(i_1, i_2), (i_2, i_3), \cdots, (i_{k-1}, i_k)\} \subseteq E(G)$, and i_j are mutually different; The number of edges in a path is called the length of the path. Usually a path is denoted by the natural sequence of vertexes in it, for example, a path from i_1 to i_k, $P = i_1 i_2 \cdots i_k$. A graph G is connected if for any two vertexes of it, there exists a path connecting them. A set of graphs $\{G_1, G_2, \cdots, G_m\}$ is jointly connected if the union of them is connected.

If $P = i_1 i_2 \cdots i_k$ is a path with $k \geq 3$, then the graph $C = P \cup i_k i_1$ is called a loop, denoted by $i_1 i_2 \cdots i_k i_1$. The number of edges or vertexes in a loop is called its length. A loop with length k is called a k-loop, denoted by C^k; A C^3 is called a triangle.

The adjacency matrix $A = A(G) = (a_{ij})_{n\times n}$ of a graph G with n vertexes is a symmetric matrix with nonnegative elements satisfying $a_{ij} \neq 0 \Leftrightarrow (i,j) \in E(G)$. The graph is called weighted whenever the elements of its adjacency matrix are other than just 0-1 elements. $a_{ij} > 0$ is called the weight of edge (i,j); $d_i = d_i(G) = \sum\limits_{j\in V} a_{ij}$ is called the degree of vertex i, which satisfies $d_i = |N_i|$ in the case of non-weighted graph; $T = \mathrm{diag}(d_1, d_2, \cdots, d_n)$ is called degree matrix; $d_{\max} = \max\{d_i : i \in V\}$ and $d_{\min} = \min\{d_i : i \in V\}$ are, respectively, called the maximum degree and the minimum degree; $\mathrm{Vol}(G) = \sum_{i=1}^n d_i$ is called the volume of graph G.

The Laplacian of a graph G is the matrix $L = T - A$. However, in discrete time problems, the so called normalized Laplacian is used more frequently. In a graph without isolated vertexes, the degree matrix T is invertible; on the other hand, in a graph with isolated vertexes,we follow the convention that the diagonal elements in $T^{-1}, T^{-\frac{1}{2}}$ corresponding to isolated vertexes take value 0. Thus, the normalized Laplacian of graph G is defined as $\mathcal{L} = T^{-\frac{1}{2}} L T^{-\frac{1}{2}}$, and $P = T^{-1}A = T^{-\frac{1}{2}}(I - \mathcal{L})T^{\frac{1}{2}}$ as the average matrix of graph G. In the following we will see that each vertex in the neighborhood graph of the Linearized Vicsek model has a loop to itself, making the degree matrix T always invertible.

From now on we will consider only non-weighted graph G with a self-loop in each vertex. The normalized Laplacian $\mathcal{L}$ is semi-definite, and thus the n eigenvalues of $\mathcal{L}$ are all nonnegative real numbers, which are denoted by $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1}$ with the corresponding orthogonal unit vector bundle $\{\phi_j, 0 \leq j \leq n-1\}$, where λ_1 is usually called the normalized algebraic connectivity. The number defined by $\bar{\lambda} = \max\limits_{1\leq j\leq n-1} |1 - \lambda_j| = \max\{|1 - \lambda_1|, |1 - \lambda_{n-1}|\}$ is called the spectral gap. The eigenvalues λ_1 and λ_{n-1} have the following Rayleigh quotient representations[18]

$$\lambda_1 = \inf_{z\perp\ T1_n} \frac{\sum\limits_{i\sim j}(z_i - z_j)^2}{\sum\limits_{j\in V(G)} z_j^2 d_j}, \quad \lambda_{n-1} = \sup_{z} \frac{\sum_{i\sim j}(z_i - z_j)^2}{\sum\limits_{j\in V(G)} z_j^2 d_j}. \tag{30}$$

Suppose $\{x_i \in \mathbb{R}^m, 1 \le i \le n\}$ are n points in the m dimensional Euclidean space, and let $x = (x_1, x_2, \cdots, x_n)^\tau \in \mathbb{R}^{mn}$, and $r > 0$ be the interaction radius. The Geometric Graph $G(x, r)$ is an undirected graph (V, E) with a self-loop in each vertex, where $V = \{1, 2, \cdots, n\}$ and $E = \{(i, j) : \|x_i - x_j\| < r, i \in V, j \in V\}$.

The geometric graph of n points in Euclidean space contains the distance information of those points. Because of the finiteness of the interaction radius r, the distance information is in some sense local and thus the random geometric graph is particularly useful in modeling systems with local interaction rules, such as Vicsek's model, Boid model, wireless sensor network, ad hoc network etc.

Suppose that $\{X_i \in \mathbb{R}^m, 1 \le i \le n\}$ are i.i.d. random vectors uniformly distributed in the unit cube $\mathcal{S}_m = \{z \in \mathbb{R}^m : 0 \le z_j \le 1, 1 \le j \le m\}$. Let $X = (X_1, X_2, \cdots, X_n)^\tau \in \mathbb{R}^{mn}$, and $r > 0$ be the interaction radius. The geometric graph $G(X, r)$ is called a Random Geometric Graph.

Random geometric graph is a newly proposed random graph model. In the modeling of some problems, it is practically much better than traditional random graphs, and thus has drawn more and more attention. The interested readers can refer to [16,26,27].

Appendix B: The Proof of Lemma 6

In this appendix, we will elaborate the detailed proof of Lemma 6, which is given by Prof. Feng TIAN and Dr. Mei LU.

In the following, all the graphs are non-weighted simple ones without self-loops. Suppose that G is a graph with n vertices and C^3 a subgraph of it. We use $G = \&C^3 + H$ to denote that after deleting the three edges of any triangle in G with a certain algorithm, the residual graph (with n vertices) is denoted by H.

Definition B.1 A graph G is called odd if the degree of each vertex is an odd number; Similarly, if the degree of each vertex is an even number, it is called an even graph.

It is quite obvious that some graphs are neither odd nor even. However, a complete graph K^n is odd when n is an odd number and is even n is an even number. We also have the following conclusion:

i) If $G = \&C^3 + H$, then graph G and H are both odd or even;

ii) If $G = \&C^3 + \emptyset$, then G is even and the number of edges satisfying $|E(G)| \equiv 0 \pmod 3$.

Lemma B.1 *If $K^n = \&C^3 + H$, then*

$$d_{\max}(H) \ge \begin{cases} 0, & n = 6k+1,\ 6k+3; \\ 1, & n = 6k,\ 6k+2; \\ 2, & n = 6k+5; \\ 3, & n = 6k+4. \end{cases} \tag{31}$$

Proof When $n = 6k + 1$ or $6k + 3$, the conclusion is obvious.

If $K^{6k} = \&C^3 + H$, since K^{6k} is odd, and H is also odd, we have $d_{\max}(H) \ge 1$. The same argument applies when $n = 6k + 2$.

If $K^{6k+5} = \&C^3 + H$, since

$$E(K^{6k+5}) = \frac{(6k+5)(6k+4)}{2} = (6k+5)(3k+2) \equiv 1 \text{ (mod 3)}$$

we have $|E(H)| \geq 1$. K^{6k+5} is even, hence $d_{\max}(H) \geq 2$.

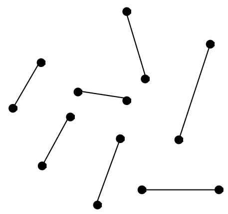


Figure 4

If $K^{6k+4} = \&C^3 + H$, then K^{6k+4} is odd; consequently H is odd; hence $d_{\max}(H) \geq 1$, $d_{\min}(H) \geq 1$. We only have to prove $d_{\max}(H) \neq 1$. If this is not true, we have $d_{\max}(H) = d_{\min}(H) = 1$, hence H could only be a graph as illustrated in Figure 4, and thus $|E(H)| = 3k+2$. However,

$$|E(K^{6k+4})| - |E(H)| = \frac{(6k+4)(6k+3)}{2} - (3k+2) = (3k+2)(6k+2) \equiv 1 \text{ (mod 3)},$$

we get a contradiction. Therefore $d_{\max}(H) \geq 3$. ∎

A subset of p elements of a nonempty set is called a p-subset. We have the following definition:

Definition B.2[28,29] Suppose $S = \{1, 2, \cdots, v\}$, then a Balanced Incomplete Block Design (BIBD) of S is a family of b k-subsets denoted by $\{B_1, B_2, \cdots, B_b\}$ satisfying the following constraints:

i) Each element of S is contained exactly in r out of the b k-subsets;

ii) Any two elements of S is contained at the same time exactly in λ out of the b k-subsets;

iii) $k < v$.

The $BIBD$ is also called (b, v, r, k, λ)-Design. But, only three of the five parameters of (b, v, r, k, λ)-Design are independent[28,29], which have the following relationships

$$bk = vr, \qquad r(k-1) = \lambda(v-1). \tag{32}$$

When $k = 3$, $\lambda = 1$, (b, v, r, k, λ)-Design is called the Steiner Triple System. According to equation (32), we have

$$r = \frac{v-1}{2}, \quad b = \frac{v(v-1)}{6}. \tag{33}$$

We have the following proposition for the existence of the Steiner Triple System:

Proposition B.1[28,29] *The Steiner Triple System exists if and only if* $v = 1, 3$ (mod 6).

Let $f(n) = \min\{d_{\max}(H) : K^n = \&C^3 + H\}$. Then the following corollary is equivalence to the above proposition.

Corollary B.1

$$f(6k+1) = f(6k+3) = 0. \tag{34}$$

Lemma B.2 *When n is an even number,*

$$f(n) = 1 \Longleftrightarrow f(n+1) = 0. \tag{35}$$

Proof Sufficiency: If $f(n) = 1$, then there exists an algorithm such that $K^n = \&C^3 + H$ with $d_{\max}(H) = 1$. Since K^n is odd, H is also odd, which means that the degree of any vertex in H satisfies $d(H) = 1$, and the structure of H is illustrated in Figure 4. We carry on the operations on a subgraph K^n of K^{n+1} directed by the algorithm to make edges of the residual graph H' pairwise match. Since each edge in H' is adjacent to the remaining vertex, all the edges together form $\frac{n}{2}$ triangles as illustrated in Figure 5. Deleting all of them, we get $f(n+1) = 0$.

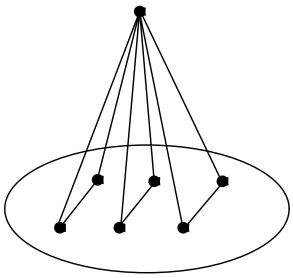


Figure 5

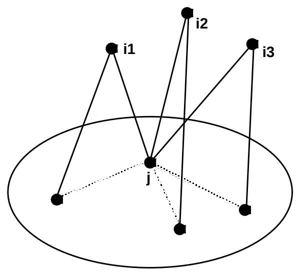


Figure 6

Necessity: If $f(n+1) = 0$, then for $S = \{1, 2, \cdots, n+1\}$, there exists a Steiner Triple System, i.e., a 3-subset family $\{B_1, B_2, \cdots, B_{\frac{n(n+1)}{6}}\}$ of S. Deleting all subsets with the form $S_i = \{i, j, n+1\}$ from the family gives a deleting algorithm for K^n, which then gives $K^n = \&C^3 + H$ with the structure of H as illustrated in Figure 4. Hence $f(n) = 1$. ∎

Combining Corollary B.1 and Lemma B.2, we get:

Corollary B.2

$$f(6k) = f(6k+2) = 1. \tag{36}$$

Lemma B.3

$$f(6k+4) = 3. \tag{37}$$

Proof Because $f(6k+7) = f(6(k+1)+1) = 0$, there exits a deleting algorithm such that $K^{6k+7} = \&C^3 + H$ and $d_{\max}(H) = 0$. Denote by $\mathcal{D}$ the set of triangles deleted by the algorithm. Pick $C_0^3 = i_1 i_2 i_3 \in \mathcal{D}$, and delete triangles $C^3 \in \mathcal{D}$ which contain i_1 or i_2 or i_3, then the residual set $\mathcal{D}'$ corresponds to a deleting algorithm for $K^{6k+4} = \frac{K^{6k+7}}{C_0^3}$. For any vertex j in $K^{6k+4} = \frac{K^{6k+7}}{C_0^3}$, the deleting algorithm will delete all edges linking to it, and those can't be deleted are indicated by dash lines in Figure 6. Thus each vertex of K^{6k+4} has at most 3 edges left, that is $f(6k+4) = 3$. ∎

Finally, combining all the above results, we have

$$f(n) = \begin{cases} 0, & n = 6k+1,\ 6k+3; \\ 1, & n = 6k,\ 6k+2; \\ 2, & n = 6k+5; \\ 3, & n = 6k+4 \end{cases} \tag{38}$$

This completes the proof of Lemma 6.

Acknowledgement

The authors would like to thank Prof. Feng TIAN and Dr. Mei LU for providing the proof of Lemma 6 in Appendix B. We would also like to thank Ms. Zhixin Liu for valuable discussions.

References

[1] E. Shaw, Fish in schools, *Natural History*, 1975, **84**(8): 40–46.

[2] B. L. Partridge, The structure and function of fish schools, *Sci. Amer.*, 1982, **246**(6): 114–123.

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

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

[5] T. Vicsek, A. Czirok, E. Jacob, I. Cohen, and O. Shochet, Novel type of phase transition in a system of self-deriven particles, *Phys. Rev. Lett.*, 1995, **75**(6): 1226–1229.

[6] A. Czirok, A. Barabasi, and T. Vicsek, Collective motion of self propelled particles: Kinetic phase transition in one dimension, *Phys. Rev. Lett.*, 1999, **82**(1): 209–212.

[7] C. W. Reynolds, Flocks, herds, and schools: A distributed behavioral model, in *Comput. Graph.* (ACM SIGGRAPH87 Conf. Proc.), 1987, **21**: 25–34.

[8] A. Jadbabaie, J. Lin, and A. S. Morse, Coordination of groups of mobile agents using nearest neighbor rules, *IEEE Trans. Autom. Control*, 2003, **48**(6): 988–1001.

[9] W. Ren and R. W. Beard, Consensus seeking in multiagent systems under dynamically changing interaction topologies, *IEEE Trans. Autom. Control*, 2005, **50**(5): 655–661.

[10] A. Jadbabaie, On distributed coordination of mobile agents with changing nearest neighbors, Technical Report, University of Pennsylvania, Philadelphia, PA, 2003.

[11] F. Cucker and S. Smale, Emergent behavior in flocks, *IEEE Trans. Autom. Control*, 2006, to appear.

[12] J. Tsitsiklis, D. Bertsekas, and M. Athans, Distributed asynchronous deterministic and stochastic gradient optimization algorithms, *IEEE Trans. Autom. Control*, 1986, **31**(9): 803–812.

[13] Z. X. Liu and L. Guo, Connectivity and synchronization of multi-agent systems, in *Proc. 25th Chinese Control Conference*, August 7–11, Harbin, China, 2006.

[14] LGuo, *Time-Varying Stochastic Systems-Stability, Estimation and Control*, Jilin Science and Technology Press, Changchun, China, 1993.

[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

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

[17] J. Diaz, J. Petit, and M. Serna, A guide to concentration bounds, in *Handbook on Randomized Computing* (Ed. by S. Rajasekaran, P. Pardalos, J. Reif, and J.Rolim), Vol. II, Chapter 12, Kluwer Press, New York, 2001, 457–507.

[18] F. R. K. Chung, *Spectral Graph Theory*, American Mathematical Society, Providence, RI, 1997.

[19] G. G. Tang and L. Guo, Convergence analysis of linearized Vicsek's model, in *Proc. 25th Chinese Control Conference*, August 7–11, Harbin, China, 2006.

[20] D. Reinhar, *Graph Theory* (Second Edition), GTM 173, Springer-Verlag, New York, 2000.

[21] B. Bollobas, *Modern Graph Theory*, GTM 184, Springer-Verlag, New York, 1998.

[22] D. B. West, *Introduction to Graph Theory* (Second Edition), Prentice Hall, Upper Saddle River, NJ, 2001.

[23] B. Bollobas, *Random Graph* (Second Edition), Cambridge University Press, Cambridge, UK, 2001.

[24] N. Biggs, *Algebraic Graph Theory*, Cambridge University Press, Cambridge, UK, 1993.

[25] C. Godsil and G. Royle, *Algebraic Graph Theory*, GTM 207, Springer-Verlag, New York, 2001.

[26] M. Penrose, *Random Geometric Graphs*, Oxford University Press, Oxford, UK, 2003.

[27] P. Gupta and P. R. Kumar, Critical power for asymptotic connectivity in wireless networks, in *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming* (Ed. by W. M. McEneany, G. Yin, and Q. Zhang), Birkhauser, Boston, MA, 1998, 547–566.

[28] Z. S. Yang, *Combinitorial Mathematics and Its Algorithms*, University of Science and Technology of China Press, Hefei, China, 1997.

[29] W. D. Wallis, *Combinatorial Designs*, Marcel Dekker, New York, 1988.

[30] D. Huang and L. Guo, Estimation of Nonstationary ARMAX Models Based on the Hannan-Rissanen Method, *The Annals of Statistics*, 1990, **18**(4): 1729–1756.