ON FLOCKS UNDER SWITCHING DIRECTED INTERACTION TOPOLOGIES

CUCKER, Felipe; DONG, Jiu-Gang

Published in:
SIAM Journal on Applied Mathematics

Published: 01/01/2019

Published version (DOI):
10.1137/18M116976X

Publication details:
CUCKER, F., & DONG, J.-G. (2019). ON FLOCKS UNDER SWITCHING DIRECTED INTERACTION TOPOLOGIES. *SIAM Journal on Applied Mathematics*, *79*(1), 95-110. https://doi.org/10.1137/18M116976X

SIAM J. APPL. MATH.
Vol. 79, No. 1, pp. 95–110

ON FLOCKS UNDER SWITCHING DIRECTED INTERACTION TOPOLOGIES*

FELIPE CUCKER† AND JIU-GANG DONG‡

Abstract. We prove unconditional and conditional flocking results for a Cucker–Smale flocking model enhanced with switching topologies. These topologies can be arbitrary; in particular, they are not necessarily symmetric. Our proofs hold for both discrete and continuous time. In both cases, the critical exponent discriminating between unconditional and conditional flocking is shown to be at most $\frac{1}{2(k-1)}$, where k is the number of agents in the population.

Key words. Cucker–Smale model, flocking, switching topology

AMS subject classifications. 93C15, 34K33

DOI. 10.1137/18M116976X

1. Introduction.

Over the past two decades, there has been considerable interest in modeling and analyzing the collective behaviors of groups of autonomous agents. From ecology and evolutionary biology to computer science and from statistical physics to systems and control theory with engineering applications, efforts have been made towards a better understanding of how a large number of agents such as flocks of birds or schools of fish can organize into an ordered motion without centralized coordination [5, 8, 12, 13, 19, 20, 23, 27]. A comprehensive survey can be found in [28].

In 1987, Reynolds introduced three heuristic rules that led to the creation of the first computer animation of flocking behavior [23]: *alignment* (the attempt to match velocity with local flockmates), *cohesion* (the attempt to stay close to local flockmates), and *separation* (the attempt to avoid collisions with local flockmates). Many models have been proposed to implement the three flocking rules of Reynolds. In its simplest form, disregarding the modeling of noise or the factors ensuring collision avoidance among others, these models follow a simple principle: each agent modifies its velocity (or its heading if speeds are fixed) by adding to it a weighted average of the differences of its velocity with the other agents' velocities. More precisely (and in continuous time), the velocity $\boldsymbol{v}_i$ of the ith agent satisfies

$$\boldsymbol{v}_i' = \sum_{j\neq i} a_{ij}(\boldsymbol{x})(\boldsymbol{v}_j - \boldsymbol{v}_i).$$

What differentiates the various models that have been studied in the literature is the nature of the weights a_{ij}. These are nonnegative real numbers quantifying the influence of agent j over agent i. The form of the *weight matrix* $A = (a_{ij})$ and the evolution of this form with time can be considered under two aspects. Firstly,

*Received by the editors February 8, 2018; accepted for publication (in revised form) December 3, 2018; published electronically January 22, 2019.
http://www.siam.org/journals/siap/79-1/M116976.html
Funding: The first author's work was supported by a GRF grant from the Research Grants Council of the Hong Kong SAR (project number CityU 11202017).
†Department of Mathematics, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong (macucker@cityu.edu.hk).
‡Department of Mathematics, Harbin Institute of Technology, Harbin 150001, People's Republic of China (jgdong@hit.edu.cn).

a *qualitative* aspect given by the consideration of which entries a_{ij} are zero. This determines a graph, called the *interaction graph*, that captures who influences whom in the population. The topology of this graph, its connectedness properties, plays a large role in the occurrence of flocking. Secondly, a *quantitative* aspect given by the magnitudes of the nonzero a_{ij}s.

A basic classification of flocking models can be done by the possible answers to the following three binary questions: (1) is the matrix A symmetric? When this is the case, the interaction graph is taken to be undirected. (2) Is the interaction graph fixed or does it change with time? (3) Are the magnitudes of the nonzero a_{ij}s fixed or do they change with time? The following table gives examples for the eight possible answers to these questions (the columns giving answers to the first two questions, regarding the interaction graph and the rows doing so for the third).

We next briefly describe the various works and flocking models occurring in Table 1. Such a description will give the appropriate frame for the understanding of the results in this paper.

Table 1
A basic classification of flocking models.

	`Sym, Fixed`	`NonSym, Fixed`	`Sym, Switching`	`NonSym, Switching`
a_{ij} fixed	[8, $\beta = 0$]	[25, 11, $\beta = 0$]	[27]	[3, 16, 17, 22, 7]
a_{ij} variable	[8, $\beta \neq 0$]	[25, 11, $\beta \neq 0$]	[9, §3.4]	[14]

One of the first models to be studied is due to Vicsek et al. [27] and is referred to in the literature as the Vicsek model. It consists of a large number of autonomous agents moving on the plane with a constant speed but with different headings. It is therefore the headings that change with time. Through computer simulations Vicsek et al. [27] found that the local rule in their model can cause all agents to eventually move in the same direction despite the fact that the neighbor set of each agent (i.e., the set of agents influencing the behavior of a given agent) changes with time: it consists of all agents that are within a fixed distance. A rigorous theoretical analysis for this behavior was given in [13] where a simplified first-order linear Vicsek model was considered. It was shown there that the reaching of a heading consensus can be guaranteed if the union of the interaction graphs is connected frequently enough as the system evolves. As weight matrices are symmetric in the Vicsek model these interaction graphs are taken to be undirected. As neighbor sets vary with time, so do the resulting interaction graphs.

Extensions of the result in [13] to flocking models with nonsymmetric weight matrices were discussed in [3, 16, 17, 22]. These studies showed that consensus can be asymptotically achieved provided a condition on the evolution of the, now directed, interaction graphs is satisfied (we will spell out this condition in the Assumption 2.1 in section 2.2).

A different extension, also in the category of nonsymmetric weight matrices with switching topology was considered in [7](its kinetic model was discussed in [2]). The neighbor set of an agent is formed now by its q closest neighbors (q being a fixed parameter of the system). The results proved in [7] establish convergence both unconditionally (when q is at least half of the number of agents minus one) or conditionally (otherwise). In the last case the condition is on the initial state only: it does not explicitly involve the evolution of the interaction graphs. But a look at the proof reveals that this condition is so demanding that it ultimately implies that the interaction graph cannot change much with time. In particular, it implies Assumption 2.1.

In 2007 another flocking model, today referred to as the Cucker–Smale (or simply, CS) model was proposed in [8]. This is a second-order nonlinear system where the interaction between two agents is characterized by a nonlinear function of their distance, which is positive but can decrease to zero as the distance between agents becomes large. In the Vicsek model the interaction graph varies with time but the strength of this interaction between any two neighbors is fixed. In contrast, in the CS model, the interaction graph is fixed (it is actually the complete graph) but the strengths of these interactions depend on the distance between agents. A notable feature of the CS model is that convergence to flocking is guaranteed under some conditions on the initial configurations and system's parameters only. More precisely, when the interaction strength decays sufficiently slow asymptotic flocking occurs unconditionally, while with a faster decaying interaction strength, convergence to flocking is guaranteed for all initial configurations satisfying an explicit constraint. This interaction strength is gauged by a system parameter denoted by β (see (2.2) below for a precise description) and slow decay for the interaction strength corresponds to $\beta = \frac{1}{2}$ (we call this value the critical exponent). The particular case $\beta = 0$ fixes the values of the weights a_{ij} to a constant one. Some extensions of the basic model were given in [9]. The weight matrices in [8, 9] are always symmetric but section 3.4 in [9] allows for switching interaction graphs.

The original result for the CS model gave rise to a stream of works extending it in several ways. One of these extensions, introduced by Shen [25], is the notion of *hierarchical leadership*. This assumes a directed interaction graph which remains fixed with time (such as the military ranks within an army during a battle). But allows the strength of these interactions to vary with time. In Shen's work the nonzero interactions follow the law in the CS model. The distinction of $\beta = 0$ or $\beta \neq 0$ thus provide perfect examples for two boxes in Table 1. Further work on the CS model with hierarchical leadership can be found in [6, 10, 15].

A further extension of the CS model having nonsymmetric matrices and fixed interaction graph was considered in [11]. It extends the results for hierarchical leadership as the interaction graphs considered here are arbitrary directed graphs. Again, both unconditional and conditional flocking results were established. The critical exponent, however, depends on some invariants of the interaction graph. Another nonsymmetric extension of the CS model with the complete graph resulting from the nonsymmetric influence between agents was introduced in [18].

Most of the existing work and extensions on the CS model are for fixed interaction graphs. Yet, studies of animals on the move suggest that switching topologies are common in nature. In [19], Nagy et al. placed ultra-light global positioning system (GPS) devices on members of an actual flock of ten pigeons and found dynamically changing but well-defined leadership networks within the flock. In [21], the homing flights of groups of up to 40 pigeons were further tracked using GPS devices. It was found that leadership hierarchies can arise from differences in the birds' typical speeds. In particular, faster individuals flying at the front of the flock tend gradually to learn navigational cues and to become leaders during the collective flight. Clearly, the switching topologies can help the flock to substantially save communication effort. These biological observations motivate the study of extensions of the CS model with switching topologies (even when this means that assumptions on the evolution of these topologies have to be made).

One such study was recently done in [14] where a discrete-time CS model under switching topologies was considered and both unconditional and conditional flocking results were shown under an assumption which is stronger than Assumption 2.1

(mentioned above). In addition, the critical exponent explicitly depends on the evolution of the interaction graph, in contrast with all existing work on the CS model on which this critical exponent is either a universal constant or depends only on the number of agents in the population. This raises the following questions: (a) can one prove a similar result based on Assumption 2.1 only? (b) Can one extend it to hold for continuous time as well? (c) Can one improve the critical exponent to depend on the number of agents only?

This paper provides positive answers to these three questions. More precisely, we prove, for both discrete time and continuous time, unconditional and conditional flocking results for the CS model under switching topologies satisfying Assumption 1. The critical exponent is shown to be (at most) $1/(2(k-1))$ (k being the number of agents). The approach adopted here is based on a blend of algebraic graph theory and nonlinear analysis tools. Our results can be understood as extensions of the known first-order consensus results listed in Table 1 to second-order nonlinear flocking systems.

2. Basic notations and statement of the main results.

2.1. Directed graphs and switching systems. We use directed graphs to model the interaction topology among agents. A *digraph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a finite set $\mathcal{V} = \{1, \ldots, k\}$ of *vertices* and a set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ of *arcs*, these being ordered pairs of vertices of $\mathcal{V}$. If $(j, i) \in \mathcal{E}$, we say that j is the *initial* vertex of the arc and that i is its *end* vertex. We also say that j is a *neighbor* of i. The *neighbor set* of vertex i is $\mathcal{N}_i := \{j : (j, i) \in \mathcal{E}\}$. For the case that $(i, i) \in \mathcal{E}$, we understand that $\mathcal{G}$ has a *self-loop* at vertex i. We say that a nonnegative matrix $A = (a_{ij})_{k\times k}$ is an *adjacency matrix* of a given digraph $\mathcal{G}$ when $a_{ij} > 0$ if and only if $(j, i) \in \mathcal{E}$. In this case, we also say that $\mathcal{G}$ is *induced* by A, and we emphasize this dependence by writing $\mathcal{G}(A)$ instead of $\mathcal{G}$. A *path* in $\mathcal{G}$ from i_0 to i_ℓ is a sequence of distinct vertices $i_0, i_1, \ldots, i_\ell$ such that each successive pair of vertices is an arc of the digraph. If there exists a path from i to j, then vertex j is said to be *reachable* from vertex i. A digraph $\mathcal{G}$ is said to be *rooted* when we can find a vertex (called a root) such that any other vertex of $\mathcal{G}$ is reachable from it.

In this paper, we assume that the interaction graph of our population of agents can switch among a finite set $\mathcal{P} = \{1, 2, \ldots, p\}$ of modes. The switching law is specified by a function $\sigma : [0, \infty) \to \mathcal{P}$, which in continuous-time context is assumed to be piecewise constant and continuous from the right, specifying which topology mode is activated at each time instant. The set of discontinuities of σ is known as the sequence of *switching instants*. This sequence may be finite or even empty but, for generality's sake, we will denote this sequence by $\{t_\ell\}_{\ell=0}^{\infty}$ with $t_0 = 0$. Denote by $\mathcal{G}_{\sigma(t)} = (\mathcal{V}, \mathcal{E}_{\sigma(t)})$ the interaction graph at time t and by $\mathcal{N}_i^{\sigma(t)}$ the neighbor set of agent i in the digraph $\mathcal{G}_{\sigma(t)}$. The *union digraph* of $\mathcal{G}_{\sigma(t)}$ during time interval $[s_1, s_2)$ is defined as $\mathcal{G}([s_1, s_2)) := \bigcup_{t\in[s_1,s_2)} \mathcal{G}_{\sigma(t)} = (\mathcal{V}, \bigcup_{t\in[s_1,s_2)} \mathcal{E}_{\sigma(t)})$.

Notation. Matrix ordering is meant componentwise, e.g., for matrices $A = (a_{ij})_{k\times k}$ and $B = (b_{ij})_{k\times k}$, $A \geq B$ stands for $a_{ij} \geq b_{ij}$ for all i, j. Let Id be the identity matrix with appropriate dimensions. For a real number c, denote by $\lfloor c \rfloor$ the floor of c, i.e., the largest integer no greater than c. Let $\mathbb{N}$ denote the set of all natural numbers (including zero). Denote by $\mathbb{E}$ the Euclidean space $\mathbb{R}^N$.

2.2. Discrete time. Consider a flock of k agents labeled $\{1, \ldots, k\}$ moving in $\mathbb{E}$, whose behavior in discrete-time at time $t \in \mathbb{N}$ is specified by

$$\begin{aligned}&\boldsymbol{x}_i[t+1] = \boldsymbol{x}_i[t] + h\boldsymbol{v}_i[t],\\ &\boldsymbol{v}_i[t+1] = \boldsymbol{v}_i[t] + h \sum_{j\in\mathcal{N}_i^{\sigma[t]}} a_{ij}(\boldsymbol{x}[t])(\boldsymbol{v}_j[t] - \boldsymbol{v}_i[t]),\end{aligned} \tag{2.1}$$

where $h > 0$ is the time step and for $1 \le i \le k$, $\boldsymbol{x}_i[t]$ and $\boldsymbol{v}_i[t] \in \mathbb{E}$ denote the position and velocity of agent i at time th respectively, and $\mathcal{N}_i^{\sigma[t]}$ is the neighbor set of agent i at time th. The interaction topology of the flock at time th is denoted by the digraph $\mathcal{G}_{\sigma[t]} = (\mathcal{V}, \mathcal{E}_{\sigma[t]})$ such that $(j,i) \in \mathcal{E}_{\sigma[t]}$ if and only if $j \in \mathcal{N}_i^{\sigma[t]}$. For convenience, we let $i \in \mathcal{N}_i^{\sigma[t]}$, i.e., $\mathcal{G}_{\sigma[t]}$ has a self-loop at each vertex i. The weight function $a_{ij}(\boldsymbol{x})$ quantifies the way the agents influence each other, which, as required in [8], takes the form

$$a_{ij}(\boldsymbol{x}) = \frac{K}{(1 + \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2)^\beta} \tag{2.2}$$

with system parameters $K > 0$ and $\beta \ge 0$.

For $1 \le i, j \le k$, define

$$\Gamma[t] := \max_{1\le i,j\le k} \|\boldsymbol{x}_i[t] - \boldsymbol{x}_j[t]\|, \qquad \Lambda[t] := \max_{1\le i,j\le k} \|\boldsymbol{v}_i[t] - \boldsymbol{v}_j[t]\|.$$

We say that the agents under system (2.1) *converge to flocking* if the following is satisfied:

$$\lim_{t\to\infty} \Lambda[t] = 0 \text{ and } \sup_{t\ge 0} \Gamma[t] < \infty.$$

Our main objective is to derive conditions on the initial state so that flocking occurs.

We now rewrite system (2.1) in a more concise form. Let $W[t] = (w_{ij}[t])_{k\times k}$ be an adjacency matrix of $\mathcal{G}_{\sigma[t]}$ such that $w_{ij}[t] = a_{ij}(\boldsymbol{x}[t])$ if $j \in \mathcal{N}_i^{\sigma[t]}$ and $w_{ij}[t] = 0$ otherwise. Note that $w_{ii}[t] = K$ for all $t \ge 0$. Let $L[t]$ be the Laplacian matrix of $W[t]$; that is, $L[t] = D[t] - W[t]$ where the diagonal matrix $D[t] = \mathrm{diag}(d_1[t], \dots, d_k[t])$ with $d_i[t] = \sum_{j=1}^k w_{ij}[t]$. Consequently, we have $L[t]\mathbf{1} = 0$ for all $t \in \mathbb{N}$ where $\mathbf{1} = (1, 1, \dots, 1)$. Then system (2.1) can be rewritten as

$$\begin{aligned}&\boldsymbol{x}[t+1] = \boldsymbol{x}[t] + h\boldsymbol{v}[t],\\ &\boldsymbol{v}[t+1] = (\mathsf{Id} - hL[t])\boldsymbol{v}[t],\end{aligned} \tag{2.3}$$

where $\boldsymbol{x}[t] = (\boldsymbol{x}_1[t], \dots, \boldsymbol{x}_k[t])$ and $\boldsymbol{v}[t] = (\boldsymbol{v}_1[t], \dots, \boldsymbol{v}_k[t])$.

Assumption 2.1. There exists an increasing time sequence $\{t_\ell^*\}_{\ell\in\mathbb{N}}$ with $t_0^* = 0$ and $t_{\ell+1}^* - t_\ell^* \le T$ for some positive constant T such that the union graph $\mathcal{G}([t_\ell^*, t_{\ell+1}^*))$ is rooted for all $\ell \in \mathbb{N}$.

Remark 2.2. Roughly speaking, Assumption 2.1 imposes some connectivity constraints for the switching interaction topologies during the system evolution. This assumption is supported by an empirical study in [4]. In that study, it is found that pigeon flocks employ an intermittent interaction mechanism where intragroup information transmission is not required at every time instant, but the union of the topologies of several consecutive interaction networks always keeps connected. Another important case with switching topologies is due to the well-known topological interactions reported in [1] where it is found that each bird interacts on average with a fixed number of closest neighbors. In [7], the closest neighbors model is analyzed in terms of the initial states of the flock only. It is shown in [7] that the switching topologies satisfy Assumption 2.1 provided that either the number of closest neighbors is large enough or the initial configurations fulfill a given condition. The existence of an upper bound T required in Assumption 2.1 means that the union of switching topologies should be connected frequently enough, which is known to be necessary for first-order consensus results [17].

In what follows we fix a solution $(\boldsymbol{x}[t], \boldsymbol{v}[t])$ of system (2.3) and an initial state $(\boldsymbol{x}[0], \boldsymbol{v}[0])$. To facilitate stating our results, we introduce the following quantities related to the initial state and the switching signal:

$$\alpha = 2T(k-1),\ \delta = h^{\alpha}\left(\frac{K}{2}\right)^{\alpha-(k-1)},$$

$$\gamma = 2\beta(k-1),\ c_1 = \frac{h\alpha\Lambda[0]}{\delta K^{k-1}},\ c_2 = \Gamma[0]+1.$$

We can now state our main result for discrete time.

THEOREM 2.3. *Consider system* (2.3) *with* $h \le \frac{1}{kK}$. *Let Assumption* 2.1 *hold. Assume that one of the following three hypotheses holds:*

(i) $\beta < 1/(2(k-1))$,
(ii) $\beta = 1/(2(k-1))$ *and* $c_1 < 1$,
(iii) $\beta > 1/(2(k-1))$ *and*

$$\left(\frac{1}{c_1}\right)^{\frac{1}{\gamma-1}}\left(\left(\frac{1}{\gamma}\right)^{\frac{1}{\gamma-1}} - \left(\frac{1}{\gamma}\right)^{\frac{\gamma}{\gamma-1}}\right) > c_2 + h\Lambda[0].$$

Then the agents converge to flocking exponentially fast.

Remark 2.4. The critical exponent here is $1/(2(k-1))$, independent of the bound T for the switching connectivity. For the case with the complete graph or hierarchical leadership, it is shown that the critical exponent is $1/2$, which is also optimal for flocking behavior [6, 8, 25]. For fixed directed topology, it is obtained in [11] that the critical exponent, however, relies on the connectivity of the interaction digraph. Here, it is depending on k, the number of agents. A natural question arising here is whether the critical exponent can be improved to be $1/2$ in the case of fixed (or even switching) directed topologies. So far, even for the fixed interaction digraph, we cannot solve this question.

2.3. Continuous time. We also consider evolution with continuous time. We model this evolution with the system, for $1 \le i \le k$,

$$\dot{\boldsymbol{x}}_i(t) = \boldsymbol{v}_i(t),$$
$$\dot{\boldsymbol{v}}_i(t) = \sum_{j\in\mathcal{N}_i^{\sigma(t)}} a_{ij}(\boldsymbol{x}(t))(\boldsymbol{v}_j(t) - \boldsymbol{v}_i(t)).$$

Here, the weight function $a_{ij}(\boldsymbol{x})$ is given by (2.2). Similar to (2.3), the continuous-time system can be rewritten in the compact form:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{v}(t),\quad \dot{\boldsymbol{v}}(t) = -L(t)\boldsymbol{v}(t). \tag{2.4}$$

In addition to Assumption 2.1, the switching signal $\sigma(t) : [0,\infty) \to \mathcal{P}$ is assumed to have a nonvanishing dwell time in the following sense.

Assumption 2.5. The sequence of switching instants $\{t_\ell\}_{\ell\in\mathbb{N}}$ satisfies that $t_{\ell+1} - t_\ell \geq \tau_0$ for some positive constant τ_0 and for all $\ell \in \mathbb{N}$.

Remark 2.6. This last assumption implies that $\lim_{\ell\to\infty} t_\ell \to \infty$. Also, we note that the time sequence $\{t_\ell^*\}_{\ell\in\mathbb{N}}$ in Assumption 2.1 can be different from the switching sequence $\{t_\ell\}_{\ell\in\mathbb{N}}$.

Without loss of generality, we assume that $\tau_0 \leq T$. It is also convenient to introduce the following quantities:

$$\alpha = 2T(k-1), \qquad \eta = e^{-kK\alpha} \min_{2\leq q\leq \lfloor \frac{2T}{\tau_0}\rfloor} \tau_0^{q(k-1)} \left(\frac{K}{2}\right)^{(q-1)(k-1)},$$

$$\gamma = 2\beta(k-1), \qquad b_1 = \frac{\alpha\Lambda(0)}{\eta K^{k-1}}, \qquad b_2 = \Gamma(0)+1.$$

We can now state our main result for continuous time.

THEOREM 2.7. *Consider system* (2.4). *Let Assumptions* 2.1 *and* 2.5 *hold. Assume that one of the following three hypotheses holds:*

(i) $\beta < 1/(2(k-1))$,
(ii) $\beta = 1/(2(k-1))$ *and* $b_1 < 1$,
(iii) $\beta > 1/(2(k-1))$ *and*

$$\left(\frac{1}{b_1}\right)^{\frac{1}{\gamma-1}} \left(\left(\frac{1}{\gamma}\right)^{\frac{1}{\gamma-1}} - \left(\frac{1}{\gamma}\right)^{\frac{\gamma}{\gamma-1}}\right) > b_2.$$

Then the agents converge to flocking exponentially fast.

Remark 2.8. We point out that Theorems 2.3 and 2.7 still hold true if the weight function is relaxed to

$$a_{ij}(\boldsymbol{x}) \geq \frac{K}{(1+\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2)^\beta}. \tag{2.5}$$

We will see that the two choices (2.2) and (2.5) make no difference in the subsequent analysis since what we need is the lower bound of the weight function $a_{ij}(\boldsymbol{x})$ rather than its specific form. In particular, (2.5) includes the nonsymmetric influence scaled by their relative distance which was considered in [18].

3. Proof of the main results.

3.1. Some technical preliminaries. A nonnegative matrix $A = (a_{ij})_{k\times k}$ is said to be *stochastic* if each of its row-sum is equal to 1. It is said to be *scrambling* if for each pair of indexes i and j there exists an index ℓ such that $a_{i\ell} > 0$ and $a_{j\ell} > 0$. Define the *ergodicity coefficient* of A to be

$$\chi(A) := \min_{i,j} \sum_{\ell=1}^{k} \min\{a_{i\ell}, a_{j\ell}\}.$$

It is easy to see that A is scrambling if and only if $\chi(A) > 0$. The importance of the ergodicity coefficient is due to the following contraction property.

LEMMA 3.1. [7, Lemma 2.1] *Assume that* $A = (a_{ij})_{k\times k}$ *is stochastic. For all* $v = (v_1, \ldots, v_k) \in \mathbb{E}^k$ *we have*

$$\max_{i,j} \|w_i - w_j\| \leq (1-\chi(A)) \max_{i,j} \|v_i - v_j\|,$$

where $w = Av$ *and* $\|\cdot\|$ *denotes the* 2*-norm in* $\mathbb{E}$.

LEMMA 3.2. [13, Lemma 2] *Let $m \geq 2$ be a positive integer and let $A_1, A_2, \ldots, A_m$ be nonnegative $k \times k$ matrices. Suppose that the diagonal elements of all of the A_i are positive and let $\underline{\lambda}$ and $\overline{\lambda}$ be the smallest and largest of these, respectively. Then*

$$A_1 A_2 \ldots A_m \geq \left(\frac{\underline{\lambda}^2}{2\overline{\lambda}}\right)^{(m-1)} (A_1 + A_2 + \cdots + A_m).$$

PROPOSITION 3.3. [29, Theorem 5.1] *If A_i are nonnegative $k \times k$ matrices with positive diagonal elements such that $\mathcal{G}(A_i)$ is rooted for all i, then $A_1 A_2 \ldots A_{k-1}$ is scrambling.*

LEMMA 3.4. [8, Lemma 2] *Let $c_1, c_2 > 0$ and $s > q > 0$. Then the equation*

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

has a unique positive zero z_*. In addition,

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

and $F(z) \leq 0$ for $0 \leq z \leq z_*$.

3.2. Proofs for discrete time. In this subsection, we present an asymptotic flocking estimate for discrete system (2.3). The key point is to establish an estimate for the decay of $\Lambda[t]$. Our strategy for this is that by Assumption 2.1, Lemma 3.2, and Proposition 3.3, we first derive a lower bound on the ergodicity coefficient of the nonnegative matrix associated to the union of interaction digraphs across some time intervals, and then we use Lemma 3.1 to establish an exponential rate of decay for $\Lambda[t]$. The following lemma is useful in our proof.

LEMMA 3.5. [11, Lemma 2] *Assume that $h \leq \frac{1}{kK}$. Then, for all $t \geq 0$, we have*

$$\Gamma[t+1] \leq \Gamma[t] + h\Lambda[t],$$
$$\Lambda[t+1] \leq \Lambda[t].$$

PROPOSITION 3.6. *Assume that $h \leq \frac{1}{kK}$. Let Assumption 2.1 hold. Then, for all $t \in \mathbb{N}$, we have*

$$\Lambda[\alpha t] \leq \left(1 - \delta \cdot \phi[\alpha t]^{k-1}\right)^t \Lambda[0],$$

where

$$\phi[\alpha t] = \min_{0 \leq s < \alpha t} \frac{K}{(1 + \Gamma[s]^2)^\beta}.$$

Proof. It follows from (2.3) that

$$\boldsymbol{v}[\alpha t] = \Phi[\alpha(t-1), \alpha t]\boldsymbol{v}[\alpha(t-1)], \tag{3.1}$$

where

$$\begin{aligned}\Phi[\alpha(t-1), \alpha t] &= \prod_{\ell=0}^{k-2} (\mathsf{Id} - hL[\alpha t - 2T\ell - 1]) \cdots (\mathsf{Id} - hL[\alpha t - 2T(\ell+1)]) \\ &=: \prod_{\ell=0}^{k-2} B_\ell.\end{aligned}$$

We need to estimate the matrix B_ℓ. Notice that, by the definition of L,

$$\begin{aligned} B_\ell &= (\mathsf{Id} - hL[\alpha t - 2T\ell - 1]) \cdots (\mathsf{Id} - hL[\alpha t - 2T(\ell+1)]) \\ &= ((\mathsf{Id} - hD[\alpha t - 2T\ell - 1]) + hW[\alpha t - 2T\ell - 1]) \\ &\quad \cdots ((\mathsf{Id} - hD[\alpha t - 2T(\ell+1)]) + hW[\alpha t - 2T(\ell+1)]) . \end{aligned} \tag{3.2}$$

By our assumption on h, for any $s \in \mathbb{N}$,

$$hD[s] \leq hkK\mathsf{Id} \leq \mathsf{Id}.$$

Therefore, (3.2) leads to

$$B_\ell \geq h^{2T} W[\alpha t - 2T\ell - 1] \cdots W[\alpha t - 2T(\ell+1)].$$

Recall that all diagonal elements of $W[s]$ are equal to K. It thus follows from Lemma 3.2 that

$$B_\ell \geq h^{2T} \left(\frac{K}{2}\right)^{2T-1} (W[\alpha t - 2T\ell - 1] + \cdots + W[\alpha t - 2T(\ell+1)]) . \tag{3.3}$$

For all $\alpha(t-1) \leq s < \alpha t$ and each $j \in \mathcal{N}_i^{\sigma[s]}$, the definition of $\phi[\alpha t]$ implies that $w_{ij}[s] \geq \phi[\alpha t]$. This, together with (3.3), yields

$$B_\ell \geq h^{2T} \left(\frac{K}{2}\right)^{2T-1} \phi[\alpha t] C_\ell,$$

where the matrix $C_\ell = (c_\ell^{ij})_{k\times k}$ is defined by $c_\ell^{ij} = 1$ if $(j,i) \in \mathcal{G}([\alpha t - 2T(\ell+1), \alpha t - 2T\ell))$, and $c_\ell^{ij} = 0$ otherwise. Note that $c_\ell^{ii} = 1$ for all $i \leq k$. Thus,

$$\prod_{\ell=0}^{k-2} B_\ell \geq h^{2T(k-1)} \left(\frac{K}{2}\right)^{(2T-1)(k-1)} \phi[\alpha t]^{k-1} \prod_{\ell=0}^{k-2} C_\ell. \tag{3.4}$$

On the other hand, using Assumption 2.1, we see that the interval $[\alpha t - 2T(\ell+1), \alpha t - 2T\ell)$ contains at least one interval $[t^*_{\ell'}, t^*_{\ell'+1})$ for some $\ell' \in \mathbb{N}$ and thus $\mathcal{G}([\alpha t - 2T(\ell+1), \alpha t - 2T\ell)) = \mathcal{G}(C_\ell)$ is rooted. Note also that $c_\ell^{ii} = 1$ for all $i \leq k$. Then Proposition 3.3 can be applied to conclude that $\prod_{\ell=0}^{k-2} C_\ell$ is scrambling. Let $\prod_{\ell=0}^{k-2} C_\ell = (e_{ij})_{k\times k}$. Since c_ℓ^{ij} is either one or zero, we have that the nonzero element e_{ij} is such that

$$e_{ij} = \sum_{1 \leq i_1, \cdots, i_{k-2} \leq k} c_0^{ii_1} c_1^{i_1 i_2} \cdots c_{k-2}^{i_{k-2} j} \geq 1.$$

Therefore,

$$\chi\left(\prod_{\ell=0}^{k-2} C_\ell\right) = \min_{i,j} \sum_{\ell=1}^{k} \min\{e_{i\ell}, e_{j\ell}\} > 0$$

implies that

$$\chi\left(\prod_{\ell=0}^{k-2} C_\ell\right) \geq 1.$$

It thus follows that

$$\chi\left(\prod_{\ell=0}^{k-2} B_\ell\right) \geq \delta \cdot \phi[\alpha t]^{k-1}.$$

This, together with (3.1) and Lemma 3.1, shows that

$$\Lambda[\alpha t] \leq \left(1 - \delta \cdot \phi[\alpha t]^{k-1}\right) \Lambda[\alpha(t-1)].$$

The statement now follows from induction on t. ∎

Proof of Theorem 2.3. For each $t \in \mathbb{N}$, define

$$\Gamma[t^*] = \max_{0 \leq s \leq t} \Gamma[s],$$

and let $\tau \in \mathbb{N}$ be such that $\alpha\tau \leq t < \alpha(\tau+1)$. By Lemma 3.5, we have

$$\Gamma[t^*] \leq \Gamma[0] + h \sum_{s=0}^{t^*-1} \Lambda[s] \leq \Gamma[0] + h \sum_{j=0}^{\tau} \sum_{s=\alpha j}^{\alpha(j+1)-1} \Lambda[s].$$

Since $\Lambda[s]$ is nonincreasing by Lemma 3.5, we deduce from Proposition 3.6 that

$$\begin{aligned} \Gamma[t^*] &\leq \Gamma[0] + h\alpha \sum_{j=0}^{\tau} \left(1 - \delta \cdot \phi[t]^{k-1}\right)^j \Lambda[0] \\ &\leq \Gamma[0] + h\alpha\Lambda[0] \sum_{j=0}^{\infty} \left(1 - \delta \cdot \phi[t]^{k-1}\right)^j \qquad (3.5) \\ &= \Gamma[0] + h\alpha\Lambda[0] \frac{\left(1 + \Gamma[t^*]^2\right)^{\beta(k-1)}}{\delta K^{k-1}}. \end{aligned}$$

Let $z[t] = \left(1 + (\Gamma[t^*])^2\right)^{1/2}$. Then (3.5) implies that

$$F(z[t]) := z[t] - c_1 z[t]^\gamma - c_2 \leq 0, \qquad (3.6)$$

where

$$c_1 = \frac{h\alpha\Lambda[0]}{\delta K^{k-1}} \text{ and } c_2 = \Gamma[0] + 1.$$

We now have three cases to consider.

(i) Assume $\beta < 1/(2(k-1))$. Then $\gamma < 1$. By Lemma 3.4, $F(z[t]) \leq 0$ implies that $1 + (\Gamma[t^*])^2 \leq \rho^2$ with

$$\rho \leq \max\left\{ \left(\frac{2h\alpha\Lambda[0]}{\delta K^{k-1}}\right)^{\frac{1}{1-\gamma}}, 2(\Gamma[0]+1) \right\}.$$

Since ρ is independent of t, we have that, for all $t \geq 0$,

$$\Gamma[t] \leq \Gamma[t^*] \leq (\rho^2 - 1)^{\frac{1}{2}}.$$

Therefore, $\phi[t] \geq \frac{K}{\rho^{2\beta}}$. It thus follows from Proposition 3.6 that

$$\Lambda[\alpha t] \leq \left(1 - \frac{\delta K^{k-1}}{\rho^\gamma}\right)^t \Lambda[0].$$

Consequently, we conclude from Lemma 3.5 that, for all $t \in \mathbb{N}$,

$$\Lambda[t] \leq \left(1 - \frac{\delta K^{k-1}}{\rho^\gamma}\right)^{\left\lfloor \frac{t}{\alpha} \right\rfloor} \Lambda[0].$$

(ii) Assume now $\beta = 1/(2(k-1))$. Then $\gamma = 1$. It now follows from (3.6) and our hypothesis that

$$1 + (\Gamma[t^*])^2 \leq \rho^2 := \left(\frac{\Gamma[0] + 1}{1 - \frac{h\alpha\Lambda[0]}{\delta K^{k-1}}} \right)^2 .$$

We now proceed as in case (i).

(iii) Assume finally $\beta > 1/(2(k-1))$. Then $\gamma > 1$. Note that the derivative of $F(z)$ is such that $F'(z) = 1 - c_1\gamma z^{\gamma-1}$ has a unique zero at $z_* = (1/(c_1\gamma))^{1/(\gamma-1)}$. Furthermore, by our hypothesis,

$$F(z_*) = \left(\frac{1}{c_1}\right)^{\frac{1}{\gamma-1}} \left(\left(\frac{1}{\gamma}\right)^{\frac{1}{\gamma-1}} - \left(\frac{1}{\gamma}\right)^{\frac{\gamma}{\gamma-1}} \right) - c_2 > 0.$$

On the other hand, since $F(0) = -c_2 < 0$ and $F(z) \to -\infty$ as $z \to +\infty$, we see that the function $F(z)$ has two zeros, say z_1 and z_2, with $0 < z_1 < z_* < z_2$ and that $F(z) > 0$ for all $z_1 < z < z_2$. For $t = 0$, we have $t^* = 0$ as well, and

$$z[0] = \left(1 + (\Gamma[0])^2\right)^{1/2} \leq c_2 < z_*.$$

Actually, $F(z[0]) \leq 0$ also implies that $z[0] \leq z_1$. We need to show that $z[t] \leq z_1$ for all $t \in \mathbb{N}$. To this end, suppose, to the contrary, that there exists a time instant $\tau > 1$ such that $z[\tau] > z_1$ and $z[\tau-1] \leq z_1$. This also implies that $z[\tau] \geq z_2$ since $F(z[\tau]) \leq 0$. Firstly, $z[\tau-1] \leq z_1$ shows that

$$\Gamma[\tau-1] \leq (z_1^2 - 1)^{1/2},$$

and $z[\tau] \geq z_2$ gives that

$$\Gamma[\tau] \geq (z_2^2 - 1)^{1/2}.$$

This, in turn, implies that

$$\begin{aligned} \Gamma[\tau] - \Gamma[\tau-1] \ &\geq \ (z_2^2-1)^{1/2} - (z_1^2-1)^{1/2} \ = \ \frac{z_2^2 - z_1^2}{(z_2^2-1)^{1/2} + (z_1^2-1)^{1/2}} \\ &\geq \ \frac{z_2^2 - z_1^2}{z_2 + z_1} \ = \ z_2 - z_1 \ \geq \ z_* - z_1. \end{aligned}$$

By the mean value theorem, there is a $\xi \in (z_1, z_*)$ such that $F(z_*) = F(z_*) - F(z_1) = F'(\xi)(z_* - z_1)$. Due to $0 \leq F'(z) \leq 1$ for all $z_1 \leq z \leq z_*$, we obtain $z_* - z_1 \geq F(z_*)$. By Lemma 3.5, we have

$$\Gamma[\tau] - \Gamma[\tau-1] \leq h\Lambda[\tau-1] \leq h\Lambda[0].$$

Therefore, we conclude that

$$h\Lambda[0] \geq F(z_*),$$

which contradicts our hypothesis. Hence, $z[t] \leq z_1 < z_*$ for all $t \in \mathbb{N}$, equivalently,

$$1 + (\Gamma[t^*])^2 \leq \rho^2 := \left(\frac{1}{c_1\gamma}\right)^{\frac{2}{\gamma-1}} .$$

We now proceed as in case (i) to complete the proof.

3.3. Proofs for continuous time. To prove Theorem 2.7 we first show some stepping stones. To state them, we need the notion of the upper-right Dini derivative of a continuous function $f(t) : \mathbb{R} \to \mathbb{R}$. This is given by

$$\mathrm{D}^+ f(t) = \limsup_{s \to t^+} \frac{f(s) - f(t)}{s - t}.$$

A detailed description of this notion can be found, e.g., in Appendix I of [24].

LEMMA 3.7. *Consider system* (2.4). *For* $t \geq 0$, *we have* $\mathrm{D}^+\Gamma(t) \leq \Lambda(t)$ *and*

$$\Gamma(t) \leq \Gamma(0) + \int_0^t \Lambda(s) ds.$$

Proof. The first statement is Lemma 3 in [11]. For the second one, we note that, for any $t \geq 0$ and $i, j \leq k$,

$$\|(\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)) - (\boldsymbol{x}_i(0) - \boldsymbol{x}_j(0))\| = \left\| \int_0^t \boldsymbol{v}_i(s) - \boldsymbol{v}_j(s) ds \right\| \leq \int_0^t \|\boldsymbol{v}_i(s) - \boldsymbol{v}_j(s)\| ds.$$

It thus follows that

$$\|\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)\| \leq \|\boldsymbol{x}_i(0) - \boldsymbol{x}_j(0)\| + \int_0^t \|\boldsymbol{v}_i(s) - \boldsymbol{v}_j(s)\| ds \leq \Gamma(0) + \int_0^t \Lambda(s) ds.$$

The statement now follows by taking indices i, j such that $\Gamma(t) = \|\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)\|$. ∎

In the following proposition, we show an exponential rate of decay for $\Lambda(t)$, which is mostly parallel to Proposition 3.6.

PROPOSITION 3.8. *Let Assumptions* 2.1 *and* 2.5 *hold. Then, for all* $t \in \mathbb{N}$, *we have*

$$\Lambda(\alpha t) \leq \left(1 - \eta \cdot \phi(\alpha t)^{k-1}\right)^t \Lambda(0),$$

where

$$\phi(\alpha t) = \min_{0 \leq s \leq \alpha t} \frac{K}{(1 + \Gamma(s)^2)^\beta}.$$

Proof. Note first that the second equation of system (2.4) is equivalent to the equation

$$\dot{\boldsymbol{v}}(t) = (A_{\sigma(t)}(t) - D_{\sigma(t)}(t))\boldsymbol{v}(t)$$

whose solution, on the interval $[\alpha t - 2T, \alpha t]$, is given by

$$\boldsymbol{v}(\alpha t) = \Psi(\alpha t, \alpha t - 2T)\boldsymbol{v}(\alpha t - 2T),$$

where $\Psi(\alpha t, \alpha t - 2T)$ is the fundamental matrix (see [26, section C.4] for this notion) of $A_{\sigma(t)}(t) - D_{\sigma(t)}(t)$ on $[\alpha t - 2T, \alpha t]$. It can be seen that $\Psi(\alpha t, \alpha t - 2T)$ is a stochastic matrix. Let $\{t_{\ell_2}, t_{\ell_3}, \ldots, t_{\ell_q}\}$ be the subsequence of $\{t_\ell\}_{\ell \in \mathbb{N}}$ containing all the switching points of $\sigma(t)$ on the interval $(\alpha t - 2T, \alpha t)$. By adding $\alpha t - 2T$ and αt to the subsequence, we obtain a new subsequence $\{t_{\ell_1}, t_{\ell_2}, \ldots, t_{\ell_q}, t_{\ell_{q+1}}\}$ with $t_{\ell_1} = \alpha t - 2T$ and $t_{\ell_{q+1}} = \alpha t$. Observe that $q \leq \lfloor 2T/\tau_0 \rfloor$. Let $\sigma(t) = p_r$ for $t \in [t_{\ell_r}, t_{\ell_{r+1}})$ and $r = 1, \ldots, q$. Actually, $\Psi(\alpha t, \alpha t - 2T)$ can be rewritten as

$$\Psi(\alpha t, \alpha t - 2T) = \Psi_{p_q}(t_{\ell_{q+1}}, t_{\ell_q}) \cdots \Psi_{p_1}(t_{\ell_2}, t_{\ell_1}), \tag{3.7}$$

where $\Psi_{p_r}(t_{\ell_{r+1}}, t_{\ell_r})$ is the fundamental matrix of $A_{p_r}(t) - D_{p_r}(t)$ on $[t_{\ell_r}, t_{\ell_{r+1}}]$ for $r = 1, \ldots, q$. We now estimate the fundamental matrix $\Psi_{p_r}(t_{\ell_{r+1}}, t_{\ell_r})$. We first note that for all $t \in [t_{\ell_r}, t_{\ell_{r+1}})$, $A_{p_r}(t) - D_{p_r}(t) \geq M_{p_r} - kK\mathsf{Id}$ where the matrix $M_{p_r} = (m_{p_r}^{ij})_{k\times k}$ is such that $m_{p_r}^{ii} = K$ for all $i \leq k$, and for $i \neq j$, $m_{p_r}^{ij} = \phi(\alpha t)$ if $j \in \mathcal{N}_i^{p_r}$ and $m_{p_r}^{ij} = 0$ otherwise. That is, M_{p_r} is an adjacency matrix of $\mathcal{G}_{p_r} = (\mathcal{V}, \mathcal{E}_{p_r})$. Let $\Phi_{p_r}(t_{\ell_{r+1}}, t_{\ell_r})$ be the fundamental matrix associated to the matrix $A_{p_r}(t) - D_{p_r}(t) + kK\mathsf{Id}$ on the interval $[t_{\ell_r}, t_{\ell_{r+1}}]$. From [7, Lemma A.1], it follows that

$$\Psi_{p_r}(t_{\ell_{r+1}}, t_{\ell_r}) = \Phi_{p_r}(t_{\ell_{r+1}}, t_{\ell_r}) e^{-kK(t_{\ell_{r+1}} - t_{\ell_r})}. \tag{3.8}$$

Also, using the Peano–Baker formula (see, e.g., [26, (C.27)]), we have

$$\begin{aligned}
&\Phi_{p_r}(t_{\ell_{r+1}}, t_{\ell_r}) \\
&\quad = \mathsf{Id} + \int_{t_{\ell_r}}^{t_{\ell_{r+1}}} (A_{p_r}(s_1) - D_{p_r}(s_1) + kK\mathsf{Id}) ds_1 + \cdots \\
&\qquad + \int_{t_{\ell_r}}^{t_{\ell_{r+1}}} \int_{t_{\ell_r}}^{s_1} \cdots \int_{t_{k_r}}^{s_{i-1}} \prod_{j=1}^{i} (A_{p_r}(s_j) - D_{p_r}(s_j) + kK\mathsf{Id}) ds_i \cdots ds_2 ds_1 + \cdots \\
&\quad \geq \mathsf{Id} + \sum_{i=1}^{\infty} \frac{((t_{\ell_{r+1}} - t_{\ell_r}) M_{p_r})^i}{i!} \\
&\quad \geq \tau_0 M_{p_r},
\end{aligned}$$

where the last inequality is from Assumption 2.5. By (3.8), we now conclude that

$$\Psi_{p_r}(t_{\ell_{r+1}}, t_{\ell_r}) \geq e^{-kK(t_{\ell_{r+1}} - t_{\ell_r})} \tau_0 M_{p_r},$$

which, together with (3.7) and Lemma 3.2, shows that

$$\begin{aligned}
\Psi(\alpha t, \alpha t - 2T) &\geq e^{-kK(t_{k_{q+1}} - t_{k_q})} \tau_0 M_{p_q} \cdots e^{-kK(t_{k_2} - t_{k_1})} \tau_0 M_{p_1} \\
&= e^{-2kKT} \tau_0^q M_{p_q} \cdots M_{p_1} \\
&\geq e^{-2kKT} \tau_0^q \left(\frac{K}{2}\right)^{q-1} (M_{p_q} + \cdots + M_{p_1}) \\
&\geq e^{-2kKT} \tau_0^q \left(\frac{K}{2}\right)^{q-1} \phi(\alpha t) C_1,
\end{aligned}$$

where C_1 is the $(0,1)$-adjacency matrix of the union graph $\mathcal{G}([\alpha t - 2T, \alpha t))$. A similar argument shows that, for $1 \leq i \leq k-1$,

$$\Psi(\alpha t - 2T(i-1), \alpha t - 2Ti) \geq e^{-2kKT} \tau_0^q \left(\frac{K}{2}\right)^{q-1} \phi(\alpha t) C_i, \tag{3.9}$$

where C_i is the $(0,1)$-adjacency matrix of the union graph $\mathcal{G}([\alpha t - 2Ti, \alpha t - 2T(i-1)))$. It follows from (3.9) that

$$\begin{aligned}
&\prod_{i=1}^{k-1} \Psi(\alpha t - 2T(i-1), \alpha t - 2Ti) \\
&\qquad \geq e^{-2k(k-1)KT} \tau_0^{q(k-1)} \left(\frac{K}{2}\right)^{(q-1)(k-1)} \phi(\alpha t)^{k-1} \prod_{i=1}^{k-1} C_i.
\end{aligned}$$

From Assumption 2.1, we see that each interval $[\alpha t-2Ti, \alpha t-2T(i-1))$ contains at least one interval $[t_\ell^*, t_{\ell+1}^*)$ for some $\ell\in\mathbb{N}$. This in turn shows that the union graph $\mathcal{G}([\alpha t-2Ti, \alpha t-2T(i-1)))$ is rooted. Applying Proposition 3.3 gives that $\prod_{i=1}^{k-1} C_i$ is scrambling. Furthermore, by the same argument as in the proof of Proposition 3.6, we have

$$\chi\left(\prod_{i=1}^{k-1} C_i\right)\geq 1,$$

which implies that

$$\chi\left(\Psi(\alpha t, \alpha(t-1)\right)=\chi\left(\prod_{i=1}^{k-1}\Psi(\alpha t-2T(i-1), \alpha t-2Ti)\right)\geq \eta\cdot\phi(\alpha t)^{k-1}.$$

Since

$$\boldsymbol{v}(\alpha t)=\Psi(\alpha t, \alpha(t-1))\boldsymbol{v}((\alpha(t-1)),$$

an application of Lemma 3.1 yields

$$\Gamma(\alpha t)\leq(1-\eta\cdot\phi(\alpha t)^{k-1})\Gamma(\alpha(t-1)).$$

We now complete the proof by induction on t. ∎

Proof of Theorem 2.7. For each $t\geq 0$, define

$$\Gamma(t^*)=\max_{0\leq s\leq t}\Gamma(s),$$

and let $\tau\in\mathbb{N}$ be such that $\alpha\tau\leq t<\alpha(\tau+1)$. By Lemma 3.7 and Proposition 3.8, we have

$$\begin{aligned}\Gamma(t^*)&\leq\Gamma(0)+\int_0^{t^*}\Lambda(s)ds\\&\leq\Gamma(0)+\sum_{i=0}^{\tau}\int_{\alpha i}^{\alpha(i+1)}\Lambda(s)ds\\&\leq\Gamma(0)+\sum_{i=0}^{\tau}\alpha\left(1-\eta\phi(t)^{k-1}\right)^i\Lambda(0)\\&\leq\Gamma(0)+\frac{\alpha\Lambda(0)}{\eta\phi(t)^{k-1}}\\&=\Gamma(0)+\alpha\Lambda(0)\frac{\left(1+(\Gamma(t^*))^2\right)^{\beta(k-1)}}{\eta K^{k-1}}.\end{aligned}$$

Let $z(t)=1+\Gamma(t^*)$. Then the above inequality implies that

$$F(z(t)):=z(t)-b_1z(t)^\gamma-b_2\leq 0, \tag{3.10}$$

where

$$b_1=\frac{\alpha\Lambda(0)}{\eta K^{k-1}}\text{ and }b_2=\Gamma(0)+1.$$

We now have three cases to consider.

(i) Assume $\beta < 1/(2(k-1))$. Then $\gamma < 1$. By Lemma 3.4, $F(z(t)) \le 0$ implies that $1+\Gamma(t^*) \le \rho$ with

$$\rho \le \max\left\{ \left(\frac{2\alpha\Lambda(0)}{\eta K^{k-1}}\right)^{\frac{1}{1-\gamma}}, 2(\Gamma(0)+1) \right\}.$$

Since ρ is independent of t, we have that, for all $t \ge 0$,

$$\Gamma(t) \le \Gamma(t^*) \le \rho - 1.$$

Therefore, $\phi(t) \ge \frac{K}{\rho^{2\beta}}$. It thus follows from Proposition 3.8 that, for all $t \in \mathbb{N}$,

$$\Lambda(\alpha t) \le \left(1 - \frac{\eta K^{k-1}}{\rho^\gamma}\right)^t \Lambda(0).$$

Consequently, we conclude Lemma 3.7 that, for all $t \ge 0$,

$$\Lambda(t) \le \left(1 - \frac{\eta K^{k-1}}{\rho^\gamma}\right)^{\lfloor \frac{t}{\alpha} \rfloor} \Lambda(0).$$

(ii) Assume now $\beta = 1/(2(k-1))$. Then $\gamma = 1$. It now follows from (3.6) and our hypothesis that

$$1 + \Gamma(t^*) \le \rho := \frac{\Gamma(0)+1}{1 - \frac{\alpha\Lambda(0)}{\eta K^{k-1}}}.$$

We now proceed as in case (i).

(iii) Assume finally $\beta > 1/(2(k-1))$. Then $\gamma > 1$. As in the proof of Theorem 2.3, we see that $F'(z)$, the derivative of $F(z)$, has a unique zero at $z_* = (1/(b_1\gamma))^{1/(\gamma-1)}$ and $F(z_*) > 0$ by our hypothesis. For $t = 0$, we have $t^* = 0$. This gives that $z(0) = b_2 < z_*$. Note also that $F(z(t))$ is continuous with respect to t and that $F(z(t)) \le 0$ for all $t \ge 0$. We can conclude that

$$1 + \Gamma(t^*) = z(t) < z_* = \left(\frac{1}{b_1\gamma}\right)^{\frac{1}{\gamma-1}} = \rho := \left(\frac{\eta K^{k-1}}{\alpha\gamma\Lambda(0)}\right)^{\frac{1}{\gamma-1}}.$$

We now proceed as in case (i) to complete the proof.

REFERENCES

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

[2] A. Blanchet and P. Degond, *Kinetic models for topological nearest-neighbor interactions*, J. Statist. Phys., 169 (2017), pp. 929–950.

[3] M. Cao, A. S. Morse, and B. D. O. Anderson, *Reaching a consensus in a dynamically changing environment: A graphical approach*, SIAM J. Control Optim., 47 (2008), pp. 575–600.

[4] D. Chen, X. Liu, B. Xu, and H.-T. Zhang, *Intermittence and connectivity of interactions in pigeon flock flights*, Sci. Rep., 7 (2017), 10452.

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

[6] F. Cucker and J.-G. Dong, *On the critical exponent for flocks under hierarchical leadership*, Math. Models Methods Appl. Sci., 19 (2009), pp. 1391–1404.

[7] F. Cucker and J.-G. Dong, *On flocks influenced by closest neighbors*. Math. Models Methods Appl. Sci., 26 (2016), pp. 2685–2708.

[8] F. Cucker and S. Smale, *Emergent behavior in flocks*, IEEE Trans. Automat. Control, 52 (2007), pp. 852–862.

[9] F. Cucker and S. Smale, *On the mathematics of emergence*, Japan. J. Math., 2 (2007), pp. 197–227.

[10] F. Dalmao and E. Mordecki, *Hierarchical Cucker-Smale model subject to random failure*, IEEE Trans. Automat. Control, 57 (2012), pp. 1124–1129.

[11] J.-G. Dong and L. Qiu, *Flocking of the Cucker-Smale model on general digraphs*, IEEE Trans. Automat. Control, 62 (2017), pp. 5234–5239.

[12] D. Floreano and R. J. Wood, *Science, technology and the future of small autonomous drones*, Nature, 521 (2015), pp. 460–466.

[13] A. Jadbabaie, J. Lin, and A. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Trans. Automat. Control, 48 (2003), pp. 988–1001.

[14] Z. Li, S.-Y. Ha, and X. Xue, *Emergent phenomena in an ensemble of Cucker-Smale particles under joint rooted leadership*, Math. Models Methods Appl. Sci., 24 (2014), pp. 1389–1419.

[15] Z. Li and X. Xue, *Cucker-Smale flocking under rooted leadership with fixed and switching topologies*, SIAM J. Appl. Math., 70 (2010), pp. 3156–3174.

[16] L. Moreau, *Stability of continuous-time distributed consensus algorithms*, in Proceedings of the 43rd IEEE Conference on Decision and Control, 2004, pp. 3998–4003.

[17] L. Moreau, *Stability of multiagent systems with time-dependent communication links*, IEEE Trans. Automat. Control, 50 (2005), pp. 169–182.

[18] S. Motsch and E. Tadmor, *A new model for self-organized dynamics and its flocking behavior*, J. Statist. Phys., 144 (2011), pp. 923–947.

[19] M. Nagy, Z. Ákos, D. Biro, and T. Vicsek, *Hierarchical group dynamics in pigeon flocks*, Nature, 464 (2010), pp. 890–893.

[20] L. Perea, P. Elosegui, and G. Gómez, *Extension of the Cucker-Smale control law to space flight formations*, J. Guidance Contr. Dyn., 32 (2009), pp. 526–536.

[21] B. Pettit, Z. Ákos, T. Vicsek, and D. Biro, *Speed determines leadership and leadership determines learning during pigeon flocking*, Current Biology, 25 (2015), pp. 3132–3137.

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

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

[24] N. Rouche, P. Habets, and M. Laloy, *Stability Theory by Liapunov's Direct Method*, Springer-Verlag, New York, 1977.

[25] J. Shen, *Cucker-Smale flocking under hierarchical leadership*, SIAM J. Appl. Math., 68 (2007), pp. 694–719.

[26] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, 2nd ed., Springer-Verlag, New York, 1998.

[27] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, *Novel type of phase transition in a system of self-driven particles*, Phys. Rev. Lett., 75 (1995), pp. 1226–1229.

[28] T. Vicsek and A. Zefeiris, *Collective motion*, Phys. Rep., 517 (2012), pp. 71–140.

[29] C. W. Wu, *Synchronization and convergence of linear dynamics in random directed networks*, IEEE Trans. Automat. Control, 51 (2006), pp. 1207–1210.

The corresponding author of this article is Jiu-Gang Dong.