SIAM J. CONTROL OPTIM.
Vol. 47, No. 2, pp. 575–600

REACHING A CONSENSUS IN A DYNAMICALLY CHANGING ENVIRONMENT: A GRAPHICAL APPROACH*

MING CAO†, A. STEPHEN MORSE†, AND BRIAN D. O. ANDERSON‡

Abstract. This paper presents new graph-theoretic results appropriate for the analysis of a variety of consensus problems cast in dynamically changing environments. The concepts of rooted, strongly rooted, and neighbor-shared are defined, and conditions are derived for compositions of sequences of directed graphs to be of these types. The graph of a stochastic matrix is defined, and it is shown that under certain conditions the graph of a Sarymsakov matrix and a rooted graph are one and the same. As an illustration of the use of the concepts developed in this paper, graph-theoretic conditions are obtained which address the convergence question for the leaderless version of the widely studied Vicsek consensus problem.

Key words. cooperative control, graph theory, switched systems, multiagent systems

AMS subject classifications. 93C05, 05C50, 05C75, 15A51, 40A20, 68W15

DOI. 10.1137/060657005

1. Introduction.

Current interest in cooperative control of groups of mobile autonomous agents has led to the rapid increase in the application of graph-theoretic ideas to problems of analyzing and synthesizing a variety of desired group behaviors such as maintaining a formation, swarming, rendezvousing, or reaching a consensus. While this in-depth assault on group coordination using a combination of graph theory and system theory is in its early stages, it is likely to significantly expand in the years to come. One line of research which illustrates the combined use of these concepts is the recent theoretical work by a number of individuals [17, 19, 22, 1, 3, 26] which successfully explains the heading synchronization phenomenon observed in simulation by Vicsek et al. [29], Reynolds [23], and others more than a decade ago. Vicsek and coauthors consider a simple discrete-time model consisting of n autonomous agents or particles all moving in the plane with the same speed but with different headings. Each agent's heading is updated using a local rule based on the average of its own heading plus the current headings of its "neighbors." Agent i's *neighbors* at time t are those agents which are either in or on a circle of prespecified radius centered at agent i's current position. In their paper, Vicsek et al. provide a variety of interesting simulation results which demonstrate that the nearest neighbor rule they are studying can cause all agents to eventually move in the same direction despite the absence of centralized coordination and despite the fact that each agent's set of nearest neighbors can change with time. A theoretical explanation for this observed behavior

*Received by the editors April 11, 2006; accepted for publication (in revised form) August 16, 2007; published electronically February 6, 2008. A preliminary version of this work can be found in A. S. Morse, *Logically switched dynamical systems*, in Nonlinear and Optimal Control Theory, Springer-Verlag, Berlin, 2008, pp. 1–84.

http://www.siam.org/journals/sicon/47-2/65700.html

†Electrical Engineering, Yale University, P.O. Box 208267, New Haven, CT 06520 (m.cao@yale.edu, morse@sysc.eng.yale.edu). The research of these authors was supported by the U.S. Army Research Office, the U.S. National Science Foundation, and a gift from the Xerox Corporation.

‡Australian National University and National ICT Australia Ltd., Locked bag 8001, Canberra ACT 2601, Australia (brian.anderson@nicta.com.au). The research of this author was supported by National ICT Australia, which is funded by the Australian Government's Department of Communications, Information Technology, and the Arts, and the Australian Research Council through the Backing Australia's Ability initiative and the ICT Centre of Excellence Program.

has recently been given in [17]. The explanation exploits ideas from graph theory [13] and from the theory of nonhomogeneous Markov chains [25, 30, 15]. Experience has shown that it is more the graph theory than the Markov chains which is key to this line of research. An illustration of this is the recent extension of the findings of [17] which explain the behavior of Reynolds' full nonlinear "boid" system [26].

Mathematically Vicsek's problem is what in statistics and computer science is called a "consensus problem" [10] or an "agreement problem" [21], although in computer science the issues tend to be concerned more with fault tolerance [12] rather than convergence. Roughly speaking, one has a group of agents which are all trying to agree on a specific value of some quantity. Each agent initially has only limited information available. The agents then try to reach a consensus by communicating what they know to their neighbors either just once or repeatedly, depending on the specific problem of interest. For the Vicsek problem, each agent knows only its own heading and the headings of its current neighbors. One feature of the Vicsek problem which sharply distinguishes it from other consensus problems is that each agent's neighbors can change with time, because all agents are in motion. The theoretical consequence of this is profound: it renders essentially useless, without elaboration, a large body of literature appropriate to the convergence analysis of "nearest neighbor" algorithms with fixed neighbor relationships. Said differently, for the linear heading update rules considered in this paper, understanding the difference between fixed neighbor relationships and changing neighbor relationships is much the same as understanding the difference between the stability of time-invariant linear systems and time-varying linear systems. Various mathematically similar versions of Vicsek's problem have been addressed in the literature [17, 19, 22, 1, 3]; some it turns out well before Vicsek's own paper was published [10, 9, 27, 28, 2].

The central aim of this paper is to establish a number of basic properties of "compositions" of sequences of directed graphs which, as shown in [7], are useful in explaining how a consensus is achieved in various settings. To motivate the graph-theoretic questions addressed and to demonstrate the utility of the answers obtained, we reconsider the version of the Vicsek consensus problem studied by Moreau [19] and Ren and Beard [22]. We derive a condition for agents to reach a consensus exponentially fast which is slightly different than but equivalent to the condition established in [19]. What this paper contributes, then, is a different approach to the understanding of the consensus phenomenon, one in which graphs and their compositions are at center stage. Of course if the consensus problem studied in [19, 22] were the only problem to which this approach were applicable, its development would have hardly been worth the effort. In a sequel to this paper [7] and elsewhere [4, 8, 6, 5] it is demonstrated that in fact the graph-theoretic approach we are advocating is applicable to a broad range of consensus problems which have so far either been only partially resolved or not studied at all.

To the best of our knowledge, all of the statements in this paper about graph compositions are original. However, because the literature on nonhomogeneous Markov chains is vast, some of these statements can undoubtedly be shown to be equivalent to statements about stochastic matrix product in the existing literature [25, 15]. The main convergence result on leaderless flocking, namely Theorem 3, is equivalent to one of the main results of [19]. Corollary 1 is in essence the main result of [17].

In section 2 we reconsider the leaderless coordination problem studied in [17] but without the assumption that the agents all have the same sensing radii. Agents are labelled 1 to n and are represented by correspondingly labelled vertices in a directed graph $\mathbb{N}$ whose arcs represent current neighbor relationships. We define the concept of a "strongly rooted graph" and show by an elementary argument that convergence to a common heading is achieved if the neighbor graphs encountered along a system trajectory are all strongly rooted. We also derive a worst case convergence rate for these types of trajectories. We next define the concept of a "rooted graph" and the operation of "graph composition." The directed graphs appropriate to the Vicsek model have self-arcs at all vertices. We prove that any composition of $(n-1)^2$ such rooted graphs is strongly rooted. Armed with this fact, we establish conditions under which consensus is achieved which are different than but equivalent to those obtained in [19, 22]. We then turn to a more in-depth study of rooted graphs. We prove that a so-called neighbor-shared graph is a special type of rooted graph and in so doing make a connection between the consensus problem under consideration and the elegant theory of "scrambling matrices" found in the literature on nonhomogeneous Markov chains [25, 15]. By exploiting this connection in [7], we are able to derive worst case convergence rate results for several versions of the Vicsek problem. The nonhomogeneous Markov chain literature also contains interesting convergence results for a class of stochastic matrices studied by Sarymsakov [24]. The class of Sarymsakov matrices is bigger than the class of all stochastic scrambling matrices. We make contact with this literature by proving that the graph of any Sarymsakov matrix is rooted and also that any stochastic matrix with a rooted graph whose vertices all have self-arcs is a Sarymsakov matrix.

2. Leaderless coordination. The system to be studied consists of n autonomous agents, labelled 1 through n, all moving in the plane with the same speed but with different headings. Each agent's heading is updated using a simple local rule based on the average of its own heading plus the headings of its "neighbors." Agent i's *neighbors* at time t are those agents, including itself, which are in a closed disk of prespecified radius r_i centered at agent i's current position. In what follows $\mathcal{N}_i(t)$ denotes the set of labels of those agents which are neighbors of agent i at time t. Agent i's heading, written θ_i, evolves in discrete time in accordance with a model of the form

$$\theta_i(t+1) = \frac{1}{n_i(t)}\left(\sum_{j\in\mathcal{N}_i(t)} \theta_j(t)\right), \tag{1}$$

where t is a discrete-time index taking values in the nonnegative integers $\{0, 1, 2, \ldots\}$, and $n_i(t)$ is the number of neighbors of agent i at time t.

2.1. Neighbor graph. The explicit form of the update equations determined by (1) depends on the relationships between neighbors which exist at time t. These relationships can be conveniently described by a directed graph $\mathbb{N}(t)$ with vertex set $\mathcal{V} = \{1, 2, \ldots, n\}$ and "arc set" $\mathcal{A}(\mathbb{N}(t)) \subset \mathcal{V} \times \mathcal{V}$ which is defined so that (i, j) is an *arc* or directed edge from i to j just in case agent i is a neighbor of agent j. Thus $\mathbb{N}(t)$ is a directed graph on n vertices with at most one arc connecting each ordered pair of distinct vertices and with exactly one self-arc at each vertex. We write $\mathcal{G}_{sa}$ for the set of all such graphs and $\mathcal{G}$ for the set of all directed graphs with vertex set $\mathcal{V}$. It is natural to call a vertex i a *neighbor* of vertex j in $\mathbb{G} \in \mathcal{G}$ if (i, j) is an arc in $\mathbb{G}$. In addition we sometimes refer to a vertex k as an *observer* of vertex j in $\mathbb{G}$ if (j, k) is an arc in $\mathbb{G}$. Thus every vertex of $\mathbb{G}$ can *observe* its neighbors, which with the interpretation of vertices as agents is precisely the kind of relationship $\mathbb{G}$ is supposed to represent.

2.2. State equation. The set of agent heading update rules defined by (1) can be written in state form. Towards this end, for each graph $\mathbb{N} \in \mathcal{G}_{sa}$, define the *flocking matrix*

$$F = D^{-1}A', \tag{2}$$

where A' is the transpose of the "adjacency matrix" of $\mathbb{N}$ and D the diagonal matrix whose jth diagonal element is the "in-degree" of vertex j within the graph.[1] The function $\mathbb{N} \longmapsto F$ is bijective. Then

$$\theta(t+1) = F(t)\theta(t), \quad t \in \{0, 1, 2, \ldots\}, \tag{3}$$

where θ is the heading vector $\theta = [\,\theta_1 \quad \theta_2 \quad \ldots \quad \theta_n\,]'$ and $F(t)$ is the flocking matrix of the *neighbor graph* $\mathbb{N}(t)$ which represents the neighbor relationships of (1) at time t. A complete description of this system would have to include a model which explains how $\mathbb{N}(t)$ changes over time as a function of the positions of the n agents in the plane. While such a model is easy to derive and is essential for simulation purposes, it would be difficult to take into account in a convergence analysis. To avoid this difficulty, we shall adopt a more conservative approach which ignores how $\mathbb{N}(t)$ depends on the agent positions in the plane and assumes instead that $t \longmapsto \mathbb{N}(t)$ might be any signal in some suitably defined set of interest.

Our ultimate goal is to show for a large class of signals $t \longmapsto \mathbb{N}(t)$ and for any initial set of agent headings that the headings of all n agents will converge to the same steady state value θ_{ss}. Convergence of the θ_i to θ_{ss} is equivalent to the state vector θ converging to a vector of the form $\theta_{ss}\mathbf{1}$, where $\mathbf{1} \stackrel{\Delta}{=} [\,1 \quad 1 \quad \ldots \quad 1\,]'_{n\times 1}$. Naturally there are situations where convergence to a common heading cannot occur. The most obvious of these is when one agent—say the ith—starts so far away from the rest that it never acquires any neighbors. Mathematically this would mean not only that $\mathbb{N}(t)$ is never strongly connected[2] at any time t but also that vertex i remains an isolated vertex of $\mathbb{N}(t)$ for all t in the sense that within each $\mathbb{N}(t)$, vertex i has no incoming arcs other than its own self-arc. This situation is likely to be encountered if the r_i are very small. At the other extreme, which is likely if the r_i are very large, all agents might remain neighbors of all others for all time. In this case, $\mathbb{N}(t)$ would remain fixed along such a trajectory as the complete graph. Convergence of θ to $\theta_{ss}\mathbf{1}$ can easily be established in this special case because with $\mathbb{N}(t)$ so fixed, (3) is a linear, time-invariant, discrete-time system. The situation of perhaps the greatest interest is between these two extremes when $\mathbb{N}(t)$ is not necessarily complete or even strongly connected for any $t \geq 0$ but when no strictly proper subset of $\mathbb{N}(t)$'s vertices is isolated from the rest for all time. Establishing convergence in this case is challenging because $F(t)$ changes with time and (3) is not time-invariant. It is this case which we intend to study.

2.3. Strongly rooted graphs. In what follows we will call a vertex i of a directed graph $\mathbb{G}$ a *root* of $\mathbb{G}$ if for each other vertex j of $\mathbb{G}$, there is a path from i to

[1] By the *adjacency matrix* of a directed graph $\mathbb{G} \in \mathcal{G}$ we mean an $n \times n$ matrix whose ijth entry is 1 if (i, j) is an arc in $\mathcal{A}(\mathbb{G})$ and 0 if it is not. The *in-degree* of vertex j in $\mathbb{G}$ is the number of arcs in $\mathcal{A}(\mathbb{G})$ of the form (i, j); thus j's in-degree is the number of *incoming* arcs to vertex j.

[2] A directed graph $\mathbb{G} \in \mathcal{G}$ with arc set $\mathcal{A}$ is *strongly connected* if it has a "path" between each distinct pair of its vertices i and j; by a *path* (of *length* m) between vertices i and j we mean a sequence of arcs in $\mathcal{A}$ of the form $(i, k_1), (k_1, k_2), \ldots, (k_{m-1}, k_m)$, where $k_m = j$ and, if $m > 1$, $i, k_1, \ldots, k_{m-1}$ are distinct vertices. $\mathbb{G}$ is *complete* if it has a path of length one (i.e., an arc) between each distinct pair of its vertices.

j. Thus i is a root of $\mathbb{G}$ if it is the root of a directed spanning tree of $\mathbb{G}$. We will say that $\mathbb{G}$ is *rooted at i* if i is in fact a root. Thus $\mathbb{G}$ is rooted at i just in case each other vertex of $\mathbb{G}$ is *reachable* from vertex i along a path within the graph. $\mathbb{G}$ is *strongly rooted at i* if each other vertex of $\mathbb{G}$ is reachable from vertex i along a path of length 1. Thus $\mathbb{G}$ is strongly rooted at i if i is a neighbor of every other vertex in the graph. A *rooted graph* $\mathbb{G}$ is a directed graph which possesses at least one root. Finally, a *strongly rooted graph* is a graph which has at least one vertex at which it is strongly rooted. It is now possible to state the following elementary convergence result which illustrates, under a restrictive assumption, the more general types of results to be derived later in the paper.

THEOREM 1. *Let $\theta(0)$ be fixed. For any trajectory of the system* (3) *along which each graph in the sequence of neighbor graphs* $\mathbb{N}(0)$, $\mathbb{N}(1), \ldots$ *is strongly rooted, there is a constant steady state heading θ_{ss} for which*

$$\lim_{t\to\infty} \theta(t) = \theta_{ss}\mathbf{1}, \tag{4}$$

where the limit is approached exponentially fast.

2.3.1. Stochastic matrices. In order to explain why Theorem 1 is true, we will make use of certain structural properties of the flocking matrices determined by the neighbor graphs in $\mathcal{G}_{sa}$. As defined, each flocking matrix F is square and nonnegative, where by a *nonnegative* matrix we mean a matrix whose entries are all nonnegative. Each F also has the property that its row sums all equal 1 (i.e., $F\mathbf{1} = \mathbf{1}$). Matrices with these two properties are called (row) *stochastic* [16]. It is easy to verify that the class of all $n \times n$ stochastic matrices is closed under multiplication. It is worth noting that because the vertices of the graphs in $\mathcal{G}_{sa}$ all have self-arcs, the F also have the property that their diagonal elements are positive. While the proof of Theorem 1 does not exploit this property, the more general results derived later in the paper depend crucially on it.

In what follows we write $M \geq N$ whenever $M - N$ is a nonnegative matrix. We also write $M > N$ whenever $M - N$ is a positive matrix where by a *positive matrix* we mean a matrix with all positive entries.

2.3.2. Products of stochastic matrices. Stochastic matrices have been extensively studied in the literature for a long time largely because of their connection with Markov chains [25, 30, 14]. One problem studied which is of particular relevance here is to describe the asymptotic behavior of products of $n \times n$ stochastic matrices of the form

$$S_j S_{j-1} \cdots S_1$$

as j tends to infinity. This is equivalent to looking at the asymptotic behavior of all solutions to the recursion equation

$$x(j+1) = S_j x(j) \tag{5}$$

since any solution $x(j)$ can be written as

$$x(j) = (S_j S_{j-1} \cdots S_1)\, x(1), \quad j \geq 1.$$

One especially useful idea, which goes back at least to [11] and has been extensively used [27], is to consider the behavior of the scalar-valued nonnegative function $V(x) =$

$\lceil x \rceil - \lfloor x \rfloor$ along solutions to (5), where $x = [\, x_1 \quad x_2 \quad \dots \quad x_n \,]'$ is a nonnegative n vector and $\lceil x \rceil$ and $\lfloor x \rfloor$ are its largest and smallest elements, respectively. The key observation is that for any $n \times n$ stochastic matrix S, the ith entry of Sx satisfies

$$\sum_{j=1}^{n} s_{ij} x_j \geq \sum_{j=1}^{n} s_{ij} \lfloor x \rfloor = \lfloor x \rfloor$$

and

$$\sum_{j=1}^{n} s_{ij} x_j \leq \sum_{j=1}^{n} s_{ij} \lceil x \rceil = \lceil x \rceil.$$

Since these inequalities hold for all rows of Sx, it must be true that $\lfloor Sx \rfloor \geq \lfloor x \rfloor$, that $\lceil Sx \rceil \leq \lceil x \rceil$, and, as a consequence, that $V(Sx) \leq V(x)$. These inequalities and (5) imply that the sequences

$$\lfloor x(1) \rfloor, \lfloor x(2) \rfloor, \dots, \qquad \lceil x(1) \rceil, \lceil x(2) \rceil, \dots, \qquad V(x(1)), V(x(2)), \dots$$

are each monotone. Thus because each of these sequences is also bounded, the limits

$$\lim_{j \to \infty} \lfloor x(j) \rfloor, \qquad \lim_{j \to \infty} \lceil x(j) \rceil, \qquad \lim_{j \to \infty} V(x(j))$$

each exist. Note that whenever the limit of $V(x(j))$ is zero, all components of $x(j)$ together with $\lfloor x(j) \rfloor$ and $\lceil x(j) \rceil$ must tend to the same constant value.

There are various different ways in which one might approach the problem of developing conditions under which $x(j)$ converges to some scalar multiple of $\mathbf{1}$ or equivalently $S_j S_{j-1} \cdots S_1$ converges to a constant matrix of the form $\mathbf{1}c$ for some constant row vector c. For example, since for any $n \times n$ stochastic matrix S, $S\mathbf{1} = \mathbf{1}$, it must be true that span $\{\mathbf{1}\}$ is an S-invariant subspace for any such S. From this and standard existence conditions for solutions to linear algebraic equation, it follows that for any $(n-1) \times n$ matrix P with kernel spanned by $\mathbf{1}$, the equation $PS = \tilde{S}P$ has unique solutions $\tilde{S}$, and, moreover, that

$$\text{spectrum } S = \{1\} \cup \text{spectrum } \tilde{S}. \tag{6}$$

As a consequence of the equation $PS_j = \tilde{S}_j P, \;\; j \geq 1$, it can easily be seen that

$$\tilde{S}_j \tilde{S}_{j-1} \cdots \tilde{S}_1 P = P S_j S_{j-1} \cdots S_1.$$

Since P has full row rank and $P\mathbf{1} = 0$, the convergence of a product $S_j S_{j-1} \cdots S_1$ to a matrix of the form $\mathbf{1}c$ is equivalent to convergence of the corresponding product $\tilde{S}_j \tilde{S}_{j-1} \cdots \tilde{S}_1$ to the zero matrix. There are two problems with this approach. First, since P is not unique, neither are the $\tilde{S}_i$. Second, it is not so clear how to go about picking P to make tractable the problem of proving that the resulting product $\tilde{S}_j \tilde{S}_{j-1} \cdots \tilde{S}_1$ tends to zero. Tractability of the latter problem generally boils down to choosing a norm for which the $\tilde{S}_i$ are all contractive. For example, one might seek to choose a suitably weighted 2-norm. This is in essence the same thing as choosing a common quadratic Lyapunov function. Although each $\tilde{S}_i$ can easily be shown to be discrete-time stable with all eigenvalues of magnitude less than 1, it is known that there are classes of S_i which give rise to $\tilde{S}_i$ for which no such common Lyapunov matrix exists [18] regardless of the choice of P. Of course there are many other possible norms to choose from other than 2-norms. In the end, success with this approach requires one to simultaneously choose *both* a suitable P and an appropriate norm with respect to which the $\tilde{S}_i$ are all contractive. In what follows we adopt a slightly different but closely related approach which ensures that we can work with what is perhaps the most natural norm for this type of convergence problem, the infinity norm.

To proceed, we need a few more ideas concerned with nonnegative matrices. For any nonnegative matrix R of any size, we write $||R||$ for the largest of the row sums of R. Note that $||R||$ is the induced infinity norm of R and consequently is submultiplicative. Note in addition that $||x|| = \lceil x \rceil$ for any nonnegative n vector x. Moreover, $||M_1|| \leq ||M_2||$ if $M_1 \leq M_2$. Observe that for any $n \times n$ stochastic matrix S, $||S|| = 1$ because the row sums of a stochastic matrix all equal 1. We extend the domain of definitions of $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ to the class of all nonnegative $n \times m$ matrix M by letting $\lfloor M \rfloor$ and $\lceil M \rceil$ now denote the $1 \times m$ row vectors whose jth entries are the smallest and largest elements, respectively, of the jth column of M. Note that $\lfloor M \rfloor$ is the largest $1 \times m$ nonnegative row vector c for which $M - \mathbf{1}c$ is nonnegative and that $\lceil M \rceil$ is the smallest nonnegative row vector c for which $\mathbf{1}c - M$ is nonnegative. Note in addition that for any $n \times n$ stochastic matrix S, one can write

$$S = \mathbf{1}\lfloor S \rfloor + \lfloor\!\lfloor S \rfloor\!\rfloor \qquad \text{and} \qquad S = \mathbf{1}\lceil S \rceil - \lceil\!\lceil S \rceil\!\rceil, \tag{7}$$

where $\lfloor\!\lfloor S \rfloor\!\rfloor$ and $\lceil\!\lceil S \rceil\!\rceil$ are nonnegative matrices defined by the equations

$$\lfloor\!\lfloor S \rfloor\!\rfloor = S - \mathbf{1}\lfloor S \rfloor \qquad \text{and} \qquad \lceil\!\lceil S \rceil\!\rceil = \mathbf{1}\lceil S \rceil - S, \tag{8}$$

respectively. Moreover, the row sums of $\lfloor\!\lfloor S \rfloor\!\rfloor$ are all equal to $1 - \lfloor S \rfloor \mathbf{1}$ and the row sums of $\lceil\!\lceil S \rceil\!\rceil$ are all equal to $\lceil S \rceil \mathbf{1} - 1$, and so

$$||\lfloor\!\lfloor S \rfloor\!\rfloor|| = 1 - \lfloor S \rfloor \mathbf{1} \qquad \text{and} \qquad ||\lceil\!\lceil S \rceil\!\rceil|| = \lceil S \rceil \mathbf{1} - 1. \tag{9}$$

In what follows we will also be interested in the matrix

$$[\![S]\!] = \lfloor\!\lfloor S \rfloor\!\rfloor + \lceil\!\lceil S \rceil\!\rceil. \tag{10}$$

This matrix satisfies

$$[\![S]\!] = \mathbf{1}(\lceil S \rceil - \lfloor S \rfloor) \tag{11}$$

because of (7).

For any infinite sequence of $n \times n$ stochastic matrices $S_1, S_2, \ldots$, we henceforth use the symbol $\lfloor \cdots S_j \cdots S_1 \rfloor$ to denote the limit

$$\lfloor \cdots S_j \cdots S_2 S_1 \rfloor = \lim_{j \to \infty} \lfloor S_j \cdots S_2 S_1 \rfloor. \tag{12}$$

From the preceding discussion it is clear that for $i \in \{1, 2, \ldots, n\}$, the limit $\lfloor \cdots S_j \cdots S_1 \rfloor e_i$ exists, where e_i is the ith unit n-vector. Thus the limit $\lfloor \cdots S_j \cdots S_1 \rfloor$ always exists, and this is true even if the product $S_j \cdots S_2 S_1$ itself does not have a limit. Two situations can occur. Either the product $S_j \cdots S_2 S_1$ converges to a rank one matrix or it does not. In fact, even if $S_j \cdots S_2 S_1$ does converge, it is quite possible that the limit is not a rank one matrix. An example of this would be a sequence in which S_1 is any stochastic matrix of rank greater than 1 and for all $i > 1$, $S_i = I_{n \times n}$. In what follows we will develop sufficient conditions for $S_j \cdots S_2 S_1$ to converge to a rank one matrix as $j \to \infty$. Note that if this occurs, then the limit must be of the form $\mathbf{1}c$, where $c\mathbf{1} = 1$ because stochastic matrices are closed under multiplication.

In what follows we will say that a matrix product $S_j S_{j-1} \cdots S_1$ *converges to* $\mathbf{1}\lfloor \cdots S_j \cdots S_1 \rfloor$ *exponentially fast at a rate no slower than* λ if there are nonnegative constants b and λ with $\lambda < 1$, such that

$$||(S_j \cdots S_1) - \mathbf{1}\lfloor \cdots S_j \cdots S_2 S_1 \rfloor|| \leq b\lambda^j, \qquad j \geq 1. \tag{13}$$

The following proposition implies that such a stochastic matrix product will so converge if $\llbracket S_j \cdots S_1 \rrbracket$ converges to 0.

PROPOSITION 1. *Let $\bar{b}$ and λ be nonnegative numbers with $\lambda < 1$. Suppose that $S_1, S_2, \ldots$ is an infinite sequence of $n \times n$ stochastic matrices for which*

$$||\llbracket S_j \cdots S_1 \rrbracket|| \leq \bar{b}\lambda^j, \quad j \geq 0. \tag{14}$$

Then the matrix product $S_j \cdots S_2 S_1$ converges to $\mathbf{1}\lfloor \cdots S_j \cdots S_1 \rfloor$ exponentially fast at a rate no slower than λ.

The proof of Proposition 1 makes use of the first of the two inequalities which follow.

LEMMA 1. *For any two $n \times n$ stochastic matrices S_1 and S_2,*

$$\lfloor S_2 S_1 \rfloor - \lfloor S_1 \rfloor \leq \lceil S_2 \rceil \llbracket S_1 \rrbracket, \tag{15}$$

$$\llbracket S_2 S_1 \rrbracket \leq \llbracket S_2 \rrbracket \, \llbracket S_1 \rrbracket. \tag{16}$$

Proof of Lemma 1. Since $S_2 S_1 = S_2(\mathbf{1}\lfloor S_1 \rfloor + \llbracket S_1 \rrbracket) = \mathbf{1}\lfloor S_1 \rfloor + S_2 \llbracket S_1 \rrbracket$ and $S_2 = \mathbf{1}\lceil S_2 \rceil - \llbracket S_2 \rrbracket$, it must be true that $S_2 S_1 = \mathbf{1}(\lfloor S_1 \rfloor + \lceil S_2 \rceil \llbracket S_1 \rrbracket) - \llbracket S_2 \rrbracket \, \llbracket S_1 \rrbracket$. Thus $\mathbf{1}(\lfloor S_1 \rfloor + \lceil S_2 \rceil \llbracket S_1 \rrbracket) - \llbracket S_2 \rrbracket \, \llbracket S_1 \rrbracket$ is nonnegative. But $\lceil S_2 S_1 \rceil$ is the smallest nonnegative row vector c for which $\mathbf{1}c - S_2 S_1$ is nonnegative. Therefore

$$\lceil S_2 S_1 \rceil \leq \lfloor S_1 \rfloor + \lceil S_2 \rceil \llbracket S_1 \rrbracket. \tag{17}$$

Moreover, $\lfloor S_2 S_1 \rfloor \leq \lceil S_2 S_1 \rceil$ because of the definitions of $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$. This and (17) imply that $\lfloor S_2 S_1 \rfloor \leq \lfloor S_1 \rfloor + \lceil S_2 \rceil \llbracket S_1 \rrbracket$ and thus that (15) is true.

Since $S_2 S_1 = S_2(\mathbf{1}\lfloor S_1 \rfloor + \llbracket S_1 \rrbracket) = \mathbf{1}\lfloor S_1 \rfloor + S_2 \llbracket S_1 \rrbracket$ and $S_2 = \lfloor S_2 \rfloor + \llbracket S_2 \rrbracket$, it must be true that $S_2 S_1 = \mathbf{1}(\lfloor S_1 \rfloor + \lfloor S_2 \rfloor \llbracket S_1 \rrbracket) + \llbracket S_2 \rrbracket \, \llbracket S_1 \rrbracket$. Thus $S_2 S_1 - \mathbf{1}(\lfloor S_1 \rfloor + \lfloor S_2 \rfloor \llbracket S_1 \rrbracket)$ is nonnegative. But $\lfloor S_2 S_1 \rfloor$ is the largest nonnegative row vector c for which $S_2 S_1 - \mathbf{1}c$ is nonnegative, and so

$$S_2 S_1 \leq \mathbf{1}\lfloor S_2 S_1 \rfloor + \llbracket S_2 \rrbracket \, \llbracket S_1 \rrbracket. \tag{18}$$

Now it is also true that $S_2 S_1 = \mathbf{1}\lfloor S_2 S_1 \rfloor + \llbracket S_2 S_1 \rrbracket$. From this and (18) it follows that (16) is true. $\square$

Proof of Proposition 1. Set $X_j = S_j \cdots S_1, j \geq 1$, and note that each X_j is a stochastic matrix. In view of (15),

$$\lfloor X_{j+1} \rfloor - \lfloor X_j \rfloor \leq \lceil S_{j+1} \rceil \llbracket X_j \rrbracket, \quad j \geq 1.$$

By hypothesis, $||\llbracket X_j \rrbracket|| \leq \bar{b}\lambda^j, j \geq 1$. Moreover, $||\lceil S_{j+1} \rceil|| \leq n$ because all entries in S_{j+1} are bounded above by 1. Therefore

$$||\lfloor X_{j+1} \rfloor - \lfloor X_j \rfloor|| \leq n\bar{b}\lambda^j, \quad j \geq 1. \tag{19}$$

Clearly

$$\lfloor X_{j+i} \rfloor - \lfloor X_j \rfloor = \sum_{k=1}^{i} (\lfloor X_{i+j+1-k} \rfloor - \lfloor X_{i+j-k} \rfloor), \;\; i, j \geq 1.$$

Thus, by the triangle inequality

$$||\lfloor X_{j+i} \rfloor - \lfloor X_j \rfloor|| \leq \sum_{k=1}^{i} ||\lfloor X_{i+j+1-k} \rfloor - \lfloor X_{i+j-k} \rfloor||, \;\; i, j \geq 1.$$

This and (19) imply that

$$||\lfloor X_{j+i} \rfloor - \lfloor X_j \rfloor|| \leq n\bar{b} \sum_{k=1}^{i} \lambda^{(i+j-k)}, \quad i, j \geq 1.$$

Now

$$\sum_{k=1}^{i} \lambda^{(i+j-k)} = \lambda^j \sum_{k=1}^{i} \lambda^{(i-k)} = \lambda^j \sum_{q=1}^{i} \lambda^{q-1} \leq \lambda^j \sum_{q=1}^{\infty} \lambda^{q-1}.$$

But $\lambda < 1$, and so

$$\sum_{q=1}^{\infty} \lambda^{q-1} = \frac{1}{1-\lambda}.$$

Therefore

$$||\lfloor X_{i+j} \rfloor - \lfloor X_j \rfloor|| \leq n\bar{b} \frac{\lambda^j}{1-\lambda}, \quad i, j \geq 1. \tag{20}$$

Set $c = \lfloor \cdots S_j \cdots S_1 \rfloor$ and note that

$$||\lfloor X_j \rfloor - c|| = ||\lfloor X_j \rfloor - \lfloor X_{i+j} \rfloor + \lfloor X_{i+j} \rfloor - c|| \leq ||\lfloor X_j \rfloor - \lfloor X_{i+j} \rfloor|| + ||\lfloor X_{i+j} \rfloor - c||, \quad i, j \geq 1.$$

In view of (20)

$$||\lfloor X_j \rfloor - c|| \leq n\bar{b} \frac{\lambda^j}{1-\lambda} + ||\lfloor X_{i+j} \rfloor - c||, \quad i, j \geq 1.$$

Since

$$\lim_{i \to \infty} ||\lfloor X_{i+j} \rfloor - c|| = 0$$

it must be true that

$$||\lfloor X_j \rfloor - c|| \leq n\bar{b} \frac{\lambda^j}{1-\lambda}, \quad j \geq 1.$$

But $||\mathbf{1}(\lfloor X_j \rfloor - c)|| = ||\lfloor X_j \rfloor - c||$ and $X_j = S_j \cdots S_1$. Therefore

$$||\mathbf{1}(\lfloor S_j \cdots S_1 \rfloor - c)|| \leq n\bar{b} \frac{\lambda^j}{1-\lambda}, \quad j \geq 1. \tag{21}$$

In view of (7)

$$S_j \cdots S_1 = \mathbf{1}\lfloor S_j \cdots S_1 \rfloor + [\![S_j \cdots S_1]\!], \quad j \geq 1.$$

Therefore

$$\begin{aligned} ||(S_j \cdots S_1) - \mathbf{1}c|| &= ||\mathbf{1}\lfloor S_j \cdots S_1 \rfloor + [\![S_j \cdots S_1]\!] - \mathbf{1}c|| \\ &\leq ||\mathbf{1}\lfloor S_j \cdots S_1 \rfloor - \mathbf{1}c|| + ||[\![S_j \cdots S_1]\!]||, \quad j \geq 1. \end{aligned}$$

From this, (14), and (21) it follows that

$$||S_j \cdots S_1 - \mathbf{1}c|| \leq \bar{b}\left(1 + \frac{n}{1-\lambda}\right)\lambda^j, \quad j \geq 1,$$

and thus that (13) holds with $b = \bar{b}(1 + \frac{n}{1-\lambda})$. □

2.3.3. Convergence. We are now in a position to make some statements about the asymptotic behavior of a product of $n \times n$ stochastic matrices of the form $S_j S_{j-1} \cdots S_1$ as j tends to infinity. Note first that (16) generalizes to sequences of stochastic matrices of any length. Thus

$$[\![S_j S_{j-1} \cdots S_2 S_1]\!] \leq [\![S_j]\!] \; [\![S_{j-1}]\!] \cdots [\![S_1]\!]. \tag{22}$$

It is therefore clear that condition (14) of Proposition 1 will hold with $\bar{b} = 1$ if

$$||[\![S_j]\!] \cdots [\![S_1]\!]|| \leq \lambda^j \tag{23}$$

for some nonnegative number $\lambda < 1$. Because $|| \cdot ||$ is submultiplicative, this means that a product of stochastic matrices $S_j \cdots S_1$ will converge to a limit of the form $\mathbf{1}c$ for some constant row vector c if each of the matrices S_i in the sequence $S_1, S_2, \ldots$ satisfies the norm bound $||[\![S_i]\!]|| \leq \lambda$. We now develop a condition, tailored to our application, for this to be so.

As a first step it is useful to characterize those stochastic matrices S for which $||[\![S]\!]|| < 1$. Note that this condition is equivalent to the requirement that the row sums of $[\![S]\!]$ are less than 1. This, in turn, is equivalent to the requirement that $\mathbf{1}\lfloor S \rfloor \neq 0$ since $[\![S]\!] = S - \mathbf{1}\lfloor S \rfloor$. Now $\mathbf{1}\lfloor S \rfloor \neq 0$ if and only if S has at least one nonzero column since the indices of the nonzero columns of S are the same as the indices of the nonzero columns of $\lfloor S \rfloor$. Thus $||[\![S]\!]|| < 1$ if and only if S has at least one nonzero column. For our purposes it proves to be especially useful to restate this condition in equivalent graph theoretic terms. For this we need the following definition.

The graph of a stochastic matrix. For any $n \times n$ stochastic matrix S, let $\gamma(S)$ denote the graph $\mathbb{G} \in \mathcal{G}$ whose adjacency matrix is the transpose of the matrix obtained by replacing all of S's nonzero entries with 1's. The graph-theoretic condition is as follows.

LEMMA 2. *A stochastic matrix S has a strongly rooted graph $\gamma(S)$ if and only if*

$$||[\![S]\!]|| < 1. \tag{24}$$

Proof. Let A be the adjacency matrix of $\gamma(S)$. Since the positions of the nonzero entries of S and A' are the same, the ith column of S will be positive if and only if A's ith row is positive. Thus (23) will hold just in case A has a positive row. But strongly rooted graphs in $\mathcal{G}$ are precisely those graphs whose adjacency matrices have at least one positive row. Therefore (23) will hold if and only if $\gamma(S)$ is strongly rooted. ∎

Lemma 2 can be used to prove the following.

PROPOSITION 2. *Let $\mathcal{S}_{sr}$ be any closed set of stochastic matrices which are all the same size and whose graphs $\gamma(S)$, $S \in \mathcal{S}_{sr}$, are all strongly rooted. Then as $j \to \infty$, any product $S_j \cdots S_1$ of matrices from $\mathcal{S}_{sr}$ converges exponentially fast to $\mathbf{1}\lfloor \cdots S_j \cdots S_1 \rfloor$ at a rate no slower than*

$$\lambda = \max_{S \in \mathcal{S}_{sr}} |||[\![S]\!]|||,$$

where λ is a nonnegative constant satisfying $\lambda < 1$.

Proof of Proposition 2. In view of Lemma 2, $|||[\![S]\!]||| < 1,\ S \in \mathcal{S}_{sr}$. Because $\mathcal{S}_{sr}$ is closed and bounded and $|||[\![\cdot]\!]|||$ is continuous, $\lambda < 1$. Clearly $|||[\![S_i]\!]||| \leq \lambda, i \geq 1$, and so (23) must hold for any sequence of matrices $S_1, S_2, \ldots$ from $\mathcal{S}_{sr}$. Therefore for any such sequence $|||[\![S_j \cdots S_1]\!]||| \leq \lambda^j, j \geq 0$. Thus by Proposition 1, the product $\Pi(j) = S_j S_{j-1} \cdots S_1$ converges to $\mathbf{1}\lfloor \cdots S_j \cdot S_1 \rfloor$ exponentially fast at a rate no slower than λ. ∎

Proof of Theorem 1. Let $\mathcal{F}_{sr}$ denote the set of flocking matrices with strongly rooted graphs. Since $\mathcal{S}_{sa}$ is a finite set, so is the set of strongly rooted graphs in $\mathcal{G}_{sa}$. Therefore $\mathcal{F}_{sr}$ is closed. By assumption, $F(t) \in \mathcal{F}_{sr},\ t \geq 0$. In view of Proposition 2, the product $F(t) \cdots F(0)$ converges exponentially fast to $\mathbf{1}\lfloor \cdots F(t) \cdots F(0) \rfloor$ at a rate no slower than

$$\lambda = \max_{F \in \mathcal{F}_{sr}} |||[\![F]\!]|||.$$

But it is clear from (3) that $\theta(t) = F(t-1) \cdots F(1)F(0)\theta(0), t \geq 1$. Therefore (4) holds with $\theta_{ss} = \lfloor \cdots F(t) \cdots F(0) \rfloor \theta(0)$ and the convergence is exponential. ∎

2.3.4. Convergence rate. Using (9) it is possible to calculate a worst case value for the convergence rate λ used in the proof of Theorem 1. Fix $F \in \mathcal{F}_{sr}$. Because $\gamma(F)$ is strongly rooted, at least one vertex—say the kth—must be a root with arcs to each other vertex. In the context of (1), this means that agent k must be a neighbor of every agent. Thus θ_k must be in each sum in (1). Since each n_i in (1) is bounded above by n, this means that the smallest element in column k of F is bounded below by $\frac{1}{n}$. Since (9) asserts that $|||[\![F]\!]||| = 1 - \lfloor F \rfloor \mathbf{1}$, it must be true that $|||[\![F]\!]||| \leq 1 - \frac{1}{n}$. This holds for all $F \in \mathcal{F}_{sr}$. Moreover, in the worst case when $\mathbb{F}$ is strongly rooted at just one vertex and all vertices are neighbors of at least one common vertex, $|||[\![F]\!]||| = 1 - \frac{1}{n}$. It follows that the worst case convergence rate is

$$\max_{F \in \mathcal{F}_{sr}} |||[\![F]\!]||| = 1 - \frac{1}{n}. \tag{25}$$

An example of a graph of a flocking matrix for which (25) holds is shown in Figure 1.

2.4. Rooted graphs. The proof of Theorem 1 depends crucially on the fact that the graphs encountered along a trajectory of (3) are all strongly rooted. It is natural to ask if this requirement can be relaxed and still have all agents' headings converge to a common value. The aim of this section is to show that this can indeed be accomplished. To do this we need to have a meaningful way of "combining" sequences of graphs so that only the combined graph need be strongly rooted but not necessarily the individual graphs making up the combination. One possible notion of combination

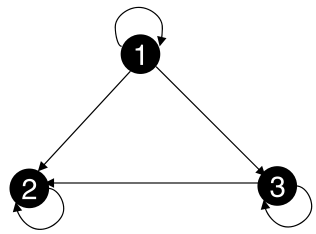

FIG. 1. *Example.*

of a sequence $\mathbb{G}_1, \mathbb{G}_2, \ldots, \mathbb{G}_k$ with the same vertex set $\mathcal{V}$ would be the graph with vertex set $\mathcal{V}$ whose arc set is the union of the arc sets of the graphs in the sequence. It turns out that because we are interested in *sequences* of graphs rather than mere *sets* of graphs, a simple union is not quite the appropriate notion for our purposes because a union does not take into account the order in which the graphs are encountered along a trajectory. What is appropriate is a slightly more general notion which we now define.

2.4.1. Composition of graphs. By the *composition* of a directed graph $\mathbb{G}_p \in \mathcal{G}$ with a directed graph $\mathbb{G}_q \in \mathcal{G}$, written $\mathbb{G}_q \circ \mathbb{G}_p$, we mean the directed graph with the vertex set $\{1, 2, \ldots, n\}$ and arc set defined in such a way so that (i, j) is an arc of the composition just in case there is a vertex k such that (i, k) is an arc of $\mathbb{G}_p$ and (k, j) is an arc of $\mathbb{G}_q$. Thus (i, j) is an arc in $\mathbb{G}_q \circ \mathbb{G}_p$ if and only if i has an observer in $\mathbb{G}_p$ which is also a neighbor of j in $\mathbb{G}_q$. Note that $\mathcal{G}$ is closed under composition and that composition is an associative binary operation; because of this, the definition extends unambiguously to any finite sequence of directed graphs $\mathbb{G}_1,\ \mathbb{G}_2, \ldots, \mathbb{G}_k$.

If we focus exclusively on graphs with self-arcs at all vertices, namely the graphs in $\mathcal{G}_{sa}$, more can be said. In this case the definition of composition implies that the arcs of $\mathbb{G}_p$ and $\mathbb{G}_q$ are arcs of $\mathbb{G}_q \circ \mathbb{G}_p$. The definition also implies in this case that if $\mathbb{G}_p$ has a directed path from i to k and $\mathbb{G}_q$ has a directed path from k to j, then $\mathbb{G}_q \circ \mathbb{G}_p$ has a directed path from i to j. Both of these implications are consequences of the requirement that the vertices of the graphs in $\mathcal{G}_{sa}$ all have self-arcs. Note in addition that $\mathcal{G}_{sa}$ is closed under composition. It is worth emphasizing that the union of the arc sets of a sequence of graphs $\mathbb{G}_1,\ \mathbb{G}_2, \ldots, \mathbb{G}_k$ in $\mathcal{G}_{sa}$ must be contained in the arc set of their composition. However, the converse is not true in general, and it is for this reason that composition rather than union proves to be the more useful concept for our purposes.

Suppose that $A_p = [\, a_{ij}(p) \,]$ and $A_q = [\, a_{ij}(q) \,]$ are the adjacency matrices of $\mathbb{G}_p \in \mathcal{G}$ and $\mathbb{G}_q \in \mathcal{G}$, respectively. Then the adjacency matrix of the composition $\mathbb{G}_q \circ \mathbb{G}_p$ must be the matrix obtained by replacing all nonzero elements in $A_p A_q$ with ones. This is because the ijth entry of $A_p A_q$, namely

$$\sum_{k=1}^{n} a_{ik}(p) a_{kj}(q),$$

will be nonzero just in case there is at least one value of k for which both $a_{ik}(p)$ and $a_{kj}(q)$ are nonzero. This of course is exactly the condition for the ijth element of the adjacency matrix of the composition $\mathbb{G}_q \circ \mathbb{G}_p$ to be nonzero. Note that if S_1 and S_2 are $n \times n$ stochastic matrices for which $\gamma(S_1) = \mathbb{G}_p$ and $\gamma(S_2) = \mathbb{G}_q$, then the matrix which results by replacing by ones all nonzero entries in the stochastic matrix S_2S_1 must be the transpose of the adjacency matrix of $\mathbb{G}_q \circ \mathbb{G}_p$. In view of the definition of $\gamma(\cdot)$, it therefore must be true that $\gamma(S_2S_1) = \gamma(S_2) \circ \gamma(S_1)$. This obviously generalizes to finite products of stochastic matrices.

LEMMA 3. *For any sequence of stochastic matrices $S_1, S_2, \ldots, S_j$ which are all the same size,*

$$\gamma(S_j \cdots S_1) = \gamma(S_j) \circ \cdots \circ \gamma(S_1).$$

2.4.2. Compositions of rooted graphs. We now give several different conditions under which the composition of a sequence of graphs is strongly rooted.

PROPOSITION 3. *Suppose $n > 1$ and let $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}$ be a finite sequence of rooted graphs in $\mathcal{G}_{sa}$.*

1. *If $m \geq (n-1)^2$, then $\mathbb{G}_{p_m} \circ \mathbb{G}_{p_{m-1}} \circ \cdots \circ \mathbb{G}_{p_1}$ is strongly rooted.*
2. *If $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}$ are all rooted at v and $m \geq n-1$, then $\mathbb{G}_{p_m} \circ \mathbb{G}_{p_{m-1}} \circ \cdots \circ \mathbb{G}_{p_1}$ is strongly rooted at v.*

The requirement of assertion 2 above that all the graphs in the sequence be rooted at a single vertex v is obviously more restrictive than the requirement of assertion 1 that all the graphs be rooted but not necessarily at the same vertex. The price for the less restrictive assumption is that the bound on the number of graphs needed in the more general case is much higher than the bound given in the case in which all the graphs are rooted at v. It is probably true that the bound $(n-1)^2$ for the more general case is too conservative, but this remains to be shown. The more special case when all graphs share a common root is relevant to the leader-follower version of the problem which will be discussed later in the paper. Proposition 3 will be proved shortly.

Note that a strongly connected graph is the same as a graph which is rooted at every vertex and that a complete graph is the same as a graph which is strongly rooted at every vertex. In view of these observations and Proposition 3 we can state the following proposition.

PROPOSITION 4. *Suppose $n > 1$ and let $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}$ be a finite sequence of strongly connected graphs in $\mathcal{G}_{sa}$. If $m \geq n-1$, then $\mathbb{G}_{p_m} \circ \mathbb{G}_{p_{m-1}} \circ \cdots \circ \mathbb{G}_{p_1}$ is complete.*

To prove Proposition 3 we will need some more ideas. We say that a vertex $v \in \mathcal{V}$ is an *observer of a subset* $\mathcal{S} \subset \mathcal{V}$ in a graph $\mathbb{G} \in \mathcal{G}$ if v is an observer of at least one vertex in $\mathcal{S}$. By the *observer function* of a graph $\mathbb{G} \in \mathcal{G}$, written $\alpha(\mathbb{G}, \cdot\,)$, we mean the function $\alpha(\mathbb{G}, \cdot\,) : 2^{\mathcal{V}} \to 2^{\mathcal{V}}$ which assigns to each subset $\mathcal{S} \subset \mathcal{V}$ the subset of vertices in $\mathcal{V}$ which are observers of $\mathcal{S}$ in $\mathbb{G}$. Thus $j \in \alpha(\mathbb{G}, i)$ just in case $(i,j) \in \mathcal{A}(\mathbb{G})$. Note that if $\mathbb{G}_p \in \mathcal{G}$ and $\mathbb{G}_q$ in $\mathcal{G}_{sa}$, then

$$\alpha(\mathbb{G}_p, \mathcal{S}) \subset \alpha(\mathbb{G}_q \circ \mathbb{G}_p, \mathcal{S}), \quad \mathcal{S} \in 2^{\mathcal{V}}, \tag{26}$$

because $\mathbb{G}_q \in \mathcal{G}_{sa}$ implies that the arcs in $\mathbb{G}_p$ are all arcs in $\mathbb{G}_q \circ \mathbb{G}_p$. Observer functions have the following important and easily proved property.

LEMMA 4. *For all $\mathbb{G}_p, \mathbb{G}_q \in \mathcal{G}$ and any nonempty subset $\mathcal{S} \subset \mathcal{V}$,*

$$\alpha(\mathbb{G}_q, \alpha(\mathbb{G}_p, \mathcal{S})) = \alpha(\mathbb{G}_q \circ \mathbb{G}_p, \mathcal{S}). \tag{27}$$

Proof. Suppose first that $i \in \alpha(\mathbb{G}_q, \alpha(\mathbb{G}_p, \mathcal{S}))$. Then (j,i) is an arc in $\mathbb{G}_q$ for some $j \in \alpha(\mathbb{G}_p, \mathcal{S})$. Hence (k,j) is an arc in $\mathbb{G}_p$ for some $k \in \mathcal{S}$. In view of the definition of composition, (k,i) is an arc in $\mathbb{G}_q \circ \mathbb{G}_p$, and so $i \in \alpha(\mathbb{G}_q \circ \mathbb{G}_p, \mathcal{S})$. Since this holds for all $i \in \mathcal{V}$, $\alpha(\mathbb{G}_q, \alpha(\mathbb{G}_p, \mathcal{S})) \subset \alpha(\mathbb{G}_q \circ \mathbb{G}_p, \mathcal{S})$.

For the reverse inclusion, fix $i \in \alpha(\mathbb{G}_q \circ \mathbb{G}_p, \mathcal{S})$ in which case (k, i) is an arc in $\mathbb{G}_q \circ \mathbb{G}_p$ for some $k \in \mathcal{S}$. By definition of composition, there exists an $j \in \mathcal{V}$ such that (k, j) is an arc in $\mathbb{G}_p$ and (j, i) is an arc in $\mathbb{G}_q$. Thus $j \in \alpha(\mathbb{G}_p, \mathcal{S})$. Therefore $i \in \alpha(\mathbb{G}_q, \alpha(\mathbb{G}_p, \mathcal{S}))$. Since this holds for all $i \in \mathcal{V}$, $\alpha(\mathbb{G}_q, \alpha(\mathbb{G}_p, \mathcal{S})) \supset \alpha(\mathbb{G}_q \circ \mathbb{G}_p, \mathcal{S})$. Therefore (27) is true. $\square$

To proceed, let us note that each subset $\mathcal{S} \subset \mathcal{V}$ induces a unique subgraph of $\mathbb{G}$ with vertex set $\mathcal{S}$ and arc set $\mathcal{A}$ consisting of those arcs (i, j) of $\mathbb{G}$ for which both i and j are vertices of $\mathcal{S}$. This, together with the natural partial ordering of $\mathcal{V}$ by inclusion, provides a corresponding partial ordering of $\mathcal{G}$. Thus if $\mathcal{S}_1$ and $\mathcal{S}_2$ are subsets of $\mathcal{V}$ and $\mathcal{S}_1 \subset \mathcal{S}_2$, then $\mathbb{G}_1 \subset \mathbb{G}_2$, where, for $i \in \{1, 2\}$, $\mathbb{G}_i$ is the subgraph of $\mathbb{G}$ induced by $\mathcal{S}_i$. For any $v \in \mathcal{V}$, there is a unique largest subgraph rooted at v, namely the graph induced by the vertex set $\mathcal{V}(v) = \{v\} \cup \alpha(\mathbb{G}, v) \cup \cdots \cup \alpha^{n-1}(\mathbb{G}, v)$, where $\alpha^i(\mathbb{G}, \cdot)$ denotes the composition of $\alpha(\mathbb{G}, \cdot)$ with itself i times. We call this graph the *rooted graph generated by* v. It is clear that $\mathcal{V}(v)$ is the smallest $\alpha(\mathbb{G}, \cdot)$-invariant subset of $\mathcal{V}$ which contains v.

The proof of Proposition 3 depends on the following lemma.

LEMMA 5. *Let $\mathbb{G}_p$ and $\mathbb{G}_q$ be graphs in $\mathcal{G}_{sa}$. If $\mathbb{G}_q$ is rooted at v and $\alpha(\mathbb{G}_p, v)$ is a strictly proper subset of $\mathcal{V}$, then $\alpha(\mathbb{G}_p, v)$ is also a strictly proper subset of $\alpha(\mathbb{G}_q \circ \mathbb{G}_p, v)$.*

Proof of Lemma 5. In general $\alpha(\mathbb{G}_p, v) \subset \alpha(\mathbb{G}_q \circ \mathbb{G}_p, v)$ because of (26). Thus if $\alpha(\mathbb{G}_p, v)$ is not a strictly proper subset of $\alpha(\mathbb{G}_q \circ \mathbb{G}_p, v)$, then $\alpha(\mathbb{G}_p, v) = \alpha(\mathbb{G}_q \circ \mathbb{G}_p, v)$, and so $\alpha(\mathbb{G}_q \circ \mathbb{G}_p, v) \subset \alpha(\mathbb{G}_p, v)$. In view of (27), $\alpha(\mathbb{G}_q \circ \mathbb{G}_p, v) = \alpha(\mathbb{G}_q, \alpha(\mathbb{G}_p, v))$. Therefore $\alpha(\mathbb{G}_q, \alpha(\mathbb{G}_p, v)) \subset \alpha(\mathbb{G}_p, v)$. Moreover, $v \in \alpha(\mathbb{G}_p, v)$ because v has a self-arc in $\mathbb{G}_p$. Thus $\alpha(\mathbb{G}_p, v)$ is a strictly proper subset of $\mathcal{V}$ which contains v and is $\alpha(\mathbb{G}_q, \cdot)$-invariant. But this is impossible because $\mathbb{G}_q$ is rooted at v. $\square$

Proof of Proposition 3. Assertion 2 will be proved first. Suppose that $m \geq n-1$ and that $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}$ are all rooted at v. In view of (26), $\mathcal{A}(\mathbb{G}_{p_k} \circ \mathbb{G}_{p_{k-1}} \circ \cdots \circ \mathbb{G}_{p_1}) \subset \mathcal{A}(\mathbb{G}_{p_m} \circ \mathbb{G}_{p_{m-1}} \circ \cdots \circ \mathbb{G}_{p_1})$ for any positive integer $k \leq m$. Thus $\mathbb{G}_{p_m} \circ \mathbb{G}_{p_{m-1}} \circ \cdots \circ \mathbb{G}_{p_1}$ will be strongly rooted at v if there exists an integer $k \leq n-1$ such that

$$\alpha(\mathbb{G}_{p_k} \circ \mathbb{G}_{p_{k-1}} \circ \cdots \circ \mathbb{G}_{p_1}, v) = \mathcal{V}. \tag{28}$$

It will now be shown that such an integer exists.

If $\alpha(\mathbb{G}_{p_1}, v) = \mathcal{V}$, set $k = 1$, in which case (28) clearly holds. If $\alpha(\mathbb{G}_{p_1}, v) \neq \mathcal{V}$, then let $i > 1$ be the greatest positive integer not exceeding $n-1$ for which $\alpha(\mathbb{G}_{p_{i-1}} \circ \cdots \circ \mathbb{G}_{p_1}, v)$ is a strictly proper subset of $\mathcal{V}$. If $i < n-1$, set $k = i$, in which case (28) is clearly true. Therefore suppose $i = n-1$; we will prove that this cannot be so. Assuming that it is, $\alpha(\mathbb{G}_{p_{j-1}} \circ \cdots \circ \mathbb{G}_{p_1}, v)$ must be a strictly proper subset of $\mathcal{V}$ for $j \in \{2, 3, \ldots, n-1\}$; by Lemma 5, $\alpha(\mathbb{G}_{p_{j-1}} \circ \cdots \circ \mathbb{G}_{p_1}, v)$ is also a strictly proper subset of $\alpha(\mathbb{G}_{p_j} \circ \cdots \circ \mathbb{G}_{p_1}, v)$ for $j \in \{2, 3, \ldots, n-1\}$. In view of this and (26), each containment in the ascending chain

$$\alpha(\mathbb{G}_{p_1}, v) \subset \alpha(\mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}, v) \subset \cdots \subset \alpha(\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_1}, v)$$

is strict. Since $\alpha(\mathbb{G}_{p_1}, v)$ has at least two vertices in it, and there are n vertices in $\mathcal{V}$, (28) must hold with $k = n-1$. Thus assertion 2 is true.

To prove assertion 1, suppose that $m \geq (n-1)^2$. Since there are n vertices in $\mathcal{V}$, the sequence $p_1, p_2, \ldots, p_m$ must contain a subsequence $q_1, q_2, \ldots, q_{n-1}$ for which the graphs $\mathbb{G}_{q_1}, \mathbb{G}_{q_2}, \ldots, \mathbb{G}_{q_{n-1}}$ all have a common root. By assertion 2, $\mathbb{G}_{q_{n-1}} \circ \cdots \circ \mathbb{G}_{q_1}$ must be strongly rooted. But $\mathcal{A}(\mathbb{G}_{q_{n-1}} \circ \cdots \circ \mathbb{G}_{q_1}) \subset \mathcal{A}(\mathbb{G}_{p_m} \circ \mathbb{G}_{p_{m-1}} \circ \cdots \circ \mathbb{G}_{p_1})$

because $\mathbb{G}_{q_1}, \mathbb{G}_{q_2}, \dots, \mathbb{G}q_{n-1}$ is a subsequence of $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \dots, \mathbb{G}_{p_m}$ and all graphs in the sequence $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \dots, \mathbb{G}_{p_m}$ have self-arcs. Therefore $\mathbb{G}_{p_m} \circ \mathbb{G}_{p_{m-1}} \circ \cdots \circ \mathbb{G}_{p_1}$ must be strongly rooted. ∎

Proposition 3 implies that every sufficiently long composition of graphs from a given subset $\widehat{\mathcal{G}} \subset \mathcal{G}_{sa}$ will be strongly rooted if each graph in $\widehat{\mathcal{G}}$ is rooted. The converse is also true. To understand why, suppose to the contrary that it is not. In this case there would have to be a graph $\mathbb{G} \in \widehat{\mathcal{G}}$, which is not rooted but for which $\mathbb{G}^m$ is strongly rooted for m sufficiently large, where $\mathbb{G}^m$ is the m-fold composition of $\mathbb{G}$ with itself. Thus $\alpha(\mathbb{G}^m, v) = \mathcal{V}$, where v is a root of $\mathbb{G}^m$. But via repeated application of (27), $\alpha(\mathbb{G}^m, v) = \alpha^m(\mathbb{G}, v)$, where $\alpha^m(\mathbb{G}, \cdot)$ is the m-fold composition of $\alpha(\mathbb{G}, \cdot)$ with itself. Thus $\alpha^m(\mathbb{G}, v) = \mathcal{V}$. But this can occur only if $\mathbb{G}$ is rooted at v because $\alpha^m(\mathbb{G}, v)$ is the set of vertices reachable from v along paths of length m. Since this is a contradiction, $\mathbb{G}$ must be rooted. We summarize.

Proposition 5. *Every possible sufficiently long composition of graphs from a given subset $\widehat{\mathcal{G}} \subset \mathcal{G}_{sa}$ is strongly rooted if and only if every graph in $\widehat{\mathcal{G}}$ is rooted.*

2.4.3. Sarymsakov graphs.

We now briefly discuss a class of graphs in $\mathcal{G}$, namely "Sarymsakov graphs," whose corresponding stochastic matrices form products which are known to converge to rank one matrices [25] even though the graphs in question need not have self-arcs at all vertices. Sarymsakov graphs are defined as follows.

First, let us agree to say that a vertex $v \in \mathcal{V}$ is a *neighbor of a subset* $\mathcal{S} \subset \mathcal{V}$ in a graph $\mathbb{G} \in \mathcal{G}$ if v is a neighbor of at least one vertex in $\mathcal{S}$. By a *Sarymsakov graph* we mean a graph $\mathbb{G} \in \mathcal{G}$ with the property that for each pair of nonempty subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ in $\mathcal{V}$ which have no neighbors in common, $\mathcal{S}_1 \cup \mathcal{S}_2$ contains a smaller number of vertices than does the set of neighbors of $\mathcal{S}_1 \cup \mathcal{S}_2$. Such seemingly obscure graphs are so named because they are the graphs of an important class of nonnegative matrices studied by Sarymsakov in [24]. In what follows we will prove that Sarymsakov graphs are in fact rooted graphs. We will also prove that the class of rooted graphs we are primarily interested in, namely those in $\mathcal{G}_{sa}$, are Sarymsakov graphs.

It is possible to characterize Sarymsakov graphs a little more concisely using the following concept. By the *neighbor function* of a graph $\mathbb{G} \in \mathcal{G}$, written $\beta(\mathbb{G}, \cdot\,)$, we mean the function $\beta(\mathbb{G}, \cdot\,) : 2^{\mathcal{V}} \to 2^{\mathcal{V}}$ which assigns to each subset $\mathcal{S} \subset \mathcal{V}$ the subset of vertices in $\mathcal{V}$ which are neighbors of $\mathcal{S}$ in $\mathbb{G}$. Thus in terms of β, a Sarymsakov graph is a graph $\mathbb{G} \in \mathcal{G}$ with the property that for each pair of nonempty subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ in $\mathcal{V}$ which have no neighbors in common, $\mathcal{S}_1 \cup \mathcal{S}_2$ contains fewer vertices than does the set $\beta(\mathbb{G}, \mathcal{S}_1 \cup \mathcal{S}_2)$. Note that if $\mathbb{G} \in \mathcal{G}_{sa}$, the requirement that $\mathcal{S}_1 \cup \mathcal{S}_2$ contain fewer vertices than $\beta(\mathbb{G}, \mathcal{S}_1 \cup \mathcal{S}_2)$ simplifies to the equivalent requirement that $\mathcal{S}_1 \cup \mathcal{S}_2$ be a strictly proper subset of $\beta(\mathbb{G}, \mathcal{S}_1 \cup \mathcal{S}_2)$. This is because every vertex in $\mathbb{G}$ is a neighbor of itself if $\mathbb{G} \in \mathcal{G}_{sa}$.

Proposition 6.

1. *Each Sarymsakov graph in $\mathcal{G}$ is rooted.*
2. *Each rooted graph in $\mathcal{G}_{sa}$ is a Sarymsakov graph.*

It follows that if we restrict attention exclusively to graphs in $\mathcal{G}_{sa}$, then rooted graphs and Sarymsakov graphs are one and the same.

In what follows $\beta^m(\mathbb{G}, \cdot)$ denotes the m-fold composition of $\beta(\mathbb{G}, \cdot)$ with itself. The proof of Proposition 6 depends on the following ideas.

Lemma 6. *Let $\mathbb{G} \in \mathcal{G}$ be a Sarymsakov graph. Let $\mathcal{S}$ be a nonempty subset of $\mathcal{V}$ such that $\beta(\mathbb{G}, \mathcal{S}) \subset \mathcal{S}$. Let v be any vertex in $\mathcal{V}$. Then there exists a nonnegative integer $m \leq n$ such that $\beta^m(\mathbb{G}, v) \cap \mathcal{S}$ is nonempty.*

Proof. If $v \in \mathcal{S}$, set $m = 0$. Suppose next that $v \notin \mathcal{S}$. Set $\mathcal{T} = \{v\} \cup \beta(\mathbb{G}, v) \cup \cdots \cup \beta^{n-1}(\mathbb{G}, v)$ and note that $\beta^n(\mathbb{G}, v) \subset \mathcal{T}$ because $\mathbb{G}$ has n vertices. Since $\beta(\mathbb{G}, \mathcal{T}) = \beta(\mathbb{G}, v) \cup \beta^2(\mathbb{G}, v) \cup \cdots \cup \beta^n(\mathbb{G}, v)$, it must be true that $\beta(\mathbb{G}, \mathcal{T}) \subset \mathcal{T}$. Therefore

$$\beta(\mathbb{G}, \mathcal{T} \cup \mathcal{S}) \subset \mathcal{T} \cup \mathcal{S}. \tag{29}$$

Suppose $\beta(\mathbb{G}, \mathcal{T}) \cap \beta(\mathbb{G}, \mathcal{S})$ is empty. Then because $\mathbb{G}$ is a Sarymsakov graph, $\mathcal{T} \cup \mathcal{S}$ contains fewer vertices than $\beta(\mathbb{G}, \mathcal{T} \cup \mathcal{S})$. This contradicts (29), and so $\beta(\mathbb{G}, \mathcal{T}) \cap \beta(\mathbb{G}, \mathcal{S})$ is not empty. In view of the fact that $\beta(\mathbb{G}, \mathcal{T}) = \beta(\mathbb{G}, v) \cup \beta^2(\mathbb{G}, v) \cup \cdots \cup \beta^n(\mathbb{G}, v)$, it must therefore be true for some positive integer $m \leq n$ that $\beta^m(\mathbb{G}, v) \cap \beta(\mathbb{G}, \mathcal{S})$ is nonempty. But by assumption $\beta(\mathbb{G}, \mathcal{S}) \subset \mathcal{S}$, and so $\beta^m(\mathbb{G}, v) \cap \mathcal{S}$ is nonempty. ∎

LEMMA 7. *Let $\mathbb{G} \in \mathcal{G}$ be rooted at r. Each nonempty subset $\mathcal{S} \subset \mathcal{V}$ not containing r is a strictly proper subset of $\mathcal{S} \cup \beta(\mathbb{G}, \mathcal{S})$.*

Proof of Lemma 7. Let $\mathcal{S} \subset \mathcal{V}$ be nonempty and not containing r. Pick $v \in \mathcal{S}$. Since $\mathbb{G}$ is rooted at r, there must be a path in $\mathbb{G}$ from r to v. Since $r \notin \mathcal{S}$ there must be a vertex $x \in \mathcal{S}$ which has a neighbor which is not in $\mathcal{S}$. Thus there is a vertex $y \in \beta(\mathbb{G}, \mathcal{S})$ which is not in $\mathcal{S}$. This implies that $\mathcal{S}$ is a strictly proper subset of $\mathcal{S} \cup \beta(\mathbb{G}, \mathcal{S})$. ∎

By a *maximal rooted subgraph of* $\mathbb{G}$ we mean a subgraph $\mathbb{G}^*$ of $\mathbb{G}$ which is rooted and which is not contained in any rooted subgraph of $\mathbb{G}$ other than itself. Graphs in $\mathcal{G}$ may have one or more maximal rooted subgraphs. Clearly $\mathbb{G}^* = \mathbb{G}$ just in case $\mathbb{G}$ is rooted. Note that if $\widehat{\mathcal{R}}$ is the set of all roots of a maximal rooted subgraph $\widehat{\mathbb{G}}$, then $\beta(\mathbb{G}, \widehat{\mathcal{R}}) \subset \widehat{\mathcal{R}}$. For if this were not so, then it would be possible to find a vertex $x \in \beta(\mathbb{G}, \widehat{\mathcal{R}})$ which is not in $\widehat{\mathcal{R}}$. This would imply the existence of a path from x to some root $\widehat{v} \in \widehat{\mathcal{R}}$; consequently the graph induced by the set of vertices along this path together with $\widehat{\mathcal{R}}$ would be rooted at $x \notin \widehat{\mathcal{R}}$ and would contain $\widehat{\mathbb{G}}$ as a strictly proper subgraph. But this contradicts the hypothesis that $\widehat{\mathbb{G}}$ is maximal. Therefore $\beta(\mathbb{G}, \widehat{\mathcal{R}}) \subset \widehat{\mathcal{R}}$. Now suppose that $\widehat{\mathbb{G}}$ is any rooted subgraph in $\mathcal{G}$. Suppose that $\widehat{\mathbb{G}}$'s set of roots $\widehat{\mathcal{R}}$ satisfies $\beta(\mathbb{G}, \widehat{\mathcal{R}}) \subset \widehat{\mathcal{R}}$. We claim that $\widehat{\mathbb{G}}$ must then be maximal. For if this were not so, there would have to be a rooted graph $\mathbb{G}^*$ containing $\widehat{\mathbb{G}}$ as a strictly proper subset. This, in turn, would imply the existence of a path from a root x^* of $\mathbb{G}^*$ to a root v of $\widehat{\mathbb{G}}$; consequently $x^* \in \beta^i(\mathbb{G}, \widehat{\mathcal{R}})$ for some $i \geq 1$. But this is impossible because $\widehat{\mathcal{R}}$ is $\beta(\mathbb{G}, \cdot\,)$ invariant. Thus $\widehat{\mathbb{G}}$ is maximal. We summarize.

LEMMA 8. *A rooted subgraph of a graph $\mathbb{G}$ generated by any vertex $v \in \mathcal{V}$ is maximal if and only if its set of roots is $\beta(\mathbb{G}, \cdot\,)$-invariant.*

Proof of Proposition 6. Write $\beta(\cdot)$ for $\beta(\mathbb{G}, \cdot)$. To prove assertion 1, pick $\mathbb{G} \in \mathcal{G}$. Let $\mathbb{G}^*$ be any maximal rooted subgraph of $\mathbb{G}$ and write $\mathcal{R}$ for its root set; in view of Lemma 8, $\beta(\mathcal{R}) \subset \mathcal{R}$. Pick any $v \in \mathcal{V}$. Then by Lemma 6, for some positive integer $m \leq n$, $\beta^m(v) \cap \mathcal{R}$ is nonempty. Pick $z \in \beta^m(v) \cap \mathcal{R}$. Then there is a path from z to v and z is a root of $\mathbb{G}^*$. But $\mathbb{G}^*$ is maximal, and so v must be a vertex of $\mathbb{G}^*$. Therefore every vertex of $\mathbb{G}$ is a vertex of $\mathbb{G}^*$, which implies that $\mathbb{G}$ is rooted.

To prove assertion 2, let $\mathbb{G} \in \mathcal{G}_{sa}$ be rooted at r. Pick any two nonempty subsets $\mathcal{S}_1, \mathcal{S}_2$ of $\mathcal{V}$ which have no neighbors in common. If $r \notin \mathcal{S}_1 \cup \mathcal{S}_2$, then $\mathcal{S}_1 \cup \mathcal{S}_2$ must be a strictly proper subset of $\mathcal{S}_1 \cup \mathcal{S}_2 \cup \beta(\mathcal{S}_1 \cup \mathcal{S}_2)$ because of Lemma 7.

Suppose next that $r \in \mathcal{S}_1 \cup \mathcal{S}_2$. Since $\mathbb{G} \in \mathcal{G}_{sa}$, $\mathcal{S}_i \subset \beta(\mathcal{S}_i)$, $i \in \{1, 2\}$. Thus $\mathcal{S}_1$ and $\mathcal{S}_2$ must be disjoint because $\beta(\mathcal{S}_1)$ and $\beta(\mathcal{S}_2)$ are. Therefore r must be in either $\mathcal{S}_1$ or $\mathcal{S}_2$ but not both. Suppose that $r \notin \mathcal{S}_1$. Then $\mathcal{S}_1$ must be a strictly proper subset of $\beta(\mathcal{S}_1)$ because of Lemma 7. Since $\beta(\mathcal{S}_1)$ and $\beta(\mathcal{S}_2)$ are disjoint, $\mathcal{S}_1 \cup \mathcal{S}_2$ must be a strictly proper subset of $\beta(\mathcal{S}_1 \cup \mathcal{S}_2)$. By the same reasoning, $\mathcal{S}_1 \cup \mathcal{S}_2$ must be a strictly proper subset of $\beta(\mathcal{S}_1 \cup \mathcal{S}_2)$ if $r \notin \mathcal{S}_2$. Thus in conclusion $\mathcal{S}_1 \cup \mathcal{S}_2$ must be a strictly proper subset of $\beta(\mathcal{S}_1 \cup \mathcal{S}_2)$ whether or not r is in $\mathcal{S}_1 \cup \mathcal{S}_2$. Since this conclusion holds for all such $\mathcal{S}_1$ and $\mathcal{S}_2$ and $\mathbb{G} \in \mathcal{G}_{sa}$, $\mathbb{G}$ must be a Sarymsakov graph. $\square$

2.5. Neighbor-shared graphs. There is a different assumption which one can make about a sequence of graphs from $\mathcal{G}$ which also ensures that the sequence's composition is strongly rooted. For this we need the concept of a "neighbor-shared graph." Let us call $\mathbb{G} \in \mathcal{G}$ *neighbor-shared* if each set of two distinct vertices shares a common neighbor. Suppose that $\mathbb{G}$ is neighbor-shared. Then both vertices in any given pair of vertices are clearly reachable from a single vertex along directed paths. Suppose that for some integer $k \in \{2, 3, \dots, n-1\}$, each subset of k vertices $\{v_1, v_2, \dots, v_k\}$ has the property that every vertex in $\{v_1, v_2, \dots, v_k\}$ is reachable from a single vertex. Let $\{v_1, v_2, \dots, v_k\}$ be any such set and v be a vertex from which all k vertices in the set can be reached. Let w be any vertex not in $\{v_1, v_2, \dots, v_k\}$. Since v and w can be reached from a common vertex y, every vertex in $\{v_1, v_2, \dots, v_k, w\}$ can be reached from y. This proves that each subset of $k+1$ vertices has the property that every vertex in the subset is reachable from a single vertex. By induction we can therefore conclude that every vertex in $\mathbb{G}$ is reachable from a single vertex. We have proved the following proposition.

PROPOSITION 7. *Each neighbor-shared graph in $\mathcal{G}$ is rooted.*

It is worth noting that although neighbor-shared graphs are rooted, the converse is not necessarily true. The reader may wish to construct a three-vertex example which illustrates this. Although rooted graphs in $\mathcal{G}_{sa}$ need not be neighbor-shared, it turns out that the composition of any $n-1$ rooted graphs in $\mathcal{G}_{sa}$ is.

PROPOSITION 8. *The composition of any set of $m \geq n-1$ rooted graphs in $\mathcal{G}_{sa}$ is neighbor-shared.*

This result is equivalent to Theorem 5.1 of [31], which was independently derived.

To prove Proposition 8 we need some more ideas. By the *reverse graph* of $\mathbb{G} \in \mathcal{G}$, written $\mathbb{G}'$, we mean the graph in $\mathcal{G}$ which results when the directions of all arcs in $\mathbb{G}$ are reversed. It is clear that $\mathcal{G}_{sa}$ is closed under the reverse operation and that if A is the adjacency matrix of $\mathbb{G}$, then A' is the adjacency matrix of $\mathbb{G}'$. It is also clear that $(\mathbb{G}_p \circ \mathbb{G}_q)' = \mathbb{G}_q' \circ \mathbb{G}_p'$, $p, q \in \mathcal{P}$, and that

$$\alpha(\mathbb{G}', \mathcal{S}) = \beta(\mathbb{G}, \mathcal{S}), \quad \mathcal{S} \in 2^{\mathcal{V}}. \tag{30}$$

LEMMA 9. *For all $\mathbb{G}_p, \mathbb{G}_q \in \mathcal{G}$ and any nonempty subset $\mathcal{S} \subset \mathcal{V}$,*

$$\beta(\mathbb{G}_q, \beta(\mathbb{G}_p, \mathcal{S})) = \beta(\mathbb{G}_p \circ \mathbb{G}_q, \mathcal{S}). \tag{31}$$

Proof of Lemma 9. In view of (27), $\alpha(\mathbb{G}_p', \alpha(\mathbb{G}_q', \mathcal{S})) = \alpha(\mathbb{G}_p' \circ \mathbb{G}_q', \mathcal{S})$. But $\mathbb{G}_p' \circ \mathbb{G}_q' = (\mathbb{G}_q \circ \mathbb{G}_p)'$, and so $\alpha(\mathbb{G}_p', \alpha(\mathbb{G}_q', \mathcal{S})) = \alpha((\mathbb{G}_q \circ \mathbb{G}_p)', \mathcal{S})$. Therefore $\beta(\mathbb{G}_p, \beta(\mathbb{G}_q, \mathcal{S})) = \beta(\mathbb{G}_q \circ \mathbb{G}_p), \mathcal{S})$ because of (30). $\square$

LEMMA 10. *Let $\mathbb{G}_p$ and $\mathbb{G}_q$ be rooted graphs in $\mathcal{G}_{sa}$. If u and v are distinct vertices in $\mathcal{V}$ for which*

$$\beta(\mathbb{G}_q, \{u, v\}) = \beta(\mathbb{G}_q \circ \mathbb{G}_p, \{u, v\}), \tag{32}$$

then u and v have a common neighbor in $\mathbb{G}_q \circ \mathbb{G}_p$.

Proof. $\beta(\mathbb{G}_q, u)$ and $\beta(\mathbb{G}_q, v)$ are nonempty because u and v are neighbors of themselves. Suppose u and v do not have a common neighbor in $\mathbb{G}_q \circ \mathbb{G}_p$. Then $\beta(\mathbb{G}_q \circ \mathbb{G}_p, u)$ and $\beta(\mathbb{G}_q \circ \mathbb{G}_p, v)$ are disjoint. But $\beta(\mathbb{G}_q \circ \mathbb{G}_p, u) = \beta(\mathbb{G}_p, \beta(\mathbb{G}_q, u))$ and $\beta(\mathbb{G}_q \circ \mathbb{G}_p, v) = \beta(\mathbb{G}_p, \beta(\mathbb{G}_q, v))$ because of (31). Therefore $\beta(\mathbb{G}_p, \beta(\mathbb{G}_q, u))$ and

$\beta(\mathbb{G}_p, \beta(\mathbb{G}_q, v))$ are disjoint. But $\mathbb{G}_p$ is rooted and thus a Sarymsakov graph because of Proposition 6. Thus $\beta(\mathbb{G}_q, \{u,v\})$ is a strictly proper subset of $\beta(\mathbb{G}_q, \{u,v\}) \cup \beta(\mathbb{G}_p, \beta(\mathbb{G}_q, \{u,v\})$. But $\beta(\mathbb{G}_q, \{u,v\}) \subset \beta(\mathbb{G}_p, \beta(\mathbb{G}_q, \{u,v\})$ because all vertices in $\mathbb{G}_q$ are neighbors of themselves and $\beta(\mathbb{G}_p, \beta(\mathbb{G}_q, \{u,v\}) = \beta(\mathbb{G}_q \circ \mathbb{G}_p, \{u,v\})$ because of (31). Therefore $\beta(\mathbb{G}_q, \{u,v\})$ is a strictly proper subset of $\beta(\mathbb{G}_q \circ \mathbb{G}_p, \{u,v\})$. This contradicts (32), and so u and v have a common neighbor in $\mathbb{G}_q \circ \mathbb{G}_p$. ∎

Proof of Proposition 8. Let u and v be distinct vertices in $\mathcal{V}$. Let $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_{n-1}}$ be a sequence of rooted graphs in $\mathcal{G}_{sa}$. Since $\mathcal{A}(\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_{n-i}}) \subset \mathcal{A}(\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_{n-(i+1)}})$ for $i \in \{1, 2, \ldots, n-2\}$, it must be true that the $\mathbb{G}_p$ yield the ascending chain

$$\beta(\mathbb{G}_{n-1}, \{u,v\}) \subset \beta(\mathbb{G}_{p_{n-1}} \circ \mathbb{G}_{pn-2}, \{u,v\}) \subset \cdots \subset \beta(\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}, \{u,v\}).$$

Because there are n vertices in $\mathcal{V}$, this chain must converge for some $i < n-1$, which means that

$$\beta(\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_{n-i}}, \{u,v\}) = \beta(\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_{n-i}} \circ \mathbb{G}_{p_{n-(i+1)}}, \{u,v\}).$$

This and Lemma 10 imply that u and v have a common neighbor in $\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_{n-i}}$ and thus in $\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}$. Since this is true for all distinct u and v, $\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}$ is a neighbor-shared graph. ∎

If we restrict attention to those rooted graphs in $\mathcal{G}_{sa}$ which are strongly connected, we can obtain a neighbor-shared graph by composing a smaller number of rooted graphs than the one claimed in Proposition 8.

PROPOSITION 9. *Let q be the integer quotient of n divided by* 2. *The composition of any set of $m \geq q$ strongly connected graphs in $\mathcal{G}_{sa}$ is neighbor-shared.*

Proof of Proposition 9. Let $k < n$ be a positive integer and let v be any vertex in $\mathcal{V}$. Let $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_k}$ be a sequence of strongly connected graphs in $\mathcal{G}_{sa}$. Since each vertex of a strongly connected graph must be a root, v must be a root of each $\mathbb{G}_{p_i}$. Note that the $\mathbb{G}_{p_i}$ yield the ascending chain

$$\{v\} \subset \beta(\mathbb{G}_{p_k}, \{v\}) \subset \beta(\mathbb{G}_{p_k} \circ \mathbb{G}_{p_{k-1}}, \{v\}) \subset \cdots \subset \beta(\mathbb{G}_{p_k} \circ \cdots \circ \mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}, \{v\})$$

because $\mathcal{A}(\mathbb{G}_{p_k} \circ \cdots \circ \mathbb{G}_{p_{k-(i-1)}}) \subset \mathcal{A}(\mathbb{G}_{p_k} \circ \cdots \circ \mathbb{G}_{p_{k-i}})$ for $i \in \{1, 2, \ldots, k-1\}$. Moreover, since $k < n$ and v is a root of each $\mathbb{G}_{p_k} \circ \cdots \circ \mathbb{G}_{p_{k-(i-1)}}$, $i \in \{1, 2, \ldots, k\}$, it must be true for each such i that $\beta(\mathbb{G}_{p_k} \circ \cdots \circ \mathbb{G}_{p_{k-(i-1)}}, v)$ contains at least $i+1$ vertices. In particular $\beta(\mathbb{G}_{p_k} \circ \cdots \circ \mathbb{G}_{p_1}, v)$ contains at least $k+1$ vertices.

Set $k = q$ and let v_1 and v_2 be any pair of distinct vertices in $\mathcal{V}$. Then there must be at least $q+1$ vertices in $\beta(\mathbb{G}_{p_q} \circ \cdots \circ \mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}, \{v_1\})$ and $q+1$ vertices in $\beta(\mathbb{G}_{p_q} \circ \cdots \circ \mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}, \{v_2\})$. But $2(q+1) > n$ because of the definition of q, and so $\beta(\mathbb{G}_{p_q} \circ \cdots \circ \mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}, \{v_1\})$ and$\beta(\mathbb{G}_{p_q} \circ \cdots \circ \mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}, \{v_2\})$ must have at least one vertex in common. Since this is true for each pair of distinct vertices $v_1, v_2 \in \mathcal{V}$, $\mathbb{G}_{p_q} \circ \cdots \circ \mathbb{G}_2 \circ \mathbb{G}_1$ must be neighbor-shared. ∎

Lemma 7 and Proposition 3 imply that any composition of $(n-1)^2$ neighbor-shared graphs in $\mathcal{G}_{sa}$ is strongly rooted. The following proposition asserts that the composition need only consist of $(n-1)$ neighbor-shared graphs and, moreover, that the graphs need only be in $\mathcal{G}$ and not necessarily in $\mathcal{G}_{sa}$.

PROPOSITION 10. *The composition of any set of $m \geq n-1$ neighbor-shared graphs in $\mathcal{G}$ is strongly rooted.*

Note that Propositions 8 and 10 imply the first assertion of Proposition 3.

To prove Proposition 10 we need a few more ideas. For any integer $1 < k \leq n$, we say that a graph $\mathbb{G} \in \mathcal{G}$ is k *neighbor-shared* if each set of k distinct vertices shares a common neighbor. Thus a neighbor-shared graph and a 2 neighbor-shared graph are one and the same. Clearly an n neighbor-shared graph is strongly rooted at the common neighbor of all n vertices.

LEMMA 11. *If $\mathbb{G}_p \in \mathcal{G}$ is a neighbor-shared graph and $\mathbb{G}_q \in \mathcal{G}$ is a k neighbor-shared graph with $k < n$, then $\mathbb{G}_q \circ \mathbb{G}_p$ is a $(k+1)$ neighbor-shared graph.*

Proof. Let $v_1, v_2, \dots, v_{k+1}$ be any distinct vertices in $\mathcal{V}$. Since $\mathbb{G}_q$ is a k neighbor-shared graph, the vertices $v_1, v_2, \dots, v_k$ share a common neighbor u_1 in $\mathbb{G}_q$ and the vertices $v_2, v_3, \dots, v_{k+1}$ share a common neighbor u_2 in $\mathbb{G}_q$ as well. Moreover, since $\mathbb{G}_p$ is a neighbor-shared graph, u_1 and u_2 share a common neighbor w in $\mathbb{G}_p$. It follows from the definition of composition that $v_1, v_2, \dots, v_k$ have w as a neighbor in $\mathbb{G}_q \circ \mathbb{G}_p$ as do $v_2, v_3, \dots, v_{k+1}$. Therefore $v_1, v_2, \dots, v_{k+1}$ have w as a neighbor in $\mathbb{G}_q \circ \mathbb{G}_p$. Since this must be true for any set of $k+1$ vertices in $\mathbb{G}_q \circ \mathbb{G}_p$, $\mathbb{G}_q \circ \mathbb{G}_p$ must be a $(k+1)$ neighbor-shared graph as claimed. $\square$

Proof of Proposition 10. The preceding lemma implies that the composition of any 2 neighbor-shared graphs is 3 neighbor-shared. From this and induction it follows that for $m < n$, the composition of m neighbor-shared graphs is $(m+1)$ neighbor-shared. Thus the composition of $(n-1)$ neighbor-shared graphs is n neighbor-shared and consequently strongly rooted. $\square$

2.6. Convergence. We are now in a position to significantly relax the conditions under which the conclusion of Theorem 1 holds. Towards this end, recall that each flocking matrix F is row stochastic. Moreover, because each vertex of each F's graph $\gamma(F)$ has a self-arc, the F have the additional property that their diagonal elements are all nonzero. Let $\mathcal{S}$ denote the set of all $n \times n$ row stochastic matrices whose diagonal elements are all positive. $\mathcal{S}$ is closed under multiplication because the class of all $n \times n$ stochastic matrices is closed under multiplication and because the class of $n \times n$ nonnegative matrices with positive diagonals is also.

THEOREM 2. *Let $\theta(0)$ be fixed. For any trajectory of the system* (3) *along which each graph in the sequence of neighbor graphs $\mathbb{N}(0),\ \mathbb{N}(1), \dots$ is rooted, there is a constant steady state heading θ_{ss} for which*

$$\lim_{t\to\infty} \theta(t) = \theta_{ss}\mathbf{1}, \tag{33}$$

where the limit is approached exponentially fast.

The theorem says that a unique heading is achieved asymptotically along any trajectory on which all neighbor graphs are rooted. It is possible to deduce an explicit convergence rate for the situation addressed by this theorem [8, 7]. The theorem's proof relies on the following generalization of Proposition 2. The proposition exploits the fact that any composition of sufficiently many rooted graphs in $\mathcal{G}_{sa}$ is strongly rooted (cf. Proposition 3).

PROPOSITION 11. *Let $\mathcal{S}_r$ be any closed set of $n \times n$ stochastic matrices with rooted graphs in $\mathcal{G}_{sa}$. There exists an integer m such that the graph of the product of every set of m matrices from $\mathcal{S}_r$ is strongly rooted. Let m be any such integer and write $\mathcal{S}_r^m$ for the set of all such matrix products. Then as $j \to \infty$, any product $S_j \cdots S_1$ of matrices from $\mathcal{S}_r$ converges exponentially fast to $\mathbf{1}\lfloor \cdots S_j \cdots S_1 \rfloor$ at a rate no slower than*

$$\lambda = \left(\max_{S \in \mathcal{S}_r^m} |||[\![S]\!]||| \right)^{\frac{1}{m}},$$

where $\lambda < 1$.

Proof of Proposition 11. By assumption, each graph $\gamma(S)$, $S \in \mathcal{S}_r$, is in $\mathcal{G}_{sa}$ and is rooted. In view of Proposition 3, $\gamma(S_q) \circ \cdots \circ \gamma(S_1)$ is strongly rooted for every list of q matrices $\{S_1, S_2, \ldots, S_q\}$ from $\mathcal{S}_r$, provided $q \geq (n-1)^2$. But $\gamma(S_q) \circ \cdots \circ \gamma(S_1) = \gamma(S_q \cdots S_1)$ because of Lemma 3. Therefore $\gamma(S_q \cdots S_1)$ is strongly rooted for all products $S_q \cdots S_1$, where each $S_i \in \mathcal{S}_r$. Thus m could be taken as q, which establishes the existence of such an integer.

Now any product $S_j \cdots S_1$ of matrices in $\mathcal{S}_r$ can be written as $S_j \cdots S_1 = \bar{S}(j)\bar{S}_k \cdots \bar{S}_1$, where $\bar{S}_i = S_{im} \cdots S_{(i-1)m+1}$, $1 \leq i \leq k$, is a product in $\mathcal{S}_r^m$, $\bar{S}(j) = S_j \cdots S_{(km+1)}$, and k is the integer quotient of j divided by m. In view of Proposition 2, $\bar{S}_k \cdots \bar{S}_1$ must converge to $\mathbf{1}\lfloor \cdots \bar{S}_k \cdots \bar{S}_1 \rfloor$ exponentially fast as $k \to \infty$ at a rate no slower than $\bar{\lambda}$, where

$$\bar{\lambda} = \max_{\bar{S} \in \mathcal{S}_r^m} |||[\![\bar{S}]\!]|||.$$

But $\bar{S}(j)$ is a product of at most m stochastic matrices, and so it is a bounded function of j. It follows that the product $S_j S_{j-1} \cdots S_1$ must converge to $\mathbf{1}\lfloor \cdots S_j \cdot S_1 \rfloor$ exponentially fast at a rate no slower than $\lambda = \bar{\lambda}^{\frac{1}{m}}$. $\square$

The proof of Proposition 11 can also be applied to any closed subset $\mathcal{S}_{ns} \subset \mathcal{S}$ of stochastic matrices with neighbor-shared graphs. In this case, one would define $m = n-1$ because of Proposition 10. Similarly, the proof also applies to any closed subset of stochastic matrices whose graphs share a common root; in this case one would define $m = n-1$ because of the first assertion of Proposition 3.

Proof of Theorem 2. Let $\mathcal{F}_r$ denote the set of flocking matrices with rooted graphs. Since $\mathcal{G}_{sa}$ is a finite set, so is the set of rooted graphs in $\mathcal{G}_{sa}$. Therefore $\mathcal{F}_r$ is closed. By assumption, $F(t) \in \mathcal{F}_r$, $t \geq 0$. In view of Proposition 11, the product $F(t) \cdots F(0)$ converges exponentially fast to $\mathbf{1}\lfloor \cdots F(t) \cdots F(0) \rfloor$ at a rate no slower than

$$\lambda = \left(\max_{S \in \mathcal{F}_r^m} |||[\![S]\!]||| \right)^{\frac{1}{m}},$$

where $m = (n-1)^2$ and $\mathcal{F}_r^m$ is the finite set of all m-term flocking matrix products of the form $F_m \cdots F_1$ with each $F_i \in \mathcal{F}_r$. But it is clear from (3) that $\theta(t) = F(t-1) \cdots F(1)F(0)\theta(0)$, $t \geq 1$. Therefore (33) holds with $\theta_{ss} = \lfloor \cdots F(t) \cdots F(0) \rfloor \theta(0)$ and the convergence is exponential. $\square$

The proof of Theorem 2 also applies to the case when all of the $\mathbb{N}(t)$, $t \geq 0$, are neighbor-shared. In this case, one would define $m = n-1$ because of Proposition 10. By similar reasoning, the proof also applies to the case when all of the $\mathbb{N}(t)$, $t \geq 0$, shared a common root; one would also define $m = n-1$ for this case because of the first assertion of Proposition 3.

2.7. Jointly rooted sets of graphs. It is possible to relax further still the conditions under which the conclusion of Theorem 1 holds. Towards this end, let us agree to say that a finite sequence of directed graphs $\mathbb{G}_{p_1}$, $\mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_k}$ in $\mathcal{G}$ is *jointly rooted* if the composition $\mathbb{G}_{p_k} \circ \mathbb{G}_{p_{k-1}} \circ \cdots \circ \mathbb{G}_{p_1}$ is rooted.

Note that since the arc sets of any graphs $\mathbb{G}_p, \mathbb{G}_q \in \mathcal{G}_{sa}$ are contained in the arc set of any composed graph $\mathbb{G}_q \circ \mathbb{G} \circ \mathbb{G}_p$, $\mathbb{G} \in \mathcal{G}_{sa}$, it must be true that if $\mathbb{G}_{p_1}$, $\mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_k}$ is a jointly rooted sequence in $\mathcal{G}_{sa}$, then so is $\mathbb{G}_q, \mathbb{G}_{p_1}$, $\mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_k}, \mathbb{G}_p$. In other words, a jointly rooted sequence of graphs in $\mathcal{G}_{sa}$ remain jointly rooted if additional graphs from $\mathcal{G}_{sa}$ are added to either end of the sequence.

There is an analogous concept for neighbor-shared graphs. We say that a finite sequence of directed graphs $\mathbb{G}_{p_1}$, $\mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_k}$ from $\mathcal{G}$ is *jointly neighbor-shared* if the composition $\mathbb{G}_{p_k} \circ \mathbb{G}_{p_{k-1}} \circ \cdots \circ \mathbb{G}_{p_1}$ is a neighbor-shared graph. Jointly neighbor-shared sequences of graphs from $\mathcal{G}_{sa}$ remain jointly neighbor-shared if additional graphs from $\mathcal{G}_{sa}$ are added to either end of the sequence. The reason for this is the same as for the case of jointly rooted sequences. Although the discussion which follows is just for the case of jointly rooted graphs, the material covered extends in the obvious way to the case of jointly neighbor-shared graphs.

In what follows we will say that an infinite sequence of graphs $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots$ in $\mathcal{G}$ is *repeatedly jointly rooted* if there is a positive integer q for which each finite sequence $\mathbb{G}_{p_{q(k-1)+1}}, \ldots, \mathbb{G}_{p_{qk}}$, $k \geq 1$, is jointly rooted. If such an integer exists, we sometimes say that $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots$ is repeatedly jointly rooted *by subsequences of length* q. We are now in a position to generalize Proposition 11.

PROPOSITION 12. *Let $\bar{\mathcal{S}}$ be any closed set of stochastic matrices with graphs in $\mathcal{G}_{sa}$. Suppose that $S_1, S_2, \ldots$ is an infinite sequence of matrices from $\bar{\mathcal{S}}$ whose corresponding sequence of graphs $\gamma(S_1)$, $\gamma(S_2), \ldots$ is repeatedly jointly rooted by subsequences of length q. Suppose that the set of all products of q matrices from $\bar{\mathcal{S}}$ with rooted graphs, written $\bar{\mathcal{S}}(q)$, is closed. There exists an integer m such that the product of every set of m matrices from $\bar{\mathcal{S}}(q)$ is strongly rooted. Let m be any such integer and write $(\bar{\mathcal{S}}(q))^m$ for the set of all such matrix products. Then as $j \to \infty$, the product $S_j \cdots S_1$ converges exponentially fast to $\mathbf{1}\lfloor \cdots S_j \cdots S_1 \rfloor$ at a rate no slower than*

$$\lambda = \left(\max_{S \in (\bar{\mathcal{S}}(q))^m} |||[\![S]\!]||| \right)^{\frac{1}{mq}},$$

where $\lambda < 1$.

It is worth pointing out that the assumption that $\bar{\mathcal{S}}(q)$ is closed is not necessarily implied by the assumption that $\bar{\mathcal{S}}$ is closed. For example, if $\bar{S}$ is the set of all 2×2 stochastic matrices whose diagonal elements are no smaller than some positive number $\alpha < 1$, then $\bar{\mathcal{S}}(2)$ cannot be closed even though $\bar{\mathcal{S}}$ is; this is because there are matrices in $\bar{S}(2)$ which are arbitrarily close (in the induced infinity norm) to the 2×2 identity which, in turn, is not in $\bar{S}(2)$. There are at least three different situations where $\bar{S}(q)$ turns out to be closed. The first is when $\bar{S}$ is a finite set, as is the case when $\bar{S}$ is all $n \times n$ flocking matrices; in this case it is obvious that for any $q \geq 1$, $\bar{S}(q)$ is closed because it is also a finite set.

The second situation arises when the simple average rule (1) is replaced by a convex combination rule as was done in [3]. In this case, the set $\bar{\mathcal{S}}$ turns out to be all $n \times n$ stochastic matrices whose diagonal entries are nonzero and whose nonzero entries (on the diagonal or not) are all underbounded by a positive number $\alpha < 1$. In this case it is easy to see that for each graph $\mathbb{G} \in \mathcal{G}_{sa}$, the subset $\bar{\mathcal{S}}(\mathbb{G})$ of $S \in \bar{\mathcal{S}}$ for which $\gamma(S) = \mathbb{G}$ is closed. Thus for any pair of graphs $\mathbb{G}_1, \mathbb{G}_2 \in \mathcal{G}_{sa}$, the subset of products $S_2 S_1$ such that $S_1 \in \bar{\mathcal{S}}(\mathbb{G}_1)$ and $S_2 \in \bar{\mathcal{S}}(\mathbb{G}_2)$ is also closed. Since $\bar{\mathcal{S}}(2)$ is the union of a finite number of sets of products of this type, namely those for which the pairs $(\mathbb{G}_1, \mathbb{G}_2)$ have rooted compositions $\mathbb{G}_2 \circ \mathbb{G}_1$, it must be that $\bar{\mathcal{S}}(2)$ is closed. Continuing this reasoning, one can conclude that for any integer $q > 0$, $\bar{\mathcal{S}}(q)$ is closed as well.

The third situation in which $\bar{\mathcal{S}}(q)$ turns out to be compact is considerably more complicated and arises in connection with an asynchronous version of the flocking problem we have been studying. In this case, the graphs of the matrices in $\bar{\mathcal{S}}$ do not have self-arcs at all vertices. We refer the reader to [6] for details.

Proof of Proposition 12. Since $\gamma(S_1), \gamma(S_2), \ldots$ is repeatedly jointly rooted by subsequences of length q, for each $k \geq 1$, the subsequence $\gamma(S_{q(k-1)+1}), \ldots, \gamma(S_{qk})$

is jointly rooted. For $k \geq 1$ define $\bar{S}_k = S_{qk} \cdots S_{q(k-1)+1}$. By Lemma 3, $\gamma(S_{qk} \cdots S_{q(k-1)+1}) = \gamma(S_{qk}) \circ \cdots \circ \gamma(S_{q(k-1)+1})$, $k \geq 1$. Therefore $\gamma(\bar{S}_k)$ is rooted for $k \geq 1$. Thus each such $\bar{S}_k$ is in the closed set $\bar{\mathcal{S}}(q)$.

By Proposition 11, there exists an integer m such that the graph of the product of every set of m matrices from $\bar{\mathcal{S}}(q)$ is strongly rooted. Moreover, since each $\bar{S}_k \in \bar{\mathcal{S}}(q)$, Proposition 11 also implies that $k \to \infty$, and the product $\bar{S}_k \cdots \bar{S}_1$ converges exponentially fast to $\mathbf{1}\lfloor \cdots \bar{S}_k \cdots \bar{S}_1 \rfloor$ at a rate no slower than

$$\bar{\lambda} = \left(\max_{S \in (\bar{\mathcal{S}}(q))^m} |||\lfloor S \rfloor||| \right)^{\frac{1}{m}},$$

where $\bar{\lambda} < 1$.

Now the product $S_j \cdots S_1$ can be written as

$$S_j \cdots S_1 = \widehat{S}(j) \bar{S}_k \cdots \bar{S}_1,$$

where k is the integer quotient of j divided by mq and $\widehat{S}(j)$ is the identity if mq is a factor of j or $\widehat{S}(j) = S_j \cdots S_{(kmq+1)}$ if it is not. But $\widehat{S}(j)$ is a product of at most mq stochastic matrices, and so it is a bounded function of j. It follows that the product $S_j S_{j-1} \cdots S_1$ must converge to $\mathbf{1}\lfloor \cdots S_j \cdot S_1 \rfloor$ exponentially fast at a rate no slower than $\lambda = \bar{\lambda}^{\frac{1}{mq}}$. $\square$

We are now in a position to apply Proposition 12 to leaderless coordination.

THEOREM 3. *Let $\theta(0)$ be fixed. For any trajectory of the system* (3) *along which each graph in the sequence of neighbor graphs $\mathbb{N}(0)$, $\mathbb{N}(1), \ldots$ is repeatedly jointly rooted, there is a constant steady state heading θ_{ss} for which*

$$\lim_{t \to \infty} \theta(t) = \theta_{ss} \mathbf{1}, \tag{34}$$

where the limit is approached exponentially fast.

Proof of Theorem 3. By hypothesis, the sequence of graphs $\gamma(F(0)), \gamma(F(1)), \ldots$ is repeatedly jointly rooted. Thus there is an integer q for which the sequence is repeatedly jointly rooted by subsequences of length q. Since the set of $n \times n$ flocking matrices $\mathcal{F}$ is finite, so is the set of all products of q flocking matrices with rooted graphs, namely $\mathcal{F}(q)$. Therefore $\mathcal{F}(q)$ is closed. Moreover, if $m = (n-1)^2$, every product of m matrices from $\mathcal{F}(q)$ is strongly rooted. It follows from Proposition 12 that the product $F(t) \cdots F(1)F(0)$ converges to $\mathbf{1}\lfloor \cdots F(t) \cdots F(1)F(0) \rfloor$ exponentially fast as $t \to \infty$ at a rate no slower than

$$\lambda = \left(\max_{S \in (\mathcal{F}(q))^m} |||\lfloor S \rfloor||| \right)^{\frac{1}{mq}},$$

where $m = (n-1)^2$, $\lambda < 1$, and $(\mathcal{F}(q))^m$ is the closed set of all products of m matrices from $\mathcal{F}(q)$. But it is clear from (3) that

$$\theta(t) = F(t-1) \cdots F(1)F(0)\theta(0), \quad t \geq 1.$$

Therefore (34) holds with $\theta_{ss} = \lfloor \cdots F_{\sigma(t)} \cdots F_{\sigma(0)} \rfloor \theta(0)$ and the convergence is exponential. $\square$

It is possible to compare Theorem 3 with similar results derived in [19, 22]. To do this it is necessary to introduce a few concepts. By the *union* $\mathbb{G}_1 \cup \mathbb{G}_2$ of two directed graphs $\mathbb{G}_1$ and $\mathbb{G}_2$ with the same vertex set $\mathcal{V}$ we mean the graph whose vertex set is $\mathcal{V}$ and whose arc set is the union of the arc sets of $\mathbb{G}_1$ and $\mathbb{G}_2$. The definition extends in the obvious way to finite sets of directed graphs with the same vertex set. Let us agree to say that a finite set of graphs $\{\mathbb{G}_{p_1},\ \mathbb{G}_{p_2},\dots,\mathbb{G}_{p_k}\}$ with the same vertex set is *collectively rooted* if the *union* of the graphs in the set is a rooted graph. In parallel with the notion of repeatedly jointly rooted, we say that an infinite sequence of graphs $\mathbb{G}_{p_1},\mathbb{G}_{p_2},\dots$ in $\mathcal{G}_{sa}$ is *repeatedly collectively rooted* if there is a positive integer q for which each finite set $\mathbb{G}_{p_{q(k-1)+1}},\dots,\mathbb{G}_{p_{qk}},\ k\geq 1$ is collectively rooted. One of the main contributions of [19] is to prove that the conclusions of Theorem 3 hold if the theorem's hypothesis is replaced by the hypothesis that the sequence of graphs $\mathbb{G}_{\sigma(0)},\mathbb{G}_{\sigma(1)},\dots$ is repeatedly collectively rooted. The two hypotheses prove to be equivalent. The reason this is so can be explained as follows.

Note first that because all graphs in $\mathcal{G}_{sa}$ have self-arcs, each arc (i,j) in the union $\mathbb{G}_2\cup\mathbb{G}_1$ of two graphs $\mathbb{G}_1,\mathbb{G}_2$ in $\mathcal{G}_{sa}$ is an arc in the composition $\mathbb{G}_2\circ\mathbb{G}_1$. While the converse is not true, the definition of composition does imply that for each *arc* (i,j) in the composition $\mathbb{G}_2\circ\mathbb{G}_1$ there is a *path* in the union $\mathbb{G}_2\cup\mathbb{G}_1$ of length at most two between i and j. More generally, simple induction proves that if (i,j) is an arc in the composition of q graphs from $\mathcal{G}_{sa}$, then the union of the same q graphs must contain a path of length at most q from i to j. These observations clearly imply that a sequence of q graphs $\mathbb{G}_{p_1},\ \mathbb{G}_{p_2},\dots,\mathbb{G}_{p_q}$ in $\mathcal{G}_{sa}$ is jointly rooted if and only if the set of graphs $\{\mathbb{G}_{p_1},\ \mathbb{G}_{p_2},\dots,\mathbb{G}_{p_q}\}$ is collectively rooted. It follows that a sequence of graphs in $\mathcal{G}_{sa}$ is repeatedly jointly rooted if and only if the set of graphs in the sequence is collectively jointly rooted.

Although Theorem 3 and the main result of [19] are equivalent, the difference between results based on unions and results based on compositions begins to emerge, when one looks deeper into the convergence question, especially when issues of convergence rate are taken into consideration. For example, if $\pi_u(m,n)$ were the number of m term sequences of graphs in $\mathcal{G}_{sa}$ whose unions are strongly rooted, and $\pi_c(m,n)$ were the number of m-term sequences of graphs in $\mathcal{G}_{sa}$ whose compositions are strongly rooted, then it is easy to see that the ratio $\rho(m,n)=\pi_c(m,n)/\pi_u(m,n)$ would always be greater than 1. In fact, $\rho(2,3)=1.04$ and $\rho(2,4)=1.96$. Moreover, probabilistic experiments suggest that this ratio can be as large as $18,181$ for $m=2$ and $n=50$. One would expect $\rho(m,n)$ to increase not only with increasing n but also with increasing m. One would also expect similar comparisons for neighbor-shared graphs rather than strongly rooted graphs. Interestingly, preliminary experimental results suggest that this is not the case, but more work needs to be done to understand why this is so. Like strongly rooted graphs, neighbor-shared graphs also play a key role in determining convergence rates [7].

3. Symmetric neighbor relations.

It is natural to call a graph in $\mathcal{G}$ *symmetric* if for each pair of vertices i and j for which j is a neighbor of i, i is also a neighbor of j. Note that $\mathbb{G}$ is symmetric if and only if its adjacency matrix is symmetric. It is worth noting that for symmetric graphs, the properties of rooted and rooted at v are both equivalent to the property that the graph is strongly connected. Within the class of symmetric graphs, neighbor-shared graphs and strongly rooted graphs are also strongly connected graphs, but in neither case is the converse true. It is possible to represent a symmetric directed graph $\mathbb{G}$ with an undirected graph $\mathbb{G}^s$ in which each self-arc is replaced with an undirected edge and each pair of directed arcs (i,j) and (j,i) for distinct vertices is replaced with an undirected edge between i and j. Notions of strongly rooted and neighbor-shared extend in the obvious way to undirected graphs. An undirected graph is said to be *connected* if there is an undirected path between each pair of vertices. Thus a strongly connected, directed graph which is symmetric is in essence the same as a connected, undirected graph. Undirected graphs are applicable when the sensing radii r_i of all agents are the same. It was the symmetric version of the flocking problem which Vicsek addressed in [29] and which was analyzed in [17] using undirected graphs.

Let $\mathcal{G}^s$ and $\mathcal{G}^s_{sa}$ denote the subsets of symmetric graphs in $\mathcal{G}$ and $\mathcal{G}_{sa}$, respectively. Simple examples show that neither $\mathcal{G}^s$ nor $\mathcal{G}^s_{sa}$ is closed under composition. In particular, composition of two symmetric directed graphs in $\mathcal{G}$ or $\mathcal{G}_{sa}$ is not typically symmetric. On the other hand, the union is. It is clear that both $\mathcal{G}^s$ and $\mathcal{G}^s_{sa}$ are closed under the union operation. It is worth emphasizing that union and composition are really quite different operations. For example, as we have already seen with Proposition 4, the composition of any $n-1$ strongly connected graphs, symmetric or not, is always complete. On the other hand, the union of $n-1$ strongly connected graphs is not necessarily complete. In terms of undirected graphs, it is simply not true that the union of $n-1$ undirected graphs with vertex set $\mathcal{V}$ is complete, even if each graph in the union has self-loops at each vertex. As noted before, the root cause of the difference between union and composition stems from the fact that the union and composition of two graphs in $\mathcal{G}$ have different arc sets—and in the case of graphs from $\mathcal{G}_{sa}$, the arc set of the union is always contained in the arc set of the composition but not conversely.

In [17] use is made of the notion of a "jointly connected set of graphs." Specifically, a set of undirected graphs with vertex set $\mathcal{V}$ is *jointly connected* if the union of the graphs in the collection is a connected graph. The notion of jointly connected also applies to directed graphs in which case the collection is jointly connected if the union is strongly connected. In what follows we will say that an infinite sequence of graphs $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \dots$ in $\mathcal{G}_{sa}$ is *repeatedly jointly connected* if there is a positive integer m for which each finite sequence $\mathbb{G}_{p_{m(k-1)+1}}, \ \dots, \mathbb{G}_{p_{mk}}, k \geq 1$, is jointly connected. The main result of [17] is, in essence, a corollary to Theorem 3.

COROLLARY 1. *Let $\theta(0)$ be fixed. For any trajectory of the system* (3) *along which each graph in the sequence of symmetric neighbor graphs* $\mathbb{N}(0)$, $\mathbb{N}(1), \dots$ *is repeatedly jointly connected, there is a constant steady state heading θ_{ss} for which*

$$\lim_{t \to \infty} \theta(t) = \theta_{ss}\mathbf{1}, \tag{35}$$

where the limit is approached exponentially fast.

4. Concluding remarks. The main goal of this paper has been to establish a number of basic properties of compositions of directed graphs which are useful in explaining how a consensus is achieved under various conditions in a dynamically changing environment. The paper brings together in one place a number of results scattered throughout the literature and at the same time presents new results concerned with compositions of graphs as well as graphical interpretations of several specially structured stochastic matrices appropriate to nonhomogeneous Markov chains.

In a sequel to this paper [7], we consider a modified version of the Vicsek consensus problem in which integer-valued delays occur in sensing the values of headings which are available to agents. In keeping with our thesis that such problems can be conveniently formulated and solved using graphs and graph operations, we analyze the sensing delay problem from mainly a graph-theoretic point of view using the tools developed in this paper. In [7] we also consider another modified version of the Vicsek problem in which each agent independently updates its heading at times determined by its own clock. We do not assume that the groups' clocks are synchronized together or that the times any one agent updates its heading are evenly spaced. Using graph-theoretic concepts from this paper we show in [7] that for both versions of the problem considered, the conditions under which a consensus is achieved are essentially the same as in the synchronized, delay-free case addressed here.

A number of questions are suggested by this work. For example, it would be interesting to have a complete characterization of those rooted graphs which are of Sarymaskov type. It would also be of interest to have convergence results for more general versions of the asynchronous consensus problem in which heading transitions occur continuously. Extensions of these results to more realistic settings such as the one considered in [26] would also be useful.

REFERENCES

[1] D. Angeli and P. A. Bliman, *Extension of a result by Moreau on stability of leaderless multi-agent systems*, in Proceedings of the 2005 IEEE CDC, 2005, pp. 759–764.
[2] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation*, Prentice–Hall, Englewood Cliffs, NJ, 1989.
[3] V. D. Blondel, J. M. Hendrichx, A. Olshevsky, and J. N. Tsitsiklis, *Convergence in multiagent coordination, consensus, and flocking*, in Proceedings of the 2005 IEEE CDC, 2005, pp. 2996–3000.
[4] M. Cao, A. S. Morse, and B. D. O. Anderson, *Coordination of an asynchronous multi-agent system via averaging*, in Proceedings of the 2005 IFAC Congress, 2005.
[5] M. Cao, A. S. Morse, and B. D. O. Anderson, *Reaching a consensus in the face of measurement delays*, in Proceedings of the 2006 Symposium on Mathematical Theory of Networks and Systems, 2006.
[6] M. Cao, A. S. Morse, and B. D. O. Anderson, *Agreeing asynchronously*, IEEE Trans. Automat. Control, to appear.
[7] M. Cao, A. S. Morse, and B. D. O. Anderson, *Reaching a consensus in a dynamically changing environment: Convergence rates, measurement delays, and asynchronous events*, SIAM J. Control Optim., 47 (2008), pp. 601–623.
[8] M. Cao, D. A. Spielman, and A. S. Morse, *A lower bound on convergence of a distributed network consensus algorithm*, in Proceedings of the 2005 IEEE CDC, 2005, pp. 2356–2361.
[9] S. Chatterjee and E. Seneta, *Towards consensus: Some convergence theorems on repeated averaging*, J. Appl. Probability, 14 (1977), pp. 89–97.
[10] M. H. De Groot, *Reaching a consensus*, J. Amer. Statist. Assoc., 69 (1974), pp. 118–121.
[11] J. L. Doob, *Markov processes—discrete parameter*, in Stochastic Processes, John Wiley and Sons, New York, 1953.
[12] M. J. Fischer, N. A. Lynch, and M. S. Paterson, *Impossibility of distributed consensus with one faulty process*, J. Assoc. Comput. Mach., 32 (1985), pp. 347–382.
[13] C. Godsil and G. Royle, *Algebraic Graph Theory*, Grad. Texts in Math. 207, Springer-Verlag, New York, 2001.
[14] D. J. Hartfiel, *Markov Set-Chains*, Springer-Verlag, Berlin, New York, 1998.
[15] D. J. Hartfiel, *Nonhomogeneous Matrix Products*. World Scientific, Singapore, 2002.
[16] R. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.
[17] A. Jadbabaie, J. Lin, and A. S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Trans. Automat. Control, 48 (2003), pp. 988–1001.
[18] C. F. Martin and W. P. Dayawansa, *On the existence of a Lyapunov function for a family of switching systems*, in Proceedings of the 1996 IEEE CDC, 1996, pp. 1820–1823.
[19] L. Moreau, *Stability of multi-agent systems with time-dependent communication links*, IEEE Trans. Automat. Control, 50 (2005), pp. 169–182.
[20] A. S. Morse, *Logically switched dynamical systems*, in Nonlinear and Optimal Control Theory, Springer-Verlag, Berlin, 2008, pp. 1–84.
[21] M. Pease, R. Shostak, and L. Lamport, *Reaching agreement in the presence of faults*, J. Assoc. Comput. Mach., 27 (1980), pp. 228–234.
[22] W. Ren and R. Beard, *Consensus seeking in multiagent systems under dynamically changing interaction topologies*, IEEE Trans. Automat. Control, 50 (2005), pp. 655–661.

[23] C. Reynolds, *Flocks, birds, and schools: A distributed behavioral model*, Comput. Graphics, 21 (1987), pp. 25–34.
[24] T. A. Sarymsakov, *Inhomogeneous Markov chains*, Theor. Probability Appl., 6 (1961), pp. 178–185.
[25] E. Seneta, *Non-negative Matrices and Markov Chains*, Springer-Verlag, New York, 1981.
[26] H. Tanner, A. Jadbabaie, and G. Pappas, *Flocking in fixed and switching networks*, IEEE Trans. Automat. Control, 52 (2007), pp. 863–868.
[27] J. N. Tsisiklis, *Problems in Decentralized Decision Making and Computation*, Ph.D thesis, MIT, Cambridge, MA, 1984.
[28] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Trans. Automat. Control, 3 (1986), pp. 803–812.
[29] 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.
[30] J. Wolfowitz, *Products of indecomposable, aperiodic, stochastic matrices*, Proc. Amer. Math. Soc., 14 (1963), pp. 733–737.
[31] C. W. Wu, *Synchronization and convergence of linear dynamics in random directed networks*, IEEE Trans. Automat. Control, 51 (2006), pp. 1207–1210.