Reaching a Consensus in a Dynamically Changing Environment

A Graphical Approach*

M. Cao  A. S. Morse  B. D. O. Anderson
Yale University  Yale University  Australia National University and
                       National ICT Australia

April 5, 2006

Abstract

This paper presents new graph-theoretic results appropriate to 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.

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 [2, 3, 4, 5, 6, 7] which successfully explains the heading synchronization phenomenon observed in simulation by Vicsek [8], Reynolds [9] and others more than a decade ago. Vicsek and co-authors 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 preliminary version of this work can be found in [1]. The research of the first two authors was supported by the US Army Research Office, the US National Science Foundation and by a gift from the Xerox Corporation. The research of the third author was supported by National ICT Australia, which is funded by the Australian Governments Department of Communications, Information Technology and the Arts and the Australian Research Council through the Backing Australias Ability initiative and the ICT Centre of Excellence Program.
a circle of pre-specified 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 has recently been given in [2]. The explanation exploits ideas from graph theory [10] and from the theory of non-homogeneous Markov chains [11, 12, 13]. 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 [2] which explain the behavior of Reynolds' full nonlinear “boid” system [7].

Vicsek’s problem is what in computer science is called a “consensus problem” [19] or an “agreement problem.” 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.

The aim of this paper is to establish a number of basic properties of “compositions” of sequences of directed graphs which, as shown in [14], 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 [3] and Beard [4]. We derive a condition for agents to reach a consensus exponentially fast which is slightly different than but equivalent to the condition established in [3]. 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 [3, 4] 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 [14] and elsewhere [15, 16, 17, 18] 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 only been partially resolved or not studied at all.

In Section 2 we reconsider the leaderless coordination problem studied in [2], 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 \( N \) 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 [3, 4]. 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 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 non-homogeneous Markov chains [11, 13]. By exploiting this connection in [14], we are able to derive worst case convergence rate results for several versions of the Vicsek problem. The non-homogeneous Markov chain literature also contains interesting convergence results for a class of stochastic matrices studied by Sarymsakov [20]. 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 either in or on a closed disk of pre-specified radius \( r_i \) centered at agent \( i \)'s current position. In the sequel \( \mathcal{N}_i(t) \) denotes the set of labels of those agents which are neighbors of agent \( i \) at time \( t \). Agent \( i \)'s heading, written \( \theta_i \), evolves in discrete-time in accordance with a model of the form

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

where \( t \) is a discrete-time index taking values in the non-negative 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 \( V = \{1, 2, \ldots n\} \) and “arc set” \( \mathcal{A}(\mathbb{N}(t)) \subset V \times V \) which is defined in such a way 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 \( V \). 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. Toward this end, for each graph \( N \in \mathcal{G}_{sa} \), define the flocking matrix

\[
F = D^{-1} A'
\]

(2)

where \( A' \) is the transpose of the “adjacency matrix” of \( N \) and \( D \) the diagonal matrix whose \( j \)-th diagonal element is the “in-degree” of vertex \( j \) within the graph\(^1\). The function \( N \mapsto F \) is bijective. Then

\[
\theta(t + 1) = F(t)\theta(t), \quad t \in \{0, 1, 2, \ldots \}
\]

(3)

where \( \theta \) is the heading vector \( \theta = [\theta_1 \  \theta_2 \  \ldots \  \theta_n]' \) and \( F(t) \) is the flocking matrix of neighbor graph \( N(t) \) 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 \( N(t) \) 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 \( N(t) \) depends on the agent positions in the plane and assumes instead that \( t \mapsto N(t) \) might be any signal in some suitably defined set of interest.

Our ultimate goal is to show for a large class of signals \( t \mapsto N(t) \) and for any initial set of agent headings, that the headings of all \( n \) agents will converge to the same steady state value \( \theta_{ss} \). Convergence of the \( \theta_t \) to \( \theta_{ss} \) is equivalent to the state vector \( \theta \) converging to a vector of the form \( \theta_{ss} \mathbf{1} \) where \( \mathbf{1} \triangleq [1 \ 1 \ \ldots \ 1]'_{n \times 1} \). Naturally there are situations where convergence to a common heading cannot occur. The most obvious of these is when one agent - say the \( \mu \)-th agent headings, that the headings of all \( \mu \) 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 \( N(t) \) depends on the agent positions in the plane and assumes instead that \( t \mapsto N(t) \) might be any signal in some suitably defined set of interest.

\(^1\)By the adjacency matrix of a directed graph \( G \in \mathcal{G} \) is meant an \( n \times n \) matrix whose \( ij \)-th entry is a 1 if \( (i,j) \) is an arc in \( A(G) \) and 0 if it is not. The in-degree of vertex \( i \) in \( G \) is the number of arcs in \( A(G) \) of the form \( (i,j) \); thus \( j \)'s in-degree is the number of incoming arcs to vertex \( j \).

\(^2\)A directed graph \( G \in \mathcal{G} \) with arc set \( A \) 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 \( j \) is meant a sequence of arcs in \( A \) 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. \( G \) is complete if it has a path of length one \{i.e., an arc\} between each distinct pair of its vertices.
2.3 Strongly Rooted Graphs

In the sequel we will call a vertex $i$ of a directed graph $G$, a root of $G$ if for each other vertex $j$ of $G$, there is a path from $i$ to $j$. Thus $i$ is a root of $G$, if it is the root of a directed spanning tree of $G$. We will say that $G$ is rooted at $i$ if $i$ is in fact a root. Thus $G$ is rooted at $i$ just in case each other vertex of $G$ is reachable from vertex $i$ along a path within the graph. $G$ is strongly rooted at $i$ if each other vertex of $G$ is reachable from vertex $i$ along a path of length $1$. Thus $G$ is strongly rooted at $i$ if $i$ is a neighbor of every other vertex in the graph. By a root graph $G$ is meant 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 $N(0), N(1), \ldots$ is strongly rooted, there is a constant steady state heading $\theta_{ss}$ for which

$$\lim_{t \to \infty} \theta(t) = \theta_{ss}1$$

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 $G_{sa}$. As defined, each flocking matrix $F$ is square and non-negative, where by a non-negative matrix is meant a matrix whose entries are all non-negative. Each $F$ also has the property that its row sums all equal 1 i.e., $F1 = 1$. Matrices with these two properties are called row stochastic [21]. 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 $G_{sa}$ 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 the sequel we write $M \geq N$ whenever $M - N$ is a non-negative matrix. We also write $M > N$ whenever $M - N$ is a positive matrix where by a positive matrix is meant 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 [11, 12, 22]. 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

$$S_jS_{j-1} \cdots S_1$$

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

$$x(j + 1) = S_jx(j)$$

(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 [23] and has been extensively used [24], is to consider the behavior of the scalar-valued non-negative function $V(x) = |x| - |x|$ along solutions to (5) where $x = [x_1 \ x_2 \ \cdots \ x_n]'$ is a non-negative $n$ vector and $|x|$ and $|x|$ are its largest and smallest elements respectively. The key observation is that for any $n \times n$ stochastic matrix $S$, the $i$th entry of $Sx$ satisfies

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

and

$$\sum_{j=1}^{n} s_{ij} x_j \leq \sum_{j=1}^{n} s_{ij} |x| = |x|$$

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

$$[x(1)], [x(2)], \ldots$$  
$$[x(1)], [x(2)], \ldots$$  
$$V(x(1)), V(x(2)), \ldots$$

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

$$\lim_{j \to \infty} |x(j)|, \quad \lim_{j \to \infty} [x(j)], \quad \lim_{j \to \infty} V(x(j))$$

each exist. Note that whenever the limit of $V(x(j))$ is zero, all components of $x(j)$ together with $|x(j)|$ and $[x(j)]$ 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 1 or equivalently $S_j S_{j-1} \cdots S_1$ converges to a constant matrix of the form $1c$. For example, since for any $n \times n$ stochastic matrix $S$, $S1 = 1$, it must be true that span $\{1\}$ is an $S$-invariant subspace for any such $S$. From this and standard existence conditions for solutions to linear algebraic equations, it follows that for any $(n - 1) \times n$ matrix $P$ with kernel spanned by 1, the equations $PS = SP$ has unique solutions $\tilde{S}$, and moreover that

$$\text{spectrum } S = \{1\} \cup \text{spectrum } \tilde{S}$$

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

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

Since $P$ has full row rank and $P1 = 0$, the convergence of a product of the form $S_j S_{j-1} \cdots S_1$ to $1c$ for some constant row vector $c$, 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 going 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
In the sequel we will also be interested in the matrix 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 non-negative matrices. For any non-negative 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 sub-multiplicative. Note in addition that $||x|| = |x|$ for any non-negative $n$ 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 $[\cdot]$ and $[\cdot]$ to the class of all non-negative $n \times m$ matrix $M$, by letting $[M]$ and $[M]$ now denote the $1 \times m$ row vectors whose $j$th entries are the smallest and largest elements respectively, of the $j$th column of $M$. Note that $[M]$ is the largest $1 \times m$ non-negative row vector $c$ for which $M - 1c$ is non-negative and that $[M]$ is the smallest non-negative row vector $c$ for which $1c - M$ is non-negative. Note in addition that for any $n \times n$ stochastic matrix $S$, one can write

$$S = 1[S] + [S] \quad \text{and} \quad S = 1[S] - [S]$$

where $[S]$ and $[S]$ are the non-negative matrices

$$[S] = S - 1[S] \quad \text{and} \quad [S] = 1[S] - S$$

respectively. Moreover the row sums of $[S]$ are all equal to $1 - [S]1$ and the row sums of $[S]$ are all equal to $[S]1 - 1$ so

$$||[S]|| = 1 - [S]1 \quad \text{and} \quad ||[S]|| = [S]1 - 1$$

In the sequel we will also be interested in the matrix

$$[S] = [S] + [S]$$

This matrix satisfies

$$[S] = 1([S] - [S])$$

because of (7).

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

$$[\cdots S_j \cdots S_2 S_1] = \lim_{j \to \infty} [S_j \cdots S_2 S_1]$$

From the preceding discussion it is clear that for $i \in \{1, 2, \ldots, n\}$, the limit $[\cdots S_j \cdots S_1]e_i$ exists, where $e_i$ is the $i$th unit $n$-vector. Thus the limit $[\cdots S_j \cdots S_1]$ 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 the sequel 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 $1c$ where $c1 = 1$ because stochastic matrices are closed under multiplication.

In the sequel we will say that a matrix product $S_jS_{j-1} \cdots S_1$ converges to $1[\cdots S_j \cdots S_1]$ exponentially fast at a rate no slower than $\lambda$ if there are non-negative constants $b$ and $\lambda$ with $\lambda < 1$, such that

$$||(S_j \cdots S_1) - 1[\cdots S_j \cdots S_1]|| \leq b\lambda^j, \quad j \geq 1$$

(13)

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

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

$$||S_j \cdots S_1|| \leq \bar{b}\lambda^j, \quad j \geq 0$$

(14)

Then the matrix product $S_j \cdots S_2S_1$ converges to $1[\cdots S_j \cdots S_1]$ exponentially fast at a rate no slower than $\lambda$.

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 $S_2$,

$$|S_2S_1| - |S_1| \leq [S_2][S_1] \leq [S_2][S_1]$$

(15)

(16)

**Proof of Lemma 1:** Since $S_2S_1 = S_2([S_1] + |S_1|) = 1[S_1] + S_2|S_1|$ and $S_2 = 1[S_2] - |S_2|$, it must be true that $S_2S_1 = 1([S_1] + |S_2||S_1|) - |S_2||S_1|$. Thus $1([S_1] + |S_2||S_1|) - |S_2||S_1|$ is non-negative. But $|S_2S_1|$ is the smallest non-negative row vector $c$ for which $1c - S_2S_1$ is non-negative. Therefore

$$|S_2S_1| \leq |S_1| + |S_2||S_1|$$

(17)

Moreover $|S_2S_1| \leq |S_2S_1|$ because of the definitions of $[-]$ and $[\cdot]$. This and (17) imply $|S_2S_1| \leq |S_1| + |S_2||S_1|$ and thus that (15) is true.

Since $S_2S_1 = S_2([S_1] + |S_1|) = 1[S_1] + S_2|S_1|$ and $S_2 = |S_2| + |S_2|$, it must be true that $S_2S_1 = 1([S_1] + |S_2||S_1|) + |S_2||S_1|$. Thus $S_2S_1 - 1([S_1] + |S_2||S_1|)$ is non-negative. But $|S_2S_1|$ is the largest non-negative row vector $c$ for which $S_2S_1 - 1c$ is non-negative so

$$S_2S_1 \leq 1|S_2S_1| + |S_2||S_1|$$

(18)

Now it is also true that $S_2S_1 = 1|S_2S_1| + |S_2S_1|$. From this and (18) it follows that (16) is true.

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

$$[X_{j+1}] - [X_j] \leq [S_{j+1}][X_j], \quad j \geq 1$$

8
By hypothesis, $\|X_j\| \leq \bar{b}\lambda^j$, $j \geq 1$. Moreover $\|S_{j+1}\| \leq n$ because all entries in $S_{j+1}$ are bounded above by 1. Therefore

$$\|X_{j+1} - X_j\| \leq n\bar{b}\lambda^j, \quad j \geq 1$$  \hspace{1cm} (19)

Clearly

$$X_{j+i} - X_j = \sum_{k=1}^{i} (X_{i+j+1-k} - X_{i+j-k}), \quad i, j \geq 1$$

Thus, by the triangle inequality

$$\|X_{j+i} - X_j\| \leq \sum_{k=1}^{i} \|X_{i+j+1-k} - X_{i+j-k}\|, \quad i, j \geq 1$$

This and (19) imply that

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

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 $\lambda < 1$ so

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

Therefore

$$\|X_{i+j} - X_j\| \leq n\bar{b} \frac{\lambda^j}{(1-\lambda)}, \quad i, j \geq 1$$  \hspace{1cm} (20)

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

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

In view of (20)

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

Since

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

it must be true that

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

But $\|1(|X_j - c|)\| = \|X_j - c\|$ and $X_j = S_j \cdots S_1$. Therefore

$$\|1(|S_j \cdots S_1 - c|)\| \leq n\bar{b} \frac{\lambda^j}{(1-\lambda)}, \quad j \geq 1$$  \hspace{1cm} (21)
In view of (7)
\[ S_j \cdots S_1 = 1[S_j \cdots S_1] + [S_j \cdots S_1], \quad j \geq 1 \]

Therefore
\[ ||(S_j \cdots S_1) - 1c|| = ||1[S_j \cdots S_1] + [S_j \cdots S_1] - 1c|| \leq ||1[S_j \cdots S_1] - 1c|| + ||[S_j \cdots S_1]||, \quad j \geq 1 \]

From this, (14) and (21) it follows that
\[ ||S_j \cdots S_1 - 1c|| \leq \bar{b} \left( 1 + \frac{n}{(1-\lambda)} \right) \lambda^j, \quad j \geq 1 \]

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

### 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
\[ [S_j S_{j-1} \cdots S_2 S_1] \leq [S_j] [S_{j-1}] \cdots [S_1] \]  

(22)

It is therefore clear that condition (14) of Proposition 1 will hold with \( \bar{b} = 1 \) if
\[ ||[S_j] \cdots [S_1]|| \leq \lambda^j \]  

(23)

for some nonnegative number \( \lambda < 1 \). Because \( ||\cdot|| \) is sub-multiplicative, this means that a product of stochastic matrices \( S_j \cdots S_1 \) will converge to a limit of the form \( 1c \) 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 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 \( 1[S] \neq 0 \) since \( [S] = S - 1[S] \). Now \( 1[S] \neq 0 \) if and only if \( S \) has at least one non-zero column since the indices of the non-zero columns of \( S \) are the same as the indices of the non-zero columns of \( [S] \). Thus \( ||[S]|| < 1 \) if and only if \( S \) has at least one non-zero column. For our purposes it proves to be especially useful to re-state 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 that graph \( \mathcal{G} \in \mathcal{G} \) whose adjacency matrix is the transpose of the matrix obtained by replacing all of \( S \)'s non-zero entries with 1s. The graph theoretic condition is as follows.

**Lemma 2** A stochastic matrix \( S \) has a strongly rooted graph \( \gamma(S) \) if and only if
\[ ||[S]|| < 1 \]  

(24)
Proof: Let $A$ be the adjacency matrix of $\gamma(S)$. Since the positions of the non-zero entries of $S$ and $A'$ are the same, $S$’s $i$th column will be positive if and only if $A$’s $i$th row is positive. Thus (23) will hold just in case $A$ has a non-zero row. But strongly rooted graphs in $\mathcal{G}$ are precisely those graphs whose adjacency matrices have at least one non-zero 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 $S_{sr}$ be any closed set of stochastic matrices which are all of the same size and whose graphs $\gamma(S)$, $S \in S_{sr}$, are all strongly rooted. Then as $j \to \infty$, any product $S_j \cdots S_1$ of matrices from $S_{sr}$ converges exponentially fast to $1[\cdots S_j \cdots S_1]$ at rate no slower than

$$\lambda = \max_{S \in S_{sr}} |||S|||$$

where $\lambda$ is a non-negative constant satisfying $\lambda < 1$.

**Proof of Proposition 2:** In view of Lemma 2, $|||S||| < 1$, $S \in S_{sr}$. Because $S_{sr}$ is closed and bounded and $|||\cdot|||$ is continuous, $\lambda < 1$. Clearly $|||S_i||| \leq \lambda$, $i \geq 1$ so (23) must hold for any sequence of matrices $S_1, S_2, \ldots$ from $S_{sr}$. 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 $1[\cdots S_j \cdots S_1]$ exponentially fast at a rate no slower than $\lambda$. ■

**Proof of Theorem 1:** Let $\mathcal{F}_{sr}$ denote the set of flocking matrices with strongly rooted graphs. Since $S_{sa}$ 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 $1[\cdots F(t) \cdots F(0)]$ 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 $\theta_{ss} = [\cdots F(t) \cdots F(0)]\theta(0)$ 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 $\lambda$ 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 $k$th – 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 $\theta_k$ 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 $|||F||| = 1 - [F]1$, 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 $F$ 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

$$\max_{F \in \mathcal{F}_{sr}} |||F||| = 1 - \frac{1}{n} \quad (25)$$
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 of a sequence $G_1, G_2, \ldots, G_k$ with the same vertex set $V$ would be that graph with vertex set $V$ 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 $G_p \in \mathcal{G}$ with a directed graph $G_q \in \mathcal{G}$, written $G_q \circ G_p$, is meant the directed graph with 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 $G_p$ and $(k, j)$ is an arc of $G_q$. Thus $(i, j)$ is an arc in $G_q \circ G_p$ if and only if $i$ has an observer in $G_p$ which is also a neighbor of $j$ in $G_q$. 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 $G_1, G_2, \ldots, G_k$.

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 $G_p$ and $G_q$ are arcs of $G_q \circ G_p$. The definition also implies in this case that if $G_p$ has a directed path from $i$ to $k$ and $G_q$ has a directed path from $k$ to $j$, then $G_q \circ G_p$ 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 $G_1, G_2, \ldots, G_k$ 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 $G_p \in \mathcal{G}$ and $G_q \in \mathcal{G}$ respectively. Then the adjacency matrix of the composition $G_q \circ G_p$ must be the matrix obtained by replacing all non-zero elements in $A_pA_q$ with ones. This is because the $ij$th entry of $A_pA_q$, namely

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

will be non-zero just in case there is at least one value of $k$ for which both $a_{ik}(p)$ and $a_{kj}(q)$ are non-zero. This of course is exactly the condition for the $ij$th element of the adjacency matrix of the composition $G_q \circ G_p$ to be non-zero. Note that if $S_1$ and $S_2$ are $n \times n$ stochastic matrices for which $\gamma(S_1) = G_p$ and $\gamma(S_2) = G_q$, then the matrix which results by replacing by ones, all non-zero entries in the stochastic matrix $S_2S_1$, must be the transpose of the adjacency matrix of $G_q \circ G_p$. 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 of 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 $G_{p_1}, G_{p_2}, \ldots, G_{p_m}$ be a finite sequence of rooted graphs in $G_{sa}$.

1. If $m \geq (n-1)^2$, then $G_{p_m} \circ G_{p_{m-1}} \circ \cdots \circ G_{p_1}$ is strongly rooted.

2. If $G_{p_1}, G_{p_2}, \ldots, G_{p_m}$ are all rooted at $v$ and $m \geq n - 1$, then $G_{p_m} \circ G_{p_{m-1}} \circ \cdots \circ G_{p_1}$ 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 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 in a moment.

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

To prove Proposition 3 we will need some more ideas. We say that a vertex $v \in V$ is an observer of a subset $S \subset V$ in a graph $G \in \mathcal{G}$, if $v$ is an observer of at least one vertex in $S$. By the observer function of a graph $G \in \mathcal{G}$, written $\alpha(G, \cdot)$ we mean the function $\alpha(G, \cdot) : 2^V \to 2^V$ which assigns to each subset $S \subset V$, the subset of vertices in $V$ which are observers of $S$ in $G$. Thus $j \in \alpha(G, i)$ just in case $(i, j) \in A(G)$. Note that if $G_p \in \mathcal{G}$ and $G_q$ in $G_{sa}$, then
\[
\alpha(G_p, S) \subset \alpha(G_q \circ G_p, S), \quad S \in 2^V
\]
(26)
because $G_q \in G_{sa}$ implies that the arcs in $G_p$ are all arcs in $G_q \circ G_p$. Observer functions have the following important and easily proved property.
Lemma 4: For all $G_p, G_q \in \mathcal{G}$ and any non-empty subset $S \subset V,$

$$\alpha(G_q, \alpha(G_p, S)) = \alpha(G_q \circ G_p, S)$$ (27)

Proof: Suppose first that $i \in \alpha(G_q, \alpha(G_p, S)).$ Then $(j, i)$ is an arc in $G_q$ for some $j \in \alpha(G_p, S).$ Hence $(k, j)$ is an arc in $G_p$ for some $k \in S.$ In view of the definition of composition, $(k, i)$ is an arc in $G_q \circ G_p$ so $i \in \alpha(G_q \circ G_p, S).$ Since this holds for all $i \in V,$ $\alpha(G_q, \alpha(G_p, S)) \subset \alpha(G_q \circ G_p, S).$

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

To proceed, let us note that each subset $S \subset V$ induces a unique subgraph of $G$ with vertex set $S$ and arc set $A$ consisting of those arcs $(i, j)$ of $G$ for which both $i$ and $j$ are vertices of $S.$ This together with the natural partial ordering of $V$ by inclusion provides a corresponding partial ordering of $G.$ Thus if $S_1$ and $S_2$ are subsets of $V$ and $S_1 \subset S_2,$ then $G_{S_1} \subset G_{S_2}$ where for $i \in \{1, 2\},$ $G_i$ is the subgraph of $G$ induced by $S_i.$ For any $v \in V,$ there is a unique largest subgraph rooted at $v,$ namely the graph induced by the vertex set $V(v) = \{v\} \cup \alpha(G, v) \cup \cdots \cup \alpha^{n-1}(G, v)$ where $\alpha(G, \cdot)$ denotes the composition of $\alpha(G, \cdot)$ with itself $i$ times. We call this graph, the rooted graph generated by $v.$ It is clear that $V(v)$ is the smallest $\alpha(G, \cdot)$ - invariant subset of $V$ which contains $v.$

The proof of Propositions 3 depends on the following lemma.

Lemma 5: Let $G_p$ and $G_q$ be graphs in $G_{sa}.$ If $G_q$ is rooted at $v$ and $\alpha(G_p, v)$ is a strictly proper subset of $V,$ then $\alpha(G_p, v)$ is also a strictly proper subset of $\alpha(G_q \circ G_p, v).$

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

Proof of Proposition 3:

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

$$\alpha(G_{p_k} \circ G_{p_{k-1}} \circ \cdots \circ G_{p_1}, v) = V$$ (28)

It will now be shown that such an integer exists.

If $\alpha(G_{p_1}, v) = V,$ set $k = 1$ in which case (28) clearly holds. If $\alpha(G_{p_1}, v) \neq V,$ then let $i > 1$ be the greatest positive integer not exceeding $n - 1$ for which $\alpha(G_{p_{i-1}} \circ \cdots \circ G_{p_1}, v)$ is a strictly proper subset of $V.$ 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, $\alpha(G_{p_{i-1}} \circ \cdots \circ G_{p_1}, v)$ must be a strictly proper subset of $V$ for $j \in \{2, 3, \ldots, n - 1\};$ By Lemma 5, $\alpha(G_{p_{i-1}} \circ \cdots \circ G_{p_1}, v)$ is also a strictly proper
subset of \( \alpha(\mathcal{G}_{p_j} \circ \cdots \circ \mathcal{G}_{p_1}, v) \) for \( j \in \{2, 3, \ldots, n-1\} \). In view of this and (26), each containment in the ascending chain
\[
\alpha(\mathcal{G}_{p_1}, v) \subset \alpha(\mathcal{G}_{p_2} \circ \mathcal{G}_{p_1}, v) \subset \cdots \subset \alpha(\mathcal{G}_{p_{n-1}} \circ \cdots \circ \mathcal{G}_{p_1}, v)
\]
is strict. Since \( \alpha(\mathcal{G}_{p_1}, v) \) 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 \( \mathcal{G}_{q_1}, \mathcal{G}_{q_2}, \ldots, \mathcal{G}_{q_{n-1}} \) all have a common root. By assertion 2, \( \mathcal{A}(\mathcal{G}_{q_{n-1}} \circ \cdots \circ \mathcal{G}_{q_1}) \) must be strongly rooted. But \( \mathcal{A}(\mathcal{G}_{q_{n-1}} \circ \cdots \circ \mathcal{G}_{q_1}) \subset \mathcal{A}(\mathcal{G}_{p_m} \circ \mathcal{G}_{p_{m-1}} \circ \cdots \circ \mathcal{G}_{p_1}) \) because \( \mathcal{G}_{q_1}, \mathcal{G}_{q_2}, \ldots, \mathcal{G}_{q_{n-1}} \) is a subsequence of \( \mathcal{G}_{p_1}, \mathcal{G}_{p_2}, \ldots, \mathcal{G}_{p_m} \) and all graphs in the sequence \( \mathcal{G}_{p_1}, \mathcal{G}_{p_2}, \ldots, \mathcal{G}_{p_m} \) have self-arcs. Therefore \( \mathcal{G}_{p_m} \circ \mathcal{G}_{p_{m-1}} \circ \cdots \circ \mathcal{G}_{p_1} \) must be strongly rooted.

Proposition 3 implies that every sufficiently long composition of graphs from a given subset \( \mathcal{G} \subset \mathcal{G}_\mathcal{sa} \) will be strongly rooted if each graph in \( \mathcal{G} \) 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 \( \mathcal{G} \in \mathcal{G} \), which is not rooted but for which \( \mathcal{G}^m \) is strongly rooted for \( m \) sufficiently large where \( \mathcal{G}^m \) is the \( m \)-fold composition of \( \mathcal{G} \) with itself. Thus \( \alpha(\mathcal{G}^m, v) = \mathcal{V} \) where \( v \) is a root of \( \mathcal{G}^m \). But via repeated application of (27), \( \alpha(\mathcal{G}^m, v) = \alpha^m(\mathcal{G}, v) \) where \( \alpha^m(\mathcal{G}, \cdot) \) is the \( m \)-fold composition of \( \alpha(\mathcal{G}, \cdot) \) with itself. Thus \( \alpha^m(\mathcal{G}, v) = \mathcal{V} \). But this can only occur if \( \mathcal{G} \) is rooted at \( v \) because \( \alpha^m(\mathcal{G}, v) \) is the set of vertices reachable from \( v \) along paths of length \( m \). Since this is a contradiction, \( \mathcal{G} \) cannot be rooted. We summarize.

**Proposition 5** Every possible sufficiently long composition of graphs from a given subset \( \mathcal{G} \subset \mathcal{G}_\mathcal{sa} \) is strongly rooted, if and only if every graph in \( \mathcal{G} \) 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 [11] 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 \( \mathcal{G} \in \mathcal{G} \), if \( v \) is a neighbor of at least one vertex in \( \mathcal{S} \). By a **Sarymsakov Graph** is meant a graph \( \mathcal{G} \in \mathcal{G} \) with the property that for each pair of non-empty 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 non-negative matrices studied by Sarymsakov in [20]. In the sequel 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 \( \mathcal{G}_\mathcal{sa} \), are Sarymsakov graphs.

It is possible to characterize Sarymsakov graph a little more concisely using the following concept. By the **neighbor function** of a graph \( \mathcal{G} \in \mathcal{G} \), written \( \beta(\mathcal{G}, \cdot) \), we mean the function \( \beta(\mathcal{G}, \cdot) : 2^\mathcal{V} \to 2^\mathcal{V} \) which assigns to each subset \( \mathcal{S} \subset \mathcal{V} \), the subset of vertices in \( \mathcal{V} \) which are neighbors of \( \mathcal{S} \) in \( \mathcal{G} \). Thus in terms of \( \beta \), a Sarymsakov graph is a graph \( \mathcal{G} \in \mathcal{G} \) with the property that for each pair of
non-empty subsets $S_1$ and $S_2$ in $\mathcal{V}$ which have no neighbors in common, $S_1 \cup S_2$ contains less vertices than does the set $\beta(G, S_1 \cup S_2)$. Note that if $G \in G_{sa}$, the requirement that $S_1 \cup S_2$ contain less vertices than $\beta(G, S_1 \cup S_2)$ simplifies to the equivalent requirement that $S_1 \cup S_2$ be a strictly proper subset of $\beta(G, S_1 \cup S_2)$. This is because every vertex in $G$ is a neighbor of itself if $G \in G_{sa}$.

**Proposition 6**

1. Each Sarymsakov graph in $G$ is rooted.

2. Each rooted graph in $G_{sa}$ is a Sarymsakov graph.

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

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

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

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

$$\beta(G, T \cup S) \subseteq T \cup S \quad (29)$$

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

**Lemma 7** Let $G \in G$ be rooted at $r$. Each non-empty subset $S \subseteq \mathcal{V}$ not containing $r$ is a strictly proper subset of $S \cup \beta(G, S)$.

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

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

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

**Proof of Proposition 6:** Write \(\beta(\cdot)\) for \(\beta(G, \cdot)\). To prove the assertion 1, pick \(G \in \mathcal{G}\). Let \(G^*\) be any maximal rooted subgraph of \(G\) and write \(R\) for its root set; in view of Lemma 8, \(\beta(R) \subset R\).

Pick any \(v \in V\). Then by Lemma 6, for some positive integer \(m \leq n\), \(\beta^m(v) \cap R\) is non-empty. Pick \(z \in \beta^m(v) \cap R\). Then there is a path from \(z\) to \(v\) and \(z\) is a root of \(G^*\). But \(G^*\) is maximal so \(v\) must be a vertex of \(G^*\). Therefore every vertex of \(G\) is a vertex of \(G^*\) which implies that \(G\) is rooted.

To prove assertion 2, let \(G \in \mathcal{G}_{sa}\) be rooted at \(r\). Pick any two non-empty subsets \(S_1, S_2\) of \(V\) which have no neighbors in common. If \(r \not\in S_1 \cup S_2\), then \(S_1 \cup S_2\) must be a strictly proper subset of \(S_1 \cup S_2 \cup \beta(S_1 \cup S_2)\) because of Lemma 7.

Suppose next that \(r \in S_1 \cup S_2\). Since \(G \in \mathcal{G}_{sa}\), \(S_1 \subset \beta(S_i), i \in \{1, 2\}\). Thus \(S_1\) and \(S_2\) must be disjoint because \(\beta(S_1)\) and \(\beta(S_2)\) are. Therefore \(r\) must be in either \(S_1\) or \(S_2\) but not both. Suppose that \(r \not\in S_1\). Then \(S_1\) must be a strictly proper subset of \(\beta(S_1)\) because of Lemma 7. Since \(\beta(S_1)\) and \(\beta(S_2)\) are disjoint, \(S_1 \cup S_2\) must be a strictly proper subset of \(\beta(S_1 \cup S_2)\). By the same reasoning, \(S_1 \cup S_2\) must be a strictly proper subset of \(\beta(S_1 \cup S_2)\) if \(r \not\in S_2\). Thus in conclusion \(S_1 \cup S_2\) must be a strictly proper subset of \(\beta(S_1 \cup S_2)\) whether \(r\) is in \(S_1 \cup S_2\) or not. Since this conclusion holds for all such \(S_1\) and \(S_2\) and \(G \in \mathcal{G}_{sa}\), \(G\) must be a Sarymsakov graph. ■

### 2.5 Neighbor-Shared Graphs

There is a different assumption which one can make about a sequence of graphs from \(G\) which also insures that the sequence’s composition is strongly rooted. For this we need the concept of a “neighbor-shared graph.” Let us call \(G \in \mathcal{G}\) neighbor shared if each set of 2 distinct vertices share a common neighbor. Suppose that \(G\) is neighbor shared. Then each pair of vertices is clearly reachable from a single vertex. Similarly each three vertices are reachable from paths starting at one of two vertices. Continuing this reasoning it is clear that each of the graph’s \(n\) vertices is reachable from paths starting at vertices in some set \(V_{n-1}\) of \(n - 1\) vertices. By the same reasoning, each vertex in \(V_{n-1}\) is reachable from paths starting at vertices in some set \(V_{n-2}\) of \(n - 2\) vertices. Thus each of the graph’s \(n\) vertices is reachable from paths starting at vertices in a set of \(n - 2\) vertices, namely the set \(V_{n-2}\). Continuing this argument we eventually arrive at the conclusion that each of the graph’s \(n\) vertices is reachable from paths starting at a single vertex, namely the one vertex in the set \(V_1\). We have proved the following.

**Proposition 7** Each neighbor-shared graph in \(G\) 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 $G_{sa}$ need not be neighbor shared, it turns out that the composition of any $n - 1$ rooted graphs in $G_{sa}$ is.

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

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

$$\alpha(G', S) = \beta(G, S), \quad S \in 2^V \quad (30)$$

**Lemma 9** For all $G_p, G_q \in \mathcal{G}$ and any non-empty subset $S \subset V$,

$$\beta(G_q, \alpha(G_p, S)) = \beta(G_p \circ G_q, S). \quad (31)$$

**Proof of Lemma 9:** In view of (27), $\alpha(G_p', \alpha(G_q', S)) = \alpha(G_p' \circ G_q', S)$. But $G_p' \circ G_q' = (G_q \circ G_p)'$, so $\alpha(G_p', \alpha(G_q', S)) = \alpha((G_q \circ G_p)', S)$. Therefore $\beta(G_p, \beta(G_q, S)) = \beta(G_q \circ G_p, S)$ because of (30).

**Lemma 10** Let $G_p$ and $G_q$ be rooted graphs in $G_{sa}$. If $u$ and $v$ are distinct vertices in $V$ for which

$$\beta(G_q, \{u, v\}) = \beta(G_q \circ G_p, \{u, v\}) \quad (32)$$

then $u$ and $v$ have a common neighbor in $G_q \circ G_p$.

**Proof:** $\beta(G_q, u)$ and $\beta(G_q, v)$ are non-empty because $u$ and $v$ are neighbors of themselves. Suppose $u$ and $v$ do not have a common neighbor in $G_q \circ G_p$. Then $\beta(G_q \circ G_p, u)$ and $\beta(G_q \circ G_p, v)$ are disjoint. But $\beta(G_q \circ G_p, u) = \beta(G_p, \beta(G_q, u))$ and $\beta(G_q \circ G_p, v) = \beta(G_p, \beta(G_q, v))$ because of (31). Therefore $\beta(G_p, \beta(G_q, u))$ and $\beta(G_p, \beta(G_q, v))$ are disjoint. But $G_p$ is rooted and thus a Sarymsakov graph because of Proposition 6. Thus $\beta(G_q, \{u, v\})$ is a strictly proper subset of $\beta(G_q, \{u, v\}) \cup \beta(G_p, \beta(G_q, \{u, v\})$. But $\beta(G_q, \{u, v\}) \subset \beta(G_p, \beta(G_q, \{u, v\})$ because all vertices in $G_q$ are neighbors of themselves and $\beta(G_p, \beta(G_q, \{u, v\}) = \beta(G_q \circ G_p, \{u, v\})$ because of (31). Therefore $\beta(G_q, \{u, v\})$ is a strictly proper subset of $\beta(G_q \circ G_p, \{u, v\})$. This contradicts (32) so $u$ and $v$ have a common neighbor in $G_q \circ G_p$.

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

$$\beta(G_{n-1}, \{u, v\}) \subset \beta(G_{n-1} \circ G_{n-2}, \{u, v\}) \subset \cdots \subset \beta(G_{n-1} \circ \cdots \circ G_{p_2} \circ G_{p_1}, \{u, v\})$$

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

$$\beta(G_{p_{n-1}} \circ \cdots \circ G_{p_{i+1}}, \{u, v\}) = \beta(G_{p_{n-1}} \circ \cdots \circ G_{p_{n-i}}, \{u, v\})$$
This and Lemma 10 imply that \( u \) and \( v \) have a common neighbor in \( G_{p_{n-1}} \circ \cdots \circ G_{p_{n-1}} \) and thus in \( G_{p_{n-1}} \circ \cdots \circ G_{p_2} \circ G_{p_1} \). Since this is true for all distinct \( u \) and \( v \), \( G_{p_{n-1}} \circ \cdots \circ G_{p_2} \circ G_{p_1} \) is a neighbor shared graph. \( \blacksquare \)

If we restrict attention to those rooted graphs in \( G_{sa} \) which are strongly connected, we can obtain a neighbor-shared graph by composing a smaller number of rooted graphs than that 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 \( G_{sa} \) is neighbor shared.

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

\[
\{v\} \subset \beta(G_{p_k}, \{v\}) \subset \beta(G_{p_k} \circ G_{p_{k-1}}, \{v\}) \subset \cdots \subset \beta(G_{p_k} \circ \cdots \circ G_{p_2} \circ G_{p_1}, \{v\})
\]

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

Set \( k = q \) and let \( v_1 \) and \( v_2 \) be any pair of distinct vertices in \( V \). Then there must be at least \( q + 1 \) vertices in \( \beta(G_{p_q} \circ \cdots \circ G_{p_{q}}, \{v_1\}) \) and \( q + 1 \) vertices in \( \beta(G_{p_q} \circ \cdots \circ G_{p_{q}}, \{v_2\}) \). But \( 2(q+1) > n \) because of the definition of \( q \), so \( \beta(G_{p_q} \circ \cdots \circ G_{p_{q}}, \{v_1\}) \) and \( \beta(G_{p_q} \circ \cdots \circ G_{p_{q}}, \{v_2\}) \) must have at least one vertex in common. Since this is true for each pair of distinct vertices \( v_1, v_2 \in V \), \( G_{p_q} \circ \cdots \circ G_2 \circ G_1 \) must be neighbor-shared. \( \blacksquare \)

Lemma 7 and Proposition 3 imply that any composition of \( (n-1)^2 \) neighbor-shared graphs in \( G_{sa} \) 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 \( G \) and not necessarily in \( G_{sa} \).

**Proposition 10** The composition of any set of \( m \geq n-1 \) neighbor-shared graphs in \( G \) 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 \( G \in G \) is \( k \)-neighbor shared if each set of \( k \) distinct vertices share a common neighbor. Thus a neighbor-shared graph and a 2 neighbor shared graph are one and the same. Clearly a \( n \) neighbor shared graph is strongly rooted at the common neighbor of all \( n \) vertices.

**Lemma 11** If \( G_p \in G \) is a neighbor-shared graph and \( G_q \in G \) is a \( k \) neighbor shared graph with \( k < n \), then \( G_q \circ G_p \) is a \( (k+1) \) neighbor shared graph.

**Proof:** Let \( v_1, v_2, \ldots, v_{k+1} \) be any distinct vertices in \( V \). Since \( G_q \) is a \( k \) neighbor shared graph, the vertices \( v_1, v_2, \ldots, v_k \) share a common neighbor \( u_1 \) in \( G_q \) and the vertices \( v_2, v_3, \ldots, v_{k+1} \) share a
common neighbor \( u_2 \) in \( G_q \) as well. Moreover, since \( G_p \) is a neighbor shared graph, \( u_1 \) and \( u_2 \) share a common neighbor \( w \) in \( G_p \). It follows from the definition of composition that \( v_1, v_2, \ldots, v_k \) have \( w \) as a neighbor in \( G_q \circ G_p \) as do \( v_2, v_3, \ldots, v_{k+1} \). Therefore \( v_1, v_2, \ldots, v_{k+1} \) have \( w \) as a neighbor in \( G_q \circ G_p \). Since this must be true for any set of \( k + 1 \) vertices in \( G_q \circ G_p \), \( G_q \circ G_p \) must be a \( k + 1 \) neighbor shared graph as claimed. □

**Proof of Proposition 10:** The preceding lemma implies that the composition of any two 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. □

### 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 non-zero. Let \( S \) denote the set of all \( n \times n \) row stochastic matrices whose diagonal elements are all positive. \( S \) is closed under multiplication because the class of all \( n \times n \) stochastic matrices is closed under multiplication and because the class of \( n \times n \) non-negative 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 \( \mathbb{N}(0), \mathbb{N}(1), \ldots \) is rooted, there is a constant steady state heading \( \theta_{ss} \) for which

\[
\lim_{t \to \infty} \theta(t) = \theta_{ss} \mathbf{1}
\]

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 [16, 14]. 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 \( G_{sa} \) is strongly rooted \{cf. Proposition 3\}.

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

\[
\lambda = \left( \max_{S \in S_r^m} ||S|| \right)^{1/m}
\]

where \( \lambda < 1 \).

**Proof of Proposition 11:** By assumption, each graph \( \gamma(S), S \in S_r \), is in \( G_{sa} \) and is rooted. In view of Proposition 3, \( \gamma(S_q) \circ \cdots \gamma(S_1) \) is strongly rooted for every list of \( q \) matrices \( \{S_1, S_2, \ldots, S_q\} \)
from \( S_r \) 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 \( S_i \in S_r \). 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 \( S_r \) can be written as \( S_j \cdots S_1 = \tilde{S}(j)\tilde{S}_k \cdots \tilde{S}_1 \) where \( \tilde{S}_i = S_i \circ \cdots \circ S_{i-1} \circ S_{(i-1)m+1} \), \( 1 \leq i \leq k \), is a product in \( S_r^m \), \( \tilde{S}(j) = S_j \circ \cdots \circ S_{(km+1)} \), and \( k \) is the integer quotient of \( j \) divided by \( m \). In view of Proposition 2, \( S_k \cdots S_1 \) must converge to \( 1[\cdots \tilde{S}_k \cdots \tilde{S}_1] \) exponentially fast as \( k \to \infty \) at a rate no slower than \( \lambda \), where

\[
\lambda = \max_{S \in \mathcal{S}_r^m} \| [S] \|
\]

But \( \tilde{S}(j) \) is a product at most \( m \) stochastic matrices, so it is a bounded function of \( j \). It follows that the product \( S_j S_{j-1} \cdots S_1 \) must converge to \( 1[\cdots S_j \cdots S_1] \) exponentially fast at a rate no slower than \( \lambda = \lambda \frac{1}{m} \). 

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 \( 1[\cdots F(t) \cdots F(0)] \) 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 \( \theta(t) = F(t-1) \cdots F(1)F(0)\theta(0), \quad t \geq 1 \). Therefore (33) holds with \( \theta_{sa} = [\cdots F(t) \cdots F(0)]\theta(0) \) and the convergence is exponential. 

The proof of Theorem 2 also applies to the case when all of the \( N(t), t \geq 0 \) 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 \( N(t), t \geq 0 \) 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 still further 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 \( G_{p_1}, G_{p_2}, \ldots, G_{p_k} \) in \( \mathcal{G} \) is *jointly rooted* if the composition \( G_{p_k} \circ G_{p_{k-1}} \circ \cdots \circ G_{p_1} \) is rooted.

Note that since the arc sets of any graphs \( G_{p}, G_{q} \in \mathcal{G}_{sa} \) are contained in the arc set of any composed graph \( G_q \circ G \circ G_p, G \in \mathcal{G}_{sa} \), it must be true that if \( G_{p_1}, G_{p_2}, \ldots, G_{p_k} \) is a jointly rooted sequence in \( \mathcal{G}_{sa} \), then so is \( G_{q}, G_{p_1}, G_{p_2}, \ldots, G_{p_k}, G_p \). 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 $G_{p_1}, G_{p_2}, \ldots, G_{p_k}$ from $\mathcal{G}$ is \textit{jointly neighbor-shared} if the composition $G_{p_k} \circ G_{p_{k-1}} \circ \cdots \circ G_{p_1}$ is a neighbor-shared graph. Jointly neighbor shared sequences of graphs from $\mathcal{G}_{sa}$ remains 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 the sequel we will say that an infinite sequence of graphs $G_{p_1}, G_{p_2}, \ldots$, in $\mathcal{G}$ is \textit{repeatedly jointly rooted} if there is a positive integer $q$ for which each finite sequence $G_{p_{q(k-1)+1}}, \ldots, G_{p_{qk}}$, $k \geq 1$ is jointly rooted. If such an integer exists we sometimes say that $G_{p_1}, G_{p_2}, \ldots$, repeatedly jointly rooted by \textit{sub-sequences of length} $q$. We are now in a position to generalize Proposition 11.

**Proposition 12** Let $\tilde{S}$ 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 $\tilde{S}$ whose corresponding sequence of graphs $\gamma(S_1), \gamma(S_2), \ldots$ is repeatedly jointly rooted by sub-sequences of length $q$. Suppose that the set of all products of $q$ matrices from $\tilde{S}$ with rooted graphs, written $\tilde{S}(q)$, is closed. There exists an integer $m$ such that the product of every set of $m$ matrices from $\tilde{S}(q)$ is strongly rooted. Let $m$ be any such integer and write $(\tilde{S}(q))^m$ for the set of all such matrix products. Then as $j \to \infty$, the product $S_j \ldots S_1$ converges exponentially fast to $1[\cdots S_j \cdots S_1]$ at a rate no slower than

$$
\lambda = \left( \max_{S \in (\tilde{S}(q))^m} \|S\| \right)^{\frac{1}{mq}}
$$

where $\lambda < 1$.

It is worth pointing out that the assumption that $\tilde{S}(q)$ is closed is not necessarily implied by the assumption that $\tilde{S}$ is closed. For example, if $\tilde{S}$ is the set of all $2 \times 2$ stochastic matrices whose diagonal elements are no smaller than some positive number $\alpha < 1$, then $\tilde{S}(2)$ cannot be closed even though $\tilde{S}$ is; this is because there are matrices in $\tilde{S}(2)$ which are arbitrarily close (in the induced infinity norm) to the $2 \times 2$ identity which, in turn, is not in $\tilde{S}(2)$. There are at least three different situations where $\tilde{S}(q)$ turns out to be closed. The first is when $\tilde{S}$ is a finite set, as is the case when $\tilde{S}$ is all $n \times n$ flocking matrices; in this case it is obvious that for any $q \geq 1$, $\tilde{S}(q)$ 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 [6]. In this case, the set $\tilde{S}$ 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 under-bounded by a positive number $\alpha < 1$. In this case it is easy to see that for each graph $G \in \mathcal{G}_{sa}$, the sub-set $\tilde{S}(G)$ of $S \in \tilde{S}$ for which $\gamma(S) = G$ is closed. Thus for any pair of graphs $G_1, G_2 \in \mathcal{G}_{sa}$, the subset of products $S_2S_1$ such that $S_1 \in \tilde{S}(G_1)$ and $S_2 \in \tilde{S}(G_2)$ is also closed. Since $\tilde{S}(2)$ is the union of a finite number of sets of products of this type, namely those for which the pairs $(G_1, G_2)$ have rooted compositions $G_2 \circ G_1$, it must be that $\tilde{S}(2)$ is closed. Continuing this reasoning, one can conclude that for any integer $q > 0$, $\tilde{S}(q)$ is closed as well.

The third situation in which $\tilde{S}(q)$ turns out to be compact is considerably more complicated and arises in connection with an asynchronous version of the flocking problem we’ve been studying. In
this case, the graphs of the matrices in $S$ do not have self-arcs at all vertices. We refer the reader to [17] 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 $S(q)$.

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

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

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

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

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

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

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 $N(0), N(1), \ldots$ is repeatedly jointly rooted, there is a constant steady state heading $\theta_{ss}$ for which

$$\lim_{t \to \infty} \theta(t) = \theta_{ss} 1$$

(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 $F$ is finite, so is the set of all products of $q$ flocking matrices with rooted graphs, namely $F(q)$. Therefore $F(q)$ is closed. Moreover, if $m = (n-1)^2$, every product of $m$ matrices from $F(q)$ is strongly rooted. It follows from Proposition 12 that the product $F(t) \cdots F(1) F(0)$ converges exponentially fast to $1[\cdots F(t) \cdots F(1) F(0)]$ exponentially fast as $t \to \infty$ at a rate no slower than

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

where $m = (n-1)^2$, $\lambda < 1$, and $(F(q))^m$ is the closed set of all products of $m$ matrices from $F(q)$.

But it is clear from (3) that

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

23
Therefore (34) holds with \( \theta_{ss} = [\cdots F_{\sigma(t)} \cdots F_{\sigma(0)}] \theta(0) \) and the convergence is exponential. 

It is possible to compare Theorem 3 with similar results derived in [3, 4]. To do this it is necessary to introduce a few concepts. By the union \( G_1 \cup G_2 \) of two directed graphs \( G_1 \) and \( G_2 \) with the same vertex set \( V \), is meant that graph whose vertex set is \( V \) and whose arc set is the union of the arc sets of \( G_1 \) and \( G_2 \). 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 \( \{G_{p_1}, G_{p_2}, \ldots, G_{p_k}\} \) 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 \( G_{p_1}, G_{p_2}, \ldots \) in \( G_{sa} \) is repeatedly collectively rooted if there is a positive integer \( q \) for which each finite set \( G_{p_{q(k-1)+1}}, \ldots, G_{p_{qk}}, k \geq 1 \) is collectively rooted. One of the main contributions of [3] is to prove that the conclusions of Theorem 3 hold if theorem’s hypothesis is replaced by the hypothesis that the sequence of graphs \( G_{\sigma(0)}, G_{\sigma(1)}, \ldots \) 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 \( G_{sa} \) have self-arcs, each arc \((i, j)\) in the union \( G_2 \cup G_1 \) of two graphs \( G_1, G_2 \) in \( G_{sa} \) is an arc in the composition \( G_2 \circ G_1 \). While the converse is not true, the definition of composition does imply that for each arc \((i, j)\) in the composition \( G_2 \circ G_1 \) there is a path in the union \( G_2 \cup G_1 \) 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 \( G_{sa} \), 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 \( G_{p_1}, G_{p_2}, \ldots, G_{p_k} \) in \( G_{sa} \) is jointly rooted if and only if the set of graphs \( \{G_{p_1}, G_{p_2}, \ldots, G_{p_q}\} \) is collectively rooted. It follows that a sequence of graphs in \( G_{sa} \) 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 [3] 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 \( G_{sa} \) whose unions are strongly rooted, and \( \pi_c(m, n) \) were the number of \( m \) term sequences of graphs in \( G_{sa} \) 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 [14].

3 Symmetric Neighbor Relations

It is natural to call a graph in \( G \) 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 \( G \) 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 \( G \) with
an undirected graph $G^*$ 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 [8] and which
was analyzed in [2] using undirected graphs.

Let $G^s$ and $G^s_{sa}$ denote the subsets of symmetric graphs in $G$ and $G_{sa}$ respectively. Simple examples
show that neither $G^s$ nor $G^s_{sa}$ is closed under composition. In particular, composition of two symmetric
directed graphs in $G$ or $G_{sa}$ is not typically symmetric. On the other hand the union is. It is clear
that both $G^s$ and $G^s_{sa}$ are closed under the union operation. It is worth emphasizing that union and
composition are really quite different operations. For example, as we’ve 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
$V$ 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 $G$ have different arc sets – and in the case of graphs from $G_{sa}$, the arc
set of the union is always contained in the arc set of the composition, but not conversely.

In [2] use is made of the notion of a “jointly connected set of graphs.” Specifically, a set of
undirected graphs with vertex set $V$ 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 the sequel we will say that
an infinite sequence of graphs $G_{p_1}, G_{p_2}, \ldots$, in $G_{sa}$ is repeatedly jointly connected if there is a positive
integer $m$ for which each finite sequence $G_{p_{mk(k-1)+1}}, \ldots, G_{p_{mk}}, \quad k \geq 1$ is jointly connected. The
main result of [2] 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 $N(0), N(1), \ldots$ is repeatedly jointly connected, there is a
constant steady state heading $\theta_{ss}$ for which

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

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 in explaining how a consensus is achieved under various condi-
tions 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 non-homogeneous Markov chains.
In a sequel to this paper [14], 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 [14] 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 [14] 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 the 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 [7] would also be useful.

References


