Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

TuA08.6

A Lower Bound on Convergence of a Distributed Network Consensus Algorithm

M. Cao[1]
Dept. of Electrical Engineering
Yale University
m.cao@yale.edu

D. A. Spielman[2]
Dept. of Computer Science
Yale University
spielman@cs.yale.edu

A. S. Morse[1]
Dept. of Electrical Engineering
Yale University
as.morse@yale.edu

Abstract— **This paper gives a lower bound on the convergence rate of a class of network consensus algorithms. Two different approaches using directed graphs as a main tool are introduced: one is to compute the "scrambling constants" of stochastic matrices associated with "neighbor shared graphs" and the other is to analyze random walks on a sequence of graphs. Both approaches prove that the time to reach consensus within a dynamic network is logarithmic in the relative error and is in worst case exponential in the size of the network.**

I. Introduction

Distributed algorithms where processors reach consensus by exchanging locally computed results have been studied extensively in the field of distributed and parallel computations [1], [2]. In recent years, network consensus algorithms have attracted more attention from control scientists and engineers interested in distributed coordination of groups of mobile autonomous agents. Some simple local rules can cause a group of agents to behave cooperatively without the existence of a central controller. The cooperative behaviors include moving in the same direction [3], [4], aggregating in swarms [5], and rendezvousing at the same location [6], [7]. Two salient features make the convergence of the consensus algorithms difficult to analyze in the context of distributed coordination within multi-agent systems: First, the final consensus value cannot be determined a priori and depends on the initial value of each agent as well as the dynamic interaction between them; and secondly, because agents move, the underlying network topology may not be fixed during system evolutions.

A special class of network consensus algorithms for flocking via distributed averaging was first studied in [8],

and convergence was proved under mild connectivity assumptions in [3] by using switching linear system theory and algebraic graph theory. Since then various works have been done to develop this class of algorithms in continuous time [9], on weakly connected graphs [4], and in asynchronous scenarios [10]. However, so far few if any specific convergence rates have been derived for this class of algorithms. This is due to the fact that tools such as joint spectrum theory cannot compute an explicit bound for this particular class of switched linear systems without equilibriums that can be determined beforehand.

This paper will study the convergence rate for the flocking process based on the encountered directed graphs. A lower bound on the convergence rate is presented using "scrambling constants" for "neighbor shared graphs". It will be shown that the time needed for all the agents' headings to converge to a ε-neighborhood of the steady state is logarithmic in ε and in worst case exponential in the total number of agents. To better understand the factors that influence the convergence rate, a separate approach using random walks on graphs is also discussed.

The rest of the paper is organized as follows. In section 2, the flocking algorithm based on distributed averaging is introduced. In section 3, the convergence is proved for the case when graphs encountered during the flocking process are all "neighbor shared". In section 4, we study the more general case when graphs encountered are all rooted. In section 5, a separate approach using random walks on graphs is introduced.

II. Flocking via Distributed Averaging

Consider a system that consists of n mobile autonomous agents, labelled 1 through n, that try to reach agreement on a scalar variable which will be called the *heading*. Each agent's heading θ_i is updated using

a simple local rule based on the average of its own heading and the headings of its "neighbors". Agent i's *neighbors* at time t are those agents, including itself, whose value at time t is available to agent i. Because of motion or other possible causes, agent i's neighbors may change with time. Let $\mathcal{N}_i(t)$ and $n_i(t)$ denote the set of labels and the number of agent i's neighbors at time t respectively. Then agent i's heading evolves in discrete-time in accordance with a model of the form

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

where t is a discrete-time index taking values in the non-negative integers $\{0, 1, 2, \ldots\}$. We assume that the time between such updates is bounded below by a positive number τ_B called a dwell time, and that all agents update their values synchronously.

The explicit form of the update equations determined by (1) depends on the neighbor relationships which exist at time t. These relationships can be conveniently described by a directed graph $\mathbb{G}$ with vertex set $\mathcal{V} = \{1, 2, \ldots n\}$ and arc set $\mathcal{A}(\mathbb{G}) \subset \mathcal{V} \times \mathcal{V}$ 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 $\mathbb{G}$ is a directed graph on n vertices with at most one arc from any vertex to another and with exactly one self-arc at each vertex. We write $\mathcal{G}$ for the set of all such graphs. We use the symbol $\mathcal{P}$ to denote a suitably defined set indexing $\mathcal{G}$. Thus $\mathcal{G} = \{\mathbb{G}_p : p \in \mathcal{P}\}$ where for $p \in \mathcal{P}$, $\mathbb{G}_p$ denotes the pth graph in $\mathcal{G}$. It is natural to call a vertex i a *neighbor* of vertex j in $\mathbb{G}$ 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 suppose to represent.

The set of agent heading update rules defined by (1) can be written in state form. Toward this end, for each $p \in \mathcal{P}$, define the *flocking matrix*

$$F_p = D_p^{-1} A_p' \qquad (2)$$

where A_p' is the transpose of the *adjacency matrix* of the graph $\mathbb{G}_p$ and D_p is the diagonal matrix whose jth diagonal element is the *in-degree* of vertex j within the graph[1]. Then

$$\theta(t+1) = F_{\sigma(t)}\theta(t), \quad t \in \{0, 1, 2, \ldots\} \qquad (3)$$

[1] By the *adjacency matrix* of a directed graph $\mathbb{G} \in \mathcal{G}$ is meant 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 vertices it observes.

where θ is the state vector $\theta = [\theta_1 \ \theta_2 \ \cdots \ \theta_n]'$ and $\sigma : \{0, 1, \ldots\} \to \mathcal{P}$ is a switching signal whose value at time t, is the index of the graph representing the agents' neighbor relationships at time t.

We will show for a large class of switching signals that all the agents will reach the same steady state value θ_{ss} given any initial set of agent headings. 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 $\mathbf{1}$ is the $n \times 1$ vector $[1 \ 1 \ \cdots \ 1]'$. In addition, we will also give a sharp lower bound on the rate of this convergence process. The main challenge lies in how to define and analyze the rate of convergence for a multi-agent system under changing neighbor relationships.

III. NEIGHBOR SHARED GRAPHS

We will first give a convergence result for the case where the graphs encountered along a trajectory of (3) are all "neighbor shared". This requirement can be relaxed significantly as explained in the next section. Let us call $\mathbb{G} \in \mathcal{G}$ *neighbor shared* if each set of 2 distinct vertices share a common neighbor. It is now possible to state the following elementary convergence result.

Theorem 1: Let $\mathcal{Q}$ denote the subset of $\mathcal{P}$ consisting of those indices q for which $\mathbb{G}_q \in \mathcal{G}$ is neighbor shared. Let $x(0)$ be fixed and let $\sigma : \{0, 1, 2, \ldots\} \to \mathcal{P}$ be a switching signal satisfying $\sigma(t) \in \mathcal{Q}, \ t \in \{0, 1, \ldots\}$. Then there is a constant steady state heading θ_{ss} depending only on $\theta(0)$ and σ for which

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

where the limit is approached exponentially fast.

In order to explain why this theorem is true, we will make use of certain structural properties of F_p. As defined, each F_p is square and non-negative, where by a *non-negative* matrix is meant a matrix whose entries are all non-negative. Each F_p also has the property that its row sums all equal 1 {i.e., $F_p\mathbf{1} = \mathbf{1}$}. Matrices with these two properties are called {row} *stochastic*. Because each vertex of each graph in $\mathcal{G}$ has a self-arc, the F_p have the additional property that their diagonal elements are all non-zero.

Stochastic matrices have been studied extensively in the literature for a long time largely because of their connection with Markov chains [11]. One problem 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

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

because 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 [12], is to consider the behavior of the scalar-valued non-negative function $V(x) = \lceil x \rceil - \lfloor x \rfloor$ along solutions to (5) where $x = [x_1 \; x_2 \; \cdots \; x_n]'$ is a non-negative $n \times 1$ 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, $\lfloor Sx \rfloor \geq \lfloor x \rfloor$, $\lceil Sx \rceil \leq \lceil x \rceil$ and, as a consequence, that $V(Sx) \leq V(x)$. These inequalities and (5) imply that the sequences

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

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

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

all exist. Note that whenever the limit of $V(x(j))$ is zero, all components of $x(j)$ must tend to the same value and moreover this value must be a constant equal to the limiting value of $\lfloor x(j) \rfloor$.

We will make use of the standard partial ordering $\geq$ on $n \times n$ nonnegative matrices by writing $B \geq A$ whenever $B - A$ is nonnegative. We extend the domain of definitions of $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ to the class of all non-negative $n \times m$ 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. 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

$$\lfloor \cdots S_j \cdots S_2 S_1 \rfloor = \lim_{j \to \infty} \lfloor S_j \cdots S_2 S_1 \rfloor \tag{6}$$

From the preceding discussion it is clear that this limit exists whether or not the product $S_j \cdots S_2 S_1$ itself has a limit. In the sequel we will be interested in the non-negative matrix

$$[S] = \mathbf{1}(\lceil S \rceil - \lfloor S \rfloor) \tag{7}$$

Note that whenever $[S] = 0$, all rows of S must be equal.

Lemma 1: For any two $n \times n$ stochastic matrices S_1 and S_2,

$$[S_2 S_1] \leq \mu(S_2)[S_1] \tag{8}$$

where for any $n \times n$ stochastic matrix S,

$$\mu(S) = \max_{i,j} \left(1 - \sum_{k=1}^{n} \min\{s_{ik}, s_{jk}\} \right) \tag{9}$$

The proof of this lemma will appear in the full length version of this paper. The quantity $\mu(S)$ has been widely studied before [13], [11] and is known as the *scrambling constant* of the stochastic matrix S. Note that since the row sums of S all equal 1, $\mu(S)$ is non-negative. It is easy to see that $\mu(S) = 0$ just in case all the rows of S are equal. Let us note that for fixed i and j, the kth term in the sum appearing in (9) will be positive just in case both s_{ik} and s_{jk} are positive. It follows that the sum will be positive if and only if for at least one k, s_{ik} and s_{jk} are both positive. Thus $\mu(S) < 1$ if and only if for each distinct i and j, there is at least one k for which s_{ik} and s_{jk} are both positive. Matrices with this property have been widely studied and are called *scrambling matrices*. Thus a stochastic matrix S is a scrambling matrix if and only if $\mu(S) < 1$. It is easy to see that the definition of a scrambling matrix also implies that S is scrambling if and only if its associated graph is neighbor shared.

Suppose that F_p is a flocking matrix for which $\mathbb{G}_p$ is neighbor shared. In view of the definition of a flocking matrix, any non-zero entry in F_p must be bounded below by $\frac{1}{n}$. Fix distinct i and j and suppose that k is a neighbor that i and j share. Then f_{ik} and f_{jk} are both non-zero, so $\min\{f_{ik}, f_{jk}\} \geq \frac{1}{n}$. This implies that the sum in (9) must be bounded below by $\frac{1}{n}$ and consequently that $\mu(F_p) \leq 1 - \frac{1}{n}$.

Now let F_p be that flocking matrix whose graph $\mathbb{G}_p \in \mathcal{G}$ is such that vertex 1 has no neighbors other than itself, vertex 2 has every vertex as a neighbor, and vertices 3 through n have only themselves and agent 1 as neighbors. Since vertex 1 has no neighbors other than itself, $f_{i,k} = 0$ for all i and for $k > 1$. Thus for all i, j, it must be true that $\sum_{k=1}^{n} \min\{f_{ik}, f_{jk}\} = \min\{f_{i1}, f_{j1}\}$. Now vertex 2 has n neighbors, so $f_{2,1} = \frac{1}{n}$. Thus $\min\{f_{i1}, f_{j1}\}$ attains its lower bound of $\frac{1}{n}$ when either $i = 2$ or $j = 2$. It thus follows that with this F_p, $\mu(F_p)$ attains its upper bound of $1 - \frac{1}{n}$. We summarize.

Lemma 2: Let $\mathcal{Q}$ be the set of indices in $\mathcal{P}$ for which $\mathbb{G}_p$ is neighbor shared. Then

$$\max_{q \in \mathcal{Q}} \mu(F_q) = 1 - \frac{1}{n} \tag{10}$$

Lemma 1 and 2 will be used in the proof of Theorem 1.

Proof of Theorem 1: By definition, the graph $\mathbb{G}_p$ of each matrix F_p in the finite set $\{F_p : p \in \mathcal{Q}\}$ is neighbor shared. By assumption, $F_{\sigma(t)} \in \{F_p : p \in \mathcal{Q}\}$ for $t \geq 0$. Let $\lambda = 1 - \frac{1}{n}$. In view of Lemma 1 and 2,

$$\lfloor F_{\sigma(t)} \cdots F_{\sigma(0)} \rfloor \leq \lfloor F_{\sigma(t)} \rfloor \cdots \lfloor F_{\sigma(0)} \rfloor \leq \lambda^{t+1}$$

and λ^{t+1} approaches zero exponentially fast as $t \to \infty$. Thus the product $F_{\sigma(t)} \cdots F_{\sigma(0)}$ converges to $\mathbf{1} \lfloor \cdots F_{\sigma(t)} \cdots F_{\sigma(0)} \rfloor$ exponentially fast at a rate no slower than λ. But it is clear from (3) that

$$\theta(t) = F_{\sigma(t-1)} \cdots F_{\sigma(1)} F_{\sigma(0)} \theta(0), \quad t \geq 1$$

Therefore (4) holds with $\theta_{ss} = \lfloor \cdots F_{\sigma(t)} \cdots F_{\sigma(0)} \rfloor \theta(0)$ and the convergence is exponential. ∎

Hence, we have proved that $1 - \frac{1}{n}$ is a worst case bound on the rate of convergence of products of flocking matrices whose graphs are all neighbor shared.

IV. Rooted graphs

The proof of Theorem 1 depends crucially on the fact that the graphs encountered along a trajectory of (3) are all neighbor shared. The aim of this section is to show that this requirement can be relaxed. To do this we need to have a meaningful way of "combining" sequences of graphs so that only the combined graph need be neighbor shared, but not necessarily the individual graphs making up the combination. Let us agree to say that the *composition* [10] of a directed graph $\mathbb{G}_{p_1} \in \mathcal{G}$ with a directed graph $\mathbb{G}_{p_2} \in \mathcal{G}$, written $\mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}$, is the directed graph with vertex set $\{1, \ldots, n\}$ 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 q such that (i, q) is an arc of $\mathbb{G}_{p_1}$ and (q, j) is an arc of $\mathbb{G}_{p_2}$. Thus (i, j) is an arc of $\mathbb{G}_{p_2} \circ \mathbb{G}_{p_1}$ if and only if i has an observer in $\mathbb{G}_{p_1}$ which is also a neighbor of j in $\mathbb{G}_{p_2}$. Note that $\mathcal{G}$ is closed under composition and that composition is an associative binary operation; because of this, the definition extend unambiguously to any finite sequence of directed graphs $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_k}$. Note that the definition of composition takes into account the order in which the graphs are encountered along a trajectory.

In the sequel we will call a vertex i of a directed graph $\mathbb{G} \in \mathcal{G}$ a *root* of $\mathbb{G}$ if for each other vertex j of $\mathbb{G}$, there is a *path* from i to j[2]. We will say that $\mathbb{G}$ is *rooted at i*

---

[2]In a directed graph $\mathbb{G} \in \mathcal{G}$, by the *path* from vertex i_1 to vertex i_k is meant a sequence of vertices $\{i_1, i_2, \ldots, i_k\}$ such that $(i_1, i_2), (i_2, i_3), \ldots, (i_{k-1}, i_k)$ are arcs of $\mathbb{G}$.

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. By a *rooted graph* $\mathbb{G} \in \bar{\mathcal{G}}$ we mean a graph which possesses at least one root.

Lemma 3: Each neighbor shared graph in $\mathcal{G}$ is rooted.

Proof: In a graph $\mathbb{G} \in \mathcal{G}$, we say v is a root for vertices $1, \ldots, k$ if each of vertices $1, \ldots, k$ is reachable from v. In a neighbor shared graph, vertices 1 and 2 have a root. One may now prove by induction that if $1, \ldots, k$ have a root for an integer $2 \leq k < n$, then $1, \ldots, k+1$ do as well: any common neighbor of vertex $k + 1$ and the root of $1, \ldots, k$ will suffice. ∎

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 $\mathcal{G}$ need not be neighbor shared, it turns out that the composition of any $n - 1$ rooted graphs in $\mathcal{G}$ is.

Lemma 4: The composition of any set of $m \geq n - 1$ rooted graphs in $\mathcal{G}$ is neighbor shared.

The proof of this lemma will appear in the full length version of this paper.

It is also possible to derive a worst case convergence rate for products of flocking matrices which have rooted rather than neighbor-shared graphs. We will use the following result without providing its proof.

Lemma 5: Let b be a positive number less than 1. Let $\mathcal{S}_r^m$, $m \geq 1$, denote the set of all m-term matrix products $S = S_m S_{m-1} \cdots S_1$ where each S_i is an $n \times n$ stochastic matrix with a rooted graph in $\mathcal{G}$ and all nonzero entries bounded below by b. Then

$$\max_{S \in \mathcal{S}_r^{n-1}} \mu(S) = 1 - b^{(n-1)}$$

It turns out that this bound is actually attained if all the S_i are the same taking the form

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ b & 1-b & 0 & 0 & \cdots & 0 \\ 0 & b & 1-b & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 1-b & 0 \\ 0 & 0 & 0 & 0 & b & 1-b \end{bmatrix} \quad (11)$$

It is possible to apply at least part of the preceding to the case when the S_i are flocking matrices. Towards this end, let $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_{n-1}}$ be any sequence of

$n - 1$ rooted graphs in $\mathcal{G}$ and let $F_{p_1}, \ldots, F_{p_{n-1}}$ be the sequence of flocking matrices associated with these graphs. Since each F_p is a flocking matrix, it must be true that any non-zero element in F_p is bounded below by $\frac{1}{n}$. Then

$$\mu(F_{p_{n-1}} \cdots F_{p_1}) \leq 1 - \left(\frac{1}{n}\right)^{(n-1)} \qquad (12)$$

Unfortunately, we cannot use the preceding reasoning to show that (12) holds with equality for some sequence of rooted graphs. This is because the stochastic matrix in (11) is not a flocking matrix when $b = \frac{1}{n}$, except in the special case when $n = 2$. Nonetheless (12) can be used as follows to develop a convergence rate for products of flocking matrices whose graphs are all rooted. Let $\mathcal{Q}$ denote the set of $p \in \mathcal{P}$ for which $\mathbb{G}_p$ is rooted and write $\mathcal{F}_r^{n-1}$ for the closed set of all products of flocking matrices of the form $F_{p_{n-1}} \cdots F_{p_1}$ where each $p_i \in \mathcal{Q}$. In view of Lemma 4, $\mathbb{G}_{p_{n-1}} \circ \cdots \circ \mathbb{G}_{p_1}$ is neighbor shared for every list of $n - 1$ indices $\{p_1, p_2, \ldots, p_{n-1}\}$ from $\mathcal{Q}$, and (12) holds for every such list. Now for any sequence $p(1), p(2), \ldots, p(j)$ of indices in $\mathcal{Q}$, the corresponding product $F_{p(j)} \cdots F_{p(1)}$ of flocking matrices can be written as

$$F_{p(j)} \cdots F_{p(1)} = \bar{S}(j)\bar{S}_k \cdots \bar{S}_1$$

where

$$\bar{S}_i = F_{p(i(n-1))} \cdots F_{p((i-1)(n-1)+1)}, \quad 1 \leq i \leq k,$$

$$\bar{S}(j) = F_{p(j)} \cdots F_{p(k(n-1)+1)},$$

and k is the integer quotient of j divided by $n - 1$. In view of (12)

$$\mu(\bar{S}_i) \leq \bar{\lambda}, \quad i \in \{1, 2, \ldots, k\}$$

where

$$\bar{\lambda} = 1 - \left(\frac{1}{n}\right)^{(n-1)}$$

It is clear that $\bar{S}_k \cdots \bar{S}_1$ 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 $\bar{\lambda}$. But $\bar{S}(j)$ is a product of at most $n - 1$ stochastic matrices, so it is a bounded function of j. It follows that the product $F_{p(j)} \cdots F_{p(1)}$ must converge to $\mathbf{1}\lfloor F_{p(j)} \cdots F_{p(1)} \rfloor$ exponentially fast at a rate no slower than $\lambda = \bar{\lambda}^{\frac{1}{n-1}}$. Using the development similar to that used in the proof of Theorem 1, we can prove the following theorem.

Theorem 2: Let $\mathcal{Q}$ denote the subset of $\mathcal{P}$ consisting of those indices q for which $\mathbb{G}_q \in \mathcal{G}$ is rooted. Let $\theta(0)$ be fixed and let $\sigma : \{0, 1, 2, \ldots\} \to \mathcal{P}$ be a switching signal satisfying $\sigma(t) \in \mathcal{Q}$, $t \in \{0, 1, \ldots\}$. Then there

is a constant steady state heading θ_{ss} depending only on $\theta(0)$ and σ for which

$$\lim_{t \to \infty} \theta(t) = \theta_{ss}\mathbf{1} \qquad (13)$$

where the limit is approached exponentially fast at a rate no slower than

$$\lambda = \left\{ 1 - \left(\frac{1}{n}\right)^{(n-1)} \right\}^{\frac{1}{n-1}}$$

It is possible to develop analogous results for *strongly connected* graphs, where by a strongly connected graph we mean a directed graph that has a path from each vertex to every other vertex. We will state the following result without proof.

Corollary 1: Under the hypotheses of Theorem 2, and the additional assumption that σ takes values only in the subset of $\mathcal{Q}$ composed of those indices for which $\mathbb{G}_p$ is strongly connected, convergence of $\theta(t)$ to $\theta_{ss}\mathbf{1}$ is exponential at a rate no slower than

$$\lambda = \left\{ 1 - \left(\frac{1}{n}\right)^{m} \right\}^{\frac{1}{m}}$$

where m is the integer quotient of n divided by 2.

V. RANDOM WALKS ON DIRECTED GRAPHS

We need some more ideas in this section. 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 $\mathcal{G}$ is closed under the reverse operation. It is also clear that $(\mathbb{G}_p \circ \mathbb{G}_q)' = \mathbb{G}'_q \circ \mathbb{G}'_p, p, q \in \mathcal{P}$. For any flocking matrix $F_p, p \in \mathcal{P}$, let $\gamma(F_p)$ denote that graph $\mathbb{G} \in \mathcal{G}$ corresponding to F_p. A different approach to analyzing the convergence of the flocking process is to multiply F_p by row vectors from the left. In this approach, we are looking at the random walk [14] where at each time step we apply a different graph $\mathbb{G}'_{\sigma(t)}, \sigma : \{0, 1, 2, \ldots\} \to \mathcal{P}$. Then the random walk will converge to some fixed distribution if the flocking process converges.

First, we will consider walks that begin at just one vertex, that is to begin with a distribution given by an elementary $1 \times n$ vector e_i with $e_i(i) = 1$ and $e_i(j) = 0$ for $j \neq i$. In the sequel, let $\mathcal{Q}$ denote the set of $p \in \mathcal{P}$ for which $\mathbb{G}_p$ is rooted.

Lemma 6: For any sequence $p(1), p(2), \ldots, p(n-1)$ of indices in $\mathcal{Q}$, let $\mathcal{R}_i$ denote the set of indices of those vertices that are reachable from vertex i after a walk on

a sequence of $n-1$ graphs $\gamma(F_{p(n-1)})', \ldots, \gamma(F_{p(1)})'$. Then for all i and $j \in \mathcal{R}_i$

$$\left(e_i F_{p(n-1)} \cdots F_{p(1)}\right)(j) \geq (\frac{1}{n})^{(n-1)} \qquad (14)$$

Proof: The vector e_i can be thought of as a unit positive mass at vertex i. From the definition of $\mathcal{R}_i$, we know vertex j can be reached from vertex i by a walk of $n-1$ steps on the sequence of $n-1$ graphs. Since in each step at least $\frac{1}{n}$ fraction of the mass is propagated along an arc, we know vertex j has at least $(\frac{1}{n})^{(n-1)}$ fraction of the unit mass. ∎

Now we will consider the difference between any pair of rows of the matrix $F_{p(n-1)} \cdots F_{p(1)}$.

Lemma 7: For any sequence $p(1), p(2), \ldots, p(n-1)$ of indices in $\mathcal{Q}$ and for each $i \neq j$

$$\|(e_i - e_j)F_{p(n-1)} \cdots F_{p(1)}\|_1 \leq 2 - 2(\frac{1}{n})^{(n-1)} \quad (15)$$

where for a vector $a = [a_1 \ a_2 \ \ldots \ a_n]'$, $\|a\|_1 = \sum_i |a_i|$.

Proof: The initial vector $e_i - e_j$ can be thought of as unit positive mass at vertex i and unit negative mass at vertex j. As the walk progresses, the sum of positive mass minus negative mass remains zero. When a positive mass meets a negative mass, they cancel each other out. Since $\gamma(F_{p(n-1)} \cdots F_{p(1)})$ by Lemma 4 is neighbor shared, the pair of vertices i and j can both reach some vertex k in $\gamma(F_{p(n-1)} \cdots F_{p(1)})'$. Hence, after $n-1$ steps of the random walk starting from vertices i and j, in view of Lemma 6 we will cancel at least $(\frac{1}{n})^{(n-1)}$ mass at vertex k. ∎

Now we are in a position to give a lower bound on the rate of convergence of the random walk process.

Lemma 8: For any sequence $p(1), p(2), \ldots, p(n-1)$ of indices in $\mathcal{Q}$ and for every row vector x such that $\sum_i x(i) = 0$,

$$\|x F_{p(n-1)} \cdots F_{p(1)}\|_1 \leq \bar{\lambda}\|x\|_1 \qquad (16)$$

where $\bar{\lambda} = 1 - (\frac{1}{n})^{(n-1)}$.

Proof: Let i be the index of the minimal non-zero value $|x_i|$. Assume without loss of generality that x_i is positive. Then pick some j such that x_j is negative. Lemma 7 says that at least a $\left(1 - (\frac{1}{n})^{(n-1)}\right)$ fraction of the mass propagated from x_i will be cancelled by mass propagated from x_j. Now, remove x_i from consideration, and x_i of the negative mass at x_j. Continuing in this way, we can pair up positive masses with negative masses, and by observing the fraction of cancellation,

we have

$$\|x F_{p(n-1)} \cdots F_{p(1)}\|_1 \leq \left(1 - (\frac{1}{n})^{(n-1)}\right)\|x\|_1 \quad ∎$$

Using Lemma 8, we can arrive at the same convergence result as that in Theorem 2. According to our experience, we can gain insight into the convergence rate by constructing special sequences of graphs on which the random walk progresses.

VI. Concluding Remarks

It is possible to relax still further the conditions under which the flocking process converges. We only require the sequence of graphs encountered during the flocking process be repeatedly "jointly rooted" where a finite sequence of directed graphs $\mathbb{G}_{p_1}, \ldots, \mathbb{G}_{p_k}$ in $\mathcal{G}$ is *jointly rooted* if the composition $\mathbb{G}_{p_k} \circ \cdots \circ \mathbb{G}_{p_1}$ is rooted.

In the future, we are interested in designing consensus algorithms with faster convergence rates once we gain deeper insight into the factors that affect the convergence process.

References

[1] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation*. Prentice Hall, 1989.

[2] N. A. Lynch. *Distributed algorithms*. Morgan Kaufmann Publishers, San Francisco, 1996.

[3] A. Jababaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48:988–1001, 2003.

[4] L. Moreau. Leaderless coordination via bidirectional and unidirectional time-dependent communication. In *Proc. of the 42th IEEE Conference on Decision and control*, pages 3070–3075, 2003.

[5] V. Gazi and K. M. Passino. Stability analysis of swarms. *IEEE Transactions on Automatic Control*, 48:692–697, 2003.

[6] J. Lin, A. S. Morse, and B. D. O. Anderson. The multi-agent rendezvous problem. In *Proc. of the 42th IEEE Conference on Decision and control*, pages 1508–1513, 2003.

[7] S. Martinez, J. Cortes, and F. Bullo. On robust rendezvous for mobile autonomous agents. In *Proc. of the IFAC world congress, Prague, Czech Republic*, 2005.

[8] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75:1226–1229, 1995.

[9] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with swiching topology and time-delays. *IEEE Transactions on Automatic Control*, 49:101–115, 2004.

[10] M. Cao, A. S. Morse, and B. D. O. Anderson. Coordination of an asynchronous multi-agent system via averaging. In *Proc. of the IFAC world congress, Prague, Czech Republic*, 2005.

[11] E. Seneta. *Non-negative matrices and Markov chains*. Springer-Verlag, New York, 1981.

[12] J. L. Doob. *Stochastic Processes*, chapter 5: Markov Processes, Discrete Parameter. John Wiley & Sons, Inc., New York, 1953.

[13] J. Hajnal. Weak ergodicity in nonhomogeneous markov chains. *Proc. Camb. Phil. Soc.*, 54:233–246, 1958.

[14] L. Lovasz. Random walks on graphs: a survey. In *Cominatorics, Paul Erdös is eighty*, pages 353–397, Budapest, 1996. Janos Bolyai Math. Soc.