Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules

Ali Jadbabaie, Jie Lin, and A. Stephen Morse, *Fellow, IEEE*

***Abstract*—In a recent *Physical Review Letters* article, Vicsek *et al.* propose a simple but compelling discrete-time model of n autonomous agents (i.e., points 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 headings of its "neighbors." In their paper, Vicsek *et al.* provide 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 change with time as the system evolves. This paper provides a theoretical explanation for this observed behavior. In addition, convergence results are derived for several other similarly inspired models. The Vicsek model proves to be a graphic example of a switched linear system which is stable, but for which there does not exist a common quadratic Lyapunov function.**

***Index Terms*—Cooperative control, graph theory, infinite products, multiagent systems, switched systems.**

I. Introduction

In [1], Vicsek *et al.* propose a simple but compelling discrete-time model of n autonomous agents (i.e., points 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 headings of its "neighbors." Agent i's *neighbors* at time t, are those agents which are either in or on a circle of pre-specified radius r centered at agent i's current position. The Vicsek model turns out to be a special version of a model introduced previously by Reynolds [2] for simulating visually satisfying flocking and schooling behaviors for the animation industry. 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 change with time as the system evolves. In this paper, we provide a theoretical explanation for this observed behavior.

Manuscript received March 4, 2002; revised December 2, 2002. Recommended by Associate Editor A. Bemporad. This work was supported by the Defense Advanced Research Project Agency (DARPA) under its SEC program and by the National Science Foundation under its KDI/LIS initiative.

A. Jadbabaie was with Yale University, New Haven, CT 06520 USA. He is now with the Department of Electrical and Systems Engineering and GRASP Laboratory, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: jadbabai@seas.upenn.edu).

J. Lin and A. S. Morse are with the Center for Computational Vision and Control, Department of Electrical Engineering, Yale University, New Haven, CT 06520 USA.

Digital Object Identifier 10.1109/TAC.2003.812781

There is a large and growing literature concerned with the coordination of groups of mobile autonomous agents. Included here is the work of Czirok *et al.* [3] who propose one-dimensional models which exhibit the same type of behavior as Vicsek's. In [4] and [5], Toner and Tu construct a continuous "hydrodynamic" model of the group of agents, while other authors such as Mikhailov and Zanette [6] consider the behavior of populations of self propelled particles with long range interactions. Schenk *et al.* determined interactions between individual self-propelled spots from underlying reaction-diffusion equation [7]. Meanwhile, in modeling biological systems, Grünbaum and Okubo use statistical methods to analyze group behavior in animal aggregations [8]. This paper and, for example, the work reported in [9]–[12] are part of a large literature in the biological sciences focusing on many aspects of aggregation behavior in different species.

In addition to these modeling and simulation studies, research papers focusing on the detailed mathematical analysis of emergent behaviors are beginning to appear. For example, Lui *et al.* [13] use Lyapunov methods and Leonard *et al.* [14] and Olfati and Murray [15] use potential function theory to understand flocking behavior, and Ögren *et al.* [16] uses control Lyapunov function-based ideas to analyze formation stability, while Fax and Murray [17] and Desai *et al.* [18] employ graph theoretic techniques for the same purpose.

The one feature which sharply distinguishes previous analyzes from that undertaken here is that this paper explicitly takes into account possible changes in nearest neighbors over time. Changing nearest neighbor sets is an inherent property of the Vicsek model and in the other models we consider. To analyze such models, it proves useful to appeal to well-known results [19], [20] characterizing the convergence of infinite products of certain types of nonnegative matrices. The study of infinite matrix products is ongoing [21]–[26], and is undoubtedly producing results which will find application in the theoretical study of emergent behaviors.

Vicsek's model is set up in Section II as a system of n simultaneous, one-dimensional recursion equations, one for each agent. A family of simple graphs on n vertices is then introduced to characterize all possible neighbor relationships. Doing this makes it possible to represent the Vicsek model as an n-dimensional switched linear system whose switching signal takes values in the set of indices which parameterize the family of graphs. The matrices which are switched within the system turn out to be nonnegative with special structural properties. By exploiting these properties and making use of a classical convergence result due to Wolfowitz [19], we prove that all n agents' headings converge to a common steady state

heading provided the n agents are all "linked together" via their neighbors with sufficient frequency as the system evolves. The model under consideration turns out to provide a graphic example of a switched linear system which is stable, but for which there does not exist a common quadratic Lyapunov function.

In Section II-B, we define the notion of an average heading vector in terms of graph Laplacians [27] and we show how this idea leads naturally to the Vicsek model as well as to other decentralized control models which might be used for the same purposes. We propose one such model which assumes each agent knows an upper bound on the number of agents in the group, and we explain why this model has convergence properties similar to Vicsek's.

In Section III, we consider a modified version of Vicsek's discrete-time system consisting of the same group of n agents, plus one additional agent, labeled 0, which acts as the group's leader. Agent 0 moves at the same constant speed as its n followers but with a fixed heading θ_0. The ith follower updates its heading just as in the Vicsek model, using the average of its own heading plus the headings of its neighbors. For this system, each follower's set of neighbors can also include the leader and does so whenever the leader is within the follower's neighborhood defining circle of radius r. We prove that the headings of all n agents must converge to the leader's provided all n agents are " linked to their leader" together via their neighbors frequently enough as the system evolves. Finally, we develop a continuous-time analog of this system and prove under condition milder than imposed in the discrete-time case, that the headings of all n agents again converge to the heading of the group's leader.

II. Leaderless Coordination

The system studied by Vicsek *et al.* [1] consists of n autonomous agents (e.g., points or particles), labeled 1 through n, all moving in the plane with the same speed but with different headings.[1] 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 which are either in or on a circle of pre-specified radius r 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 θ_i, evolves in discrete-time in accordance with a model of the form

$$\theta_i(t+1) = \langle \theta_i(t) \rangle_r \tag{1}$$

where t is a discrete-time index taking values in the nonnegative integers $\{0, 1, 2, \ldots\}$, and $\langle \theta_i(t) \rangle_r$ is the average of the headings of agent i and agent i's neighbors at time t; that is

$$\langle \theta_i(t) \rangle_r = \frac{1}{1+n_i(t)} \left(\theta_i(t) + \sum_{j \in \mathcal{N}_i(t)} \theta_j(t) \right) \tag{2}$$

where $n_i(t)$ is the number of neighbors of agent i at time t. Observe that the preceding heading update rule maps headings with values $[0, 2\pi)$ into a heading with a value also in $[0, \pi)$. Because of this, it makes sense to represent headings at any finite time t, as real numbers in $[0, 2\pi)$. Of course it is entirely possible that in the limit as $t \to \infty$, a heading might approach the value 2π; any such limiting value is interpreted as a heading of 0. Analogous statement apply to all other models considered in the sequel. Accordingly, throughout the paper headings at any finite time t, are represented as real numbers in $[0, 2\pi)$.

The explicit form of the update equations determined by (1) and (2) depends on the relationships between neighbors which exist at time t. These relationships can be conveniently described by a simple, undirected graph[2] with vertex set $\{1, 2, \ldots, n\}$ which is defined so that (i, j) is one of the graph's edges just in case agents i and j are neighbors. Since the relationships between neighbors can change over time, so can the graph which describes them. To account for this we will need to consider all possible such graphs. In the sequel we use the symbol $\mathcal{P}$ to denote a suitably defined set, indexing the class of all simple graphs $\mathbb{G}_p$ defined on n vertices.

The set of agent heading update rules defined by (1) and (2), can be written in state form. Toward this end, for each $p \in \mathcal{P}$, define

$$F_p = (I + D_p)^{-1}(A_p + I) \tag{3}$$

where A_p is the adjacency matrix of graph $\mathbb{G}_p$ and D_p the diagonal matrix whose ith diagonal element is the valence of vertex i within the graph. Then

$$\theta(t+1) = F_{\sigma(t)}\theta(t), \qquad t \in \{0, 1, 2, \ldots\} \tag{4}$$

where θ is the heading vector $\theta = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_n]'$ and σ: $\{0, 1, \ldots\} \to \mathcal{P}$ is a switching signal whose value at time t, is the index of the graph representing the agents' neighbor relationships at time t. A complete description of this system would have to include a model which explains how σ 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 σ depends on the agent positions in the plane and assumes instead that σ might be any switching signal in some suitably defined set of interest.

Our goal is to show for a large class of switching signals and for any initial set of agent headings that the headings of all n agents will converge to the same steady state value θ_{ss}. Convergence of the θ_i to θ_{ss} is equivalent to the state vector θ converging to a vector of the form $\theta_{ss}\mathbf{1}$ where $\mathbf{1} \triangleq [1 \quad 1 \quad \cdots \quad 1]'_{n\times 1}$. Naturally, there are situations where convergence to a common heading cannot occur. The most obvious of these is when one agent—say the ith—starts so far away from the rest that it never acquires any neighbors. Mathematically, this would mean not only that $\mathbb{G}_{\sigma(t)}$ is never

[1]The Vicsek system also includes noise input signals, which we ignore in this paper.

[2]By an undirected *graph* $\mathbb{G}$ on vertex set $\mathcal{V} = \{1, 2, \ldots n\}$ is meant $\mathcal{V}$ together with a set of unordered pairs $\mathcal{E} = \{(i, j)\colon i, j \in \mathcal{V}\}$ which are called $\mathbb{G}$'s edges. Such a graph is *simple* if it has no self-loops [i.e., $(i, j) \in \mathcal{E}$ only if $i \neq j$] or repeated edges (i.e., $\mathcal{E}$ contains only distinct elements). By the *valence* of a vertex v of $\mathbb{G}$ is meant the number of edges of $\mathbb{G}$ which are "incident" on v where by an *indicant* edge on v is meant an edge (i, j) of $\mathbb{G}$ for which either $i = v$ or $j = v$. The *adjacency matrix* of $\mathbb{G}$ is an $n \times n$ matrix of whose ijth entry is 1 if (i, j) is one of $\mathbb{G}$'s edges and 0 if it is not.

connected[3] at any time t, but also that vertex i remains an isolated vertex of $\mathbb{G}_{\sigma(t)}$ for all t. This situation is likely to be encountered if r is very small. At the other extreme, which is likely if r is very large, all agents might remain neighbors of all others for all time. In this case, σ would remain fixed along such a trajectory at that value in $p \in \mathcal{P}$ for which $\mathbb{G}_p$ is a complete graph. Convergence of θ to $\theta_{ss}\mathbf{1}$ can easily be established in this special case because with σ so fixed, (4) is a linear, time-invariant, discrete-time system. The situation of perhaps the greatest interest is between these two extremes when $\mathbb{G}_{\sigma(t)}$ is not necessarily complete or even connected for any $t \geq 0$, but when no strictly proper subset of $\mathbb{G}_{\sigma(t)}$'s vertices is isolated from the rest for all time. Establishing convergence in this case is challenging because σ changes with time and (4) is not time-invariant. It is this case which we intend to study. Toward this end, we denote by $\mathcal{Q}$ the subset of $\mathcal{P}$ consisting of the indices of the connected graphs in $\{\mathbb{G}_p\colon\ p \in \mathcal{P}\}$. Our first result establishes the convergence of θ for the case when σ takes values only in $\mathcal{Q}$.

Theorem 1: Let $\theta(0)$ be fixed and let $\sigma\colon \{0, 1, 2, \ldots\} \to \mathcal{P}$ be a switching signal satisfying $\sigma(t) \in \mathcal{Q}, t \in \{0, 1, \ldots\}$. Then

$$\lim_{t\to\infty} \theta(t) = \theta_{ss}\mathbf{1} \tag{5}$$

where θ_{ss} is a number depending only on $\theta(0)$ and σ.

It is possible to establish convergence to a common heading under conditions which are significantly less stringent that those assumed in Theorem 1. To do this we need to introduce several concepts. By the *union* of a collection of simple graphs, $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$, each with vertex set $\mathcal{V}$, is meant the simple graph $\mathbb{G}$ with vertex set $\mathcal{V}$ and edge set equaling the union of the edge sets of all of the graphs in the collection. We say that such a collection is *jointly connected* if the union of its members is a connected graph. Note that if such a collection contains at least one graph which is connected, then the collection must be jointly connected. On the other hand, a collection can be jointly connected even if none of its members are connected.

It is natural to say that the n agents under consideration are *linked together* across a time interval $[t, \tau]$ if the collection of graph $\{\mathbb{G}_{\sigma(t)}, \mathbb{G}_{\sigma(t+1)}, \ldots, \mathbb{G}_{\sigma(\tau)}\}$ encountered along the interval, is jointly connected. Theorem 1 says, in essence, that convergence of all agents' headings to a common heading is for certain provided all n agents are linked together across each successive interval of length one (i.e., all of the time). Of course there is no guarantee that along a specific trajectory the n agents will be so linked. Perhaps a more likely situation, at least when r is not too small, is when the agents are linked together across contiguous intervals of arbitrary but finite length. If the lengths of such intervals are uniformly bounded, then in this case too convergence to a common heading proves to be for certain.

[3]A simple graph $\mathbb{G}$ with vertex set $\mathcal{V} = \{1, 2, \ldots, n\}$ and edge set $\mathcal{E}$ is *connected* if has a "path" between each distinct pair of its vertices i and j where by a *path* (of *length* m) between vertices i and j is meant a sequence of distinct edges of $\mathbb{G}$ of the form $(i, k_1), (k_1, k_2), \ldots (k_m, j)$. $\mathbb{G}$ is *complete* if has a path of length one (i.e., an edge) between each distinct pair of its vertices.

Theorem 2: Let $\theta(0)$ be fixed and let $\sigma\colon \{0, 1, 2, \ldots\} \to \mathcal{P}$ be a switching signal for which there exists an infinite sequence of contiguous, nonempty, bounded, time-intervals $[t_i, t_{i+1})$, $i \geq 0$, starting at $t_0 = 0$, with the property that across each such interval, the n agents are linked together. Then

$$\lim_{t\to\infty} \theta(t) = \theta_{ss}\mathbf{1} \tag{6}$$

where θ_{ss} is a number depending only on $\theta(0)$ and σ.

The hypotheses of Theorem 2 require each of the collections $\{\mathbb{G}_{\sigma(t_i)}, \mathbb{G}_{\sigma(t_i+1)}, \ldots, \mathbb{G}_{\sigma(t_{i+1}-1)}\}$, $i \geq 0$, to be jointly connected. Although no constraints are placed on the intervals $[t_i, t_{i+1}), i \geq 0$, other than that they be of finite length, the constraint on σ is more restrictive than one might hope for. What one would prefer instead is to show that (6) holds for every switching signal σ for which there is an infinite sequence of bounded, nonoverlapping {but not necessarily contiguous} intervals across which the n agents are linked together. Whether or not this is true remains to be seen.

A sufficient but not necessary condition for σ to satisfy the hypotheses of Theorem 2 is that on each successive interval $[t_i, t_{i+1})$, σ take on at least one value in $\mathcal{Q}$. Theorem 1 is thus an obviously a consequence of Theorem 2 for the case when all intervals are of length 1. For this reason we need only develop a proof for Theorem 2. To do this we will make use of certain structural properties of the F_p. As defined, each F_p is square and nonnegative, where by a *nonnegative* matrix is meant a matrix whose entries are all nonnegative. 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 *stochastic* [28]. The F_p have the additional property that their diagonal elements are all nonzero. For the case when $p \in \mathcal{Q}$ (i.e., when $\mathbb{G}_p$ is connected), it is known that $(I + A_p)^m$ becomes a matrix with all positive entries for m sufficiently large [28]. It is easy to see that if $(I + A_p)^m$ has all positive entries, then so does F_p^m. Such $(I + A_p)$ and F_p are examples of "primitive matrices" where by a *primitive* matrix is meant any square, nonnegative matrix M for which M^m is a matrix with all positive entries for m sufficiently large [28]. It is known [28] that among the n eigenvalues of a primitive matrix, there is exactly one with largest magnitude, that this eigenvalue is the only one possessing an eigenvector with all positive entries, and that the remaining $n - 1$ eigenvalues are all strictly smaller in magnitude than the largest one. This means that for $p \in \mathcal{Q}$, 1 must be F_p's largest eigenvalue and all remaining eigenvalues must lie within the open unit circle. As a consequence, each such F_p must have the property that $\lim_{i\to\infty} F_p^i = \mathbf{1}c_p$ for some row vector c_p. Any stochastic matrices M for which $\lim_{i\to\infty} M^i$ is a matrix of rank 1 is called *ergodic* [28]. Primitive stochastic matrices are thus ergodic matrices. To summarize, each F_p is a stochastic matrix with positive diagonal elements and if $p \in \mathcal{Q}$ then F_p is also primitive and, as a result, ergodic. The crucial convergence result upon which the proof of Theorem 2 depends is classical [19] and is as follows.

Theorem 3 (Wolfowitz): Let $M_1, M_2, \ldots, M_m$ be a finite set of ergodic matrices with the property that for each sequence $M_{i_1}, M_{i_2}, \ldots, M_{i_j}$ of positive length, the matrix

product $M_{i_j}M_{i_{j-1}},\ldots,M_{i_1}$ is ergodic. Then for each infinite sequence, $M_{i_1}, M_{i_2}, \ldots$ there exists a row vector c such that

$$\lim_{j\to\infty} M_{i_j}M_{i_{j-1}}\cdots M_{i_1} = \mathbf{1}c.$$

The finiteness of the set $M_1, M_2, \ldots, M_m$ is crucial to Wolfowitz's proof. This finiteness requirement is also the reason why we've needed to assume *contiguous*, bounded intervals in the statement of Theorem 2.

In order to make use of Theorem 3, we need a few facts concerning products of the types of matrices we are considering. First, we point out that the class of $n \times n$ stochastic matrices with positive diagonal elements is closed under matrix multiplication. This is because the product of two nonnegative matrices with positive diagonals is a matrix with the same properties and because the product of two stochastic matrices is stochastic. Second, we will use the following key result.[4]

Lemma 1: Let $\{p_1, p_2, \ldots, p_m\}$ be a set of indices in $\mathcal{P}$ for which $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$ is a jointly connected collection of graphs. Then the matrix product $F_{p_1}F_{p_2},\ldots,F_{p_m}$ is ergodic.

Proof of Theorem 2[5]*:* Let T denote the least upper bound on the lengths of the intervals $[t_i, t_{i+1})$, $i \geq 0$. By assumption $T < \infty$. Let $\Phi(t, t) = I, t \geq 0$, and $\Phi(t, \tau) \triangleq F_{\sigma(t-1)}\cdots F_{\sigma(\tau+1)}F_{\sigma(\tau)}$, $t > \tau \geq 0$. Clearly $\theta(t) = \Phi(t, 0)\theta(0)$. To complete the theorem's proof, it is, therefore, enough to show that

$$\lim_{t\to\infty} \Phi(t, 0) = \mathbf{1}c \tag{7}$$

for some row vector c since this would imply (6) with $\theta_{ss} \triangleq c\theta(0)$. In view of Lemma 1, the constraints on σ imply that each such matrix product $\Phi(t_{j+1}, t_j)$, $j \geq 0$, is ergodic. Moreover the set of possible $\Phi(t_{j+1}, t_j)$, $j \geq 0$, must be finite because each $\Phi(t_{j+1}, t_j)$ is a product of at most T matrices from $\{F_p\colon\ p \in \mathcal{P}\}$ which is a finite set. But $\Phi(t_j, 0) = \Phi(t_j, t_{j-1})\Phi(t_{j-1}, t_{j-2}),\ldots,\Phi(t_1, t_0)$. Therefore, by Theorem 3

$$\lim_{j\to\infty} \Phi(t_j, 0) = \mathbf{1}c. \tag{8}$$

For each $t \geq 0$, let j_t be the largest nonnegative integer such that $t_{j_t} \leq t$. Then, $\Phi(t, 0) = \Phi(t, t_{j_t})\Phi(t_{j_t}, 0)$ and $\Phi(t, t_{j_t})\mathbf{1} = \mathbf{1}$ so

$$\Phi(t, 0) - \mathbf{1}c = \Phi(t, t_{j_t})(\Phi(t_{j_t}, 0) - \mathbf{1}c). \tag{9}$$

Note that $t \longmapsto \Phi(t, t_{j_t})$ is a bounded function because $\Phi(t, t_{j_t})$ is the product of at most $T-1$ matrices F_p which come from a bounded set. Moreover $(\Phi(t_{j_t}, 0) - \mathbf{1}c) \to 0$ as $t \to \infty$ because of (8). From this and (9), it follows that $(\Phi(t, 0) - \mathbf{1}c) \to 0$ as $t \to \infty$. Therefore, (7) holds. ∎

To prove Lemma 1 we shall make use of the standard partial ordering $\geq$ on $n \times n$ nonnegative matrices by writing $B \geq A$ whenever $B - A$ is nonnegative. Let us note that if A is a primitive matrix and if $B \geq A$, then B is primitive as well. Lemma 1 is a simple consequence of the following result.

Lemma 2: Let $m \geq 2$ be a positive integer and let $A_1, A_2, \ldots A_m$ be nonnegative $n \times n$ matrices. Suppose that the diagonal elements of all of the A_i are positive and let μ and ρ denote the smallest and largest of these, respectively. Then

$$A_1A_2\cdots A_m \geq \left(\frac{\mu^2}{2\rho}\right)^{(m-1)}(A_1 + A_2 + \cdots + A_m). \tag{10}$$

Proof: Set $\delta = \mu^2/2\rho$. It will be shown by induction that

$$A_1A_2\cdots A_i \geq \delta^{(i-1)}(A_1 + A_2 + \cdots + A_i) \tag{11}$$

holds for $i \in \{2, 3, \ldots, m\}$. Toward this end. note that it is possible to write each A_i as $A_i = \mu I + B_i$ where B_i is nonnegative. Then, for any $j, k \in \{1, 2, \ldots, m\}$

$$A_jA_k = (\mu I + B_j)(\mu I + B_k) = \mu^2 I + \mu(B_j + B_k) + B_jB_k.$$

Hence

$$\begin{aligned} A_jA_k &\geq \mu^2 I + \mu(B_j + B_k) \geq \mu^2 I + \frac{\mu^2}{2\rho}(B_j + B_k) \\ &= \delta((\rho I + B_j) + (\rho I + B_k)). \end{aligned}$$

Since $(\rho I + B_j) \geq A_j$ and $(\rho I + B_k) \geq A_k$ it follows that

$$A_jA_k \geq \delta(A_j + A_k) \qquad \forall j, k \in \{1, 2, \ldots, m\}. \tag{12}$$

Setting $j = 1$ and $k = 2$ proves that (11) holds for $i = 2$. If $m = 2$, the proof is complete.

Now, suppose that $m > 2$ and that (11) holds for $i \in \{2, 3, \ldots l\}$ where l is some integer in $\{2, 3, \ldots, m-1\}$. Then, $A_1A_2\cdots A_{l+1} = (A_1,\ldots,A_l)A_{l+1}$ so by the inductive hypothesis

$$A_1A_2\cdots A_{l+1} \geq \delta^{(l-1)}(A_1 + A_2 + \cdots + A_l)A_{l+1}. \tag{13}$$

However, using (12) l times, we can write

$$\begin{aligned} &(A_1 + A_2 + \cdots + A_l)A_{l+1} \\ &\quad \geq \delta\{(A_1 + A_{l+1}) + (A_2 + A_{l+1}) + \cdots + (A_l + A_{l+1})\}. \end{aligned}$$

Thus

$$(A_1 + A_2 + \cdots + A_l)A_{l+1} \geq \delta(A_1 + A_2 + \cdots + A_{l+1}).$$

This and (13) imply that (11) holds for $i = l + 1$. Therefore, by induction (11) is true for all $i \in \{2, 3, \ldots, m\}$. ∎

Proof of Lemma 1: Set $F = (I + D)^{-1}(I + A)$ where A and D are respectively the adjacency matrix and diagonal valence matrix of the union of the collection of graphs $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$. Since the collection is jointly connected, its union is connected which means that F is primitive. By Lemma 2

$$F_{p_1}F_{p_2}\cdots F_{p_m} \geq \gamma(F_{p_1} + F_{p_2} + \cdots + F_{p_m}) \tag{14}$$

where γ is a positive constant depending on the matrices in the product. Since for $i \in \{1, 2, \ldots, m\}$, $F_{p_i} = (I + D_{p_i})^{-1}(I +$

[4]We are indebted to M. Artzrouni, University of Pau, France, for his help with the proof of an earlier version of this lemma.

[5]The authors thank D. Liberzon for pointing out a flaw in the original version of this proof, and S. Meyn for suggesting how to fix it.

A_{p_i}) and $D \geq D_{p_i}$, it must be true that $F_{p_i} \geq (I+D)^{-1}(I+A_{p_i})$, $i \in \{1, 2, \ldots, m\}$. From this and (14) it follows that

$$F_{p_1}F_{p_2}\cdots F_{p_m} \geq \gamma(I+D)^{-1}(mI+A_{p_1}+A_{p_2}+\cdots+A_{p_m}). \tag{15}$$

However, $A_{p_1}+A_{p_2}+\cdots+A_{p_m} \geq A$ and $mI \geq I$ so

$$F_{p_1}F_{p_2}\cdots F_{p_m} \geq \gamma F.$$

Since the product $F_{p_1}F_{p_2}\cdots F_{p_m}$ is bounded below by a primitive matrix, namely γF, the product must be primitive as well. Since $F_{p_1}F_{p_2}\cdots F_{p_m}$ is also a stochastic matrix, it must therefore be ergodic. ■

As we have already noted, $F_p\mathbf{1}=\mathbf{1}$, $p \in \mathcal{P}$. Thus, span $\{\mathbf{1}\}$ is an F_p-invariant subspace. 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 $\mathbf{1}$, the equations

$$PF_p = \tilde{F}_p P, \qquad p \in \mathcal{P} \tag{16}$$

have unique solutions $\tilde{F}_p$, $p \in \mathcal{P}$, and moreover that

$$\text{spectrum } F_p = \{1\} \cup \text{spectrum } \tilde{F}_p, \qquad p \in \mathcal{P}. \tag{17}$$

As a consequence of (16) it can easily be seen that for any sequence of indexes $p_0, p_1, \ldots, p_i$ in $\mathcal{P}$

$$\tilde{F}_{p_i}\tilde{F}_{p_{i-1}},\ldots,\tilde{F}_{p_0}P = PF_{p_i}F_{p_{i-1}},\ldots,F_{p_0}. \tag{18}$$

Since P has full-row rank and $P\mathbf{1}=0$, the convergence of a product of the form $F_{p_i}F_{p_{i-1}},\ldots,F_{p_0}$ to $\mathbf{1}c$ for some row vector c, is equivalent to convergence of the corresponding product $\tilde{F}_{p_i}\tilde{F}_{p_{i-1}},\ldots,\tilde{F}_{p_0}$ to the zero matrix. Thus, for example, if $p_0, p_1, \ldots$ is an infinite sequence of indices in $\mathcal{Q}$, then, in view of Theorem 3

$$\lim_{i\to\infty} \tilde{F}_{p_i}\tilde{F}_{p_{i-1}},\ldots,\tilde{F}_{p_0} = 0. \tag{19}$$

Some readers might be tempted to think, as we first did, that the validity of (19) could be established directly by showing that the $\tilde{F}_p$ in the product share a common quadratic Lyapunov function. More precisely, (19) would be true if there were a single positive–definite matrix M such that all of the matrices $\tilde{F}_p'M\tilde{F}_p - M$, $p \in \mathcal{Q}$ were negative definite. Although each $\tilde{F}_p$, $p \in \mathcal{Q}$ can easily be shown to be discrete-time stable, there are classes of F_p for which that no such common Lyapunov matrix M exists. While we have not been able to construct a simple analytical example which demonstrates this, we have been able to determine, for example, that no common quadratic Lyapunov function exists for the class of all F_p whose associated graphs have 10 vertices and are connected. One can verify that this is so by using semidefinite programming and restricting the check to just those connected graphs on ten vertices with either nine or ten edges.

It is worth noting that existence of a common quadratic Lyapunov function for all discrete time stable $m \times m$ matrices $M_1, M_2, \ldots$ in some given finite set $\mathcal{M}$, is a much stronger condition than is typically needed to guarantee that all infinite products of the M_i converge to zero. It is known [29] that convergence to zero of all such infinite products is in fact equivalent to the "joint spectral radius" of $\mathcal{M}$ being strictly less than 1 where by *joint spectral radius of* $\mathcal{M}$ is meant

$$\rho_{\mathcal{M}} := \limsup_{k\to\infty}\left\{\max_{M_{i_1}\in\mathcal{M}}\max_{M_{i_2}\in\mathcal{M}} \cdots \max_{M_{i_k}\in\mathcal{M}} \|M_{i_1}M_{i_2},\ldots,M_{i_k}\|\right\}^{1/k}.$$

Here, $\|\cdot\|$ is any norm on the space of real $m\times m$ matrices. It turns out that $\rho_{\mathcal{M}}$ does not depend on the choice of norm [21, p. 237]. On the other hand, a "tight" sufficient condition for the existence of a common quadratic Lyapunov function for the matrices in $\mathcal{M}$, is $\rho_{\mathcal{M}} \in [0, 1/\sqrt{m})$ [30]. This condition is *tight* in the sense that one can find a finite set of $m\times m$ matrices with joint spectral radius $\rho = 1/\sqrt{m}$, whose infinite products converge to zero despite the fact that there does not exist common quadratic Lyapunov function for the set. From this one can draw the conclusion that sets of matrices with "large" m are not likely to possess a common quadratic, even though all infinite products of such matrices converge to zero. This can in turn help explain why it has proved to be necessary to go as high as $n=10$ to find a case where a common quadratic Lyapunov function for a family of F_p does not exist.

A. Generalization

It is possible to interpret the Vicsek model analyzed in the last section as the closed-loop system which results when a suitably defined decentralized feedback law is applied to the n-agent heading model

$$\theta(t+1) = \theta(t) + u(t) \tag{20}$$

with open-loop control u. To end up with the Vicsek model, u would have to be defined as

$$u(t) = -(I+D_{\sigma(t)})^{-1}e(t) \tag{21}$$

where e is the *average heading error* vector

$$e(t) \triangleq L_{\sigma(t)}\theta(t) \tag{22}$$

and, for each $p \in \mathcal{P}$, L_p is the symmetric matrix

$$L_p = D_p - A_p \tag{23}$$

known in graph theory as the *Laplacian* of $\mathbb{G}_p$ [27], [31]. It is easily verified that equations (20) to (21) do indeed define the Vicsek model. We have elected to call e the average heading error because if $e(t)=0$ at some time t, then the heading of each agent with neighbors at that time will equal the average of the headings of its neighbors.

In the present context, Vicsek's control (21) can be viewed as a special case of a more general decentralized feedback control of the form

$$u(t) = -G_{\sigma(t)}^{-1}L_{\sigma(t)}\theta(t) \tag{24}$$

where for each $p \in \mathcal{P}$, G_p is a suitably defined, nonsingular diagonal matrix with ith diagonal element g_p^i. This, in turn, is an

abbreviated description of a system of n individual agent control laws of the form

$$u_i(t) = -\frac{1}{g_i(t)}\left(n_i(t)\theta_i(t) + \sum_{j\in\mathcal{N}_i(t)}\theta_j(t)\right) \quad i \in \{1, 2, \ldots, n\} \tag{25}$$

where for $i \in \{1, 2, \ldots, n\}$, $u_i(t)$ is the ith entry of $u(t)$ and $g_i(t) \triangleq g^i_{\sigma(t)}$. Application of this control to (20) would result in the closed-loop system

$$\theta(t+1) = \theta(t) - G^{-1}_{\sigma(t)}L_{\sigma(t)}\theta(t). \tag{26}$$

Note that the form of (26) implies that if θ and σ were to converge to a constant values $\overline{\theta}$, and $\overline{\sigma}$, respectively, then $\overline{\theta}$ would automatically satisfy $L_{\overline{\sigma}}\overline{\theta} = 0$. This means that control (24) automatically forces each agent's heading to converge to the average of its neighbors, if agent headings were to converge at all. In other words, the choice of the G_p does not effect the requirement that each agent's heading equal the average of the headings of its neighbors, if there is convergence at all.

The preceding suggests that there might be useful choices for the G_p alternative to those considered by Vicsek, which also lead to convergence. One such choice turns out to be

$$G_p = gI, \qquad p \in \mathcal{P} \tag{27}$$

where g is any number greater than n. Our aim is to show that with the G_p so defined, Theorem 2 continues to be valid. In sharp contrast with the proof technique used in the last section, convergence will be established here using a common quadratic Lyapunov function.

As before, we will use the model

$$\theta(t+1) = F_{\sigma(t)}\theta(t) \tag{28}$$

where, in view of the definition of the G_p in (27), the F_p are now symmetric matrices of the form

$$F_p = I - \frac{1}{g}L_p, \qquad p \in \mathcal{P}. \tag{29}$$

To proceed we need to review a number of well known and easily verified properties of graph Laplacians relevant to the problem at hand. For this, let $\mathbb{G}$ be any given simple graph with n vertices. Let D be a diagonal matrix whose diagonal elements are the valences of $\mathbb{G}$'s vertices and write A for $\mathbb{G}$'s adjacency matrix. Then, as noted before, the Laplacian of $\mathbb{G}$ is the symmetric matrix $L = D - A$. The definition of L clearly implies that $L\mathbf{1} = 0$. Thus, L must have an eigenvalue at zero and $\mathbf{1}$ must be an eigenvector for this eigenvalue. Surprisingly L is always a positive semidefinite matrix [31]. Thus, L must have a real spectrum consisting of nonnegative numbers and at least one of these numbers must be 0. It turns out that the number of connected components of $\mathbb{G}$ is exactly the same as the multiplicity of L's eigenvalue at 0 [31]. Thus, $\mathbb{G}$ is a connected graph just in case L has exactly one eigenvalue at 0. Note that the trace of L is the sum of the valences of all vertices of $\mathbb{G}$. This number can never exceed $(n-1)n$ and can attain this high value only for a complete graph. In any event, this property implies that the maximum eigenvalue of L is never larger that $n(n-1)$. Actually, the largest eigenvalue of L can never be larger than n [31]. This means that the eigenvalues of $(1/g)L$ must be smaller than 1 since $g > n$. From these properties it clearly follows that the eigenvalues of $(I - (1/g)L)$ must all be between 0 and 1, and that if $\mathbb{G}$ is connected, then all will be strictly less than 1 except for one eigenvalue at 1 with eigenvector $\mathbf{1}$. Since each F_p is of the form $(I-(1/g)L)$, each F_p possesses all of these properties.

Let σ be a fixed switching signal with value $p_t \in \mathcal{Q}$ at time $t \geq 0$. What we'd like to do is to prove that as $i \to \infty$, the matrix product $F_{p_i}F_{p_{i-1}}, \ldots, F_{p_0}$ converges to $\mathbf{1}c$ for some row vector c. As noted in the Section II-A, this matrix product will so converge just in case

$$\lim_{i\to\infty}\tilde{F}_{p_i}\tilde{F}_{p_{i-1}}, \ldots, \tilde{F}_{p_0} = 0 \tag{30}$$

where as in Section II-A, $\tilde{F}_p$ is the unique solution to $PF_p = \tilde{F}_pP$, $p \in \mathcal{P}$ and P is any full rank $(n-1)\times n$ matrix satisfying $P\mathbf{1} = 0$. For simplicity and without loss of generality we shall henceforth assume that the rows of P form a basis for the orthogonal complement of the span of $\mathbf{e}$. This means that PP' equals the $(n-1)\times(n-1)$ identity $\tilde{I}$, that $\tilde{F}_p = PF_pP'$, $p \in \mathcal{P}$ and, thus, that each $\tilde{F}_p$ is symmetric. Moreover, in view of (17) and the spectral properties of the F_p, $p \in \mathcal{Q}$, it is clear that each $\tilde{F}_p$, $p \in \mathcal{Q}$, must have a real spectrum lying strictly inside of the unit circle. This plus symmetry means that for each $p \in \mathcal{Q}$, $\tilde{F}_p - \tilde{I}$ is negative definite, that $\tilde{F}_p'\tilde{F}_p - \tilde{I}$ is negative definite and thus, that $\tilde{I}$ is a common discrete-time Lyapunov matrix for all such $\tilde{F}_p$. Using this fact it is straight forward to prove that Theorem 1 holds for system (26) provided the G_p are defined as in (27) with $g > n$.

In general, each $\tilde{F}_p$ is a discrete-time stability matrix for which $\tilde{F}_p'\tilde{F}_p - \tilde{I}$ is negative definite only if $p \in \mathcal{Q}$. To craft a proof of Theorem 2 for the system described by (26) and (27), one needs to show that for each interval $[t_i, t_{i+1})$ on which $\{\mathbb{G}_{\sigma(t_{i+1}-1)}, \ldots \mathbb{G}_{\sigma(t_i+1)}, \mathbb{G}_{\sigma(t_i)}\}$ is a jointly connected collection of graphs, the product $\tilde{F}_{\sigma(t_{i+1}-1)}\cdots\tilde{F}_{\sigma(t_i+1)}\tilde{F}_{\sigma(t_i)}$ is a discrete-time stability matrix and

$$(\tilde{F}_{\sigma(t_{i+1}-1)}, \ldots, \tilde{F}_{\sigma(t_i+1)}\tilde{F}_{\sigma(t_i)})'(\tilde{F}_{\sigma(t_{i+1}-1)}\cdots\tilde{F}_{\sigma(t_i)}) - \tilde{I}$$

is negative definite. This is a direct consequence of the following proposition.

Proposition 1: If $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$ is a jointly connected collection of graphs, then

$$(\tilde{F}_{p_1}\tilde{F}_{p_2}, \ldots, \tilde{F}_{p_m})'(\tilde{F}_{p_1}\tilde{F}_{p_2}, \ldots, \tilde{F}_{p_m}) - \tilde{I}$$

is a negative–definite matrix.

In the light of Proposition 1, it is clear that the conclusion Theorem 2 is also valid for the system described by (26) and (27). A proof of this version of Theorem 2 will not be given.

To summarize, both the Vicsek control defined by $u = -(I + D_{\sigma(t)})^{-1}e(t)$ and the simplified control given by $u = -(1/g)e(t)$ achieve the same emergent behavior. While latter is much easier to analyze than the former, it has the disadvantage of not being a true decentralized control because each agent must know an upper bound (i.e., g) on the total number of agents within the group. Whether or not this is really a disadvantage, of course depends on what the models are to be used for.

The proof of Proposition 1 depends on two lemmas. In the sequel, we state the lemmas, use them to prove Proposition 1, and then conclude this section with proofs of the lemmas themselves.

Lemma 3: If $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$ is a jointly connected collection of graphs with Laplacians $L_{p_1}, L_{p_2}, \ldots, L_{p_m}$, then

$$\bigcap_{i=1}^{m} \text{kernel } L_{p_i} = \text{span } \{\mathbf{1}\}. \tag{31}$$

Lemma 4: Let $M_1, M_2, \ldots, M_m$ be a set of $n \times n$ real symmetric, matrices whose induced 2-norms are all less than or equal to 1. If

$$\bigcap_{i=1}^{m} \text{kernel } (I - M_i) = 0 \tag{32}$$

then the induced two-norm of $M_1 M_2, \ldots, M_m$ is less than 1.

Proof of Proposition 1: The definition of the F_p in (29) implies that $I - F_p = (1/g)L_p$. Hence, by Lemma 3 and the hypothesis that $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$ is a jointly connected collection

$$\bigcap_{i=1}^{m} \text{kernel } (I - F_{p_i}) = \text{span } \{\mathbf{1}\}. \tag{33}$$

We claim that

$$\bigcap_{i=1}^{m} \text{kernel } (\tilde{I} - \tilde{F}_{p_i}) = 0. \tag{34}$$

To establish this fact, let $\overline{x}$ be any vector such that $(\tilde{I} - \tilde{F}_{p_i})\overline{x} = 0, i \in \{1, 2, \ldots, m\}$. Since P has independent rows, there is a vector x such that $\overline{x} = Px$. But $P(I - F_{p_i}) = (\tilde{I} - \tilde{F}_{p_i})P$, so $P(I - F_{p_i})x = 0$. Hence, $(I - F_{p_i})x = a_i\mathbf{1}$ for some number a_i. But $\mathbf{1}'(I - F_{p_i}) = (1/g)\mathbf{1}'L_{p_i} = 0$, so $a_i\mathbf{1}'\mathbf{1} = 0$. This implies that $a_i = 0$ and, thus, that $(I - F_{p_i})x = 0$. However, this must be true for all $i \in \{1, 2, \ldots, m\}$. It follows from (33) that $x \in \text{span } \{\mathbf{1}\}$ and, since $\overline{x} = Px$, that $\overline{x} = 0$. Therefore, (34) is true.

As defined, the $\tilde{F}_p$ are all symmetric, positive semi-definite matrices with induced 2-norms not exceeding 1. This and (34) imply that the family of matrices $\tilde{F}_{p_1}, \tilde{F}_{p_2}, \ldots, \tilde{F}_{p_m}$ satisfy the hypotheses of Lemma 4. It follows that Proposition 1 is true. ■

Proof of Lemma 3: In the sequel we write $L(\mathbb{G})$ for the Laplacian of a simple graph $\mathbb{G}$. By the *intersection* of a collection of simple graphs, $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$, each with vertex set $\mathcal{V}$, is meant the simple graph $\mathbb{G}$ with vertex set $\mathcal{V}$ and edge set equaling the intersection of the edge sets of all of the graphs in the collection. It follows at once from the definition of a Laplacian that

$$L(\mathbb{G}_p) + L(\mathbb{G}_q) = L(\mathbb{G}_p \cap \mathbb{G}_q) + L(\mathbb{G}_p \cup \mathbb{G}_q)$$

for all $p, q \in \mathcal{P}$. Repeated application of this identity to the set $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$ yields the relation

$$\sum_{i=1}^{m} L(\mathbb{G}_{p_i}) = L\left(\bigcup_{i=1}^{m} \mathbb{G}_{p_i}\right) + \sum_{i=1}^{m-1} L\left(\mathbb{G}_{p_{i+1}} \bigcap \left\{\bigcup_{j=1}^{i} \mathbb{G}_{p_j}\right\}\right) \tag{35}$$

which is valid for $m > 1$. Since all matrices in (35) are positive semidefinite, any vector x which makes the quadratic form $x'\{L(\mathbb{G}_{p_1}) + L(\mathbb{G}_{p_2}) + \cdots + L(\mathbb{G}_{p_m})\}x$ vanish, must also make the quadratic form $x'L(\mathbb{G}_{p_1} \cup \mathbb{G}_{p_2} \cup \cdots \cup \mathbb{G}_{p_m})x$ vanish. Since any vector in the kernel of each matrix $L(\mathbb{G}_{p_i})$ has this property, we can draw the following conclusion:

$$\bigcap_{i=1}^{m} \text{kernel } L(\mathbb{G}_{p_i}) \subset \text{kernel } L\left(\bigcup_{i=1}^{m} \mathbb{G}_{p_i}\right).$$

Suppose now that $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$ is a jointly connected collection. Then the union $\mathbb{G}_{p_1} \cup \mathbb{G}_{p_2} \cup \cdots \cup \mathbb{G}_{p_m}$ is connected so its Laplacian must have exactly span $\{\mathbf{1}\}$ for its kernel. Hence, the intersection of the kernels of the $L(\mathbb{G}_{p_i})$ must be contained in span $\{\mathbf{1}\}$. But span $\{\mathbf{1}\}$ is contained in the kernel of each matrix $L(\mathbb{G}_{p_i})$ in the intersection and, therefore, in the intersection of the kernels of these matrices as well. It follows that (31) is true. ■

Proof of Lemma 4: In the sequel we write $|x|$ for the 2-norm of a real n-vector x and $|M|$ for the induced 2-norm of a real $n \times n$ matrix. Let $x \in \mathbb{R}^n$ be any real, nonzero n-vector. It is enough to show that

$$|M_1 M_2 \cdots M_m x| < |x|. \tag{36}$$

In view of (32) and the assumption that $x \neq 0$, there must be a largest integer $k \in \{1, 2, \ldots, m\}$ such that $x \notin \text{kernel } (M_k - I)$. We claim that

$$|M_k x| < |x|. \tag{37}$$

To show that this is so we exploit the symmetry of M_k to write x as $x = \alpha_1 y_1 + \alpha_2 y_2 + \cdots + \alpha_n y_n$ where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are real numbers and $\{y_1, y_2, \ldots, y_n\}$ is an orthonormal set of eigenvectors of M_k with real eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$. Note that $|\lambda_i| \leq 1, i \in \{1, 2, \ldots, n\}$, because $|M_k| \leq 1$. Next, observe that since $M_k x = \alpha_1\lambda_1 y_1 + \alpha_2\lambda_2 y_2 + \cdots + \alpha_n\lambda_n y_n$ and $M_k x \neq x$, there must be at least one integer j such that $\alpha_j\lambda_j \neq \alpha_j$. Hence, $|\alpha_j\lambda_j y_j| < |\alpha_j y_j|$. However, $|M_k x|^2 = |\alpha_1\lambda_1 y_1|^2 + \cdots + |\alpha_j\lambda_j y_j|^2 + \cdots + |\alpha_n\lambda_n y_n|^2$ so

$$|M_k x|^2 < |\alpha_1\lambda_1 y_1|^2 + \cdots + |\alpha_j y_j|^2 + \cdots + |\alpha_n\lambda_n y_n|^2.$$

Moreover

$$|\alpha_1\lambda_1 y_1|^2 + \cdots + |\alpha_j y_j|^2 + \cdots + |\alpha_n\lambda_n y_n|^2 \leq |\alpha_1 y_1|^2 + \cdots + |\alpha_j y_j|^2 + \cdots + |\alpha_n y_n|^2 = |x|^2$$

so $|M_k x|^2 < |x|^2$; therefore, (37) is true.

In view of the definition of k, $M_j x = x$, $j \in \{k+1, \ldots, m\}$. From this and (37), it follows that $|M_1 \cdots M_m x| = |M_1 \cdots M_k x| \leq |M_1 \cdots M_{k-1}||M_k x| < |M_1 \cdots M_{k-1}||x|$. However, $|M_1 \cdots M_{k-1}| \leq 1$ because each M_i has an induced two-norm not exceeding 1. Therefore, (36a) is true. ■

III. Leader Following

In this section, we consider a modified version of Vicsek's discrete-time system consisting of the same group of n agents as before, plus one additional agent, labeled 0, which acts as the group's leader. Agent 0 moves at the same constant speed as its n followers but with a fixed heading θ_0. The ith follower updates

its heading just as before, using the average of its own heading plus the headings of its neighbors. The difference now is that each follower's set of neighbors can include the leader and does so whenever the leader is within the follower's neighborhood defining circle of radius r. Agent i's update rule, thus, is of the form

$$\theta_i(t+1) = \frac{1}{1+n_i(t)+b_i(t)} \cdot \left(\theta_i(t) + \sum_{j\in\mathcal{N}_i(t)} \theta_j(t) + b_i(t)\theta_0\right) \quad (38)$$

where as before, $\mathcal{N}_i(t)$ is the set of labels of agent i's neighbors from the original group of n followers, and $n_i(t)$ is the number of labels within $\mathcal{N}_i(t)$. Agent 0's heading is accounted for in the ith average by defining $b_i(t)$ to be 1 whenever agent 0 is a neighbor of agent i and 0 otherwise.

The explicit form of the n update equations exemplified by (38), depends on the relationships between neighbors which exist at time t. Like before, each of these relationships can be conveniently described by a simple undirected graph. In this case, each such graph has vertex set $\{0, 1, 2, \ldots, n\}$ and is defined so that (i, j) is one of the graph's edges just in case agents i and j are neighbors. For this purpose we consider an agent—say i—to be a neighbor of agent 0 whenever agent 0 is a neighbor of agent i. We will need to consider all possible such graphs. In the sequel we use the symbol $\overline{\mathcal{P}}$ to denote a set indexing the class of all simple graphs $\overline{\mathbb{G}}_p$ defined on vertices $0, 1, 2, \ldots, n$. We will also continue to make reference to the set of all simple graphs on vertices $1, 2, \ldots, n$. Such graphs are now viewed as subgraphs of the $\overline{\mathbb{G}}_p$. Thus, for $p \in \overline{\mathcal{P}}$, $\mathbb{G}_p$ now denotes the subgraph obtained from $\overline{\mathbb{G}}_p$ by deleting vertex 0 and all edges incident on vertex 0.

The set of agent heading update rules defined by (38) can be written in state form. Toward this end, for each $p \in \overline{\mathcal{P}}$, let A_p denote the $n \times n$ adjacency matrix of the n-agent graph $\mathbb{G}_p$ and let D_p be the corresponding diagonal matrix of valences of $\mathbb{G}_p$. Then, in matrix terms, (38) becomes

$$\theta(t+1) = (I + D_{\sigma(t)} + B_{\sigma(t)})^{-1}((I + A_{\sigma(t)})\theta(t) + B_{\sigma(t)}\mathbf{1}\theta_0), \qquad t \in \{0, 1, 2, \ldots\} \quad (39)$$

where σ: $\{0, 1, \ldots\} \rightarrow \overline{\mathcal{P}}$ is now a switching signal whose value at time t, is the index of the graph $\overline{\mathbb{G}}_p$ representing the agent system's neighbor relationships at time t and for $p \in \overline{\mathcal{P}}$, B_p is the $n \times n$ diagonal matrix whose ith diagonal element is 1, if $(i, 0)$ is one of $\overline{\mathbb{G}}_p$'s edges and 0, otherwise.

Much like before, our goal here is to show for a large class of switching signals and for any initial set of follower agent headings, that the headings of all n followers converge to the heading of the leader. For convergence in the leaderless case we required all n-agents to be linked together across each interval within an infinite sequence of contiguous, bounded intervals. We will need a similar requirement in the leader following case under consideration. Let us agree to say that the n agents are *linked to the leader* across an interval $[t, \tau]$ if the collection of graphs $\{\overline{\mathbb{G}}_{\sigma(t)}, \overline{\mathbb{G}}_{\sigma(t+1)}, \ldots, \overline{\mathbb{G}}_{\sigma(\tau)}\}$ encountered along the interval is jointly connected. In other words, the n agents are linked to their leader across an interval $\mathcal{I}$ just when the $n+1$-member group consisting of the n agents and their leader is linked together across $\mathcal{I}$. Note that for the n-agent group to be linked to its leader across $\mathcal{I}$ does not mean that the n-agent group must be linked together across $\mathcal{I}$. Nor is the n-agent group necessarily linked to its leader across $\mathcal{I}$ when it is linked together across $\mathcal{I}$. Our main result on discrete-time leader following is next.

Theorem 4: Let $\theta(0)$ and θ_0 be fixed and let σ: $\{0, 1, 2, \ldots\} \rightarrow \overline{\mathcal{P}}$ be a switching signal for which there exists an infinite sequence of contiguous, nonempty, bounded, time-intervals $[t_i, t_{i+1})$, $i \geq 0$, starting at $t_0 = 0$, with the property that across each such interval, the n-agent group of followers is linked to its leader. Then

$$\lim_{t\rightarrow\infty} \theta(t) = \theta_0\mathbf{1}. \quad (40)$$

The theorem says that the members of the n-agent group all eventually follow their leader provided there is a positive integer T which is large enough so that the n-agent group is linked to its leader across each contiguous, nonempty time-interval of length at most T. In the sequel, we outline several ideas upon which the proof of Theorem 4 depends.

To begin, let us note that to prove that (40) holds is equivalent to proving that $\lim_{t\rightarrow\infty} \epsilon(t) \rightarrow 0$ where ϵ is the heading error vector $\epsilon(t) \triangleq \theta(t) - \theta_0\mathbf{1}$. From (39) it is easy to deduce that ϵ satisfies the equation

$$\epsilon(t+1) = F_{\sigma(t)}\epsilon(t) \quad (41)$$

where for $p \in \overline{\mathcal{P}}$, F_p is

$$F_p = (I + D_p + B_p)^{-1}(I + A_p). \quad (42)$$

Note that the partitioned matrices

$$\overline{F}_p \triangleq \begin{bmatrix} F_p & H_p\mathbf{1} \\ 0 & 1 \end{bmatrix}, \qquad p \in \overline{\mathcal{P}} \quad (43)$$

are stochastic where, for $p \in \overline{\mathcal{P}}$

$$H_p \triangleq (I + D_p + B_p)^{-1}B_p. \quad (44)$$

To proceed, we need a few more ideas concerned with nonnegative matrices. In the sequel, we write $M > N$ whenever $M - N$ is a positive matrix, where by a *positive matrix* is meant a matrix with all positive entries. For any nonnegative matrix R of any size, we write $\|R\|$ for the largest of the row sums of R. Note that $\|R\|$ is the induced infinity norm of R and consequently is sub-multiplicative. We denote by $\lceil R \rceil$, the matrix obtained by replacing all of R's nonzero entries with a 1. Note that $R > 0$ if and only if $\lceil R \rceil > 0$. It is also true for any pair of $n \times n$ nonnegative matrices A and B with positive diagonal elements, that $\lceil AB \rceil = \lceil \lceil A \rceil \lceil B \rceil \rceil$. Moreover, in view of Lemma 2, any such pair of matrices must also satisfy $\lceil AB \rceil \geq \lceil B \rceil$ and $\lceil BA \rceil \geq \lceil B \rceil$.

Let $p_1, p_2, \ldots, p_m$ be a given set of indices in $\overline{\mathcal{P}}$. It is possible to relate the connectedness of the collection $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_m}\}$ to properties of the matrix pairs $(F_{p_i}, H_{p_i}\mathbf{1})$, $i \in \{1, 2, \ldots, m\}$. Let us note first that for any $p \in \overline{P}$, the indices of the nonzero rows of $B_p\mathbf{1}$ are precisely the labels of vertices in $\overline{\mathbb{G}}_p$ which are connected to vertex 0

by paths of length 1. More generally, for any integer $j > 0$, the indices of the nonzero rows of $(I + A_p)^{(j-1)}B_p\mathbf{1}$ are the labels of vertices in $\overline{\mathbb{G}}_p$ connected to vertex 0 by paths of length less than or equal to j. Hence, for such j, the nonzero rows of the sum $\sum_{k=1}^{m}(I + A_{p_k})^{(j-1)}B_{p_k}\mathbf{1}$ must be the labels of vertices in the union of the collection $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_m}\}$ which are connected to vertex 0 by paths of length less than or equal to j. It follows that if $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_m}\}$ is jointly connected, there must be a value of j sufficiently large so that $\sum_{k=1}^{m}(I + A_{p_k})^{(j-1)}B_{p_k}\mathbf{1} > 0$. Since any vertex in a connected graph with $n + 1$ vertices is reachable from any other vertex along a path of length of at most $n + 1$, it follows that if $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_m}\}$ is jointly connected, then $\sum_{k=1}^{m}(I + A_{p_k})^{(i-1)}B_{p_k}\mathbf{1} > 0, \forall\, i > n$. Now it is easy to see from the definitions of the F_p and H_p in (42) and (44) respectively, that $\lceil F_p^j H_p\mathbf{1}\rceil = \lceil (I + A_p)^j B_p\mathbf{1}\rceil$, $j \geq 0$. We have proved the following lemma.

Lemma 5: Let $\{p_1, p_2, \ldots, p_m\}$ be any set of indices in $\overline{\mathcal{P}}$ for which $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_m}\}$ is a jointly connected collection of graphs. Then

$$\sum_{k=1}^{m} F_{p_k}^{(i-1)} H_{p_k}\mathbf{1} > 0, \qquad i > n. \tag{45}$$

Now, consider the partitioned matrices $\overline{F}_p$ defined by (43). Since each of these matrices is stochastic and products of stochastic matrices are also stochastic, for each $p \in \overline{\mathcal{P}}$ and each $i \geq 1$, $\overline{F}_p^{\,i}$ is stochastic. However

$$\overline{F}_p^{\,i} = \begin{bmatrix} F_p^i & \displaystyle\sum_{j=1}^{i} F_p^{(j-1)} H_p\mathbf{1} \\ 0 & 1 \end{bmatrix}, \qquad p \in \overline{\mathcal{P}}.$$

Moreover, if $\overline{\mathbb{G}}_p$ is connected, then

$$\sum_{j=1}^{i} F_p^{(j-1)} H_p\mathbf{1} > 0, \qquad i > n \tag{46}$$

because of Lemma 5. It follows that if $\overline{\mathbb{G}}_p$ is connected and $i > n$, the row sums of F_p^i must all be less that 1. In other words

$$\|F_p^i\| < 1, \qquad i > n,\ p \in \overline{\mathcal{Q}}. \tag{47}$$

The following proposition generalizes (47) and is central to the proof of Theorem 4.

Proposition 2: Let T be a finite positive integer. There exists a positive number $\lambda < 1$, depending only on T, for which

$$\|F_{p_{\overline{t}}} F_{p_{\overline{t}-1}} \cdots F_{p_1}\| < \lambda \tag{48}$$

for every sequence $p_1, p_2, \ldots p_{\overline{t}} \in \overline{\mathcal{P}}$ of at length at most T possessing values $q_1, q_2, \ldots, q_m$ which each occur in the sequence at least $n + 1$ times and for which $\{\overline{\mathbb{G}}_{q_1}, \overline{\mathbb{G}}_{q_2}, \ldots, \overline{\mathbb{G}}_{q_m}\}$ is a jointly connected collection of graphs.

The proof of this proposition depends on the following basic property of nonnegative matrices.

Lemma 6: Let $M_1, M_2, \ldots, M_k$ be a finite sequence of $n \times n$ nonnegative matrices whose diagonal entries are all positive. Suppose that M is a matrix which occurs in the sequence at least $m > 0$ times. Then

$$\lceil M_1 M_2 \cdots M_k\rceil \geq \lceil M^m\rceil. \tag{49}$$

Proof: We claim that for $j \geq 1$

$$\lceil M_1 M_2 \cdots M_{k_j}\rceil \geq \lceil M^j\rceil \tag{50}$$

provided $M_1 M_2 \cdots M_{k_j}$ is a product within which M occurs at least j times. Suppose $M_1 M_2 \cdots M_{k_1}$ is a product within which M occurs at least once. Then $M_1 M_2 \cdots M_{k_1} = AMB$ where A and B are nonnegative matrices with positive diagonal elements. By Lemma 2, $\lceil AMB\rceil \geq \lceil MB\rceil$ and $\lceil MB\rceil \geq \lceil M\rceil$. Thus, $\lceil AMB\rceil \geq \lceil M\rceil$ which proves that (50) is true for $j = 1$.

Now suppose that (50) holds for $j \in \{1, 2, \ldots i\}$ and let $M_1 M_2 \cdots M_{k_{i+1}}$ be a product within which M occurs at least $i + 1$ times. We can write $M_1 M_2 \cdots M_{k_{i+1}} = AMB$ where A and B are nonnegative matrices with positive diagonal elements and A is a product within which M occurs at least i times. By the inductive hypothesis, $\lceil A\rceil \geq \lceil M^i\rceil$. By Lemma 2, $\lceil AMB\rceil \geq \lceil AM\rceil$. It follows that $\lceil AM\rceil = \lceil\lceil A\rceil\lceil M\rceil\rceil \geq \lceil\lceil M^i\rceil\lceil M\rceil\rceil = \lceil M^{i+1}\rceil$ and thus that (50) holds for $j = i+1$. By induction, (50) therefore holds for all $i \in \{1, 2, \ldots, m\}$. Hence, the lemma is true. ■

Proof of Proposition 2: It will be enough to prove that

$$\|F_{p_{\overline{t}}} F_{p_{\overline{t}-1}} \cdots F_{p_1}\| < 1 \tag{51}$$

for every sequence $p_1, p_2, \ldots p_{\overline{t}}$ of length at most T possessing values $q_1, q_2, \ldots, q_m$ which each occur in the sequence at least $n + 1$ times and for which $\{\overline{\mathbb{G}}_{q_1}, \overline{\mathbb{G}}_{q_2}, \ldots, \overline{\mathbb{G}}_{q_m}\}$ is a jointly connected collection of graphs. For if this is so, then one can define the uniform bound

$$\lambda \triangleq \max_{\mathcal{S}} \|F_{p_{\overline{t}}} F_{p_{\overline{t}-1}} \cdots F_{p_1}\|$$

where $\mathcal{S}$ is the set of all such sequences. Note that $\lambda < 1$ if (51) holds, because $\mathcal{S}$ is a finite set.

Let $p_1, p_2, \ldots p_{\overline{t}}$ be a sequence of at length at most T possessing values $q_1, q_2, \ldots, q_m$ which each occur in the sequence at least $n + 1$ times and for which $\{\overline{\mathbb{G}}_{q_1}, \overline{\mathbb{G}}_{q_2}, \ldots, \overline{\mathbb{G}}_{q_m}\}$ is a jointly connected collection of graphs. The definition of the $\overline{F}_p$ in (43) implies that

$$\overline{F}_{p_{\overline{t}}} \overline{F}_{p_{\overline{t}-1}} \cdots \overline{F}_{p_1} = \begin{bmatrix} F_{p_{\overline{t}}} F_{p_{\overline{t}-1}} \cdots F_{p_1} & \displaystyle\sum_{j=1}^{\overline{t}} \Phi_{\overline{t}j} H_{p_j}\mathbf{1} \\ 0 & 1 \end{bmatrix}$$

where $\Phi_{\overline{t}\overline{t}} = I$ and $\Phi_{\overline{t}j} = F_{p_{\overline{t}}} F_{p_{\overline{t}-1}} \cdots F_{p_{j+1}}$ for $j < \overline{t}$. Since the $\overline{F}_p$ are all stochastic, $\overline{F}_{p_{\overline{t}}} \overline{F}_{p_{\overline{t}-1}} \cdots \overline{F}_{p_1}$ must be stochastic as well. Thus, to establish (51) it is sufficient to prove that

$$\sum_{j=1}^{\overline{t}} \Phi_{\overline{t}j} H_{p_j}\mathbf{1} > 0. \tag{52}$$

By assumption, each member of $\{q_1, q_2, \ldots, q_m\}$ occurs in the sequence $p_1, p_2, \ldots p_{\overline{t}}$ at least $n + 1$ times. For $k \in \{1, 2, \ldots m\}$, let i_k be the smallest integer such that $p_{i_k} = q_k$. Since each q_k occurs at least $n + 1$ times, each q_k must occur at least n times in the subsequence $p_{i_k+1}, p_{i_k+2}, \ldots p_{\overline{t}}$. It follows

from Lemma 6 and the definition of $\Phi_{\overline{t}j}$ that $\lceil \Phi_{\overline{t}i_k} \rceil \geq \lceil F^n_{q_k} \rceil$. Thus, $\lceil \Phi_{\overline{t}i_k} H_{q_k} \mathbf{1} \rceil \geq \lceil F^n_{q_k} H_{p_{i_k}} \mathbf{1} \rceil$. Since this hold for all $k \in \{1, 2, \ldots, m\}$,

$$\sum_{k=1}^{m} \lceil \Phi_{\overline{t}i_k} H_{q_k} \mathbf{1} \rceil \geq \sum_{k=1}^{m} \lceil F^n_{q_k} H_{q_k} \mathbf{1} \rceil.$$

From this and (45) it follows that $\sum_{k=0}^{m} \Phi_{\overline{t}i_k} H_{q_k} \mathbf{1} > 0$. But $\sum_{j=1}^{\overline{t}} \Phi_{\overline{t}j} H_{p_j} \mathbf{1} \geq \sum_{k=1}^{m} \Phi_{\overline{t}i_k} H_{q_k} \mathbf{1}$, so (52) is true. ∎

Proposition 2 actually implies that any finite product $F_{p_1} F_{p_2} \cdots F_{p_j}$ will be a discrete-time stability matrix provided there is a set of indices $\{q_1, q_2, \ldots q_m\}$ for which i) each q_k occurs in the set $\{p_1, p_2, \cdots p_j\}$ at least $n+1$ times and ii) $\{\overline{\mathbb{G}}_{q_1}, \overline{\mathbb{G}}_{q_2}, \ldots, \overline{\mathbb{G}}_{q_m}\}$ is a jointly connected collection of graphs. From this it is not difficult to see that any finite product $\overline{F}_{q_1} \overline{F}_{q_2} \cdots \overline{F}_{q_m}$ will be ergodic provided $\{\overline{\mathbb{G}}_{q_1}, \overline{\mathbb{G}}_{q_2}, \ldots, \overline{\mathbb{G}}_{q_m}\}$ is a jointly connected collection of graphs.[6] It is possible to use this fact together with Wolfowitz's theorem (Theorem 3) to devise a proof of Theorem 4, much like the proof of Theorem 2 given earlier. On the other hand, it is also possible to give a simple direct proof of Theorem 4, without using Theorem 3, and this is the approach we take.

Proof of Theorem 4: Let $\mathcal{J}$ denote the set of all subsets $\{p_1, p_2, \ldots, p_m\}$ of $\overline{\mathcal{P}}$ with the property that $\{\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \ldots, \mathbb{G}_{p_m}\}$ is a jointly connected collection. The constraints on σ imply that $\sigma(t)$ takes on every value in one such subset on every interval $[t_i, t_{i+1})$, $i \geq 0$. Let $\overline{n}$ be the number of elements in $\mathcal{J}$. Then for any integer $i > 0$ there must be at least one subset in $\mathcal{J}$ whose elements are each values of σ at least i times on any sequence of $i\overline{n}$ contiguous time-intervals. Set $\overline{t}_i = t_{i(n+1)\overline{n}}$, $i \geq 0$ and let T be the least upper bound on the lengths of the intervals, $[\overline{t}_i, \overline{t}_{i+1})$, $i \geq 0$. By assumption, $T < \infty$. Let $\Phi(t, s)$ denote the state transition matrix defined by $\Phi(t, t) = I$, $t \geq 0$ and $\Phi(t, s) \triangleq F_{\sigma(t-1)} F_{\sigma(t-2)} \cdots F_{\sigma(s)}$, $t > s \geq 0$,. Then $\epsilon(t) = \Phi(t, 0)\epsilon(0)$. To complete the theorem's proof, it is, therefore, enough to show that

$$\lim_{j \to \infty} \Phi(\overline{t}_j, \overline{t}_0) = 0. \tag{53}$$

Clearly $\Phi(\overline{t}_j, 0) = \Phi(\overline{t}_j, \overline{t}_{j-1}) \cdots \Phi(\overline{t}_2, \overline{t}_1)\Phi(\overline{t}_1, \overline{t}_0)$. Moreover, for $i \geq 0$, $[\overline{t}_i, \overline{t}_{i+1})$ is an interval of length at most $(n+1)\overline{n}T$ on which $\sigma(t)$ takes on at least $n+1$ times, every value p_i in some subset $\{p_1, p_2, \ldots, p_m\}$ in $\mathcal{J}$. It follows from Proposition 2 and the definition of Φ that $\|\Phi(\overline{t}_j, \overline{t}_{j-1})\| \leq \lambda$, $j \geq 1$ where λ is a positive number depending only on $(n+1)\overline{n}T$ which satisfies $\lambda < 1$. Hence, $\|\Phi(\overline{t}_j, 0)\| \leq \lambda^j$, $j \geq 1$ from which (53) follows at once. ∎

IV. Leader Following in Continuous Time

Our aim here is to study the convergence properties of the continuous-time version of the leader-follower model discussed in the last section. We begin by noting that the update rule for agent i's heading, defined by (38), is what results when the local feedback law

$$u_i(t) = -\frac{1}{1 + n_i(t) + b_i(t)} \cdot \left((n_i(t) + b_i(t))\theta_i(t) - \sum_{j \in \mathcal{N}_i(t)} \theta_j(t) - b_i(t)\theta_0 \right) \tag{54}$$

is applied to the open-loop discrete-time heading model

$$\theta_i(t+1) = \theta_i(t) + u_i(t). \tag{55}$$

The continuous-time analog of (55) is the integrator equation

$$\dot{\theta}_i = u_i \tag{56}$$

where now t takes values in the real half interval $[0, \infty)$. On the other hand, the continuous time analog of (54) has exactly the same form as (54), except in the continuous time case, $n_i(t)$, $b_i(t)$, and $\theta_i(t)$ are continuous-time variables. Unfortunately, in continuous time control laws of this form can lead to chattering because neighbor relations can change abruptly with changes in agents' positions. One way to avoid this problem is to introduce dwell time, much as was done in [32]. What this means in the present context is that each agent is constrained to change its control law only at discrete times. In particular, instead of using (54), to avoid chatter agent i would use a hybrid control law of the form

$$u_i(t) = -\frac{1}{1 + n_i(t_{ik}) + b_i(t_{ik})} \left((n_i(t_{ik}) + b_i(t_{ik}))\theta_i(t) - \sum_{j \in \mathcal{N}_i(t_{ik})} \theta_j(t) - b_i(t_{ik})\theta_0 \right), \qquad t \in [t_{ik}, t_{ik} + \tau_i) \tag{57}$$

where τ_i is a pre-specified positive number called a *dwell time* and $t_0, t_1, \ldots$ is an infinite time sequence such that $t_{i(k+1)} - t_{ik} = \tau_i$, $k \geq 0$. In the sequel we will analyze controls of this form subject to two simplifying assumptions. First we will assume that all n agents use the same dwell time which we henceforth denote by τ_D. Second we assume the agents are synchronized in the sense that $t_{ik} = t_{jk}$ for all $i, j \in \{1, 2, \ldots, n\}$ and all $k \geq 0$. These assumptions enable us to write u as

$$u = -(I + D_\sigma + B_\sigma)^{-1}((L_\sigma + B_\sigma)\theta - B_\sigma \mathbf{1}\theta_0) \tag{58}$$

where $\overline{\mathcal{P}}$, D_p, B_p and A_p are as before, $L_p = D_p - A_p$ is the Laplacian of $\mathbb{G}_p$, and σ: $[0, \infty) \to \overline{\mathcal{P}}$ is a piecewise constant switching signal with successive switching times separated by τ_D time units. Application of this control to the vector version of (56) results in the closed-loop continuous-time leader-follower model

$$\dot{\theta} = -(I + D_\sigma + B_\sigma)^{-1}((L_\sigma + B_\sigma)\theta - B_\sigma \mathbf{1}\theta_0). \tag{59}$$

In analogy to the discrete-time case, let us agree to say that the n agents are *linked to the leader* across an interval $[t, \tau)$ between switching times t and τ, if the collection of graphs $\{\overline{\mathbb{G}}_{\sigma(t)}, \overline{\mathbb{G}}_{\sigma(t+1)}, \ldots, \overline{\mathbb{G}}_{\sigma(\tau-1)}\}$ encountered along the interval, is jointly connected. Much like before, our goal

[6] Using this fact and the structure of the $\overline{F}_p$ it is also not difficult to show that any finite product $F_{p_1} F_{p_2} \cdots F_{p_j}$ will be a discrete-time stability matrix provided only that $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_j}\}$ is a jointly connected collection of graphs.

here is to show for a large class of switching signals and for any initial set of follower agent headings, that the headings of all n followers converge to the heading of the leader. For convergence, we shall continue to require there to exist infinite sequence of bounded, nonoverlapping time-intervals across which the n-agent group is linked to its leader. However, unlike the discrete-time case we shall not require this sequence of intervals to be contiguous.

Theorem 5: Let $\tau_D > 0$, $\theta(0)$ and θ_0 be fixed and let σ: $[0, \infty) \rightarrow \overline{\mathcal{P}}$ be a piecewise-constant switching signal whose switching times $t_1, t_2, \ldots$ satisfy $t_{i+1} - t_i \geq \tau_D, i \geq 1$. If there is an infinite sequence of bounded, nonoverlapping time-intervals $[t_{i_j}, t_{i_j+k_j}), j \geq 1$, with the property that across each such interval the n-agent group of followers is linked to its leader, then

$$\lim_{t\rightarrow\infty} \theta(t) = \theta_0 \mathbf{1}. \tag{60}$$

Theorem 5 states that θ will converge to $\theta_0\mathbf{1}$, no matter what the value of τ_D, so long as τ_D is greater than zero. This is in sharp contrast to other convergence results involving dwell time switching such as those given in [33], which hold only for sufficiently large values of τ_D. Theorem 5 is a more or less obvious consequence of the following lemma.

Lemma 7: Let

$$M_p \triangleq -(I + D_p + B_p)^{-1}(L_p + B_p), \qquad p \in \overline{\mathcal{P}}. \tag{61}$$

Then

$$\|e^{M_p t}\| \leq 1 \qquad \forall\, t \geq 0, p \in \overline{\mathcal{P}}. \tag{62}$$

Moreover, for each finite set of indices $p_1, p_2, \ldots, p_m$ in $\overline{\mathcal{P}}$ for which $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_m}\}$ is jointly connected, and each set of finite, positive times $t_1, t_2, \ldots, t_m$,

$$\|e^{M_{p_m} t_m} \cdots e^{M_{p_2} t_2} e^{M_{p_1} t_1}\| < 1. \tag{63}$$

Proof of Theorem 5: Let $M_p \triangleq -(I+D_p+B_p)^{-1}(L_p + B_p)$, $p \in \overline{\mathcal{P}}$ and for $i \geq 1$, set

$$N_i = e^{M_{\sigma(t_i)}(t_{i+1} - t_i)}. \tag{64}$$

From inequality (62) in Lemma 7 it follows that

$$\|e^{N_i}\| \leq 1, \qquad i \geq 1. \tag{65}$$

By assumption there is a finite upper bound T on the lengths of the intervals $[t_{i_j}, t_{i_j+k_j})$ across which the n agents are linked to their leader. This and the assumption that $t_{i+1} - t_i \geq \tau_D$, $i \geq 0$, imply that $k_j \leq m, j \geq 1$, where m is the smallest positive integer such that $T \leq m\tau_D$. Let $\mathcal{J}$ be the set of all sequences $p_1, p_2, \ldots p_l \in \overline{\mathcal{P}}$ of at length at most m for which $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_l}\}$ is jointly connected. Define

$$\lambda = \max_{\tau_1 \in [\tau_D, T]} \max_{\tau_2 \in [\tau_D, T]} \cdots \max_{\tau_l \in [\tau_D, T]} \max_{\mathcal{J}} \cdot \|e^{M_{p_l}\tau_l} \cdots e^{M_{p_2}\tau_2} e^{M_{p_1}\tau_1}\|. \tag{66}$$

Note that $\lambda < 1$ because the inequality in (63) is strict, because $\mathcal{J}$ is a finite set, because $[\tau_D, T]$ is compact and because the matrix exponentials in (66) depend continuously on the τ_i. In view of the definition of λ and the definitions of the N_i in (64),

$$\|e^{N_{i_j+k_j-1}} \cdots e^{N_{i_j+1}} e^{N_{i_j}}\| \leq \lambda, \qquad j \geq 1. \tag{67}$$

However

$$e^{N_{i_{j+1}-1}} \cdots e^{N_{i_j+1}} e^{N_{i_j}} = \left(e^{N_{i_{j+1}-1}} \cdots e^{N_{i_j+k_j}}\right) \cdot \left(e^{N_{i_j+k_j-1}} \cdots e^{N_{i_j+1}} e^{N_{i_j}}\right), \qquad j \geq 1.$$

This, (65), (67) and the sub-multiplicative property of the induced infinity norm imply that

$$\left\|e^{N_{i_{j+1}-1}} \cdots e^{N_{i_j+1}} e^{N_{i_j}}\right\| \leq \lambda, \qquad j \geq 1. \tag{68}$$

Set $\overline{\theta}(t) = \theta(t) - \mathbf{1}\theta_0$ and note that

$$\dot{\overline{\theta}} = -(I + D_\sigma + B_\sigma)^{-1}(L_\sigma + B_\sigma)\overline{\theta}$$

because of (59). Let $\Phi(t, \mu)$ be the state transition matrix of $-(I + D_{\sigma(t)} + B_{\sigma(t)})^{-1}(L_{\sigma(t)} + B_{\sigma(t)})$. Then $\overline{\theta}(t) = \Phi(t, 0)\overline{\theta}(0)$. To complete the proof it is, therefore, enough to show that

$$\|\Phi(t_{i_j}, t_{i_1})\| \leq \lambda^{j-1}, \qquad j \geq 1. \tag{69}$$

In view of the definitions of the N_i,

$$\Phi(t_{i_{j+1}}, t_{i_j}) = e^{N_{i_{j+1}-1}} \cdots e^{N_{i_j+1}} e^{N_{i_j}}, \qquad j \geq 1.$$

From this and (68) it follows that

$$\|\Phi(t_{i_{j+1}}, t_{i_j})\| \leq \lambda, \qquad j \geq 1. \tag{70}$$

However

$$\Phi(t_{i_j}, t_1) = \Phi(t_{i_j}, t_{i_{j-1}}) \cdots \Phi(t_{i_2}, t_{i_1})$$

so

$$\|\Phi(t_{i_j}, t_1)\| \leq \|\Phi(t_{i_j}, t_{i_{j-1}})\| \cdots \|\Phi(t_{i_2}, t_{i_1})\|.$$

From this and (70), it now follows that (69) is true. ■

Proof of Lemma 7: Fix $t > 0$ and $p \in \overline{\mathcal{P}}$. Observe first that

$$M_p = F_p - I \tag{71}$$

where F_p is the matrix $F_p = (I + D_p + B_p)^{-1}(I + A_p)$. As noted previously, the partitioned matrix

$$\overline{F}_p \triangleq \begin{bmatrix} F_p & H_p\mathbf{1} \\ 0 & 1 \end{bmatrix} \tag{72}$$

originally defined in (43), is stochastic with positive diagonal elements as are the matrices

$$\overline{F}_p^{\,i} = \begin{bmatrix} F_p^i & \displaystyle\sum_{j=1}^{i} F_p^{(j-1)} H_p \mathbf{1} \\ 0 & 1 \end{bmatrix}, \qquad i \geq 1. \tag{73}$$

Since

$$e^{\overline{F}_p t} = \sum_{i=0}^{\infty} \frac{(t\overline{F}_p)^i}{i!} \tag{74}$$

$e^{\overline{F}_p t}$ must also be nonnegative with positive diagonal elements. But $e^{(\overline{F}_p-\overline{I})t} = e^{-t}e^{\overline{F}_p t}$, where $\overline{I}$ is the $(n+1)\times(n+1)$ identity, so the same must be true of $e^{(\overline{F}_p-\tilde{I})t}$. Moreover $(\overline{F}_p - \overline{I})\mathbf{1} = 0$ which means that $e^{(\overline{F}_p-\overline{I})t}\mathbf{1} = e^0\mathbf{1} = \mathbf{1}$ and, thus, that $e^{(\overline{F}_p-\overline{I})t}$ is stochastic. In summary, $e^{(\overline{F}_p-\overline{I})t}$ is a stochastic matrix with positive diagonal entries.

Equations (72)–(74) imply that

$$e^{\overline{F}_p t} = \begin{bmatrix} e^{F_p t} & k_p \\ 0 & e^t \end{bmatrix}$$

where

$$k_p = \sum_{i=0}^{\infty} \frac{t^i}{i!} \sum_{j=1}^{i} F_p^{(j-1)} H_p \mathbf{1}. \tag{75}$$

Therefore

$$e^{(\overline{F}_p-\overline{I})t} = \begin{bmatrix} e^{(F_p-I)t} & k_p \\ 0 & 1 \end{bmatrix}. \tag{76}$$

However, $e^{(\overline{F}_p-\overline{I})t}$ is row-stochastic, so $e^{(F_p-I)t}$ must have its row sums all bounded above by 1. From this and (71) it follows that (62) is true.

Now suppose that $\{\overline{\mathbb{G}}_{p_1}, \overline{\mathbb{G}}_{p_2}, \ldots, \overline{\mathbb{G}}_{p_m}\}$ is a jointly connected collection of graphs. Then by Lemma 5

$$\sum_{j=1}^{m} F_{p_j}^{(i-1)} H_{p_j} \mathbf{1} > 0, \qquad i > n. \tag{77}$$

This and (75) imply that

$$\sum_{j=1}^{m} k_{p_j} > 0 \tag{78}$$

because all of the matrices in the sums defining the k_p in (75) are nonnegative.

Using (76) and the definition of M_p in (71) we can write

$$e^{(\overline{F}_{p_m}-\overline{I})t_m} e^{(\overline{F}_{p_{m-1}}-\overline{I})t_{m-1}} \cdots e^{(\overline{F}_{p_1}-\overline{I})t_1} = \begin{bmatrix} e^{M_{p_m}t_m} e^{M_{p_{m-1}}t_{m-1}} \cdots e^{M_{p_1}t_1} & \displaystyle\sum_{j=1}^{m} \Phi_{mj} k_{p_j} \\ 0 & 1 \end{bmatrix} \tag{79}$$

where $\Phi_{mm} = I$ and $\Phi_{mj} = e^{M_{p_m}t_m} \cdots e^{M_{p_{j+1}}t_{j+1}}$, $j < m$. Since the matrix on the right in (79) is stochastic, its row sums all equal one. To complete the proof it is, therefore, enough to show that

$$\sum_{j=1}^{m} \Phi_{mj} k_{p_j} > 0. \tag{80}$$

Note that for any nonnegative $n \times n$ matrix N, $e^N \geq I$ because the matrix $\sum_{i=1}^{\infty}(1/i!)N^i$ in the definition $e^N = \sum_{i=0}^{\infty}(1/i!)N^i$ is nonnegative. Thus, for $j \in \{1, 2, \ldots, m\}$, $e^{M_{p_j}t_j} \geq I$ and consequently $\Phi_{mj} \geq I$. Therefore, $\Phi_{mj}k_j \geq k_j$, $j \in \{1, 2, \ldots, m\}$. From this and (78) it follows that (80) is true and, thus, that the inequality in (63) is correct. ■

V. Concluding Remarks

As stated in the abstract, the main objective of this paper has been to provide a theoretical explanation for behavior observed in the simulation studies reported in [1]. We refer the interested reader to Vicsek's paper and references cited therein, for a more thorough description of the model considered and for data documenting the simulation studies performed.

The theorems in this paper all provide convergence results for rich classes of switching signals and arbitrary initial heading vectors. Of course as soon as one elects to interpret these results in the context of heading models for mobile autonomous agents, one needs to add qualifications, because the actual switching signal σ generated along a particular model's trajectory would have to depend on the model's initial heading vector. To make such a dependence explicit (and to run meaningful simulations) more complete models would have to be defined. In carrying out this step, one can expect to obtain a variety of results. For example, with very large agent sensing regions (i.e., r very large) and agents initially near each other, one would expect enough connectivity along resultant trajectories for convergence to a common heading to occur. On the other hand, with widely distributed initial agent positions and r very small, one would expect to see a bifurcation of the group into distinct subgroups with different steady state headings. In other words, a complete deterministic understanding of the flocking problems we've considered would require both more complete agent motion models as well as an understanding of the nonlinear feedback process upon which σ actually would depend. An alternative probabilistic approach might enable one to circumvent or at least simplify the analysis of the feedback process.

Some versions of the Vicsek model and others considered in this paper may ultimately find application in coordination of groups of mobile autonomous agents. Of course before this can happen many compelling issues such as collision avoidance, and the effects of disturbances, noise, sensing errors, vehicle modeling errors, etc. would have to be satisfactorily addressed. For example, the collision avoidance question might also be approached by replacing the point models implicitly used in this paper, with the model of a bumper-like "virtual shell" within which each agent vehicle is constrained to remain [34].

While the analysis in this paper is deterministic and does not address the noise issue, the results obtained suggest that to understand the effect of additive noise, one should focus on how noise inputs effect connectivity of the associated neighbor graphs. Simulation results presented in [1] indicate that when noise intensity in the system is fixed, there is a *phase transition* as the density of the agents is increased, i.e., there is a critical density after which all agents eventually become aligned. It is possible that this phenomenon can be adequately explained using percolation theory of random graphs [35].

The results of this paper have been extended to the case where there are inter-agent forces due to attraction, repulsion and alignment [36]. The new results indicate that the convergence arguments used in this paper also apply to the more general problem considered in [36] under similar assumptions regarding the connectivity of the graph representing the nearest neighbor relationships.

The convergence proof for Vicsek's model presented in Section II relies heavily on Wolfowitz's theorem. By generalizing some of the constructions Wolfowitz used in his proof, it is possible to develop a convergence result for a continuous-time analog of the Vicsek model which is quite similar to Theorem 5.

In studying continuous-time leader-following, we imposed the requirement that all followers use the same dwell time. This is not really necessary. In particular, without much additional effort it can be shown that Theorem 5 remains true under the relatively mild assumption that all agents use dwell times which are rationally related. In contrast, the synchronization assumption may be more difficult to relax. Although convergence is still likely without synchronization, the aperiodic nature of σ's switching times which could result, make the analysis problem more challenging.

The use of simply averaging rules such as those discussed in this paper can sometimes have counter-intuitive consequences which may be undesirable in some applications. For example the average of headings 0.01 and $2\pi - 0.01$ is π so this might cause two agents with headings both close to 0, to both approximately reverse direction to a heading of π on the next step. It would be of interest to determine how update rules might be modified to avoid this type of behavior. Of course issues along these lines would not arise at all if the systems we've considered were modeling other physically significant variables such as agent speed or temperature where one could take all of $\mathbb{R}$ rather than just $[0,\ 2\pi)$ as the set in which the θ_i take values.

The models we have analyzed are of course very simple and as a consequence, they are probably not really descriptive of actual bird-flocking, fish schooling, or even the coordinated movements of envisioned groups of mobile robots. Nonetheless, these models do seem to exhibit some of the rudimentary behaviors of large groups of mobile autonomous agents and for this reason they serve as a natural starting point for the analytical study of more realistic models. It is clear from the developments in this paper, that ideas from graph theory and dynamical system theory will play a central role in both the analysis of such biologically inspired models and in the synthesis of provably correct distributed control laws which produce such emergent behaviors in man-made systems.

REFERENCES

[1] T. Vicsek, A. Czirok, E. Ben Jacob, I. Cohen, and O. Schochet, "Novel type of phase transitions in a system of self-driven particles," *Phys. Rev. Lett.*, vol. 75, pp. 1226–1229, 1995.

[2] C. Reynolds, "Flocks, birds, and schools: A distributed behavioral model," *Comput. Graph.*, vol. 21, pp. 25–34, 1987.

[3] A. Czirok, A. L. Barabasi, and T. Vicsek, "Collective motion of self-propelled particles: Kinetic phase transition in one dimension," *Phys. Rev. Lett.*, vol. 82, pp. 209–212, 1999.

[4] J. Toner and Y. Tu, "Flocks, herds, and schools: A quantitative theory of flocking," *Phys. Rev. E*, vol. 58, pp. 4828–4858, 1998.

[5] ——, "Long range order in a two dimensional xy model: How birds fly together," *Phys. Rev. Lett.*, vol. 75, pp. 4326–4329, 1995.

[6] A. S. Mikhailov and D. Zannette, "Noise induced breakdown of collective coherent motion in swarms," *Phys. Rev. E*, vol. 60, pp. 4571–4575, 1999.

[7] C. P. Schenk, P. Schutz, M. Bode, and H. G. Purwins, "Interaction of self organized quasi particles in a two dimensional reaction-diffusion system: The formation of molecules," *Phys. Rev. E*, vol. 58, pp. 6480–6486, 1998.

[8] D. Grunbaum and A. Okubo, "Modeling social animal aggregations," in *Frontiers in Theoretical Biology*. New York: Springer-Verlag, 1994, 100 of Lecture Notes in Biomathematics, pp. 296–325.

[9] C. M. Breder, "Equations descriptive of fish schools and other animal aggregations," *Ecology*, vol. 35, pp. 361–370, 1954.

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

[11] K. Warburton and J. Lazarus, "Tendency-distance models of social cohesion in animal groups," *J. Theoret. Biol.*, vol. 150, pp. 473–488, 1991.

[12] G. Flierl, D. Grunbaum, S. Levin, and D. Olson, "From individuals to aggregations: The interplay between behavior and physics," *J. Theoret. Biol.*, vol. 196, pp. 397–454, 1999.

[13] Y. Liu, K. M. Passino, and M. Polycarpou, "Stability analysis of one-dimensional asynchronous swarms," in *Proc. Amer. Control Conf.*, Arlington, VA, June 2001, pp. 716–721.

[14] N. Leonard and E. Friorelli, "Virtual leaders, artificial potentials and coordinated control of groups," presented at the IEEE Conf. Decision Control, Orlando, FL, 2001.

[15] R. Olfati and R. M. Murray, "Distributed structural stabilization and tracking for formations of dynamic multi-agents," presented at the IEEE Conf. Decision Control, Las Vegas, NV, 2002.

[16] P. Ogren, M. Egerstedt, and X. Hu, "A control Lyapunov function approach to multi-agent coordination," *IEEE Trans. Robot. Automat.*, p. 18, Oct. 2002.

[17] J. A. Fax and R. M. Murray, "Graph Laplacians and stabilization of vehicle formations," presented at the 15th IFAC Congr., Barcelona, Spain, 2002.

[18] J. P. Desai, J. P. Ostrowski, and V. Kumar, "Modeling and control of formations of nonholonomic mobile robots," *IEEE Trans. Robot. Automat.*, vol. 17, pp. 905–908, June 2001.

[19] J. Wolfowitz, "Products of indecomposable, aperiodic, stochastic matrices," in *Proc. Amer. Mathematical Soc.*, vol. 15, 1963, pp. 733–736.

[20] E. Seneta, *Non-Negative Matrices and Markov Chains*. New York: Springer-Verlag, 1981.

[21] I. Daubechies and J. C. Lagarias, "Sets of matrices all infinite products of which converge," *Linear Alg. Applicat.*, vol. 161, pp. 227–263, 1992.

[22] ——, "Corrigendum/addendum to 'Sets of matrices all infinite products of which converge'," *Linear Alg. Applicat.*, vol. 327, pp. 69–83, 2001.

[23] R. Bru, L. Elsner, and M. Neumann, "Convergence of infinite products of matrices and inner-outer iteration schemes," *Electron. Trans. Numer. Anal.*, vol. 2, pp. 183–193, 1994.

[24] W. J. Beyn and L. Elsner, "Infinite products and paracontracting matrices," *Electron. J. Linear Alg.*, vol. 2, pp. 1–8, 1997.

[25] A. Rhodius, "On the maximum of ergodicity coefficients, the Dobrushin ergodicity coefficient, and products of stochastic matrices," *Linear Alg. Applicat.*, vol. 253, pp. 141–154, 1997.

[26] A. Vladimirov, E. Elsner, and W. J. Beyn, "Stability and paracontractivity of discrete linear inclusions," *Linear Alg. Applicat.*, vol. 312, pp. 125–134, 2000.

[27] B. Mohar, "The Laplacian spectrum of graphs," in *Graph Theory, Combin. Applicat.*, Y. A. G. Chartrand, O. R. Ollerman, and A. J. Schwenk, Eds., 1991, vol. 2, pp. 871–898.

[28] R. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

[29] M. Berger and Y. Wang, "Bounded semigroups of matrices," *Linear Alg. Applicat.*, vol. 166, pp. 21–27, 1992.

[30] M. Ando and M. H. Shih, "Simultaneous contractibility. Siam journal on matrix analysis and applications," *Linear Alg. Applicat.*, vol. 19, pp. 487–498, 1998.

[31] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York: Springer-Verlag, 2001.

[32] A. S. Morse, "Dwell-time switching," in *Proc. 2nd Eur. Control Conf.*, 1993, pp. 176–181.

[33] ——, "A bound for the disturbance-to-tracking error gain of a supervised set-point control system," in *Perspectives in Control—Theory and Applications*, D. N. Cyrot, Ed. New York: Springer-Verlag, 1998, pp. 23–41.

[34] ——, (2001, July) Virtual shells for avoiding collisions. [Online]. Available: http://entity.eng.yale.edu/controls/publications.html.

[35] B. Bollabas, *Random Graphs*, 2nd ed. New York: Cambridge Univ. Press, 2001.

[36] H. Tanner, A. Jadbabaie, and G. Pappas, "Distributed coordination strategies for groups of mobile autonomous agents," ESE Dept., Univ. Pennsylvania, Tech. Rep., Dec. 2002.

Ali Jadbabaie received the B.S. degree (with high honors) from Sharif University of Technology, Tehran, Iran, the M.S. degree in electrical engineering from the University of New Mexico, Albuquerque, and the Ph.D. degree in control and dynamical systems (CDS) from the California Institute of Technology (Caltech), Pasadena, in 1995, 1997, and 2000, respectively.

From October 2000 to July 2001, he was a Postdoctoral Scholar at the CDS Department at Caltech. From July 2001 to August 2002, he was a Postdoctoral Scholar at the Electrical Engineering Department, Yale University, New haven, CT. Since August 2002, he has been an Assistant Professor in the department of Electrical and Systems Engineering and the General Robotics Automation, Sensing, and Perception (GRASP) laboratory at the University of Pennsylvania, Philadelphia. His research interests are in real-time optimization-based control with applications to Unmanned Aerial Vehicles, distributed coordination of mobile autonomous vehicles, and robust and optimal control.

Jie Lin was born in Shangrao, P.R. China, in 1976. He received the undergraduate degree in physics from Fudan University, Shanghai, P.R. China, the M.S. degree in physics from Creighton University, Omaha, NE, and the M.S. degree in electrical engineering from Yale University, New Haven, CT, in 1997, 1999, and 2000, respectively. He is currently working toward the Ph.D. degree at Yale University.

His research interests are in the areas of analysis and synthesis of multiagent coordination control, as well as the applications to mobile sensor networks and robotics.

A. Stephen Morse (S'62–M'67–SM'78–F'84) was born in Mt. Vernon, NY. He received the B.S.E.E. degree from Cornell University, Ithaca, NY, the M.S. degree from the University of Arizona, Tucson, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 1962, 1964, and 1967, respectively.

From 1967 to 1970, he was associated with the Office of Control Theory and Application (OCTA) at the NASA Electronics Research Center, Cambridge, MA. Since 1970, he has been with Yale University, New Haven, CT, where he is currently a Professor of Electrical Engineering. His main interest is in system theory, and he has done research in network synthesis, optimal control, multivariable control, adaptive control, urban transportation, vision-based control, hybrid and nonlinear systems, and, most recently, coordination and control of large grouping of mobile autonomous agents.

Dr. Morse is a Distinguished Lecturer of the IEEE Control Systems Society, and a corecipient of the Society's George S. Axelby Outstanding Paper Award. He has twice received the American Automatic Control Council's Best Paper Award, and is the 1999 recipient of the IEEE Technical Field Award for Control Systems. He is a member of the National Academy of Engineering.