

REACHING A CONSENSUS IN A DYNAMICALLY CHANGING ENVIRONMENT: CONVERGENCE RATES, MEASUREMENT DELAYS, AND ASYNCHRONOUS EVENTS*

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

Abstract. This paper uses recently established properties of compositions of directed graphs together with results from the theory of nonhomogeneous Markov chains to derive worst case convergence rates for the headings of a group of mobile autonomous agents which arise in connection with the widely studied Vicsek consensus problem. The paper also uses graph-theoretic constructions to solve modified versions of the Vicsek problem in which there are measurement delays, asynchronous events, or a group leader. In all three cases the conditions under which consensus is achieved prove to be almost the same as the conditions under which consensus is achieved in the synchronous, delay-free, leaderless case.

Key words. cooperative control, graph theory, switched systems, convergence rates, delays, asynchronism

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

DOI. 10.1137/060657029

1. Introduction. In a recent paper [6] the present authors defined the notion of “graph composition” and established a number of basic properties of compositions of directed graphs which are useful in explaining how a consensus might be reached by a group of mobile autonomous agents in a dynamically changing environment. The aim of this paper is to use the graph-theoretic findings of [6] to address several issues related to the well-known Vicsek consensus problem [20] which have either not been considered before or have been only partially resolved.

The paper begins with a brief review in section 2 of the basic leaderless consensus problem treated in [6, 14, 16]. Section 3 exploits the connection between “neighbor-shared” graphs and the elegant theory of “scrambling matrices” found in the literature on nonhomogeneous Markov chains [17, 9] to help in the derivation of worst case agent heading convergence rates for the leaderless version of the Vicsek problem. Section 4 addresses a modified version of the consensus problem in which integer-valued delays occur in the values of the headings which agents measure. In keeping with the overall theme of this paper, the effect of measurement delays is analyzed from a mainly graph-theoretic point of view. This enables us to significantly relax previously derived conditions [18, 19, 3] under which consensus can be achieved in the face of measurement delays. A comparison is made between the results of [3] and

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

<http://www.siam.org/journals/sicon/47-2/65702.html>

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

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

the main result of this paper on measurement delays, namely Theorem 2. To model dynamics when delays are present requires a somewhat different type of stochastic “flocking matrix” than the one which is appropriate in the delay-free case. The graphs of the type of matrices to which we are referring are directed, just as in the delay-free case, but do not have self-arcs at every vertex. As a result, the set of such graphs, denoted by \mathcal{D} , is not closed under composition. The smallest set of directed graphs which contains \mathcal{D} and which is closed under composition is called the set of “extended delay graphs.” This class is explicitly characterized. Section 4 then develops the requisite properties of extended delay graphs needed to prove Theorem 2.

Section 5 considers a modified version of the flocking problem in which each agent independently updates its heading at times determined by its own clock. It is not assumed that the groups’ clocks are synchronized together or that the times any one agent updates its heading are evenly spaced. In this case, the deriving of conditions under which all agents eventually move with the same heading requires the analysis of the asymptotic behavior of an overall *asynchronous* process which models the n -agent system. The analysis is carried out by first embedding this process in a suitably defined *synchronous* discrete-time, hybrid dynamical system \mathbb{S} . This is accomplished using the concept of *analytic synchronization* outlined previously in [12, 13]. This enables us to bring to bear results derived earlier in [6] to characterize a rich class of system trajectories under which consensus is achieved.

In section 6 we briefly consider a modified version of the consensus problem for the same group of n agents as before but now with one of the group’s members (say agent 1) acting as the group’s *leader*. The remaining agents, called *followers* and labelled 2 through n , do not know who the leader is or even if there is a leader. Accordingly they continue to function as if there was no leader using the same update rules as are used in the leaderless case. The leader, on the other hand, acting on its own, ignores these update rules and moves with a constant heading. Using the main result on leaderless consensus summarized in section 2, we then develop conditions under which all follower agents eventually move in the same direction as the leader. These conditions correct prior findings on leader following in [11] which are in error.

2. Background. As in [6], the system of interest consists of n autonomous agents, labelled 1 through n , all moving in the plane with the same speed but with different headings. Each agent’s heading is updated using a simple local rule based on the average of its own heading plus the headings of its “neighbors.” Agent i ’s *neighbors* at time t are those agents, including itself, which are in a closed disk of prespecified radius r_i centered at agent i ’s current position. In what follows $\mathcal{N}_i(t)$ denotes the set of labels of those agents which are neighbors of agent i at time t . Agent i ’s heading, written θ_i , evolves in discrete time in accordance with a model of the form

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

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

2.1. Neighbor graph. The explicit form of the update equations determined by (1) depends on the relationships between neighbors which exist at time t . These relationships can be conveniently described by a directed graph $\mathbb{N}(t)$ with vertex set

$\mathcal{V} = \{1, 2, \dots, n\}$ and arc set $\mathcal{A}(\mathbb{N}(t)) \subset \mathcal{V} \times \mathcal{V}$ which is defined so that (i, j) is an arc or directed edge from i to j just in case agent i is a neighbor of agent j at time t . Thus $\mathbb{N}(t)$ is a directed graph on n vertices with at most one arc connecting each ordered pair of distinct vertices and with exactly one self-arc at each vertex. We write \mathcal{G}_{sa} for the set of all such graphs and \mathcal{G} for the set of all directed graphs with vertex set \mathcal{V} . It is natural to call a vertex i a *neighbor* of vertex j in a graph $\mathbb{G} \in \mathcal{G}$ if (i, j) is an arc in \mathbb{G} .

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

$$(2) \quad F = D^{-1}A',$$

where A' is the transpose of the adjacency matrix of \mathbb{N} and D the diagonal matrix whose j th diagonal element is the in-degree of vertex j within \mathbb{N} . Then

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

where θ is the heading vector $\theta = [\theta_1 \ \theta_2 \ \dots \ \theta_n]'$ and $F(t)$ is the flocking matrix of the *neighbor graph* $\mathbb{N}(t)$.

2.3. Leaderless consensus. To proceed, we need to recall a few definitions from [6]. We call a vertex i of a directed graph \mathbb{G} a *root* of \mathbb{G} if for each other vertex j of \mathbb{G} , there is a path from i to j . Thus i is a root of \mathbb{G} if it is the root of a directed spanning tree of \mathbb{G} . We say that \mathbb{G} is *rooted at i* if i is in fact a root. Thus \mathbb{G} is rooted at i just in case each other vertex of \mathbb{G} is *reachable* from vertex i along a path within the graph. \mathbb{G} is *strongly rooted at i* if each other vertex of \mathbb{G} is reachable from vertex i along a path of length 1. Thus \mathbb{G} is strongly rooted at i if i is a neighbor of every other vertex in the graph. A *rooted graph* \mathbb{G} is a graph which possesses at least one root. Finally, a *strongly rooted graph* is a graph which has at least one vertex at which it is strongly rooted.

By the *composition* of two directed graphs $\mathbb{G}_p, \mathbb{G}_q$ with the same vertex set \mathcal{V} we mean the graph $\mathbb{G}_q \circ \mathbb{G}_p$ with the same vertex set \mathcal{V} and arc set defined such that (i, j) is an arc of $\mathbb{G}_q \circ \mathbb{G}_p$ if for some vertex k , (i, k) is an arc of \mathbb{G}_p and (k, j) is an arc of \mathbb{G}_q . A finite sequence of directed graphs $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_q$ with the same vertex set is *jointly rooted* if the composition $\mathbb{G}_q \circ \mathbb{G}_{q-1} \circ \dots \circ \mathbb{G}_1$ is rooted. An infinite sequence of graphs $\mathbb{G}_1, \mathbb{G}_2, \dots$ with the same vertex set is *repeatedly jointly rooted by subsequences of length q* if there is a positive integer q for which each finite sequence $\mathbb{G}_{qk+1}, \dots, \mathbb{G}_{q(k+1)}$, $k \geq 0$, is jointly rooted. The main result on leaderless consensus [14, 16] is equivalent to the following result from [6].

THEOREM 1. *Let $\theta(0)$ be fixed. For any trajectory of the system determined by (1) along which the sequence of neighbor graphs $\mathbb{N}(0), \mathbb{N}(1), \dots$ is repeatedly jointly rooted by sequences of length q , there is a constant θ_{ss} , depending only on $\theta(0)$, for which*

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

where the limit is approached exponentially fast.

3. Convergence rates. The aim of this section is to derive a bound on the rate at which θ converges.¹ There are two distinct ways to go about this, and below we

¹This section summarizes and extends some of the key findings of [7].

describe both. To do this we will make use of certain structural properties of the F . As defined, each F is square and nonnegative, where by a *nonnegative* matrix we mean a matrix whose entries are all nonnegative. Each F also has the property that its row sums all equal 1 (i.e., $F\mathbf{1} = \mathbf{1}$). Matrices with these two properties are called (row) *stochastic* [10]. It is easy to verify that the class of all $n \times n$ stochastic matrices is closed under multiplication. It is worth noting that because the vertices of the graphs in \mathcal{G}_{sa} all have self-arcs, the F also have the property that their diagonal elements are positive.

In what follows we write $M \geq N$ whenever $M - N$ is a nonnegative matrix. We also write $M > N$ whenever $M - N$ is a positive matrix, where by a *positive matrix* we mean a matrix with all positive entries. For any nonnegative matrix R of any size, we write $\|R\|$ for the largest of the row sums of R . Note that $\|R\|$ is the induced infinity norm of R and consequently is submultiplicative. Moreover, $\|M_1\| \leq \|M_2\|$ if $M_1 \leq M_2$. Observe that for any $n \times n$ stochastic matrix S , $\|S\| = 1$ because the row sums of a stochastic matrix all equal 1. As in [6] we write $\lfloor M \rfloor$ and $\lceil M \rceil$ for the $1 \times m$ row vectors whose j th entries are the smallest and largest elements, respectively, of the j th column of M . Note that $\lfloor M \rfloor$ is the largest $1 \times m$ nonnegative row vector c for which $M - \mathbf{1}c$ is nonnegative and that $\lceil M \rceil$ is the smallest nonnegative row vector c for which $\mathbf{1}c - M$ is nonnegative. Note in addition that for any $n \times n$ stochastic matrix S , one can write

$$(5) \quad S = \mathbf{1}\lfloor S \rfloor + \lceil S \rceil \quad \text{and} \quad S = \mathbf{1}\lceil S \rceil - \lfloor S \rfloor,$$

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

$$(6) \quad \lfloor S \rfloor = S - \mathbf{1}\lceil S \rceil \quad \text{and} \quad \lceil S \rceil = \mathbf{1}\lceil S \rceil - S,$$

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

$$(7) \quad \|\lfloor S \rfloor\| = 1 - \lceil S \rceil \mathbf{1} \quad \text{and} \quad \|\lceil S \rceil\| = \lceil S \rceil \mathbf{1} - 1.$$

In what follows we will also be interested in the matrix

$$(8) \quad \llbracket S \rrbracket = \lfloor S \rfloor + \lceil S \rceil.$$

This matrix satisfies

$$(9) \quad \llbracket S \rrbracket = \mathbf{1}(\lceil S \rceil - \lfloor S \rfloor)$$

because of (5).

To prove that all θ_i converge to a common heading, it is necessary to prove that θ converges to a vector of the form $\theta_{ss}\mathbf{1}$, where $\mathbf{1}$ is the $n \times 1$ vector of 1's. It is clear from (3) that θ will converge to such a vector just in case, as $t \rightarrow \infty$, the matrix product $F(t) \cdots F(0)$ converges to a rank one matrix of the form $\mathbf{1}c$ for some $n \times 1$ row vector c . Thus to study how such matrix products converge it is sufficient to study how products of stochastic matrices of the form $S_j \cdots S_1$ converge as $j \rightarrow \infty$. As in [6], we say that a matrix product $S_j S_{j-1} \cdots S_1$ *converges exponentially fast at a rate no slower than* λ to a matrix of the form $\mathbf{1}c$ if there are nonnegative constants b and λ with $\lambda < 1$, such that

$$(10) \quad \|(S_j \cdots S_1) - \mathbf{1}c\| \leq b\lambda^j, \quad j \geq 1.$$

The following fact is proved in [6].

PROPOSITION 1. *If an infinite sequence of stochastic matrices S_1, S_2, \dots satisfies*

$$(11) \quad \|\|S_j \cdots S_1\|\| \leq \bar{b}\lambda^j, \quad j \geq 0,$$

for some nonnegative constants \bar{b} and $\lambda < 1$, then the product $S_j S_{j-1} \cdots S_1$ converges exponentially fast at a rate no slower than λ to a matrix of the form $\mathbf{1}c$.

We will exploit this inequality in deriving specific convergence rates.

Any $n \times n$ stochastic matrix S determines a directed graph $\gamma(S)$ with the vertex set $\{1, 2, \dots, n\}$ and arc set defined in such a way so that (i, j) is an arc of $\gamma(S)$ from i to j just in case the j th entry of S is nonzero. Note that the graph of any stochastic matrix with positive diagonal elements must be in \mathcal{S}_{sa} . Since flocking matrices have this property, their graphs must be in \mathcal{G}_{sa} . It is known [6] that for the set of $n \times n$ stochastic matrices S_1, S_2, \dots, S_p

$$(12) \quad \gamma(S_p \cdots S_2 S_1) = \gamma(S_p) \circ \cdots \circ \gamma(S_2) \circ \gamma(S_1).$$

We will make use of the fact that for any two $n \times n$ stochastic matrices S_1 and S_2 ,

$$(13) \quad \phi(S_2 S_1) \geq \phi(S_2)\phi(S_1),$$

where for any nonnegative matrix M , $\phi(M)$ denotes the smallest nonzero element of M . To prove that this is so, note first that any stochastic matrix S can be written as $S = \phi(S)\bar{S}$, where \bar{S} is a nonzero matrix whose nonzero entries are all bounded below by 1; moreover, if $S = \hat{\phi}(S)\hat{S}$, where $\hat{\phi}(S)$ is a number and \hat{S} is also a nonzero matrix whose nonzero entries are all bounded below by 1, then $\phi(S) \geq \hat{\phi}(S)$. Accordingly, write $S_i = \phi(S_i)\bar{S}_i$, $i \in \{1, 2\}$, where each \bar{S}_i is a nonzero matrix whose nonzero entries are all bounded below by 1. Since $S_2 S_1 = \phi(S_2)\phi(S_1)\bar{S}_2\bar{S}_1$ and $S_2 S_1$ is nonzero, $\bar{S}_2\bar{S}_1$ must be nonzero as well. Moreover, the nonzero entries of $\bar{S}_2\bar{S}_1$ must be bounded below by 1 because the product of any two $n \times n$ matrices with all nonzero entries bounded below by 1 must be a matrix with the same property. Therefore $\phi(S_2 S_1) \geq \phi(S_2)\phi(S_1)$ as claimed. An important consequence of (13) is that for any set of stochastic matrices S_1, S_2, \dots, S_m for which each $\phi(S_i)$ is bounded below by a positive number b ,

$$(14) \quad \phi(S_m \cdots S_1) \geq b^m.$$

Our goal is now to use these facts to derive an explicit convergence rate for the situation considered by Theorem 1. We will do this in two different ways. The first way is based on properties of stochastic matrices with strongly rooted graphs.

3.1. Strongly rooted graphs. Let $\mathcal{F}(q)$ denote the set of all products of q flocking matrices whose corresponding sequences of q graphs are each jointly rooted. In view of (12), each matrix in $\mathcal{F}(q)$ must have a rooted graph in \mathcal{G}_{sa} . In other words, each matrix in $\mathcal{F}(q)$ has a rooted graph and is a product of q flocking matrices. Since the set of all flocking matrices is finite, so is $\mathcal{F}(q)$. It is shown in [6] that the composition of any set of at least $(n - 1)^2$ rooted graphs in \mathcal{G}_{sa} is strongly rooted. This and (12) imply that the product of any $(n - 1)^2$ matrices in $\mathcal{F}(q)$ must have a strongly rooted graph in \mathcal{G}_{sa} . Thus if we set $m = (n - 1)^2$ and write $(\mathcal{F}(q))^m$ for the set of all products of m matrices from $\mathcal{F}(q)$, then each matrix in $(\mathcal{F}(q))^m$ must have a

strongly rooted graph. Moreover, $(\mathcal{F}(q))^m$ must be a finite set because $\mathcal{F}(q)$ is. It is shown in [6] that convergence of the θ_i in Theorem 1 occurs at a rate no slower than

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

Our goal is to derive an explicit bound for λ .

As a first step towards this end, let S be any stochastic matrix with a strongly rooted graph. We claim that

$$(15) \quad \|\|S\|\| \leq 1 - \phi(S).$$

To understand why this is so, note first that because $\gamma(S)$ is strongly rooted, at least one vertex—say the k th—must be a root with arcs to every other vertex. This means that the k th column of S must be positive. Since $\phi(S)$ is a lower bound on all nonzero elements in S , the smallest element in the k th column of S is bounded below by $\phi(S)$. Therefore $\lfloor S \rfloor \mathbf{1} \geq \phi(S)$. But from (7) $\|\|S\|\| = 1 - \lfloor S \rfloor \mathbf{1}$. This implies that (15) is true.

As a second step, let us note that the definition of a flocking matrix implies that all nonzero entries are bounded below by $\frac{1}{n}$. In other words, $F \in \mathcal{F}$ implies that $\phi(F) \geq \frac{1}{n}$. But the flocking matrix $F = \frac{1}{n} \mathbf{1}\mathbf{1}'$ is in \mathcal{F} , and for this matrix $\phi(F) = \frac{1}{n}$. Therefore

$$(16) \quad \min_{F \in \mathcal{F}} \phi(F) = \frac{1}{n}.$$

Now suppose that $S \in \mathcal{F}(q)$. Thus S is the product of q matrices from \mathcal{F} . From this, (14), and (16) it follows that $\phi(S) \geq \frac{1}{n^q}$. Therefore

$$(17) \quad \min_{S \in \mathcal{F}(q)} \phi(S) \geq \frac{1}{n^q}.$$

Next suppose that $S \in (\mathcal{F}(q))^m$. Thus S is the product of m matrices from $\mathcal{F}(q)$. From this, (14), and (17) it follows that $\phi(S) \geq \frac{1}{n^{qm}}$. This and (15) thus imply that $\|\|S\|\| \leq 1 - \frac{1}{n^{qm}}$ and thus that

$$\max_{S \in (\mathcal{F}(q))^{(n-1)^2}} \|\|S\|\| \leq 1 - \frac{1}{n^{qm}}.$$

Therefore, since $m = (n-1)^2$,

$$(18) \quad \lambda \leq \left(1 - \frac{1}{n^{q(n-1)^2}} \right)^{\frac{1}{q(n-1)^2}}.$$

The derivation of this particular upper bound on the rate at which the θ_i converge to θ_{ss} ultimately depends on two facts established in [6]. First, as we said before, the composition of at most $(n-1)^2$ rooted graphs is strongly rooted. Second, for any infinite sequence of stochastic matrices S_1, S_2, \dots , with strongly rooted graphs which come from a compact set \mathcal{S}_{sr} , the product $S_j \cdots S_1$ converges exponentially fast, as $j \rightarrow \infty$, to a rank one matrix $\mathbf{1}c$ at a rate no slower than

$$\max_{S \in \mathcal{S}_{sr}} \|\|S\|\|.$$

It turns out that by exploiting two different but corresponding facts about stochastic matrices with “neighbor-shared” graphs we can obtain a significantly smaller bound than the one given by (18).

3.2. Neighbor-shared graphs. By a neighbor-shared graph we mean any graph with two or more vertices with the property that each pair of vertices in the graph shares a common neighbor. Every neighbor-shared graph is rooted, but the converse is false [6]. The convergence rate bounds we are about to derive depend on two facts. First, the composition of at most $(n - 1)$ rooted graphs is neighbor shared [6]. Second, for any infinite sequence of stochastic matrices S_1, S_2, \dots , with neighbor-shared graphs which come from a compact set \mathcal{S}_{ns} , the product $S_j \cdots S_1$ converges exponentially fast, as $j \rightarrow \infty$, to a rank one matrix $\mathbf{1}c$ at a rate no slower than

$$\max_{S \in \mathcal{S}_{ns}} \mu(S),$$

where $\mu(S)$ is a positive number, called a *scrambling constant*, which is defined by the formula

$$(19) \quad \mu(S) = \max_{i,j} \left(1 - \sum_{k=1}^n \min\{s_{ik}, s_{jk}\} \right).$$

In what follows we will make use of some well-known ideas and constructions from the theory of nonhomogeneous Markov chains [17] to explain why this second statement is true.

Scrambling constants. Let S be any $n \times n$ stochastic matrix. Observe that for any nonnegative n -vector x , the i th minus the j th entries of Sx can be written as

$$\sum_{k=1}^n (s_{ik} - s_{jk})x_k = \sum_{k \in \mathcal{K}} (s_{ik} - s_{jk})x_k + \sum_{k \in \bar{\mathcal{K}}} (s_{ik} - s_{jk})x_k,$$

where

$$\mathcal{K} = \{k : s_{ik} - s_{jk} \geq 0, k \in \{1, 2, \dots, n\}\}$$

and

$$\bar{\mathcal{K}} = \{k : s_{ik} - s_{jk} < 0, k \in \{1, 2, \dots, n\}\}.$$

Therefore

$$\sum_{k=1}^n (s_{ik} - s_{jk})x_k \leq \left(\sum_{k \in \mathcal{K}} (s_{ik} - s_{jk}) \right) [x] + \left(\sum_{k \in \bar{\mathcal{K}}} (s_{ik} - s_{jk}) \right) [x].$$

But

$$\sum_{k \in \mathcal{K} \cup \bar{\mathcal{K}}} (s_{ik} - s_{jk}) = 0,$$

and so

$$\sum_{k \in \bar{\mathcal{K}}} (s_{ik} - s_{jk}) = - \sum_{k \in \mathcal{K}} (s_{ik} - s_{jk}).$$

Thus

$$\sum_{k=1}^n (s_{ik} - s_{jk})x_k \leq \left(\sum_{k \in \mathcal{K}} (s_{ik} - s_{jk}) \right) ([x] - [x]).$$

Now

$$\sum_{k \in \mathcal{K}} (s_{ik} - s_{jk}) = 1 - \sum_{k \in \bar{\mathcal{K}}} s_{ik} - \sum_{k \in \mathcal{K}} s_{jk}$$

because the row sums of S are all one. Moreover,

$$s_{ik} = \min\{s_{ik}, s_{jk}\}, \quad k \in \bar{\mathcal{K}},$$

$$s_{jk} = \min\{s_{ik}, s_{jk}\}, \quad k \in \mathcal{K},$$

and so

$$\sum_{k \in \mathcal{K}} (s_{ik} - s_{jk}) = 1 - \sum_{k=1}^n \min\{s_{ik}, s_{jk}\}.$$

It follows that

$$\sum_{k=1}^n (s_{ik} - s_{jk})x_k \leq \left(1 - \sum_{k=1}^n \min\{s_{ik}, s_{jk}\}\right) (\lceil x \rceil - \lfloor x \rfloor).$$

Hence with μ as defined by (19),

$$\sum_{k=1}^n (s_{ik} - s_{jk})x_k \leq \mu(S)(\lceil x \rceil - \lfloor x \rfloor).$$

Since this holds for all i, j , it must hold for the i and j for which

$$\sum_{k=1}^n s_{ik}x_k = \lceil Sx \rceil \quad \text{and} \quad \sum_{k=1}^n s_{jk}x_k = \lfloor Sx \rfloor.$$

Therefore

$$(20) \quad \lceil Sx \rceil - \lfloor Sx \rfloor \leq \mu(S)(\lceil x \rceil - \lfloor x \rfloor).$$

Now let S_1 and S_2 be any two $n \times n$ stochastic matrices and let e_i be the i th unit n -vector. Then from (20),

$$(21) \quad \lceil S_2 S_1 e_i \rceil - \lfloor S_2 S_1 e_i \rfloor \leq \mu(S_2)(\lceil S_1 e_i \rceil - \lfloor S_1 e_i \rfloor).$$

Meanwhile, from (9),

$$\lceil S_2 S_1 \rceil e_i = \mathbf{1}(\lceil S_2 S_1 \rceil - \lfloor S_2 S_1 \rfloor)e_i$$

and

$$\lceil S_1 \rceil e_i = \mathbf{1}(\lceil S_1 \rceil - \lfloor S_1 \rfloor)e_i.$$

But for any nonnegative matrix M , $\lceil M \rceil e_i = \lceil M e_i \rceil$ and $\lfloor M \rfloor e_i = \lfloor M e_i \rfloor$, and so

$$\lceil S_2 S_1 \rceil e_i = \mathbf{1}(\lceil S_2 S_1 e_i \rceil - \lfloor S_2 S_1 e_i \rfloor)$$

and

$$\lceil S_1 \rceil e_i = \mathbf{1}(\lceil S_1 e_i \rceil - \lfloor S_1 e_i \rfloor).$$

From these expressions and (21) it follows that

$$\llbracket S_2 S_1 \rrbracket e_i \leq \mu(S_2) \llbracket S_1 \rrbracket e_i.$$

Since this is true for all i , we arrive at the following fact.

LEMMA 1. *For any two stochastic matrices in \mathcal{S} ,*

$$(22) \quad \llbracket S_2 S_1 \rrbracket \leq \mu(S_2) \llbracket S_1 \rrbracket.$$

Note that since the row sums of S all equal 1, $\mu(S)$ is nonnegative. 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 k th term in the sum appearing in (19) 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* [17]. 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 graph $\gamma(S)$ is neighbor-shared.

As before, let \mathcal{S}_{n_s} be a closed subset consisting of stochastic matrices whose graphs are all neighbor-shared. Then the scrambling constant $\mu(S)$ defined in (19) satisfies $\mu(S) < 1$, $S \in \mathcal{S}_{n_s}$ because each such S is a scrambling matrix. Let

$$\bar{\mu} = \max_{S \in \mathcal{S}_{n_s}} \mu(S).$$

Then $\bar{\mu} < 1$ because \mathcal{S}_{n_s} is closed and bounded and because $\mu(\cdot)$ is continuous. In view of Lemma 1,

$$\|\llbracket S_2 S_1 \rrbracket\| \leq \bar{\mu} \|\llbracket S_1 \rrbracket\|, \quad S_1, S_2 \in \mathcal{S}_{n_s}.$$

Hence by induction, for any sequence of matrices S_1, S_2, \dots in \mathcal{S}_{n_s}

$$\|\llbracket S_j \cdots S_1 \rrbracket\| \leq \bar{\mu}^{j-1} \|\llbracket S_1 \rrbracket\|, \quad S_i \in \mathcal{S}_{n_s}.$$

But from (8), $\llbracket S \rrbracket \leq \llbracket S \rrbracket, S \in \mathcal{S}$, and so $\|\llbracket S \rrbracket\| \leq \|\llbracket S \rrbracket\|, S \in \mathcal{S}$. Therefore for any sequence of stochastic matrices S_1, S_2, \dots with neighbor-shared graphs

$$(23) \quad \|\llbracket S_j \cdots S_1 \rrbracket\| \leq \bar{\mu}^{j-1} \|\llbracket S_1 \rrbracket\|.$$

Therefore from Proposition 1, any such product $S_j \cdots S_1$ converges exponentially at a rate no slower than $\bar{\mu}$ as $j \rightarrow \infty$. This establishes the validity of the statement about convergence of products of stochastic matrices made at the beginning of section 3.2.

Suppose now that F is a flocking matrix for which $\gamma(F)$ is neighbor-shared. In view of the definition of a flocking matrix, any nonzero entry in F 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 nonzero, and so $\min\{f_{ik}, f_{jk}\} \geq \frac{1}{n}$. This implies that the sum in (19) must be bounded below by $\frac{1}{n}$ and consequently that $\mu(F) \leq 1 - \frac{1}{n}$.

Now let F be that flocking matrix whose graph $\gamma(F)$ 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, $f_{1k} = 0$ for all $k > 1$. 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, and 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 , $\mu(F)$ attains its upper bound of $1 - \frac{1}{n}$. We summarize.

LEMMA 2. *Let \mathcal{F}_{ns} be the set of $n \times n$ flocking matrices with neighbor-shared graphs. Then*

$$(24) \quad \max_{F \in \mathcal{F}_{ns}} \mu(F) = 1 - \frac{1}{n}.$$

Thus $1 - \frac{1}{n}$ is a tight bound on the convergence rate for an infinite product of flocking matrices with neighbor-shared graphs. In [6] it is shown that

$$\max_{F \in \mathcal{F}_{sr}} \mu(F) = 1 - \frac{1}{n},$$

where \mathcal{F}_{sr} is the set of flocking matrices with strongly rooted graphs. Thus $1 - \frac{1}{n}$ is also a tight bound on the convergence rate for an infinite product of flocking matrices with strongly rooted graphs. Of course a strongly rooted graph is a more special type of graph than a neighbor-shared graph because strongly rooted graphs are neighbor shared but not conversely.

We now use the preceding to derive a better convergence rate bound than the one in (18) for the type of trajectory addressed by Theorem 1. As a first step towards this end, we exploit the fact that for any $n \times n$ stochastic scrambling matrix S , the scrambling constant of $\mu(S)$ satisfies the inequality

$$(25) \quad \mu(S) \leq 1 - \phi(S).$$

To understand why this is so, assume that S is any given scrambling matrix. Note that for any distinct i and j , there must be a k for which $\min\{s_{ik}, s_{jk}\}$ is nonzero and bounded below by $\phi(S)$. Thus

$$\sum_{k=1}^n \min\{s_{ik}, s_{jk}\} \geq \phi(S),$$

and so

$$1 - \sum_{k=1}^n \min\{s_{ik}, s_{jk}\} \leq 1 - \phi(S).$$

But this holds for all distinct i and j . In view of the definition of $\mu(S)$ in (19), (25) must therefore be true.

As before, let $\mathcal{F}(q)$ denote the set of products of q flocking matrices F_1, F_2, \dots, F_q from \mathcal{F} for which $\{\gamma(F_1), \gamma(F_2), \dots, \gamma(F_q)\}$ is a jointly rooted set. Then as noted before, each matrix in $\mathcal{F}(q)$ is rooted. Set $p = (n - 1)$ and let $(\mathcal{F}(q))^p$ now denote the set of all products of p matrices from $\mathcal{F}(q)$. Then each matrix in $(\mathcal{F}(q))^p$ is neighbor-shared. Let S be any such matrix. Then S is a product of qp flocking matrices. But each such flocking matrix F satisfies (16). Because of this and (14), it must be true that $\phi(S) \geq \frac{1}{n^{qp}}$. Therefore $\mu(S) \leq 1 - \frac{1}{n^{qp}}$ because of (25). Since this is true for all $S \in (\mathcal{F}(q))^p$, $1 - \frac{1}{n^{qp}}$ must be a convergence rate upper bound for all infinite products of matrices from $(\mathcal{F}(q))^p$. Therefore, since $p = n - 1$,

$$(26) \quad \left(1 - \frac{1}{n^{q(n-1)}}\right)^{\frac{1}{q(n-1)}}$$

must be an upper bound on the convergence rate for all infinite products of flocking matrices F_1, F_2, \dots which have the property that the sequence of graphs $\gamma(F_1), \gamma(F_2), \dots$ is repeatedly jointly rooted by subsequences of length q . Since this is precisely the type of sequence of flocking matrices which arise under the assumptions of Theorem 1, (26) is a convergence rate bound for the type of trajectory addressed by the theorem. Note that this convergence rate upper bound is much smaller (i.e., faster) than the one given by (18). We refer the reader to [7] for additional convergence rate calculations along these lines.

4. Measurement delays. In this section we consider a modified version of the flocking problem in which integer-valued delays occur in sensing the values of headings which are available to agents. More precisely we suppose that at each time $t \in \{0, 1, 2, \dots\}$, the value of neighboring agent j 's headings which agent i may sense is $\theta_j(t - d_{ij}(t))$, where $d_{ij}(t)$ is a delay whose value at t is some integer between 0 and $m_j - 1$; here m_j is a prespecified positive integer. While well-established principles of feedback control would suggest that delays should be dealt with using dynamic compensation, in this paper we will consider the situation in which the delayed value of agent j 's heading sensed by agent i at time t is the value which will be used in the heading update law for agent i . Thus

$$(27) \quad \theta_i(t + 1) = \frac{1}{n_i(t)} \left(\sum_{j \in \mathcal{N}_i(t)} \theta_j(t - d_{ij}(t)) \right),$$

where $d_{ij}(t) \in \{0, 1, \dots, (m_j - 1)\}$ if $j \neq i$ and $d_{ij}(t) = 0$ if $i = j$. Our main result is the following theorem, which states in essence that the conclusions of Theorem 1 continue to hold for the update model described by (27).

THEOREM 2. *Let $\theta(0)$ be fixed. For any trajectory of the system determined by (27) along which the sequence of neighbor graphs $\mathbb{N}(0), \mathbb{N}(1), \dots$ is repeatedly jointly rooted, there is a constant θ_{ss} , depending only on $\theta(0)$, for which*

$$(28) \quad \lim_{t \rightarrow \infty} \theta(t) = \theta_{ss} \mathbf{1},$$

where the limit is approached exponentially fast.

As noted in the introduction, the consensus problem with measurement delays we have been discussing has been considered previously in [3]. It is possible to compare the hypotheses of Theorem 2 with the corresponding hypotheses for exponential convergence stated in [3], namely assumptions 2 and 3 of that paper. To do this, let us agree, as before, to say that the *union* of a set of graphs $\mathbb{G}_{r_1}, \mathbb{G}_{r_2}, \dots, \mathbb{G}_{r_k}$ with vertex set \mathcal{V} is the graph with the vertex set \mathcal{V} and arc set consisting of the union of the arcs of all of the graphs $\mathbb{G}_{r_1}, \mathbb{G}_{r_2}, \dots, \mathbb{G}_{r_k}$. Taken together, assumptions 2 and 3 of [3] are essentially equivalent to assuming that there are finite positive integers q and s such that the *union*

$$\mathbb{G}(k) \triangleq \mathbb{N}((k + 1)q - 1) \cup \mathbb{N}((k + 1)q - 2) \cup \dots \cup \mathbb{N}(kq)$$

is strongly connected and independent of k for $k \geq s$. By way of comparison, the hypothesis of Theorem 2 is equivalent to assuming that there is a finite positive integer q such that the *composition*

$$\bar{\mathbb{G}}(k) \triangleq \mathbb{N}((k + 1)q - 1) \circ \mathbb{N}((k + 1)q - 2) \circ \dots \circ \mathbb{N}(kq)$$

is rooted for $k \geq 0$. The latter assumption is weaker than the former for several reasons. First, the arc set of $\mathbb{G}(k)$ is always a subset of the arc set of $\bar{\mathbb{G}}(k)$, and in some cases the containment may be strict. Second, $\bar{\mathbb{G}}(k)$ is not assumed to be independent of k , even for k sufficiently large, whereas $\mathbb{G}(k)$ is; in other words, $\bar{\mathbb{G}}(k)$ is not assumed to converge, whereas $\mathbb{G}(k)$ is. Third, each $\mathbb{G}(k)$ is assumed to be strongly connected, whereas each $\bar{\mathbb{G}}(k)$ need only be rooted; note that a strongly connected graph is a special type of rooted graph in which every vertex is a root. Perhaps what is most important about Theorem 2 and the development which justifies it is that the underlying structural properties of the graphs involved required for consensus are explicitly determined.

4.1. State space system. Using standard lifting techniques for dealing with delays in discrete-time systems, it is possible to represent the agent system defined by (27) as a state space model similar to the model discussed earlier for the delay-free case. Our first objective is to characterize the class of graphs \mathcal{D} of the stochastic matrices which result from this lifting process. Towards this end, let $\bar{\mathcal{G}}$ denote the set of all directed graphs with vertex set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2 \cup \dots \cup \mathcal{V}_n$, where $\mathcal{V}_i = \{v_{i1}, \dots, v_{im_i}\}$. Here vertex v_{ij} labels the j th possible delay value of agent i , namely $j - 1$. We sometimes write i for v_{i1} , $i \in \{1, 2, \dots, n\}$, write \mathcal{V} for the subset of vertices $\{v_{11}, v_{21}, \dots, v_{n1}\}$, and think of v_{i1} as an alternative label of agent i .

To take account of the fact that each agent can use its own current heading in its update formula (27), we will utilize those graphs in $\bar{\mathcal{G}}$ which have self-arcs at each vertex in \mathcal{V} . We will also require the arc set of each such graph to have, for $i \in \{1, 2, \dots, n\}$, an arc from each vertex $v_{ij} \in \mathcal{V}_i$ except the last to its successor $v_{i(j+1)} \in \mathcal{V}_i$. Finally we stipulate that for each $i \in \{1, 2, \dots, n\}$, each vertex v_{ij} with $j > 1$ has in-degree of exactly 1. In what follows we call any such graph a *delay graph* and write \mathcal{D} for the subset of all such graphs. Note that unlike the class of graphs \mathcal{G}_{sa} considered before, there are graphs in \mathcal{D} possessing vertices without self-arcs. Nonetheless each vertex of each graph in \mathcal{D} has positive in-degree. An example of a delay graph for a three-agent system is shown in Figure 1.

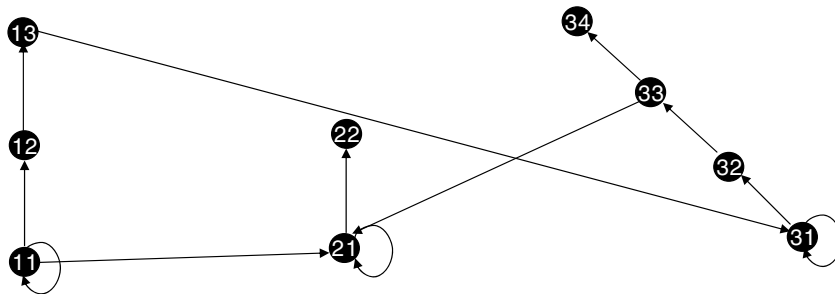


FIG. 1. *Delay graph.*

The specific delay graph representing the sensed headings the agents use at time t to update their own headings according to (27) is the graph $\mathbb{D}(t) \in \mathcal{D}$ whose arc set contains an arc from $v_{ik} \in \mathcal{V}_i$ to $v_{j1} \in \mathcal{V}$ if agent j uses $\theta_i(t + 1 - k)$ to update. There is a simple relationship between $\mathbb{D}(t)$ and the neighbor graph $\mathbb{N}(t)$ defined earlier. In particular,

$$(29) \quad \mathbb{N}(t) = Q(\mathbb{D}(t)),$$

where $Q(\mathbb{D}(t))$ is the “quotient graph” of $\mathbb{D}(t)$. By the *quotient graph* of any $\mathbb{G} \in \bar{\mathcal{G}}$, written $Q(\mathbb{G})$, we mean the directed graph in \mathcal{G} with vertex set \mathcal{V} whose arc set consists of those arcs (i, j) for which \mathbb{G} has an arc from some vertex in \mathcal{V}_i to some vertex in \mathcal{V}_j . The quotient graph of $\mathbb{D}(t)$ thus models which headings are being used by each agent in updates at time t without describing the specific delayed headings actually being used. The quotient graph of the delay graph in Figure 1 is shown in Figure 2.

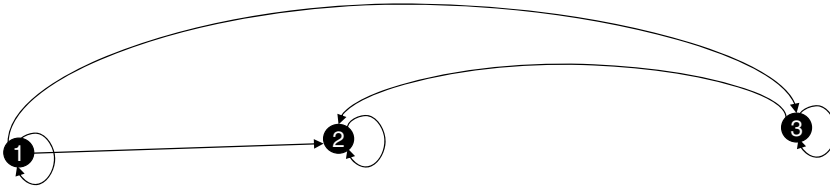


FIG. 2. Quotient graph.

The set of agent heading update rules defined by (27) can be written in state form. Towards this end, define $\theta(t)$ to be the $(m_1 + m_2 + \dots + m_i)$ vector whose first m_1 elements are $\theta_1(t)$ to $\theta_1(t+1-m_1)$, whose next m_2 elements are $\theta_2(t)$ to $\theta_2(t+1-m_2)$, and so on. Order the vertices of \mathcal{V} as $v_{11}, \dots, v_{1m_1}, v_{21}, \dots, v_{2m_2}, \dots, v_{n1}, \dots, v_{nm_n}$, and with respect to this ordering define for each graph $\mathbb{D} \in \mathcal{D}$ the flocking matrix

$$(30) \quad F = D^{-1}A',$$

where A' is the transpose of the adjacency matrix of \mathbb{D} and D the diagonal matrix whose ij th diagonal element is the in-degree of vertex v_{ij} within the graph. Then $\gamma(F) = \mathbb{D}$ and

$$(31) \quad \theta(t + 1) = F(t)\theta(t), \quad t \in \{0, 1, 2, \dots\}.$$

Let $\bar{\mathcal{F}}$ denote the set of all such F . As before our goal is to characterize the sequences of neighbor graphs $\mathbb{N}(0), \mathbb{N}(1), \dots$ for which all entries of $\theta(t)$ converge to a common steady state value.

There are a number of similarities and a number of differences between the situation under consideration here and the delay-free situation considered in [6]. For example, the notion of graph composition defined earlier can be defined in the obvious way for graphs in $\bar{\mathcal{G}}$. On the other hand, unlike the situation in the delay-free case, the set of graphs used to model the system under consideration, namely the set of delay graphs \mathcal{D} , is not closed under composition except in the special case when all of the delays are at most 1, i.e., when all of the $m_i \leq 2$. In order to characterize the smallest subset of $\bar{\mathcal{G}}$ containing \mathcal{D} which is closed under composition, we will need several new concepts.

4.2. Hierarchical graphs. As before, let \mathcal{G} be the set of all directed graphs with vertex set $\mathcal{V} = \{1, 2, \dots, n\}$. Let us agree to say that a rooted graph $\mathbb{G} \in \mathcal{G}$ is a *hierarchical graph* with *hierarchy* $\{v_1, v_2, \dots, v_n\}$ if it is possible to relabel the vertices in \mathcal{V} as v_1, v_2, \dots, v_n in such a way so that v_1 is a root of \mathbb{G} with a self-arc and for $i > 1$, v_i has a neighbor v_j “lower” in the hierarchy where by *lower* we mean $j < i$. It is clear that any graph in \mathcal{G} with a root possessing a self-arc is hierarchical. Note that a graph may have more than one hierarchy and two graphs with the same hierarchy need not be equal. Note also that even though rooted graphs with the same hierarchy

share a common root, examples show that the composition of hierarchical graphs in \mathcal{G} need not be hierarchical or even rooted. On the other hand, the composition of two rooted graphs in \mathcal{G} with the same hierarchy is always a graph with the same hierarchy. To understand why this is so, consider two graphs \mathbb{G}_1 and \mathbb{G}_2 in \mathcal{G} with the same hierarchy $\{v_1, v_2, \dots, v_n\}$. Note first that v_1 has a self-arc in $\mathbb{G}_2 \circ \mathbb{G}_1$ because v_1 has self-arcs in \mathbb{G}_1 and \mathbb{G}_2 . Next pick any vertex v_i in \mathcal{V} other than v_1 . By definition, there must exist vertex v_j lower in the hierarchy than v_i such that (v_j, v_i) is an arc of \mathbb{G}_2 . If $v_j = v_1$, then (v_1, v_i) is an arc in $\mathbb{G}_2 \circ \mathbb{G}_1$ because v_1 has a self-arc in \mathbb{G}_1 . On the other hand, if $v_j \neq v_1$, then there must exist a vertex v_k lower in the hierarchy than v_j such that (v_k, v_j) is an arc of \mathbb{G}_1 . It follows from the definition of composition that in this case (v_k, v_i) is an arc in $\mathbb{G}_2 \circ \mathbb{G}_1$. Thus v_i has a neighbor in $\mathbb{G}_2 \circ \mathbb{G}_1$ which is lower in the hierarchy than v_i . Since this is true for all v_i , $\mathbb{G}_2 \circ \mathbb{G}_1$ must have the same hierarchy as \mathbb{G}_1 and \mathbb{G}_2 . This proves the claim that composition of two rooted graphs with the same hierarchy is a graph with the same hierarchy.

Our objective is to show that the composition of a sufficiently large number of graphs in \mathcal{G} with the same hierarchy is strongly rooted. Note that the fact that the composition of $(n-1)^2$ graphs in \mathcal{G}_{sa} is rooted [6] cannot be used to reach this conclusion because the v_i in the graphs under consideration here do not all necessarily have self-arcs.

As before, let \mathbb{G}_1 and \mathbb{G}_2 be two graphs in \mathcal{G} with the same hierarchy $\{v_1, v_2, \dots, v_n\}$. Let v_i be any vertex in the hierarchy and suppose that v_j is a neighbor vertex of v_i in \mathbb{G}_2 . If $v_j = v_1$, then v_i retains v_1 as a neighbor in the composition $\mathbb{G}_2 \circ \mathbb{G}_1$ because v_1 has a self-arc in \mathbb{G}_1 . On the other hand, if $v_j \neq v_1$, then v_j has a neighboring vertex v_k in \mathbb{G}_1 which is lower in the hierarchy than v_j . Since v_k is a neighbor of v_i in the composition $\mathbb{G}_2 \circ \mathbb{G}_1$, we see that in this case v_i has acquired a neighbor in $\mathbb{G}_2 \circ \mathbb{G}_1$ lower in the hierarchy than a neighbor it had in \mathbb{G}_2 . In summary, any vertex $v_i \in \mathcal{V}$ either has v_1 as neighbor in $\mathbb{G}_2 \circ \mathbb{G}_1$ or has a neighbor in $\mathbb{G}_2 \circ \mathbb{G}_1$ which is at least one vertex lower in the hierarchy than any neighbor it had in \mathbb{G}_2 .

Now consider three graphs $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3$ in \mathcal{G} with the same hierarchy. By the same reasoning as above, any vertex $v_i \in \mathcal{V}$ either has v_1 as neighbor in $\mathbb{G}_3 \circ \mathbb{G}_2 \circ \mathbb{G}_1$ or has a neighbor in $\mathbb{G}_3 \circ \mathbb{G}_2 \circ \mathbb{G}_1$ which is at least one vertex lower in the hierarchy than any neighbor it had in $\mathbb{G}_3 \circ \mathbb{G}_2$. Similarly v_i either has v_1 as neighbor in $\mathbb{G}_3 \circ \mathbb{G}_2$ or has a neighbor in $\mathbb{G}_3 \circ \mathbb{G}_2$ which is at least one vertex lower in the hierarchy than any neighbor it had in \mathbb{G}_3 . Combining these two observations we see that any vertex $v_i \in \mathcal{V}$ either has v_1 as neighbor in $\mathbb{G}_3 \circ \mathbb{G}_2 \circ \mathbb{G}_1$ or has a neighbor in $\mathbb{G}_3 \circ \mathbb{G}_2 \circ \mathbb{G}_1$ which is at least two vertices lower in the hierarchy than any neighbor it had in \mathbb{G}_3 . This clearly generalizes, and so after the composition of m such graphs $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_m$, v_i either has v_1 as neighbor in $\mathbb{G}_m \circ \dots \circ \mathbb{G}_2 \circ \mathbb{G}_1$ or has a neighbor in $\mathbb{G}_m \circ \dots \circ \mathbb{G}_2 \circ \mathbb{G}_1$ which is at least $m-1$ vertices lower in the hierarchy than any neighbor it had in \mathbb{G}_m . It follows that if $m \geq n$, then v_i must be a neighbor of v_1 . Since this is true for all vertices, we have proved the following.

PROPOSITION 2. *Let $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_m$ denote a set of rooted graphs in \mathcal{G} which all have the same hierarchy. If $m \geq n-1$, then $\mathbb{G}_m \circ \dots \circ \mathbb{G}_2 \circ \mathbb{G}_1$ is strongly rooted.*

4.3. The Closure of \mathcal{D} . We now return to the study of the graphs in \mathcal{D} . As before \mathcal{D} is the subset of $\bar{\mathcal{G}}$ consisting of those graphs which (i) have self-arcs at each vertex in $\mathcal{V} = \{v_{11}, v_{21}, \dots, v_{n1}\}$, (ii) for each $i \in \{1, 2, \dots, n\}$ have an arc from each vertex $v_{ij} \in \mathcal{V}_i$ except the last to its successor $v_{i(j+1)} \in \mathcal{V}_i$, and (iii) for each $i \in \{1, 2, \dots, n\}$, each vertex v_{ij} with $j > 1$ has in-degree of exactly 1. It can easily be shown by example that \mathcal{D} is not closed under composition. We deal with this

problem as follows. First, let us agree to say that a vertex v in a graph $\mathbb{G} \in \bar{\mathcal{G}}$ is a *neighbor* of a subset of \mathbb{G} 's vertices \mathcal{U} if v is a neighbor of at least one vertex in \mathcal{U} . Next we say that a graph $\mathbb{G} \in \bar{\mathcal{G}}$ is an *extended delay graph* if for each $i \in \{1, 2, \dots, n\}$, (i) every neighbor of \mathcal{V}_i which is not in \mathcal{V}_i is a neighbor of v_{i1} and (ii) the subgraph of \mathbb{G} induced by \mathcal{V}_i has $\{v_{i1}, \dots, v_{im_i}\}$ as a hierarchy. We write $\bar{\mathcal{D}}$ for the set of all extended delay graphs in $\bar{\mathcal{G}}$. It is easy to see that every delay graph is an extended delay graph. The converse, however, is not true. The set of extended delay graphs has the following property.

PROPOSITION 3. $\bar{\mathcal{D}}$ is closed under composition.

In light of this proposition it is natural to call $\bar{\mathcal{D}}$ the *closure* of \mathcal{D} . To prove the proposition, we will need the following fact.

LEMMA 3. Let $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_q$ be any sequence of $q > 1$ directed graphs with vertex set \mathcal{V} . For $i \in \{1, 2, \dots, q\}$, let $\bar{\mathbb{G}}_i$ be the subgraph of \mathbb{G}_i induced by $\mathcal{U} \subset \mathcal{V}$. Then $\bar{\mathbb{G}}_q \circ \dots \circ \bar{\mathbb{G}}_2 \circ \bar{\mathbb{G}}_1$ is contained in the subgraph of $\mathbb{G}_q \circ \dots \circ \mathbb{G}_2 \circ \mathbb{G}_1$ induced by \mathcal{U} .

Proof of Lemma 3. It will be enough to prove the lemma for $q = 2$, since the proof for $q > 2$ would then directly follow by induction. Suppose $q = 2$. Let (i, j) be in $\mathcal{A}(\bar{\mathbb{G}}_2 \circ \bar{\mathbb{G}}_1)$. Then $i, j \in \mathcal{U}$ and there exists an integer $k \in \mathcal{U}$ such that $(i, k) \in \mathcal{A}(\bar{\mathbb{G}}_1)$ and $(k, j) \in \mathcal{A}(\bar{\mathbb{G}}_2)$. Therefore $(i, k) \in \mathcal{A}(\mathbb{G}_1)$ and $(k, j) \in \mathcal{A}(\mathbb{G}_2)$. Thus $(i, j) \in \mathcal{A}(\mathbb{G}_2 \circ \mathbb{G}_1)$. But $i, j \in \mathcal{S}$, and so (i, j) must be an arc in the subgraph of $\mathbb{G}_2 \circ \mathbb{G}_1$ induced by \mathcal{U} . Since this clearly is true for all arcs in $\mathcal{A}(\bar{\mathbb{G}}_2 \circ \bar{\mathbb{G}}_1)$, the proof is complete. \square

Proof of Proposition 3. Let \mathbb{G}_1 and \mathbb{G}_2 be two extended delay graphs in $\bar{\mathcal{D}}$. It will first be shown that for each $i \in \{1, 2, \dots, n\}$, every neighbor of \mathcal{V}_i which is not in \mathcal{V}_i is a neighbor of v_{i1} in $\mathbb{G}_2 \circ \mathbb{G}_1$. Fix $i \in \{1, 2, \dots, n\}$ and let v be a neighbor of \mathcal{V}_i in $\mathbb{G}_2 \circ \mathbb{G}_1$ which is not in \mathcal{V}_i . Then $(v, k) \in \mathcal{A}(\mathbb{G}_2 \circ \mathbb{G}_1)$ for some $k \in \mathcal{V}_i$. Thus there is a $s \in \bar{\mathcal{V}}$ such that $(v, s) \in \mathcal{A}(\mathbb{G}_1)$ and $(s, k) \in \mathcal{A}(\mathbb{G}_2)$. If $s \notin \mathcal{V}_i$, then $(s, v_{i1}) \in \mathcal{A}(\mathbb{G}_2)$ because \mathbb{G}_2 is an extended delay graph. Thus in this case $(v, v_{i1}) \in \mathcal{A}(\mathbb{G}_2 \circ \mathbb{G}_1)$ because of the definition of composition. If, on the other hand, $s \in \mathcal{V}_i$, then $(v, v_{i1}) \in \mathcal{A}(\mathbb{G}_1)$ because \mathbb{G}_1 is an extended delay graph. Thus in this case $(v, v_{i1}) \in \mathcal{A}(\mathbb{G}_2 \circ \mathbb{G}_1)$ because v_{i1} has a self-arc in \mathbb{G}_2 . This proves that every neighbor of \mathcal{V}_i which is not in \mathcal{V}_i is a neighbor of v_{i1} in $\mathbb{G}_2 \circ \mathbb{G}_1$. Since this must be true for each $i \in \{1, 2, \dots, n\}$, $\mathbb{G}_2 \circ \mathbb{G}_1$ has the first property defining extended delay graphs in $\bar{\mathcal{D}}$.

To establish the second property, we exploit the fact that the composition of two graphs with the same hierarchy is a graph with the same hierarchy. Thus for any integer $i \in \{1, 2, \dots, n\}$, the composition of the subgraphs of \mathbb{G}_1 and \mathbb{G}_2 , respectively, induced by \mathcal{V}_i must have the hierarchy $\{v_{i1}, v_{i2}, \dots, v_{im_i}\}$. But by Lemma 3, for any integer $i \in \{1, 2, \dots, n\}$, the composition of the subgraphs of \mathbb{G}_1 and \mathbb{G}_2 , respectively, induced by \mathcal{V}_i is contained in the subgraph of the composition of \mathbb{G}_1 and \mathbb{G}_2 induced by \mathcal{V}_i . This implies that for $i \in \{1, 2, \dots, n\}$, the subgraph of the composition of \mathbb{G}_1 and \mathbb{G}_2 induced by \mathcal{V}_i has $\{v_{i1}, v_{i2}, \dots, v_{im_i}\}$ as a hierarchy. \square

Our main result regarding extended delay graphs is as follows.

PROPOSITION 4. Let m be the largest integer in the set $\{m_1, m_2, \dots, m_n\}$. The composition of any set of at least $m(n-1)^2 + m - 1$ extended delay graphs will be strongly rooted if the quotient graph of each of the graphs in the composition is rooted.

To prove this proposition we will need several more concepts. Let us agree to say that an extended delay graph $\mathbb{G} \in \bar{\mathcal{D}}$ has *strongly rooted hierarchies* if for each $i \in \mathcal{V}$, the subgraph of \mathbb{G} induced by \mathcal{V}_i is strongly rooted. Proposition 2 states that a hierarchical graph on m_i vertices will be strongly rooted if it is the composition of at least $m_i - 1$ rooted graphs with the same hierarchy. This and Lemma 3 imply that

the subgraph of the composition of at least $m_i - 1$ extended delay graphs induced by \mathcal{V}_i will be strongly rooted. We are led to the following lemma.

LEMMA 4. *Any composition of at least $m - 1$ extended delay graphs in $\bar{\mathcal{D}}$ has strongly rooted hierarchies.*

To proceed we will need one more type of graph which is uniquely determined by a given graph in $\bar{\mathcal{G}}$. By the *agent subgraph* of $\mathbb{G} \in \bar{\mathcal{G}}$ we mean the subgraph of \mathbb{G} induced by \mathcal{V} . Note that while the quotient graph of \mathbb{G} describes relations between distinct agent hierarchies, the agent subgraph of \mathbb{G} captures only the relationships between the roots of the hierarchies. Note in addition that both the agent subgraph of \mathbb{G} and the quotient graph of \mathbb{G} are graphs in \mathcal{G}_{sa} because all n vertices of \mathbb{G} in \mathcal{V} have self-arcs. The agent subgraph of the graph in Figure 1 is shown in Figure 3.



FIG. 3. *Agent subgraph.*

LEMMA 5. *Let \mathbb{G}_p and \mathbb{G}_q be extended delay graphs in $\bar{\mathcal{D}}$. If \mathbb{G}_p has a strongly rooted agent subgraph and \mathbb{G}_q has strongly rooted hierarchies, then the composition $\mathbb{G}_q \circ \mathbb{G}_p$ is strongly rooted.*

Proof of Lemma 5. Let v_{i1} be a root of the agent subgraph of \mathbb{G}_p and let v_{jk} be any vertex in $\bar{\mathcal{V}}$. Then $(v_{i1}, v_{j1}) \in \mathcal{A}(\mathbb{G}_p)$ because the agent subgraph of \mathbb{G}_p is strongly rooted. Moreover, $(v_{j1}, v_{jk}) \in \mathcal{A}(\mathbb{G}_q)$ because \mathbb{G}_q has strongly rooted hierarchies. Therefore, in view of the definition of graph composition, $(v_{i1}, v_{jk}) \in \mathcal{A}(\mathbb{G}_q \circ \mathbb{G}_p)$. Since this must be true for every vertex $v_{jk} \in \bar{\mathcal{V}}$, $\mathbb{G}_q \circ \mathbb{G}_p$ is strongly rooted. \square

LEMMA 6. *The agent subgraph of any composition of at least $(n - 1)^2$ extended delay graphs in $\bar{\mathcal{D}}$ will be strongly rooted if the agent subgraph of each of the graphs in the composition is rooted.*

Proof of Lemma 6. Let $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_q$ be any sequence of $q \geq (n - 1)^2$ extended delay graphs in $\bar{\mathcal{D}}$ whose agent subgraphs, $\bar{\mathbb{G}}_i$, $i \in \{1, 2, \dots, q\}$, are all rooted. Since the $\bar{\mathbb{G}}_i$ are in \mathcal{G}_{sa} , Proposition 3 of [6] applies, and it can therefore be concluded that $\bar{\mathbb{G}}_q \circ \dots \circ \bar{\mathbb{G}}_2 \circ \bar{\mathbb{G}}_1$ is strongly rooted. But $\bar{\mathbb{G}}_q \circ \dots \circ \bar{\mathbb{G}}_2 \circ \bar{\mathbb{G}}_1$ is contained in the agent subgraph of $\mathbb{G}_q \circ \dots \circ \mathbb{G}_2 \circ \mathbb{G}_1$ because of Lemma 3. Therefore the agent subgraph of $\mathbb{G}_q \circ \dots \circ \mathbb{G}_2 \circ \mathbb{G}_1$ is strongly rooted. \square

LEMMA 7. *Let \mathbb{G}_p and \mathbb{G}_q be extended delay graphs in $\bar{\mathcal{D}}$. If \mathbb{G}_p has strongly rooted hierarchies and \mathbb{G}_q has a rooted quotient graph, then the agent subgraph of the composition $\mathbb{G}_q \circ \mathbb{G}_p$ is rooted.*

Proof of Lemma 7. Let (i, j) be any arc in the quotient graph of \mathbb{G}_q with $i \neq j$. This means that $(v_{ik}, v_{js}) \in \mathcal{A}(\mathbb{G}_q)$ for some $v_{ik} \in \mathcal{V}_i$ and $v_{js} \in \mathcal{V}_j$. Clearly $(v_{i1}, v_{ik}) \in \mathcal{A}(\mathbb{G}_p)$ because \mathbb{G}_p has strongly rooted hierarchies. Moreover, since $i \neq j$, v_{ik} is a neighbor of \mathcal{V}_j which is not in \mathcal{V}_j . From this and the definition of an extended delay graph, it follows that v_{ik} is a neighbor of v_{j1} . Therefore $(v_{ik}, v_{j1}) \in \mathcal{A}(\mathbb{G}_q)$. Thus $(v_{i1}, v_{j1}) \in \mathcal{A}(\mathbb{G}_q \circ \mathbb{G}_p)$. We have therefore proved that for any path of length one between any two distinct vertices i, j in the quotient graph of \mathbb{G}_q , there is a corresponding path between vertices v_{i1} and v_{j1} in the agent subgraph of $\mathbb{G}_q \circ \mathbb{G}_p$. This implies that for any path of any length between any two distinct vertices i, j in the quotient graph of \mathbb{G}_q , there is a corresponding path between vertices v_{i1} and v_{j1} in the agent subgraph of $\mathbb{G}_q \circ \mathbb{G}_p$. Since by assumption the quotient graph of \mathbb{G}_q is rooted, the agent subgraph of $\mathbb{G}_q \circ \mathbb{G}_p$ must be rooted as well. \square

Proof of Proposition 4. Let $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_s$ be a sequence of at least $m(n-1)^2 + m - 1$ extended delay graphs with rooted quotient graphs. The graph $\mathbb{G}_s \circ \dots \circ \mathbb{G}_{(m(n-1)^2+1)}$ is composed of at least $m - 1$ extended delay graphs. Therefore $\mathbb{G}_s \circ \dots \circ \mathbb{G}_{(m(n-1)^2+1)}$ must have strongly rooted hierarchies because of Lemma 4. In view of Lemma 5, to complete the proof it is enough to show that $\mathbb{G}_{m(n-1)^2} \circ \dots \circ \mathbb{G}_1$ has a strongly rooted agent subgraph. But $\mathbb{G}_{m(n-1)^2} \circ \dots \circ \mathbb{G}_1$ is the composition of $(n - 1)^2$ graphs, each itself a composition of m extended delay graphs with rooted quotient graphs. In view of Lemma 6, to complete the proof it is enough to show that the agent subgraph of any composition of m extended delay graphs is rooted if each quotient graph of each extended delay graph in the composition is rooted. Let $\mathbb{H}_1, \mathbb{H}_2, \dots, \mathbb{H}_m$ be such a family of extended delay graphs. By assumption, \mathbb{H}_m has a rooted quotient graph. In view of Lemma 7, the agent subgraph of $\mathbb{H}_m \circ \mathbb{H}_{m-1} \circ \dots \circ \mathbb{H}_1$ will be rooted if $\mathbb{H}_{m-1} \circ \dots \circ \mathbb{H}_1$ has strongly rooted hierarchies. But $\mathbb{H}_{m-1} \circ \dots \circ \mathbb{H}_1$ has this property because of Lemma 4. \square

Finally we will need the following fact.

PROPOSITION 5. *Let $\mathbb{G}_1, \dots, \mathbb{G}_r$ be a sequence of extended delay graphs in $\bar{\mathcal{D}}$. If the composition $Q(\mathbb{G}_r) \circ \dots \circ Q(\mathbb{G}_1)$ is rooted, then so is the quotient graph $Q(\mathbb{G}_r \circ \dots \circ \mathbb{G}_1)$.*

This proposition is a direct consequence of the following lemma.

LEMMA 8. *Let $\mathbb{G}_p, \mathbb{G}_q$ be two extended delay graphs in $\bar{\mathcal{D}}$. For each arc (i, j) in the composition $Q(\mathbb{G}_q) \circ Q(\mathbb{G}_p)$, there is a path from i to j in the quotient graph $Q(\mathbb{G}_q \circ \mathbb{G}_p)$.*

Proof of Lemma 8. Fix $(i, j) \in \mathcal{A}(Q(\mathbb{G}_q) \circ Q(\mathbb{G}_p))$. If $i = j$, then $(i, j) \in \mathcal{A}(Q(\mathbb{G}_q \circ \mathbb{G}_p))$ because $Q(\mathbb{G}_q \circ \mathbb{G}_p) \in \mathcal{G}_{sa}$. Thus in this case there is a path of length 1 from i to j in $Q(\mathbb{G}_q \circ \mathbb{G}_p)$.

Suppose $i \neq j$. Since $(i, j) \in \mathcal{A}(Q(\mathbb{G}_q) \circ Q(\mathbb{G}_p))$, there exists an integer $k \in \mathcal{V}$ such that $(i, k) \in \mathcal{A}(Q(\mathbb{G}_p))$ and $(k, j) \in \mathcal{A}(Q(\mathbb{G}_q))$. Thus there are integers $v_{is}, v_{kt}, v_{ku}, v_{jw} \in \mathcal{V}$ such that $(v_{is}, v_{kt}) \in \mathcal{A}(\mathbb{G}_p)$ and $(v_{ku}, v_{jw}) \in \mathcal{A}(\mathbb{G}_q)$. Since $\mathbb{G}_p \in \bar{\mathcal{D}}$, \mathbb{G}_p has a hierarchy rooted at v_{k1} . This means that there must be a vertex v_{kx} no higher in this hierarchy than v_{ku} such that $(v_{kx}, v_{ku}) \in \mathcal{A}(\mathbb{G}_p)$. Therefore $(v_{kx}, v_{jw}) \in \mathcal{A}(\mathbb{G}_q \circ \mathbb{G}_p)$. If $k = i$, then $(v_{ix}, v_{jw}) \in \mathcal{A}(\mathbb{G}_q \circ \mathbb{G}_p)$, and so $(i, j) \in \mathcal{A}(Q(\mathbb{G}_q \circ \mathbb{G}_p))$. Thus in this case there is a path of length 1 from i to j in $Q(\mathbb{G}_q \circ \mathbb{G}_p)$.

Suppose $k \neq i$. Since $(v_{is}, v_{kt}) \in \mathcal{A}(\mathbb{G}_p)$ and $\mathbb{G}_p \in \bar{\mathcal{D}}$, $(v_{is}, v_{k1}) \in \mathcal{A}(\mathbb{G}_p)$. But \mathbb{G}_q must have a self-arc at v_{k1} because $\mathbb{G}_q \in \bar{\mathcal{D}}$. Therefore $(v_{is}, v_{k1}) \in \mathcal{A}(\mathbb{G}_q \circ \mathbb{G}_p)$. Moreover, there must be a path in $\mathbb{G}_q \circ \mathbb{G}_p$ from v_{k1} to v_{kx} because v_{kx} is in the hierarchy rooted at v_{k1} . But both (v_{is}, v_{k1}) and (v_{kx}, v_{jw}) are arcs in $\mathbb{G}_q \circ \mathbb{G}_p$, and so there must be a path in $\mathbb{G}_q \circ \mathbb{G}_p$ from v_{is} to v_{jw} . This implies that there must be a path in $Q(\mathbb{G}_q \circ \mathbb{G}_p)$ from i to j . \square

Proof of Proposition 5. To prove the proposition it is enough to show that if $Q(\mathbb{G}_r) \circ \dots \circ Q(\mathbb{G}_1)$ contains a path from some $i \in \mathcal{V}$ to some $j \in \mathcal{V}$, then $Q(\mathbb{G}_r \circ \dots \circ \mathbb{G}_1)$ also contains a path from i to j . As a first step towards this end, we claim that if $\mathbb{G}_p, \mathbb{G}_q$ are graphs in $\bar{\mathcal{D}}$ for which $Q(\mathbb{G}_q) \circ Q(\mathbb{G}_p)$ contains a path from u to v , for some $u, v \in \mathcal{V}$, then $Q(\mathbb{G}_q \circ \mathbb{G}_p)$ also contains a path from u to v . To prove that this is so, fix $u, v \in \mathcal{V}$ and $\mathbb{G}_p, \mathbb{G}_q \in \bar{\mathcal{D}}$ and suppose that $Q(\mathbb{G}_q) \circ Q(\mathbb{G}_p)$ contains a path from u to v . Then there must be a positive integer s and vertices k_1, k_2, \dots, k_s ending at $k_s = v$ for which $(u, k_1), (k_1, k_2), \dots, (k_{s-1}, k_s)$ are arcs in $Q(\mathbb{G}_q) \circ Q(\mathbb{G}_p)$. In view of Lemma 8, there must be paths in $Q(\mathbb{G}_q \circ \mathbb{G}_p)$ from u to k_1, k_1 to k_2, \dots, k_{s-1} to k_s . It follows that there must be a path in $Q(\mathbb{G}_q \circ \mathbb{G}_p)$ from u to v . Thus the claim is established.

It will now be shown by induction for each $s \in \{2, \dots, m\}$ that if $Q(\mathbb{G}_s) \circ \dots \circ Q(\mathbb{G}_1)$ contains a path from i to some $j_s \in \mathcal{V}$, then $Q(\mathbb{G}_r \circ \dots \circ \mathbb{G}_1)$ also contains a path from i to j_s . In view of the claim just proved above, the assertion is true if $s = 2$. Suppose the assertion is true for all $s \in \{2, 3, \dots, t\}$, where t is some integer in $\{2, \dots, r - 1\}$. Suppose that $Q(\mathbb{G}_{t+1}) \circ \dots \circ Q(\mathbb{G}_1)$ contains a path from i to j_{t+1} . Then there must be an integer k such that $Q(\mathbb{G}_t) \circ \dots \circ Q(\mathbb{G}_1)$ contains a path from i to k and $Q(\mathbb{G}_{t+1})$ contains a path from k to j_{t+1} . In view of the inductive hypothesis, $Q(\mathbb{G}_t \circ \dots \circ \mathbb{G}_1)$ contains a path from i to k . Therefore $Q(\mathbb{G}_{t+1}) \circ Q(\mathbb{G}_t \circ \dots \circ \mathbb{G}_1)$ has a path from i to j_{t+1} . Hence the claim established at the beginning of this proof applies, and it can be concluded that $Q(\mathbb{G}_{t+1} \circ \mathbb{G}_t \circ \dots \circ \mathbb{G}_1)$ has a path from i to j_{t+1} . Therefore by induction the aforementioned assertion is true. \square

4.4. Proof of convergence. Our aim is to make use of the properties of extended delay graphs just derived to prove Theorem 2. We will also need the following result from [6].

PROPOSITION 6. *Let \mathcal{S}_{sr} be any closed set of stochastic matrices which are all of the same size and whose graphs $\gamma(S)$, $S \in \mathcal{S}_{sr}$, are all strongly rooted. As $j \rightarrow \infty$, any product $S_j \dots S_1$ of matrices from \mathcal{S}_{sr} converges exponentially fast to a matrix of the form $\mathbf{1}c$ at a rate no slower than λ , where c is a nonnegative row vector depending on the sequence and λ is a nonnegative constant less than 1 depending only on \mathcal{S}_{sr} .*

Proof of Theorem 2. In view of (31), $\theta(t) = F(t-1) \dots F(0)\theta(0)$. Thus to prove the theorem it suffices to prove that as $t \rightarrow \infty$ the matrix product $F(t) \dots F(0)$ converges exponentially fast to a matrix of the form $\mathbf{1}c$.

By hypothesis, the sequence of neighbor graphs $\mathbb{N}(0), \mathbb{N}(1), \dots$ is repeatedly jointly rooted by subsequences of length q . This means that each of the sequences $\mathbb{N}(kq), \dots, \mathbb{N}((k+1)q-1)$, $k \geq 0$, is jointly rooted. Let $\mathbb{D}(t) = \gamma(F(t))$, $t \geq 0$. In view of (29), $\mathbb{N}(t) = Q(\mathbb{D}(t))$, $t \geq 0$. Thus each of the sequences $Q(\mathbb{D}(kq)), \dots, Q(\mathbb{D}((k+1)q-1))$, $k \geq 0$, is jointly rooted, and so each composition $Q(\mathbb{D}((k+1)q-1)) \circ \dots \circ Q(\mathbb{D}(kq))$ is a rooted graph. In view of Proposition 5, each graph $Q(\mathbb{D}((k+1)q-1)) \circ \dots \circ \mathbb{D}(kq)$, $k \geq 0$, is also rooted.

Set $p = (m(n-1)^2 + m - 1)q$, where m is the largest integer in the set $\{m_1, m_2, \dots, m_n\}$. In view of Proposition 4, each of the graphs $\mathbb{D}((k+1)p-1) \circ \dots \circ \mathbb{D}(kp)$, $k \geq 0$, is strongly rooted. Let $\mathcal{F}(p)$ denote the set of all products of p matrices from $\bar{\mathcal{F}}$ which have the additional property that each such product has a strongly rooted graph. Then $\mathcal{F}(p)$ is finite and therefore compact, because $\bar{\mathcal{F}}$ is.

For $k \geq 0$, define

$$(32) \quad S(k) = F((k+1)p-1) \dots F(kp).$$

In view of (12) and the fact that $\gamma(F(t)) = \mathbb{D}(t)$, $t \geq 0$, it must be true that $\gamma(S(k)) = \mathbb{D}((k+1)p-1) \circ \dots \circ \mathbb{D}(kp)$, $k \geq 0$. Thus each $S(k)$ has a strongly rooted graph. Moreover, each such $S(k)$ is the product of p matrices from $\bar{\mathcal{F}}$. Therefore $S(k) \in \mathcal{F}(p)$, $k \geq 0$. Therefore Proposition 6 applies with $\mathcal{S}_{sr} = \mathcal{F}(p)$, and so it can be concluded that the matrix product $S(k) \dots S(0)$ converges exponentially fast as $k \rightarrow \infty$ to a matrix of the form $\mathbf{1}c$ as $k \rightarrow \infty$.

In view of the definition of $S(k)$ it is clear that for any t , there is an integer $k(t)$ and a stochastic matrix $\hat{S}(t)$ composed of the product of at most $p-1$ matrices from $\bar{\mathcal{F}}$ such that

$$F(t) \dots F(1) = \hat{S}(t)S(k(t)) \dots S(0).$$

Moreover, $t \mapsto k(t)$ must be an unbounded, strictly increasing function; because of this the product $S(k(t)) \cdots S(0)$ must converge exponentially fast as $t \rightarrow \infty$ to a limit of the form $\mathbf{1}c$. Since $\widehat{S}(t)\mathbf{1}c = \mathbf{1}c$, $t \geq 0$, the product $F(t) \cdots F(1)$ must also converge exponentially fast as $t \rightarrow \infty$ to the same limit $\mathbf{1}c$. \square

5. Asynchronous flocking. In this section we consider a modified version of the consensus problem treated in [6] in which each agent independently updates its heading at times determined by its own clock.² We do not assume that the groups' clocks are synchronized or that the times any one agent updates its heading are evenly spaced. Updating of agent i 's heading is done as follows. At its k th *sensing event time* t_{ik} , agent i senses the headings $\theta_j(t_{ik})$, $j \in \mathcal{N}_i(t_{ik})$, of its current neighbors (which includes itself) and from this data computes its k th "way-point" $w_i(t_{ik})$. In what follows we will consider way-points based on averaging. In particular, agent i 's k th *way-point* is defined by the rule

$$(33) \quad w_i(t_{ik}) = \frac{1}{n_i(t_{ik})} \left(\sum_{j \in \mathcal{N}_i(t_{ik})} \theta_j(t_{ik}) \right), \quad i \in \{1, 2, \dots, n\},$$

where $n_i(t_{ik})$ is the number of neighbor elements in the neighbor index set $\mathcal{N}_i(t_{ik})$. After computing $w_i(t_{ik})$, agent i changes its heading from $\theta_i(t_{ik})$ to $w_i(t_{ik})$ on the interval $(t_{ik}, t_{i(k+1)}]$. In this paper we will consider the case when each agent updates its heading instantaneously at its own event times and holds its heading fixed between event times. More precisely, we will assume that agent i 's heading $\theta_i(t)$ takes on its agent i 's k th way-point value $w_i(t_{ik})$ immediately after its k th event time t_{ik} and that $\theta_i(t)$ is constant on each continuous-time interval $(t_{i(k-1)}, t_{ik}]$, $k \geq 1$, where $t_{i0} = 0$ is agent i 's zeroth event time. In other words for $k \geq 0$, agent i 's heading satisfies

$$(34) \quad \theta_i(t_{i(k+1)}) = \frac{1}{n_i(t_{ik})} \left(\sum_{j \in \mathcal{N}_i(t_{ik})} \theta_j(t_{ik}) \right),$$

$$(35) \quad \theta_i(t) = \theta_i(t_{ik}), \quad t_{i(k-1)} < t \leq t_{ik}.$$

5.1. Analytic synchronization. To develop conditions under which all agents eventually move with the same heading requires the analysis of the asymptotic behavior of the *asynchronous* process which the $2n$ heading equations of the form (34), (35) define. Despite the apparent complexity of this process, it is possible to capture its salient features using a suitably defined *synchronous* discrete-time, hybrid dynamical system \mathbb{S} . The sequence of steps involved in defining \mathbb{S} has been discussed before and is called *analytic synchronization* [12, 13]. Analytic synchronization is applicable to any finite family of continuous or discrete-time dynamical processes $\{\mathbb{P}_1, \mathbb{P}_2, \dots, \mathbb{P}_n\}$ under the following conditions. First, each process \mathbb{P}_i must be a dynamical system whose inputs consist of functions of the states of the other processes as well as signals which are exogenous to the entire family. Second, each process \mathbb{P}_i must have associated with it an ordered sequence of event times $\{t_{i1}, t_{i2}, \dots\}$ defined in such a way so that the state of \mathbb{P}_i at event time $t_{i(k_i+1)}$ is uniquely determined by values of the exogenous signals and states of the \mathbb{P}_j , $j \in \{1, 2, \dots, n\}$, at event times t_{jk_j} which occur

²A preliminary version of the material in this section was presented at the 2005 IFAC congress [4].

prior to $t_{i(k_i+1)}$ but in the finite past. Event time sequences for different processes need not be synchronized. Analytic synchronization is a procedure for creating a single synchronous process for purposes of analysis which captures the salient features of the original n asynchronously functioning processes. As a first step, all n event time sequences are merged into a single ordered sequence of event times \mathcal{T} . (This clever idea has been used before in [2] to study the convergence of totally asynchronous iterative algorithms.) The “synchronized” state of \mathbb{P}_i is then defined to be the original state of \mathbb{P}_i at \mathbb{P}_i ’s event times $\{t_{i1}, t_{i2}, \dots\}$ plus possibly some additional variables; at values of $t \in \mathcal{T}$ between event times t_{ik_i} and $t_{i(k_i+1)}$, the synchronized state of \mathbb{P}_i is taken to be the same as the value of its original state at time t_{ik} . Although it is not always possible to carry out all of these steps, when it is what ultimately results is a synchronous dynamical system \mathbb{S} evolving on the index set of \mathcal{T} , with the state composed of the synchronized states of the n individual processes under consideration. We now use these ideas to develop such a synchronous system \mathbb{S} for the asynchronous process under consideration.

5.2. Definition of \mathbb{S} . As a first step, let \mathcal{T} denote the set of all event times of all n agents. Relabel the elements of \mathcal{T} as t_0, t_1, t_2, \dots in such a way so that $t_j < t_{j+1}$, $j \in \{1, 2, \dots\}$. Next define

$$(36) \quad \bar{\theta}_i(\tau) = \theta_i(t_\tau), \quad \tau \geq 0, \quad i \in \{1, 2, \dots, n\}.$$

In view of (34), it must be true that if t_τ is an event time of agent i , then

$$\bar{\theta}_i(\tau') = \frac{1}{\bar{n}_i(\tau)} \left(\sum_{j \in \bar{\mathcal{N}}_i(\tau)} \bar{\theta}_j(\tau) \right),$$

where $\bar{\mathcal{N}}_i(\tau) = \mathcal{N}_i(t_\tau)$, $\bar{n}_i(\tau) = n_i(t_\tau)$, and $t_{\tau'}$ is the next event time of agent i after t_τ . But $\bar{\theta}_i(\tau') = \bar{\theta}_i(\tau + 1)$ because $\theta_i(t)$ is constant for $t_\tau < t \leq t_{\tau'}$ (approximately (35)). Therefore

$$(37) \quad \bar{\theta}_i(\tau + 1) = \frac{1}{\bar{n}_i(\tau)} \left(\sum_{j \in \bar{\mathcal{N}}_i(\tau)} \bar{\theta}_j(\tau) \right)$$

if t_τ is an event time of agent i . Meanwhile if t_τ is not an event time of agent i , then

$$(38) \quad \bar{\theta}_i(\tau + 1) = \bar{\theta}_i(\tau),$$

again because $\theta_i(t)$ is constant between event times. Note that if we define $\bar{\mathcal{N}}_i(\tau) = \{i\}$ and $\bar{n}_i(\tau) = 1$ for every value of τ for which t_τ is not an event time of agent i , then (38) can be written as

$$(39) \quad \bar{\theta}_i(\tau + 1) = \frac{1}{\bar{n}_i(\tau)} \left(\sum_{j \in \bar{\mathcal{N}}_i(\tau)} \bar{\theta}_j(\tau) \right).$$

Doing this enables us to combine (37) and (39) into a single formula valid for all $\tau \geq 0$. In other words, agent i ’s heading satisfies

$$(40) \quad \bar{\theta}_i(\tau + 1) = \frac{1}{\bar{n}_i(\tau)} \left(\sum_{j \in \bar{\mathcal{N}}_i(\tau)} \bar{\theta}_j(\tau) \right), \quad \tau \geq 0,$$

where

$$(41) \quad \bar{\mathcal{N}}_i(\tau) = \left\{ \begin{array}{ll} \mathcal{N}_i(t_\tau) & \text{if } t_\tau \text{ is an event time of agent } i \\ \{i\} & \text{if } t_\tau \text{ is not an event time of agent } i \end{array} \right\}$$

and $\bar{n}_i(\tau) = 1$ if t_τ is not an event time of agent i . Thus for all τ , $\bar{n}_i(\tau)$ is the number of indices in $\bar{\mathcal{N}}_i(\tau)$. For purposes of analysis, it is useful to interpret (41) as meaning that between agent i 's event times, its only neighbor is itself. There are n equations of the form in (40), and together they define a synchronous system \mathbb{S} which models the evolutions of the n agents' headings at event times.

5.3. State space model. As before, we can represent the neighbor relationships associated with (41) using a directed graph \mathbb{N} with vertex set $\mathcal{V} = \{1, 2, \dots, n\}$ and arc set $\mathcal{A}(\mathbb{N}) \subset \mathcal{V} \times \mathcal{V}$ which is defined in such a way so that (i, j) is an arc from i to j just in case agent i is a neighbor of agent j . Thus as before, \mathbb{N} 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 continue to write \mathcal{G}_{sa} for the set of all such graphs.

For each graph $\mathbb{N} \in \mathcal{G}_{sa}$ let $F = D^{-1}A'$, where A' is the transpose of the adjacency matrix of \mathbb{N} and D the diagonal matrix whose j th diagonal element is the in-degree of vertex j within the graph. The set of agent heading update rules defined by (41) can be written in state form as

$$(42) \quad \bar{\theta}(\tau + 1) = F(\tau)\bar{\theta}(\tau), \quad \tau \in \{0, 1, 2, \dots\},$$

where $\bar{\theta}$ is the heading vector $\bar{\theta} = [\bar{\theta}_1 \quad \bar{\theta}_2 \quad \dots \quad \bar{\theta}_n]'$, and $F(\tau)$ is the flocking matrix determined by neighbor graph $\mathbb{N}(\tau)$ at event time t_τ .

Up to this point the development is essentially the same as in the leaderless consensus problem discussed in section 2. But when one considers the type of graphs in \mathcal{G}_{sa} which are likely to be encountered along a given trajectory, things are quite different. Note, for example, that the only vertices of $\mathbb{N}(\tau)$ which can have more than one incoming arc are those of agents for whom t_τ is an event time. Thus in the most likely situation when distinct agents have only distinct event times, there will be at most one vertex in each graph $\mathbb{N}(\tau)$ which has more than one incoming arc. It is this situation we want to explore further. Towards this end, let $\mathcal{G}_{sa}^* \subset \mathcal{G}_{sa}$ denote the subclass of all graphs which have at most one vertex with more than one incoming arc. Note that for $n > 2$, there is no rooted graph in \mathcal{G}_{sa}^* . Nonetheless, in light of Theorem 1 it is clear that convergence to a common steady state heading will occur if the infinite sequence of graphs $\mathbb{N}(0), \mathbb{N}(1), \dots$ is repeatedly jointly rooted. This of course would require that there exist a jointly rooted sequence of graphs from \mathcal{G}_{sa}^* . We will now explain why such sequences do in fact exist.

Let us agree to call a graph $\mathbb{G} \in \mathcal{G}_{sa}$ an *all neighbor graph centered at v* if every vertex of \mathbb{G} is a neighbor of v . Note that every all neighbor graph in \mathcal{G}_{sa} is also in \mathcal{G}_{sa}^* . Note also that all neighbor graphs are maximal in \mathcal{G}_{sa}^* with respect to the partial ordering of \mathcal{G}_{sa}^* by inclusion. Note also the composition of any all neighbor graph with itself is itself. On the other hand, because the union of two graphs in \mathcal{G}_{sa} is always contained in the composition of the two graphs, the composition of n all neighbor graphs with distinct centers must be a graph in which each vertex is a neighbor of every other, i.e., the complete graph. Thus the composition of n all neighbor graphs with distinct centers is strongly rooted. In summary, the hypothesis of Theorem 1 is not vacuous for the asynchronous problem under consideration. When that hypothesis is satisfied, convergence to a common steady state heading will occur.

6. Leader following. In this section we consider a modified version of the flocking problem for the same group of n agents as before but now with one of the group's members (say agent 1) acting as the group's *leader* [11, 8]. The remaining agents, henceforth called *followers* and labelled 2 through n , do not know who the leader is or even if there is a leader. Accordingly they continue to use the same heading update rule (1) as before. The leader, on the other hand, acting on its own, ignores update rule (1) and moves with a constant heading $\theta_1(0)$. Thus

$$(43) \quad \theta_1(t+1) = \theta_1(t).$$

The situation just described can be modelled as a state space system

$$(44) \quad \theta(t+1) = F(t)\theta(t), \quad t \geq 0,$$

just as before, except now agent 1 is constrained to have no neighbors other than itself. The neighbor graphs \mathbb{N} which model neighbor relations accordingly all have a distinguished *leader vertex* which has no incoming arcs other than its own.

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. Convergence in the leaderless case under the most general conditions required the sequence of neighbor graphs $\mathbb{N}(0), \mathbb{N}(1), \dots$ encountered along a trajectory to be repeatedly jointly rooted. For the leader-follower case now under consideration, what is required is exactly the same. However, since the leader vertex has only one incoming arc which is a self-arc, the only way $\mathbb{N}(0), \mathbb{N}(1), \dots$ can be repeatedly jointly rooted is that the sequence be "rooted at the leader vertex $v = 1$." More precisely, an infinite sequence of graphs $\mathbb{G}_1, \mathbb{G}_2$ in \mathcal{G}_{sa} is *repeatedly jointly rooted at v* if there is a positive integer m for which each finite sequence $\mathbb{G}_{m(k-1)+1}, \dots, \mathbb{G}_{mk}, k \geq 1$, is "jointly rooted at v "; a finite sequence of directed graphs $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_k$ is *jointly rooted at v* if the composition $\mathbb{G}_k \circ \mathbb{G}_{k-1} \circ \dots \circ \mathbb{G}_1$ is rooted at v . Our main result on discrete-time leader following is next.

THEOREM 3. *Let $\theta(0)$ be fixed. For any trajectory of the system determined by (1) along which the sequence of neighbor graphs $\mathbb{N}(0), \mathbb{N}(1), \dots$ is repeatedly jointly rooted at vertex 1, there is a constant θ_{ss} , depending only on $\theta(0)$, for which*

$$\lim_{t \rightarrow \infty} \theta(t) = \theta_1(0)\mathbf{1},$$

where the limit is approached exponentially fast.

Proof of Theorem 3. Since any sequence which is repeatedly jointly rooted at v is repeatedly jointly rooted, Theorem 1 is applicable. Therefore the headings of all n agents converge exponentially fast to a single common steady state heading θ_{ss} . But since the heading of the leader is fixed, θ_{ss} must be the leader's heading $\theta_1(0)$. \square

7. Concluding remarks. The main goal of this paper has been to study various versions the flocking problem considered in [14, 16, 3, 11, 1] and elsewhere from a single point of view which emphasizes the underlying graphical structures for which consensus can be reached. The paper brings together in one place a number of results scattered throughout the literature and at the same time presents new results concerned with convergence rates, asynchronous operation, sensing delays, and graphical interpretations of several specially structured stochastic matrices appropriate to nonhomogeneous Markov chains.

The approach taken in this paper to analyze consensus in the face of measurement delays first goes through a lifting process and then focuses on the resulting state space

model. As we have explained, the lifting process determines stochastic matrices whose graphs do not have self-arcs at all vertices. Nonetheless the graphs which result, namely delay graphs, have special structure, which we have exploited. One is able to associate with each such graph two special graphs, namely a quotient graph and an agent subgraph. These graphs play roles in the analysis of the consensus problem with delays which are similar to the roles played by corresponding quotient graphs and the “injected subgraph” used in the analysis of the asynchronous flocking problem treated in [5]. Although the corresponding graphs which arise in the two problems are completely different, there does seem to be a general pattern of use appropriate to both problems. This suggests that quotient graphs and graphs similar to injected graphs or agent subgraphs may be useful in analyzing other problems as well.

REFERENCES

- [1] D. ANGELI AND P. A. BLIMAN, *Extension of a result by Moreau on stability of leaderless multi-agent systems*, in Proceedings of the 2005 IEEE CDC, 2005, pp. 759–764.
- [2] D. P. BERTSEKAS AND J. N. TSITSIKLIS, *Parallel and Distributed Computation*, Prentice–Hall, Englewood Cliffs, NJ, 1989.
- [3] V. D. BLONDEL, J. M. HENDRICHX, A. OLSHEVSKY, AND J. N. TSITSIKLIS, *Convergence in multiagent coordination, consensus, and flocking*, in Proceedings of the 2005 IEEE CDC, 2005, pp. 2996–3000.
- [4] M. CAO, A. S. MORSE, AND B. D. O. ANDERSON, *Coordination of an asynchronous multi-agent system via averaging*, in Proceedings of the 2005 IFAC Congress, 2005.
- [5] M. CAO, A. S. MORSE, AND B. D. O. ANDERSON, *Agreeing asynchronously*, IEEE Trans. Automat. Control, to appear.
- [6] M. CAO, A. S. MORSE, AND B. D. O. ANDERSON, *Reaching a consensus in a dynamically changing environment: A graphical approach*, SIAM J. Control Optim., 47 (2008), pp. 575–600.
- [7] M. CAO, D. A. SPIELMAN, AND A. S. MORSE, *A lower bound on convergence of a distributed network consensus algorithm*, in Proceedings of the 2005 IEEE CDC, 2005, pp. 2356–2361.
- [8] L. GAO AND D. CHENG, *Comment on: “Coordination of groups of mobile autonomous agents using nearest neighbor rules,”* IEEE Trans. Automat. Control, 50 (2005), pp. 1913–1916.
- [9] D. J. HARTFIEL, *Nonhomogeneous Matrix Products*, World Scientific, Singapore, 2002.
- [10] R. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, New York, 1985.
- [11] A. JADBABAIE, J. LIN, AND A. S. MORSE, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Trans. Automat. Control, 48 (2003), pp. 988–1001.
- [12] J. LIN, A. S. MORSE, AND B. D. O. ANDERSON, *The multi-agent rendezvous problem—the asynchronous case*, in Proceedings of the 2004 IEEE CDC, 2004, pp. 1926–1931.
- [13] J. LIN, A. S. MORSE, AND B. D. O. ANDERSON, *The multi-agent rendezvous problem. Part 2: The asynchronous case*, SIAM J. Control Optim., 46 (2007), pp. 2120–2147.
- [14] L. MOREAU, *Stability of multi-agent systems with time-dependent communication links*, IEEE Trans. Automat. Control, 50 (2005), pp. 169–182.
- [15] A. S. MORSE, *Logically switched dynamical systems*, in Nonlinear and Optimal Control Theory, Springer-Verlag, Berlin, 2008, pp. 1–84.
- [16] W. REN AND R. BEARD, *Consensus seeking in multiagent systems under dynamically changing interaction topologies*, IEEE Trans. Automat. Control, 50 (2005), pp. 655–661.
- [17] E. SENETA, *Non-negative Matrices and Markov Chains*, Springer-Verlag, New York, 1981.
- [18] J. N. TSITSIKLIS, *Problems in Decentralized Decision Making and Computation*, Ph.D. thesis, MIT, Cambridge, MA, 1984.
- [19] J. N. TSITSIKLIS, D. P. BERTSEKAS, AND M. ATHANS, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Trans. Automat. Control, 31 (1986), pp. 803–812.
- [20] T. VICSEK, A. CZIRÓK, E. BEN-JACOB, I. COHEN, AND O. SHOCHET, *Novel type of phase transition in a system of self-driven particles*, Phys. Rev. Lett., 75 (1995), pp. 1226–1229.