

CONVERGENCE RESULTS FOR THE LINEAR CONSENSUS PROBLEM UNDER MARKOVIAN RANDOM GRAPHS*

ION MATEI[†], JOHN S. BARAS[†], AND CHRISTOFOROS SOMARAKIS[†]

Abstract. This paper discusses the linear asymptotic consensus problem for a network of dynamic agents whose communication network is modeled by a randomly switching graph. The switching is determined by a finite state Markov process, each topology corresponding to a state of the process. We address the cases where the dynamics of the agents is expressed both in continuous time and in discrete time. We show that, if the consensus matrices are doubly stochastic, average consensus is achieved in the mean square sense and the almost sure sense if and only if the graph resulting from the union of graphs corresponding to the states of the Markov process is strongly connected. The aim of this paper is to show how techniques from the theory of Markovian jump linear systems, in conjunction with results inspired by matrix and graph theory, can be used to prove convergence results for stochastic consensus problems.

Key words. consensus, Markovian random graphs, stochastic systems

AMS subject classifications. 93A14, 93C05, 93E03

DOI. 10.1137/100816870

1. Introduction. A consensus problem consists of a group of dynamic agents that seek to agree upon certain quantities of interest by exchanging information according to a set of rules. This problem can model many phenomena involving information exchange between agents such as cooperative control of vehicles, formation control, flocking, synchronization, and parallel computing. Distributed computation over networks has a long history in control theory starting with the work of Borkar and Varaiya [1] and Tsitsiklis, Bertsekas, and Athans (see [22, 23]) on asynchronous agreement problems and parallel computing. A theoretical framework for solving consensus problems was introduced by Olfati-Saber and Murray [16, 17], while Jadbabaie, Lin, and Morse studied alignment problems [7] for reaching an agreement. Relevant extensions of the consensus problem were done by Ren and Beard [19], by Moreau [11], and, more recently, by Nedic and Ozdaglar (see [12, 14]).

Typically agents are connected via a network that changes with time due to link failures, packet drops, node failure, etc. Such variations in topology can happen randomly, which motivates the investigation of consensus problems under a stochastic framework. Hatano and Mesbahi consider in [6] an agreement problem over random information networks, where the existence of an information channel between a pair of elements at each time instance is probabilistic and independent of other channels. In [18], Porfiri and Stilwell provide sufficient conditions for reaching consensus almost surely in the case of a discrete linear system, where the communication flow is given by a directed graph derived from a random graph process, independent of other time instances. Under a similar communication topology model, Tahbaz-Salehi and Jadbabaie give necessary and sufficient conditions for almost sure convergence to

*Received by the editors December 3, 2010; accepted for publication (in revised form) November 28, 2012; published electronically April 10, 2013. This material is based in part upon work supported in part by U.S. Army Research Office MURI award grant W911NF-08-1-0238, by National Science Foundation (NSF) award grant CNS-1035655, and by U.S. AFOSR MURI award grant FA9550-09-1-0538.

<http://www.siam.org/journals/sicon/51-2/81687.html>

[†]Institute for Systems Research, University of Maryland, College Park, MD 20742 (imatei@umd.edu, baras@umd.edu, csomarak@umd.edu).

consensus in [20], while in [21], they extend the applicability of their necessary and sufficient conditions to strictly stationary ergodic random graphs. Another recent result on the consensus problem under random topologies can be found in [8].

This paper deals with the linear consensus problem for a group of dynamic agents. We assume that the communication flow between agents is modeled by a (possibly directed) randomly switching graph. The switching is determined by a homogeneous, finite-state Markov chain, each communication pattern corresponding to a state of the Markov process. We address the cases where the dynamics of the agents is expressed both in continuous and in discrete time, and, under certain assumptions on the consensus matrices, we give necessary and sufficient conditions to guarantee convergence to average consensus in the mean square sense and in the almost sure sense. The Markovian switching model goes beyond the common independent and identically distributed (i.i.d.) assumption on the random communication topology and appears in cases where Rayleigh fading channels are considered.

The aim of this paper is to show how mathematical techniques used in the stability analysis of Markovian jump linear systems, together with results inspired by matrix and graph theory, can be used to prove (intuitively clear) convergence results for the (linear) stochastic consensus problem. Another avenue of approach in proving (part of the) convergence analysis of the consensus algorithm could be to use ergodic theory of Markov chains [2].

1.1. Notation and definitions. We denote by $\mathbf{1}$ the column vector of all ones. If the dimension of the vector needs to be emphasized, an index will be added for clarity (for example, if $\mathbf{1}$ is an n dimensional vector, we will explicitly mark this by using $\mathbf{1}_n$). Let x be a vector in \mathbb{R}^n . By $av(x)$ we denote the quantity $av(x) = x' \mathbf{1} / \mathbf{1}' \mathbf{1}$. The symbols \otimes and \oplus represent the Kronecker product and sum, respectively. Given a matrix A , $\text{Null}(A)$ designates the nullspace of the considered matrix. If \mathcal{X} is some finite dimensional space, $\dim(\mathcal{X})$ gives us the dimension of \mathcal{X} . We denote by $col(A)$ a vector containing the columns of matrix A .

Let \mathcal{M} be a finite set of matrices, and let A be some matrix. By \mathcal{M}' we denote the set of the transpose matrices of \mathcal{M} , i.e., $\mathcal{M}' = \{M' \mid M \in \mathcal{M}\}$. By $\mathcal{M} \otimes A$ we understand the following set of matrices: $\mathcal{M} \otimes A = \{M \otimes A \mid M \in \mathcal{M}\}$. By writing that $A\mathcal{M} = \mathcal{M}$ we understand that the set \mathcal{M} is invariant to left multiplication by matrix A , i.e., given $M_i \in \mathcal{M}$, $AM_i = \tilde{M}_i$ and $\{M_i\}_{i=1}^{|\mathcal{M}|} = \{\tilde{M}_i\}_{i=1}^{|\mathcal{M}|} = \mathcal{M}$, where $|\mathcal{M}|$ is the cardinality of set \mathcal{M} .

Let P be a probability transition matrix corresponding to a discrete-time, homogeneous, finite state Markov chain. We denote by \mathbf{P}_∞ the limit set of the sequence $\{P^k\}_{k \geq 0}$, i.e., all matrices L for which there exists a sequence $\{t_k\}_{k \geq 0}$ in \mathbb{N} such that $\lim_{k \rightarrow \infty} P^{t_k} = L$. Note that if the matrix P corresponds to an ergodic Markov chain, the cardinality of \mathbf{P}_∞ is one, with the limit point $\mathbf{1}\pi'$, where π is the stationary distribution. If the Markov chain is periodic with period m , the cardinality of \mathbf{P}_∞ is m . Let $d(M, \mathbf{P}_\infty)$ denote the distance from M to the set \mathbf{P}_∞ , that is, the smallest distance from M to any matrix in \mathbf{P}_∞ :

$$d(M, \mathbf{P}_\infty) = \inf_{L \in \mathbf{P}_\infty} \|L - M\|,$$

where $\|\cdot\|$ is a matrix norm.

DEFINITION 1.1. Let A be a matrix in $\mathbb{R}^{n \times n}$, and let $G = (V, E)$ be a graph of order n , with no self-loops. We say that matrix A corresponds to graph G or that graph G corresponds to matrix A if an edge e_{ij} belongs to E if and only if the (i, j) entry of A is nonzero. The graph corresponding to A will be denoted by G_A .

DEFINITION 1.2. Let s be a positive integer, and let $\mathcal{A} = \{A_i\}_{i=1}^s$ be a set of matrices with a corresponding set of graphs $\mathcal{G} = \{G_{A_i}\}_{i=1}^s$. We say that the graph $G_{\mathcal{A}}$ corresponds to the set \mathcal{A} if it is given by the union of graphs in \mathcal{G} , i.e.,

$$G_{\mathcal{A}} \triangleq \bigcup_{i=1}^s G_{A_i}.$$

In this paper we will mainly use two type of matrices: *probability transition matrices* (rows sum up to one) and *generator matrices* (rows sum up to zero). A generator matrix whose rows and columns both sum up to zero will be called a *doubly stochastic generator matrix*.

To simplify the exposition we will sometimes characterize a probability transition/generator matrix as being irreducible or strongly connected, and by this we understand that the corresponding Markov chain (directed graph) is irreducible (strongly connected).

DEFINITION 1.3. Let $A \in \mathbb{R}^{n \times n}$ be a probability transition/generator matrix. We say that A is block diagonalizable if there exists a similarity transformation P , encapsulating a number of row permutations, such that PAP' is a block diagonal matrix with irreducible blocks on the main diagonal.

For simplicity, the time index for both the continuous- and discrete-time cases is denoted by t .

Paper organization. In section 2 we present the setup and formulation of the problem, and we state our main convergence theorem. In section 3 we derive a number of results which constitute the core of the proof of our main result, which is given in section 4. Section 5 contains a discussion of our convergence result.

2. Problem formulation and statement of the convergence result. We assume that a group of n agents, labeled 1 through n , is organized in a communication network whose topology is given by a time varying graph $\mathbf{G}(t) = (V, E(t))$ (with no self-loops), where V is the set of n vertices and $E(t)$ is the time varying set of edges. The graph $\mathbf{G}(t)$ has an underlying random process governing its evolution, given by a homogeneous, continuous- or discrete-time Markov chain $\theta(t)$, taking values in the finite set $\{1, \dots, s\}$, for some positive integer s . In the case when $\theta(t)$ is a discrete-time Markov chain, its probability transition matrix is $P = (p_{ij})$ (rows sum up to one), while if $\theta(t)$ is a continuous-time Markov chain, its generator matrix is denoted by $\Lambda = (\lambda_{ij})$ (rows sum up to zero). The random graph $\mathbf{G}(t)$ takes values in a finite set of graphs $\mathcal{G} = \{G_i\}_{i=1}^s$ with probability $Pr(\mathbf{G}(t) = G_i) = Pr(\theta(t) = i)$ for $i = 1, \dots, s$. We denote by $q = (q_i)$ the initial distribution of $\theta(t)$, i.e., $P(\theta(0) = i) = q_i, i = 1, \dots, s$.

Letting $x(t)$ denote the state of the n agents, in the case when $\theta(t)$ is a discrete-time Markov chain, we model the dynamics of the agents by the following linear stochastic difference equation:

$$(1) \quad x(t+1) = \mathbf{D}_{\theta(t)}x(t), \quad x(0) = x_0,$$

where $\mathbf{D}_{\theta(t)}$ is a random matrix taking values in the finite set $\mathcal{D} = \{D_i\}_{i=1}^s$, with probability distribution $Pr(\mathbf{D}_{\theta(t)} = D_i) = Pr(\theta(t) = i)$. The matrices D_i are stochastic matrices (rows sum up to one) with positive diagonal entries and correspond to the graphs G_i for $i = 1, \dots, s$.

In the case when $\theta(t)$ is a continuous-time Markov chain, we model the dynamics of the agents by the following linear stochastic equation:

$$(2) \quad dx(t) = \mathbf{C}_{\theta(t)}x(t)dt, \quad x(0) = x_0,$$

where $\mathbf{C}_{\theta(t)}$ is a random matrix taking values in the finite set $\mathcal{C} = \{C_i\}_{i=1}^s$, with probability distribution $Pr(\mathbf{C}_{\theta(t)} = C_i) = Pr(\theta(t) = i)$. The matrices C_i are generator-like matrices (rows sum up to zero, off-diagonal entries are nonnegative) and correspond to the graphs G_i for $i = 1, \dots, s$. The initial state $x(0) = x_0$, for both continuous and discrete models, is assumed to be deterministic. We will sometimes refer to the matrices belonging to the sets \mathcal{D} and \mathcal{C} as *consensus matrices*. The underlying probability space (for both models) is denoted by $(\Omega, \mathcal{F}, \mathcal{P})$, and the solution process $x(t, x_0, \omega)$ (or simply, $x(t)$) of (1) or (2) is a random process defined on $(\Omega, \mathcal{F}, \mathcal{P})$. We note that the stochastic dynamics (1) and (2) represent *Markovian jump linear systems* for discrete and continuous time, respectively. For a comprehensive study of the theory of (discrete-time) Markovian jump linear systems, the reader can, for example, refer to [3].

ASSUMPTION 2.1. *Throughout this paper we assume that the matrices belonging to the set \mathcal{D} (\mathcal{C}) are doubly stochastic, that is, both rows and columns sum up to one (zero). Moreover, the elements of \mathcal{D} have positive diagonal entries. We also assume that the Markov chain $\theta(t)$ is irreducible.*

Remark 2.1. Consensus matrices that satisfy Assumption 2.1 can be constructed, for instance, by using a Laplacian-based scheme in the case where the communication graph is undirected or balanced (for every node, the inner degree is equal to the outer degree) and possibly weighted. If L_i denotes the Laplacian of the graph G_i , we can choose $D_i = I - \epsilon L_i$ and $C_i = -L_i$, where $\epsilon > 0$ is chosen such that D_i is stochastic, with positive diagonal elements.

DEFINITION 2.1. *We say that $x(t)$ converges to average consensus*

(a) *in the mean square sense if for any $x_0 \in \mathbb{R}^n$ and initial distribution $q = (q_1, \dots, q_s)$ of $\theta(t)$,*

$$\lim_{t \rightarrow \infty} E[\|x(t) - av(x_0)\mathbf{1}\|^2] = 0;$$

(b) *in the almost sure sense if for any $x_0 \in \mathbb{R}^n$ and initial distribution $q = (q_1, \dots, q_s)$ of $\theta(t)$,*

$$Pr\left(\lim_{t \rightarrow \infty} \|x(t) - av(x_0)\mathbf{1}\|^2 = 0\right) = 1.$$

Remark 2.2. Under the homogeneous assumption on the Markov chain $\theta(t)$, in both the discrete- and the continuous-time cases, it suffices to prove convergence in the mean square sense since this implies stability in the almost sure sense as well (see, for instance, Corollary 3.46 of [3] for the discrete-time case or Theorem 2.1 of [4] for the continuous-time case).

Assumption 2.1 guarantees reaching average consensus, which is desirable in important distributed computing applications such as distributed estimation [15] or distributed optimization [13]. Any other scheme can be used as long as it produces matrices with the properties stated above and reflects the communication structures among agents.

PROBLEM 2.1. *Given the random processes $\mathbf{D}(t)$ and $\mathbf{C}(t)$, together with Assumption 2.1, we derive necessary and sufficient conditions such that the state vector $x(t)$, evolving according to (1) or (2), converges to average consensus in the sense of Definition 2.1.*

In the following we state the convergence result for the linear consensus problem under Markovian random communication topology.

THEOREM 2.2. *The state vector $x(t)$, evolving according to the dynamics (1) (or (2)), converges to average consensus in the sense of Definition 2.1(a) and (b) if and only if $G_{\mathcal{D}}$ (or $G_{\mathcal{C}}$) is strongly connected.*

Theorem 2.2 formulates an intuitively obvious condition for reaching consensus under linear scheme (1) or (2) and under the Markovian assumption on the communication patterns. Namely, it expresses the need for persistent communication paths among all agents. We defer to section 4 the proof of this theorem and provide here an intuitive and nonrigorous interpretation. Since $\theta(t)$ is irreducible, with probability one all states are visited infinitely many times. But since the graph $G_{\mathcal{D}}$ ($G_{\mathcal{C}}$) is strongly connected, communication paths between all agents are formed infinitely many times, which allows consensus to be achieved. Conversely, if the graph $G_{\mathcal{D}}$ ($G_{\mathcal{C}}$) is not strongly connected, then there exist at least two agents such that for any sample path of $\theta(t)$, no communication path among them (direct or indirect) is ever formed. Consequently, consensus cannot be reached. Our main contribution is to prove Theorem 2.2 using an approach based on the stability theory of Markovian jump linear systems, in conjunction with results on matrix and graph theory.

3. Preliminary results. This section starts with a set of general preliminary results after which it continues with results characteristic to the cases where the dynamics of the agents is expressed in discrete and continuous time. The proof of Theorem 2.2 is mainly based on four lemmas (Lemmas 3.10 and 3.11 for the discrete-time case and Lemmas 3.12 and 3.13 for the continuous-time case) which state properties of some matrices that appear in the dynamic equations of the first and second moment of the state vector. The proofs of these lemmas are based on results introduced in the next subsection.

3.1. General preliminary results. This subsection contains the statement of a number of preliminary results that are needed in the proofs of the auxiliary results corresponding to the discrete- and continuous-time cases and in the proof of the main theorem.

The following theorem introduces a convergence result for an infinite product of ergodic matrices, whose proof can be found in [24].

THEOREM 3.1 (see [24]). *Let s be a positive integer, and let $\{A_i\}_{i=1}^s$ be a finite set of $n \times n$ ergodic matrices, with the property that any finite product $A_{i_1} A_{i_2} \dots A_{i_m}$ is ergodic. Then, for any map $r : \mathbb{N} \rightarrow \{1, \dots, s\}$, there exists a vector c with nonnegative entries (summing up to one) such that*

$$(3) \quad \lim_{j \rightarrow \infty} \prod_{i=1}^j A_{r(i)} = \mathbf{1}c'.$$

COROLLARY 3.2. *Under the same assumptions as in Theorem 3.1, if in addition the matrices in the set $\{A_i\}_{i=1}^s$ are doubly stochastic, then*

$$(4) \quad \lim_{j \rightarrow \infty} \prod_{i=1}^j A_{r(i)} = \frac{1}{n} \mathbf{1}\mathbf{1}'.$$

Proof. By Theorem 3.1 we have that

$$\lim_{j \rightarrow \infty} \prod_{i=1}^j A_{r(i)} = \mathbf{1}c'.$$

Since the matrices considered are doubly stochastic and ergodic, their transposes are ergodic as well. Hence, by again applying Theorem 3.1 on the transpose versions of $\{A_i\}_{i=1}^s$, we obtain that there exists a vector d such that

$$\lim_{j \rightarrow \infty} \left(\prod_{i=1}^j A_{r(i)} \right)' = \mathbf{1}d'.$$

But since the stochastic matrix $\mathbf{1}c'$ must be equal to $d\mathbf{1}'$, the result follows. \square

Remark 3.1. The homogeneous finite-state Markov chain corresponding to a doubly stochastic transition matrix P cannot have transient states. Indeed, since P is doubly stochastic, the same is true for P^t for all $t \geq 1$. Assuming that there exists a transient state i , then $\lim_{t \rightarrow \infty} (P^t)_{ji} = 0$ for all j ; i.e., all entries on column i converge to zero. But this means that there exists some t^* for which $\sum_j (P^{t^*})_{ji} < 1$, which contradicts the fact that P^{t^*} must be doubly stochastic. An important implication is that we can relabel the vertices of the Markov chain such that P is block diagonalizable.

Remark 3.2. Since the Markov chain corresponding to a doubly stochastic transition/generator matrix cannot have transient states, the Markov chain (seen as a graph) has a spanning tree if and only if it is irreducible (strongly connected).

The next lemma gives an upper bound on a finite product of nonnegative matrices in terms of the sum of matrices that appear in the product.

LEMMA 3.3 (see [7]). *Let $m \geq 2$ be a positive integer, and let $\{A_i\}_{i=1}^m$ be a set of nonnegative $n \times n$ matrices with positive diagonal elements; then*

$$\prod_{i=1}^m A_i \geq \gamma \sum_{i=1}^m A_i,$$

where $\gamma > 0$ depends on the matrices A_i , $i = 1, \dots, m$.

In the following proposition we study the convergence properties of a particular sequence of matrices.

PROPOSITION 3.4. *Consider a matrix $Q \in \mathbb{R}^{n \times n}$ such that $\|Q\|_1 \leq 1$ and a set of matrices $\mathcal{S} = \{S_1, \dots, S_m\}$ for some positive integer $m \leq n$. Assume that there exists a subsequence $\{t_k\} \subset \mathbb{N}$ such that \mathcal{S} is a limit set of the sequence $\{Q^{t_k}\}_{k \geq 0}$ and that, for any $S \in \mathcal{S}$, $QS \in \mathcal{S}$ as well. Then, \mathcal{S} is a limit set of the sequence $\{\bar{Q}^k\}_{k \geq 0}$, i.e.,*

$$(5) \quad \lim_{k \rightarrow \infty} d(Q^k, \mathcal{S}) = 0,$$

where $d(Q, \mathcal{S}) = \min_{S \in \mathcal{S}} \|Q - S\|$ and $\|\cdot\|$ is some arbitrary matrix norm.

Proof. We will prove (5) for the particular case of matrix norm one, and the general result will follow from the equivalence of norms. Pick a subsequence $\{t'_k\}_{k \geq 0}$ given by $t'_k = t_k + \delta_k$, where $\delta_k \in \mathbb{N}$. It follows that

$$d(Q^{t'_k}, \mathcal{S}) = \min_{S \in \mathcal{S}} \|Q^{\delta_k} Q^{t_k} - Q^{\delta_k} S\|_1 \leq \|Q^{\delta_k}\|_1 \min_{S \in \mathcal{S}} \|Q^{t_k} - S\|_1 \leq d(Q^{t_k}, \mathcal{S}).$$

Therefore, we get that \mathcal{S} is a limit set for the sequence $\{Q^{t'_k}\}_{k \geq 0}$, and the result follows since we can make $\{t'_k\}_{k \geq 0}$ arbitrary. \square

The next lemma states a property of the nullspace of the sum of two block diagonalizable generator matrices.

LEMMA 3.5. *Let $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ be two block diagonalizable generator matrices. Then*

$$\text{Null}(A + B) = \text{Null}(A) \cap \text{Null}(B).$$

Proof. Obviously, $\text{Null}(A) \cap \text{Null}(B) \subset \text{Null}(A + B)$. In the following we show the opposite inclusion. Since A is block diagonalizable, there exists a permutation matrix T such that $\bar{A} = TAT'$ is a block diagonal generator matrix (with irreducible blocks). Let $\bar{A}_i \in \mathbb{R}^{n_i \times n_i}$, $i = 1, \dots, m$, denote the irreducible blocks on the main diagonal of \bar{A} , where m is the number of such blocks and $\sum_{i=1}^m n_i = n$. Since each block on the main diagonal is irreducible, the nullspace of \bar{A} can be expressed as

$$\text{Null}(\bar{A}) = \left\{ \left(\begin{array}{c} \alpha_1 \mathbf{1}_{n_1} \\ \vdots \\ \alpha_m \mathbf{1}_{n_m} \end{array} \right) \mid \alpha_l \in \mathbb{R}, l = 1, \dots, m \right\}.$$

We assumed that B is block diagonalizable, which means that G_B is a union of isolated, strongly connected subgraphs, property which remains valid for the graph corresponding to $\bar{B} = TBT'$, since $G_{\bar{B}}$ is just a relabeled version of G_B . By adding \bar{B} to \bar{A} one of two phenomena can happen: either we can leave the graph $G_{\bar{A}}$ unchanged or we can create new connections among the vertices of $G_{\bar{A}}$. In the first case, $G_{\bar{B}} \subset G_{\bar{A}}$, and therefore $\text{Null}(\bar{A} + \bar{B}) = \text{Null}(\bar{A})$. In the second case we create new connections among the blocks of \bar{A} . But since all the subgraphs of \bar{B} are strongly connected, this means that if \bar{A}_i becomes connected to \bar{A}_j , then necessarily \bar{A}_j becomes connected to \bar{A}_i , and hence \bar{A}_i and \bar{A}_j form an irreducible (strongly connected) new block, whose nullspace is spanned by the vectors of all ones. Assuming that these are the only new connections that are added to $G_{\bar{A}}$, the nullspace of $\bar{A} + \bar{B}$ will have an expression similar to that of the nullspace of \bar{A} with the main difference being that the coefficients α_i and α_j will be equal. Therefore, in this particular case, the nullspace of $\bar{A} + \bar{B}$ can be expressed as

$$\text{Null}(\bar{A} + \bar{B}) = \left\{ \left(\begin{array}{c} \alpha_1 \mathbf{1}_{n_1} \\ \vdots \\ \alpha_m \mathbf{1}_{n_m} \end{array} \right) \mid \alpha_l \in \mathbb{R}, \alpha_i = \alpha_j, l = 1, \dots, m \right\}.$$

In general all blocks \bar{A}_i which become interconnected after adding \bar{B} will have equal coefficients in the expression of the nullspace of $\bar{A} + \bar{B}$, compared to the nullspace of \bar{A} . Therefore, $\text{Null}(\bar{A} + \bar{B}) \subset \text{Null}(\bar{A})$, which also means that $\text{Null}(A + B) \subset \text{Null}(A)$. Therefore, if $(A + B)v = 0$, then $Av = 0$, which also implies that $Bv = 0$ or $v \in \text{Null}(B)$. Hence if $v \in \text{Null}(A + B)$, then $v \in \text{Null}(A) \cap \text{Null}(B)$, which concludes the proof. \square

In the following corollary we present a property of the eigenspaces corresponding to the eigenvalue one of a set of probability transition matrices.

COROLLARY 3.6. *Let s be a positive integer, and let $\mathcal{A} = \{A_i\}_{i=1}^s$ be a set of doubly stochastic, probability transition matrices. Then,*

$$\text{Null} \left(\sum_{i=1}^s (A_i - I) \right) = \bigcap_{i=1}^s \text{Null}(A_i - I),$$

and $\dim(\text{Null}(\sum_{i=1}^s (A_i - I))) = 1$ if and only if $G_{\mathcal{A}}$ is strongly connected.

Proof. Since $A_i, i = 1, \dots, s$, are doubly stochastic, $A_i - I$ are block diagonalizable doubly stochastic generator matrices. Therefore, by recursively applying Lemma 3.5 $s - 1$ times, the first part of the corollary follows. For the second part of the corollary, note that by Corollary 3.5 of [19], $\frac{1}{N} \sum_{i=1}^s A_i$ has the algebraic multiplicity of its eigenvalue $\lambda = 1$ equal to one if and only if the graph associated with $\frac{1}{N} \sum_{i=1}^s A_i$ has a spanning tree or, in our case, is strongly connected. This in turn implies that $\dim(\text{Null}(\sum_{i=1}^s (A_i - I))) = 1$ if and only if $G_{\mathcal{A}}$ is strongly connected. \square

The following corollary is an immediate consequence of Corollary 3.5 of [19].

COROLLARY 3.7. *A generator matrix Λ has the algebraic multiplicity of its eigenvalue $\lambda = 0$ equal to one if and only if the graph associated with the matrix has a spanning tree.*

Proof. The proof follows immediately from Corollary 3.5 of [19] by forming the probability transition matrix $P = I + \epsilon\Lambda$ for some appropriate $\epsilon > 0$ and noting that $\text{Null}(P - I) = \text{Null}(\Lambda)$. \square

The following corollary is the counterpart of Lemma 3.7 of [19], in the case of generator matrices.

COROLLARY 3.8. *Let $\Lambda \in \mathbb{R}^{n \times n}$ be a generator matrix. If Λ has an eigenvalue $\lambda = 0$ with algebraic multiplicity equal to one, then $\lim_{t \rightarrow \infty} e^{\Lambda t} = \mathbf{1}v'$, where v is a nonnegative vector satisfying $\Lambda'v = 0$ and $v'\mathbf{1} = 1$.*

Proof. Choose $h_1 > 0$, and let $\{t_k^1\}_{k \geq 0}$ be a sequence given by $t_k^1 = h_1 k$ for all $k \geq 0$. Then

$$\lim_{k \rightarrow \infty} e^{\Lambda t_k^1} = \lim_{k \rightarrow \infty} e^{h_1 k \Lambda} = \lim_{k \rightarrow \infty} P_{h_1}^k,$$

where we defined $P_{h_1} \triangleq e^{h_1 \Lambda}$. From the theory of continuous-time Markov chains we know that P_{h_1} is a stochastic matrix with positive diagonal entries and that, given a vector $x \in \mathbb{R}^n$, $x'P_{h_1} = x'$ if and only if $x'\Lambda = 0$. This means that the algebraic multiplicity of the eigenvalue $\lambda = 1$ of P_{h_1} is one. By Lemma 3.7 of [19], we have that $\lim_{k \rightarrow \infty} P_{h_1}^k = \mathbf{1}v'_{h_1}$, where v_{h_1} is a nonnegative vector satisfying $P_{h_1}'v_{h_1} = v_{h_1}$ and $v'_{h_1} \mathbf{1} = 1$. Also, $\Lambda'v_{h_1} = 0$. Choose another $h_2 > 0$, and let $P_{h_2} \triangleq e^{h_2 \Lambda}$. As above, we have that

$$\lim_{k \rightarrow \infty} P_{h_2}^k = \mathbf{1}v'_{h_2},$$

where v_{h_2} satisfy properties similar to those satisfied by v_{h_1} . But since both vectors belong to the nullspace of Λ' of dimension one, they must be equal. Indeed, if x is a left eigenvector of Λ , then v_{h_1} and v_{h_2} can be written as $v_{h_1} = \alpha_1 x$ and $v_{h_2} = \alpha_2 x$. However, since $\mathbf{1}'v_{h_1} = 1$ and $\mathbf{1}'v_{h_2} = 1$, it follows that $\alpha_1 = \alpha_2$. We have shown that for any choice of $h > 0$,

$$\lim_{k \rightarrow \infty} e^{\Lambda t_k} = e^{hk \Lambda} = \mathbf{1}v',$$

where v is a nonnegative vector satisfying $\Lambda'v = 0$ and $\mathbf{1}'v = 1$, and, therefore, the result follows. \square

3.2. Preliminary results for the case where the agents' dynamics is expressed in discrete time. In this subsection we state and prove a set of results used to prove Theorem 2.2 in the case where the agents' dynamics is expressed in discrete time. Basically these results study the convergence properties of a sequence of matrices $\{Q^k\}_{k \geq 0}$, where Q has a particular structure which comes from the analysis of the first and second moments of the state vector $x(t)$.

LEMMA 3.9. *Let s be a positive integer, and let $\{A_{ij}\}_{i,j=1}^s$ be a set of $n \times n$ doubly stochastic, ergodic matrices which satisfy the assumptions of Theorem 3.1. Let $P = (p_{ij})$ be an $s \times s$ stochastic matrix corresponding to an irreducible, homogeneous Markov chain, and let \mathbf{P}_∞ be the limit set of the sequence $\{P^k\}_{k \geq 0}$. Consider the $ns \times ns$ dimensional matrix Q whose (i, j) th block is defined by $Q_{ij} \triangleq p_{ji}A_{ij}$. Then $\mathbf{P}'_\infty \otimes (\frac{1}{n}\mathbf{1}\mathbf{1}')$ is the limit set of the matrix sequence $\{Q^k\}_{k \geq 1}$, i.e.,*

$$(6) \quad \lim_{k \rightarrow \infty} d \left(Q^k, \mathbf{P}'_\infty \otimes \left(\frac{1}{n}\mathbf{1}\mathbf{1}' \right) \right) = 0.$$

Proof. The proof of this lemma is based on Corollary 3.2. The (i, j) th block entry of the matrix Q^k can be expressed as

$$(7) \quad (Q^k)_{ij} = \sum_{1 \leq i_1, \dots, i_{k-1} \leq s} p_{ji} p_{i_1 i_2} \dots p_{i_{k-1} i} A_{i_1 i_2} \dots A_{i_{k-1} j}.$$

Let p_{ji}^∞ be the (j, i) th entry of an arbitrary matrix in \mathbf{P}_∞ ; i.e., there exists a sequence $\{t_k\}_{k \geq 1} \subset \mathbb{N}$ such that $\lim_{k \rightarrow \infty} (P^{t_k})_{ji} = p_{ji}^\infty$.

We have that

$$\begin{aligned} \left\| (Q^k)_{ij} - p_{ji}^\infty \frac{1}{n} \mathbf{1}\mathbf{1}' \right\| &\leq \sum_{1 \leq i_{t_1}, \dots, i_{k-1} \leq s} (p_{ji} \dots p_{i_{k-1} i}) \left\| A_{i_{t_1} i_{t_2}} \dots A_{i_{k-1} j} - \frac{1}{n} \mathbf{1}\mathbf{1}' \right\| \\ &\quad + \left| \sum_{1 \leq i_1, \dots, i_{k-1} \leq s} p_{ji} \dots p_{i_{k-1} i} - p_{ji}^\infty \right| \left\| \frac{1}{n} \mathbf{1}\mathbf{1}' \right\| \\ &\leq \max_{i_1, \dots, i_{k-1}} \left\{ \left\| A_{i_1 i_2} \dots A_{i_{k-1} j} - \frac{1}{n} \mathbf{1}\mathbf{1}' \right\| \right\} \sum_{1 \leq i_1, \dots, i_{k-1} \leq s} p_{ji} \dots p_{i_{k-1} i} \\ &\quad + \left\| \frac{1}{n} \mathbf{1}\mathbf{1}' \right\| \left| \sum_{1 \leq i_1, \dots, i_{k-1} \leq s} p_{ji} \dots p_{i_{k-1} i} - p_{ji}^\infty \right|, \end{aligned}$$

where $\|\cdot\|$ was used to denote some matrix norm. Consider the limit of the left-hand side of the above inequality for the sequence $\{t_k\}_{k \geq 0}$. By Corollary 3.2 we know that

$$\lim_{k \rightarrow \infty} A_{i_{t_1} i_{t_2}} \dots A_{i_{t_{k-1}} j} = \frac{1}{n} \mathbf{1}\mathbf{1}'$$

for all sequences $i_{t_1}, \dots, i_{t_{k-1}}$, and since, obviously,

$$\lim_{k \rightarrow \infty} \sum_{1 \leq i_{t_1}, \dots, i_{t_{k-1}} \leq s} p_{ji} \dots p_{i_{t_{k-1}} i} = p_{ji}^\infty,$$

it results that

$$\lim_{k \rightarrow \infty} (Q^{t_k})_{ij} = p_{ji}^\infty \frac{1}{n} \mathbf{1}\mathbf{1}'.$$

Therefore $\mathbf{P}'_\infty \otimes (\frac{1}{n}\mathbf{1}\mathbf{1}')$ is the limit set for the sequence of matrices $\{Q^k\}_{k \geq 1}$. □

LEMMA 3.10. *Let s be a positive integer, and consider a set of doubly stochastic matrices with positive diagonal entries, $\mathcal{D} = \{D_i\}_{i=1}^s$, such that the corresponding graph $G_{\mathcal{D}}$ is strongly connected. Let P be the $s \times s$ dimensional probability transition matrix of an irreducible, homogeneous Markov chain, and let \mathbf{P}_{∞} be the limit set of the sequence $\{P^k\}_{k \geq 0}$. Consider the $ns \times ns$ matrix Q whose blocks are given by $Q_{ij} \triangleq p_{ji}D_j$. Then $\mathbf{P}_{\infty} \otimes (\frac{1}{n}\mathbb{1}\mathbb{1}')$ is the limit set of the sequence of matrices $\{Q^k\}_{k \geq 1}$, i.e.,*

$$(8) \quad \lim_{k \rightarrow \infty} d \left(Q^k, \mathbf{P}_{\infty} \otimes \left(\frac{1}{n} \mathbb{1}\mathbb{1}' \right) \right) = 0.$$

Proof. Our strategy consists in showing that there exists a $k \in \mathbb{N}$ such that each (i, j) th block matrix of Q^k becomes a weighted ergodic matrix, i.e., $(Q^k)_{ij} = p_{ji}^{(k)} A_{ij}^{(k)}$, where $A_{ij}^{(k)}$ is ergodic and $p_{ji}^{(k)} = (P^k)_{ji}$. If this is the case, we can apply Lemma 3.9 to obtain (8). The (i, j) th block matrix of Q^k looks as it does in (7), with the difference that in the current case $A_{ij} = D_j$:

$$(9) \quad (Q^k)_{ij} = \sum_{1 \leq i_1, \dots, i_{k-1} \leq s} p_{ji} p_{i_1 i_2} \dots p_{i_{k-1} i} D_j D_{i_1} \dots D_{i_{k-1}} = p_{ji}^{(k)} A_{ij}^{(k)},$$

where

$$A_{ij}^{(k)} \triangleq \sum_{1 \leq i_1, \dots, i_{k-1} \leq s} \alpha_{i_1, \dots, i_{k-1}} D_j D_{i_1} \dots D_{i_{k-1}},$$

with

$$\alpha_{i_1, \dots, i_{k-1}} \triangleq \begin{cases} p_{ji} p_{i_1 i_2} \dots p_{i_{k-1} i} / p_{ji}^{(k)}, & p_{ji}^{(k)} > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Note that each matrix product $D_j D_{i_1} \dots D_{i_{k-1}}$ appearing in $A_{ij}^{(k)}$ corresponds to a path from node j to node i in $k-1$ steps. Therefore, by the irreducibility assumption of P , there exists a k such that each matrix in the set \mathcal{D} appears at least once in one of the terms of the sum (9); i.e., $\{1, \dots, s\} \subseteq \{i_1, \dots, i_{k-1}\}$. Using an idea similar to that in Lemma 1 of [7] or Lemma 3.9 of [19], by Lemma 3.3, a lower bound for $D_j D_{i_1} \dots D_{i_{k-1}}$ is given by

$$(10) \quad D_j D_{i_1} \dots D_{i_{k-1}} \geq \gamma \sum_{l=1}^s D_l = \gamma \bar{D},$$

where $\gamma > 0$ depends on the matrices in \mathcal{D} and \bar{D} is a doubly stochastic matrix with positive entries,

$$\bar{D} = \sum_i D_i.$$

Since $G_{\mathcal{D}}$ is strongly connected, the same is true for $G_{\bar{D}}$. Therefore, \bar{D} corresponds to an irreducible, aperiodic (\bar{D} has positive diagonal entries), and hence ergodic, Markov chain. By inequality (10), it follows that the matrix product $D_j D_{i_1} \dots D_{i_{k-1}}$ is ergodic. This is enough to infer that $A_{ij}^{(k)}$ is ergodic as well, since it is a result of a

convex combination of (doubly) stochastic matrices with at least one ergodic matrix in the combination. Choose a k^* large enough such that for all nonzero $p_{ij}^{(k^*)}$, the matrices $A_{ij}^{(k^*)}$ are ergodic for all i, j . Such a k^* always exists due to the irreducibility assumption on P . In addition, by applying Lemma 3.3 again, any finite matrix product of the form $A_{i_1 i_1}^{(k^*)} A_{i_1 i_2}^{(k^*)} \dots A_{i_{m-1} i_m}^{(k^*)}$ is ergodic as well. Then according to Lemma 3.9, we have that for the subsequence $\{t_m\}_{m \geq 0}$, with $t_m = mk^*$,

$$(11) \quad \lim_{m \rightarrow \infty} d \left(Q^{t_m}, \mathbf{P}'_\infty \otimes \left(\frac{1}{n} \mathbf{1} \mathbf{1}' \right) \right) = 0.$$

The result follows by Proposition 3.4 since $\|Q\|_1 \leq 1$ and since $Q (\mathbf{P}'_\infty \otimes (\frac{1}{n} \mathbf{1} \mathbf{1}')) = \mathbf{P}'_\infty \otimes (\frac{1}{n} \mathbf{1} \mathbf{1}')$. \square

LEMMA 3.11. *Under the same assumptions as in Lemma 3.10, if we define the matrix blocks of Q as $Q_{ij} \triangleq p_{ji} D_j \otimes D_j$, then $\mathbf{P}'_\infty \otimes (\frac{1}{n^2} \mathbf{1} \mathbf{1}')$ is the limit set of the sequence $\{Q^k\}_{k \geq 1}$, i.e.,*

$$\lim_{k \rightarrow \infty} d \left(Q^k, \mathbf{P}'_\infty \otimes \left(\frac{1}{n^2} \mathbf{1}_{n^2} \mathbf{1}'_{n^2} \right) \right) = 0.$$

Proof. In the current setup (9) becomes

$$(12) \quad (Q^k)_{ij} = \sum_{1 \leq i_1, \dots, i_{k-1} \leq s} p_{j i_1} p_{i_1 i_2} \dots p_{i_{k-1} i} (D_j \otimes D_j) (D_{i_1} \otimes D_{i_1}) \dots (D_{i_{k-1}} \otimes D_{i_{k-1}}).$$

The result follows from the same arguments used in Lemma 3.10 together with the fact that the matrix products in (12) can be written as $(D_j \otimes D_j) (D_{i_1} \otimes D_{i_1}) \dots (D_{i_{k-1}} \otimes D_{i_{k-1}}) = (D_j D_{i_1} \dots D_{i_{k-1}}) \otimes (D_j D_{i_1} \dots D_{i_{k-1}})$ and with the observation that the Kronecker product of an ergodic matrix with itself produces an ergodic matrix as well. \square

3.3. Preliminary results for the case where the agents' dynamics is expressed in continuous time. The following two lemmas emphasize geometric properties of two matrices arising from the linear dynamics of the first and second moments of the state vector, in the continuous-time case.

LEMMA 3.12. *Let s be a positive integer, and let $\mathcal{C} = \{C_i\}_{i=1}^s$ be a set of $n \times n$ doubly stochastic matrices such that $G_{\mathcal{C}}$ is strongly connected. Consider also an $s \times s$ generator matrix $\Lambda = (\lambda_{ij})$ corresponding to an irreducible Markov chain with stationary distribution $\pi = (\pi_i)$. Define the matrices $A \triangleq \text{diag}(C'_i, i = 1, \dots, s)$ and $B \triangleq \Lambda \otimes I$. Then $A + B$ has an eigenvalue $\lambda = 0$ with algebraic multiplicity one and with corresponding right and left eigenvectors given by $\mathbf{1}_{ns}$ and $(\pi_1 \mathbf{1}'_n, \pi_2 \mathbf{1}'_n, \dots, \pi_s \mathbf{1}'_n)$, respectively.*

Proof. We first note that $A + B$ is a generator matrix and that both A and B are block diagonalizable (indeed, A has doubly stochastic matrices on its main diagonal, and B contains n copies of the irreducible Markov chain corresponding to Λ). Therefore, $A + B$ has an eigenvalue $\lambda = 0$ with algebraic multiplicity at least one.

Let v be a vector in the nullspace of $A + B$. By Lemma 3.5, we have that $v \in \text{Null}(A)$ and $v \in \text{Null}(B)$. Given the structure of B , v must respect the following pattern:

$$v' = \underbrace{(u', \dots, u')}_{s \text{ times}} \text{ for some } u \in \mathbb{R}^n.$$

But since $v \in \text{Null}(A)$, we have that $C'_i u = 0$, $i = 1, \dots, s$, or $\mathbf{C}u = 0$, where $\mathbf{C} = \sum_{i=1}^s C'_i$. Since G_C was assumed strongly connected, \mathbf{C} corresponds to an irreducible Markov chain, and it follows that u must be of the form $u = \alpha \mathbf{1}$ for some $\alpha \in \mathbb{R}$. By backtracking, we get that $v = \alpha \mathbf{1}$ for some $\alpha \in \mathbb{R}$ and consequently $\text{Null}(A + B) = \text{span}(\mathbf{1})$. Therefore, $\lambda = 0$ has algebraic multiplicity one, with right eigenvector given by $\mathbf{1}$. By simple verification we note that $(\pi_1 \mathbf{1}', \pi_2 \mathbf{1}', \dots, \pi_s \mathbf{1}')$ is a left eigenvector corresponding to the eigenvalue $\lambda = 0$. \square

LEMMA 3.13. *Let s be a positive integer, and let $\mathcal{C} = \{C'_i\}_{i=1}^s$ be a set of $n \times n$ doubly stochastic matrices such that G_C is strongly connected. Consider also an $s \times s$ generator matrix $\Lambda = (\lambda_{ij})$ corresponding to an irreducible Markov chain with stationary distribution $\pi = (\pi_i)$. Define the matrices $A \triangleq \text{diag}(C'_i \oplus C'_i, i = 1, \dots, s)$ and $B \triangleq \Lambda \otimes I$. Then $A + B$ has an eigenvalue $\lambda = 0$ with algebraic multiplicity one, with corresponding right and left eigenvectors given by $\mathbf{1}_{n^2 s}$ and $(\pi_1 \mathbf{1}'_{n^2}, \pi_2 \mathbf{1}'_{n^2}, \dots, \pi_s \mathbf{1}'_{n^2})$, respectively.*

Proof. It is not difficult to check that $A + B$ is a generator matrix. Also, we note that $C'_i \oplus C'_i = C'_i \otimes I + I \otimes C'_i$ is block diagonalizable since both $C'_i \otimes I$ and $I \otimes C'_i$ are block diagonalizable. Indeed, since C_i is doubly stochastic, it is block diagonalizable. The matrix $C'_i \otimes I$ contains n isolated copies of C'_i , and therefore it is block diagonalizable. Also, $I \otimes C'_i$ has n blocks on its diagonal, each block being given by C'_i , and it follows that it is block diagonalizable as well.

Let v be a vector in the nullspace of $A + B$. By Lemma 3.5, $v \in \text{Null}(A)$ and $v \in \text{Null}(B)$. From the structure of B we note that v must be of the form

$$v' = \underbrace{(u', \dots, u')}_{s \text{ times}} \text{ for some } u \in \mathbb{R}^{n^2}.$$

Consequently we have that $(C'_i \oplus C'_i)u = 0$, $i = 1, \dots, s$, or $(\mathbf{C} \oplus \mathbf{C})u = 0$, where $\mathbf{C} = \sum_{i=1}^s C'_i$. Since G_C is strongly connected, \mathbf{C} is a generator matrix corresponding to an irreducible Markov chain. By again applying Lemma 3.5 for the matrix $\mathbf{C} \oplus \mathbf{C} = I \otimes \mathbf{C} + \mathbf{C} \otimes I$, we get that u must have the form

$$u' = \underbrace{(\bar{u}', \dots, \bar{u}')}_{n \text{ times}},$$

where $\bar{u} \in \mathbb{R}^n$ and $\mathbf{C}\bar{u} = 0$. But \mathbf{C} is irreducible and therefore $\bar{u} = \alpha \mathbf{1}_n$, or $u = \alpha \mathbf{1}_{n^2}$, or finally $v = \alpha \mathbf{1}_{n^2 s}$, where $\alpha \in \mathbb{R}$. Consequently, $\text{Null}(A + B) = \text{span}(\mathbf{1})$, which means the eigenvalue $\lambda = 0$ has algebraic multiplicity one. By simple verification, we note that $(\pi_1 \mathbf{1}'_{n^2}, \pi_2 \mathbf{1}'_{n^2}, \dots, \pi_s \mathbf{1}'_{n^2})$ is a left eigenvector corresponding to the zero eigenvalue. \square

4. Proof of the convergence theorem. The proof will focus on showing that the state vector $x(t)$ converges in the mean square sense to average consensus. Equivalently, by making the change of variable $z(t) = x(t) - av(x_0)\mathbf{1}$, we will actually show that $z(t)$ is mean square stable for the initial condition $z(0) = x_0 - av(x_0)\mathbf{1}$, where $z(t)$ respects the same dynamic equation as $x(t)$.

Using results from the stability theory of Markovian jump linear systems, for constant probability transition matrices (generator matrices), mean square stability also implies stability in the almost sure sense (see, for instance, Corollary 3.46 of [3] for the discrete-time case or Theorem 2.1 of [4] for the continuous-time case, bearing in mind that we are interested in the stability property being satisfied for a specific initial condition, rather than for any initial condition). For time varying probability

transition matrices (generator matrices), without additional assumptions, the two modes of convergence are not necessarily equivalent.

We first prove the discrete-time case after which we continue with the proof for the continuous-time case.

4.1. Discrete-time case—sufficiency. Let $V(t)$ denote the second moment of the state vector:

$$V(t) \triangleq E[x(t)x(t)^T],$$

where we used E to denote the expectation operator. The matrix $V(t)$ can be expressed as

$$(13) \quad V(t) = \sum_{i=1}^s V_i(t),$$

where $V_i(t)$ is given by

$$(14) \quad V_i(t) \triangleq E[x(t)x(t)^T \chi_{\{\theta(t)=i\}}], \quad i = 1, \dots, s,$$

with $\chi_{\{\theta(t)=i\}}$ being the indicator function of the event $\{\theta(t) = i\}$.

The set of discrete coupled Lyapunov equations governing the evolution of the matrices $V_i(t)$ is given by

$$(15) \quad V_i(t+1) = \sum_{j=1}^s p_{ji} D_j V_j(t) D_j^T, \quad i = 1, \dots, s,$$

with initial conditions $V_i(0) = q_i x_0 x_0^T$. By defining $\eta(t) \triangleq \text{col}(V_i(t), i = 1, \dots, s)$, we obtain a vectorized form of equations (15),

$$(16) \quad \eta(t+1) = \Gamma_d \eta(t),$$

where Γ_d is an $n^2 s \times n^2 s$ matrix given by

$$(17) \quad \Gamma_d = \begin{pmatrix} p_{11} D_1 \otimes D_1 & \cdots & p_{s1} D_s \otimes D_s \\ \vdots & \ddots & \vdots \\ p_{1s} D_1 \otimes D_1 & \cdots & p_{ss} D_s \otimes D_s \end{pmatrix} \quad \text{and} \quad \eta_0 = \begin{pmatrix} q_1 \text{col}(x_0 x_0^T) \\ \vdots \\ q_s \text{col}(x_0 x_0^T) \end{pmatrix}.$$

We note that Γ_d satisfies all the assumptions of Lemma 3.11, and hence we get

$$\lim_{k \rightarrow \infty} d \left(\Gamma_d^k, \mathbf{P}'_\infty \otimes \left(\frac{1}{n^2} \mathbf{1} \mathbf{1}' \right) \right) = 0,$$

where \mathbf{P}_∞ is the limit set of the matrix sequence $\{P^k\}_{k \geq 0}$. Using the observation that

$$\frac{1}{n^2} \mathbf{1} \mathbf{1}' \text{col}(x_0 x_0^T) = av(x_0)^2 \mathbf{1},$$

the limit of the sequence $\{\eta(t_k)\}_{k \geq 0}$, where $\{t_k\}_{k \geq 0}$ is such that $\lim_{k \rightarrow \infty} (P^{t_k})_{ij} = p_{ij}^\infty$, is

$$\lim_{k \rightarrow \infty} \eta(t_k)' = av(x_0)^2 \begin{pmatrix} \sum_{j=1}^s p_{j1}^\infty q_j \mathbf{1} \\ \vdots \\ \sum_{j=1}^s p_{js}^\infty q_j \mathbf{1}' \end{pmatrix}.$$

By collecting the entries of $\lim_{k \rightarrow \infty} \eta(t_k)$, we obtain

$$\lim_{k \rightarrow \infty} V_i(t_k) = av(x_0)^2 \left(\sum_{j=1}^s p_{ji}^\infty q_j \right) \mathbf{1}\mathbf{1}',$$

and from (13) we get

$$(18) \quad \lim_{k \rightarrow \infty} V(t_k) = av(x_0)^2 \mathbf{1}\mathbf{1}'$$

since $\sum_{i,j=1}^s p_{ji}^\infty q_j = 1$. By repeating the previous steps for all subsequences generating limit points for $\{P^k\}_{k \geq 0}$, we obtain that (18) holds for any sequence in \mathbb{N} .

Through a process similar to that in the case of the second moment (instead of Lemma 3.11 we use Lemma 3.10), we show that

$$(19) \quad \lim_{k \rightarrow \infty} E[x(t)] = av(x_0)\mathbf{1}.$$

From (18) and (19) we have that

$$\begin{aligned} \lim_{t \rightarrow \infty} E[\|x(t) - av(x_0)\mathbf{1}\|^2] &= \lim_{t \rightarrow \infty} \text{tr}(E[(x(t) - av(x_0)\mathbf{1})(x(t) - av(x_0)\mathbf{1})']) \\ &= \lim_{t \rightarrow \infty} \text{tr}(E[x(t)x(t)'] - av(x_0)\mathbf{1}E[x(t)'] - av(x_0)E[x(t)]\mathbf{1}' + av(x_0)^2\mathbf{1}\mathbf{1}') = 0. \end{aligned}$$

Therefore, $x(t)$ converges to average consensus in the mean square sense and consequently in the almost sure sense as well. \square

4.2. Discrete-time case—necessity. If G_A is not strongly connected, then by Corollary 3.6, $\dim(\bigcap_{i=1}^s \text{Null}(A_i - I)) > 1$. Consequently, there exists a vector $v \in \bigcap_{i=1}^s \text{Null}(A_i - I)$ such that $v \notin \text{span}(\mathbf{1})$. If we choose v as initial condition, for every realization of $\theta(t)$, we have that

$$x(t) = v \quad \forall t \geq 0,$$

and therefore consensus cannot be reached in the sense of Definition 2.1. \square

4.3. Sufficiency—continuous time. Using the same notation as in the discrete-time case, the dynamic equations describing the evolution of the second moment of $x(t)$ are given by

$$(20) \quad \frac{d}{dt} V_i(t) = C_i V_i(t) + V_i(t) C_i' + \sum_{j=1}^s \lambda_{ji} V_j(t), \quad i = 1, \dots, s,$$

equations whose derivation is treated in [5]. By defining the vector $\eta(t) \triangleq \text{col}(V_i(t), i = 1, \dots, s)$, the vectorized equivalent of (20) is given by

$$(21) \quad \frac{d}{dt} \eta(t) = \Gamma_c \eta(t),$$

where

$$\Gamma_c = \begin{pmatrix} C_1 \oplus C_1 & 0 & \dots & 0 \\ 0 & C_2 \oplus C_2 & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & C_s \oplus C_s \end{pmatrix} + \Lambda' \otimes I \quad \text{and} \quad \eta_0 = \begin{pmatrix} q_1 \text{col}(x_0 x_0') \\ q_2 \text{col}(x_0 x_0') \\ \vdots \\ q_s \text{col}(x_0 x_0') \end{pmatrix}.$$

By Lemma 3.13, the eigenspace corresponding to the zero eigenvalue of Γ_c has dimension one, with unique (up to the multiplication by a scalar) left and right eigenvectors given by $\mathbb{1}_{n^2_s}$ and $\frac{1}{n^2}(\pi_1 \mathbb{1}'_{n^2}, \pi_2 \mathbb{1}'_{n^2}, \dots, \pi_s \mathbb{1}'_{n^2})$, respectively. Since Γ'_c is a generator matrix with an eigenvalue zero of algebraic multiplicity one, by Corollary 3.8 we have that $\lim_{t \rightarrow \infty} e^{\Gamma'_c t} = v \mathbb{1}'$, where $v' = \frac{1}{n^2}(\pi_1 \mathbb{1}', \pi_2 \mathbb{1}', \dots, \pi_s \mathbb{1}')$. Therefore, as t goes to infinity, we have that

$$\lim_{t \rightarrow \infty} \eta(t) = \begin{pmatrix} \pi_1 \frac{\mathbb{1}\mathbb{1}'}{n^2} & \cdots & \pi_1 \frac{\mathbb{1}\mathbb{1}'}{n^2} \\ \vdots & \ddots & \vdots \\ \pi_s \frac{\mathbb{1}\mathbb{1}'}{n^2} & \cdots & \pi_s \frac{\mathbb{1}\mathbb{1}'}{n^2} \end{pmatrix} \begin{pmatrix} q_1 \text{col}(x_0 x'_0) \\ \vdots \\ q_s \text{col}(x_0 x'_0) \end{pmatrix}.$$

By noting that

$$\frac{\mathbb{1}\mathbb{1}'}{n^2} \text{col}(x_0 x'_0) = av(x_0)^2 \mathbb{1}_{n^2},$$

we further get

$$\lim_{t \rightarrow \infty} \eta(t) = av(x_0)^2 \begin{pmatrix} \pi_1 \mathbb{1}_{n^2} \\ \vdots \\ \pi_s \mathbb{1}_{n^2} \end{pmatrix}.$$

Rearranging the columns of $\lim_{t \rightarrow \infty} \eta(t)$, we get

$$\lim_{t \rightarrow \infty} V_i(t) = av(x_0)^2 \pi_i \mathbb{1}\mathbb{1}'$$

or

$$\lim_{t \rightarrow \infty} V(t) = av(x_0)^2 \mathbb{1}\mathbb{1}'.$$

Through a similar process (this time using Lemma 3.12), we can show that

$$\lim_{t \rightarrow \infty} E[x(t)] = av(x_0) \mathbb{1}.$$

Therefore, $x(t)$ converges to average consensus in the mean square sense and consequently in the almost surely sense.

4.4. Necessity—continuous time. The proof follows the same lines as in the discrete-time case.

5. Discussion. In the previous sections we proved a convergence result for the stochastic, linear consensus problem for the cases where the dynamics of the agents was expressed in both discrete and continuous time. Our main contributions consist of considering a Markovian process, not necessarily ergodic, as underlying process for the random communication graph, and using a Markovian jump system theory inspired approach to prove this result. In what we have shown, we assumed that the Markov process $\theta(t)$ was irreducible and that the matrices D_i and C_i were doubly stochastic. We can assume, for instance, that $\theta(t)$ is not irreducible (i.e., $\theta(k)$ may have transient states). We treated this case in [9] (only for discrete-time dynamics), and we showed that convergence in the sense of Definition 2.1 is achieved if and only if the union of graphs corresponding to each of the irreducible closed sets of states of the Markov chain produces a strongly connected graph. This should be intuitively clear since the

probability of returning to a transient state converges to zero as time goes to infinity, and therefore the influence of the matrices D_i (or C_i), corresponding to the transient states, is canceled. We can also assume that D_i and C_i are not necessarily doubly stochastic. We treated this case (again only for the discrete-time dynamics) in [10], and we showed that the state converges in the mean square sense and in the almost sure sense to consensus, and not necessarily to average consensus. From the technical point of view, the difference consists of the fact that the $n^2 \times n^2$ block matrices of $\{\Gamma_d^t\}_{t \geq 0}$ (or $\{e^{t\Gamma_c}\}_{t \geq 0}$) no longer converge to $\pi_i \frac{1}{n^2} \mathbf{1}\mathbf{1}'$ but to $\pi_i \mathbf{1}c'$, for some vector $c \in \mathbb{R}^{n^2}$ with nonnegative entries summing up to one, a vector c which in general cannot be a priori determined. In relevant distributed computation applications (such as distributed state estimation or distributed optimization), however, convergence to average consensus is desirable, and therefore the assumption that D_i or C_i are doubly stochastic makes sense.

The proof of Theorem 2.2 was based on the analysis of two matrix sequences $\{e^{\Gamma_c t}\}_{t \geq 0}$ and $\{\Gamma_d^t\}_{t \geq 0}$ arising from the dynamic equations of the state's second moment for continuous and discrete time, respectively. The reader may have noted that we approached the analysis of the two sequences in different ways. In the case of continuous-time dynamics, our approach was based on showing that the left and right eigenspaces induced by the zero eigenvalue of Γ_c have dimension one, and we provided the left and right eigenvectors (bases of the respective subspaces). The convergence of $\{e^{\Gamma_c t}\}_{t \geq 0}$ followed from Corollary 3.8. In the case of the discrete-time dynamics, we analyzed the sequence $\{\Gamma_d^t\}_{t \geq 0}$ by looking at how the matrix blocks of Γ_d^t evolve as t goes to infinity. Although, similarly to the continuous-time case, we could have proved properties of Γ_d related to the left and right eigenspaces induced by the eigenvalue one, this would not have been enough in the discrete-time case. This is because, through $\theta(t)$, Γ_d can be periodic, in which case the sequence $\{\Gamma_d^t\}_{t \geq 0}$ does not converge (remember that in the discrete-time consensus problems, a common assumption is that the consensus matrices have positive diagonal entries which prevent them from being periodic).

Other approaches to the problem we considered can be found in [20, 21]. A first difference between these works and ours is that there the various weights would vary with time, while in our case, so long as there is a connection, the weight is fixed. Moreover, in the case of i.i.d. random graphs [20], or, more generally, in the case of strictly stationary, ergodic random graphs [21], a necessary and sufficient condition for reaching consensus almost surely (in the discrete-time case) is $|\lambda_2(E[\mathbf{D}_{\theta(t)}])| < 1$, where λ_2 denotes the eigenvalue with second largest modulus. In the case of Markovian random topology, however, such a condition may not hold, neither for each time t , nor in the limit. Take, for instance, two (symmetric) stochastic matrices D_1 and D_2 such that each of the graphs G_{D_1} and G_{D_2} , respectively, is not strongly connected, but their union is. If the two-state Markov chain $\theta(t)$ is periodic, with transitions given by $p_{11} = p_{22} = 0$ and $p_{12} = p_{21} = 1$, we note that $\lambda_2(E[\mathbf{D}_{\theta(t)}]) = 1$ for all $t \geq 0$, although the sequence $\{E[\mathbf{D}_{\theta(t)}]\}_{t \geq 0}$ does not have a limit. Yet, consensus is reached. The assumption that allowed the aforementioned necessary and sufficient condition to hold is that $\theta(t)$ is a stationary process (which in turn implies that $E[\mathbf{D}_{\theta(t)}]$ is constant for all $t \geq 0$). However, this is not necessarily true if $\theta(t)$ is a (homogeneous) irreducible Markov chain, *unless* the initial distribution is the stationary distribution.

For the discrete-time case we can formulate a result involving the second largest eigenvalue of the time average expectation of $\mathbf{D}_{\theta(t)}$, i.e., $\lim_{N \rightarrow \infty} \frac{\sum_{t=1}^N E[\mathbf{D}_{\theta(t)}]}{N}$, which reflects the proportion of time $\mathbf{D}_{\theta(t)}$ spends in each state of the Markov chain.

PROPOSITION 5.1. *Consider the stochastic system (1). Then, under Assumption 2.1, the state vector $x(t)$ converges to average consensus in the sense of Definition 2.1 if and only if*

$$\left| \lambda_2 \left(\lim_{N \rightarrow \infty} \frac{\sum_{t=0}^N E[\mathbf{D}_{\theta(t)}]}{N} \right) \right| < 1.$$

Proof. The time average of $E[\mathbf{D}_{\theta(t)}]$ can be explicitly written as

$$\lim_{N \rightarrow \infty} \frac{\sum_{t=0}^N E[\mathbf{D}_{\theta(t)}]}{N} = \sum_{i=1}^s \pi_i D_i = \bar{D},$$

where $\pi = (\pi_i)$ is the stationary distribution of $\theta(t)$. By Corollary 3.5 in [19], $|\lambda_2(\bar{D})| < 1$ if and only if the graph corresponding to \bar{D} has a spanning tree or, in our case, is strongly connected. But the graph corresponding to \bar{D} is the same as $G_{\mathcal{D}}$, and the result follows from Theorem 2.2. \square

Unlike in the discrete-time case, in the case of continuous-time dynamics, we know that if there exists a stationary distribution π (under the irreducibility assumption), the probability distribution of $\theta(t)$ converges to π ; hence the time averaging is not necessary. In the following we introduce (without proof since basically it is similar to the proof of Proposition 5.1) a necessary and sufficient condition for reaching average consensus, involving the expected value of the second largest eigenvalue of $\mathbf{C}_{\theta(t)}$, for the continuous-time dynamics.

PROPOSITION 5.2. *Consider the stochastic system (2). Then, under Assumption 2.1, the state vector $x(t)$ converges to average consensus in the sense of Definition 2.1 if and only if*

$$\operatorname{Re} \left(\lambda_2 \left(\lim_{t \rightarrow \infty} E[\mathbf{C}_{\theta(t)}] \right) \right) < 0.$$

Our analysis also provides estimates on the rate of convergence to average consensus in the mean square sense. From linear dynamic equations of the state's second moment we notice that the eigenvalues of Γ_d and Γ_c dictate how fast the second moment converges to average consensus. Since Γ'_d is a probability transition matrix and since Γ'_c is a generator matrix, an estimate of the rate of convergence of the second moment of $x(t)$ to average consensus is given by the second largest eigenvalue (in modulus) of Γ_d for the discrete-time case, and by the real part of the second largest eigenvalue of Γ_c for the continuous-time case.

6. Conclusion. In this paper we analyzed the convergence properties of the linear consensus problem when the communication topology is modeled as a directed random graph with an underlying Markovian process. We addressed the cases where the dynamics of the agents is expressed both in continuous and in discrete time. Under some assumptions on the communication topologies, we provided a rigorous mathematical proof for the intuitive necessary and sufficient conditions for reaching average consensus in the mean square sense and the almost sure sense. These conditions are expressed in terms of connectivity properties of the union of graphs corresponding to the states of the Markov process. The aim of this paper is to show how mathematical techniques from the stability theory of the Markovian jump systems, in conjunction with results from the matrix and graph theory, can be used to prove convergence results for consensus problems under a stochastic framework.

REFERENCES

- [1] V. BORKAR AND P.P. VARAIYA, *Asymptotic agreement in distributed estimation*, IEEE Trans. Automat. Control, 27 (1982), pp. 650–655.
- [2] R. COGBURN, *On products of random stochastic matrices*, in Random Matrices and Their Applications (Proceedings of the AMS-IMS-SIAM Joint Summer Research Conference, Brunswick, ME, 1984), Contemp. Math. 50, AMS, Providence, RI, 1986, pp. 199–213.
- [3] O.L.V. COSTA, M.D. FRAGOSO, AND R.P. MARQUES, *Discrete-Time Markov Jump Linear Systems*, Springer-Verlag, London, 2005.
- [4] X. FENG AND K.A. LOPARO, *Stability of linear Markovian jump systems*, in Proceedings of the 29th IEEE Conference on Decision and Control, Honolulu, HI, 1990, pp. 1408–1413.
- [5] M.D. FRAGOSO AND O.L.V. COSTA, *A unified approach for stochastic and mean square stability of continuous-time linear systems with Markovian jumping parameters and additive disturbances*, in SIAM J. Control Optim., 44 (2005), pp. 1165–1191.
- [6] Y. HATANO AND M. MESBAHI, *Agreement over random networks*, IEEE Trans. Automat. Control, 50 (2005), pp. 1867–1872.
- [7] A. JADBABAIE, J. LIN, AND A.S. MORSE, *Coordination of groups of mobile autonomous agents using nearest neighbor*, IEEE Trans. Automat. Control, 48 (2003), pp. 998–1001.
- [8] S. KAR AND J.M.F. MOURA, *Sensor networks with random links: Topology design for distributed consensus*, IEEE Trans. Signal Process., 56 (2008), pp. 3315–3326.
- [9] I. MATEI, N. MARTINS, AND J. BARAS, *Almost sure convergence to consensus in Markovian random graphs*, in Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico, 2008, pp. 3535–3540.
- [10] I. MATEI, N. MARTINS, AND J. BARAS, *Consensus problems with directed Markovian communication patterns*, in Proceedings of the 2009 American Control Conference, St. Louis, MO, 2009, pp. 1298–1303.
- [11] L. MOREAU, *Stability of multiagent systems with time-dependent communication links*, IEEE Trans. Automat. Control, 50 (2005), pp. 169–182.
- [12] A. NEDIĆ AND A. OZDAGLAR, *Convergence rate for consensus with delays*, J. Global Optim., 47 (2010), pp. 437–456.
- [13] A. NEDIĆ AND A. OZDAGLAR, *Distributed subgradient methods for multi-agent optimization*, IEEE Trans. Automat. Control, 54 (2009), pp. 48–61.
- [14] A. NEDIĆ, A. OZDAGLAR, AND P.A. PARRILO, *Constrained consensus and optimization in multi-agent networks*, IEEE Trans. Automat. Control, 55 (2010), pp. 922–938.
- [15] R. OLFATI-SABER, *Distributed Kalman filter with embedded consensus filters*, in Proceedings of the 44th IEEE Conference on Decision and Control, 2005, pp. 8179–8184.
- [16] R. OLFATI-SABER AND R.M. MURRAY, *Consensus protocols for networks of dynamic agents*, in Proceedings of the 2003 American Control Conference, 2003, pp. 951–956.
- [17] R. OLFATI-SABER AND R.M. MURRAY, *Consensus problem in networks of agents with switching topology and time-delays*, IEEE Trans. Automatic Control, 49 (2004), pp. 1520–1533.
- [18] M. PORFIRI AND D.J. STILWELL, *Consensus seeking over random directed weighted graphs*, IEEE Trans. Automat. Control, 52 (2007), pp. 1767–1773.
- [19] W. REN AND R.W. BEARD, *Consensus seeking in multiagent systems under dynamically changing interaction topologies*, IEEE Trans. Automat. Control, 50 (2005), pp. 655–661.
- [20] A. TAHAZ-SALEHI AND A. JADBABAIE, *A necessary and sufficient condition for consensus over random networks*, IEEE Trans. Automat. Control, 53 (2008), pp. 791–795.
- [21] A. TAHAZ-SALEHI AND A. JADBABAIE, *Consensus over ergodic stationary graph processes*, IEEE Trans. Automat. Control, 55 (2010), pp. 225–230.
- [22] J.N. TSITSIKLIS, *Problems in Decentralized Decision Making and Computation*, Ph.D. dissertation, Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 1984.
- [23] 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.
- [24] J. WOLFOWITZ, *Product of indecomposable aperiodic stochastic matrices*, Proc. Amer. Math. Soc., 15 (1963), pp. 733–736.