THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2009-18

Convergence results for the linear consensus problem under Markovian random graphs

Ion Matei, John S. Baras

Convergence results for the linear consensus problem under Markovian random graphs

Ion Matei and John S. Baras

Abstract

This note discusses the linear discrete and continuous time consensus problem for a network of dynamic agents with directed information flows and random switching topologies. The switching is determined by a Markov chain, each topology corresponding to a state of the Markov chain. We show that, under doubly stochastic assumption on the matrices involved in the linear consensus scheme, average consensus is achieved in the mean square sense and almost surely if and only if the graph resulted from the union of graphs corresponding to the states of the Markov chain is strongly connected. The aim of this note is to show how techniques from Markovian jump linear systems theory, in conjunction with results inspired by matrix and graph theory, can be used to prove convergence results for stochastic consensus problems.

I. Introduction

A consensus problem consists of a group of dynamic agents who seek to agree upon certain quantities of interest by exchanging information among them 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, parallel computing, etc. Distributed computation over networks has a long history in control theory starting with the work of Borkar and Varaiya [1], Tsitsikils, Bertsekas and Athans [23], [24] on asynchronous agreement problems and parallel computing. A theoretical framework for solving consensus problems was introduced by Olfati-Saber and Murray in [17], [18], while Jadbabaie et al. studied alignment problems [8] for reaching an agreement. Relevant extensions of the consensus problem were done by Ren and Beard [16], by Moreau in [11] or, more recently, by Nedic and Ozdaglar in [13], [12].

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 [9] 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 [15], 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 model of the communication topology, Tahbaz-Salehi and Jadbabaie give necessary and sufficient conditions for almost sure convergence to consensus in [21], while in [22], the authors extend the applicability of their necessary and sufficient conditions to strictly stationary ergodic random graphs.

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 (possible directed) random 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 both the continuous and discrete time cases and, under certain assumption on the matrices involved in the linear scheme, we give necessary and sufficient conditions such that average consensus is achieved in the mean square sense and in the almost sure sense. The Markovian switching model goes beyond the common i.i.d. assumption on the random communication topology and appears in the 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.

A. Basic notations and definitions

We will denote by $\mathbb{1}$ the vector of all ones. If the dimension of the vector needs to be emphasize, and index will be added for clarity (for example, if $\mathbb{1}$ is an n dimensional vector, we will explicitly mark this by using $\mathbb{1}_n$). Let x be a vector in $\mathbb{R}^n$. By $av(x)$ we denote the quantity $av(x) = x'\mathbb{1}/\mathbb{1}'\mathbb{1}$. The symbols $\otimes$ and $\oplus$ represent the Kronecker product and sum, respectively.

Given a matrix A, $Null(A)$ designates the nullspace of the considered matrix. If $\mathscr{X}$ represents some finite dimensional (sub-) space, $dim(\mathscr{X})$ gives us the dimension of $\mathscr{X}$.

Let $\mathscr{M}$ be a set of matrices. By $\mathscr{M} \otimes A$, where A is some arbitrary matrix, we understand the following matrix set: $\mathscr{M} \otimes A = \{M \otimes A \mid M \in \mathscr{M}\}$. Also by $\mathscr{M}'$ we denote the set of the transpose matrices of $\mathscr{M}$, i.e. $\mathscr{M}' = \{M' \mid M \in \mathscr{M}\}$. By writing that $A\mathscr{M} = \mathscr{M}$ we understand that $AM \in \mathscr{M}$, for any $M \in \mathscr{M}$.

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

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

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. We say that matrix A *corresponds* 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 non-zero. The graphs corresponding to A will be denoted by G_A.

Definition 1.2: Let s be a positive integer and let $\mathscr{A} = \{A_1, \ldots, A_s\}$ be a set of matrices with a corresponding set of graphs $\mathscr{G} = \{G_{A_1}, \ldots, G_{A_s}\}$. We say that the graph $G_{\mathscr{A}}$ *corresponds* to the set $\mathscr{A}$ if it is given by the union of graphs in $\mathscr{G}$, i.e.

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

In this note we will use mainly to type of matrices: probability transition matrices (row sum up to one) and rate transition matrices (row sum up to zero). Using an abuse of designation, a rate matrix whose both rows and columns sum up to zero will be called *doubly stochastic rate transition matrix*.

To simplify the exposition we will sometimes characterize a probability/rate transition 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/rate transition matrix. We say that A is *block diagonalizable* if there exists a similarity transformation P, encapsulating a successions of row permutations, such that PAP' is a bloc diagonal matrix with irreducible blocks on the main diagonal.

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

Paper organization: This paper has five sections besides the introduction: In Section II we present the setup and formulation of the problem and we state our main convergence theorem. In Section III we derive a number of results which will constitute the core of the proof of our main result, presented in Section IV. We continue with a discussion of the convergence result in Section V, and we end the paper with some conclusions.

II. Problem formulation and statement of the convergence result

We assume a group of n agents, labeled 1 through n, organized in a communication network whose topology is give by a time varying graph $\mathbf{G}(t) = (V, E(t))$, 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 which governs its evolution, given by a homogeneous continuous or discrete time Markov chain $\theta(t)$, taking values in the finite set $\{1,\ldots,s\}$. In the case of a discrete-time Markov chain, $\theta(t)$ has a transition probability matrix $P = (p_{ij})$ (rows sum up to one), while in the case of a continuous time Markov chain it has a rate transition matrix $\Lambda = (\lambda_{ij})$ (rows sum up to zero). The random graph $\mathbf{G}(t)$ takes values is a finite set of graphs $\mathscr{G} = \{G_1,\ldots,G_s\}$ with probability $Pr(\mathbf{G}(t) = G_i) = Pr(\theta(t) = i)$, for $i = 1\ldots s$. We denote by $q = (q_i)$ the initial distribution of $\theta(t)$.

Letting $x(t)$ denote the state of the n agents, we model the discrete-time dynamics of the agents by the following linear stochastic difference equation

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

where $\mathbf{D}_{\theta(t)}$ is a random matrix taking values in the finite set $\mathscr{D} = \{D_1,\ldots,D_s\}$, with probability distribution $Pr(\mathbf{D}_{\theta(t)} = D_i) = Pr(\theta(t) = i)$. The matrices D_i are stochastic matrices (row sum up to one) with positive diagonal entries and correspond to the graphs G_i, for $i = 1\ldots s$.

The continuous-time dynamics is modeled by the following linear stochastic equation

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

where $\mathbf{C}_{\theta(t)}$ is a random matrix taking values in the finite set $\mathscr{C} = \{C_1,\ldots,C_s\}$, with probability distribution $Pr(\mathbf{C}_{\theta(t)} = C_i) = Pr(\theta(t) = i)$. The matrices C_i are rate transition like matrices (row sum up to zero) and correspond to the graphs G_i, for $i = 1\ldots s$. The initial state $x(0) = x_0$, for both continuous and discrete models, is assumed deterministic. The underlying probability space (for both models) is denoted by $(\Omega,\mathscr{F},P)$ and the solution process $x(t,x_0,\omega)$ (or simply, $x(t)$) of (1) or (2) is a random process defined on $(\Omega,\mathscr{F},P)$. 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 consult [2] for example.

Assumption 2.1: Throughout this paper we assume that the matrices belonging to the sets $\mathscr{S}$ and $\mathscr{C}$ are doubly stochastic (row and column sum up to one or zero) and in the case of the set $\mathscr{D}$ have positive diagonal entries. We assume also that the Markov chain $\theta(t)$ is irreducible.

We can use for instance a Laplacian based schemes to construct the matrices in the aforementioned sets in the case of (possible weighted) undirected or balanced (for every node, the inner degree is equal to the outer degree) communication graphs. If L_i denotes the Laplacian of the graph G_i, we can choose $A_i = I - \varepsilon L_i$ and $C_i = -L_i$, where $\varepsilon > 0$ is chosen such that A_i is stochastic. The above assumption will ensure reaching average consensus, desirable in important distributed computing applications such as distributed estimation [19], distributed optimization [14] , etc. Any other scheme can be used as long as it produces matrices with the properties stated above and it reflects the communication structures among agents.

Definition 2.1: We say that $x(t)$ converges to average consensus

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

$$\lim_{t\to\infty} E[\|x(t) - av(x_0)\mathbb{1}\|^2] = 0.$$

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

$$P(\lim_{t\to\infty} \|x(t) - av(x_0)\mathbb{1}\|) = 1.$$

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

The above theorem formulates an intuitively obvious condition for reaching consensus under the 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 for Section IV the proof of this theorem and provide here an intuitive and non-rigorous interpretation. Since $\theta(t)$ is irreducible, with probability one, all states in $\mathscr{S}$ are visited infinitely many times. But since the graphs in $\mathscr{G}$ are jointly, strongly connected, communication paths between all agents are formed infinitely many times, which allows for consensus to be achieved. Conversely, if the graphs in $\mathscr{G}$ are not strongly connected, , then (as we will see later) under Assumption 2.1 it does not have a spanning tree either. Therefore, there exists 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 can not be reached. Our main contribution in this note is to prove Theorem 2.1 using an approach based on the Markovian jump linear system stability theory in conjunction with a set of results we derive based on matrix and graph theory.

III. Preliminary results

In this section we introduce a set of results the proof of Theorem 2.1 is based on. We start with a number of general results, whose proofs can be found in the Appendix. Then, we will continue with results characteristic to the discrete and continuous time cases.

A. General preliminary results

Theorem 3.1: ([25]) Let s be a positive integer and let $\{A_i\}_{i=1}^{s}$ be a finite set of $n \times n$ ergodic matrices. Consider a map $r : \mathbb{N} \to \{1, \ldots, s\}$ such that for any finite sequence $\{r(i)\}_{i=1}^{j}$, the matrix product $\prod_{i=1}^{j} A_{r(i)}$ is ergodic. Then, there exists a vector c with non-negative entries (summing up to one), such that:

$$\lim_{j\to\infty}\prod_{i=1}^{j} A_{r(i)} = \mathbb{1}c'. \tag{3}$$

From the above Theorem we can immediately obtain the following corollary.

Corollary 3.1: Under the assumptions of *Theorem 3.1*, if the matrices in the set $\{A_i\}_{i=1}^{s}$ are doubly stochastic, then

$$\lim_{j\to\infty}\prod_{i=1}^{j} A_{r(i)} = \frac{1}{n}\mathbb{1}\mathbb{1}'. \tag{4}$$

Lemma 3.1: [8] 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, $i = 1,\dots,m$.

The following Corollary is an immediate consequence of Corollary 3.5 of [16].

Corollary 3.2: A rate transition matrix G has algebraic multiplicity equal to one for its eigenvalue $\lambda = 0$ if and only if the graph associated with the matrix has a spanning tree.

Remark 3.1: The homogeneous finite state Markov chain corresponding to a doubly stochastic transition matrix P can not have transient states. Indeed, since P is doubly stochastic, the same is true for P^t, for all $t \geq 1$. Assuming that there exist a transient state i, then $\lim_{t\to\infty}(P^t)_{ji} = 0$ for all j. But this means that there exist some t^* for which $\sum_j (P^{t^*})_{ji} < 1$ which contradicts the fact that P^{t^*} must be doubly stochastic. This means that we can relabel de vertices of the Markov chain such that P is block diagonalizable.

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

Lemma 3.2: Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ be two block diagonalizable rate transition matrices. Then

$$Null(A+B) = Null(A) \cap Null(B).$$

Corollary 3.3: Let s be a positive integer and let $\mathscr{A} = \{A_i\}_{i=1}^{s}$ be a set of doubly stochastic (probability transition) matrices. Then,

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

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

The following Corollary is the counterpart of Lemma 3.7 of [16], in the case of rate transition matrices.

Corollary 3.4: Let $G \in \mathbb{R}^{n\times n}$ be a rate transition matrix. If G has an eigenvalue $\lambda = 0$ with algebraic multiplicity equal to one, then $\lim_{t\to\infty} e^{Gt} = \mathbb{1}v'$, where v is a nonnegative vector satisfying $G'v = 0$ and $v'\mathbb{1} = 1$.

Proposition 3.1: Consider a matrix $Q \in \mathbb{R}^{n\times n}$ such that $\|Q\|_1 \leq 1$ and a set of matrices $\mathscr{S} = \{S_1, \dots S_m\}$, for some positive integer $m \leq n$. Assume that there exist a subsequence $\{t_k\} \subset \mathbb{Z}^+$ such that $\mathscr{S}$ is a limit set for the sequence $\{Q^{t_k}\}_{k\geq 0}$ and that for any $S \in \mathscr{S}$, $QS \in \mathscr{S}$, as well. Then, $\mathscr{S}$ is a limit set for the sequence $\{Q^k\}_{k\geq 0}$, i.e.

$$\lim_{k\to\infty} d(Q^k, \mathscr{S}) = 0,$$

where $d(Q, \mathscr{S}) = \min_{S\in\mathscr{S}} \|Q - S\|_1$.

B. *Preliminary results for the discrete-time case*

Lemma 3.3: 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. Let P be a $s \times s$ stochastic matrix corresponding to an irreducible, homogeneous Markov chain and consider the $ns \times ns$ dimensional matrix Q whose $(i, j)^{th}$ block is defined by $Q_{ij} \triangleq p_{ji}A_{ij}$. Then $\mathscr{P}'_\infty \otimes \left(\frac{1}{n}\mathbb{1}\mathbb{1}'\right)$ is the limit set for the matrix sequence $\{Q^k\}_{k\geq 1}$, i.e.:

$$\lim_{k\to\infty} d\left(Q^k, \mathscr{P}'_\infty \otimes \left(\frac{1}{n}\mathbb{1}\mathbb{1}'\right)\right) = 0. \quad (5)$$

Proof: The proof of this lemma is based on *Corollary 3.1*. The $(i, j)^{th}$ block entry of the matrix Q^k can be expressed as follows:

$$(Q^k)_{ij} = \sum_{1\leq i_1,\dots i_{k-1}\leq s} p_{ji_1}p_{i_1i_2}\dots p_{i_{k-1}i}A_{ii_1}A_{i_1i_2}\dots A_{i_{k-1}j}. \quad (6)$$

Let p^∞_{ji} be the (j, i) entry of some matrix in $\mathscr{P}_\infty$, i.e. there exist a sequence $\{t_k\}_{k\geq 1} \subset \mathbb{N}$ such that $\lim_{k\to\infty}(P^{t_k})_{ji} = p^\infty_{ji}$.

We have that

$$\|(Q^k)_{ij} - p_{ji}^{\infty}\frac{1}{n}\mathbb{1}\mathbb{1}'\| \leq \sum_{1\leq i_{t_1},\dots i_{k-1}\leq s}(p_{ji_1}\dots p_{i_{k-1}i})\|A_{ii_1}\dots A_{i_{k-1}j} - \frac{1}{n}\mathbb{1}\mathbb{1}'\| +$$
$$+ \sum_{1\leq i_1,\dots i_{k-1}\leq s}(p_{ji_1}\dots p_{i_{k-1}i} - p_{ji}^{\infty}) \leq$$
$$\leq \max_{i_1,\dots i_{k-1}}\{\|A_{ii_1}\dots A_{i_{k-1}j} - \frac{1}{n}\mathbb{1}\mathbb{1}'\|\}\sum_{1\leq i_1,\dots i_{k-1}\leq s}(p_{ji_1}\dots p_{i_{k-1}i}) +$$
$$+ \sum_{1\leq i_1,\dots i_{k-1}\leq s}(p_{ji_1}\dots p_{i_{k-1}i} - p_{ji}^{\infty}),$$

where $\|\cdot\|$ was used to denote some matrix norm. Consider the limit of the above left hand side for the sequence $\{t_k\}_{k\geq 1}$. By Corollary 3.1 we know that

$$\lim_{k\to\infty} A_{ii_{t_1}}\dots A_{i_{t_{k-1}}j} = \frac{1}{n}\mathbb{1}\mathbb{1}'$$

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

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

it results

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

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

Lemma 3.4: Let s be a positive integer and consider a set of doubly stochastic matrices with positive diagonal entries, $\mathscr{D} = \{D_i\}_{i=1}^{s}$, such that the corresponding graph $G_{\mathscr{D}}$ is strongly connected. Consider also P the $s\times s$ probability transition matrix of an irreducible, homogeneous Markov chain and consider the $ns\times ns$ matrix Q whose blocks are given by $Q_{ij} \triangleq p_{ji}W_j$. Then $\mathscr{P}'_{\infty}\otimes\left(\frac{1}{n}\mathbb{1}\mathbb{1}'\right)$ is the limit set of the sequence of matrices $\{Q^k\}_{k\geq 1}$, i.e.:

$$\lim_{k\to\infty} d\left(Q^k, \mathscr{P}'_{\infty}\otimes\left(\frac{1}{n}\mathbb{1}\mathbb{1}'\right)\right) = 0. \tag{7}$$

Proof: Our strategy consists in showing that there exist a $k\in\mathbb{N}$, such that each $(i,j)^{\text{th}}$ 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 3.3* to obtain (7). The $(i,j)^{\text{th}}$ block matrix of Q^k looks as in (6), with the difference that in the current case $A_{ij} = D_j$:

$$(Q^k)_{ij} = \sum_{1 \le i_1, \ldots i_{k-1} \le s} p_{ji_1} p_{i_1 i_2} \ldots p_{i_{k-1} i} D_j D_{i_1} \ldots D_{i_{k-1}} = p_{ji}^{(k)} A_{ij}^{(k)} \tag{8}$$

where

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

with

$$\alpha_{i_1,\ldots i_{k-1}} \triangleq \begin{cases} p_{ji_1} p_{i_1 i_2} \ldots p_{i_{k-1} i} / p_{ji}^{(k)}, & p_{ji}^{(k)} > 0 \\ 0, & \text{otherwise} \end{cases}$$

Notice that each of the matrix product $D_j D_{i_1} \ldots D_{i_{k-1}}$ 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 $\mathscr{D}$ appears at least once in one of the terms of the sum (8), i.e. $\{1,...,s\} \subseteq \{i_1, \ldots i_{k-1}\}$. Using a similar idea as in Lemma 1 in [8] or Lemma 3.9 in [16], by Lemma 3.1, we upper bound such term

$$D_j D_{i_1} \ldots D_{i_{k-1}} \ge \gamma \sum_{l=1}^{s} D_l = \gamma s \bar{D}, \tag{9}$$

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

$$\bar{D} = \frac{1}{s} \sum_{i}^{s} D_i.$$

Since $G_{\mathscr{D}}$ 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 (9), it follows that the matrix product $D_j D_{i_1} \ldots D_{i_{k-1}}$ is ergodic. This is enough to infer that $A_{ij}^{(k)}$ is ergodic as well, since 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 non-zero $p_{ij}^{(k^*)}$, the matrices $A_{ij}^{(k^*)}$ are ergodic $\forall i,j$. Such k^* always exists due to irreducibility assumption on P. Then according to Lemma 3.3, we have that for the subsequence $\{t_m\}_{m \ge 0}$, with $t_m = mk^*$

$$\lim_{m \to \infty} d\left(Q^{t_m}, \mathscr{P}'_\infty \otimes \left(\frac{1}{n} \mathbb{1}\mathbb{1}'\right)\right) = 0. \tag{10}$$

The results follows by Proposition 3.1 since $\|Q\|_1 \le 1$ and since $Q\left(\mathscr{P}'_\infty \otimes \left(\frac{1}{n}\mathbb{1}\mathbb{1}'\right)\right) = \mathscr{P}'_\infty \otimes \left(\frac{1}{n}\mathbb{1}\mathbb{1}'\right)$.

■

Corollary 3.5: Under the same assumptions as in Lemma 3.4, if we define the matrix blocks of Q as $Q_{ij} \triangleq p_{ji} D_j \otimes D_j$, then $\mathscr{P}'_\infty \otimes \left(\frac{1}{n^2}\mathbb{1}\mathbb{1}'\right)$ is the limit set of the sequence $\{Q^k\}_{k\geq 1}$, i.e.

$$\lim_{k\to\infty} d\left(Q^k, \mathscr{P}'_\infty \otimes \left(\frac{1}{n^2}\mathbb{1}\mathbb{1}'\right)\right),$$

where the vector $\mathbb{1}$ above has dimension n^2.

Proof: In the current setup (8) becomes:

$$(Q^k)_{ij} = \sum_{1\leq i_1,\dots i_{k-1}\leq s} p_{ji_1} p_{i_1 i_2} \dots p_{i_{k-1} i} (D_j \otimes D_j)(D_{i_1} \otimes D_{i_1}) \dots (D_{i_{k-1}} \otimes D_{i_{k-1}}). \tag{11}$$

The result follows from the same arguments used in Lemma 3.4 together with the fact that the matrix products in (11) 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 a ergodic matrix with itself produces an ergodic matrix as well. ■

C. Preliminary results for the continuous time case

The following two corollaries emphasize geometric properties of two matrices arising from the linear dynamics of the first and second moment of the state vector.

Corollary 3.6: Let s be a positive integer and let $\mathscr{C} = \{C_1, \dots, C_s\}$ be a set of $n \times n$ doubly stochastic matrices such that $G_{\mathscr{C}}$ is strongly connected. Consider also a $s \times s$ rate transition matrix $\Lambda = (\lambda_{ij})$ corresponding to an irreducible Markov chain with stationary distribution $\pi = (\pi_i)$. Define the matrices $A \triangleq diag(C'_i, i = 1\dots s)$ and $B \triangleq \Lambda \otimes I$. Then $A + B$ has an eigenvalue $\lambda = 0$ with algebraic multiplicity one, with corresponding right and left eigenvectors given by $\mathbb{1}_{ns}$ and $(\pi_1 \mathbb{1}'_n, \pi_2 \mathbb{1}'_n, \dots, \pi_s \mathbb{1}'_n)$, respectively.

Proof: We first note that $A + B$ is a rate transition 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 null space of $A + B$. By Lemma 3.2, we have that $v \in Null(A)$ and $v \in Null(B)$. Given the structure of B, v must respect the following pattern $v' = \{(\underbrace{u' \ u' \dots u'}_{s\ times}) \mid u \in \mathbb{R}^n\}$. But since $u \in Null(A)$, we have that $C'_i u = 0$, $i = 1\dots s$, or $\mathbf{C}u = 0$, where $\mathbf{C} = \sum_{i=1}^s C'_i$. But since $G_{\mathscr{C}}$ was assumed strongly connected, $\mathbf{C}$ corresponds to an irreducible Markov chain, and it follows that u must be of the form $u = \alpha\mathbb{1}$, for some $\alpha \in \mathbb{R}$. By backtracking, we get that $v = \alpha\mathbb{1}$, for some $\alpha \in \mathbb{R}$ and consequently $Null(A+B) = span(\mathbb{1})$. Therefore, $\lambda = 0$ has algebraic multiplicity one, with right eigenvector given by $\mathbb{1}$. By simple verification we note that $(\pi_1\mathbb{1}', \pi_2\mathbb{1}', \ldots, \pi_s\mathbb{1}')$ is a left eigenvector corresponding to the eigenvalue $\lambda = 0$. ■

Corollary 3.7: Let s be a positive integer and let $\mathscr{C} = \{C_1, \ldots, C_s\}$ be a set of $n \times n$ doubly stochastic matrices such that $G_{\mathscr{C}}$ is strongly connected. Consider also a $s \times s$ rate transition matrix $\Lambda = (\lambda_{ij})$ corresponding to an irreducible Markov chain with stationary distribution $\pi = (\pi_i)$. Define the matrices $A \triangleq diag(C_i \oplus C_i', i = 1 \ldots s)$ 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 $\mathbb{1}_{n^2 s}$ and $(\pi_1\mathbb{1}_{n^2}', \pi_2\mathbb{1}_{n^2}', \ldots, \pi_s\mathbb{1}_{n^2}')$, respectively.

Proof: It is not difficult to check that $A+B$ is a rate transition 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 then it is block diagonalizable. The matrix $C_i' \otimes I$ contains n isolated copies of C_i' and therefore it is block diagonalizable. Also, $I \otimes C_i'$ it has a number of n block on its diagonal, each block being given by C_i', and it follows is block diagonalizable as well.

Let v be a vector in the nullspace of $A+B$. By Lemma 3.2, $v \in Null(A)$ and $v \in Null(B)$. From the structure of B we note that v must be of the form $v' = (\underbrace{u', \ldots, u'}_{s\ times})' \mid u \in \mathbb{R}^{n^2}$. Consequently we have that $(C_i' \oplus C_i')u = 0,\ i = 1, \ldots s$, or $(\mathbf{C} \otimes \mathbf{C})u = 0$, where $\mathbf{C} = \sum_{i=1}^{s} C_i'$. Since, $G_{\mathscr{G}}$ is strongly connected, $\mathbf{C}$ is a rate transition matrix corresponding to an irreducible Markov chain. By applying again Lemma 3.2 for the matrix $\mathbf{C} \oplus \mathbf{C} = I \otimes \mathbf{C} + \mathbf{C} \otimes I$, we get that u must have the form $u' = (\underbrace{\bar{u}', \ldots, \bar{u}'}_{n\ times})'$, where $\bar{u} \in \mathbb{R}^n$ and $\mathbf{C}\bar{u} = 0$. But $\mathbf{C}$ is irreducible and therefore $\bar{u} = \alpha\mathbb{1}_n$, or $u = \alpha\mathbb{1}_{n^2}$, or finally $v = \alpha\mathbb{1}_{n^2 s}$, where $\alpha \in \mathbb{R}$. Consequently, $Null(A+B) = span(\mathbb{1})$ which means the eigenvalue $\lambda = 0$ has algebraic multiplicity one. By simple verification, we note that $(\pi_1\mathbb{1}_{n^2}', \pi_2\mathbb{1}_{n^2}', \ldots, \pi_s\mathbb{1}_{n^2}')$ is a left eigenvector corresponding to the zero eigenvalue. ■

IV. Proof of the convergence theorem

The proof will focus on showing that the state vector $x(t)$ converges in mean square sense to average consensus. Or, by making the change of variable $z(t) = x(t) - av(x_0)\mathbb{1}$, we will actually show that $z(t)$ is mean square stable for the initial condition $z(0) = x_0 - av(x_0)\mathbb{1}$, where $z(t)$

respects the same dynamic equation as $x(t)$. Using results for the stability theory of Markovian jump linear systems, mean square stability also imply stability in the almost sure sense (see for instance Corollary 3.46 of [2] for discrete-time case or Theorem 2.1 of [5] for continuous-time case, with the remark that we are interested for the stability property to be satisfied for a specific initial condition, rather then for any initial condition), which for us imply that $x(t)$ converges almost surely to average consensus.

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

A. Discrete-time case - Sufficiency

Proof:

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

$$V(t) = \sum_{i=1}^{s} V_i(t), \tag{12}$$

where $V_i(t)$ is given by

$$V_i(t) \triangleq E[x(t)x(t)^T \chi_{\{\theta(t)=i\}}] \ i = 1\ldots s, \tag{13}$$

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 $V_i(t)$ are given by

$$V_i(t+1) = \sum_{j=1}^{s} p_{ji} D_j V_j(t) D_j^T, \ i = 1\ldots s, \tag{14}$$

with initial conditions $V_i(0) = q_i x_0 x_0^T$. By defining $\eta(t) \triangleq col(V_i(t), i = 1\ldots s)$, we obtain a vectorized form of equations (14)

$$\eta(t+1) = \Gamma_d \eta(t), \tag{15}$$

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

$$\Gamma_d = \begin{pmatrix} p_{11}D_1 \otimes D_1 & \dots & p_{s1}D_s \otimes D_s \\ \vdots & \ddots & \vdots \\ p_{1s}D_1 \otimes D_1 & \dots & p_{ss}D_s \otimes D_s \end{pmatrix} \text{ and } \eta_0 = \begin{pmatrix} q_1 col(x_0 x_0') \\ \vdots \\ q_s col(x_0 x_0') \end{pmatrix}. \quad (16)$$

We notice that Γ_d satisfies all the assumptions of *Corollary 3.5* and hence we get

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

Using the observation that

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

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

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

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

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

and from (12) we get

$$\lim_{k\to\infty} V(t_k) = av(x_0)^2 \mathbb{1}\mathbb{1}' \quad (17)$$

since $\sum_{i,j=1}^s p_{ji}^\infty q_j = 1$. By repeating the previous steps for all subsequences generating limit points of $\{P^t\}_{t\geq 0}$ we obtain that (17) holds for any sequence in $\mathbb{N}$.

Through a similar process as in the case of the case of the second moment (in stead of Corollary 3.5 we use Lemma 3.4), we show that:

$$\lim_{k\to\infty} E[x(t)] = av(x_0)\mathbb{1}. \quad (18)$$

From (17) and (18) we have that

$$\lim_{t\to\infty} E[\|x(t) - av(x_0)\mathbb{1}\|^2] = \lim_{t\to\infty} trace(E[(x(t) - av(x_0)\mathbb{1})(x(t) - av(x_0)\mathbb{1})']) =$$

$$= \lim_{t\to\infty} trace(E[x(t)x(t)'] - av(x_0)\mathbb{1}E[x(t)'] - av(x_0)E[x(t)]\mathbb{1}' + av(x_0)^2\mathbb{1}\mathbb{1}') = 0.$$

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

■

B. Discrete-time case - Necessity

Proof: If $G_{\mathscr{A}}$ is not strongly connected then by Corollary 3.3, $dim(\bigcap_{i=1}^{s} Null(A_i - I)) > 1$. Consequently, there exist a vector $v \in \bigcap_{i=1}^{s} Null(A_i - I))$ such that $v \notin span(\mathbb{1})$. If we choose v as initial condition, for every realization of $\theta(t)$, we have that

$$x(t) = v, \text{ for all } t \geq 0,$$

and therefore consensus can not be reached in the sense of Definition 2.1. ■

C. Sufficiency - Continuous time

Using the same notations as in the discrete-time case, the dynamic equations describing the evolution of the second moment of $x(t)$ are given by

$$\frac{d}{dt}V_i(t) = C_i V_i(t) + V_i(t) C_i' + \sum_{j=1}^{s} \lambda_{ji} V_j(t), \; i = 1 \dots s, \tag{19}$$

equations whose derivation is treated in [6]. By defining the vector $\eta(t) \triangleq col(V_i(t), i = 1 \dots s)$, the vectorized equivalent of equations (19) is given by

$$\frac{d}{dt}\eta(t) = \Gamma_c \eta(t), \tag{20}$$

where

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

By Corollary 3.6, the eigenspace corresponding to the zero eigenvalue 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 rate transition matrix with an eigenvalue zero of algebraic multiplicity one, by Corollary 3.4 we have that $\lim_{t\to\infty} e^{\Gamma_c' t} = v\mathbb{1}'$, 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

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

By noting that

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

we farther get

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

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

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

or

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

Through a similar process (using this time Corollary 3.6), we can show that

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

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

D. Necessity - Continuous time

Follows the same lines as in the discrete-time case.

V. DISCUSSION

In the previous sections we proved a convergence result for the discrete and continuous time linear, stochastic consensus problem. Our main contributions consist of considering a Markovian process, not necessarily ergodic, as underlying process for the random communication graph and of the Markovian Jump theory inspired approach we took for proving 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 [10] (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 to return 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 cancelled. We can also assume that D_i and C_i are not necessarily doubly connected. We treated this case (again only for the discrete-time dynamics and without being completely rigorous) in [xxx] and we showed that the state converges in mean square sense and in almost sure sense to consensus, and not necessarily average consensus. From the technical point view, the difference consist of the fact that the $n^2 \times n^2$ block matrices of Γ_d (or Γ_c) no longer converge to $\pi_i \frac{1}{n^2}\mathbb{1}\mathbb{1}'$ but to $\pi_i \mathbb{1}c'$, for some vector $c \in \mathbb{R}^{n^2}$ with non-negative entries summing up to one, vector c which in general can no be a priori determined. In relevant distributed computation application (such as distributed state estimation or distributed optimization) however, convergence to average consensus is required, and therefore the assumption, that D_i or C_i are doubly stochastic, makes sense.

The proof of Theorem 2.1 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 the continuous and discrete time, respectively. The reader may have noted that we approached differently the analysis of the two sequences. 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 3.4. 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, similar 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, the stochastic matrices are assumed to have positive diagonal entries, to avoid the possibility of being periodic).

In the case of i.i.d. random graphs [21], or more general, in the case of strictly stationary, ergodic random graphs [22] , 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 a similar condition, does not necessarily 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, are 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 $t \geq 0$. Also note that $\lambda_2(\lim_{t\to\infty} E[\mathbf{D}_{\theta(t)}])$ does not exist since the sequence $\{E[\mathbf{D}_{\theta(t)}]\}_{t\geq 0}$ does not have a limit. Yet, consensus is reached. The assumption that allowed for the aforementioned necessary and sufficient condition to hold, was that $\theta(t)$ is a stationary process (which in turn implies that $E[D_{\theta(t)}]$ 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 value of the time average expectation of $\mathbf{D}_{\theta(t)}$, i.e. $\lim_{N\to\infty} \frac{\sum_{t=1}^{N} E[\mathbf{D}_{\theta(t)}]}{N}$, which reflects the proportion of time $\mathbf{D}_{\theta(t)}$ spends in each state of the Markov chain.

Corollary 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

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

Proof:

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

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

where $\pi = (\pi_i)$ is the stationary distribution of $\theta(t)$. By Corollary 3.5 in [16], $|\lambda_2(\bar{D})| < 1$ 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 $G_{\mathscr{D}}$, and the result follows from Theorem 2.1.

■

Unlike the discrete-time, 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 its similar to the proof of Corollary 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.

Corollary 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

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

Our analysis provides also 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 dictates how fast the second moment converges to average consensus. Since Γ_d' is a probability transition matrix and since Γ_c is a rate transition 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.

VI. Conclusion

In this note we treated the continuous and discrete time stochastic consensus problem. 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. Under additional assumptions on the directed communication topologies, we provided a rigorous mathematical proof for the intuitive necessary and sufficient conditions for reaching average consensus in the mean square and almost surely sense. These conditions are expressed in terms of connectivity properties of the union of graphs corresponding to the states of the Markov chain. The aim of this note has been 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. Varaya, "Asymptotic agreement in distributed estimation", *IEEE Trans. Autom. Control*, vol. AC-27, no. 3, pp. 650-655, Jun 1982.

[2] O.L.V. Costa, M.D. Fragoso and R.P. Marques, *Discrete-Time Markov Jump Linear Systems*, Springer-Verlag, London 2005.

[3] J.A. Fax, "Optimal and cooperative control of vehicles formations", Ph.D. dissertation, Control Dynamical Syst., California Inst. Technol., Pasadena, CA, 2001.

[4] J.A. Fax and R.M. Murray, "Information flow and cooperative control of vehicles formations", *IEEE Trans. Automatic Control*, vol. 49, no. 9, pp. 1456-1476, Sept. 2004.

[5] X. Feng and K.A. Loparo, ”Stability of linear Markovian systems,“ *Proceedings of the* 29^{th} *Conference on Decision and Control*, pp. 1408-1413, Honolulu, Hawaii, December 1990.

[6] 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,“ *SIAM Journal on Control and Optimization*, vol. 44, no. 4, pp. 1165-1191.

[7] G. Grimmet and D. Stirzaker, *Probability and Random Processes*, Third Edition, Oxford University Press, 2001.

[8] A. Jadbabaie, J. Lin and A.S. Morse, ”Coordination of groups of mobile autonomous agents using nearest neighbor”, *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 998-1001, Jun 2003.

[9] Y. Hatano and M. Mesbahi, ”Agreement over Random Networks”, *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 18671872, 2005.

[10] I. Matei, N. Martins and J. Baras, ”Almost sure convergence to consensus in Markovian random graphs”, *the* 47^{th} *Conference on Decision and Control*, to be presented,2008.

[11] L. Moreau, ”Stability of multi-agents systems with time-dependent communication links”, *IEEE Trans. Automatic Control*, vol. 50, no. 2, pp. 169-182, Feb. 2005.

[12] A. Nedic, A. Ozdaglar, and P.A. Parrilo, ”Constrained Consensus and Optimization in Multi-Agent Networks,”, to appear in IEEE Transactions on Automatic Control, 2009.

[13] A. Nedic and A. Ozdaglar, ”Convergence Rate for Consensus with Delays,” to appear in Journal of Global Optimization, 2008.

[14] A. Nedic and A. Ozdaglar, Distributed Subgradient Methods for Multi-agent Optimization, *IEEE Transactions on Automatic Control* Volume 54, Issue 1, Jan. 2009 Page(s):48 - 61.

[15] M. Porfiri and D.J. Stilwell, ”Consensus seeking over Random Directed Weighted Graphs”, *IEEE Trans. Autom. Control*, vol. 52, no. 9, Sept. 2007.

[16] W. Ren and R.W. Beard, ”Consensus seeking in multi-agents systems under dynamically changing interaction topologies”, IEEE Trans. Autom. Control, vol. 50, no. 5, pp. 655-661, May 2005.

[17] R.O. Saber and R.M. Murray, ”Consensus protocols for networks of dynamic agents”, in *Proc. 2003 Am. Control Conf.*, 2003, pp. 951-956.

[18] R.O. Saber and R.M. Murray, ”Consensus problem in networks of agents with switching topology and time-delays”, *IEEE Trans. Automatic Control*, vol. 49, no. 9, pp. 1520-1533, Sept. 2004.

[19] R.O. Saber, ”Distributed Kalman Filter with Embedded Consensus Filters,”, *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005.

[20] R.O. Saber, J.A. Fax and R.M. Murray, ”Consensus and Cooperation in Networked Multi-Agent Systems”, *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215-233, Jan. 2007.

[21] A. Tahbaz Salehi and A. Jadbabaie, ”Necessary and Sufficient Conditions for Consensus over random networks”, *IEEE Transactions on Automatic Control*, accepted, August 2007.

[22] A. Tahbaz Salehi and A. Jadbabaie, ”Consensus Over Ergodic Stationary Graph Processes,“ *IEEE Transactions on Automatic Control*, to appear

[23] J.N. Tsitsiklis, ”Problems in decentralized decision making and computation”, Ph.D. dissertation, Dept. Electr. Eng., Massachusetts Inst. Technol., Cambridge, MA, Nov. 1984.

[24] J.N. Tsitsiklis, D.P. Bertsekas and M. Athans, ”Distributed asynchronous deterministic and stochastic gradient optimization algorithms”, *IEEE Trans. Automatic Control*, vol. 31, no. 9, pp. 803-812, Sept. 1986.

[25] J. Wolfowitz, "Product of indecomposable aperiodic stochastic matrices", in *Proc. Am. Math. Soc.*, 1963, vol. 15, pp. 733-736.

Appendix

A. Proof of Corollary 3.1

By *Theorem 3.1* we have that

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

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

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

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

B. Proof of Corollary 3.2

Follows immediately from Corollary 3.5 of [16], by forming the probability transition matrix $P = I + \varepsilon G$, for some appropriate $\varepsilon > 0$, and noting that $Null(P-I) = Null(G)$.

C. Proof of Lemma 3.2

Obviously, $v \in Null(A)\cap Null(B)$ implies $v \in Null(A+B)$. Since A is block diagonalizable, then there exists a similarity transformation T such that $\bar{A} = TAT'$ is a block diagonal rate transition matrix (with irreducible blocks). Let $\bar{A}_i \in \mathbb{R}^{n_i\times n_i}$, $i = 1\ldots m$ denote the irreducible blocks on the main diagonal of $\bar{A}$, where m is the number of such blocks and $\sum_{i=1}^{m} n_i = n$. The nullspace of $\bar{A}$ can be expressed as

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

We assumed that B is block diagonalizable, which mean that G_B is a union of isolated, strongly connected subgraphs, property which remains valid for the graph corresponding to $\bar{B}$. By adding $\bar{B}$ to $\bar{A}$ two phenomena can happen: we can either 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 $Null(\bar{A}+\bar{B}) = Null(\bar{A})$. 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 $\bar{A}_i$, 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 a similar expression to the nullspace of $\bar{A}$ with the main difference 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

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

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, $Null(\bar{A}+\bar{B}) \subset Null(\bar{A})$, which means also that $Null(A+B) \subset Null(A)$. Therefore, if $(A+B)v = 0$, then $Av = 0$ which implies also that $Bv = 0$ or $v \in Null(B)$. Hence if $v \in Null(A+B)$ then $v \in Null(A) \cap Null(B)$, which concludes the proof.

D. *Proof of Corollary 3.3*

Since A_i, $i = 1 \ldots s$ are doubly stochastic then $A_i - I$ are block diagonalizable doubly stochastic rate transition matrices. Therefore, by recursively applying Lemma 3.2 $s-1$ times, the first part of the Corollary follows. For the second part of the Corollary, note that, by Corollary 3.5 of [16], $\frac{1}{N}\sum_{i=1}^{s} A_i$ has the algebraic multiplicity equal to one, of its eigenvalue $\lambda = 1$ if and only if the graph associated to $\frac{1}{N}\sum_{i=1}^{s} A_i$ has a spanning tree, or in our case is strongly connected, which in turn implies that $dim(Null(\sum_{i=1}^{s}(A_i - I))) = 1$ if and only if $G_{\mathscr{A}}$ is strongly connected.

E. *Proof of Corollary 3.4*

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

$$\lim_{k\to\infty} e^{Gt_k^1} = e^{h_1 kG} = P_{h_1}^k,$$

where we defined $P_{h_1} \triangleq e^{h_1 G}$. From the theory of continuous-time Markov chain 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 $x'G = 0$. This means that the algebraic multiplicity of the eigenvalue $\lambda = 1$ of P_{h_1} is one. By Lemma 3.7 of [16], we have that $\lim_{k\to\infty} P_{h_1}^k = \mathbb{1} v_{h_1}'$, where v_{h_1} is a nonnegative vector satisfying $P_{h_1}' v_{h_1} = v_{h_1}$ and $v_{h_1}' \mathbb{1} = 1$. Also $G' v_{h_1} = 0$. Choose another $h_2 > 0$ and let $P_{h_2} \triangleq e^{h_2 G}$. Similarly as above, we have that

$$\lim_{k\to\infty} P_{h_2}^k = \mathbb{1} v_{h_2}',$$

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

$$\lim_{k\to\infty} e^{Gt_k} = e^{hkG} = \mathbb{1} v',$$

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

F. Proof of Proposition 3.1

Pick a subsequence $\{t_k'\}_{k\geq 0}$ given by $t_k' = t_k + \delta_k$, where $\delta_k \in \mathbb{Z}^+$. It follows that

$$d(Q^{t_k'}, \mathscr{S}) = \min_{S\in\mathscr{S}} \|Q^{\delta_k} Q^{t_k}, Q^{\delta^k} S\|_1 \leq \|Q^{\delta_k}\|_1 \min_{S\in\mathscr{S}} \|Q^{t_k}, S\|_1 \leq d(Q^{t_k}, \mathscr{S}).$$

Therefore, we get $\mathscr{S}$ 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.