Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

Convergence results for the agreement problem on Markovian random topologies

Ion Matei * **John S. Baras** *

* *Department of Electrical and Computer Engineering and Institute for System Research, University of Maryland, College Park, USA (e-mail: imatei@ umd.edu, baras@umd.edu).*

Abstract: We study the linear distributed asymptotic agreement(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 both the cases where the dynamics of the agents is expressed in continuous and discrete time. We show that, if the consensus matrices are doubly stochastic, convergence to average consensus is achieved in the mean square and almost sure sense, if and only if the graph resulted 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.

Keywords: distributed, agreement, random topologies, Markovian jump linear systems.

1. 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 Borkar and Varaya (1982), Tsitsikils, Bertsekas and Athans Tsitsiklis (1984); Tsitsiklis et al. (1986) on asynchronous agreement problems and parallel computing. A theoretical framework for solving consensus problems was introduced by Olfati-Saber and Murray in Saber and Murray (2003, 2004), while Jadbabaie et al. studied alignment problems Jadbabaie et al. (2004) for reaching an agreement. Relevant extensions of the consensus problem were done by Ren and Beard Ren and Beard (2005), by Moreau in Moreau (2005) or, more recently, by Nedic and Ozdaglar in Nedic and Ozdaglar (2010); Nedic et al. (2010).

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 Hatano and Mesbahi (2005) 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 Porfiri and Stilwell (2007), 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 consensus in Salehi and Jadbabaie (2008), while in Salehi and Jadbabaie (2010), the authors 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 Kar and Moura (2008).

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) 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 both the cases where the dynamics of the agents is expressed in *continuous and discrete time* and, under certain assumptions on the consensus matrices, we give necessary and sufficient conditions to guarantee convergence to average consensus in mean square and in almost sure sense. The Markovian switching model goes *beyond the common 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.

⋆ This material is based upon work supported by the US Air Force Office of Scientific Research MURI award FA9550-09-1-0538, by the Defence Advanced Research Projects Agency (DARPA) under award number 013641-001 for the Multi-Scale Systems Center (MuSyC) through the FRCP of SRC and DARPA, and by BAE Systems award number W911NF-08-2-0004 (the ARL MAST CTA).

1.1 Basic notations and definitions

We denote by $\mathbb{1}$ the vector of all ones. If the dimension of the vector needs to be emphasized, an 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 $\mathscr{X}$ is some finite dimensional space, $dim(\mathscr{X})$ gives us the dimension of $\mathscr{X}$. We denote by $col(A)$ a vector containing the columns of matrix A, by $\otimes$ the Kronecker product and by $\oplus$ the Kronecker sum.

Let $\mathscr{M}$ be a set of matrices and let A be some matrix. 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 $\mathscr{M} \otimes A$ we understand the following matrix set: $\mathscr{M} \otimes A = \{M \otimes A \mid M \in \mathscr{M}\}$. 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 $\mathbf{P}_\infty$ 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 $\lim_{k\to\infty} P^{t_k} = L$. Note that if the matrix P corresponds to an ergodic Markov chain, the cardinality of $\mathbf{P}_\infty$ 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 $\mathbf{P}_\infty$ is m. Let $d(M,\mathbf{P}_\infty)$ denote the distance from M to the set $\mathbf{P}_\infty$, that is the smallest distance from M to any matrix in $\mathbf{P}_\infty$:

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

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. 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 non-zero. The graph corresponding to A will be denoted by G_A.

Definition 1.2. Let s be a positive integer and let $\mathscr{A} = \{A_i\}_{i=1}^s$ be a set of matrices with a corresponding set of graphs $\mathscr{G} = \{G_{A_i}\}_{i=1}^s$. 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 paper we will use mainly to type of matrices: *probability transition matrices* (row sum up to one) and *generator* matrices (row sum up to zero). A generator matrix whose both rows and columns sum up to zero will be called *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 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 constitute the core of the proof of our main results; proof which is given in Section IV. Section V contains a discussion of our convergence results.

2. PROBLEM FORMULATION AND STATEMENT OF THE CONVERGENCE RESULTS

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 $\mathbf{G}(t) = (V,E(t))$, 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 $\{1,\ldots,s\}$, for some positive integer s. In the case $\theta(t)$ 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 $\mathscr{G} = \{G_i\}_{i=1}^s$ 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, in the case $\theta(t)$ is a discrete-time Markov chain, we model the dynamics of the agents by the following linear stochastic difference equation

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

where $\mathbf{D}_{\theta(t)}$ is a random matrix taking values in the finite set $\mathscr{D} = \{D_i\}_{i=1}^s$, 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 $i = 1\ldots s$.

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

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

where $\mathbf{C}_{\theta(t)}$ is a random matrix taking values in the finite set $\mathscr{C} = \{C_i\}_{i=1}^s$, 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) 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. We will sometimes refer to the matrices belonging to the sets $\mathscr{D}$ and $\mathscr{C}$ as *consensus matrices*. The underlying probability space (for both models) is denoted by $(\Omega,\mathscr{F},\mathscr{P})$ and the solution process $x(t,x_0,\omega)$ (or simply, $x(t)$) of (1) or (2) is a random process defined on $(\Omega,\mathscr{F},\mathscr{P})$. 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 refer to Costa et al. (2005) for example.

Assumption 2.1. Throughout this paper we assume that the matrices belonging to the sets $\mathscr{D}$ and $\mathscr{C}$ are *doubly stochastic* (rows and columns sum up to one and zero, respectively) and in the case of the set $\mathscr{D}$ have *positive diagonal entries*. We assume also 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 possible weighted. 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.

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

Assumption 2.1 will guarantee reaching *average consensus*, desirable in important distributed computing applications such as distributed estimation Saber (2005) or distributed optimization Nedic and Ozdalgar (2009). 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.

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 results for the linear consensus problem under Markovian random communication topology.

Theorem 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 are visited infinitely many times. But since the graph $G_{\mathscr{D}}$ (or $G_{\mathscr{C}}$) is strongly connected, communication paths between all agents are formed infinitely many times, which allows for consensus to be achieved. Conversely, if the graph $G_{\mathscr{D}}$ (or $G_{\mathscr{C}}$) is not strongly connected, then 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 is to prove Theorem 1 using an approach based on the stability theory of Markovian jump linear systems, in conjunction with a set of results based 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, respectively. The proof of Theorem 1 is mainly based on four lemmas (Lemmas 3.1 and 3.2 for discrete-time case and Lemmas 3.3 and 3.4 for 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 preliminary results are stated *without proofs*; proofs which can be founded in our 2009 ISR Technical Report which can be accessed through the link provided in the reference list (Matei and Baras (2009)).

3.1 General preliminary results

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

Corollary 3.1. 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 Ren and Beard (2005), in the case of generator matrices.

Corollary 3.2. 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 $v'\mathbb{1} = 1$.

3.2 Preliminary results - discrete-time dynamics

In this subsection we state a set of results used to prove Theorem 1 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 moment of the state vector $x(t)$.

Lemma 3.1. 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. Let P be the $s\times s$ dimensional probability transition matrix of an irreducible, homogeneous Markov chain and let $\mathbf{P}_\infty$ 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 $\mathbf{P}'_\infty \otimes \left(\frac{1}{n}\mathbb{1}\mathbb{1}'\right)$ is the limit set of the sequence of matrices $\{Q^k\}_{k\geq 1}$, i.e.:
$$\lim_{k\to\infty} d\left(Q^k, \mathbf{P}'_\infty \otimes \left(\frac{1}{n}\mathbb{1}\mathbb{1}'\right)\right) = 0. \quad (3)$$

Lemma 3.2. Under the same assumptions as in Lemma 3.1, if we define the matrix blocks of Q as $Q_{ij} \triangleq p_{ji}D_j \otimes D_j$, then $\mathbf{P}'_\infty \otimes \left(\frac{1}{n^2}\mathbb{1}\mathbb{1}'\right)$ is the limit set of the sequence $\{Q^k\}_{k\geq 1}$, i.e.
$$\lim_{k\to\infty} d\left(Q^k, \mathbf{P}'_\infty \otimes \left(\frac{1}{n^2}\mathbb{1}\mathbb{1}'\right)\right),$$
where the vector $\mathbb{1}$ above has dimension n^2.

3.3 Preliminary results - continuous-time dynamics

The following two lemmas emphasize geometric properties of two matrices arising from the linear dynamics of the first and second moment of the state vector, in the continuous-time case.

Lemma 3.3. Let s be a positive integer and let $\mathscr{C} = \{C_i\}_{i=1}^s$ be a set of $n \times n$ doubly stochastic matrices such that $G_{\mathscr{C}}$ is strongly connected. Consider also a $s \times s$ generator 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 and 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.

Lemma 3.4. Let s be a positive integer and let $\mathscr{C} = \{C_i\}_{i=1}^s$ be a set of $n \times n$ doubly stochastic matrices such that $G_{\mathscr{C}}$ is strongly connected. Consider also a $s \times s$ generator matrix $\Lambda = (\lambda_{ij})$ corresponding to an irreducible Markov chain with stationary distribution $\pi = (\pi_i)$. Define the matrices $A \triangleq diag(C_i' \oplus C_i', i = 1 \dots s)$ 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 $\mathbb{1}_{n^2 s}$ and $(\pi_1 \mathbb{1}_{n^2}', \pi_2 \mathbb{1}_{n^2}', \dots, \pi_s \mathbb{1}_{n^2}')$, respectively.

4. 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. Equivalently, 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 from 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 Costa et al. (2005) for discrete-time case or Theorem 2.1 of Feng and Loparo (1990) 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.

4.1 Sufficiency - discrete-time case

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

$$V(t) = \sum_{i=1}^{s} V_i(t), \tag{4}$$

where $V_i(t)$ is given by

$$V_i(t) \triangleq E[x(t)x(t)^T \chi_{\{\theta(t)=i\}}] \; i = 1 \dots s, \tag{5}$$

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

$$V_i(t+1) = \sum_{j=1}^{s} p_{ji} D_j V_j(t) D_j^T, \; i = 1 \dots s, \tag{6}$$

with initial conditions $V_i(0) = q_i x_0 x_0^T$. By defining $\eta(t) \triangleq col(V_i(t), i = 1 \dots s)$, we obtain a vectorized form of equations (6)

$$\eta(t+1) = \Gamma_d \eta(t), \tag{7}$$

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

$$\Gamma_d = \begin{pmatrix} p_{11} D_1 \otimes D_1 & \dots & p_{s1} D_s \otimes D_s \\ \vdots & \ddots & \vdots \\ p_{1s} D_1 \otimes D_1 & \dots & p_{ss} D_s \otimes D_s \end{pmatrix} \tag{8}$$

and $\eta_0' = (q_1 col(x_0 x_0')', \dots, q_s col(x_0 x_0')')'$.

We note that Γ_d satisfies all the assumptions of *Lemma 3.2* and hence we get

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

where $\mathbf{P}_\infty$ is the limit set of the matrix sequence $\{P^k\}_{k \geq 0}$. 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 (4) we get

$$\lim_{k \to \infty} V(t_k) = av(x_0)^2 \mathbb{1}\mathbb{1}' \tag{9}$$

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 (9) holds for any sequence in $\mathbb{N}$.

Through a similar process as in the case of the second moment (in stead of Lemma 3.2 we use Lemma 3.1), we show that:

$$\lim_{k \to \infty} E[x(t)] = av(x_0)\mathbb{1}. \tag{10}$$

From (9) and (10) 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.

4.2 Necessity - discrete-time case

Proof. If $G_{\mathscr{A}}$ is not strongly connected then by Corollary 3.1, $dim(\bigcap_{i=1}^s Null(A_i - I)) > 1$. 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.

4.3 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

$$\frac{d}{dt}V_i(t) = C_iV_i(t) + V_i(t)C_i' + \sum_{j=1}^{s} \lambda_{ji}V_j(t),\ i = 1\dots s, \quad (11)$$

equations whose derivation is treated in Fragoso and Costa (2005). By defining the vector $\eta(t) \triangleq col(V_i(t), i = 1\dots s)$, the vectorized equivalent of equations (11) is given by

$$\frac{d}{dt}\eta(t) = \Gamma_c\eta(t), \quad (12)$$

where

$$\Gamma_c = \begin{pmatrix} C_1 \oplus C_1 & 0 & \cdots & 0 \\ 0 & C_2 \oplus C_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_s \oplus C_s \end{pmatrix} + \Lambda' \otimes I$$

and $\eta_0' = (q_1 col(x_0x_0')', \dots, q_s col(x_0x_0')')'$.
By Lemma 3.4, 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 $\mathbb{1}_{n^2s}$ 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 3.2 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

$$\lim_{t\to\infty} \eta(t) = \begin{pmatrix} \pi_1\frac{\mathbb{1}\mathbb{1}'}{n^2} & \cdots & \pi_1\frac{\mathbb{1}\mathbb{1}'}{n^2} \\ \vdots & \ddots & \vdots \\ \pi_s\frac{\mathbb{1}\mathbb{1}'}{n^2} & \cdots & \pi_s\frac{\mathbb{1}\mathbb{1}'}{n^2} \end{pmatrix} \begin{pmatrix} q_1 col(x_0x_0') \\ \vdots \\ q_s col(x_0x_0') \end{pmatrix}.$$

By noting that

$$\frac{\mathbb{1}\mathbb{1}'}{n^2} col(x_0x_0') = av(x_0)^2\mathbb{1}_{n^2},$$

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 Lemma 3.3), 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.

4.4 Necessity proof - continuous time

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

5. DISCUSSION

The proof of Theorem 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 3.2. 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 Salehi and Jadbabaie (2008), or more general, in the case of strictly stationary, ergodic random graphs Salehi and Jadbabaie (2010) , 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 eigenvalue 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.

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

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

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 Ren and Beard (2005), $|\lambda_2(\bar{D})| < 1$ 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 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 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

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

Our analysis provides also estimates on the rate of convergence to average consensus in the mean square sense. From the 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 both the cases where the dynamics of the agents is expressed in continuous and 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 and 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

Borkar, V. and Varaya, P. (1982). Asymptotic agreement in distributed estimation. *IEEE Trans. Autom. Control*, AC-27(3), 650–655.

Costa, O., Fragoso, M., and Marques, R. (2005). *Discrete-Time Markov Jump Linear Systems*. Springer-Verlag, London.

Feng, X. and Loparo, K. (1990). Stability of linear markovian systems. *Proceedings of the 29th IEEE Conference on Decision and Control*, 1408–1413.

Fragoso, M. and Costa, O. (2005). 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*, 44(4), 1165–1191.

Hatano, Y. and Mesbahi, M. (2005). Agreement over random networks. *IEEE Trans. Autom. Control*, 50(11), 1867–1872.

Jadbabaie, A., Lin, J., and Morse, A. (2004). Coordination of groups of mobile autonomous agents using nearest neighbor. *IEEE Trans. Autom. Control*, 48(6), 998–1001.

Kar, S. and Moura, J. (2008). Sensor networks with random links; topology design for dostributed consensus. *IEEE Trans. Signal Process.*, 56(7), 3315–3326.

Matei, I. and Baras, J. (2009). Convergence results for the linear consensus problem under markovian random graphs. *ISR Technical Report, http://drum.lib.umd.edu/bitstream /1903/9693/5/IMateiJBaras.pdf*, (TR-2009-18).

Moreau, L. (2005). Stability of multi-agents systems with time-dependent communication links. *IEEE Trans. Autom. Control*, 50(2), 169–182.

Nedic, A. and Ozdaglar, A. (2010). Convergence rate for consensus with delays. *Journal of Global Optimization*, 47(3), 437–456.

Nedic, A., Ozdaglar, A., and Parrilo, P. (2010). Constrained consensus and optimization in multi-agent networks. *IEEE Trans. Autom. Control*, 55(4), 922–938.

Nedic, A. and Ozdalgar, A. (2009). Distributed subgradient methods for multi-agent optimization. *IEEE Trans. Autom. Control*, 54(1), 48–61.

Porfiri, M. and Stilwell, D. (2007). Consensus seeking over random directed weighted graphs. *IEEE Trans. Autom. Control*, 52(9), 1767–1773.

Ren, W. and Beard, R. (2005). Consensus seeking in multi-agents systems under dynamically changing interaction topologies. *IEEE Trans. Autom. Control*, 50(5), 655–661.

Saber, R. (2005). Distributed kalman filter with embedded consensus filters. *Proceedings of the 44th IEEE Conference on Decision and Control*.

Saber, R. and Murray, R. (2003). Consensus protocols for networks of dynamic agents. *Proceedings of the 2003 IEEE American Control Conference*, 951–956.

Saber, R. and Murray, R. (2004). Consensus problem in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control*, 49(9), 1520–1533.

Salehi, A.T. and Jadbabaie, A. (2008). Necessary and sufficient conditions for consensus over random networks. *IEEE Trans. Autom. Control*, 53(3), 791–795.

Salehi, A.T. and Jadbabaie, A. (2010). Consensus over ergodic stationary graph processes. *IEEE Trans. Autom. Control*, 55(1), 225–230.

Tsitsiklis, J. (1984). Problems in decentralized decision making and computation. *Ph.D. dissertation, Dept. Electr. Eng.*

Tsitsiklis, J., Bertsekas, D., and Athans, M. (1986). Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Trans. Autom. Control*, 31(9), 803–812.