Necessary and Sufficient Conditions for Consensus Over Random Independent and Identically Distributed Switching Graphs

Alireza Tahbaz-Salehi and Ali Jadbabaie

Abstract— **In this paper we consider the consensus problem for stochastic switched linear dynamical systems. For discrete-time and continuous-time stochastic switched linear systems, we present necessary and sufficient conditions under which such systems reach a consensus almost surely. In the discrete-time case, our assumption is that the underlying graph of the system at any given time instance is derived from a random graph process, independent of other time instances. These graphs can be weighted, directed and with dependent edges. For the continuous-time case, we assume that the switching is governed by a Poisson point process and the graphs characterizing the topology of the system are independent and identically distributed over time. For both such frameworks, we present necessary and sufficient conditions for almost sure asymptotic consensus using simple ergodicity and probabilistic arguments. These easily verifiable conditions depend on the spectrum of the average weight matrix and the average Laplacian matrix for the discrete-time and continuous-time cases, respectively.**

I. Introduction

Decentralized iterative schemes such as agreement and consensus problems have an old history [1]–[3]. Over the past few years they have attracted a significant amount of attention in various contexts, such as motion coordination of autonomous agents [4], distributed computation of averages and least squares among sensors [5]–[7], and rendezvous problems [8]. In all these cases the dynamical system under study is deterministic. More recently, there has been some interest in the stochastic variants of the problem. In [9], the authors study the linear dynamical system $x(k) = W_k x(k-1)$, where the weight matrices W_k are i.i.d. stochastic matrices. It is shown that all the entries of $x(k)$ converge to a common value almost surely, if each edge of $\mathbf{G}(W_k)$, the graph corresponding to matrix W_k, is chosen independently with the same probability (Erdős-Rényi random graph model). A more general model appeared in [10], where the edges of $\mathbf{G}(W_k)$ are directed and not necessarily independent. However, the author proves only convergence to a consensus in probability, rather than the more general notion of almost sure convergence. Moreover, the assumption in [10] is the occurrence of scrambling matrices with positive probability, which can be weakened.

The purpose of this paper is to provide necessary and sufficient conditions for almost sure consensus in stochastic linear dynamical systems, both discrete-time and continuous-time. In the discrete-time case we assume that the weight matrices are general i.i.d. stochastic matrices. Our results contain the results of [9] and [10] as special cases. This necessary and sufficient condition is easily verifiable and only depends on the spectrum of the average weight matrix $\mathbb{E}W_k$. More precisely, we show that such a stochastic system reaches a consensus with probability one if and only if the expected weighted graph of the system has a directed spanning tree. We state a similar result for continuous-time stochastic linear systems with switching governed by a Poisson process and i.i.d. graphs. We show that for such a system, the set *span*$\{\mathbf{1}\}$ is globally asymptotically attractive almost surely if and only if the expected Laplacian matrix of the underlying graphs has exactly one zero eigenvalue.

Even though our results can be derived by combining results from ergodic theory of Markov chains in random environments [11], [12], our proofs are only based on simple linear algebra machinery and the concept of *coefficients of ergodicity*, as introduced by Dobrushin [13].

The paper is organized as follows: The problem setup for the discrete-time stochastic switched linear systems is presented in section II. Weakly ergodic sequences of matrices and coefficients of ergodicity are defined in section III. Section IV contains the necessary and sufficient condition for almost sure consensus of discrete-time switched linear system. Continuous-time stochastic linear systems with Poisson switchings are presented in section V followed by the necessary and sufficient condition for their asymptotic consensus. Finally, section VI contains our conclusions.

II. Discrete-Time Stochastic Switched Linear Systems

Let $(\Omega_0, \mathcal{B}, \mu)$ be a probability space, where $\Omega_0 = \mathrm{S}_n = \{$set of stochastic matrices of order n with strictly positive diagonal entries$\}$, $\mathcal{B}$ is the Borel σ-algebra of Ω_0, and μ is a probability measure defined on Ω_0. Define the product probability space as $(\Omega, \mathcal{F}, \mathbb{P}) = \prod_{k=1}^{\infty}(\Omega_0, \mathcal{B}, \mu)$. By definition, the elements of the product space have the following forms:

$$\begin{aligned}
&\Omega = \{(\omega_1, \omega_2, \cdots) : \omega_k \in \Omega_0\} \\
&\mathcal{F} = \mathbb{B} \times \mathbb{B} \times \cdots \\
&\mathbb{P} = \mu \times \mu \times \cdots
\end{aligned}$$

which means that the coordinates of the infinite dimensional vector $\omega \in \Omega$ are independent and identically distributed (i.i.d.) stochastic matrices with positive diagonals.

Now consider the following random discrete-time dynamical system:

$$x(k) = W_k(\omega)x(k-1), \tag{1}$$

This research is supported in parts by the following grants: ARO/MURI W911NF-05-1-0381, ONR/YIP N00014-04-1-0467 and NSF-ECS-0347285.

Alireza Tahbaz-Salehi and Ali Jadbabaie are with GRASP Laboratory and Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104-6228, USA. {atahbaz,jadbabai}@seas.upenn.edu

where $k \in \{1, 2, \cdots\}$ is the discrete time index, $x(k) \in \mathbb{R}^n$ is the state vector at time k and the mapping $W_k : \Omega \to S_n$ is the kth coordinate function, which for all $\omega = (\omega_1, \omega_2, \cdots) \in \Omega$ is defined as

$$W_k(\omega) = \omega_k.$$

As a result, (1) defines a stochastic linear dynamical system, in which the weight matrices are drawn independently from the common distribution μ. For notational simplicity, we denote $W_k(\omega)$ by W_k through the rest of the note.

For a general weight matrix W, one can define the corresponding graph $\mathbf{G}(W)$ as a weighted directed graph with an edge (i, j) from vertex i to vertex j with weight W_{ji} if and only if $W_{ji} \neq 0$. In this case we say vertex j has *access* to vertex i and we say vertices i and j *communicate* if both (i, j) and (j, i) are edges of $\mathbf{G}(W)$. Note that the communication relation is an equivalence relation and defines equivalence classes on the set of vertices. If no vertex in a specific communication class has access to any vertex outside that class, such a class is called *initial*.

For the given dynamical system, we now define the notions of reaching state consensus in probability and almost surely.

Definition 1: Dynamical system (1) reaches consensus in probability, if for any initial state value $x(0)$ and any $\epsilon > 0$,

$$\mathbb{P}(|x_i(k) - x_j(k)| > \epsilon) \to 0$$

as $k \to \infty$ for all $i, j = 1, \cdots, n$.
This notion of reaching state agreement asymptotically, which is addressed in [10], is a special case of reaching consensus almost surely, defined below.

Definition 2: Dynamical system (1) reaches consensus almost surely, if for any initial state value $x(0)$,

$$|x_i(k) - x_j(k)| \to 0 \quad \textit{almost surely}$$

as $k \to \infty$ for all $i, j = 1, \cdots, n$.
This stronger notion of consensus, not only requires that the probability of the event $\{|x_i(k) - x_j(k)| > \epsilon\}$ goes to zero for an arbitrary $\epsilon > 0$ as time goes by, but also guarantees that such events occur only finitely many times [14].

III. Ergodicity

Given (1), if $x(0)$ is the initial state value, one can write the state vector at time k as

$$x(k) = W_k \cdots W_2 W_1 x(0). \tag{2}$$

As it is evident from (2), one needs to investigate the behavior of infinite products of stochastic matrices in order to check for an asymptotic consensus. This motivates us to borrow the concept of *weak ergodicity* of a sequence of stochastic matrices from the theory of Markov chains.

Definition 3: The sequence $\{W_k\}_{k=1}^{\infty} = W_1, W_2, \cdots$ of $n \times n$ stochastic matrices is weakly ergodic, if for all $i, j, s = 1, \cdots, n$ and all p,

$$\left(U_{i,s}^{(k,p)} - U_{j,s}^{(k,p)}\right) \to 0$$

as $k \to \infty$, where $U^{(k,p)} = W_{k+p} \cdots W_{2+p} W_{1+p}$ is the left product of the matrices in the sequence.

As the definition suggests, if a sequence of stochastic matrices is weakly ergodic, then the rows of the product matrix converge to each other, as the number of terms in the product grows. In other words, the weak ergodicity of the sequence of matrices is a subset of the event that the linear dynamical system (1) reaches a consensus asymptotically.

A closely related concept is *strong ergodicity* of a matrix sequence.

Definition 4: A sequence of $n \times n$ stochastic matrices $\{W_k\}_{k=1}^{\infty}$ is strongly ergodic, if for all $i, s = 1, \cdots, n$ and all p

$$U_{i,s}^{(k,p)} \to d_s^p$$

as $k \to \infty$, where $U^{(k,p)}$ is the left product and d_s^p is a constant not depending on i.

One can easily see that weak and strong ergodicity[1] both describe a tendency to consensus. If either type of ergodicity (weak or strong) holds for the matrix sequence $\{W_k\}_{k=1}^{\infty} = W_1, W_2, \cdots$, the pairwise differences between rows of the product matrix $U^{(k,0)}$ converge to zero, and therefore system (1) reaches a consensus asymptotically.

At the first glance, it may seem that there exists a difference between weak and strong ergodicity. In the case of weak ergodicity, every two entries of vector $x(k)$ converge to each other, but each entry does not necessarily converge to some limit. On the other hand, in the presence of strong ergodicity, not only the difference between any two entries converges to zero, but also all of them enjoy a common limit. Although one may consider the difference to be an important one, that is not the case here because of the following theorem [2], [15]:

Theorem 1: Given a matrix sequence $\{W_k\}_{k=1}^{\infty}$ and their left products $U^{(k,p)} = W_{k+p} \cdots W_{1+p}$, weak and strong ergodicity are equivalent.

Proof: We only need to prove that weak ergodicity implies strong ergodicity. For any $\epsilon > 0$, weak ergodicity implies that if k is large enough, $-\epsilon \le U_{i,s}^{(k,p)} - U_{j,s}^{(k,p)} \le \epsilon$ uniformly for all i, j, s. Since $U^{(k+1,p)} = W_{k+p+1} U^{(k,p)}$, we have

$$U_{i,s}^{(k,p)} - \epsilon \le U_{h,s}^{(k+1,p)} \le U_{i,s}^{(k,p)} + \epsilon,$$

which by induction implies that

$$U_{i,s}^{(k,p)} - \epsilon \le U_{h,s}^{(k+r,p)} \le U_{i,s}^{(k,p)} + \epsilon,$$

for all $i, s, h = 1, \cdots, n$ and $r \ge 0$. By setting $i = h$, it is evident that $U_{i,s}^{(k,p)}$ is a Cauchy sequence and therefore, $\lim_{k\to\infty} U_{i,s}^{(k,p)}$ exists. ∎

Therefore, weak ergodicity is equivalent to the existence of a vector d satisfying $U^{(k,p)} \to \mathbf{1} d^T$, in which $\mathbf{1}$ is a vector with all entries equal to one. We now define the concept of the *coefficients of ergodicity* which are of central importance in proving weak ergodicity results.

[1]To be more precise, we have stated the definitions of weak and strong ergodicity *in the backward direction*. But since this is the only type of ergodicity that we are concerned with, we simply refer to these properties as ergodicity.

Definition 5: The scalar continuous function $\tau(\cdot)$ defined on the set of $n \times n$ stochastic matrices is called a coefficient of ergodicity if it satisfies $0 \leq \tau(\cdot) \leq 1$. A coefficient of ergodicity is said to be *proper* if

$$\tau(W) = 0 \qquad \text{if and only if} \qquad W = \mathbf{1}d^T,$$

where d is a vector of size n satisfying $d^T\mathbf{1} = 1$.

Two examples of coefficients of ergodicity that we use in this paper are

$$\kappa(W) = \frac{1}{2}\max_{i,j}\sum_{s=1}^{n}|W_{is} - W_{js}|,$$
$$\nu(W) = 1 - \max_{j}(\min_{i} W_{ij}).$$

Note that $\nu(\cdot)$ is an improper coefficient of ergodicity while $\kappa(\cdot)$ is proper, and for any stochastic matrix W they satisfy

$$\kappa(W) \leq \nu(W). \tag{3}$$

Given the above definition, it is straightforward to show that weak ergodicity is equivalent to

$$\tau(U^{(k,p)}) \to 0 \qquad \forall p$$

as $k \to \infty$ for a proper coefficient of ergodicity τ. Therefore, we can state the following theorem:

Theorem 2: Suppose $\tau_1(\cdot)$ and $\tau_2(\cdot)$ are proper coefficients of ergodicity that for any $m \geq 1$ stochastic matrices W_k, $k = 1, 2, \cdots, m$ satisfy

$$\tau_1(W_m \cdots W_2 W_1) \leq \prod_{k=1}^{m} \tau_2(W_k). \tag{4}$$

Then the sequence $\{W_k\}_{k=1}^{\infty}$ is weakly ergodic if and only if there exists a strictly increasing sequence of integers k_r, $r = 1, 2, \cdots$ such that

$$\sum_{r=1}^{\infty} \left(1 - \tau_2(W_{k_{r+1}} \cdots W_{k_r+1})\right) = \infty. \tag{5}$$

Proof: Since only the sufficiency part of the above theorem will be used in this note, we only prove that (5) implies weak ergodicity of the sequence. A proof for the reverse implication can be found in [15, theorem 4.18].

Suppose that there exists a strictly increasing sequence of positive integers k_r such that (5) holds. Then, the inequality $\log x \leq x - 1$ implies that

$$\sum_{r=1}^{\infty} \log\left(\tau_2(W_{k_{r+1}} \cdots W_{k_r+1})\right) = -\infty,$$

and as a result, $\prod_{r=1}^{\infty} \tau_2(W_{k_{r+1}} \cdots W_{k_r+1}) = 0$. Because we assumed that τ_1 is a proper coefficient, (4) guarantees weak ergodicity of the sequence. ∎

IV. Necessary and Sufficient Condition for Almost Sure Ergodicity

In this section we study the necessary and sufficient conditions for ergodicity of an i.i.d. sequence of stochastic matrices based on the framework presented in section II.

Lemma 1: The weak ergodicity of the sequence $W_1, W_2, \cdots$ is a trivial event.

Proof: Let k be a positive integer. Define the event

$$A_k = \{\text{The sequence } W_k, W_{k+1}, \cdots \text{ is weakly ergodic.}\}$$

which is an event in $\mathcal{F}'_k = \sigma(W_k, W_{k+1}, \cdots)$. These events form a *decreasing sequence* of events, satisfying $A_1 \supseteq A_2 \supseteq \cdots$. As a result,

$$\bigcap_{k=1}^{\infty} A_k \in \mathcal{T},$$

where $\mathcal{T}$ is the tail σ-field of the sequence of stochastic matrices defined as $\mathcal{T} = \bigcap_{k=1}^{\infty} \mathcal{F}'_k$. Therefore, by Kolmogorov's 0-1 law [14], $\mathbb{P}(\cap_k A_k) = 0$ or 1. Now we have,

$$\mathbb{P}(\bigcap_{k=1}^{\infty} A_k) = \lim_{j\to\infty} \mathbb{P}(\bigcap_{k=1}^{j} A_k) = \lim_{j\to\infty} \mathbb{P}(A_j) = \mathbb{P}(A_1),$$

where the first equality is because of continuity of the probability measure and the last one holds because the distribution of W_k does not depend on k. Therefore, weak ergodicity of $W_1, W_2, \cdots$ is trivial. ∎

This lemma indicates that weak ergodicity of random i.i.d. weight matrices obeys a 0-1 law. In other words, the sequence of matrices is weakly ergodic either almost surely or almost never, indicating a discontinuous behavior. In order to find a criterion to distinguish between these two cases, we need another lemma, the proof of which can be found in [16].

Lemma 2: Suppose that W is a stochastic matrix for which its corresponding graph has s communication classes named $\alpha_1, \cdots, \alpha_s$. Class α_r is initial, if and only if the spectral radius of $\alpha_r[W]$ equals to one, where $\alpha_r[W]$ is the submatrix of W corresponding to the vertices in the class α_r.

Finally, suppose that the average weight matrix $\mathbb{E}W_k$ has n eigenvalues satisfying

$$0 \leq |\lambda_n(\mathbb{E}W_k)| \leq \cdots \leq |\lambda_2(\mathbb{E}W_k)| \leq |\lambda_1(\mathbb{E}W_k)| = 1.$$

At this point we state our main theorem.

Theorem 3: The random sequence $\{W_k\}_{k=1}^{\infty} = W_1, W_2, \cdots$ of stochastic matrices with positive diagonals is (weakly) ergodic almost surely if and only if $|\lambda_2(\mathbb{E}W_k)| < 1$ and in the case that $|\lambda_2(\mathbb{E}W_k)| = 1$, the sequence is weakly ergodic almost never. Moreover, in both cases, the probability of (1) reaching a consensus asymptotically is equal to the probability of weak ergodicity of $\{W_k\}_{k=1}^{\infty}$.

Proof: To show the necessity of the given condition, we first assume that $|\lambda_2(\mathbb{E}W_k)| = 1$. Since all $\omega_k \in \Omega_0$ have positive diagonals, $\mathbb{E}W_k$ has strictly positive diagonal entries as well. Hence, if $\mathbb{E}W_k$ is irreducible, then it is primitive and as a result of the Perron-Frobenius theorem [16], $|\lambda_2(\mathbb{E}W_k)| < 1$ which is in contradiction with our assumption. Therefore, $|\lambda_2(\mathbb{E}W_k)| = 1$ implies reducibility of $\mathbb{E}W_k$. As a result, without loss of generality, one can label the vertices such that $\mathbb{E}W_k$ gets the following block triangular form

$$\mathbb{E}W_k = \begin{bmatrix} Q_{11} & 0 & \cdots & 0 \\ Q_{21} & Q_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ Q_{s1} & Q_{s2} & \cdots & Q_{ss} \end{bmatrix}, \qquad (6)$$

where each Q_{ii} is an irreducible matrix and represents the vertices in the equivalence class α_i. Since $|\lambda_2(\mathbb{E}W_k)| = 1$, submatrices corresponding to at least two of the classes have unit spectral radii (note that because of irreducibility and aperiodicity of Q_{ii}'s, the multiplicity of the unit-modulus eigenvalue of each one of them cannot be more than one). Hence, lemma 2 implies,

$$\exists i \neq j \text{ s.t. } \alpha_i \text{ and } \alpha_j \text{ are both initial classes,}$$

or equivalently, $Q_{ir} = 0$ for all $r \neq i$ and $Q_{jl} = 0$ for all $l \neq j$. In other words, the matrix $\mathbb{E}W_k$ has two orthogonal rows. Since Ω_0 is a subset of nonnegative matrices, W_k has the same type (zero block pattern) as $\mathbb{E}W_k$ for all time k with probability one. Therefore, $U^{(k,0)} = W_k \cdots W_2 W_1$ has two orthogonal rows almost surely for any k which means that the random discrete-time dynamical system (1) reaches a consensus with probability zero. Since weak ergodicity of $\{W_k\}_{k=1}^{\infty}$ is a subset of convergence of (1) to a consensus, the random sequence of weight matrices is weakly ergodic almost never.

The reverse implication can be proved by combining results from the ergodic theory of Markov chains in random environments, more specifically by using [11], [12]. However, here we provide a simpler proof based on theorem 2. In order to do so, we assume that $|\lambda_2(\mathbb{E}W_k)|$ is subunit. Lemma 2 implies that $\mathbf{G}(\mathbb{E}W_k)$ has exactly one initial class. We investigate the two cases of $\mathbb{E}W_k$ being irreducible and reducible separately.

1) Irreducibility: Suppose $\mathbb{E}W_k$ is irreducible. Since it has only one unit-modulus eigenvalue, $\mathbb{E}W_k$ is primitive [16]. Hence,

$$\exists m \text{ s.t. } [\mathbb{E}W_k]^m > 0,$$

where by > 0 for a matrix, we mean the entry-wise positivity of that matrix. Independence over time implies

$$\mathbb{E}(W_m \cdots W_1) = [\mathbb{E}W_k]^m > 0.$$

As a result, for all $i, j = 1, \cdots, n$ the (i, j) entry of $U^{(m,0)} = W_m \cdots W_2 W_1$ is positive with nonzero probability, say $p_{ij} > 0$. Therefore, since the weight matrices are i.i.d. with positive diagonals, the matrix $W_{n^2 m} \cdots W_2 W_1$ is completely entry-wise positive with at least probability $\prod_{i,j} p_{ij} > 0$, i.e. the event $\{W_{n^2 m} \cdots W_2 W_1 > 0\}$ has nonzero probability. [2] As a result, if we define $\delta(W) = 1 - \nu(W) = \max_j(\min_i W_{ij})$, there exists $\epsilon > 0$ such that

$$\mathbb{P}(\delta(W_{n^2 m} \cdots W_2 W_1) > \epsilon) > 0.$$

[2] Selecting $n^2 m$ as the number of matrices in the product in order to make the product matrix entry-wise positive is a conservative pick. The product matrix becomes positive with nonzero probability with less terms, but such a pick suffices our purpose.

Hence, by the second Borel-Cantelli lemma [14, page 49], we have

$$\mathbb{P}(\delta(W_{(r+1)n^2 m} \cdots W_{rn^2 m+1}) > \epsilon \text{ for infinitly many } r) = 1.$$

Once we set $k_r = rn^2 m$, we have

$$\delta(W_{k_{r+1}} \cdots W_{k_r+1}) > \epsilon \quad \textit{i.o. a.s.}$$

Also note that (4) holds for $\tau_1(\cdot) = \tau_2(\cdot) = \kappa(\cdot)$. Therefore, this together with (3) results in

$$\sum_{r=1}^{\infty} \big(1 - \kappa(W_{k_{r+1}} \cdots W_{k_r+1})\big) = \infty \quad \textit{a.s.},$$

which is exactly (5), the sufficient condition for weak ergodicity. Therefore, the sequence is weakly ergodic almost surely.

2) Reducibility: When $\mathbb{E}W_k$ is reducible, without loss of generality, it can be written as (6), where all Q_{ii} are irreducible matrices. Since α_1 (the class corresponding to submatrix Q_{11}) is the only initial class of $\mathbf{G}(\mathbb{E}W_k)$, there exists a directed path from a vertex in α_1 (e.g. say vertex labeled 1) to any vertex of $\mathbf{G}(\mathbb{E}W_k)$, such that the length of the path is at most some positive integer m. In other words, any vertex of $\mathbf{G}(\mathbb{E}W_k)$ is at most an m-hop neighbor of vertex 1. This combined with the fact that $\mathbb{E}W_k$ has strictly positive diagonals guarantees that the first column of $[\mathbb{E}W_k]^m$ is strictly positive. Therefore, as in case 1, independence implies the positivity of the first column of $\mathbb{E}(W_m \cdots W_1)$.

As a result, for $j = 1, \cdots, n$ the $(j, 1)$-entry of the matrix $W_m \cdots W_1$ is nonzero with positive probability $\hat{p}_{j1}$. Hence, identical to the discussion of case 1, we have

$$\exists \epsilon > 0 \text{ s.t. } \mathbb{P}\left(\delta(W_{nm} \cdots W_2 W_1) > \epsilon\right) \geq \prod_{j=1}^{n} \hat{p}_{j1} > 0.$$

Now if we set $k_r = rnm$, once again the second Borel-Cantelli lemma guarantees that

$$\mathbb{P}\left(\delta(W_{k_{r+1}} \cdots W_{k_r+1}) > \epsilon \text{ for infinitly many } r\right) = 1.$$

Therefore, the sum $\sum_{r=1}^{\infty} \left(1 - \kappa(W_{k_{r+1}} \cdots W_{k_r+1})\right)$ diverges to infinity with probability one and theorem 2 implies that the random sequence $\{W_k\}_{k=1}^{\infty}$ is weakly ergodic almost surely. And moreover as a result, (1) reaches a consensus asymptotically with probability one. ■

Theorem 3 provides a simple criterion to distinguish between the two cases of asymptotically reaching consensus almost surely and almost never. It suggests that the information in the average weight matrix $\mathbb{E}W_k$ suffices to predict the long-run behavior of the linear dynamical system (1). This should not come as a surprise to the reader. Intuitively, when $|\lambda_2(\mathbb{E}W_k)|$ is subunit, there exists a sequence of integer numbers k_r, $r = 1, 2, \cdots$ such that the graph collection $\{\mathbf{G}(W_{k_r+1}), \cdots, \mathbf{G}(W_{k_{r+1}})\}$ is jointly connected (i.e. the graph constructed by unioning the edge sets of the graphs in the collection contains a directed spanning tree) [4]. This infinite often connectivity over time guarantees the possibility of information flow in the graph over time, and therefore results in asymptotic consensus with probability one. On the other hand, when $|\lambda_2(\mathbb{E}W_k)| = 1$, no such sequence exists and therefore, there are at least two classes of vertices in the graph such that they never have access to each other, and hence, no consensus.

Moreover, theorem 3 contains the results of [9] and [10] as special cases. Since in [9] the authors use Erdős-Rényi as their random graph model, the matrix $\mathbb{E}W_k$ is completely entry-wise positive, which results in $|\lambda_2(\mathbb{E}W_k)| < 1$, and hence almost sure consensus. On the other hand, when the weight matrices are scrambling with positive probability, as in [10], $\mathbb{E}W_k$ is also scrambling and as a result, its unit-modulus eigenvalue has multiplicity one. Hence, (1) reaches an asymptotic consensus almost surely (and therefore in probability).

V. Continuous-Time Stochastic Switched Linear Systems

In this section we investigate the continuous-time variant of stochastic linear dynamical systems and provide a necessary and sufficient condition for almost sure asymptotic consensus in such systems.

Consider the following switched linear dynamical system

$$\dot{x}(t) = -L_{t_k}x(t) \qquad t \in [t_k, t_{k+1}), \tag{7}$$

where the switching time sequence $\{t_k\}_{k=0}^{\infty}$ is a strictly increasing sequence of real numbers with $t_0 = 0$ and L_{t_k} is the weighted *Laplacian* of the graph which appears over the interval $[t_k, t_{k+1})$. The weighted Laplacian of the graph $\mathbf{G}$ corresponding to weight matrix $W \in \Omega_0$ is defined as

$$L = diag(W\mathbf{1}) - W.$$

In other words, $L_{ij} = -W_{ij}$ if $i \neq j$ and $L_{ii} = \sum_{j=1}^{n} W_{ij}$ for all $1 \leq i \leq n$. Note that by construction, L is a singular matrix with at least one zero eigenvalue corresponding to eigenvector $\mathbf{1}$.

We address the consensus problem for the linear dynamical system (7) under the following two assumptions.

Assumption 1: The switching time sequence $t_1, t_2, \cdots$ forms a Poisson point process with rate γ.

Assumption 2: Given the switching time signal $t_1, t_2, \cdots$ the graphs $\mathbf{G}_{t_k}$ with Laplacians L_{t_k} are independent and identically distributed.

Assumption 1 states that the intervals $[t_k, t_{k+1})$ over which the graphs remain unchanged are independent and identically distributed. More precisely, the length of each such interval has an exponential distribution with mean γ:

$$\mathbb{P}(t_{k+1} - t_k \leq T) = 1 - e^{-\gamma T}.$$

Note that we are not assuming any minimum dwell-time (or average dwell-time) on the switching times. On the other hand, assumption 2 guarantees that whenever a switching occurs, the emerging graph is drawn from the set of possible graphs independent of other graphs and its distribution remains the same over time. Assumption 2 is exactly the same i.i.d. assumption we made for the discrete-time case.

Given the above assumptions, we now present our consensus theorem.

Theorem 4: Under assumptions 1 and 2, the continuous-time linear dynamical system (7) converges to a consensus almost surely for all initial conditions if and only if

$$\lambda_{n-1}(\mathbb{E}L_{t_k}) \neq 0. \tag{8}$$

Moreover, in the case of $\lambda_{n-1}(\mathbb{E}L_{t_k}) = 0$, the system reaches consensus with probability zero.

This theorem states that similar to the discrete-time case, the connectivity of the expected graph is necessary and sufficient for almost sure convergence. Note that the edges of the underlying graph at a given time can be dependent. All we are assuming is independence between graphs over time and a Poisson switching signal.

Proof: First we suppose that $\lambda_{n-1}(\mathbb{E}L_{t_k}) = 0$ and show the necessity of (8) for an almost sure consensus. Write the expected Laplacian matrix as $\mathbb{E}L_{t_k} = I - \tilde{A}$. Clearly, $\tilde{A}$ is a non-negative matrix and all its diagonal entries are positive. The assumption $\lambda_{n-1}(\mathbb{E}L_{t_k}) = 0$ implies that $\lambda_2(\tilde{A}) = 1$. Based on an argument similar to proof of theorem 3, $\tilde{A}$ has a block triangular form as in (6) with at least two orthogonal rows. Therefore, the graph corresponding to $\mathbb{E}L_{t_k}$ has at least two initial communication classes. Since all the entries of a Laplacian are sign-definite, the matrix L_{t_k} is of the same type of $\mathbb{E}L_{t_k}$ for all t_k almost surely. Thus,

$$\exists a \notin span\{\mathbf{1}\} \text{ s.t. } \mathbb{P}\left(a \in \bigcap_{j \geq k} \ker L_{t_j}\right) = 1 \text{ for all } k$$

In other words, the vector $a \notin span\{\mathbf{1}\}$ is in the kernel of Laplacian matrices of all times almost surely. As a result, given the initial condition $x(0) = a$, the solution of the linear dynamical system (7) is $x(t) = a$ for all t almost surely, i.e. no consensus with probability one. This implies that (8) is a necessary condition for almost sure consensus and if it does not hold, there exists an initial condition for which the system reaches consensus with probability zero.

Note that so far we have not used assumption 1. In fact, we do not need to assume anything about the switching signal in order to prove necessity, as long as the switching is well-defined (i.e. there are finitely many switchings in any finite interval). In the case of a switching signal governed by a Poisson process, the probability of having d switchings in an interval of length T is $e^{-\gamma T}\frac{(\gamma T)^d}{d!}$. Therefore, the switching system experiences no *Zenoness* almost surely and the probability of having infinitely many switchings in a finite interval is zero[3].

To prove sufficiency, assume that (8) holds. Since there are only finite switchings in any finite interval almost surely, the solution of the dynamical system (7) can be written as $x(t) = \Phi(t, 0)x(0)$, where

$$\Phi(t, 0) = \left(e^{-(t-t_k)L_{t_k}}\right) \cdots \left(e^{-(t_1-t_0)L_{t_0}}\right) x(0)$$

[3]This also implies that the number of intervals over which the graph remains unchanged is countable.

for $t \in [t_k, t_{k+1})$. What we need to show is that $\tau(\Phi(t,0))$ converges to zero almost surely, for some proper coefficient of ergodicity $\tau(\cdot)$. For this purpose, we use κ as defined in section III.

For a given k, define the stochastic matrix

$$F_k = e^{-(t_{k+1}-t_k)L_{t_k}}.$$

which is the state transition matrix of (7) between two consecutive switchings t_k and t_{k+1}. Since the Laplacians are i.i.d. and the switching signal is a Poisson process, all matrices F_k are independent and identically distributed. Moreover, $\lambda_{n-1}(\mathbb{E}L_{t_k}) \neq 0$ implies that $|\lambda_2(\mathbb{E}F_k)| < 1$.[4] Hence, matrices F_k satisfy all the properties of the weight matrices in the discrete-time case. More precisely, they are i.i.d. stochastic matrices with positive diagonals and their average has only one unit-modulus eigenvalue. Therefore, similar to the discrete-time case, the events

$$\begin{aligned} B_r(T,\epsilon,m) = &\left\{\delta(F_{(r+1)n^2m}\cdots F_{rn^2m+1}) > \epsilon\right\} \cap \\ &\left\{t_{k+1}-t_k < T, \text{ for all } rn^2m < k \leq (r+1)n^2m\right\}. \end{aligned}$$

are independent (because of Poisson switching) and there exists $\epsilon > 0$ such that $\mathbb{P}(B_r(T,\epsilon,m)) > 0$, for some integer m and any $T > 0$. Hence, second Borel-Cantelli lemma implies that such events occur infinitely many times almost surely.

On the other hand, the submultiplicativity inequality (4) holds for $\tau_1 = \tau_2 = \kappa$ and as a result

$$\kappa(\Phi(t,0)) \leq (1-\epsilon)^{c(t)}$$

where $c(t)$ is the number of events $B_r(T,\epsilon,m)$ that have occurred up to time t. Since the events B_r occur infinitely often almost surely, $c(t)$ is an unbounded sequence and therefore, $\kappa(\Phi(t,0))$ converges to zero with probability one, proving almost sure asymptotic consensus. ■

Similar to the discrete-time case, theorem 4 provides a simple criterion to distinguish between the two cases of asymptotically reaching consensus almost surely and almost never. It states that the information in the average Laplacian matrix $\mathbb{E}L_{t_k}$ suffices to predict the long-run behavior of the linear switched system (7). The joint connectivity interpretation is also valid for this case. In the case that $\lambda_{n-1}(\mathbb{E}L_{t_k}) \neq 0$, the graphs of the network are jointly connected over time infinitely often almost surely (i.e. the union graph has a directed spanning tree) and as a result, a consensus is achieved. But note that the results of [4], [17], [18] are not directly applicable to this case, because they all assume a lower bound on the dwell-time (or an average dwell-time) in the switching, while the time interval between two switchings can be arbitrarily small for a Poisson process.

[4]To see the reason, suppose $|\lambda_2(\mathbb{E}F_k)| = 1$. Therefore, as in section IV, there are two orthogonal rows in every matrix F_k with probability one. By writing the Laplacian in the form of $L_{t_k} = I - \tilde{A}_k$ where $\tilde{A}_k$ is a non-negative matrix and using the expansion of the matrix exponential, one can see that our assumption implies $\lambda_{n-1}(\mathbb{E}L_{t_k}) = 0$.

VI. Conclusions

In this paper, we showed how the problem of reaching consensus over a switched linear system can be reduced to the problem of weak ergodicity of a sequence of matrices. In particular, for the case of i.i.d. weight matrices, we showed that ergodicity is a trivial event. Moreover, we showed that the discrete-time linear dynamical system $x(k) = W_k x(k-1)$ reaches state consensus almost surely if and only if $\mathbb{E}W_k$ has exactly one eigenvalue with unit modulus, or equivalently, the deterministic system $x(k) = (\mathbb{E}W_k)x(k-1)$ reaches a consensus. Similarly, we showed that for the continuous-time switched linear system $\dot{x}(t) = -L_{t_k}x(t)$ with i.i.d. Laplacian matrices and a Poisson switching reaches an asymptotic state consensus almost surely if and only if the expected Laplacian matrix has exactly one eigenvalue equal to zero, or equivalently, the deterministic system $\dot{x}(t) = -(\mathbb{E}L_{t_k})x(t)$ reaches a consensus.

References

[1] M. H. DeGroot, "Reaching a consensus," *Journal of American Statistical Association*, vol. 69, no. 345, pp. 118–121, Mar. 1974.

[2] S. Chatterjee and E. Seneta, "Towards consensus: Some convergence theorems on repeated averaging," *Journal of Applied Probability*, vol. 14, no. 1, pp. 89–97, Mar. 1977.

[3] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1984.

[4] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[5] S. Boyd, A. Gosh, B. Prabhakar, and D. Shah, "Gossip algorithms: Design, analysis and applications," in *Proceedings of IEEE INFOCOM 2005*, vol. 3, Miami, Mar. 2005, pp. 1653–1664.

[6] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proceedings of the 4th International Conference on Information Processing in Sensor Networks*, Apr. 2005, pp. 63–70.

[7] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, Sept. 2004.

[8] J. Cortes, S. Martinez, and F. Bullo, "Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions," *IEEE Transactions on Automatic Control*, vol. 51, no. 8, 2006.

[9] Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Transactions on Automatic Control*, vol. 50, pp. 1867–1872, 2005.

[10] C. W. Wu, "Synchronization and convergence of linear dynamics in random directed networks," *IEEE Transactions on Automatic Control*, vol. 51, no. 7, pp. 1207–1210, July 2006.

[11] R. Cogburn, "The ergodic theory of Markov chains in random environments," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 66, no. 1, pp. 109–128, 1984.

[12] ——, "On products of random stochastic matrices," in *Random matrices and their applications, Proceedings of American Mathematical Society Summer Research Conference*, vol. 50, Providence, 1986, pp. 199–213.

[13] R. L. Dobrushin, "Central limit theorem for non-stationary Markov chains, I, II," *Theory of Probability and its Applications*, vol. 1, pp. 65–80, 329–383, 1956, (English Translation).

[14] R. Durrett, *Probability: Theory and Examples*, 3rd ed. Belmont, CA: Duxbury Press, 2005.

[15] E. Seneta, *Non-negative Matrices and Markov Chains*, 2nd ed. New York: Springer, 1981.

[16] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York: Academic Press, 1979.

[17] A. S. Morse, "Dwell-time switching," in *Proceedings of second European Control Conference*, 1993, pp. 176–181.

[18] J. Hespanha, "Uniform stability of switched linear systems: extensions of LaSalle's invariance principle," *IEEE Transactions on Automatic Control*, vol. 49, no. 4, pp. 470–482, 2004.