Consensus Over Random Networks
Alireza Tahbaz-Salehi and Ali Jadbabaie

Abstract—We consider the consensus problem for stochastic discrete-time linear dynamical systems. The underlying graph of such systems at a given time instance is derived from a random graph process, independent of other time instances. For such a framework, we present a necessary and sufficient condition for almost sure asymptotic consensus using simple ergodicity and probabilistic arguments. This easily verifiable condition uses the spectrum of the average weight matrix. Finally, we investigate a special case for which the linear dynamical system converges to a fixed vector with probability one.

Index Terms—Consensus problem, random graphs, weak ergodicity, tail events.

I. INTRODUCTION

Decentralized iterative schemes such as agreement and consensus problems have an old history [1]–[4]. Over the past few years they have attracted a significant amount of attention in various contexts, such as motion coordination of autonomous agents [5], [6], distributed computation of averages and least squares among sensors [7]–[9], and rendezvous problems [10]. 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 [11], 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 \( G(W_k) \), 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 [12], where the edges of \( G(W_k) \) are directed and not necessarily independent. However, the authors prove only convergence to a consensus in probability, rather than the more general notion of almost sure convergence.

The purpose of this note is to provide a necessary and sufficient condition for an almost sure consensus in linear dynamical systems, when the weight matrices are general i.i.d. stochastic matrices. Our results contain the results of [11] and [12] as special cases. Furthermore, the necessary and sufficient condition is easily verifiable and only depends on the spectrum of the average weight matrix \( E W_k \). Finally, we state a variant of a theorem in [2] which provides the consensus value that the state of the system converge to for a special case.

II. PROBLEM SETUP

Let \((\Omega_0, \mathcal{B}, \mu)\) be a probability space, where \(\Omega_0 = S_n = \{\text{set of stochastic matrices of order } n\} \) with strictly positive diagonal entries, \(\mathcal{B}\) is the Borel \(\sigma\)-algebra of \(\Omega_0\), and \(\mu\) is a probability measure defined on \(\Omega_0\). 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:

\[
\Omega = \{ (\omega_1, \omega_2, \ldots) : \omega_k \in \Omega_0 \}
\]

\[
\mathcal{F} = \mathcal{B} \times \mathcal{B} \times \cdots
\]

\[
\mathbb{P} = \mu \times \mu \times \cdots
\]

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 autonomous dynamical system:

\[
x(k) = W_k(\omega)x(k-1),
\]

where \(k \in \{1, 2, \ldots\}\) is the discrete time index, \(x(k) \in \mathbb{R}^n\) is the state vector at time \(k\) and the mapping \(W_k : \Omega \rightarrow S_n\) is \(k\)th 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 \(\mu\). For notation 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 \(G(W)\) as a weighted directed graph with an edge \((i, j)\) from vertex \(i\) to vertex \(j\) with weight \(W_{ij}\) if and only if \(W_{ij} \neq 0\). In this case we say vertex \(j\) has access to vertex \(i\). We say vertices \(i\) and \(j\) communicate if both \((i, j)\) and \((j, i)\) are edges of \(G(W)\). Note that the communication relation is an equivalence relation and defines equivalence classes on the set of vertices. If all the vertices outside a specific communication class, have no access to any vertex in that class, such a class is called final.

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 \(\epsilon > 0\),

\[
\mathbb{P}\{|x_i(k) - x_j(k)| > \epsilon\} \rightarrow 0
\]

as \(k \rightarrow \infty\) for all \(i, j = 1, \ldots, n\).

This notion of reaching an eventual state agreement, which is addressed in [12], 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 \(x(0)\),

\[
|x_i(k) - x_j(k)| \rightarrow 0 \text{ almost surely}
\]

as \(k \rightarrow \infty\) for all \(i, j = 1, \ldots, n\).

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 [13].

III. ERGODICITY

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

\[
x(k) = W_k \cdots W_2 W_1 x(0).
\]

As it is evident from (2), one needs to investigate the behavior of infinite products of stochastic matrices in order to check for reaching an eventual 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 \(i, j, s = 1, \ldots, n\)

\[
(U_{i,s}^{(k)} - U_{j,s}^{(k)}) \rightarrow 0
\]

as \(k \rightarrow \infty\), where \(U_{i,s}^{(k)} = W_k \cdots W_2 W_1\) is the left product of the matrices in the sequence.

As the definition suggests, a sequence of stochastic matrices is weakly ergodic if the rows of the product matrix converge to each
other, as the number of terms in the product grows. 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 $i, s = 1, \ldots, n$

$$U^{(k)}_{i,s} \to d_s$$

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

One can easily see that weak and strong ergodicity both describe a tendency to consensus. If either type of ergodicity (weak or strong) holds for the matrix sequence $\{W_k\}_{k=1}^\infty = W_1, W_2, \ldots$, the pairwise differences between rows of the product matrix $U^{(k)}$ converge to zero. Conversely, there would be no consensus for all initial state values $x(0)$, unless the sequence of weight matrices is (weakly or strongly) ergodic.

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:

Theorem 1: A matrix sequence $\{W_k\}_{k=0}^\infty$ and their left products $U^{(k)} = W_k \cdots W_1$, weak and strong ergodicity are equivalent.

Proof: See [2].

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

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

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

d where $d$ is a vector of size $n$ satisfying $d^T 1 = 1$.

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

$$\tau(U^{(k)}) \to 0$$

as $k \to \infty$ for a proper coefficient of ergodicity $\tau$. Therefore, we can state the following theorem, the proof of which can be found in [14, theorem 4.8].

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

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

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

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

Two examples of coefficients of ergodicity used in this note 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

$$\kappa(W) \leq \nu(W).$$

IV. MAIN RESULTS: NECESSARY AND SUFFICIENT CONDITIONS FOR 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 $W_1, W_2, \ldots$ is a trivial event.

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

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

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

$$\kappa(W) \leq \nu(W).$$

Lemma 2: Suppose that $W$ is a stochastic matrix for which its corresponding graph has $s$ communication classes named $\alpha_1, \alpha_2, \ldots, \alpha_s$. Class $\alpha_s$ is final, if and only if the spectral radius of $[W] \alpha_s$ equals to one, where $[W] \alpha_s$ is the submatrix of $W$ corresponding to the vertices in the class $\alpha_s$.

Our next theorem indicates a criteria for occurrence of weak ergodicity with probability zero.

Theorem 3: Suppose the average weight matrix $EW_k$ has $n$ eigenvalues satisfying

$$0 \leq |\lambda_n(EW_k)| \leq \cdots \leq |\lambda_2(EW_k)| \leq |\lambda_1(EW_k)| = 1.$$

The random sequence $\{W_k\}_{k=0}^\infty = W_1, W_2, \ldots$ of stochastic matrices with positive diagonals is (weakly) ergodic almost never, if $|\lambda_2(EW_k)| = 1$.

Proof: Since all $\omega_k \in \Omega_0$ have positive diagonals, $EW_k$ has strictly positive diagonal entries as well. Hence, if $EW_k$ is irreducible, then it is aperiodic and as a result of the Perron-Frobenius theorem [16, theorem 6.6.1], $|\lambda_2(EW_k)| < 1$ which is in contradiction with our assumption. Therefore, $|\lambda_2(EW_k)| = 1$ implies reducibility of $EW_k$. As a result, without loss of generality, one can label the vertices such that $EW_k$ gets the following block triangular form

$$EW_k = \begin{bmatrix}
Q_{11} & 0 & \cdots & 0 \\
Q_{21} & Q_{22} & \cdots & 0 \\
& & \ddots & \vdots \\
Q_{s1} & Q_{s2} & \cdots & Q_{ss}
\end{bmatrix},$$

where $Q_{ii} = 0$ for $i = 1, 2, \ldots, s$.
where each $Q_{kl}$ is an irreducible matrix and represents the vertices in the equivalence class $\alpha_l$. Since $|\lambda_2(EW_k)| = 1$, submatrices corresponding to at least two of the classes have unit spectral radii (note that because of irreducibility and aperiodicity of $Q_{kl}$’s, the multiplicity of the unit-modulus eigenvalue of each one of them cannot be more than one). Therefore, lemma 2 implies,

$$\exists i \neq j \text{ s.t. } \alpha_i \text{ and } \alpha_j \text{ are both final classes, or equivalently, } Q_{ir} = 0 \text{ for all } r \neq i \text{ and } Q_{ij} = 0 \text{ for all } t \neq j.$$  

Since $\Omega_0$ is a subset of nonnegative matrices, $W_k$ has the same type (zero block pattern) as $E W_k$ for all time $k$ with probability one. Therefore, $U^{(k)} = W_k \cdots W_2 W_1$ has two orthogonal rows almost surely for any $k$ which means that the random sequence $(W_k)_{k=0}^{\infty}$ is weakly ergodic almost never.

Theorem 3 shows that a necessary condition for weak ergodicity (or equivalently, reaching an eventual consensus) with positive probability is $|\lambda_2(EW_k)| < 1$. Our next theorem suggests that this condition is also sufficient and therefore, by lemma 1, guarantees almost sure ergodicity.

**Theorem 4:** The random sequence $(W_k)_{k=0}^{\infty}$ of stochastic matrices with positive diagonals is (weakly) ergodic almost surely, if $|\lambda_2(EW_k)| < 1$.

**Proof:** Since $|\lambda_2(EW_k)|$ is subunit, lemma 2 implies that $G(EW_k)$ has exactly one final class. We investigate the two cases of $EW_k$ being irreducible and reducible separately.

1) Irreducibility: Suppose $EW_k$ is irreducible. Since it has only one unit-modulus eigenvalue, $EW_k$ is primitive [15]. Hence,

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

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

$$E(W_m \cdots W_1) = [EW_k]^m > 0.$$

As a result, for all $i, j = 1, \cdots, n$ the $(i,j)$ entry of $U^{(m)} = W_m \cdots W_2 W_1$ is positive with nonzero probability, say $\rho_{ij} > 0$. Therefore, since the weight matrices are i.i.d. with positive diagonals, the matrix $W_n \cdots W_2 W_1$ is completely entry-wise positive with at least probability $\prod_{i=1}^{n} \rho_{ij} > 0$, i.e. the event $(W_n \cdots W_2 W_1 > 0)$ has nonzero probability. Hence, by the second Borel-Cantelli lemma [13, page 49], we have

$$P(W_r > 0 \text{ for all } r) = 1.$$  

As a result, if we define $\delta(W) = 1 - \nu(W) = \max_i (\min_j W_{ij})$ and set $\kappa = r n^2 m$, we will have

$$\delta(W_{k+1} \cdots W_{k+1}) > 0 \quad \text{i.o. a.s.}$$

Also note that (3) hold for $\theta_i(\cdot) = \tau_i(\cdot) = \kappa(\cdot)$. Therefore, this together with (5) results in

$$\sum_{r=1}^{\infty} (1 - \kappa(W_{k+1} \cdots W_{k+1})) = \infty \quad \text{a.s.,}$$

which is exactly (4), the sufficient condition for weak ergodicity. Therefore, the sequence is weakly ergodic almost surely.

2) Reducibility: When $EW_k$ is reducible, without loss of generality, it can be written as (6), where all $Q_{kl}$ are irreducible matrices. Since $\alpha_1$ (the class corresponding to submatrix $Q_{11}$) is the only final class of $G(EW_k)$, there exists a directed path between any vertex of $G(EW_k)$ and a vertex in $\alpha_1$ (e.g. say vertex labeled 1), such that the length of the path is at most some positive integer $m$. In other words, any vertex of $G(EW_k)$ is at most an $m$-hop neighbor of vertex 1. This combined with the fact that $EW_k$ has strictly positive diagonals guarantees that the first column of $[EW_k]^m$ is strictly positive. Therefore, as in case 1, independence implies the positivity of the first column of $E(W_m \cdots W_1)$.

As a result, if $j = 1, \cdots, n$ the $(j,1)$-entry of the matrix $W_m \cdots W_1$ is nonzero with positive probability $\rho_{j1}$. Hence, again in parallel to the discussion of case 1, we have

$$P(\delta(W_{m+1} \cdots W_{m+1}) > 0) \geq \prod_{j=1}^{n} \rho_{j1} > 0.$$  

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

$$P(\delta(W_{k+1} \cdots W_{k+1}) > 0 \text{ for infinitely many } r) = 1.$$  

As a result, the sum $\sum_{r=1}^{\infty} (1 - \kappa(W_{k+1} \cdots W_{k+1}))$ diverges to infinity with probability one. Now, theorem 2 implies that the random sequence $(W_k)_{k=0}^{\infty}$ is weakly ergodic almost surely.

Theorems 3 and 4 together provide a simple criteria to distinguish between the two cases of almost sure and almost never eventual consensus. They suggest that the information in the average weight matrix $EW_k$, rather than the whole information in distribution $\mu$, is sufficient to predict the long-run behavior of the linear dynamical system (1). This should not come as a surprise to the reader. When $|\lambda_2(EW_k)|$ is subunit, there exists a sequence of integer numbers $k_r, r = 1, 2, \cdots$ such that the graph collection $\{G(W_{k+1}), \cdots, G(W_{k+1})\}$ is jointly connected (i.e. the graph constructed by unioning the edge sets of the graphs in the collection contains a spanning tree) [5]. This infinite often connectivity over time guarantees the possibility of information flow in the graph over time, and therefore results in eventual consensus with probability one. On the other hand, when $|\lambda_2(EW_k)| = 1$, 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, the theorems of this section contain the results of [11] and [12] as special cases. Since in [11] the authors use Erdős-Rényi as their random graph model, the matrix $EW_k$ is completely entry-wise positive, which results in $|\lambda_2(EW_k)| < 1$, and hence almost sure consensus. On the other hand, when the weight matrices are scrambling with positive probability, as in [12], $EW_k$ is also scrambling and as a result, its unit-modulus eigenvalue has multiplicity one. Hence, (1) reaches an eventual consensus almost surely (and therefore in probability).

V. CONSENSUS VALUE

As shown in the previous section, if $|\lambda_2(EW_k)|$ is subunit, then the linear dynamical system (1) converges with probability one and $x(k) \overset{d}{\rightarrow} c \mathbf{1}$, where $c$ is a scalar random variable depending on the initial state value $x(0)$ and the random sequence of weight matrices. The following theorem states that the distribution of $c$ is concentrated at a point for the special case that all weight matrices have the same left eigenvector corresponding to their unit eigenvalue.

**Theorem 5:** For $y \in \mathbb{R}^n$, set $S(y) = \{W \in S_n | y^T W = y^T\}$. For a given initial state value vector $x(0)$, if $|\lambda_2(EW_k)| < 1$ and
\( \mu(S_n - S(y)) = 0 \) hold, then

\[
\lim_{k \to \infty} x(k) = \left( y^T x(0) \right) 1 \text{ almost surely.}
\]

\textbf{Proof:} A variant of this theorem is proved as theorem 3 in [2]. A similar proof can be used here as well.

One special case of the above theorem is when the weight matrices are doubly stochastic almost surely. In such a case, the weight matrices have vector 1 as their common left eigenvector at all times and therefore, all the entries of the state vector converge to \( \frac{1}{n} \left( 1^T x(0) \right) 1 \), the average of the initial state values, with probability one. This special case is exactly theorem 1 of [17].

VI. CONCLUSION

In this note, we showed how the problem of reaching consensus 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 \( E(W_k) \) has exactly one eigenvalue with unit modulus. And finally we showed how our theorems simply recover the other known results in the field of consensus over random graphs.

REFERENCES