($\mathbf{1}$) First we obtain the reduced form of system Σ by Algorithm 3.2. ($\mathbf{2}$) Obtain characteristic polynomial $q(z) = \chi(A_{22})$ of the unreachable part. ($\mathbf{3}$) Check if $f(z)$ is a multiple of $q(z)$. ($\mathbf{4}$) If answer to (3) is YES then ($\mathbf{4}$) obtain polynomials $\chi(A_{r:r}), \chi(A_{r-1:r}, \ldots, \chi(A_{1:r})$ and ($\mathbf{5}$) solve modular identities arising from Theorem 2.3.

Collecting the costs of all procedures give us a total of $\mathcal{O}(rn^3)$ arithmetic operations in R plus the calculation of the characteristic polynomial $\chi(A_{22})$ of degree $n-r$. This characteristic polynomial can be computed deterministically (see [15, Th. 5.1]) up to a cost of $\mathcal{O}(((n-r)^{3+1/3})^{1+o(1)})$ arithmetic operations in R.

If R is an infinite domain then generic case for a single input system is rank $r = n$. Hence neither line (3) in the above procedure nor the calculation of $\chi(A_{22})$ are needed in most cases. The generic case on an infinite Euclidean domain involves a cost of $\mathcal{O}(n^4)$ arithmetic operations in R.

REFERENCES

[1] J. Assan, J. F. Lafay, and A. M. Perdon, "Weak feedback cyclizability and coefficient assignment for weakly reachable systems over a principal ideal domain," in *Proc. 37th. IEEE Conf. Decision Control*, 1998, vol. 3, pp. 3563–3568.

[2] J. W. Brewer, J. W. Bunce, and F. S. Van Vleck, *Linear Systems Over Commutative Rings*. New York: Marcel Dekker, 1986.

[3] W. C. Brown, *Matrices Over Commutative Rings*. New York: Marcel Dekker, 1993.

[4] R. Bumby, E. D. Sontag, H. J. Sussmann, and W. V. Vasconcelos, "Remarks on the pole-shifting problem over rings," *J. Pure Appl. Algebra*, vol. 20, pp. 113–127, 1981.

[5] M. Carriegos and J. A. Hermida-Alonso, "Canonical forms for single input linear systems," *Syst. Control Lett.*, vol. 49, no. 2, pp. 99–110, Jun. 2003.

[6] M. Carriegos, J. A. Hermida-Alonso, and T. Sánchez-Giralda, "Pole-shifting for linear systems over commutative rings," *Lin. Algebra App.*, vol. 346, pp. 97–107, 2002.

[7] F. Castro, *Calculs effectifs pour les idéaux d'opérateurs différentiels (Travaux en Cours, vol. 24)*. Paris, France: Herman, 1987.

[8] F. Chyzak, A. Quadrat, and D. Robertz, "Effective algorithms for parametrizing linear control systems over Ore algebras," *Applicable Algebra Eng., Commun. Comput.*, vol. 16, pp. 319–376, 2005.

[9] T. Cluzeau and A. Quadrat, "Factoring and decomposing a class of linear functional systems," *Lin. Algebra App.*, vol. 428, pp. 324–381, 2008.

[10] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*. Cambridge, MA: MIT Press, 1990.

[11] J. Ferrer, M. I. García, and F. Puerta, "Differentiable families of subspaces," *Lin. Algebra App.*, vol. 199, pp. 229–252, 1994.

[12] J. Ferrer, M. I. García, and F. Puerta, "Brunowsky local form of a holomorphic familiy of pairs of matrices," *Lin. Algebra App.*, vol. 253, pp. 175–198, 1997.

[13] H. Gluesing-Luerssen, *Linear Delay-Differential Systems with Commensurate Delays: An Algebraic Approach*. Berlin, Germany: Springer-Verlag, 2002.

[14] R. Hutchinson, J. Rosenthal, and R. Smarandache, "Convolutional codes with maximum distance profile," *Syst. Control Lett.*, vol. 54, pp. 53–63, 2005.

[15] E. Kaltofen and G. Villard, "On the complexity of computing determinants," *Comput. Complexity*, vol. 13, pp. 91–130, 2004.

[16] M. Kuijper, R. Pinto, and J. W. Polderman, "The predictable degree property and row reducedness for systems over a finite ring," *Linear Algebra Appl.*, vol. 425, pp. 776–796, 2007.

[17] E. Lee and W. Lu, "Coefficient assignability for linear systems with delays," *IEEE Trans. Autom. Control*, vol. AC-29, no. 11, pp. 1048–1052, Nov. 1984.

[18] V. Lomadze and E. Zerz, "Fractional representations of linear system," *Syst. Control Lett.*, vol. 39, pp. 275–281, 2000.

[19] V. Lomadze, "State and internal variables for linear systems," *Linear Algebra Appl.*, vol. 425, pp. 534–547, 2000.

[20] B. R. McDonald, *Linear Algebra over Commutative Rings*. New York: Marcel Dekker, 1984.

[21] A. S. Morse, "Ring models for delay-differential systems," *Automatica*, vol. 12, pp. 529–531, 1976.

[22] J. O'Halloran, "Feedback equivalence of constant linear systems," *Syst. Control Lett.*, vol. 8, pp. 241–246, 1987.

[23] J. Rosenthal, J. M. Schumacher, and E. V. York, "On behaviors and convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pt. 1, pp. 1881–1891, Nov. 1996.

[24] J. Rosenthal and E. V. York, "BCH convolutional codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pt. 1, pp. 1833–1844, Sep. 1999.

[25] A. Sáez-Schwedt and T. Sánchez-Giralda, "Coefficient assignability and a block decomposition for systems over rings," *Linear Algebra Appl.*, vol. 429, pp. 1277–1287, 2008.

[26] W. Schmale, "Computation of maximal reachability submodules," *Syst. Control Lett.*, vol. 42, pp. 229–232, 2001.

[27] E. Zerz, "Discrete multidimensional systems over $\mathbb{Z}_n$," *Syst. Control Lett.*, vol. 56, pp. 702–708, 2007.

Consensus Over Ergodic Stationary Graph Processes

Alireza Tahbaz-Salehi and Ali Jadbabaie

***Abstract*—In this technical note, we provide a necessary and sufficient condition for convergence of consensus algorithms when the underlying graphs of the network are generated by an ergodic and stationary random process. We prove that consensus algorithms converge almost surely, if and only if, the expected graph of the network contains a directed spanning tree. Our results contain the case of independent and identically distributed graph processes as a special case. We also compute the mean and variance of the random consensus value that the algorithm converges to and provide a necessary and sufficient condition for the distribution of the consensus value to be degenerate.**

***Index Terms*—Consensus algorithm, ergodic stationary process, random graph.**

I. INTRODUCTION

Due to their wide range of applications, distributed consensus algorithms have attracted significant amount of attention over the past few years. The main focus in the study of these algorithms is to derive conditions under which a group of agents in a network with *local* communication capabilities can reach *global* agreement, using simple, linear information exchange protocols. Applications include distributed and parallel computing [1], motion coordination of autonomous agents [2], [3], distributed sensor fusion [4], as well as opinion dynamics [5] and belief formation in social networks [6], [7].

More recently, there has been much interest in understanding the behavior of consensus algorithms in random settings. The randomness can be either due to the choice of a randomized network communication protocol or, simply caused by the potential unpredictability of the environment in which the distributed consensus algorithm is implemented [8]. Hatano and Mesbahi [9] provide one of the earliest studies

Manuscript received July 15, 2008; revised January 08, 2009. First published November 24, 2009; current version published January 13, 2010. This work is supported in part by the following grants: ARO/MURI W911NF-05-1-0381, DARPA/DSO SToMP, ONR MURI N000140810747, and NSF-ECS-0347285. Recommended by Associate Editor C. J. Tomlin.

A. Tahbaz-Salehi is with the Department of Electrical and Systems Engineering and Department of Economics, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: atahbaz@seas.upenn.edu).

A. Jadbabaie is with the General Robotics, Automation, Sensing and Perception (GRASP) Laboratory, Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: jadbabai@seas.upenn.edu).

Digital Object Identifier 10.1109/TAC.2009.2034054

of consensus algorithms over random networks. They prove that if communication links between any pair of agents are activated independently with some exogenously specified probability p (what is known as the Erdős-Rényi random graph model), then agents reach agreement asymptotically. Wu [10] and Porfiri and Stilwell [11] extend the results of [9] to more general settings. In [12], Tahbaz-Salehi and Jadbabaie study the asymptotic properties of random consensus algorithms over the general class of i.i.d. weighted and directed random graph sequences, where different communication links at a given time are correlated, even though realizations of the network at two different time steps are independent. They prove that randomized consensus algorithms over i.i.d. networks converge to consensus if and only if $|\lambda_2(\mathbb{E}W_k)| < 1$, where matrix $\mathbb{E}W_k$ captures the expected weights that agents assign to one another's states, with λ_2 representing its eigenvalue with the second largest modulus. Put differently, over i.i.d. random networks, asymptotic consensus is achieved if and only if the graph of the network contains a directed spanning tree in expectation. In a related paper, Fagnani and Zampieri [8] study the speed of convergence to consensus and provide concentration results for general i.i.d. network processes.

The common crucial assumption of the works mentioned above is that the realizations of underlying communication network among agents at different time steps are independent and identically distributed. However, in many realistic applications, this is too strong of an assumption. For example, the existence of a communication link in a wireless network at a given time is strongly correlated with its existence at previous time steps.

In this note, we relax the independence assumption and assume that the graphs representing the communication network among agents are generated by an ergodic and stationary process. Building on the results of Picci and Taylor [13] and by applying Birkhoff's ergodic theorem, we show that condition $|\lambda_2(\mathbb{E}W_k)| < 1$ appeared in [12] is a necessary and sufficient condition for almost sure convergence to consensus. This condition implies that existence of a directed path in the expected graph of the network from some node to all other nodes is both necessary and sufficient for reaching consensus with probability one. The results presented in this note are more general than [8]–[12], which assume independence over time. Also contrary to Picci and Taylor [13], who consider unweighted edges one at a time, we consider a general ergodic stationary process of stochastic matrices.

As a second contribution, we characterize the mean and variance of the random consensus value that the algorithm converges to in terms of the first and second moments of weight matrices W_k. We also provide a necessary and sufficient condition for the distribution of the random consensus value to be degenerate.

II. Ergodic Stationary Matrix Processes

Let $(\Omega_0, \mathcal{B})$ be a measurable space, where $\Omega_0 = \{\text{set of stochastic matrices of order } n \text{ with strictly positive diagonal entries}\}$ and $\mathcal{B}$ is the Borel σ-algebra on Ω_0. Consider probability measure $\mathbb{P}$ defined on the sequence space $(\Omega, \mathcal{F})$

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

such that $(\Omega, \mathcal{F}, \mathbb{P})$ forms a probability space. Let $\varphi : \Omega \to \Omega$ be the shift operator defined as $\varphi(\omega_1, \omega_2, \ldots) = (\omega_2, \omega_3, \ldots)$ and define the first coordinate map $W : \Omega \to \Omega_0$ as $W(\omega) = \omega_1$. For $\omega \in \Omega$, we define the sequence of stochastic matrices $\{W_k(\omega) : k \geq 1\}$, where $W_k(\omega) \triangleq W(\varphi^{k-1}\omega) = \omega_k$. For notational simplicity, we denote $W_k(\omega)$ by W_k.

Definition 1: A sequence of random stochastic matrices $W_1, W_2, \ldots$ is *stationary* if the families $\{W_{k_1}, W_{k_2}, \ldots, W_{k_r}\}$ and $\{W_{k_1+h}, W_{k_2+h}, \ldots, W_{k_r+h}\}$ have the same joint distribution for all $k_1, k_2, \ldots, k_r$ and all $h > 0$.

In other words, $\{W_k : k \geq 1\}$ is a stationary process if all of its finite dimensional distributions are invariant under time shifts. Equivalently, the process is stationary if the shift operator is a *measure-preserving* transformation, i.e., $\mathbb{P}(\varphi B) = \mathbb{P}(B)$ for all sets $B \in \mathcal{F}$.

Definition 2: Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and suppose that the shift operator $\varphi : \Omega \to \Omega$ is measure-preserving. φ is said to be *ergodic* if every invariant set $B \in \mathcal{F}$ is trivial.

In other words, transformation φ is ergodic if for every $B \in \mathcal{F}$ satisfying $\mathbb{P}(B \Delta \varphi^{-1} B) = 0$, we have $\mathbb{P}(B) \in \{0, 1\}$, where Δ denotes the symmetric difference between the two sets.[1]

Finally, we say random matrix process $\{W_k : k \geq 1\}$ is ergodic stationary, if the shift operator defined over $(\Omega, \mathcal{F}, \mathbb{P})$ is measure-preserving and ergodic. For example, a time-invariant Markov chain with its unique stationary distribution as the initial distribution forms a stationary and ergodic process. Clearly, any i.i.d. sequence of matrices is also both ergodic and stationary. We have the following lemma for ergodic stationary processes.

Lemma 1: Suppose $W_1, W_2, \ldots$ is an ergodic stationary process of stochastic $n \times n$ matrices. If the event $\{W_k \in A\}$ has positive probability $p > 0$, then such events occur infinitely often almost surely, that is, $\mathbb{P}(W_k \in A \text{ for infinitely many } k) = 1$.

Proof: Since the process $\{W_k : k \geq 1\}$ is ergodic stationary, so is the process $\{\mathbb{I}_{\{W_k \in A\}} : k \geq 1\}$, where $\mathbb{I}$ is the indicator function. Therefore, by Birkhoff's ergodic theorem [14], [15]

$$\frac{1}{T} \sum_{k=1}^{T} \mathbb{I}_{\{W_k \in A\}} \to \mathbb{P}\{W_1 \in A\} = p \text{ almost surely}$$

which implies

$$\mathbb{P}\left(\sum_{k=1}^{\infty} \mathbb{I}_{\{W_k \in A\}} = \infty\right) = 1.$$

Thus, the events $\{W_k \in A\}$ occur infinitely often almost surely. ∎

III. Consensus Over Random Networks

In this section, we present our framework for consensus algorithms over ergodic and stationary graph processes.

Consider the discrete-time autonomous dynamical system

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

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 $\{W_k(\omega) : k \geq 1\}$ is an *ergodic stationary* process of stochastic matrices with strictly positive diagonals, defined in Section II. We say dynamical system (1) reaches *consensus* asymptotically on some path $\omega \in \Omega$, if along that path, $|x_i(k) - x_j(k)| \to 0$ as $k \to \infty$ for all $i, j = 1, \ldots, n$. In other words, the system reach consensus on some path, if the difference between any two elements of the state vector, on that path, converges to zero. We now define almost sure convergence to consensus:

Definition 3: Dynamical system (1) reaches consensus almost surely, if for any initial state value $x(0)$

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

as $k \to \infty$ for all $i, j = 1, \ldots, n$.

A. Random Graph Interpretation

One can interpret linear dynamical system (1) as a randomized distributed scheme where a collection of agents, labeled 1 through n, up-

[1]The symmetric difference between two sets X and Y is defined as $X \Delta Y = (X \setminus Y) \cup (Y \setminus X)$.

date their state values as a convex combination of the state values of their neighbors at the previous time step. In this interpretation, $x_i(k)$ corresponds to the state value of agent i at time k, and W_k captures the neighborhood relations between different agents at time k. To further clarify this point, we define the graph corresponding to weight matrix W_k, denoted by $\mathbf{G}(W_k)$, as a weighted directed graph on n vertices, with an edge (i,j) from vertex i to vertex j with weight W_{ji} if and only if $W_{ji} \neq 0$. Given this definition, linear update $x(k) = W_k x(k-1)$ represents a distributed update scheme over the vertices of $\mathbf{G}(W_k)$, where the value of $x_i(k)$ only depends on the elements of the set $\{x_j(k-1) : (j,i) \text{ is an edge of } \mathbf{G}(W_k)\}$, which is the set of *neighbors* of agent i at time k.

If both (i,j) and (j,i) are edges of $\mathbf{G}(W_k)$, we say vertices i and j *communicate* at time k. Communication relation is an equivalence relation and defines equivalence classes on the set of vertices. If the vertices in a specific communication class have no neighbors outside of that class, such a class is called *initial*. Later in the note, we use the following lemma, the proof of which can be found in [16].

Lemma 2: Suppose that W is a stochastic matrix for which its corresponding graph $\mathbf{G}(W)$ has s communication classes named $\alpha_1, \ldots, \alpha_s$. 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, notice that the assumption that the sequence $\{W_k : k \geq 1\}$ is a general stationary and ergodic process implies that the edges of $\{\mathbf{G}(W_k) : k \geq 1\}$ are not necessarily independent over time. Instead, the existence of an edge in the network at some time step k_1 might be correlated with the weights of other edges at some other time k_2.

B. Weak Ergodicity

Given (1), the state vector at time k can be written as

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

where $x(0)$ denotes the vector of initial state values. Equation (2) suggests that asymptotic behavior of linear dynamical system (1) depends on the behavior of infinite products of stochastic matrices W_k. This motivates us to borrow the concept of weak ergodicity of a sequence of stochastic matrices from the theory of Markov chains.

Definition 4: The sequence $\{W_k\}_{k=1}^{\infty} = W_1, W_2, \ldots,$ of $n \times n$ stochastic matrices is *weakly ergodic*, if for all $i, j, s = 1, \ldots, n$ and all integer $p \geq 0$

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

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

As the definition suggests, a sequence of stochastic matrices is weakly ergodic if the difference between any two rows of the product matrix converges to zero, as the number of terms in the product grows.[2] Note that weak ergodicity does not require the left products $U^{(k,p)}$ to converge as $k \to \infty$. The following theorem, however, shows that in the presence of weak ergodicity, all infinite left products converge [17], [18].

Theorem 1: Suppose that matrix sequence $\{W_k\}_{k=0}^{\infty}$ is weakly ergodic. For all $i, s = 1, \ldots, n$ and all integers $p \geq 0$ there exist vectors $d^{(p)}$ not depending on i such that

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

as $k \to \infty$.

Proof: For any $\epsilon > 0$, weak ergodicity implies that for large enough k, we have $-\epsilon \leq U_{i,s}^{(k,p)} - U_{j,s}^{(k,p)} \leq \epsilon$ uniformly for all $i, j, s = 1, \ldots, n$. Since $U^{(k+1,p)} = W_{k+p+1} U^{(k,p)}$, we have

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

which by induction implies that

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

for all $i, s, h = 1, \ldots, n$ and $r \geq 0$. 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. ■

The above theorem implies that whenever $\{W_k : k \geq 1\}$ is weakly ergodic, a non-negative vector d exists such that $U^{(k,0)} \to \mathbf{1} d^T$, where $\mathbf{1}$ denotes a vector with all entries equal to one. Therefore, almost sure weak ergodicity of $\{W_k : k \geq 1\}$ guarantees that linear dynamical system (1) reaches consensus almost surely, with the asymptotic consensus value equal to $d^T x(0)$. We use this fact as the basis of our proofs for convergence to consensus. It is important to note that the converse of this statement is not true in general. In other words, the event of weak ergodicity of the sequence of matrices is a subset of the event that (1) reaches consensus asymptotically for all initial state values $x(0)$. For instance, the existence of a rank one matrix in the sequence implies asymptotic consensus, while it does not guarantee weak ergodicity.

We now define the *coefficient of ergodicity* which is an extremely useful and effective tool in dealing with infinite products of stochastic matrices.

Definition 5: The scalar continuous function $\tau(\cdot)$ defined on the set of $n \times n$ stochastic matrices is called a (proper) coefficient of ergodicity if $\tau(\cdot) \in [0,1]$ and

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

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

It is straightforward to show that weak ergodicity is equivalent to

$$\tau(U^{(k,p)}) \to 0 \quad \forall\, p \in \mathbb{N} \cup \{0\}$$

as $k \to \infty$ for some coefficient of ergodicity τ. We have the following theorem [18].

Theorem 2: Suppose $\tau(\cdot)$ is a coefficient of ergodicity that for any $m \geq 1$ stochastic matrices F_s, $s = 1, 2, \ldots, m$ satisfies

$$\tau(F_m \ldots F_2 F_1) \leq \prod_{s=1}^{m} \tau(F_s). \tag{3}$$

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

$$\sum_{r=1}^{\infty} \left(1 - \tau(W_{k_{r+1}} \ldots W_{k_r+1})\right) = \infty. \tag{4}$$

Proof: Suppose that there exists a strictly increasing sequence of positive integers k_r such that (4) holds. Then, inequality $\log x \leq x - 1$ implies that

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

and as a result, $\prod_{r=1}^{\infty} \tau(W_{k_{r+1}} \ldots W_{k_r+1}) = 0$. Therefore, (3) guarantees the sequence is weakly ergodic. ■

[2]To be more precise, we have stated the definition of weak ergodicity *in the backward direction*.

IV. Convergence of Consensus Algorithms Over Ergodic Stationary Graph Processes

In this section, we prove a necessary and sufficient condition for linear dynamical system (1) to converge to consensus almost surely, when weight matrix process $\{W_k : k \geq 1\}$ is ergodic and stationary. Our results contain the results of [9]–[12] as special cases, which simply assume an i.i.d. matrix process.

Theorem 3: Let $\{W_k : k \geq 1\} = W_1, W_2, \ldots$ denote a sequence of stochastic matrices with positive diagonals generated by an ergodic stationary process. Linear dynamical system (1) reaches consensus almost surely, if and only if $|\lambda_2(\mathbb{E}W_k)| < 1$, where λ_2 is the eigenvalue with the second largest modulus.[3]

Proof: Suppose $|\lambda_2(\mathbb{E}W_k)| = 1$, which implies that $\mathbb{E}W_k$ is reducible [16].[4] Therefore, without loss of generality, $\mathbb{E}W_k$ has the following block triangular form:

$$\mathbb{E}W_k = \begin{bmatrix} Q_{11} & 0 & \ldots & 0 \\ Q_{21} & Q_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ Q_{s1} & Q_{s2} & \ldots & Q_{ss} \end{bmatrix}$$

where each Q_{ii} is an irreducible matrix and represents the vertices in the i-th communication class of $\mathbb{E}W_k$, denoted by α_i. Since $\lambda_1(\mathbb{E}W_k) = |\lambda_2(\mathbb{E}W_k)| = 1$, submatrices corresponding to at least two classes have unit spectral radii. Therefore, by Lemma 2

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

or equivalently, $Q_{ir} = 0$ for all $r \neq i$ and $Q_{jl} = 0$ for all $l \neq j$. In other words, matrix $\mathbb{E}W_k$ has two orthogonal rows. This, and the non-negativity of the matrices in $\{W_k : k \geq 1\}$ imply that $U^{(k,0)} = W_k \ldots W_2 W_1$ has two orthogonal rows almost surely for any k. Therefore, there are initial conditions for which random discrete-time dynamical system (1) reaches consensus with probability zero.

We now prove the reverse implication. When $|\lambda_2(\mathbb{E}W_k)| < 1$, Lemma 2 implies that $\mathbf{G}(\mathbb{E}W_k)$ has exactly one initial class; that is, there exists a vertex i such that for all $j \neq i$ there is a sequence of vertices $i = j(0), j(1), \ldots, j(s_j) = j$ for which $(\mathbb{E}W_k)_{j(q),j(q-1)} > 0$. In other words, there exists a path of length s_j from some node i to any other node j in the expected graph of the network. As a result, there exists $\epsilon > 0$ such that

$$\mathbb{P}\left[(W_k)_{j(q),j(q-1)} > \epsilon\right] > 0 \text{ for all } q = 1, 2, \ldots, s_j$$

for all vertices j. Hence, Lemma 1 implies

$$\mathbb{P}\left[(W_k)_{j(q),j(q-1)} > \epsilon \text{ infinitely often}\right] = 1 \quad 1 \leq q \leq s_j$$

for all $j \neq i$. Since finite intersections of these events also occurs with probability one, there exists a deterministic time T for which

$$\mathbb{P}\left[\delta(W_T \ldots W_2 W_1) > \zeta\right] > 0$$

for some $\zeta > 0$, where $\delta(W) = \max_j(\min_i W_{ij})$. In other words, there exists a deterministic time T, for which all entries of at least one column of the matrix product $W_T \ldots W_2 W_1$ is bounded away from zero with positive probability. Now, once again the ergodicity and stationarity of sequence $\{W_k : k \geq 1\}$ implies that such an event occurs infinitely often almost surely, i.e.,

$$\mathbb{P}\left(\delta(W_{(r+1)T} \ldots W_{rT+1}) > \zeta \text{ for infinitely many } r\right) = 1.$$

Therefore, by defining $k_r = rT$, we have

$$\delta(W_{k_{r+1}} \ldots W_{k_r+1}) > \zeta \text{ infinitely often almost surely.}$$

Notice that $1 - \delta(W) \geq \tau_1(W) \triangleq (1/2)\max_{i,j} \sum_{s=1}^{n} |(W)_{is} - (W)_{js}|$. It is straightforward to verify that τ_1 is a coefficient of ergodicity that satisfies (3). Therefore

$$\sum_{r=1}^{\infty} \left(1 - \tau_1(W_{k_{r+1}} \ldots W_{k_r+1})\right) = \infty \text{ almost surely}$$

which is exactly (4), the sufficient condition for weak ergodicity. Thus, $\{W_k : k \geq 1\}$ is weakly ergodic almost surely, which implies that linear dynamical system (1) reaches consensus with probability one. ■

Theorem 3 establishes that $|\lambda_2(\mathbb{E}W_k)| < 1$ is a necessary and sufficient condition for almost sure asymptotic consensus in (1) when the weight matrices (and hence, their corresponding graphs) are generated by an ergodic stationary process. Therefore, asymptotic consensus over ergodic stationary graph processes is guaranteed if and only if the expected graph of the network contains a directed spanning tree. This result is a generalization of our results in [12], which provides a similar criterion for the i.i.d. case.

The ergodicity of the graph process can be interpreted as the property that the ensemble average coincides with the time average. In other words, when the expected graph of the network contains a directed spanning tree, then there exists a time sequence $\{k_r : r \geq 1\}$ such that collection of graphs $\{\mathbf{G}(W_{k_r+1}), \ldots, \mathbf{G}(W_{k_{r+1}})\}$ are infinitely often *jointly strongly rooted* with probability one, and therefore, asymptotic consensus is guaranteed almost surely [2].

Theorem 3 also states that depending on the second largest eigenvalue modulus of the expected weight matrix, weak ergodicity occurs with either probability 1 or 0. This was to be expected, as the event $B = \{W_1, W_2, \ldots \text{ is weakly ergodic}\}$ satisfies $B = \varphi B$ and therefore, is invariant, i.e., $\mathbb{P}(B \Delta \varphi B) = 0$. Due to ergodicity of φ, such an event must be trivial.

In order to illustrate the results presented in this section we provide a simple example.

Example 1: Consider a graph on n vertices with its potential undirected edges numbered 1 through $n(n-1)/2$. We assume that the realization of the graph at time k contains edge e with weight $1/n$ if and only if the e-th entry of the random vector $z_k \in \mathbb{R}^{n(n-1)/2}$ is non-negative, where z_k is generated by an autoregressive process of order one. More precisely, for any $i \neq j$ and edge $e = (i,j)$

$$(W_k)_{ij} = \frac{1}{n}\mathbb{I}_{\{z_{k_e} \geq 0\}} \tag{5}$$

$$z_k = \gamma z_{k-1} + (1-\gamma)\epsilon_k \mathbf{1} \tag{6}$$

where $\gamma \in [0,1)$ is a constant, $z_0 \sim \mathcal{N}(0, \frac{1-\gamma}{1+\gamma}\mathbf{1}\mathbf{1}^T)$, and $\{\epsilon_k : k \geq 1\}$ is a sequence of i.i.d. unit normal random variables independent from z_0. The diagonal elements of W_k are defined such that the matrix is stochastic. Note that z_k is a convex combination of its value at the previous time step and an independent noise term ϵ_k. This means that existence of an edge at time k is correlated with existence of other edges at the same time (due to the common noise term), as well as with the realization of the random vector z at all other times (as long as $\gamma \neq 0$).

Equations (5) and (6) together define an ergodic stationary weight matrix process $\{W_k : k \geq 1\}$. Therefore, we can apply Theorem 3. It is easy to verify that $\mathbb{P}(z_{k_e} \geq 0) = 1/2$, and as a consequence, $\mathbb{E}W_k =$

[3]Note that $\mathbb{E}W_k$ is time-invariant because of the stationarity assumption.

[4]Note that all diagonal entries of $\mathbb{E}W_k$ are strictly positive.

$(1/2)I + (1/2n)\mathbf{1}\mathbf{1}^T$, which is irreducible. Thus, linear dynamical system (1) reaches consensus with probability one.

V. Asymptotic Distribution of the Consensus Value

As stated in the previous section, the ergodic stationary matrix process $\{W_k : k \geq 1\}$ of stochastic matrices is weakly ergodic almost surely, if and only if, $|\lambda_2(\mathbb{E}W_k)| < 1$. Therefore, if the expected weight matrix has a unique unit-modulus eigenvalue, linear dynamical system (1) converges to consensus almost surely, and all agents agree on the random value $x^* = d^T x(0)$, where d is a unit vector such that $U^{(k,0)} \to \mathbf{1}d^T$.

A natural question to ask is whether one can determine the distribution of this random consensus value. However, except for some very special cases, computing the distribution of the consensus value is far from trivial, even when the weight matrices are independent and identically distributed. In this section, we investigate a special case, for which one can compute the distribution analytically. More specifically, we provide a necessary and sufficient condition for the random consensus value to be degenerate, i.e., a condition under which the consensus algorithm in (1) converges to a deterministic constant almost surely. We also compute the mean and variance of the random consensus value x^* for the case of i.i.d. weight matrices.

A. Convergence to a Degenerate Distribution

The next theorem provides a necessary and sufficient condition for the distribution of the asymptotic consensus value to be degenerate.

Theorem 4: Let $\{W_k : k \geq 1\} = W_1, W_2, \ldots$ denote a sequence of stochastic matrices with positive diagonals generated by an ergodic stationary process with $|\lambda_2(\mathbb{E}W_k)| < 1$. Also consider the deterministic vector a satisfying $\mathbf{1}^T a = 1$. Then, the left product $U^{(k,0)} = W_k \ldots W_1$ converges to $\mathbf{1}a^T$ almost surely, if and only if a is a left eigenvector of W_k corresponding to the unit eigenvalue, with probability one.

Proof: The sufficiency proof is trivial and quite well-known [4], [17]: since $|\lambda_2(\mathbb{E}W_k)| < 1$, Theorem 3 guarantees that the product in (2) converges to a rank one matrix with probability one, i.e., $W_k \ldots W_2 W_1 \to \mathbf{1}d^T$ almost surely, for some random vector d. In the case that almost all weight matrices share the same left eigenvector a corresponding to the unit eigenvalue with probability one,[5], any product $U^{(k,0)}$ has also the same left eigenvector, and so does its limit as $k \to \infty$. Therefore, $W_k \ldots W_2 W_1 \to \mathbf{1}a^T$ almost surely, which means that $\mathbb{P}(d = a) = 1$.

To prove the reverse implication assume $|\lambda_2(\mathbb{E}W_k)| < 1$. Also, suppose that there exists a non-random stochastic vector a such that $U^{(k,0)} = W_k \ldots W_1 \to \mathbf{1}a^T$ almost surely. Since the sequence $\{W_k : k \geq 1\}$ is stationary, $U^{(k,1)} = W_k \ldots W_2$ should also converge to $\mathbf{1}a^T$ almost surely. Combining the above, we have

$$U^{(k,0)} = U^{(k,1)} W_1 \to \mathbf{1}a^T W_1 \text{ almost surely.}$$

As a consequence, $\mathbb{P}\left(a^T W_1 = a^T\right) = 1$, which means that almost all weight matrices have the same common left eigenvector a corresponding to the unit eigenvalue, with probability one. ■

A special case of interest is when all matrices that can appear with positive probability are doubly-stochastic. In this special case, $a = (1/n)\mathbf{1}$ is a common left eigenvector of all matrices in the sequence. Theorem 4 states that this is a necessary and sufficient condition for the limiting consensus value to be equal to the average of initial values $x(0)$ almost surely. In such a case, we say the linear dynamical system reaches an *average consensus* with probability one.

As a final remark, note that stationarity of the matrix process plays a crucial role in proving the necessity part of the above theorem. In fact, if the weight matrix process is not stationary, having a common left eigenvector corresponding to the unit eigenvalue is not necessary anymore. For instance, consider the following two stochastic matrices:

$$W_1 = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}, \quad W_2 = \begin{bmatrix} \frac{1}{5} & \frac{4}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix}.$$

It is easy to verify that neither matrix is doubly stochastic. However, the product $W_2 W_1$ is a doubly stochastic matrix. Therefore, if matrices W_k that appear in the sequence are doubly stochastic for $k \geq 3$, the linear dynamical system converges to the average consensus, even though W_1 and W_2 are not doubly stochastic.

B. Computing First and Second Moments

As stated before, computing the distribution of the consensus value in terms of the distribution of the weight matrices remains an open problem. Nonetheless, it is possible to compute the first two moments of the random consensus value. In the remainder of this section, we compute the mean and variance of the random consensus value $x^* = d^T x(0)$ for a general i.i.d. process.

Computing the mean of the consensus value is straightforward. We showed that whenever $|\lambda_2(\mathbb{E}W_k)|$ is subunit, $W_k \ldots W_2 W_1 \to \mathbf{1}d^T$ almost surely, for some random stochastic vector d. By taking expectations and applying the dominated convergence theorem [14], one obtains

$$\mathbb{E}[W_k \ldots W_2 W_1] \to \mathbb{E}[\mathbf{1}d^T]$$

which implies $[\mathbb{E}W_1]^k \to \mathbf{1}(\mathbb{E}d^T)$, due to independence. Therefore, by the Perron-Frobenius theorem, $\mathbb{E}d$ is simply equal to the normalized left eigenvector of $\mathbb{E}W_k$, corresponding to its unit eigenvalue.[6] Thus, the mean of the asymptotic consensus value x^* conditional on the initial condition $x(0)$ is given by $\mathbb{E}x^* = x(0)^T \mathbf{v}_1(\mathbb{E}W_k)$, where $\mathbf{v}_1(\cdot)$ denotes the normalized left eigenvector corresponding to the unit eigenvalue. For example, if the expected weight matrix is symmetric (and hence, doubly stochastic), then the expected consensus value is equal to the average of the initial conditions.

In order to compute the variance, first note that

$$\frac{1}{n}(W_k \ldots W_1)^T (W_k \ldots W_1) \to dd^T \text{ almost surely}$$

which can be rewritten as

$$\begin{aligned} &\frac{1}{n}\text{vec}\left[(W_k \ldots W_1)^T (W_k \ldots W_1)\right] = \\ &\quad = \frac{1}{n}\left(W_1^T \otimes W_1^T\right)\left(W_2^T \otimes W_2^T\right)\ldots\left(W_k^T \otimes W_k^T\right)\text{vec}(I_n) \\ &\quad \longrightarrow \text{vec}(dd^T) \text{ almost surely} \end{aligned}$$

where vec is the vectorization operator, $\otimes$ denotes the Kronecker product and I_n is the identity matrix of size n.[7] By applying the dominated convergence theorem once again, and using the assumption that the weight matrices are independent, we get

$$\frac{1}{n}\left[\mathbb{E}\left(W_1^T \otimes W_1^T\right)\right]^k \text{vec}(I_n) \longrightarrow \mathbb{E}[\text{vec}(dd^T)] = \mathbb{E}(d \otimes d).$$

Hence, by the Perron-Frobenius theorem

$$\begin{aligned} \mathbb{E}(d \otimes d) &= \frac{1}{n}\mathbf{v}_1(\mathbb{E}[W_k \otimes W_k])\left(\mathbf{1}_{2n}^T \text{vec}(I_n)\right) \\ &= \mathbf{v}_1(\mathbb{E}[W_k \otimes W_k]) \end{aligned}$$

[5]Note that since $|\lambda_2(\mathbb{E}W_k)|$ is subunit, there is only one such vector a.

[6]The assumption $|\lambda_2(\mathbb{E}W_k)| < \lambda_1(\mathbb{E}W_k) = 1$ guarantees that such an eigenvector exists and is unique.

[7]In deriving this expression, we have used the identity $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$.

where $\mathbf{v}_1(\cdot)$ denotes the normalized left eigenvector corresponding to the unit eigenvalue. Therefore, the covariance matrix of the random vector d satisfies

$$\begin{aligned}\mathrm{vec}(\mathrm{cov}(d)) &= \mathrm{vec}(\mathbb{E}dd^T) - \mathrm{vec}(\mathbb{E}d\mathbb{E}d^T)\\ &= \mathbb{E}\left(d \otimes d\right) - \mathbb{E}d \otimes \mathbb{E}d\\ &= \mathbf{v}_1\left(\mathbb{E}\left[W_k \otimes W_k\right]\right) - \mathbf{v}_1\left[\mathbb{E}W_k\right] \otimes \mathbf{v}_1\left[\mathbb{E}W_k\right].\end{aligned}$$

By combining all the above, one can compute the conditional variance of the random consensus value $x^* = d^T x(0)$ in terms of the moments of the weight matrices

$$\mathrm{var}\, x^* = [x(0)\otimes x(0)]^T \mathbf{v}_1(\mathbb{E}\left[W_k \otimes W_k\right]) - \left[x(0)^T \mathbf{v}_1(\mathbb{E}W_k)\right]^2. \quad (7)$$

It is easy to verify that the variance is equal to zero, if and only if almost all weight matrices share the same left eigenvector corresponding to the unit eigenvalue. Therefore, the distribution of the consensus value is degenerate if and only if there exists a vector a such that $a^T W_k = a^T$ with probability one, as shown in Theorem 4.

VI. Conclusion

In this note, we proved a necessary and sufficient condition for almost sure convergence of consensus algorithms over general weighted and directed stationary and ergodic random graph processes. We showed that 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. Our results contain the cases of i.i.d. and (ergodic and stationary) Markovian graph processes as special cases. We also showed that, given the assumptions of ergodicity and stationarity, the asymptotic consensus value has a degenerated distribution, if and only if almost all weight matrices share a common left eigenvector corresponding to their unit eigenvalue. Finally, we provided expressions for the mean and variance of the consensus value for i.i.d. random networks, in terms of the first two moments of weight matrices W_k.

Acknowledgment

The authors would like to thank A. Ozdaglar and F. Schofheide for helpful comments and discussions.

References

[1] J. N. Tsitsiklis, "Problems in Decentralized Decision Making and Computation," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1984.

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

[3] J. Cortes, S. Martinez, and F. Bullo, "Analysis and design tools for distributed motion coordination," in *Proc. Amer. Control Conf.*, Portland, OR, Jun. 2005, pp. 1680–1685.

[4] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4th Int. Conf. Inform. Processing Sensor Networks*, Apr. 2005, pp. 63–70.

[5] M. H. DeGroot, "Reaching a consensus," *J. Amer. Stat. Assoc.*, vol. 69, no. 345, pp. 118–121, Mar. 1974.

[6] B. Golub and M. O. Jackson, "Naïve Learning in Social Networks: Convergence, Influence, and the Wisdom of Crowds," *Amer. Econ. J.: Microecon.*, to be published.

[7] D. Acemoglu, A. Ozdaglar, and A. ParandehGheibi, "Spread of (Mis)Information in Social Networks," Massachusetts Institute of Technology, Cambridge, MA, LIDS Rep. 2812, 2009.

[8] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," *IEEE J. Selected Areas Commun.*, vol. 26, no. 5, pp. 634–649, May 2008.

[9] Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1867–1872, Nov. 2005.

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

[11] M. Porfiri and D. J. Stilwell, "Consensus seeking over random weighted directed graphs," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1767–1773, Sep. 2007.

[12] A. Tahbaz-Salehi and A. Jadbabaie, "A necessary and sufficient condition for consensus over random networks," *IEEE Trans. Autom. Control*, vol. 53, no. 3, pp. 791–795, Apr. 2008.

[13] G. Picci and T. Taylor, "Almost sure convergence of random gossip algorithms," in *Proc. 46th IEEE Conf. Decision Control*, New Orleans, LA, Dec. 2007, pp. 282–287.

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

[15] I. P. Cornfeld, S. V. Fomin, and Y. G. Sinai, *Ergodic Theory*, ser. Grundlehren der mathematischen Wissenschaften. New York: Springer, 1982, vol. 245.

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

[17] S. Chatterjee and E. Seneta, "Towards consensus: Some convergence theorems on repeated averaging," *J. Appl. Prob.*, vol. 14, no. 1, pp. 89–97, Mar. 1977.

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

Modelling and Control of Bi-Directional Discrete Linear Repetitive Processes

Jacek Bochniak, Krzysztof Galkowski, and Eric Rogers

***Abstract*—Repetitive processes are characterized by a series of sweeps or passes through a set of dynamics defined over a finite duration where the output produced on any pass acts as a forcing function on, and hence contributes to, the dynamics of the next pass. The resulting control problem is that the output sequence of pass profiles can contain oscillations that increase in amplitude in the pass-to-pass direction. This paper considers bi-directional operation, i.e. a pass is completed and at the end the next one begins but in the opposite direction. In particular, a model for such a process in the case of discrete dynamics is first proposed and new results on stability and control law design for stabilization and performance developed.**

***Index Terms*—Control, modelling, stability, uni- and bi-directional repetitive process.**

I. Introduction

The unique characteristic of a repetitive, or multipass, process is a series of sweeps, termed passes, through a set of dynamics defined over a fixed finite duration known as the pass length. On each pass an output,

Manuscript received April 10, 2008; revised May 14, 2009, May 19, 2009, and October 15, 2009. First published December 04, 2009; current version published January 13, 2010. This work was supported in part by the Ministry of Science and Higher Education in Poland under Project N N514 293235. Recommended by Associate Editor K. Fujimoto.

J. Bochniak and K. Galkowski are with the Faculty of Electrical Engineering, Computer Science and Telecommunications, University of Zielona, Gora 65-246, Poland (e-mail: j.bochniak@weit.uz.zgora.pl; k.galkowski@issi.uz.zgora.pl).

E. Rogers is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: etar@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/TAC.2009.2036288