6-2012

Consensus Over Martingale Graph Processes

Arastoo Fazeli
University of Pennsylvania

Ali Jadbabaie
University of Pennsylvania, jadbabai@seas.upenn.edu

Fazeli, A. & Jadbabaie, A. (2012). Consensus over martingale graph processes. *American Control Conference (ACC), 2012*, pp.845-850.
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6315532&isnumber=6314593

Consensus Over Martingale Graph Processes

Abstract
In this paper, we consider a consensus seeking process based on repeated averaging in a randomly changing network. The underlying graph of such a network at each time is generated by a martingale random process. We prove that consensus is reached almost surely if and only if the expected graph of the network contains a directed spanning tree. We then provide an example of a consensus seeking process based on local averaging of opinions in a dynamic model of social network formation which is a martingale. At each time step, individual agents randomly choose some other agents to interact with according to some arbitrary probabilities. The interaction is one-sided and results in the agent averaging her opinion with those of her randomly chosen neighbors based on the weights she assigns to them. Once an agent chooses a neighbor, the weights are updated in such a way that the expected values of the weights are preserved. We show that agents reach consensus in this random dynamical network almost surely. Finally, we demonstrate that a Polya Urn process is a martingale process, and our prior results in [1] is a special case of the model proposed in this paper.

Keywords
Consensus, Stochastic Process, Social Networks, Random Networks, Graph Theory

Disciplines
Controls and Control Theory | Dynamic Systems | Other Applied Mathematics | Probability

Comments
Fazeli, A. & Jadbabaie, A. (2012). Consensus over martingale graph processes. *American Control Conference (ACC), 2012*, pp.845-850. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6315532&isnumber=6314593

Consensus Over Martingale Graph Processes

Arastoo Fazeli and Ali Jadbabaie

Abstract— **In this paper, we consider a consensus seeking process based on repeated averaging in a randomly changing network. The underlying graph of such a network at each time is generated by a martingale random process. We prove that consensus is reached almost surely if and only if the expected graph of the network contains a directed spanning tree. We then provide an example of a consensus seeking process based on local averaging of opinions in a dynamic model of social network formation which is a martingale. At each time step, individual agents randomly choose some other agents to interact with according to some arbitrary probabilities. The interaction is one-sided and results in the agent averaging her opinion with those of her randomly chosen neighbors based on the weights she assigns to them. Once an agent chooses a neighbor, the weights are updated in such a way that the expected values of the weights are preserved. We show that agents reach consensus in this random dynamical network almost surely. Finally, we demonstrate that a Polya Urn process is a martingale process, and our prior results in [1] is a special case of the model proposed in this paper.**

I. INTRODUCTION

Consensus algorithms based on local averaging have attracted a significant amount of attention in a diverse set of applications and contexts. These applications range from parallel and distributed computation [2], distributed control and coordination [3]–[5] and robotics [6], to opinion dynamics and belief formation in social networks [7], [8]. Recently, however, there has been a growing interest in studying consensus algorithms in a probabilistic setting. This randomness can be due to the unpredictability of the environment in which the communication between agents occurs or due to the inherent probabilistic characteristic of the communication among agents [9].

Existing results on random consensus include the case of Erdos-Reyni model (where edges are independent from each other and also independent over time) [10], as well as [11] and [12]. In [13], the authors study consensus over randomly switching networks where the graphs are i.i.d. over time, but graph edges could be correlated. The authors show that consensus is reached almost surely in i.i.d. networks if and only if the graphs of the networks contain a directed spanning tree in expectation. Closed form formulas for the mean and the variance of consensus value in i.i.d. networks, in terms of the first two moments of the i.i.d. weight matrices, is found in [14].

In all of the studies mentioned above a common crucial assumption is that the realizations of the network are independent and identically distributed over time. This is in fact a very strong assumption; since in many realistic applications the realizations of a network at different time steps are correlated. For instance, in wireless networks links formed at each time step are strongly correlated with links formed at previous time steps. Also in social network settings, the interactions between people at each time depend on the social interactions in the past and will influence the future formation of the social network. Therefore, network models which allow for correlation in time seem to be more realistic.

In certain special cases (such as the case of an ergodic stationary process [14] or a Markov process [15]) similar results can be derived for non-i.i.d. processes. In [14], the authors show that the independence assumption can be replaced by ergodicity and stationarity. More recently, the authors in [15] have shown similar results (albeit under more stringent conditions) for the average consensus problem when the network change is governed by a Markov chain.

One property that both i.i.d. and ergodic stationary processes have in common is that the expected graph remains fixed over time. In fact, the necessary and sufficient condition for consensus in these two settings is provided for this constant expected graph of the network. Therefore, one is tempted to conclude that a process which preserves the expectation over time is likely to reach consensus. For example a social network where the interactions between agents change over time, while the expectation of the interactions stays constant.

One of the most well known class of stochastic processes in which the expectation is preserved is a martingale process. In a martingale process, the realizations of the process can be correlated over time, although the expectation of the process does not change. In fact the main property of a martingale process is that the observed value of the process at each time step is equal to the expectation of the process at the next time step with respect to the available information so far.

In this paper we provide a necessary and sufficient condition for almost sure convergence to consensus in the linear dynamical system $x(k+1) = W(k)x(k)$, where the weight matrices are generated by a martingale stochastic process. We show that we reach consensus with probability one if the expected graph of the network contains a directed spanning tree. Algebraically, this condition is equivalent to $\lambda_2(\mathbb{E}(W(k))) < 1$. This is easily verifiable and only depends on the spectrum of the average weight matrix $\mathbb{E}(W(k))$. It is in fact the same condition for consensus in both i.i.d. and ergodic-stationary graph processes.

The rest of this paper is organized as follows: In section II, we rigorously define a martingale graph process . In section III, we introduce the notion of consensus over random graph processes and we define the coefficient of ergodicity. In section IV, we study distributed consensus algorithms over martingale random networks and provide necessary and sufficient conditions for almost sure consensus. In section V, we propose a consensus seeking process based on local averaging of opinions in a dynamic model of social network

The authors are with the Department of Electrical and Systems Engineering and GRASP Laboratory, University of Pennsylvania, Philadelphia, PA 19104-6228 USA. {arastoo, jadbabai}@seas.upenn.edu

This research was supported by ONR MURI HUNT, ONR MURI N000140810747, ARO MURI SWARMS, and AFOSR Complex Networks Program.

formation which is a martingale and also is the generalization of our proposed model in [1]. Finally, in section VI, we conclude the paper.

II. Martingale Matrix Processes

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

$$\Omega = \{(w_1, w_2, \ldots) : w_k \in \Omega_0\}$$
$$\mathcal{F} = \mathcal{B} \times \mathcal{B} \times \ldots$$

such that $(\Omega, \mathcal{F}, \mathbb{P})$ forms a probability space. Let $\varphi : \Omega \to \Omega$ be the shift operator defined as $\varphi(w_1, w_2, \ldots) = (w_2, w_3, \ldots)$ and define the first coordinate map $W : \Omega \to \Omega_0$ as $W(w) = w_1$. For $w \in \Omega$, we define the sequence of stochastic matrices $\{W_k(w) : k \geq 1\}$, where $W_k(w) \triangleq W(\varphi^{k-1}w) = w_k$. For notational simplicity, we denote $W_k(w)$ by $W(k)$.

Definition 1: A sequence $\{\mathcal{F}_k, k \geq 0\}$ of σ-fields is called a filtration on $(\Omega, \mathcal{F})$ if $\mathcal{F}_k \subseteq \mathcal{F}_{k+1} \subseteq \mathcal{F}$ for all $\{k \geq 0\}$. A sequence $\{W(k)\}$ of random matrices on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be adapted to the filtration $\{\mathcal{F}_k\}$ if $W(k) \in \mathcal{F}_k$ for each k. In other words, an adopted sequence at each time is measurable with respect to the information of the process so far.

Definition 2: For a sequence of matrices, a submartingale process is an adapted sequence to the filtration $\{\mathcal{F}_k\}$ satisfying the inequality

$$\mathbb{E}(W_{ij}(k+1)|\mathcal{F}_k) \geq W_{ij}(k).$$

If the inequality holds with strict equality

$$\mathbb{E}(W_{ij}(k+1)|\mathcal{F}_k) = W_{ij}(k),$$

the sequence of matrices is a martingale process. In other words, $\{W(k)\}$ is a martingale process if the observed value of the process at each time step is equal to the expectation of the process at the next time step with respect to the available information so far. Obviously a martingale process is submartingale as well. By taking conditional expectation from both sides of a martingale process we have

$$\mathbb{E}(W_{ij}(k+1)) = \mathbb{E}(W_{ij}(k)).$$

Therefore, the expectation of a martingale process is preserved over time. This is in fact a very useful property of a martingale process and is the same property as that of i.i.d. and ergodic stationary processes. Having the conservation of expectation in both i.i.d. and martingale processes, we can predict that these two processes might have the same important properties as well, which we will discuss in the following chapters.

III. Consensus Over Random Networks

In this section, we present our framework for consensus algorithms over martingale graph processes. Consider the discrete-time dynamical system

$$x(k+1) = W(k)x(k), \tag{1}$$

where $k \in \{0, 1, 2, ...\}$ is the discrete time index, $x(k) \in \mathbb{R}^n$ is the state vector at time k, and $\{W(k)\}_{k=0}^{\infty}$ is a martingale sequence of stochastic matrices with strictly positive diagonals, defined in the section II. In social network settings, we can think of $x_i(k) \in \mathbb{R}$ as the belief or opinion of agent i at time k, where at each time each agent takes an average of her belief and the belief of the agents she has access to. In other words, linear dynamical system (1) can be viewed as a distributed averaging scheme over the set of n vertices, where at each time step each agent updates her belief as a convex combination of the beliefs of her neighbors and her belief at the previous time step. The neighborhood relation between different agents at each time step is captured by the weight matrix $W(k)$ where $W_{ij}(k)$ is the weight that agent i assigns to agent j. From the graph theory point of view, the weight matrix $W(k)$ corresponds to the weighted graph $G(W(k))$ defined on n vertices, where an edge (i, j) from vertex i to vertex j exists with weight W_{ji} if and only if $W_{ji} \neq 0$. 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 $G(W(k))$. Note that 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, i.e. $W_{ir} = 0$ for all vertices i inside the communication class and all vertices r outside the communication class, such a class is called initial. For communication classes of a stochastic matrix we have the following lemma (the proof can be found in [16]) which we use in the next section.

Lemma 1: Suppose that W is a stochastic matrix for which its corresponding graph 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 1, where $\alpha_r[W]$ is the submatrix of W corresponding to the vertices in the class α_r.

We say dynamical system (1) reaches consensus asymptotically on some path $w \in \Omega$, if along that path, $|x_i(k) - x_j(k)| \to 0$ as $k \to \infty$ for all $i, j \in \{1, \ldots, n\}$. In other words, the system reaches consensus on some path, if the difference between any two elements of the state vector, on that path, converges to zero. Almost sure convergence to consensus can also be defined as follows.

Definition 3: Dynamical system (1) reaches consensus almost surely, if for any initial state x_0 and all $i, j \in \{1, \ldots, n\}$

$$\mathbb{P}(\lim_{k\to\infty} |x_i(k) - x_j(k)| = 0) = 1.$$

In other words, the dynamical system reaches consensus almost surely, if for all sample path $w \in \Omega$, the system reaches consensus asymptotically.

We now define the *coefficient of ergodicity* which is a useful tool in dealing with infinite products of stochastic matrices and as a result with consensus in dynamical systems.

Definition 4: The scalar continuous function $\tau(.)$ defined on the set of $n \times n$ stochastic matrices is called a coefficient of ergodicity if it satisfies $0 \leq \tau(.) \leq 1$. A coefficient of ergodicity is said to be proper if

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

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

Two examples of coefficients of ergodicity 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 $\nu(W)$ is an improper coefficient of ergodicity, while $\kappa(W)$ is proper, and for any stochastic matrix W, they satisfy

$$\kappa(W) \leq \nu(W). \tag{2}$$

The coefficient of ergodicity is submultiplicative for row-stochastic matrices, i.e.

$$\tau(W(n)\cdots W(2)W(1)) \leq \prod_{k=1}^{n}\tau(W(k)). \tag{3}$$

Therefore, if $\tau(W(k)) < 1-\epsilon < 1$ for all k and some constant $\epsilon > 0$, then $\lim_{n\to\infty}\prod_{k=1}^{n}W(k) = \mathbf{1}_n d^T$ and the dynamical system (1) reaches consensus.

IV. Convergence Of Consensus Algorithms Over Martingale Graph Processes

In this section, we provide a necessary and sufficient condition for linear dynamical system (1) to reach consensus almost surely, when weight matrix process $\{W(k)\}$ is generated according to a martingale process. Our results contain the result of our previous work in [1] as a special case, where the number of links for each agent at each time step is constant. We show that the dynamical system (1) reaches consensus in three steps. In the first step, we show that a (sub)martingale sequence of stochastic matrices converges to a random limit almost surely. In the second step, we show that if the expected graph of the network contains a directed spanning tree, then the network's graph's random limit contains a directed spanning tree almost surely as well. In the third step, utilizing the results of the previous steps, we show that a dynamical system containing a directed spanning tree in expectation, and as a result in the random limit, reaches consensus almost surely.

A. Convergence to a Random Limit

In this subsection, we show that a martingale sequence of stochastic matrices, and its composition with power function, converges to a random limit almost surely. For this purpose and to develop some basic tools for our analysis, we review the concept of convexity with respect to the non-negative matrix cone.

Lemma 2: The power of a non-negative matrix is convex with respect to the non-negative matrix cone.

Proof: We prove by induction that $(\alpha X+(1-\alpha)Y)^n \leq \alpha X^n + (1-\alpha)Y^n$. For $k=2$ it can be easily seen that the inequality holds:

$$\begin{aligned}
&(\alpha X + (1-\alpha)Y)^2\\
&= \alpha^2X^2 + (1-\alpha)^2Y^2 + \alpha(1-\alpha)(XY+YX)\\
&\leq \alpha^2X^2 + (1-\alpha)^2Y^2 + \alpha(1-\alpha)(X^2+Y^2)\\
&= \alpha X^2 + (1-\alpha)Y^2.
\end{aligned}$$

The last inequality holds since $0 \leq (X-Y)^2$, and therefore, $XY+YX \leq X^2+Y^2$. Now let us assume that it holds for $k=n$. Then for $k=n+1$ we have

$$\begin{aligned}
&(\alpha X + (1-\alpha)Y)^{n+1}\\
&\leq (\alpha X^n + (1-\alpha)Y^n)(\alpha X + (1-\alpha)Y)\\
&= \alpha^2X^{n+1} + (1-\alpha)^2Y^{n+1} + \alpha(1-\alpha)(X^nY+Y^nX)\\
&\leq \alpha^2X^{n+1} + (1-\alpha)^2Y^{n+1} + \alpha(1-\alpha)(X^{n+1}+Y^{n+1})\\
&= \alpha X^{n+1} + (1-\alpha)Y^{n+1}.
\end{aligned}$$

The last step is true since if $X \leq Y$ then $X^n \leq Y^n$ and as a result

$$\begin{aligned}
&0 \leq (X-Y)(X^n-Y^n) => \\
&X^nY + Y^nX \leq X^{n+1} + Y^{n+1}.
\end{aligned}$$

Since the inequality holds for $k=n+1$, the proof is complete by induction. This weighted average can be generalized to the expectation of matrices since the expectation is simply the weighted average of the outcomes of a random variable, where the weights are based on the probabilities of those outcomes. ■

From lemma 2 we can easily obtain a matrix inequality, similar to Jenson's inequality for the scalar power function, which is with respect to the non-negative matrix cone

$$(\mathbb{E}(W))^n \leq \mathbb{E}(W^n). \tag{4}$$

Having defined a martingale and a submartingale process in section II and using the convexity of power function for the space of matrices in lemma 2, now we state a lemma for the composition of a convex function and a martingale process. The proof of this lemma can be found in [17].

Lemma 3: If $\{W(k)\}$ is a martingale sequence of matrices with respect to $\mathcal{F}_k$ and φ is a convex function with respect to the non-negative matrix cone and $\mathbb{E}|\varphi(W_{ij}(k))| < \infty$ for all k, then the sequence of matrices $\{\varphi(W(k))\}$ is a submartingale with respect to $\mathcal{F}_k$.

Note that if a matrix sequence is stochastic, then it is bounded and therefore satisfies the requirement of lemma 3.

Now in the next theorem (the proof can be found in [17]), we show that if a sequence of stochastic matrices is a martingale (or more generally submartingale) process then it converges to a random limit almost surely.

Theorem 1: If $\{W(k)\}$ is a (sub)martingale sequence of matrices with $\sup_k\mathbb{E}(W_{ij}(k)) < \infty$, then as $k\to\infty$, $W(k)$ converges almost surely to a random limit matrix W with $\mathbb{E}|W_{ij}| < \infty$.

Lemma 2, lemma 3 and theorem 1 imply the following corollary.

Corollary 1: A sequence of power of martingale stochastic matrices, i.e. $\{W^n(k)\}$, is bounded and submartingale, therefore, converges to a random limit matrix almost surely.

B. Directed Spanning Tree in the Expected Graph and in the Random Limit

In this subsection, we show that if the expected graph of a martingale process contains a directed spanning tree, then the limit of the process contains a directed spanning tree almost surely as well. For this purpose, we first introduce some definitions which we will use later in this subsection.

Definition 5: A directed graph is called strongly connected if there is a path from each vertex in the graph to every other vertex. A matrix W is irreducible if and only if its associated graph $G(W)$ is strongly connected.

Definition 6: Let W be a non-negative matrix. For an index i the period of i is the greatest common divisor of all natural numbers k such that $(W^k)_{ii} > 0$. When W is irreducible, the period of every index is the same and is called the period of W. If the period is 1, the irreducible matrix W is aperiodic. It is obvious that if the irreducible matrix W has positive diagonals, i.e. the strongly connected graph $G(W)$ has self loops in all vertices, then the matrix W is aperiodic.

Definition 7: A matrix W is primitive if it is non-negative and its n-th power is positive ($W^n > 0$) for some natural number n. Note that it can be proved that primitive matrices are the same as irreducible aperiodic non-negative matrices. In the next theorem (refer to [16] for its proof) we show the effect of the primitivity of a matrix on its largest eigenvalue.

Perron Frobenius Theorem: A stochastic, primitive (irreducible with positive diagonals) matrix W has simple leading eigenvalue of $\lambda_1(W) = 1$. Equivalently, $\lambda_2(W) < 1$ where $\lambda_2(W)$ is the eigenvalue with the second largest modulus.

Definition 8: A sequence of random matrices $\{W(k)\}$ is uniformly integrable if for each ij-th entry of matrices

$$\lim_{M\to\infty}(\sup_k \mathbb{E}(|W_{ij}(k)|; |W_{ij}(k)| > M)) = 0.$$

In other words, a sequence of random matrices is uniformly integrable if its expected value has no mass in infinity. As an example, a collection of submartingale stochastic matrices, which are obviously bounded, are uniformly integrable.

Now we define the convergence in ℓ_1 norm which is different from almost sure convergence.

Definition 9: A sequence of random matrices $\{W(k)\}$ converges to a random matrix W in ℓ_1 norm if for each ij-th entry of matrices we have

$$\lim_{k\to\infty} \mathbb{E}(|W_{ij}(k) - W_{ij}|) = 0.$$

Introducing basic definitions for our analysis, we state a lemma (the proof can be found in [17]) relating the uniform integrability (e.g. a bounded process) and convergence in ℓ_1.

Lemma 4: For a submartingale sequence of matrices $\{W(k)\}$, uniform integrability and convergence in ℓ_1 are equivalent.

Remark 1: If we have almost sure convergence for a bounded process, then convergence in ℓ_1 can be concluded from dominated convergence theorem as well.

Dominated Convergence Theorem: If $X_k \to X$ a.s., $|X_k| \le Y$ for all k, and $\mathbb{E}(Y) < \infty$, then $\mathbb{E}(X_k) \to \mathbb{E}(X)$. Therefore, if we have a submartingale sequence of stochastic matrices (which is obviously bounded), from theorem 1 we know that the sequence converges almost surely, and as a result of dominated convergence theorem the convergence in ℓ_1 occurs as well. However, if we don't have uniform integrability (or boundedness), we may not have convergence in ℓ_1, even if almost sure convergence occurs. Constructing all necessary steps, we state a theorem for our main result in this subsection.

Theorem 2: For a martingale sequence of non-negative, stochastic matrices $\{W(k)\}$, if $\lambda_2(\mathbb{E}(W(k))) < 1$ (the expected graph contains a directed spanning tree), then $\lambda_2(W) < 1$ almost surely, where $W = \lim_{k\to\infty} W(k)$.

Proof: We prove for both cases where $\mathbb{E}(W(k))$ is irreducible and reducible. First suppose $\mathbb{E}(W(k))$ is irreducible. Since $W(k)$ has positive diagonals, $\mathbb{E}(W(k))$ is primitive. Therefore, $(\mathbb{E}(W(k)))^n > 0$. We now prove by contradiction that $W = \lim_{k\to\infty} W(k)$ is primitive as well. Let us define $Z(k) \triangleq W^n(k)$. Note that $W(k) \ge 0$, hence, $Z(k) \ge 0$. Corollary 1 implies that the sequence $\{Z(k)\}$ is submartingale and converges to the random limit $Z = W^n = \lim_{k\to\infty} W^n(k)$ almost surely. Also, since stochastic matrices are bounded and as a result uniformly integrable, by lemma 4 the submartingale sequence $\{Z(k)\}$ also converges in ℓ_1. Now let us assume $Z = W^n > 0$ does not hold. Therefore, for some ij-th entry we have $Z_{ij} = 0$. From convergence in ℓ_1 for the sequence $\{Z(k)\}$ we have

$$\begin{aligned}&\lim_{k\to\infty} \mathbb{E}(|Z_{ij}(k) - Z_{ij}|) = \\ &\lim_{k\to\infty} \mathbb{E}(|Z_{ij}(k)|) = \lim_{k\to\infty} \mathbb{E}(Z_{ij}(k)) = 0.\end{aligned} \tag{5}$$

However, if we have $(\mathbb{E}(W(k)))^n > 0$, then from equation (4) we have

$$\mathbb{E}(Z_{ij}(k)) = \mathbb{E}((W^n(k))_{ij}) \ge (\mathbb{E}(W(k)))^n_{ij} > 0, \tag{6}$$

and since the expectation is preserved in a martingale process, by taking the limit we have

$$\lim_{k\to\infty} \mathbb{E}(Z_{ij}(k)) > 0,$$

which is a contradiction with equation (5). Therefore, W is primitive and Perron Frobenius theorem implies that $\lambda_2(W) < 1$ almost surely. This proves the result for the case when $\mathbb{E}(W(k))$ is irreducible.

Now suppose $\mathbb{E}(W(k))$ is reducible. Without the loss of generality, one can label the vertices such that $\mathbb{E}(W(k))$ gets the following block triangular form

$$\mathbb{E}(W(k)) = \begin{pmatrix} Q_{11} & 0 & \cdots & 0 \\ Q_{21} & Q_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ Q_{s1} & Q_{s2} & \cdots & Q_{ss} \end{pmatrix} \tag{7}$$

where each Q_{ii} is an irreducible matrix and represents the vertices in the i-th communication class of $\mathbb{E}(W(k))$. Since $\lambda_2(\mathbb{E}(W(k)))$ is subunit, lemma 1 from section III implies that $G(\mathbb{E}(W(k)))$ has exactly one initial class. Assume α_1 (the class corresponding to submatrix Q_{11}) is the only initial class of $G(\mathbb{E}(W(k)))$. Therefore, there exists a directed path from a vertex in α_1 (e.g., say, vertex labeled 1) to any vertex of $G(\mathbb{E}(W(k)))$, such that the length of the path is at most some positive integer $m < n$. In other words, any vertex of $G(\mathbb{E}(W(k)))$ 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. Since equation (6) holds element-wise, it can be seen that the first column of $Z = W^n = \lim_{k\to\infty} W^n(k)$ should be strictly positive as well. Therefore, $\sigma(W^n) = 1 - \nu(W^n) = \max_j(\min_i W^n_{ij}) > \epsilon$. This and equation (2) implies that $\kappa(W^n) < 1 - \epsilon$. Hence, from equation (3) it follows that $\kappa(W^{kn}) \le \kappa^k(W^n) < (1-\epsilon)^k$, and as a result, $\lim_{k\to\infty} \kappa(W^k) = 0$. Since $\kappa(.)$ is a proper coefficient of ergodicity, we have $\lim_{k\to\infty} W^k = \mathbf{1}_n d^T$. Therefore, for all $x \in \mathbb{R}^n$

$$\lim_{k\to\infty} W^k x = (d^T x)\mathbf{1}_n. \tag{8}$$

Now if we represent x on the basis of eigenvectors of W we get

$$Wx = W(\alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_n v_n) =$$
$$\alpha_1 \lambda_1(W) v_1 + \alpha_2 \lambda_2(W) v_2 + \cdots \alpha_n \lambda_n(W) v_n.$$

Using the fact that for the stochastic matrix W, $\lambda_1(W) = 1$ and $v_1 = \mathbf{1}_n$ and employing equation (8) we obtain

$$(d^T x)\mathbf{1}_n = \lim_{k\to\infty} W^k x = \lim_{k\to\infty} (\alpha_1 \lambda_1^k(W) v_1 + \alpha_2 \lambda_2^k(W) v_2 + \cdots \alpha_n \lambda_n^k(W) v_n) = \alpha_1 \mathbf{1}_n + \lim_{k\to\infty} (\alpha_2(W) \lambda_2^k(W) v_2 + \cdots \alpha_n \lambda_n^k(W) v_n).$$

Therefore, $\lim_{k\to\infty} \lambda_2^k(W) = 0$. As a result, $|\lambda_2(W)| < 1$ almost surely. This proves the result for the case when $\mathbb{E}(W(k))$ is reducible. ■

C. Directed Spanning Tree in the Random Limit and Consensus

In this subsection, we show that a dynamical system reaches consensus almost surely if and only if it contains a directed spanning tree in expectation. We discuss both irreducible and reducible cases of the expected graph. In order to obtain this result, we use the following theorem which relates the consensus in a dynamical system to the paracontraction of the limit of the process.

Definition 10: Let $||.||$ denote a vector norm in $\mathbb{R}^n$. A $n \times n$ matrix W is nonexpansive with respect to $||.||$ if for all $x \in \mathbb{R}^n$,

$$||Wx|| \le ||x||.$$

W is called paracontracting with resect to $||.||$ if for all $x \in \mathbb{R}^n$,

$$Wx \neq x \Leftrightarrow ||Wx|| < ||x||.$$

It can be seen that any stochastic matrix W is nonexpansive. Also if $\lambda_2(W) < 1$, the stochastic matrix W is paracontracting as well.

Lemma 5: Let $\{W(k)\}$ be a sequence of matrices which are nonexpansive with respect to the same vector norm. If there exists a subsequence of $\{W(k)\}$ converging to a limit matrix W and W is paracontracting with respect to some norm, then $x(k)$ defined by (1) reaches consensus.

The proof of this lemma can be found in [18].

Exploiting the results of the preceding three subsections, we state the main result of this paper for consensus in martingale graph processes in the following theorem.

Theorem 3: For a martingale sequence of non-negative, stochastic matrices $\{W(k)\}$, with positive diagonals, $x(k)$ defined by equation (1) reaches consensus almost surely, if and only if $\lambda_2(\mathbb{E}(W(k))) < 1$ (the expected graph of the network contains a directed spanning tree).

Proof: Assume that $\lambda_2(\mathbb{E}(W(k))) = 1$. Since all $w_k \in \Omega_0$ 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 Perron Frobenius theorem $\lambda_2(\mathbb{E}(W(k))) < 1$, 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, it can be written as (7), where all Q_{ii} are irreducible matrices. Since $\lambda_2(\mathbb{E}(W(k))) = 1$, submatrices corresponding to at least two classes of $\mathbb{E}(W(k))$ have unit spectral radius (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 1). Therefore, lemma 1 implies

$$\exists i \neq j \quad \text{such that } \alpha_i \text{ and } \alpha_j \text{ are both initial classes}$$

or equivalently, $Q_{ir} = 0$ for all $r \neq i$ and $Q_{jl} = 0$ for all $j \neq l$. In other words, the 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 $\prod_{m=1}^{k} W(m)$ 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.

Now we prove the reverse implication. Since $\lambda_2(\mathbb{E}(W(k)))$ is subunit, theorem 2 implies that $\lambda_2(W) < 1$. Therefore, W is paracontracting with respect to some norm. From lemma 5 we can see that the dynamical system defined by equation (1) reaches consensus almost surely. ■

V. A Martingale Model of Social Network Formation

In our previous work in [1], we showed consensus happens in a Polya Urn model of social networks where at each time step, each agent takes an average of her belief and the belief of her neighbor. This model is an interesting model of social network formation, however, it requires that at each time step the current link associated to each agent is removed and is replaced by another link, therefore, $\Delta W_{ij}(k) = 0$ for all k. We can relax this requirement to a constraint in which the number of links associated to each agent is not constant over time, however, its expectation with respect to the available information so far is constant, i.e. $\mathbb{E}(\Delta W_{ij}(k)|\mathcal{F}_{k-1}) = 0$ for all k. Also, in the Polya Urn model all agents form a link with some other agent at each time step. This condition can also be generalized in a model where only some subset of agents form a link with some other subset of agents. Another assumption in the Polya Urn model which can be relaxed is the probability of link formation with each agent. In the Polya Urn model this probability is proportional to the assigned weight of that agent, which makes the model possible to analyze by the exchangeability property of the resulting process. A more general model would be the one which allows arbitrary probability of link formation with each agent at each time step.

It can be seen that the Polya Urn process, besides being an exchangeable process, is in fact a martingale process. From this observation and previous results of this paper, a more generalized version of the Polya Urn model, which is still a martingale process, should reach consensus as well. For this purpose, we build a martingale model of social networks which contains the Polya-Urn model as a special case. Let us consider the social network $G = (V, E)$ with a fixed set of agents, $V = \{1, \ldots, n\}$, and directed edges between them. Let $W_{ij}(k)$ be the weight that agent i assigns to agent j at time k. Let $\bar{W}_i(k) = (\bar{W}_{i1}(k), \ldots, \bar{W}_{in}(k))$ be the normalized weight vector, where

$$\bar{W}_{ij}(k) = \frac{W_{ij}(k)}{\sum_l W_{il}(k)}.$$

At each time step, a subset of agents form links with some other subset of agents according to some arbitrary probabilities of $p_{ij}(k)$. In contrast to the Polya Urn model, these probabilities do not need to be proportional to the assigned weights of the agents with whom links are formed. The link formed with each agent is independent of the link formed with any other agent. While link formation each agent increases or decreases the weight associated with the agents she formed a link with by $\Delta W_{ij}(k+1)$ conditioned on $\mathbb{E}(\Delta W_{ij}(k+1)|\mathcal{F}_k) = 0$ and $W_{ij}(k+1) \geq 0$. In other words

$$W_{ij}(k+1) = W_{ij}(k) + \Delta W_{ij}(k+1).$$

In the meantime, agents update their beliefs as a convex combination of the beliefs of their neighbors and their beliefs at the previous time step

$$x(t+1) = \bar{W}(t)x(t).$$

In other words

$$x_i(t+1) = \sum_{j \in \mathcal{N}_i(t)} \bar{W}_{ij}(t)x_j(t). \tag{9}$$

where $\mathcal{N}_i(t)$ is the set of agent i and her neighbors at time t. From the definition of a martingale process we know that a network is martingale if and only if the "normalized weight matrix" i.e. $\bar{W}(k)$, is martingale or if we have

$$\mathbb{E}(\bar{W}_{ij}(k+1)|\mathcal{F}_k) = \bar{W}_{ij}(k).$$

We now show that for the network to be martingale it suffices that the "unnormalized weight matrix" i.e. $W(k)$, to be martingale or

$$\mathbb{E}(W_{ij}(k+1)|\mathcal{F}_k) = W_{ij}(k)$$

which is equal to

$$\mathbb{E}(\Delta W_{ij}(k+1)|\mathcal{F}_k) = 0.$$

If $\mathbb{E}(\Delta W_{ij}(k+1)|\mathcal{F}_k) = 0$ we have

$$\begin{aligned}
&\mathbb{E}(\bar{W}_{ij}(k+1)|\mathcal{F}_k) = \\
&\mathbb{E}[p_{ij}(k)(\frac{W_{ij}(k) + \Delta W_{ij}(k+1)}{\sum_l (W_{il}(k) + \Delta W_{il}(k+1))})|\mathcal{F}_k] + \\
&\mathbb{E}[(1-p_{ij}(k))(\frac{W_{ij}(k)}{\sum_l W_{il}(k)})|\mathcal{F}_k] = \\
&p_{ij}(k)(\frac{W_{ij}(k) + \mathbb{E}(\Delta W_{ij}(k+1)|\mathcal{F}_k)}{\sum_l (W_{il}(k) + \mathbb{E}(\Delta W_{il}(k+1)|\mathcal{F}_k))} + \\
&(1-p_{ij}(k))(\frac{W_{ij}(k)}{\sum_l W_{il}(k)}) = \frac{W_{ij}(k)}{\sum_l W_{il}(k)} = \bar{W}_{ij}(k).
\end{aligned}$$

Therefore, instead of $\mathbb{E}(\Delta \bar{W}_{ij}(k+1)|\mathcal{F}_k) = 0$, if $\mathbb{E}(\Delta W_{ij}(k+1)|\mathcal{F}_k) = 0$ the resulting social network, i.e. $\bar{W}(k)$, would be martingale and if it contains a directed spanning tree in expectation as well, then we can use the results from section IV to show that consensus happens in the distributed update scheme (9) almost surely. Also note that in the Polya Urn model the unnormalized weight matrix, $W(k)$, is initialized as a connected graph, i.e. $W(0) = \mathbf{1}_n, \mathbf{1}_n^T - I_n$ and as a result of the process it will be a complete graph for all time steps almost surely. Therefore, the expectation of the social network, i.e. $\mathbb{E}(\bar{W}(k))$, is a complete graph for all time steps as well. This condition is relaxed to a directed spanning tree in the expectation of the social network in this proposed model. This is in fact the main requirement for consensus in martingale random networks.

VI. Conclusion

In this paper, we provided a necessary and sufficient condition for almost sure convergence of consensus algorithms over general weighted and directed martingale random graph processes. We showed that linear dynamical system $x(k+1) = W(k)x(k)$ reaches consensus almost surely if $\mathbb{E}(W(k)$ has exactly one eigenvalue with unit modulus or if the expected graph of the network contains a directed spanning tree. We also constructed a model of social network formation which was the generalization of a Polya Urn model of social network formation proposed in [1]. We showed a consensus seeking process based on local averaging of opinions in this model converges to a random consensus value almost surely. A direction for future research is to find closed form formula for the statistics of this random consensus value as in the case of an i.i.d. graph process.

References

[1]
[2] J. Tsitsiklis, "Problems in decentralized decision making and computation," *Ph.D. dissertation*, 1984.
[3] A. Jadbabaie, J. Lin, and A. 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.
[4] 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, 2004.
[5] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Transactions on Automatic Control*, vol. 50, no. 2, pp. 169–182, 2005.
[6] J. Cortes, S. Martinez, and F. Bullo, "Analysis and design tools for distributed motion coordination," in *American Control Conference, 2005. Proceedings of the 2005*, pp. 1680–1685, 2005.
[7] M. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.
[8] B. Golub and M. Jackson, "Naive Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.
[9] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," in *Information Theory and Applications Workshop, 2007*, pp. 150–159, 2007.
[10] Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 1867–1872, 2005.
[11] C. Wu, "Synchronization and convergence of linear dynamics in random directed networks," *IEEE Transactions on Automatic Control*, vol. 51, no. 7, pp. 1207–1210, 2006.
[12] M. Porfiri and D. Stilwell, "Consensus seeking over random weighted directed graphs," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1767–1773, 2007.
[13] A. Tahbaz-Salehi and A. Jadbabaie, "A necessary and sufficient condition for consensus over random networks," *IEEE Transactions on Automatic Control*, vol. 53, no. 3, pp. 791–795, 2008.
[14] A. Tahbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Transactions on automatic Control*, vol. 55, no. 1, pp. 225–230, 2010.
[15] M. Ion and B. John, "Convergence results for the linear consensus problem under Markovian random graphs," 2009.
[16] A. Berman and R. Plemmons, *Nonnegative matrices in the mathematical sciences*, vol. 9. Society for Industrial Mathematics, 1994.
[17] R. Durrett and R. Durrett, *Probability: theory and examples*. Cambridge Univ Pr, 2010.
[18] R. Bru, L. Elsner, and M. Neumann, "Convergence of infinite products of matrices and inner-outer iteration schemes," *Electron. Trans. Numer. Anal.*, vol. 2, no. 11, p. 183193, 1994.