Products of Stochastic Matrices: Exact Rate for Convergence in Probability for Directed Networks

Dragana Bajović, *Student Member, IEEE,* João Xavier, *Member, IEEE,* and Bruno Sinopoli, *Member, IEEE*

Abstract—We study the products $W_k \cdots W_1$ of random stochastic, not necessarily symmetric matrices. It is known that, under certain conditions, the product $W_k \cdots W_1$ converges almost surely (a.s.) to a random rank-one matrix; the latter is equivalent to $|\lambda_2(W_k \cdots W_1)| \to 0$ a.s., where $\lambda_2(\cdot)$ is the second largest (in modulus) eigenvalue. In this paper, we show that the probability that $|\lambda_2(W_k \cdots W_1)|$ stays above $\epsilon \in (0,1]$ in the long run decays to zero exponentially fast $\sim e^{-k\,I}$. Furthermore, we explicitly characterize the rate of this convergence I and show that it depends only on the underlying graphs that support the matrices W_k's. Our results reveal that the rate I is essentially determined by the most likely way in which the union (over time) of the support graphs fails to form a directed tree.

Keywords— Consensus, Convergence in probability, Directed networks, Exponential rate, Stochastic matrices.

I. INTRODUCTION

Many applications in distributed systems, such as rendezvous of robots or distributed inference in sensor networks, have natural representation as the linear systems driven by stochastic[1] matrices $x_{k+1} = W_k x_k$, where x_k is the system state at time k, and W_k is the system (possibly time varying or random) stochastic matrix; see, e.g., [1] for distributed inference in sensor networks. The matrices W_k respect the sparsity pattern of the underlying communication network; in many applications, like with wireless sensor networks, the W_k's are random, due to packet dropouts that occur at random times. Also, physical communication links may be asymmetric, which necessitates the study of the systems with non-symmetric W_k's.

The key in analyzing the performance of such systems is in the study of the long term behavior of the products $W_k \cdots W_1$. In deterministic time invariant systems, where $W_k = A$ for all k (and A is a stochastic matrix), it is well known that $W_k \cdots W_1$ converges to a (deterministic) rank one matrix $1v^\top$, if $|\lambda_2(A)|$ – the modulus of the second largest (in magnitude) eigenvalue of A is strictly less than 1 [2]. It is also known that the convergence of $W_k \cdots W_1 - 1v^\top$ to zero occurs at a geometric rate (in k), equal to $|\lambda_2(A)|$. In the time varying case, under the assumption that the unions of graphs over finite time windows are connected (bounded intercommunication intervals), the rate of convergence is also geometric [3]. In contrast, with random W_k's, a geometric rate of convergence is no longer guaranteed for every sample path. In this paper, we are interested in finding the large deviation probabilities that the convergence rate is sub-exponential. With the random W_k's, it is known that the product $W_k \cdots W_1$ converges almost surely (a.s.) to a random, rank one matrix $1v^\top$ (the vector v is random) [4], when the W_k's have positive diagonals and $|\lambda_2(\mathbb{E}[W_k])|$ is strictly less than one. Further, the path-wise convergence of $W_k \cdots W_1$ to $1v^\top$ is equivalent to the path-wise convergence of $|\lambda_2(W_k \cdots W_1)|$ to zero.

We consider independent and identically distributed (i.i.d.) asymmetric stochastic matrices W_k with positive diagonals. As a measure of the size of the residual error $W_k \cdots W_1 - 1v^\top$, we naturally adopt $|\lambda_2(W_k \cdots W_1)|$. We show that the probability that $|\lambda_2(W_k \cdots W_1)|$ converges to zero sub-exponentially, i.e., the probability that $|\lambda_2(W_k \cdots W_1)| \geq e^{-\zeta_k}$, for a sequence $0 \leq \zeta_k = o(k)$, decays to zero exponentially fast $\sim e^{-k\,I}$. Furthermore, we solve with equality (rather than with upper and lower bounds) the large deviation limit:

$$\lim_{k \to +\infty} \frac{1}{k} \log \mathbb{P}\left(|\lambda_2(W_k \cdots W_1)| \geq e^{-\zeta_k}\right) = -I, \quad (1)$$

where $I \geq 0$. Note that a special case is to take $e^{-\zeta_k} = \epsilon, \forall k$, with $\epsilon \in (0,1]$. Interestingly, our results reveal that I depends on the distribution of matrices W_k only through the distribution of their induced graphs. Intuitively, we show that the rate I is determined by the probability of the most likely way in which the union (over time) of the graphs that support the W_k's does not contain a directed tree. (See Theorem 4 for a precise statement of this result.)

We further calculate the rate I for a commonly used broadcast gossip protocol [5] in sensor networks, where (only one) node u activates at a time with probability p_u, and broadcasts its state to all single-hop neighbors. The rate $I = |\log 1 - p_{\min}|$, where $p_{\min}$ is the probability of the most rarely active node.

We now distinguish the current paper from our previous work on the products of stochastic matrices [6], [7]. In [6],

Corresponding author Dragana Bajović is with the Institute for Systems and Robotics (ISR), Instituto Superior Técnico (IST), Lisbon, Portugal, and with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA dragana@isr.ist.utl.pt, dbajovic@andrew.cmu.edu

João Xavier is with the Institute for Systems and Robotics (ISR), Instituto Superior Técnico (IST), Lisbon, Portugal jxavier@isr.ist.utl.pt

Bruno Sinopoli is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA brunos@ece.cmu.edu

[1]The rows of the matrix sum to one and all the entries are nonnegative.

[7], we studied the case when W_k's are both stochastic and *symmetric* and we derived a counterpart of the limit in (1). With respect to [6], [7], here we significantly extend our result, as we remove the assumption that the W_k's need to be symmetric, thus allowing for directed underlying networks. Directed networks are highly relevant in practice, as practical communication links can be asymmetric: a directed link from i to j can be strong enough to support communication, but the link from j to i may not be strong enough. Mathematically, the analysis with directed networks (asymmetric W_k's) introduces two new major challenges. First, we no longer have convergence of the product $W_k \cdots W_1$ to the deterministic consensus matrix $J = \frac{1}{N} 11^\top$ as in the symmetric matrices case; rather, the product converges to a random rank one matrix $1 v^\top$. Second, the error matrix was of the form $W_k \cdots W_1 - J$, which allowed for the following factorization $W_k \cdots W_1 - J = (W_k - J) \cdots (W_1 - J)$. With directed networks, we no longer have the deterministic limit J nor the factorization.

The remainder of the paper is organized as follows. Section II describes the model of random matrices W_k that we assume, introduces certain concepts needed for subsequent analysis, and provides some preliminary results. Section III states our main result on the exponential decay rate I, and gives the key steps for the proof of the result. Section IV illustrates with examples the computation of the rate I. Finally, Section V concludes the paper and provides directions for future research.

Throughout the paper we use the following notation. We denote by: A_{ij} or $[A]_{ij}$ the entry in i-th row and j-th column of a matrix A; 1 the vector with unit entries; $J = \frac{1}{N} 11^\top$ the ideal consensus matrix; $\mathbb{S}^N$ the set of stochastic N by N matrices; $\mathbb{G}^N$ the set of all directed graphs (without self-loops) on the set of vertices $V = \{1, ..., N\}$; and $\mathbb{E}$ the expectation operator.

II. PROBLEM SETUP

Let $\{W_k\}_{k \geq 1}$ be a sequence of random matrices defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in $\mathbb{S}^N$. We assume that W_k, $k \geq 1$, are i.i.d. and have almost surely positive diagonals; that is, we assume that, for every i, $[W_k]_{ii} > 0$ with probability one. The last assumption assures that, if an entry ij of the product $W_k \cdots W_1$ becomes positive at some time k, then, due to positivity of the diagonals, this entry remains positive for all $k' \geq k$; or, intuitively, it guarantees continuous flow of information. In the sequel, we denote by $\Phi(k, 0) = W_k \cdots W_1$.

Let $G : \mathbb{S}^N \mapsto \mathbb{G}^N$ denote the induced graph map defined by $G(W) = (V, \{\{i, j\} : W_{ij} > 0, i \neq j\})$. We define the sequence of random directed graphs G_k, $k \geq 1$, through the sequence of random matrices W_k, $k \geq 1$, by $G_k = G(W_k)$, for $k \geq 1$. Remark that the sequence G_k, $k \geq 1$, is i.i.d., as the matrix sequence is i.i.d. For a graph $H \in \mathbb{G}^N$, we denote with p_H the probability that the realization of the induced graph equals H, $p_H = \mathbb{P}(G(W_k) = H)$. Similarly, for a collection of graphs $\mathcal{H} \subseteq \mathbb{G}^N$, we let $p_{\mathcal{H}} = \mathbb{P}(G(W_k) \in \mathcal{H})(= \sum_{H \in \mathcal{H}} p_H)$. We collect all realizations of the induced graphs that occur with positive probability in the set $\mathcal{G}$:

$$\mathcal{G} := \{H \in \mathbb{G}^N : p_H > 0\}. \tag{2}$$

A. Accumulation graph $\Gamma(k, 0)$ and tree-free collections

For a collection of graphs $\mathcal{H} \subseteq \mathbb{G}^N$, let $\Gamma(\mathcal{H})$ denote the graph that contains all the edges of all the graphs in $\mathcal{H}$, $\Gamma(\mathcal{H}) = (V, \cup_{H \in \mathcal{H}} E(H))$, where $E(H)$ denotes the set of edges of a graph H.

An object that will be particularly relevant in our analysis is the graph that collects all the links that appeared from time 1 until some time k. We refer to it as the accumulation graph from time 1 until time k and formally define it as

$$\Gamma(k, 0) = (V, \cup_{t=1}^k E(G_t)). \tag{3}$$

We study connectivity properties of $\Gamma(k, 0)$ using the concept of strong components of directed graphs [8], [2]. Two nodes i, j in a directed graph H are said to communicate if H contains both a directed path from i to j and a directed path from j to i. Communication relation is an equivalence relation and it induces the equivalence classes on H called the communication classes of H. If a communication class of H has only outgoing edges, this class is called initial.

We now give the properties that shed light on the relation between the accumulation graph $\Gamma(k, 0)$ and the product matrix $\Phi(k, 0)$. The first step is a simple, but important relation between $\Gamma(k, 0)$ and the induced graph of $\Phi(k, 0)$ which we state without proof.

Observation 1 Two nodes communicate in $\Gamma(k, 0)$ if and only if they communicate in $G(\Phi(k, 0))$.

Note that the induced graph of the product matrix, which contains both one-hop information flows and their superpositions in time, contains in general more links than the graph $\Gamma(k, 0)$ that registers only one-hop information flows. However, Observation 1 assures that the communication classes of these two graphs are nevertheless the same. We use this observation to derive the key relation between the product matrix $\Phi(k, 0)$ and the accumulation graph $\Gamma(k, 0)$ which we state in Lemma 2.

Lemma 2 $|\lambda_2(\Phi(k, 0))| < 1$ if and only if $\Gamma(k, 0)$ contains a directed tree.

Proof: We use (without proof, which can be derived from Lemma (3.20) on page 224, in [8]) the fact that, for every stochastic matrix W with positive diagonals, $|\lambda_2(W)| < 1$ if and only if $G(W)$ has exactly one initial class. Combining this with Observation 1, it follows that $|\lambda_2(\Phi(k, 0))| < 1$ if and only $\Gamma(k, 0)$ has exactly one initial class. Finally, the last condition is equivalent to the condition that $\Gamma(k, 0)$ contains a directed tree. ∎

We say that a collection of directed graphs $\mathcal{H} \subseteq \mathcal{G}$ is a tree-free collection on $\mathcal{G}$ if the accumulation graph $\Gamma(\mathcal{H})$ does not contain a directed tree. Denote with $\Pi(\mathcal{G})$ the set that collects all tree-free collections on $\mathcal{G}$:

$$\Pi(\mathcal{G}) = \{\mathcal{H} \subseteq \mathcal{G} : \Gamma(\mathcal{H}) \text{ is tree} - \text{free}\}. \tag{4}$$

Next subsection illustrates the introduced concepts of accumulation graph and tree-free collection on the example of broadcast gossip algorithm.

B. Example: broadcast gossip

Let $\widehat{G} = (V, \widehat{E})$ be a directed graph that collects all the available communication links in a network. At each time k, a node is chosen at random in V according to the probability mass function $p_u > 0$, $u \in V$, $\sum_{u \in V} p_u = 1$. Denote by u_k the node chosen at time k. With broadcast gossip [5], the weight matrix W_k has the sparsity pattern of a directed star graph centered at u_k and is given by: $[W_k]_{u_k v} = [W_k]_{vv} = 1/2$, for all v such that $\{u_k, v\} \in \widehat{E}$ (out-neighbors of u_k), $[W_k]_{vv} = 1$ otherwise, and the rest of the entries are zero. The graph realization at time k is then $G_k = (V, \{(u_k, v) : v \in V, (u_k, v) \in \widehat{E}\})$. Since each node in V has a positive probability of being chosen, we conclude that the collection of realizable graphs with broadcast gossip is the collection of all star subgraphs of $\widehat{G}$ centered at its nodes:

$$\mathcal{G}^{\mathrm{B-gossip}} = \{H_u : u \in V\}, \qquad (5)$$

where $H_u = \left(V, \{(u, v) : v \in V, (u, v) \in \widehat{E}\}\right)$.

For concreteness, we consider now a simple case when $\widehat{G}$ is a four node graph, $V = \{1, 2, 3, 4\}$, and with the set of edges $\widehat{E} = \{(1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 3)\}$, as shown in Figure 1a. We can see that, for node 1,
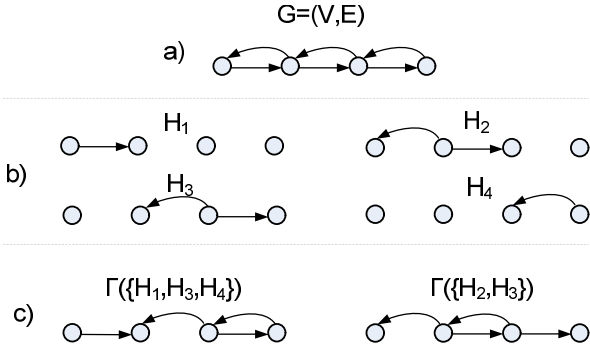

G=(V,E)

a)

b)

c)

$\Gamma(\{H_1, H_3, H_4\})$ $\Gamma(\{H_2, H_3\})$

Fig. 1. Example of a broadcast gossip on a 4-node chain; a) $\widehat{G} = (V, \widehat{E})$ is the total budget of communication links; b) $\mathcal{G} = \{H_1, H_2, H_3, H_4\}$ is the set of realizable graphs; c) $\mathcal{H} = \{H_1, H_3, H_4\}$ is a tree-free collection, whereas $\mathcal{H}' = \{H_2, H_3\}$ is not.

its only out-neighbor in $\widehat{G}$ is node 2 – thus, H_1 is the single arc graph $(V, (1, 2))$, as shown in Figure 1b. Similarly, node 2 has two out-neighbors in $\widehat{G}$, node 1 and node 3, and thus $H_2 = (V, \{(2, 1), (2, 3)\})$. Checking the out-neighborhoods of the remaining two nodes, 3 and 4, we conclude that the set of all realizable graphs is $\mathcal{G} = \{H_1, H_2, H_3, H_4\}$, with H_i as in Figure 1b, $i = 1, ..., 4$. To find the tree-free collections on $\mathcal{G}$, we notice first that the removal of the edges of H_2 and H_3 makes $\widehat{G}$ tree-free. Thus, any subset of $\mathcal{G} \setminus \{H_2\}$ is a tree-free collection, and similarly for $\mathcal{G} \setminus \{H_3\}$; for example, $\mathcal{H} = \{H_1, H_3, H_4\}$ shown in Figure 1c (left) is one such a collection (it is easy to see from Figure 1c that $\Gamma(\{H_1, H_3, H_4\})$ does not have a directed spanning

tree). On the other hand, simultaneous removal of edges of H_1 and H_4 "does no harm", as $\Gamma(\{H_2, H_3\})$ still contains a directed spanning tree (in fact, it contains two directed spanning trees, as can be seen from Figure 1c (right)). Summing up, the set of all tree-free collections on $\mathcal{G}$ is $\Pi(\mathcal{G}) = 2^{\mathcal{G} \setminus \{H_2\}} \cup 2^{\mathcal{G} \setminus \{H_3\}}$, where $2^{\mathcal{G} \setminus \{H_2\}} = \{\emptyset, \{H_1\}, \{H_3\}, \{H_4\}, \{H_1, H_3\}, \{H_1, H_4\}, \{H_3, H_4\}, \{H_1, H_3, H_4\}\}$ is the power set of $\mathcal{G} \setminus \{H_2\}$, and similarly for $\mathcal{G} \setminus \{H_3\}$.

III. MAIN RESULT

This section states the main result of this paper on the large deviation limit (1) for the sequence of random stochastic matrices, and it provides key steps of the proof.

Theorem 3 For any sequence of nonnegative numbers $\{\zeta_k\}_{k \geq 1}$ such that $\zeta_k = o(k)$

$$\lim_{k \to +\infty} \frac{1}{k} \log \mathbb{P}\left(|\lambda_2(\Phi(k, 0))| \geq e^{-\zeta_k}\right) = -I \qquad (6)$$

where

$$I = \begin{cases} |\log p_{\max}|, & \text{If } \Pi(\mathcal{G}) \neq \emptyset \\ +\infty, & \text{otherwise} \end{cases}$$

and $p_{\max} = \max_{\mathcal{H} \in \Pi(\mathcal{G})} p_{\mathcal{H}}$.

Theorem 3 can be proven by proving separately the lower and the upper large deviation bounds: $\liminf_{k \to \infty} \frac{1}{k} \log \mathbb{P}\left(|\lambda_2(\Phi(k, 0))| \geq e^{-\zeta_k}\right) \geq \log p_{\max}$ and $\limsup_{k \to \infty} \frac{1}{k} \log \mathbb{P}\left(|\lambda_2(\Phi(k, 0))| \geq e^{-\zeta_k}\right) \leq \log p_{\max}$. In the next subsection, we provide a complete proof for the lower bound. The proof of the upper bound is omitted due to lack of space and will be presented elsewhere.

A. Lower bound

If all induced graphs G_t from time $t = 1$ until time $t = k$ belong to some tree-free collection on $\mathcal{G}$, then it must be that $\Gamma(k, 0) = \Gamma(G_1, ..., G_k)$ is tree-free. In other words, for any $\mathcal{H} \in \Pi(\mathcal{G})$, the following inclusion relation holds between the two events:

$$\{G_1 \in \mathcal{H}, ..., G_k \in \mathcal{H}\} \subseteq \{\Gamma(k, 0) \text{ is tree} - \text{free}\}.$$

Computing now the probabilities of the events above and using the fact that G_t, $1 \leq t \leq k$, are i.i.d., it follows

$$p_{\mathcal{H}}^k \leq \mathbb{P}\left(\Gamma(k, 0) \text{ is tree} - \text{free}\right). \qquad (7)$$

Since equation (7) holds for every $\mathcal{H} \in \Pi(\mathcal{G})$, it also holds for the tree-free collection $\mathcal{H}^\star \in \Pi(\mathcal{G})$ that has the highest probability $p_{\mathcal{H}^\star} = p_{\max}$, and therefore:

$$p_{\max}^k \leq \mathbb{P}\left(\Gamma(k, 0) \text{ is tree} - \text{free}\right). \qquad (8)$$

To relate $p_{\max}^k$ and the probability of our event of interest $\{|\lambda_2(\Phi(k, 0))| \geq e^{-\zeta_k}\}$, we observe that, by the fact that ζ_k is nonnegative, we have:

$$\mathbb{P}\left(|\lambda_2(\Phi(k, 0))| \geq e^{-\zeta_k}\right) \geq \mathbb{P}\left(|\lambda_2(\Phi(k, 0))| = 1\right). \qquad (9)$$

Using now Lemma 2, the "only-if" part, to link (9) and (8) yields

$$p_{\max}^k \leq \mathbb{P}\left(|\lambda_2(\Phi(k, 0))| \geq e^{-\zeta_k}\right). \qquad (10)$$

Taking the $\log$, dividing by k and taking the $\liminf_{k \to +\infty}$, the large deviation lower bound follows.

IV. COMPUTATION OF I

In this section we illustrate computation of rate I with two random models: leader following with link failures and broadcast gossip.

A. Leader following with link failures

Let $T = (V, E)$ be a directed tree on N nodes and let r denote the root node in T. We assume that each link e in T may fail with probability $1 - p_e$ and also that links fail independently in time and in space. With leader following algorithm on T, every node in T at each time k transmits its state to all of its children; however, due to link failures, some if its children may not receive the sent information. The corresponding matrix of interactions W_k is then given by $[W_k]_{u t_u} = \alpha$ and $[W_k]_{uu} = 1 - \alpha$ if the node u received the information from its parent denoted by t_u (that is, if the link $(u, t_u) \in E$ was online at time k) and $[W_k]_{u t_u} = 0$ and $[W_k]_{uu} = 1$, otherwise; remark also that $[W_k]_{uv} = 0$ with probability one for all $v \neq t_u$. Using the fact that the links fail independently, we get that each subgraph of T occurs with positive probability and thus the collection of realizable graphs is the collection of all subgraphs of T:

$$\mathcal{G}^{\text{Leader}} = \left\{ (V, E') : E' \subseteq E \right\}. \tag{11}$$

Now, since $T = (V, E)$ is a directed tree, the easiest way to make T tree-free is to remove an arbitrary arc from T. Thus, a candidate for the most likely tree-free collection on $\mathcal{G}^{\text{Leader}}$ has to be of the following form: $\{(V, E') : E' \subseteq E \setminus e\}$, for some $e \in E$. The probability of such a collection equals $1 - p_e$, $e \in E$. Thus, the one with the highest probability is the one for which p_e is minimal. We conclude that

$$p_{\max}^{\text{Leader}} = 1 - p_{e^\star}, \quad I^{\text{Leader}} = |\log(1 - p_{e^\star})|,$$

where $e^\star$ is the "weakest" link in T that has the lowest probability of occurrence.

B. Broadcast gossip on a tree

In Subsection II-B we explained the broadcast gossip algorithm running on a generic network that is defined by graph $\widehat{G} = (V, \widehat{E})$. We consider now the case when $\widehat{G}$ is a symmetric tree (the undirected graph of $\widehat{G}$ is a tree and for every arc in $\widehat{G}$ its inverted arc also belongs to $\widehat{G}$). Note that the four node graph from the example in Figure 1 is of this type. Similarly as in the case of the four node graph, we can see that inhibition of any internal node (non-leaf node) $u \in V$ yields a tree-free network, that is, $\Gamma(\mathcal{G} \setminus \{H_u\})$ is tree-free. Thus, it suffices to remove just one graph of this type from $\mathcal{G}^{\text{B-gossip}}$, and therefore the most likely tree-free collection must be of the form $\mathcal{G} \setminus \{H_u\}$, where u is some internal node in $\widehat{G}$. The probability of such a collection is $1 - p_u$, and the most likely one is the one for which p_u is minimal:

$$p_{\max}^{\text{B-gossip}} = 1 - \min_{u \in V : u \text{ is internal node}} p_u$$

$$I^{\text{B-gossip}} = \left| \log p_{\max}^{\text{B-gossip}} \right|.$$

For the simplest case when all the nodes have the same probability of activation, equal to $\frac{1}{N}$, the rate $I^{\text{B-gossip}} = \left| \log \left(1 - \frac{1}{N} \right) \right| \sim \frac{1}{N}$, for large N.

V. CONCLUSION AND FUTURE WORK

In this paper, we studied the products $W_k \cdots W_1$ of random stochastic matrices W_k that are not necessarily symmetric. Asymmetric random matrices are highly relevant in applications, like with wireless sensor networks where asymmetric link failures may occur at random, and the study of their behavior is much more challenging than the study of symmetric matrices. It is well known that, when the diagonal entries of the W_k's are positive almost surely, and $|\lambda_2(\mathbb{E}[W_k])|$ – the modulus of the second largest (in modulus) eigenvalue of $\mathbb{E}[W_k]$ is strictly less than one, the product $W_k \cdots W_1$ converges to a rank-one random matrix, which is equivalent to $|\lambda_2(W_k \cdots W_1)|$ converging to zero. In this paper, we show that the probability that $|\lambda_2(W_k \cdots W_1)| \geq e^{-\zeta_k}$, with $0 \leq \zeta_k = o(k)$, converges to zero exponentially fast$\sim e^{-kI}$. Further, we explicitly characterize the rate I, and we show that I depends only on the graphs that support the W_k's. Intuitively, our result shows that I is determined by the most likely way in which the union of the support graphs over time does not contain a directed spanning tree.

As a future work, we plan to address the computation of I for generic random models (for example, the link failure model on a general directed graph, as opposed to a tree graph studied in Section IV). Guided by the previous findings on the relation of I with the minimum cut for the case of undirected networks [6], [7], we expect a qualitatively similar result for directed networks as well, but now, the minimum cut would be understood in the sense of "cutting" all the directed spanning trees.

REFERENCES

[1] D. Bajović, D. Jakovetić, J. Xavier, B. Sinopoli, and J. M. F. Moura, "Distributed detection via Gaussian running consensus: Large deviations asymptotic analysis," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4381–4396, Sep. 2011.

[2] E. Seneta, *Nonnegative Matrices and Markov Chains*. New York: Springer, 1981.

[3] A. Olshevsky and J. N. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM Rev.*, vol. 53, Nov. 2011.

[4] A. Tahbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 225–230, January 2010.

[5] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2748–2761, July 2009.

[6] D. Bajović, J. Xavier, J. M. F. Moura, and B. Sinopoli, "Consensus and products of random stochastic matrices: Exact rate for convergence in probability," *submitted to IEEE Transactions on Signal Processing*, Feb. 2012.

[7] D. Bajović, J. Xavier, J. M. F. Moura, and B. Sinopoli, "Exact rate for convergence in probability of averaging processes via generalized min-cut," in *Accepted for presentation at CDC '12, 51st IEEE Conference on Decision and Control*, Hawaii, USA, December 2012.

[8] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. NY: Academic Press, INC., 1970.