Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Consensus and synchronization in discrete-time networks of multi-agents with stochastically switching topologies and time delays

by

Wenlian Lu, Fatihcan M. Atay, and Jürgen Jost

Preprint no.: 30 2011

Manuscript submitted to
AIMS' Journals
Volume , Number **0**,

Website: http://AIMsciences.org

pp. **000–000**

CONSENSUS AND SYNCHRONIZATION IN DISCRETE-TIME NETWORKS OF MULTI-AGENTS WITH STOCHASTICALLY SWITCHING TOPOLOGIES AND TIME DELAYS

WENLIAN LU

Center for Computational Systems Biology, School of Mathematical Sciences,
Fudan University, Shanghai, China
and Max Planck Institute for Mathematics in the Sciences,
Inselstr. 22, 04103 Leipzig, Germany

FATIHCAN M. ATAY

Max Planck Institute for Mathematics in the Sciences,
Inselstr. 22, 04103 Leipzig, Germany

JÜRGEN JOST

Max Planck Institute for Mathematics in the Sciences,
Inselstr. 22, 04103 Leipzig, Germany
and Santa Fe Institute for the Sciences of Complexity
1399 Hyde Park Road, Santa Fe, NM 87501, USA

Preprint. Final version in *Networks and Heterogeneous Media*, 6(2):329–349, 2011.

ABSTRACT. We analyze stability of consensus algorithms in networks of multi-agents with time-varying topologies and delays. The topology and delays are modeled as induced by an adapted process and are rather general, including i.i.d. topology processes, asynchronous consensus algorithms, and Markovian jumping switching. In case the self-links are instantaneous, we prove that the network reaches consensus for all bounded delays if the graph corresponding to the conditional expectation of the coupling matrix sum across a finite time interval has a spanning tree almost surely. Moreover, when self-links are also delayed and when the delays satisfy certain integer patterns, we observe and prove that the algorithm may not reach consensus but instead synchronize at a periodic trajectory, whose period depends on the delay pattern. We also give a brief discussion on the dynamics in the absence of self-links.

1. **Introduction.**

Consensus problems have been recognized as important in distribution coordination of dynamic agent systems, which is widely applied in distributed computing [21], management science [5], flocking/swarming theory [32], distributed control [10], and sensor networks [26]. In these applications, the multi-agent systems need to agree on a common value for a certain quantity of interest that depends on the states of the interests of all agents or is a preassigned value. The interaction rule for each agent specifying the information communication between itself and its neighborhood is called the *consensus protocol/algorithm*. A related concept of consensus, namely *synchronization*, is considered as "coherence of different processes", and is a widely existing phenomenon in physics and

2000 *Mathematics Subject Classification.* Primary: 93C05, 37H10, 15A51, 40A20; Secondary: 05C50, 60J10.

Key words and phrases. Consensus, synchronization, delay, network of multi-agents, adapted process, switching topology.

biology. Synchronization of interacting systems has been one of the focal points in many research and application fields [33, 16, 29]. For more details on consensus and the relation between consensus and synchronization, the reader is referred to the survey paper [27] and the references therein.

A basic idea to solve the consensus problem is updating the current state of each agent by averaging the previous states of its neighborhood and its own. The question then is whether or under which circumstances the multi-agent system can reach consensus by the proposed algorithm. In the past decade, the stability analysis of consensus algorithms has attracted much attention in control theory and mathematics [27]. The core purpose of stability analysis is not only to obtain the algebraic conditions for consensus, but also to get the consensus properties of the topology of the network. The basic discrete-time consensus algorithm can be formulated as follows:

$$x_i^{t+1} = x_i^t + \epsilon \sum_{j\in\mathcal{N}_i} (x_j^t - x_i^t),\ i = 1,\dots,m, \tag{1}$$

where $x_i^t \in \mathbb{R}$ denotes the state variable of the agent i, t is the discrete-time, $\mathcal{N}_i$ denotes the neighborhood of the agent i, and ϵ is the coupling strength. Define $\mathcal{L} = [l_{ij}]_{i,j=1}^m$ as the Laplacian of the graph of the network in the manner that $l_{ij} = 1$ if $i \neq j$ and a link from j to i exists, $l_{ij} = 0$ if that $i \neq j$ and no link from j to i exists, and $l_{ii} = -\sum_{j\neq i} l_{ij}$. With $G = I - \epsilon\mathcal{L}$, (1) can be rewritten as

$$x^{t+1} = Gx^t, \tag{2}$$

where $x^t = [x_1^t,\dots,x_m^t]^\top$. If the diagonal elements in G are nonnegative, i.e., $0 \le \epsilon \le 1/\max_i l_{ii}$, then G is a stochastic matrix. Eq. (2) is a general model of the synchronous consensus algorithm on a network with fixed topology. The network can be a directed graph, for example, the leader-follower structure [22], and may have weights.

In many real-world applications, the connection structure may change in time, for instance when the agents are moving in physical space. One must then consider time-varying topologies under link failure or creation. The asynchronous consensus algorithm also indicates that the updating rule varies in time [9]. Thus, the consensus algorithm becomes

$$x^{t+1} = G(t)x^t, \tag{3}$$

where the time-varying coupling matrix $G(t)$ expresses to the time-varying topology. We associate $G(t)$ with a directed graph at time t (see Sec. 2), in which $G_{ij}(t) > 0$ implies that there is a link from j to i at time t, which may be a *self-link* if $i = j$. Note that the self links in G arise from the presence of the x_i on the right hand side of (1); they do not necessarily mean that the physical network of multi-agents have self-loops.

Furthermore, delays occur inevitably due to limited information transmission speed. The consensus algorithm with transmission delays can be described as

$$x_i^{t+1} = \sum_{j=1}^m G_{ij}(t) x_j^{t-\tau_{ij}^t}, \tag{4}$$

where $\tau_{ij}^t \in \mathbb{N}$, $i,j = 1,\dots,m$, denotes the time-dependent delay from vertex j to i. A link from j to i is called *instantaneous* if $\tau_{ij}^t = 0\ \forall t$, and *delayed* otherwise.

In this paper, we study a general consensus problem in networks with time-varying topologies and time delays described by

$$x_i^{t+1} = \sum_{j=1}^m G_{ij}(\sigma^t) x_j^{t-\tau_{ij}(\sigma^t)}, \quad i = 1,\dots,m, \tag{5}$$

as well as the more general form

$$x_i^{t+1} = \sum_{\tau=0}^{\tau_M} \sum_{j=1}^{m} G_{ij}^{\tau}(\sigma^t) x_j^{t-\tau}, \quad i = 1, \dots, m. \tag{6}$$

Note that (5) can be put into the form (6) by partitioning the inter-links according to delays, where τ_M is the maximum delay. However, (6) is more general, as it in principle allows for multiple links with different delays between the same pair of vertices. In particular, there may exist both instantaneous and delayed self-links, which may naturally arise in a model like (1) where the term x_i appears both by itself as well as under the summation sign. In reference to (6), we talk about *self-link(s)* when $G_{ii}^{\tau} \neq 0$, which may be *instantaneous* or *delayed* depending on whether $\tau = 0$ or $\tau > 0$, respectively. In equations (5)–(6), σ^t denotes a stochastic process, $G(\sigma^t) = [G_{ij}(\sigma^t)]_{i,j=1}^n = [\sum_{\tau=0}^{\tau_M} G_{ij}^{\tau}(\sigma^t)]_{i,j=1}^n$ is a stochastic matrix, $\tau_{ij}(\sigma^t) \in \mathbb{N}$ is the stochastically-varying transmission delay from agent j to agent i. This model can describe, for instance, communications between randomly moving agents, where the current locations of the agents, and hence the links between them, are regarded as stochastic. Furthermore, the delays are also stochastic since they arise due to the distances between agents. In this paper, $\{\sigma^t\}$ is assumed to be an adapted stochastic process.

Definition 1.1. (*Adapted process*) Let $\{A_k\}$ be a stochastic process defined on the basic probability space $\{\Omega, \mathcal{F}, P\}$, with the state space Ω, the σ-algebra $\mathcal{F}$, and the probability $\mathbb{P}$. Let $\{\mathcal{F}^k\}$ be a filtration, i.e., a sequence of nondecreasing sub-σ-algebras of $\mathcal{F}$. If A_k is measurable with respect to (w.r.t.) $\mathcal{F}^k$, then the sequence $\{A_k, \mathcal{F}^k\}$ is called an adapted process.

Via a standard transformation, any stochastic process can be regarded as an adapted process. Let $\{\xi_t\}$ be a stochastic process in probability spaces $\{\Omega^t, \mathcal{H}^t, \mathbb{P}^t\}$. Define $\Omega = \prod_t \Omega^t$, $\mathcal{F}$ and $\mathbb{P}$ are both induced by $\prod_t \mathcal{H}^t$ and $\prod_t \mathbb{P}^t$, where $\prod$ stands for the Cartesian product. Let $\sigma^t = [\xi^k]_{k=1}^t$ and $\mathcal{F}^t$ be the minimal σ-algebra induced by $\prod_{k=1}^t \mathcal{H}^t$. Then $\mathcal{F}^t$ is a filtration. Thus, it is clear that the notion of an adapted process is rather general, and it contains i.i.d. processes, Markov chains, and so on, as special cases.

Related work. Many recent papers address the stability analysis of consensus in networks of multi-agents. However, the model (5) with delays we have proposed above is more general than the existing models in the literature. We first mention some papers where models of the form (3) are treated. A result from [25] shows that (3) can reach consensus uniformly if and only if there exists $T > 0$ such that the union graph across any T-length time period has a spanning tree. Ref. [2] derived a similar condition for reaching a consensus via an equivalent concept: *strongly rooted* graph. Our previous papers [19, 20] studied synchronization of nonlinear dynamical systems of networks with time-varying topologies by a similar method. Ref. [36] has pointed out that under the assumption that self-links always exist and are instantaneous (i.e. without delays), the condition presented in Ref. [25] also guarantees consensus with arbitrary bounded multiple delays. However, this criterion may not work when the time-varying topology involves randomness, because for any $T > 0$, it might occur with positive probability that the union graph across some T-length time period does not have a spanning tree for any T. Refs. [14, 35, 31] studied the consensus in networks under the circumstance that the processes $\{G(t)\}_{t \geq 0}$ are independently and identically distributed (i.i.d.) and [38] also investigated the stability of consensus of multi-agent systems with Markovian switching topology with finite states. In these papers, consensus is considered in the almost sure sense. Ref. [8] studied a particular situation with packet drop communication. The most related literature to the current paper is [18], where a general stochastic process, an adapted process, was introduced to model the switching topology, which generalized the existing works including i.i.d. and Markovian jumping topologies as special cases. The authors proved that, if the δ-graph (see its definition in Sec. 2.2) corresponding to the conditional expectation of the coupling matrix sum across a finite time interval has a spanning tree almost surely, then the system reaches consensus. However, none of those works considered the stochastic delays but rather assumed that self-links always exist. There are also many papers concerned with the continuous-time consensus algorithm on networks of agents with time-varying topologies or delays. See Ref. [28] for a framework and Ref. [27] for a survey, as well as Refs. [24, 1, 37, 23], among others. Also, there are papers concerned with nonlinear coupling functions [6] and general coordination [17].

Statement of contributions. In the following sections, we study the stability of the consensus of the delayed system (5), where σ^t is an adapted process. First, we consider the case that each agent contains an instantaneous self-link. In this case, we show that the same conditions enabling the consensus of algorithms without transmission delays, as mentioned in Ref. [18], can also guarantee consensus for the case of arbitrary bounded delays. Second, in case that delays also occur at the self-links (for example, when it costs time for each agent to process its own information), and only certain delay patterns can occur, we show that the algorithm does not necessarily reach consensus but may synchronize to a periodic trajectory instead. As we show, the period of the synchronized state depends on the possible delay patterns. Finally, we briefly study the situation without self-links, and present consensus conditions based on the graph topology and the product of coupling matrices.

The basic tools we use are theorems about product of stochastic matrices and the results from probability theory. Ref. [3] has proved a necessary and sufficient condition for the convergence of infinite stochastic matrix products, which involves the concept of scramblingness. Ref. [34] provided a means to get scrambling matrices (defined in Sec. 2.2) from products of finite stochastic indecomposable aperiodic (SIA) matrices and Ref. [36] showed that an SIA matrix can be guaranteed if the corresponding graph has a spanning tree and one of the roots has a self-link. The Borel-Cantelli lemma [7] indicates that if the conditional probability of the occurrence of SIA matrices in a product of stochastic matrices is always positive, then it occurs infinitely often. These previous results give a bridge connecting the properties of stochastic matrices, graph topologies, and probability theory which we will call upon in the present paper.

The paper is organized as follows. Introductory notations, definitions, and lemmas are given in Sec. 2. The dynamics of the consensus algorithms in networks of multi-agents with switching topologies and delays, which are modeled as adapted processes, are studied in Sec. 3. Applications of the results are provided in Sec. 4 to i.i.d. and Markovian jumping switching. Proofs of theorem are presented in Sec. 5. Conclusions are drawn in Sec. 6.

2. **Preliminaries.**

This paper is written in terms of stochastic process and algebraic graph theory. For the reader's convenience, we present some necessary notations, definitions and lemmas in this section. In what follows, $\underline{N}$ denotes the integers from 1 to N, i.e., $\underline{N} = \{1, \ldots, N\}$. For a vector $v = [v_1, \ldots, v_n]^\top \in \mathbb{R}^n$, $\|v\|$ denotes some norm to be specified, for instance, the L^1 norm $\|v\|_1 = \sum_{i=1}^n |v_i|$. $\mathbb{N}$ denotes the set of positive integers and $\mathbb{Z}$ denotes the integers. For two integers i and j, we denote by $\langle i \rangle_j$ the quotient integer set $\{kj + i : \ k \in \mathbb{Z}\}$. The greatest common divisor of the integers $i_1, \ldots, i_K$ is denoted $\gcd(i_1, \ldots, i_K)$. The product $\prod_{k=1}^n B_k$ of matrices denotes the left matrix product $B_n \times \cdots \times B_1$. For a matrix A, A_{ij} or $[A]_{ij}$ denotes the entry of A on the ith row and jth column. In a block matrix B, B_{ij} or $[B]_{ij}$ can also stand for its i,j-th block. For two matrices A, B of the same dimension, $A \geq B$ means $A_{ij} \geq B_{ij}$ for all i, j, and the relations $A > B$, $A < B$, and $A \leq B$ are defined similarly. I_m denotes the identity matrix of dimension m.

2.1. **Probability theory.** $\{\Omega, \mathcal{F}, \mathbb{P}\}$ is our general notation for a probability space, which may be different in different contexts. In this notation, Ω stands for the state space, $\mathcal{F}$ the Borel σ-algebra, and $\mathbb{P}\{\cdot\}$ the probability on Ω. $\mathbb{E}_{\mathbb{P}}\{\cdot\}$ is the expectation with respect to P (sometimes $\mathbb{E}$ for simplicity, if no ambiguity arises). For any σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, $\mathbb{E}\{\cdot|\mathcal{G}\}$ ($\mathbb{P}\{\cdot|\mathcal{G}\}$) is the conditional expectation (probability, respectively) with respect to $\mathcal{G}$. It should be noted that both $\mathbb{E}\{\cdot|\mathcal{G}\}$ and $\mathbb{P}\{\cdot|\mathcal{G}\}$ are actually random variables measurable w.r.t. $\mathcal{G}$. The following lemma provides the general statement of the principle of large numbers.

Lemma 2.1. [7] *(The Second Borel-Cantelli Lemma) Let $\mathcal{F}_n$, $n \geq 0$ be a filtration with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and C_n, $n \geq 1$ a sequence of events with $C_n \in \mathcal{F}_n$. Then*

$$\{C_n \text{ infinitely often }\} = \Big\{\sum_{n=1}^{+\infty} P\{C_n|\mathcal{F}_{n-1}\} = +\infty\Big\}$$

with a probability 1, where "infinitely often" means that an infinite number of $\{C_n\}_{n=1}^{\infty}$ occur.

2.2. **Stochastic matrices and graphs.** An $m \times m$ matrix $A = [a_{ij}]_{i,j=1}^m$ is said to be a *stochastic matrix* if $a_{ij} \geq 0$ for all $i, j = 1, \ldots, m$ and $\sum_{j=1}^m a_{ij} = 1$ for all $i = 1, \ldots, m$. A matrix $A \in \mathbb{R}^{m,m}$ is said to be *SIA* if A is stochastic, indecomposable, and aperiodic, i.e., $\lim_{n\to\infty} A^n$ converges to a matrix with identical rows. The *Hajnal diameter* is introduced in Ref. [12, 13] to describe the compression rate of a stochastic matrix. For a matrix A with row vectors $a_1, \ldots, a_m$ and a vector norm $\|\cdot\|$ in $\mathbb{R}^m$, the Hajnal diameter of A is defined by $\mathrm{diam}(A) = \max\limits_{i,j} \|a_i - a_j\|$. The *scramblingness* η of a stochastic matrix A is defined as

$$\eta(A) = \min_{i,j} \|a_i \wedge a_j\|_1, \tag{7}$$

where $a_i \wedge a_j = [\min(a_{i1}, a_{j1}), \ldots, \min(a_{im}, a_{jm})]$. The stochastic matrix A is said to be *scrambling* if $\eta(A) > 0$. The Hajnal inequality estimates the Hajnal diameter of the product of stochastic matrices. For two stochastic matrices A and B of the same order, the inequality

$$\mathrm{diam}(AB) \leq (1 - \eta(A))\mathrm{diam}(B) \tag{8}$$

holds for any matrix norm [30]. It can be seen from (8) that the diameter of the product AB is strictly less than that of B if A is scrambling.

The link between stochastic matrices and graphs is an essential feature of this paper. A stochastic (or simply nonnegative) matrix $A = [a_{ij}]_{i,j=1}^m \in \mathbb{R}^{m,m}$ defines a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1, \ldots, m\}$ denotes the *vertex set* with m vertices and $\mathcal{E}$ denotes the *link set* where there exists a directed link from vertex j to i, i.e., $e(i,j)$ exists, if and only if $a_{ij} > 0$. We denote this graph corresponding to the stochastic matrix A by $\mathcal{G}(A)$. For a directed link $e(i,j)$, we say that j is the *start* of the link and i is the *end* of the link. The vertex i is said to be self-linked if $e(i,i)$ exists, i.e., $a_{ii} > 0$. $\mathcal{G}$ is said to be a *bigraph* if the existences of $e(i,j)$ and $e(j,i)$ are equivalent. Otherwise, $\mathcal{G}$ is said to a *digraph*. An L-length *path* in the graph denotes a vertex sequence $(v_i)_{i=1}^L$ satisfying that the link $e(v_{i+1}, v_i)$ exists for all $i = 1, \ldots, L-1$. The vertex i can *access* the vertex j, or equivalently, the vertex j is *accessible* from the vertex i, if there exists a path from the vertex i to j. The graph $\mathcal{G}$ has a *spanning tree* if there exists a vertex i which can access all other vertices, and the set of vertices that can access all other vertices is named the *root set*. The graph $\mathcal{G}$ is said to be *strongly connected* if each vertex is a root. We refer interested readers to the book [11] for more details. Due to the relationship between nonnegative matrices and graphs, we can call on the properties of nonnegative matrices, or equivalently, those of their corresponding graphs. For example, the indecomposability of a nonnegative matrix A is equivalent to that $\mathcal{G}(A)$ has a spanning tree, and the aperiodicity of a graph is associated with the aperiodicity of its corresponding matrix [15]. We say that $\mathcal{G}$ is *scrambling* if for each pair of vertices $i \neq j$, there exists a vertex k such that both $e(i,k)$ and $e(j,k)$ exist, which can be seen to be equivalent to the definition of scramblingness for stochastic matrices. For two matrices $A = [a_{ij}]_{i,j=1}^n, B = [b_{ij}]_{i,j=1}^n \in \mathbb{R}^{n,n}$, we say A is an *analog* of B and write $A \approx B$, in case that $a_{ij} \neq 0$ if and only if $b_{ij} \neq 0$, $\forall i, j = 1, \dots, n$, that is, when their corresponding graphs are identical.

Furthermore, for a nonnegative matrix A and a given $\delta > 0$, the δ-*matrix* of A, denoted by A^δ, is defined as

$$[A^\delta]_{ij} = \begin{cases} \delta, & \text{if } A_{ij} \geq \delta; \\ 0, & \text{if } A_{ij} < \delta. \end{cases}$$

The δ-*graph* of A is the directed graph corresponding to the δ-matrix of A. We denote by $\mathcal{N}_i^\delta$ the neighborhood set of the vertex v_i in the δ-graph: $\mathcal{N}_i^\delta = \{v_j : A_{ij} \geq \delta\}$.

2.3. **Convergence of products of stochastic matrices.** Here, we provide the definition of consensus and synchronization of the system (5). Suppose the delays are bounded, namely, $\tau_{ij}(\sigma^k) \leq \tau_M$ for all $i, j = 1, \dots, m$ and $\sigma^k \in \Omega$.

Definition 2.2. The multi-agent system is said to *reach consensus* via the algorithm (5) if for any essentially bounded random initial data $x_\tau^0 \in \mathbb{R}^m$, $\tau = 0, 1, \dots, \tau_M$, (that is, x_τ^0 is bounded with probability one), and almost every sequence $\{\sigma^t\}$, there exists a number $\alpha \in \mathbb{R}$ such that $\lim_{t\to\infty} x^t = \alpha \mathbf{1}$ with $\mathbf{1} = [1, 1, \dots, 1]^\top$. The multi-agent system is said to *synchronize* via the algorithm (5) if for any initial essentially bounded random $x^0 \in \mathbb{R}^m$ and almost every sequence $\{\sigma^t\}$, $\lim_{t\to\infty} |x_i(t) - x_j(t)| = 0,\ i, j = 1, \dots, m$. In particular, if for any initial essentially bounded random $x_\tau^0 \in \mathbb{R}^m$, $\tau = 0, 1, \dots, \tau_M$, and almost every sequence, there exists a P-periodic trajectory $s(t)$ (P independent of the initial values and the sequence) such that $\lim_{t\to\infty} |x_i(t) - s(t)| = 0$ holds for all $i = 1, \dots, m$, then the multi-agent system is said to *synchronize to a P-periodic trajectory* via the algorithm (5).

In general, consensus can be regarded as a special case of synchronization, where the multi-agent system synchronizes at an equilibrium. As shown in Ref. [3], in the absence of delays, consensus and synchronization are equivalent w.r.t. the product of infinite stochastic matrices; that is, whenever a system synchronizes, it also reaches consensus. However, we will show in the following sections that, under transmission delays, consensus and synchronization of the algorithm (5) are not equivalent. Thus, a system can synchronize without necessarily reaching consensus.

Consider the model where the topologies are induced by a stochastic process:

$$x_i^{t+1} = \sum_{j=1}^m G_{ij}(\xi^t) x_j^t,\ i = 1, \dots, m, \tag{9}$$

where $\{\xi^t\}_{t\in\mathbb{N}}$ is a stochastic process with a probability distribution of the sequence $\mathbb{P}$. The results of this paper are based on the following lemma, which is a consequence of Theorem 2 in Ref. [3].

Lemma 2.3. *Let $\eta(\cdot)$ denote the scramblingness, as defined in (7). The multi-agent system via the algorithm (9) reaches consensus if and only if for $\mathbb{P}$-almost every sequence there exist infinitely many disjoint integer intervals $I_i = [a_i, b_i]$ such that*

$$\sum_{i=1}^{\infty} \eta\bigg(\prod_{k=a_i}^{b_i} G(\xi^k)\bigg) = \infty.$$

As a trivial extension to a set of SIA matrices, we have the next lemma on how to obtain scramblingness.

Lemma 2.4. [34] *Let $\Theta \subset \mathbb{R}^{m,m}$ be a set of SIA matrices. There exists an integer N such that any n-length matrix sequence with $n > N$ picked from Θ: $G^1, G^2, \ldots, G^n$ satisfies*

$$\eta\bigg(\prod_{k=1}^{n} G^k\bigg) > 0.$$

The following result provides a relation between SIA matrices and spanning trees.

Lemma 2.5. *(Lemma 1 in Ref.* [36]*) If the graph corresponding to a stochastic matrix A has a spanning tree and a self-link at one of its root vertices, then A is SIA.*

3. Main results.

We first consider the multi-agent network without transmission delays:

$$x_i^{t+1} = \sum_{j=1}^{n} G_{ij}(\sigma^t) x_j^t, \; i = 1, \ldots, m. \tag{10}$$

The following theorem is the main tool for the proofs of the main results and it can be regarded as a realization of Lemma 2.3 and an extension from Ref. [18] without assuming self-links.

Theorem 3.1. *For the system (10), if there exist $L \in \mathbb{N}$ and $\delta > 0$ such that the δ-graph of the matrix product*

$$\mathbb{E}\bigg\{\prod_{k=n+1}^{n+L} G(\sigma^k) | \mathcal{F}^n\bigg\} \tag{11}$$

has a spanning tree and is aperiodic for all $n \in \mathbb{N}$ almost surely, then the multi-agent system reaches a consensus.

The proof is given in Sec. 5.1. The main result of [18] can be regarded as a consequence of Theorem 3.1, where each node in the graph was assumed to have a self-link. In the following, we first study the multi-agent systems with transmission delays such that each agent is linked to itself without delay and then investigate the general situation where delays may occur also on the self-links. Finally, we give a brief discussion on the consensus algorithms without self-links. All proofs in this section are placed in Sec. 5.

3.1. Consensus and synchronization with transmission delays.

Consider the consensus algorithm (6), which we rewrite in matrix form as

$$x^{t+1} = \sum_{\tau=0}^{\tau_M} G^{\tau}(\sigma^t) x^{t-\tau}, \tag{12}$$

where $G(\sigma^t) = [G_{ij}^{\tau}(\sigma^t)]_{i,j=1}^n$. We assume the following for the matrices $G^{\tau}(\cdot)$.

A: Each $G^{\tau}(\sigma^t)$, $\tau \in \underline{\tau_M}$, is a measurable map from Ω to the set of nonnegative matrices with respect to $\mathcal{F}^t$.

Letting $y^t = [{x^t}^\top, {x^{t-1}}^\top, \dots, {x^{t-\tau_M}}^\top]^\top \in \mathbb{R}^{m\times(\tau_M+1)}$, we can write (12) as

$$y^{t+1} = B(\sigma^t)y^t, \tag{13}$$

where $B(\sigma^t) \in \mathbb{R}^{(\tau_M+1)\times m,(\tau_M+1)\times m}$ has the form

$$B(\sigma^t) = \left[\begin{array}{lllll} G^0(\sigma^t) & G^1(\sigma^t) & \cdots & G^{\tau_M-1}(\sigma^t) & G^{\tau_M} \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \end{array}\right].$$

Thus, the consensus of (6) is equivalent to that of (13). As a default labeling, let us consider the corresponding graph $\mathcal{G}(B(\sigma^t))$, which has $(\tau_M+1)m$ vertices, which we denote by $\{v_{i,j},\ i \in \underline{\tau_M+1},\ j \in \underline{m}\}$, where $v_{i,j}$ corresponds to the $((i-1)m+j)$th row (or column) of the matrix $B(\sigma^t)$.

Theorem 3.2. *Assume the conditions* $\mathbf{A}$, *and suppose there exist* $\mu > 0$, $L \in \mathbb{N}$, *and* $\delta > 0$ *such that* $G^0(\sigma) > \mu I_m$ *for all* $\sigma \in \Omega$ *and the* δ*-graph of* $\mathbb{E}\{\sum_{k=n+1}^{n+L} G(\sigma^k)|\mathcal{F}^n\}$ *has a spanning tree for all* $n \in \mathbb{N}$ *almost surely. Then the delayed multi-agent system (6) reaches consensus.*

The proof is given in Sec 5.2. In the case that the topological switching is deterministic, a similar result is obtained in the literature [24, 36].

Example 3.3. We give a simple example to illustrate Theorem 3.2. Consider a delayed multi-agent system on a network with 2 vertices and the maximum delay is 1. The system can be written as

$$x^{t+1} = G^0(\sigma^t)x^t + G^1(\sigma^t)x^{t-1},$$

which can further be put into a form without delays $y^{t+1} = B(\sigma^t)y^t$ with

$$B(\sigma^t) = \left(\begin{array}{ll} G^0(\sigma^t) & G^1(\sigma^t) \\ I_m & 0 \end{array}\right).$$

Let us consider the product of two matrices B^1 and B^2:

$$B^1 = \left(\begin{array}{llll} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right), \quad B^2 = \left(\begin{array}{llll} 1/2 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right).$$

In the absence of delays, they correspond to $G_1 = \left(\begin{array}{ll} 1 & 0 \\ 0 & 1 \end{array}\right)$ and $G_2 = \left(\begin{array}{ll} 1/2 & 1/2 \\ 0 & 1 \end{array}\right)$. One can see that the union of the graphs $\mathcal{G}(G_1)$ and $\mathcal{G}(G_2)$ has spanning trees and self-connections. Then the proof of Theorem 3.2 says that for some integer L, the product of L successive matrices corresponds to a graph which has a spanning tree and a self-link on the root node. For example, we consider the following matrix product:

$$B^1B^2 = \left(\begin{array}{llll} 1/2 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 \end{array}\right).$$

The corresponding graph has four vertices, which we label as $v_{1,1}$, $v_{1,2}$, $v_{2,1}$, and $v_{2,2}$ following the scheme defined below Eq. (13). From Figure 1, it can be seen that the graph corresponding to B^1B^2 has spanning trees with $v_{1,2}$ being the root vertex which has a self-link. So, by Theorem 3.2, the system reaches consensus.

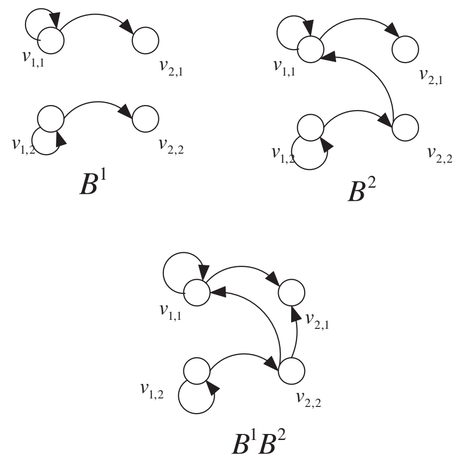

FIGURE 1. The graphs corresponding to the matricesB^1, B^2, and the matrix product B^1B^2, respectively.

In some cases delays occur at self-links, for example, when it takes time for each agent to process its own information. Suppose that the self-linking delay for each vertex is identical, that is, $\tau_{ii} = \tau_0 > 0$. We classify each integer t in the discrete-time set $\mathbb{N}$ (or the integer set $\mathbb{Z}$) via $\mod (t+1, \tau_0+1)$ as the quotient group of $(\mathbb{Z}+1)/(\tau_0+1)$. As a default set-up, we denote $\langle i\rangle_{\tau_0+1}$ by $\langle i\rangle$. Let $\hat{G}^i(\cdot) = \sum_{j\in\langle i\rangle} G^j(\cdot)$. For a simplified statement of the result, we provide the following condition **B**:

***B**.1 There exist an integer $\tau_0 > 0$ and a number $\mu > 0$ such that $G^{\tau_0}(\sigma_1) > \mu I_m$ for all $\sigma_1 \in \Omega$;*

***B**.2 There exist $\tau_1, \dots, \tau_K$ excluding the integers in $\langle 0\rangle$ with gcd$(\tau_0+1, \tau_1+1, \dots, \tau_K+1) = P > 1$ such that $\hat{G}^j(\sigma_1) = 0$ for all $j \notin \{\tau_1, \dots, \tau_K\}$ and all $\sigma_1 \in \Omega$ and the δ-matrix of $\mathbb{E}\{\hat{G}^{\tau_k}(\sigma^{n+1})|\mathcal{F}^n\}$ is nonzero for all $n \in \mathbb{N}$ and $k = 1, \dots, K$ almost surely.*

Theorem 3.4. *Assume that the conditions* **A** *and* **B** *hold, and suppose there exist $L \in \mathbb{N}$ and $\delta > 0$ such that the δ-graph of $\mathbb{E}\{\sum_{k=n+1}^{n+L} \hat{G}^0(\sigma^k)|\mathcal{F}^n\}$ is strongly connected for all $n \in \mathbb{N}$ almost surely. Then the system (6) synchronizes to a P-periodic trajectory. In particular, if $P = 1$, then (6) reaches consensus.*

The proof is given in Sec. 5.3. From this theorem, one can see that under self-linking delays, consensus is not equivalent to synchronization. In fact, the delays that occur on self-links are essential for the failure to reach consensus.

Example 3.5. Theorem 3.4 demands that the δ-graph corresponding to the matrix $\mathbb{E}\{\sum_{t=n+1}^{n+L} \hat{G}^0(\sigma^t)|\mathcal{F}^n\}$ is strongly connected. This is stronger than the condition in Theorem 3.2, which demands that the corresponding graph has a spanning tree. We give an example to show that the strong connectivity is necessary for the reasoning in the proof. Consider a delayed multi-agent system on a network with two vertices and a maximum

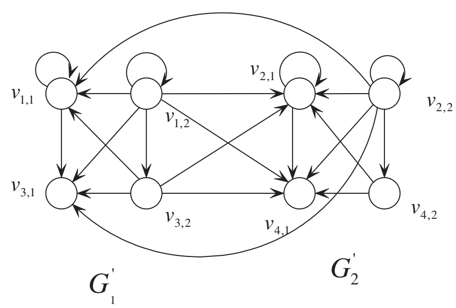

FIGURE 2. The graph corresponding to the matrix product (14).

delay of 3. Consider the form (13) and the matrix $B(\cdot)$. Suppose that the state space only contains one state σ_1 as follows:

$$B(\sigma_1)=\begin{bmatrix} 0 & 0 & 1/3 & 0 & 0 & 1/3 & 0 & 1/3\\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Here, $\tau_0 = 1$. It is clear that the subgraph corresponding to each $\hat{G}^0_{1,2}$ has spanning trees but is not strongly connected, and that there is a link between the subgraphs corresponding to $\langle 1\rangle$ and $\langle 0\rangle$. For the word $\sigma_1\sigma_1\cdots\sigma_1\sigma_1$, direct calculations show that the corresponding matrix product is an analog of the following matrix if the length of the word is sufficiently long:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1\\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1\\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{14}$$

The corresponding graph is shown in Figure 2, using the labeling scheme for the vertices as defined below Eq. (13). One can see that it does not have a spanning tree since the vertices $v_{1,2}$ and $v_{2,2}$ do not have incoming links other than self-links. In fact, the set of eigenvalues of the matrix $B(\sigma_1)$ contains 1 and -1, which implies that (12) with $B(\sigma^t)$ can not reach consensus even though the condition in Theorem 3.2 is satisfied.

3.2. **Consensus and synchronization without self-links.** So far the stability result is based on the assumption that each agent takes its own state into considerations when updating. In other words, the coupling matrix has positive diagonals (possibly with delays). There also exist consensus algorithms that are realized by updating each agent's state via averaging its neighbor's states and possibly *excluding* its own [9]. In [5], it is shown that consensus can be reached in a static network if each agent can communicate with others by a directed graph and the coupling graph is aperiodic, which can be proved by nonnegative matrix theory [15]. In the following, we briefly discuss the general consensus algorithms in networks of stochastically switching topologies that do not necessarily have self-links for all vertices.

When transmission delays occur, the general algorithm (6) can be regarded as increasing dimensions as in (13). Thus, one can similarly associate (13) with a new graph on $m \times (\tau_M + 1)$ vertices $\{v_{ij} : \ i \in \underline{\tau_M + 1}, j \in \underline{m}\}$, denoted by $\mathcal{G}^{'}(\cdot)$, where $B(\cdot)$ denotes the link set of $\mathcal{G}^{'}(\cdot)$, by which v_{ij} corresponds to the $(i-1) \times (\tau_M + 1) + j$ column and row of B. $\hat{B}_p(\sigma_1)$ as the matrix corresponding the vertices $\{v_{ij} : \ i \in \langle p \rangle, \ j \in \underline{m}\}$. Based on theorem 3.1, we have the following results, which can be proved similarly to Theorems 3.2 and 3.4.

Proposition 3.6. *Assume* $\mathbf{A}$ *holds, and suppose there exist* $L \in \mathbb{N}$ *and* $\delta > 0$ *such that the* δ*-graph of* $\mathbb{E}\{\prod_{k=u+1}^{u+L} B(\sigma^k)|\mathcal{F}^u\}$ *has a spanning tree and self-link at one root vertex for all* $n \in \mathbb{N}$ *almost surely. Then the algorithm (10) reaches consensus.*

In fact, under the stated conditions, each product $\mathbb{E}\{\prod_{k=u+1}^{u+L} B(\sigma^k)|\mathcal{F}^u\}$ is SIA almost surely; so, this proposition is a direct consequence of Theorem 3.1.

In the possible absence of self-links, the following is a consequence of Proposition 3.6.

Proposition 3.7. *Assume* $\mathbf{A}$ *and* $\mathbf{B}.2$ *hold (*$\mathbf{B}.1$ *need not hold). Suppose there exist* $L \in \mathbb{N}$ *and* $\delta > 0$ *such that the* δ*-graph of* $\mathbb{E}\{\prod_{k=n+1}^{n+L} \hat{B}_p(\sigma^k)|\mathcal{F}^n\}$ *is strongly connected and has at least one self-link for all* $n \in \mathbb{N}$ *and* $p \in \underline{P}$ *almost surely, where* $\hat{B}_p$ *is defined in the proof of Theorem 3.4, for example, (15) in Sec. 5.3. Then the algorithm (6) synchronizes to a* P*-periodic trajectory. In particular if* $P = 1$*, then the algorithm (6) reaches consensus.*

4. **Applications.** Adapted processes are rather general and include i.i.d processes and Markov chains as two special cases. Therefore, the results obtained above can be directly utilized to derive sufficient conditions for the cases where the topology switching and delays are i.i.d. or Markovian.

First, by a standard construction as mentioned in Sec. 1, from the property of i.i.d. it follows that $\mathbb{E}\{G(\sigma^{k+1})|\mathcal{F}^k\} = \mathbb{E}\{G(\sigma^{k+1})\}$ is a constant stochastic matrix. Then, we have the following results.

Corollary 4.1. *Assume that* $\mathbf{A}$ *holds and* $\{\sigma^t\}$ *is an i.i.d. process. Suppose there exist* $\mu > 0$*,* $L \in \mathbb{N}$*, and* $\delta > 0$ *such that* $G^0(\sigma) > \mu I_m$ *for all* $\sigma \in \Omega$ *and the* δ*-graph of* $\mathbb{E}\{G(\sigma^1)\}$ *has a spanning tree. Then the delayed multi-agent system via algorithm (6) reaches consensus.*

Corollary 4.2. *Assume that* $\mathbf{A}$ *and* $\mathbf{B}$ *hold and* $\{\sigma^t\}$ *is an i.i.d. process. Suppose there exist* $L \in \mathbb{N}$ *and* $\delta > 0$ *such that the* δ*-graph of* $\mathbb{E}\{\hat{G}^0(\sigma^1)\}$ *is strongly connected for all* $n \in \mathbb{N}$ *almost surely. Then the system (6) synchronizes to a* P*-periodic trajectory. In particular, if* $P = 1$*, then (6) reaches consensus.*

Second, we consider the Markovian switching topologies, namely, the graph sequence is induced by a homogeneous Markov chain with a stationary distribution and the property of uniform ergodicity, which is defined as follows.

Definition 4.3. [4] A Markov chain $\{\sigma^t\}$, defined on $\{\Omega, \mathcal{F}\}$, with a stationary distribution π and a transition probability $\mathbb{T}(x, A)$ is called uniformly ergodic if

$$\sum_{x \in \Omega} \|\mathbb{T}^k(x, \cdot) - \pi(\cdot)\| \to 0 \text{ as } k \to +\infty,$$

where $\mathbb{T}^k(\cdot,\cdot)$ denotes the k-th iteration of the transition probability $\mathbb{T}(\cdot,\cdot)$, for two probability measures μ and ν on $\{\Omega,\mathcal{F})\}$, and $\|\mu-\nu\|=\sup_{\mathcal{A}\in\mathcal{F}}|\mu(\mathcal{A})-\nu(\mathcal{A})|$.

From the Markovian property, we have the following results.

Corollary 4.4. *Assume that* $\mathbf{A}$ *holds. Let* $\{\sigma^t\}$ *be an irreducible and aperiodic Markov chain with a unique invariant measure* π*. Suppose* $\{\sigma^t\}$ *is uniformly ergodic and there exist* $\mu>0$ *and* $\delta>0$ *such that* $G^0(\sigma)>\mu I_m$ *for all* $\sigma\in\Omega$ *and the* δ*-graph of* $\mathbb{E}_\pi\{G(\sigma^1)\}$ *has a spanning tree. Then the delayed multi-agent system(6) reaches consensus.*

Proof. From the Markovian property, we have

$$E\{\frac{1}{L}\sum_{t=n+1}^{n+L}G(\sigma^t)|\mathcal{F}^n\}=E\{\frac{1}{L}\sum_{t=n+1}^{n+L}G(\sigma^t)|\sigma^n\}.$$

If $\{\sigma^t\}$ is uniformly ergodic, then

$$\lim_{L\to+\infty}\mathbb{E}\{\frac{1}{L}\sum_{1=n+1}^{n+L}G(\sigma^t)|\sigma^n\}=\lim_{L\to+\infty}\frac{1}{L}\sum_{i=1}^{L}\int_\Omega G(y)\mathbb{T}^i(\sigma^n,dy)=\int_\Omega G(y)\pi(dy)=\mathbb{E}_\pi[G(\sigma^1)].$$

Since the convergence is uniform, there exits L such that the $\delta/2$-graph corresponding to $\mathbb{E}\{(1/L)\sum_{t=n+1}^{n+L}G(\sigma^t)|\mathcal{F}^n\}$ has a spanning tree almost surely. From Theorem 3.2, the conclusion can be derived. □

Corollary 4.5. *Assume that* $\mathbf{A}$ *and* $\mathbf{B}$ *hold, and let* $\{\sigma^t\}$ *be an irreducible and aperiodic Markov chain with a unique invariant measure* π*. Suppose that* $\{\sigma^t\}$ *is uniformly ergodic and there exists* $\delta>0$ *such that the* δ*-graph of* $\mathbb{E}_\pi\{\hat{G}^0(\sigma^1)\}$ *is strongly connected. Then the system (6) synchronizes to a* P*-periodic trajectory. In particular, if* $P=1$*, then (6) reaches consensus.*

These corollaries can be proved directly from Theorems 3.4 in the same way as Corollary 4.4. It can be seen that the a homogeneous Markov chain with finite state space and unique invariant distribution is uniformly ergodic. Hence, the results of Corollaries 4.4 and 4.5 hold for this scenario.

5. **Proofs of the main results.** In the following, the coupling matrix $B(\cdot)$ in the delayed system (13) is written in the following block form:

$$B(\sigma^t)=\left[\begin{array}{cccc} B_{1,1}(\sigma^t) & B_{1,2}(\sigma^t) & \cdots & B_{1,\tau_M+1}(\sigma^t)\\ B_{2,1}(\sigma^t) & B_{2,2}(\sigma^t) & \cdots & B_{2,\tau_M+1}(\sigma^t)\\ \vdots & \vdots & \ddots & \vdots\\ B_{\tau_M+1,1}(\sigma^t) & B_{\tau_M+1,2}(\sigma^t) & \cdots & B_{\tau_M+1,\tau_M+1}(\sigma^t)\end{array}\right]\in\mathbb{R}^{(\tau_M+1)m,(\tau_M+1)m}$$

with $B_{ij}(\sigma^t)\in\mathbb{R}^{m,m}$, $i,j\in\underline{\tau_M+1}$. For two index sets I and J, we denote by $[B(\sigma^t)]_{I,J}$ the sub-matrix of $B(\sigma^t)$ with row index set I and column index set J. For an n-length word $\sigma=(\sigma^k)_{k=1}^n$ in the stochastic process, we use $B(\sigma)$ to represent the matrix product $\prod_{i=1}^n B(\sigma^i)$. One can see that the structure of the matrix $B(\sigma^t)$ has the following properties: (1). Each $B_{i,i-1}=I_m$ for all $i\geq 2$; (2). $B_{i,j}=0$ for all $i\geq 2$ and $j\neq i-1$. These properties are essential for the following proofs.

As the same way defined below Eq. (13), let us consider the corresponding graph $\mathcal{G}(B(\sigma^t))$, which has $(\tau_M+1)m$ vertices, which we denote by $\{v_{i,j},\ i\in\underline{\tau_M+1},\ j\in\underline{m}\}$, where $v_{i,j}$ corresponds to the $(i-1)m+j$ row of the matrix $B(\sigma)$.

We denote the following finitely generated periodic group:

$$\langle i_1, i_2, \ldots, i_K\rangle_j := \{p:\ p = \sum_{l=k}^{K} i_k p_k \bmod j,\ p_k \in \mathbb{Z}\}.$$

If these numbers are be picked in a finite integer set, for instance, $\{1, \ldots, \tau_M + 1\}$ in the present paper, then $\langle i_1, i_2, \ldots, i_K\rangle_j$ denotes the set $\langle i_1, i_2, \ldots, i_K\rangle_j \bigcap \underline{\tau_M + 1}$ unless specified otherwise. As a default setup, $\langle i\rangle$ denotes $\langle i\rangle_{\tau_0+1}$ where τ_0 is the self-linking delay as in (12). We will sometimes be interested in whether an element in a matrix is zero or not, regardless of its actual value.

5.1. **Proof of Theorem 3.1.** From the condition in this theorem, we can see that the δ-matrix of $\mathbb{E}\{\prod_{k=n+1}^{n+L} G(\sigma^k)|\mathcal{F}^n\}$ is SIA for all $n \in \mathbb{N}$. Lemma 2.4 states that there exists $N \in \mathbb{N}$ such that the product of any N SIA matrices in $\mathbb{R}^{m,m}$ is scrambling. Note that

$$\mathbb{E}\Big\{\prod_{t=n+1}^{n+NL} G(\sigma^t)|\mathcal{F}^n\Big\} = \mathbb{E}\Big\{\cdots\mathbb{E}\Big\{\mathbb{E}\Big\{\prod_{t_L=n+(N-1)L+1}^{n+NL} G(\sigma^{t_L})|\mathcal{F}^{n+(N-1)L}\Big\}$$
$$\prod_{t_{L-1}=n+(N-2)L+1}^{n+(N-1)L} G(\sigma^{t_{L-1}})|\mathcal{F}^{n+(N-2)L}\Big\}\cdots\prod_{t_1=n+1}^{n+L} G(\sigma^{t_1})|\mathcal{F}^n\Big\},$$

since $\{\mathcal{F}^t\}$ is a filtration. This implies that there exists a positive constant $\delta_1 < \delta^N$ such that the δ_1-graph of $\mathbb{E}\{\prod_{t=n+1}^{n+NL} G(\sigma^t)|\mathcal{F}^n\}$ is scrambling. So, from Lemma 3.12 in Ref. [18], there exist $\delta' > 0$ and $M_1 \in \mathbb{N}$ such that

$$\mathbb{P}\Big\{\eta\Big(\prod_{t=n+1}^{n+M_1NL} G(\sigma^t)\Big) > \delta'|\mathcal{F}^n\Big\} > \delta',\ \forall\, n \in \mathbb{N}.$$

Let $C_k = \prod_{t=kM_1NL+1}^{(k+1)M_1NL} G(\sigma^t)$. We can conclude that for almost every sequence of $\{\sigma^t\}$, it holds that

$$\lim_{K\to\infty}\sum_{k=1}^{K}\mathbb{P}\Big\{\eta(C_k) > \delta'|\mathcal{F}^{kNL}\Big\} > \lim_{K\to\infty} K \times \delta' = +\infty.$$

From Lemma 2.1, we can conclude that the events $\{\eta(C_k) > \delta'\}$, $k = 1, 2, \ldots,$ occur infinitely often almost surely. Therefore, we can complete the proof directly from Lemma 2.3.

5.2. **Proof of Theorem 3.2.** The proof of this theorem is based on the structural characteristics of the product of matrices $B(\cdot)$. We denote by $[B(\cdot)]_{ij}$ the $\mathbb{R}^{m,m}$ sub-matrix of $B(\sigma)$ in the position (i, j). We first show by the following lemma that the graph corresponding to the product of more than $\tau_M + 1$ successive matrices $B(\sigma^t)$, as defined by (13), has a spanning tree and self-link at one root vertex. Thus, we can prove Theorem 3.2 by employing Theorem 3.1.

Lemma 5.1. *Under the conditions in Theorem 3.2, for any n-length word $\sigma = (\sigma_i)_{i=1}^n$ with $n \geq \tau_M + 1$, there exists $\mu_1 > 0$ such that*

(i). $[B(\sigma)]_{i,1} \geq \mu_1^n I_m$;

(ii). $\sum_{j=1}^{\tau_M+1}[B(\sigma)]_{1,j} \geq \mu_1^n \sum_{j=1}^{\tau_M+1}\sum_{k=1}^{n} G^j(\sigma^k)$.

Proof. We choose $0 < \mu_1 < \mu$, where μ is defined in Theorem 3.2. (i). For a word $\sigma = (\sigma_i)_{i=1}^n$ with $n \geq \tau_M + 1$,

$$\begin{aligned}[B(\sigma)]_{i,1} &= \sum_{i_1,\dots,i_n} [B(\sigma_n)]_{i,i_1}[B(\sigma_{n-1})]_{i_1,i_2}\cdots[B(\sigma_1)]_{i_n,1} \\ &\geq \left(\prod_{k=n-i+2}^{n} [B(\sigma_k)]_{k+i-n,k+i-n-1}\right)\left(\prod_{k=1}^{n-i+1} [B(\sigma_k)]_{1,1}\right) \\ &= \prod_{k=1}^{n-i+1} [B(\sigma_k)]_{1,1} \geq \mu_1^n I_m\end{aligned}$$

since $[B(\varpi)]_{k+i-n,k+i-n-1} = I_m$ for all $k \geq n-i+2$ and $[B(\varpi)]_{1,1} \geq \mu I_m \geq \mu_1 I_m$ for all $\varpi \in \Omega$.

(ii). Let $j \in \underline{\tau_M + 1}$ and $t_0 \in \underline{n}$. If $t_0 \geq j$, we have

$$\begin{aligned}\sum_l [B(\sigma)]_{1,l} &= \sum_{i_1,\dots,i_n,l} [B(\sigma_n)]_{1,i_1}[B(\sigma_{n-1})]_{i_1,i_2}\cdots[B(\sigma_1)]_{i_n,l} \\ &\geq \left(\prod_{k=t_0+1}^{n} [B(\sigma_k)]_{1,1}\right)[B(\sigma_{t_0})]_{1,j}\left(\prod_{l=t_0-j+2}^{t_0-1} [B(\sigma_l)]_{l-t_0+j,l-t_0+j-1}\right) \\ &\quad \left(\prod_{p=1}^{t_0-j+1} [B(\sigma_p)]_{1,1}\right) \\ &\geq \mu_1^n [B(\sigma_{t_0})]_{1,j},\end{aligned}$$

since $[B(\varpi)]_{1,1} \geq \mu_1 I_m$, $[B(\varpi)]_{l-t_0+j,l-t_0+j-1} = I_m$ for all $l \geq t_0 - j + 2$ for all $\varpi \in \Omega$; whereas if $j > t_0$, we similarly have

$$\begin{aligned}\sum_l [B(\sigma)]_{1,l} &\geq \left(\prod_{k=t_0+1}^{n} [B(\sigma_n)]_{1,1}\right)[B(\sigma_{t_0})]_{1,j}\left(\prod_{l=1}^{t_0-1} [B(\sigma_l)]_{l+j-t_0+1,l+j-t_0}\right) \\ &\geq \mu_1^n [B(\sigma_{t_0})]_{1,j}.\end{aligned}$$

Summing the right-hand side of the above inequality with respect to t_0 and $j$proves (ii). □

Proof of Theorem 3.2. Let us consider the μ_1^n-graph of $B(\sigma)$ for all $\sigma = (\sigma^t)_{t=1}^n$ with $n \geq \tau_M + 1$, as defined in Lemma 5.1. The item (i) in Lemma 5.1 indicates that for each vertex $v_{i,j}$ with $i \geq 2$ and $j \in \underline{m}$, there exist a path from vertex $v_{1,j}$ to $v_{i,j}$: $(v_{1,j}, v_{2,j}, \dots, v_{i,j})$.

From item (ii) in Lemma 5.1 and the conditions in Theorem 3.2, one can see that there exits $\delta > 0$ and $L \in \mathbb{N}$ such that the δ-graph of $\sum_l [\mathbb{E}\{\prod_{t=n+1}^{n+L} B(\sigma^t)|\mathcal{F}^n\}]_{1,l}$ has spanning trees and self-links. Let $\mathcal{G}$ be the random variable corresponding to the δ-graph of $\mathbb{E}\{\prod_{t=n+1}^{n+L} B(\sigma^t)|\mathcal{F}^n\}$ and $\mathcal{G}'$ be the random variable corresponding to the δ-graph of $\sum_l [\mathbb{E}\{\prod_{t=n+1}^{n+L} B(\sigma^t)|\mathcal{F}^n\}]_{1,l}$. Then, for almost every graph $\mathcal{G}'$, there exists an index $j_0 \in \underline{m}$ such that for any j, there exists a path $(j_0, j_1, \dots, j_{K-1}, j)$ to access j. This implies that for almost every graph $\mathcal{G}$, there exists a path from v_{1,j_0} to $v_{1,j}$. Thus, v_{1,j_0} can access all vertices $v_{i,j}$, $i = 1, \dots, \tau_M + 1$, since $v_{1,j}$ can access all $v_{i,j}$ for $\tau_M + 1 \geq i \geq 2$ by a directed link and v_{1,j_0} and has self-link, noting that $G^0(\cdot)$ has positive diagonals. Therefore, for almost every graph $\mathcal{G}$, it has a spanning tree and the vertex v_{1,j_0} is one of the roots. From Lemma 2.5, one can see that $\mathbb{E}\{\prod_{t=n+1}^{n+L} B(\sigma^t)|\mathcal{F}^n\}$ is SIA almost surely. According to Theorem 3.1, the system (10) reaches consensus. This proves the theorem.

5.3. **Proof of Theorem 3.4.** *Outline of the proof:* For a better understanding of the proof, we first give the following sketch. We start the proof by defining a permutation matrix $Q \in \mathbb{R}^{\tau_M+1,\tau_M+1}$ corresponding to the permutation sequence from $(1, 2 \ldots, \tau_M + 1)$ to $(\langle 1\rangle, \langle 2\rangle, \ldots, \langle P\rangle)$. Then we show by the lemma that follows that the matrix $B(\sigma^t)$ can be transformed into the following form:

$$[Q \otimes I_m]B(\sigma^t)[Q \otimes I_m]^\top = \begin{bmatrix} \hat{B}_1(\sigma^t) & 0 & \cdots & 0 \\ 0 & \hat{B}_2(\sigma^t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{B}_P(\sigma^t) \end{bmatrix}, \tag{15}$$

where $\otimes$ stands for the Kronecker product and $\hat{B}_p(\sigma^t) = B_{\langle\langle p\rangle|_P\rangle,\langle\langle p\rangle|_P\rangle}(\sigma^t)$. By the permutation Q, we can rewrite the coupled system (5) as

$$\hat{y}^{t+1} = \hat{B}(\sigma^t)\hat{y}^t, \tag{16}$$

where $\hat{y}^t = [Q \otimes I_m]y^t$ and $\hat{B}(\cdot) = [Q \otimes I_m]B(\cdot)[Q \otimes I_m]^\top$. This system can be divided into P subsystem as

$$\hat{y}_p^{t+1} = \hat{B}_p(\sigma^t)\hat{y}_p^t, \; p \in \underline{P}, \tag{17}$$

where $\hat{y}_p^t$ corresponds to $[y^t]_{\langle\langle p\rangle|P\rangle}$. So, it is sufficient to prove the following claim to complete this proof from Lemma 3.1:

Claim 1: For each $p \in \underline{P}$, there exists $\delta' > 0$ and $L \in \mathbb{N}$ such that the δ'-graph of the matrix

$$\mathbb{E}\left\{ \prod_{t=n+1}^{n+L} \hat{B}_p(\sigma^t)|\mathcal{F}^n \right\} \tag{18}$$

has a spanning tree for all $n \in \mathbb{N}$ almost surely.

The proof of this theorem is also based on the structural characteristics of the product of matrices $B(\cdot)$. By the lemmas below, we are to show the permutation form (15) can be guaranteed.

Lemma 5.2. *Under the conditions of Theorem 3.4, for any $(\tau_0 + 1)$-length word $\sigma = (\sigma_k)_{k=1}^{\tau_0+1}$, there exists some $\mu_1 > 0$ such that the following hold:*

(i). $[B(\sigma)]_{i,i} \geq \mu_1^{\tau_0+1} I_m$ *for all* $i \in \underline{\tau_0 + 1}$*;*

(ii). $[B(\sigma)]_{j,j-(\tau_0+1)} \geq I_m$ *for all* $j \geq \tau_0 + 2$*;*

(iii). $\sum_{l\in\langle 1-j\rangle}[B(\sigma)]_{\tau_0+2-j,l} \geq \mu_1^{\tau_0+1}\hat{G}^0(\sigma_j)$ *for all* $j \in \underline{\tau_0 + 1}$*;*

(iv). $\sum_{l\in\langle i+(\tau+1)\rangle}[B(\sigma)]_{i,l} \geq \mu_1^{\tau_0+1}[B(\sigma_{\tau_0+2-i})]_{1,\tau+1}$ *for all* $i \in \underline{\tau_0 + 1}$ *and* $\tau \in \underline{\tau_M}$*.*

Proof. We choose $0 < \mu_1 < \mu$. (i). For any $i \in \underline{\tau_0 + 1}$, we have

$$\begin{aligned}
[B(\sigma)]_{i,i} &= \sum_{i_1,\ldots,i_{\tau_0}} [B(\sigma_{\tau_0+1})]_{i,i_1}[B(\sigma_{\tau_0})]_{i_1,i_2}\cdots[B(\sigma_1)]_{i_{\tau_0},i} \\
&\geq \left(\prod_{p=\tau_0+3-i}^{\tau_0+1} [B(\sigma_p)]_{p+i-1-\tau_0,p+i-2-\tau_0} \right)[B(\sigma_{\tau_0-i+2})]_{1,\tau_0+1} \\
&\left(\prod_{q=1}^{\tau_0-i+1} [B(\sigma_q)]_{q+i,q+i-1} \right) \geq \mu I_m \geq \mu_1^{\tau_0+1} I_m
\end{aligned}$$

since $[B(\varpi)]_{i+1,i} = I_m$ and $[B(\varpi)]_{1,\tau_0+1} \geq \mu I_m$ for all $\varpi \in \Omega$ and $i \in \underline{\tau_M}$.

(ii). For any $j \geq \tau_0 + 2$, we have

$$[B(\sigma)]_{j,j-(\tau_0+1)} = \sum_{i_1,\ldots,\tau_0} [B(\sigma_{\tau_0+1})]_{j,i_1}[B(\sigma(\tau_0))]_{i_1,i_2}\cdots[B(\sigma_1)]_{i_{\tau_0},j-(\tau_0+1)}$$

$$\geq \prod_{k=1}^{\tau_0+1}[B(\sigma_k)]_{k+j-\tau_0-1,k+j-\tau_0-2} = I_m$$

since $[B(\varpi)]_{i+1,i} = I_m$ for all $i \geq 2$ and $\varpi \in \Omega$.

(iii). For any $i \in \underline{\tau_0+1}$, we have

$$\sum_{i_1,\ldots,i_{\tau_0},k}[B(\sigma)]_{i,i+(\tau_0+1)k} = \sum_k [B(\sigma_{\tau_0+1})]_{i,i_1}[B(\sigma_{\tau_0})]_{i_1,i_2}\cdots[B(\sigma_1)]_{i_{\tau_0},(\tau_0+1)k+i}$$

$$\geq \left(\prod_{k=2}^{i}[B(\sigma_{\tau_0-i+k+1})]_{k,k-1}\right)[B(\sigma_{\tau_0-i+2})]_{1,(k+1)(\tau_0+1)}$$

$$\left(\prod_{l=i+(\tau_0+1)(k+1)-\tau_0}^{(k+1)(\tau_0+1)}[B(\sigma_{l-k(\tau_0+1)-i})]_{l,l-1}\right) \geq [B(\sigma_{\tau_0+2-i})]_{1,(k+1)(\tau_0+1)}$$

for all $k \geq 0$. Summing the right-hand side with respect to k and letting $j = \tau_0 + 2 - i$, we have $\sum_{l\in\langle 1-j\rangle}[B(\sigma)]_{\tau_0+2-j,l} \geq \sum_{l\in\langle\tau_0+1\rangle}[B(\sigma_j)]_{1,l}$.

(iv). Let $j = \tau_0 + 2 - i$. If $j \geq \tau$,

$$\sum_k [B(\sigma)]_{\tau_0+2-j,\tau_0+2-j+(\tau+1)+(\tau_0+1)k} \geq \left(\prod_{p=j+1}^{\tau_0+1}[B(\sigma_p)]_{p-j+1,p-j}\right)$$

$$[B(\sigma_j)]_{1,\tau+1}\left(\prod_{q=j-\tau}^{j-1}[B(\sigma_q)]_{q+\tau+2-j,q+\tau+1-j}\right)[B(\sigma_{j-\tau-1})]_{1,\tau_0+1}$$

$$\left(\prod_{l=1}^{j-\tau-2}[B(\sigma_l)]_{l+\tau+\tau_0+3-j,l+\tau+\tau_0+2-j}\right) \geq \mu[B(\sigma_j)]_{1,\tau+1};$$

whereas if $j < \tau$,

$$\sum_k [B(\sigma)]_{\tau_0+2-j,\tau_0+2-j+(\tau+1)+(\tau_0+1)k} \geq \left(\prod_{p=j+1}^{\tau_0+1}[B(\sigma_p)]_{p-j+1,p-j}\right)[B(\sigma_j)]_{1,\tau+1}$$

$$\left(\prod_{q=1}^{j-1}[B(\sigma_q)]_{q+\tau+2-j,q+\tau+1-j}\right) \geq [B(\sigma_j)]_{1,\tau+1}.$$

These calculations complete the proof of the lemma. □

Lemma 5.3. *Under the conditions of Theorem 3.4, consider an $L(\tau_0+1)$-length word $\tilde{\sigma} = (\tilde{\sigma}_1, \ldots, \tilde{\sigma}_L)$, where each $\tilde{\sigma}_l = (\sigma_{l,i})_{i=1}^{\tau_0+1}$ is a (τ_0+1)-length word. If $L \geq \tau_M + 1$, then there exists $\mu_1 > 0$ such that*

(i). $[B(\tilde{\sigma})]_{j,i} \geq \mu_1^{(\tau_0+1)L} I_m$ *for all $j \in \langle i\rangle$ and $i \in \underline{\tau_0+1}$;*

(ii). $\sum_{l\in\langle i\rangle}[B(\tilde{\sigma})]_{\tau_0+2-j,l} \geq \mu_1^{(\tau_0+1)L}\sum_k \hat{G}^0(\sigma_{k,j})$ *for all $j \in \underline{\tau_0+1}$;*

(iii). $\sum_{j\in\langle i+\tau+1\rangle}[B(\tilde{\sigma})]_{i,j} \geq \mu_1^{\tau_0+1}\sum_{l\in\langle\tau+1\rangle}[B(\tilde{\sigma}_{\tau_0+2-i})]_{1,l}$ *for all $i \in \underline{\tau_0+1}$ and $\tau \in \underline{\tau_M}$;*

(iv). If τ' is such that $\tau'+1 \notin \langle\tau_0+1, \tau_1+1, \ldots, \tau_K+1\rangle$ and $[B(\sigma_1)]_{1,\langle\tau'+1\rangle} = 0$ for all $\sigma_1 \in \Omega$, then $[B(\tilde{\sigma})]_{i,\langle i+\tau'+1\rangle} = 0$ for all $i \geq 1$.

Proof. We pick some $\mu_1 < \mu$. (i). For $j \le \tau_0 + 1$, the proof is similar to the proof of item (i) of Lemma 5.2. For $j \ge \tau_0 + 2$, we have

$$[B(\tilde{\sigma})]_{j,i} \ge \Bigg(\prod_{l=l_1}^{L} [B(\tilde{\sigma}_l)]_{j-(L-l)(\tau_0+1),j-(L-l+1)(\tau_0+1)} \Bigg) \Bigg(\prod_{p=1}^{l_1} [B(\tilde{\sigma}_p)]_{i,i} \Bigg) \ge \mu_1^{(\tau_0+1)L} I_m,$$

where $l_1 = L + 1 - (j - i)/(\tau_0 + 1)$ is an integer (noting $j \in \langle i \rangle$), since $[B(\tilde{\sigma}_l)]_{j-(L-l)(\tau_0+1),j-(L-l+1)(\tau_0+1)} \ge I_m$ holds here, as mentioned in Lemma 5.2 (ii).

The items (ii) and (iii) can be proved by similar arguments as in the proof of items (iii) and (iv) of Lemma 5.2. It remains to prove item (iv). In the following, we will prove a slightly more general result, namely that $[B(\sigma)]_{i,\langle i+\tau'+1\rangle} = 0$ for all words σ having length $L \in \langle \tau_0 + 1 \rangle$. Let $\sigma = (\sigma_i)_{i=1}^{L}$ be an arbitrary L-length word. We calculate $[B(\sigma)]_{i,j}$ with $j \in \langle i + \tau' + 1 \rangle$ as a sum of several matrix product terms:

$$[B(\sigma)]_{i,j} = \sum_{i_1,\dots,i_{L-1}} [B(\sigma_L)]_{i,i_1} [B(\sigma_{L-1})]_{i_1,i_2} \cdots [B(\sigma_1)]_{i_{L-1},j}.$$

Since any zero factor yields zero product, we avoid zero factors in the calculations. That is, in the expression above, only factors of the form $[B(\sigma_l)]_{i+1,i}$ and $[B(\sigma_l)]_{1,j}$ can occur where $j \in \langle i + \tau' + 1 \rangle$ and $\tau' + 1 \notin \langle \tau_0 + 1, \tau_1 + 1, \dots, \tau_K + 1 \rangle$. Thus, letting $j_1 = i$, we have

$$[B(\sigma)]_{i,j} = \sum_{j_1,\dots,j_V,V} \Bigg\{ \Bigg[\prod_{l=1}^{V} \Bigg(\prod_{k_l=L-\sum_{p=1}^{l} j_p+2}^{L-\sum_{p=1}^{l-1} j_p} [B(\sigma_{k_l})]_{\sum_{p=1}^{l} j_p+k_l-L,\sum_{p=1}^{l} j_p+k_l-L-1} \Bigg)$$
$$[B(\sigma_{L-\sum_{p=1}^{l} j_p+1})]_{1,j_p+1} \Bigg] \Bigg(\prod_{k_{V+1}=1}^{L-\sum_{p=1}^{V} j_p} [B(\sigma_{k_{V+1}})]_{L-\sum_{p=1}^{V} j_p+k_{V+1},\sum_{p=1}^{V} L-\sum_{p=1}^{V} j_p+k_{V+1}-1} \Bigg) \Bigg\},$$

where each $j_p \in \langle \tau_0 + 1, \tau_1 + 1, \dots, \tau_K + 1 \rangle$. Suppose that the matrix product is nonzero. Then $j = \sum_{p=1}^{V} j_p - L$, i.e., $\langle (i + \tau' + 1) - (\sum_{p=1}^{V} j_p - L) \rangle = 0$, which implies $\langle \tau' + 1 - \sum_{p=2}^{V} j_p + L \rangle = 0$. This means that $\tau' + 1 \in \langle \tau_0 + 1, \tau_k + 1 : \ k = 1, \dots, K \rangle$, which contradicts the condition $\tau' + 1 \notin \langle \tau_0 + 1, \tau_k + 1 : \ k = 1, \dots, K \rangle$. The lemma is proved. $\square$

Proof of Theorem 3.4. Consider the graph $\hat{\mathcal{G}}^{\delta}(\sigma^t) = \{\hat{\mathcal{V}}, \hat{\mathcal{E}}(\sigma^t)\}$ on $(\tau_M + 1)m$ vertices corresponding to the δ-graph of the matrix $B(\sigma^t)$ as defined at the beginning of this section. For $L \in \mathbb{N}$ as fixed in the main condition of Theorem 3.4 and an arbitrary fixed $m \in \mathbb{N}$, let $\mathbb{B} = \mathbb{E}\{\prod_{t=n+1}^{n+L} B(\sigma^t) | \mathcal{F}^n\}$ and $\hat{\mathcal{G}}^{\delta}$ be the random variable picked in the δ-graphs of $\mathbb{B}$.

First, we divide the graph $\hat{\mathcal{G}}^{\delta}$ into $\tau_M + 1$ subgraphs: $\mathcal{G}_k^{\delta} = \{\mathcal{V}_k, \mathcal{E}_k(\sigma^t)\}$, $k \in \underline{\tau_M + 1}$, where $\mathcal{V}_k = \{v_{k,i} : \ i \in \underline{m}\}$ corresponds to the rows or columns of $\mathbb{B}_{k,k}$ and the vertex $v_{k,i}$ corresponds the i-th row or column of the matrix $\mathbb{B}_{k,k}$. Then, integrate the subgraphs $\{\mathcal{G}_k^{\delta}\}_{k=1}^{\tau_M+1}$ into $\tau_0 + 1$ subgraphs: ${\mathcal{G}'}_l^{\delta} = \{\mathcal{V}'_l, \mathcal{E}'_l\}$, $l \in \underline{\tau_0 + 1}$, where $\mathcal{V}'_l = \bigcup_{k \in \langle l \rangle} \mathcal{V}_k$, $l \in \underline{\tau_0 + 1}$ and $\mathcal{E}'_l$ corresponds to the intra-links in $\mathcal{V}'_l$. Let $\mathcal{E}_{l_1,l_2}$ denote the inter-links from the subgraph of $\mathcal{V}'_{l_2}$ to the subgraph $\mathcal{V}'_{l_1}$. Lemma 5.3 (i) implies that for each $l \in \underline{\tau_0 + 1}$, there must exist a link from $v_{l,i}$ to $v_{k,i}$ in the subgraph ${\mathcal{G}'}_l^{\delta}(\cdot)$, for each vertex $v_{k,i} \in \mathcal{V}_k$ with $k > l$ and $k \in \langle l \rangle$. Similarly to the the proof of Theorem 3.2, the main condition of Theorem 3.4 and items (ii) and (iii) in Lemma 5.3 imply that there exist $\delta_1 > 0$ and $L \in \mathbb{N}$ such that the subgraph ${\mathcal{G}'}_l^{\delta_1}$ is strongly connected, consequently having a spanning tree, and each vertex in $\mathcal{V}_l$ is one of the roots in ${\mathcal{G}'}_l^{\delta_1}$ and has a self-link almost surely for all $l \in \underline{\tau_0 + 1}$.

Second, according to $\gcd(\tau_0+1, \tau_k+1:\ k\in\underline{K}) = P$, we integrate the subgraphs ${\mathcal{G}'}_l^{\delta_1}$ for all $l \in \underline{\tau_0+1}$, into P subgraphs, denoted by $\tilde{\mathcal{G}}_p^{\delta_1} = \{\tilde{\mathcal{V}}_p, \tilde{\mathcal{E}}_p\}$, $p\in\underline{P}$ by $\tilde{\mathcal{V}}_p = \{\mathcal{V}'_j:\ \mathcal{E}_{j,p}\neq\emptyset\}$. The items (ii) and (iii) in Lemma 5.3 and the second item in condition $\mathbf{B}$ indicate that the δ_1-matrix of $\sum_{j\in\langle\tau_k+1:\ k=0,1,\dots,K\rangle}\mathbb{B}_{l,l+j}$ is positive for all $l\in\underline{\tau_0+1}$. This implies that there exists at least one link from ${\mathcal{G}'}_{l+j}^{\delta_1}$ to ${\mathcal{G}'}_l^{\delta_1}$ and this link end in $\mathcal{V}_l$. So, in the graph ${\mathcal{G}'}^{\delta_1}$, the root vertex in ${\mathcal{G}'}_{l+j}^{\delta_1}$ can reach all vertices in ${\mathcal{G}'}_l^{\delta_1}$ since each vertex in V_l is a root vertex in ${\mathcal{G}'}_l^{\delta_1}$. This leads to the conclusion that $\mathcal{V}'_j\subset\tilde{\mathcal{V}}_l$ provided $j-l\in\langle\tau_k+1: k=0,1,\dots,K\rangle$. Also, we can conclude that each root vertex in ${\mathcal{G}'_{l+j}}^{\delta_1}$ can reach all vertices in ${\mathcal{G}'}_l^{\delta_1}$, by item (i) in Lemma 5.3. Therefore, we can conclude that $\tilde{\mathcal{V}}_p = \bigcup_{l\in\langle p\rangle_P}\mathcal{V}'_p$ and each $\tilde{\mathcal{G}}_p$ has a spanning tree almost surely. This proves Claim 1. Moreover, there exists a vertex with self-link in $\mathcal{V}_i$, $i\in\underline{\tau_0+1}$ and $i\in\langle p\rangle_P$, as one of its roots, in $\hat{\mathcal{G}}^{\delta_1}$. So, according to the arbitrariness of integer n, we can conclude that the δ_1-graph of $\mathbb{E}\{\prod_{t=n+1}^{n+L}\hat{B}_p(\sigma^t)|\mathcal{F}^n\}$ is SIA almost surely for all $n\in\mathbb{N}$.

Finally, according to the second item in condition $\mathbf{B}$ and the (iv) item in Lemma 5.3, one can conclude that there are no links between the graph $\tilde{\mathcal{G}}_p^{\delta}$ for different $p\in\underline{P}$ for any $\delta\geq 0$. So, by a permutation matrix Q corresponding to the permutation sequence from $(1,2\dots,\tau_M+1)$ to $(\langle 1\rangle,\langle 2\rangle,\dots,\langle P\rangle)$, $[Q\otimes I_m]B(\sigma^t)[Q\otimes I_m]^\top$ has the form (15).

By Theorem 3.1, we can conclude that (17) reaches consensus for all $p=1,\dots,P$, but converges to different values except for initial values in a set of Lebesgue measure zero. Therefore, x^t can synchronize and converge to a P-periodic trajectory. This completes the proof of Theorem 3.4.

6. **Conclusions.**

In this paper we have studied the convergence of the consensus algorithm in multi-agent systems with stochastically switching topologies and time delays. We have shown that consensus can be obtained if the graph corresponding to the conditional expectations of the coupling matrix product in consecutive times has spanning trees almost surely and self-links are possible. With multiple delays, if self-links always exist and are instantaneous (undelayed), then consensus can be guaranteed for arbitrary bounded delays. Moreover, when the self-links are also delayed, we have shown the phenomenon that the algorithm may not reach consensus but instead may synchronize to a periodic trajectory according to the delay patterns. Finally, we have briefly studied consensus algorithms without self-links. We have presented several results for i.i.d. and Markovian switching topologies as special cases.

Acknowledgments. W. Lu is supported by the National Key Basic Research and Development Program (No. 2010CB731403), the National Natural Sciences Foundation of China under Grant No. 60804044, the Foundation for the Author of National Excellent Doctoral Dissertation of China No. 200921, SGST 09DZ2272900, and LMNS, Fudan University.

REFERENCES

[1] P.-A. Bliman and G. Ferrari-Trecate, *Average consensus problems in networks of agents with delayed communications*, Automatica, **44** (2008), 1985–1995.
[2] M. Cao, A. S. Morse and B. D. O. Anderson, *Reaching a consensus in a dynamically changing environment: a graphical approach*, SIAM J. Control Optim., **47** (2008), 575–600.
[3] S. Chatterjee and E. Seneta, *Towards consensus: Some convergence theorems on repeated averaging*, J. Appl. Prob., **14** (1977), 89–97.
[4] O. Chilina, "f-uniform ergodicity of Markov chains", Supervised Project, Unversity of Toronto, 2006.
[5] M. H. DeGroot, *Reaching a consensus*, J. Amer. Statist. Assoc., **69** (1974), 118–121.

[6] D. V. Dimarogonasa and K. H. Johansson, *Stability analysis for multi-agent systems using the incidence matrix: Quantized communication and formation control*, Automatica, **46** (2010), 695–700.

[7] R. Durrett, "Probability: Theory and Examples", 3^{rd} edition, Belmont, CA: Duxbury Press, 2005.

[8] F. Fagnani and S. Zampieri, *Average consensus with packet drop communication*, SIAM J. Control Optim., **48** (2009), 102–133.

[9] L. Fang, P. J. Antsaklis and A. Tzimas, *Asynchronous consensus protocols: Preliminary results, simulations and open questions*, Proceedings of the 44th IEEE Conf. Decision and Control, the Europ. Control Conference (2005), 2194–2199.

[10] J. A. Fax and R. M. Murray, *Information flow and cooperative control of vehicle formations*, IEEE Trans. Autom. Control, **49** (2004), 1465–1476.

[11] C. Godsil and G. Royle, "Algebraic graph theory", Springer-Verlag, New York, 2001.

[12] J. Hajnal, *The ergodic properties of non-homogeneous finite Markov chains*, Proc. Camb. Phil. Soc., **52** (1956), 67–77.

[13] J. Hajnal, *Weak ergodicity in non-homogeneous Markov chains*, Proc. Camb. Phil. Soc., **54** (1958), 233–246.

[14] Y. Hatano and M. Mesbahi, *Agreement over random networks*, IEEE Trans. Autom. Control, **50** (2005), 1867–1872.

[15] R. A. Horn and C. R. Johnson, "Matrix analysis", Cambridge University Press, 1985.

[16] Y. Kuramoto, "Chemical Oscillations, waves, and turbulence", Springer-Verlag, New York, 1984.

[17] J. Lin, A. S. Morse and B. D. O. Anderson, *The multi-agent rendezvous problem Part 2: The asynchronous case*, SIAM J. Control Optim., **46** (2007), 2120–2147.

[18] B. Liu, W. Lu and T. Chen, *Consensus in networks of multiagents with switching topologies modeled as adapted stochastic processes*, SIAM J. Control Optim., **49** (2011), 227–253.

[19] W. Lu, F. M. Atay and J. Jost, *Synchronization of discrete-time networks with time-varying couplings*, SIAM J. Math. Analys., **39** (2007), 1231–1259.

[20] W. Lu, F. M. Atay and J. Jost, *Chaos synchronization in networks of coupled maps with time-varying topologies*, Eur. Phys. J. B, **63** (2008), 399–406.

[21] N. A. Lynch, "Distributed algorithms", CA: Morgan Kaufmann, San Francisco, 1996.

[22] W. Ni and D. Z. Cheng, *Leader-following consensus of multi-agent systems under fixed and switching topologies*, Systems & Control Letters, **59** (2010), 209–217.

[23] W. Michiels, C.-I. Morărescu and S.-I. Niculescu, *Consensus problems with distributed delays, with application to traffic flow models*, SIAM J. Control Optim., **48** (2009), 77–101.

[24] L. Moreau, *Stability of continuous-time distributed consensus algorithms*, 43rd IEEE Conference on Decision and Control, **4** (2004), 3998–4003.

[25] L. Moreau, *Stability of multiagent systems with time-dependent communication links*, IEEE Trans. Autom. Control, **50** (2005), 169–182.

[26] R. Olfati-Saber and J. S. Shamma, *Consensus filters for sensor networks and distributed sensor fusion*, 44th IEEE Conference on Decision and Control 2005, and 2005 European Control Conference CDC-ECC '05. 6698–6703.

[27] R. Olfati-Saber, J. A. Fax and R. M. Murray, *Consensus and cooperation in networked multi-agent systems*, Proceedings of the IEEE, **95** (2007), 215–233.

[28] R. Olfati-Saber and R. M. Murray, *Consensus problems in networks of agents with switching topology and time-delays*, IEEE Trans. Autom. Control, **49** (2004),1520–1533.

[29] A. Pikovsky, M. Rosenblum and J. Kurths, *Synchronization: A universal concept in nonlinear sciences*, Cambridge University Press, 2001.

[30] J. Shen, *A geometric approach to ergodic non-homogeneous Markov chains*, Wavelet Anal. Multi. Meth., LNPAM, **212** (2000), 341–366.

[31] A. Tahbaz-Salehi and A. Jadbabaie, *A necessary and sufficient condition for consensus over random networks*, IEEE Trans. Autom. Control, **53** (2008), 791–795.

[32] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen and O. Shochet, *Novel type of phase transition in a system of self-driven particles*, Phys. Rev. Lett, **75** (1995), 1226–1229.

[33] A. T. Winfree, "The Geometry of biological time", Springer Verlag, New York, 1980.

[34] J. Wolfowitz, *Products of indecomposable, aperiodic, stochastic matrices*, Proceedings of AMS, **14** (1963), 733–737.

[35] C. W. Wu, *Synchronization and convergence of linear dynamics in random directed networks*, IEEE Trans. Autom. Control, **51**(2006), 1207–1210.

[36] F. Xiao and L. Wang, *Consensus protocols for discrete-time multi-agent systems with time-varying delays*, Automatica, **44** (2008), 2577–2582.

[37] F. Xiao, L. Wang, *Asynchronous consensus in continuous-time multi-agent systems with switching topology and time-varying delays*, IEEE Transactions on Automatic Control, **53** (2008), 1804–1816.

[38] Y. Zhang and Y.-P. Tian, *Consentability and protocol design of multi-agent systems with stochastic switching topology*, Automatica, **45** (2009), 1195–1201.

E-mail address: wenlian@fudan.edu.cn
E-mail address: atay@mis.mpg.de
E-mail address: jost@mis.mpg.de