Distributed coordination and partial synchronization in complex networks

Qin, Yuzhen

DOI:
10.33612/diss.108085222

4 Stochastic Distributed Coordination Algorithms

In this chapter, we deal with several stochastic distributed coordination algorithms, which is the central aim of Part I. The new stochastic Lyapunov criteria developed in Chapter 3 will be used to prove the convergence of these stochastic algorithms.

4.1 Introduction

Distributed coordination algorithms, known as distributed weighted averaging algorithms, have been playing crucial roles in various distributed systems and algorithms, including distributed optimization [25, 26], distributed control of networked robots [112], opinion dynamics [6, 32, 115, 116], and many other distributed algorithms [8, 9, 9–11, 35, 36]. In order to analyze such systems and algorithms, one frequently encounters the need to prove convergence of inhomogeneous Markov chains, or equivalently the convergence of backward products of random sequences of stochastic matrices $\{W(k)\}$. Most of the existing results assume exclusively that all the $W(k)$ in the sequence have all positive diagonal entries, see e.g., [73, 128, 129]. This assumption simplifies the analysis of convergence significantly; moreover, without this assumption, the existing results do not always hold. For example, from [35, 36] one knows that the product of $W(k)$ converges to a rank-one matrix almost surely if exactly one of the eigenvalues of the expectation of $W(k)$ has the modulus of one, which can be violated if $W(k)$ has zero diagonal elements. Note also that most of the existing results are confined to special random sequences, e.g., independently distributed sequences [35], stationary ergodic sequences [36], or independent sequences [75, 76]. In the first part of this chapter, we work on more general classes of random sequences of stochastic matrices without the assumption of non-zero diagonal entries. Using the novel Lyapunov criteria we developed in Chapter 3, we show that if there exists a fixed length such that the product of any successive subsequence of matrices of this length has the *scrambling* property (see the definition in Section 2.3) with positive probability, the convergence to a rank-one matrix for the infinite product can be guaranteed almost surely. We also prove that the convergence can be exponentially fast if this probability is lower bounded by some positive number, and the greater the lower bound is, the faster the convergence becomes. For some particular random sequences, we further relax this "scrambling" condition. If the random sequence is driven by a *stationary* process, the almost sure convergence can be ensured as long as the product of any successive subsequence of finite length has a positive probability to be indecomposable and aperiodic (SIA). The exponential convergence rate follows without other assumptions if the random process that governs the evolution of the sequence is a *stationary ergodic* process.

Using these results on products of random stochastic matrices, we then investigate a classic agreement problem, in which agents coupled by a network repeatedly update their states to the weighted average of their neighbors' states and their own. This problem is usually modeled by a linear recursion equation $x(k) = Wx(k-1)$ with W a stochastic matrix describing the interaction structure. The agreement problem is equivalent to studying whether W^k converge to a rank-one matrix. Usually, W is required to be indecomposable and aperiodic matrix (SIA) [68, 71]. However, the case when W is not an SIA matrix has not been studied before. For example, a periodic W leads to oscillating behaviors. We address the agreement problem when W is periodic in Section 4.3. We show that, instead of oscillation, agreement takes place if the agents update asynchronously. Specifically, we assume that each agent has access to its own state while executing averaging actions at every time instant. In other words, at each time step, a random number of agents are activated and then update. In sharp contrast to the existing works, e.g. [130, 131] and [129], agents do not need to use their own states to update. The obtained results reveal that asynchrony can play a very important role in giving rise to an agreement.

We then investigate another distributed coordination algorithm for solving linear algebraic equations of the form $Ax = b$, as another application of the finite-step stochastic Lyapunov criteria in Chapter 3. The problem is to design a distributed algorithm such that the equations are solved in parallel by n agents, each of whom just knows a subset of the rows of the matrix $[A, b]$. Each agent recursively updates its estimate of the solution using the current estimates from its neighbors. Recently several solutions under different sufficient conditions have been proposed [29, 30, 77], and particularly in [77], the sequence of the neighbor relationship graphs $\mathcal{G}(k)$ is required to be repeated jointly strongly connected. We show that a much weaker condition is sufficient to solve the problem almost surely, namely the algorithm in [77] works if there exists a fixed length such that any subsequence of $\{\mathcal{G}(k)\}$ at this length is jointly strongly connected with positive probability. The proof also relies on the new Lyapunov criteria we developed in the previous section.

Outline

The remainder of this chapter is structured as follows. Products of random sequences of stochastic matrices are studied in Section 4.2. We investigate asynchronous updating induced agreement problem in Section 4.3. A distributed algorithm to solve linear equation is studied in Section 4.4. Concluding remarks appear in Section 4.5.

4.2 Products of Random Sequences of Stochastic Matrices

In this section, we study the convergence of products of stochastic matrices, where the obtained results on finite-step Lyapunov functions are used for analysis. Let $\Omega_0 := \{1, 2, \dots, m\}$ be the state space and $\mathcal{M} := \{F_1, F_2, \dots, F_m\}$ be the set of m stochastic matrices $F_i \in \mathbb{R}^{n\times n}$. Consider a random sequence $\{W_\omega(k) : k \in \mathbb{N}\}$ on the probability space $(\Omega, \mathcal{F}, \Pr)$, where Ω is the collection of all infinite sequences $\omega = (\omega_1, \omega_2, \dots)$ with $\omega_k \in \Omega_0$, and we define $W_\omega(k) := F_{\omega_k}$. For notational simplicity, we denote $W_\omega(k)$ by $W(k)$. For the backward product of stochastic matrices

$$W(t + k, t) = W(t + k) \cdots W(t + 1), \tag{4.1}$$

where $k \in \mathbb{N}, t \in \mathbb{N}_0$, we are interested in establishing conditions on $\{W(k)\}$, under which it holds that $\lim_{k\to\infty} W(k, 0) = L$ for a random matrix $L = \mathbf{1}\xi^\top$ where $\xi \in \mathbb{R}^n$ satisfies $\xi^\top \mathbf{1} = 1$.

Before proceeding, let us introduce some concepts in probability. Let $\mathcal{F}_k = \sigma(W(1), \dots, W(k))$, so that evidently $\{\mathcal{F}_k\}$, $k = 1, 2, \dots$, is an increasing sequence of σ-fields. Let $\chi : \Omega \to \Omega$ be the shift operator, i.e., $\chi(\omega_1, \omega_2, \dots) = (\omega_2, \omega_3, \dots)$. A random sequence of stochastic matrices $\{W(1), W(2), \dots, W(k), \dots\}$ is said to be *stationary* if the shift operator is measure-preserving. In other words, for any $k_1, k_2, \dots, k_r$ and $\tau \in \mathbb{N}$, the sequence

$$\{W(k_1 + \tau), W(k_2 + \tau), \dots, W(k_r + \tau)\}$$

has the same joint distribution as $\{W(k_1), W(k_2), \dots, W(k_r)\}$. Moreover, a sequence is said to be *stationary ergodic* if it is stationary, and every invariant set $\mathcal{B}$ is trivial, i.e., for every $A \in \mathcal{B}$, $\Pr[A] \in \{0, 1\}$. Here by an invariant set $\mathcal{B}$, we mean $\chi^{-1}\mathcal{B} = \mathcal{B}$.

4.2.1 Convergence Results

In this subsection, we provide some sufficient conditions such that the backward product of the sequence $\{W(k)\}$ converges to a rank one matrix.

We first recall three classes of stochastic matrices defined in Section 2.3, denoted by $\mathcal{M}_1, \mathcal{M}_2$, and $\mathcal{M}_3$, respectively. Give a stochastic matric $A \in \mathbb{R}^{n\times n}$, we say $A \in \mathcal{M}_1$ if A is SIA (stochastic, indecomposable, and aperiodic); $A \in \mathcal{M}_2$ if A is scrambling; and $A \in \mathcal{M}_3$ if A is Markov.

Coefficients of ergodicity serve as a fundamental tool in analyzing the convergence of products of stochastic matrices. In this chapter, we employ a standard one. For a stochastic matrix $A \in \mathbb{R}^{n\times n}$, the coefficient of ergodicity $\tau(A)$ is defined by

$$\tau\left(A\right) = 1 - \min\nolimits_{i,j} \sum\nolimits_{s=1}^{n} \min(a_{is}, a_{js}). \tag{4.2}$$

It is known that this coefficient of ergodicity satisfies $0 \leq \tau(A) \leq 1$, and $\tau(A)$ is proper since $\tau(A) = 0$ if and only if all the rows of A are identical. Importantly, it holds that

$$\tau(A) < 1 \tag{4.3}$$

if and only if $A \in \mathcal{M}_2$ (see [71, p.82]). For any two stochastic matrices A, B, there is an important property for this coefficient of ergodicity

$$\tau(AB) \leq \tau(A)\tau(B). \tag{4.4}$$

This property will be used also in the proof in Section 4.6. Before providing our first results in this subsection, we make the following assumption for the random sequence $\{W(k)\}$.

Assumption 4.1. *Suppose the sequence of stochastic matrices $\{W(k) : k \in \mathbb{N}\}$ is driven by a random process satisfying the following conditions:*

a) There exists an integer $h > 0$ such that for any $k \in \mathbb{N}_0$, it holds that

$$\Pr\left[W(k+h,k) \in \mathcal{M}_2\right] > 0, \tag{4.5}$$

$$\sum_{i=1}^{\infty} \Pr\left[W\big(k+ih, k+(i-1)\,h\big) \in \mathcal{M}_2\right] = \infty; \tag{4.6}$$

b) There is a positive number α such that $W_{ij}(k) \geq \alpha$ for any $i, j \in \mathbf{N}, k \in \mathbb{N}_0$ satisfying $W_{ij}(k) > 0$.

In other words, Assumption 4.1 requires that any corresponding matrix product of length h becomes a scrambling matrix with positive probability, and the positive elements for any matrix in $\mathcal{M}$ are uniformly lower bounded away from some positive value. Now we are ready to provide our main result on the convergence of stochastic matrices' products.

Theorem 4.1. *Under Assumption 4.1, the product of the random sequence of stochastic matrices* $W(k, 0)$ *converges to a random matrix* $L = \mathbf{1}\xi^\top$ *almost surely.*

To prove Theorem 4.1, consider the stochastic discrete-time dynamical system described by

$$x_{k+1} = F_{y(k+1)}x_k := W(k+1)x_k, \qquad k \in \mathbb{N}_0, \tag{4.7}$$

where $x_k \in \mathbb{R}^n$; the initial state x_0 is a constant with probability one; $y(k) \in \{1, \dots, m\}$ is regarded as the randomly switching signal; and $\{W(1), W(2), \dots\}$ is the random process of stochastic matrices we are interested in. One knows that x_k is adapted to $\mathcal{F}_k$. Thus, to investigate the limiting behavior of the product (4.1), it is sufficient to study the limiting behavior of system dynamics (4.7). We say the state of system (4.7) reaches an *agreement* state if $\lim_{k\to\infty} x_k = \mathbf{1}\xi$ for some $\xi \in \mathbb{R}$. Then, from [75] one can say that the agreement of system (4.7) for any initial state x_0 implies that $W(k, 0)$ converges to a rank-one matrix as $k \to \infty$.

To investigate the agreement problem, we define

$$\lceil x_k \rceil := \max_{i\in\mathbf{N}} x_k^i, \qquad \lfloor x_k \rfloor := \min_{i\in\mathbf{N}} x_k^i$$

where $k \in \mathbb{N}_0$, and

$$v_k = \lceil x_k \rceil - \lfloor x_k \rfloor. \tag{4.8}$$

For any $k \in \mathbb{N}$, v_k is adapted to $\mathcal{F}_k$ since x_k is. The agreement is said to be reached asymptotically almost surely if $v_k \xrightarrow{a.s.} 0$ as $k \to \infty$; and it is said to be reached exponentially almost surely with convergence rate no slower than γ^{-1} for some $\gamma > 1$ if $\gamma^k v_k \xrightarrow{a.s.} y$ for some finite $y \geq 0$. The random variable v_k has some important properties given by the following proposition.

Proposition 4.1. *Consider a system* $x_{k+1} = Ax_k, k \in \mathbb{N}_0$*, where* A *is a stochastic matrix. For* v_k *defined in* (4.8)*, it follows that* $v_{k+1} \leq v_k$*, and the strict inequality holds for any* $x_k \notin \operatorname{span}(\mathbf{1})$ if and only if A *is scrambling.*

Proof. It is shown in [71] that $v_{k+1} \leq \tau(A)v_k$ with $\tau(\cdot)$ defined in (4.2). Therefore, the sufficiency follows from (4.3) straightforwardly. We then prove the necessity by contradiction. Suppose A is not scrambling, and then there must exist at least two rows, denoted by i, j, that are orthogonal. Define the two sets $\mathbf{i} := \{l : a_{il} > 0, l \in \mathbf{N}\}$ and $\mathbf{j} := \{m : a_{jm} > 0, m \in \mathbf{N}\}$, respectively. It follows then from the scrambling property that $\mathbf{i} \cap \mathbf{j} = \emptyset$. Let $x_k^q = 1$ for all $q \in \mathbf{i}$, $x_k^q = 0$ for all $q \in \mathbf{j}$, and let x_k^m be any arbitrary positive number less than 1 for all $m \in \mathbf{N}\backslash(\mathbf{i} \cup \mathbf{j})$ if $\mathbf{N}\backslash(\mathbf{i} \cup \mathbf{j})$ is not empty. Then the states of i and j at time $k+1$ become

$$x_{k+1}^{i} = \sum\nolimits_{l=1}^{n} a_{il} x_k^l = \sum\nolimits_{l\in\mathbf{i}} a_{il} x_k^l = 1,$$
$$x_{k+1}^{j} = \sum\nolimits_{l=1}^{n} a_{jl} x_k^l = \sum\nolimits_{l\in\mathbf{j}} a_{jl} x_k^l = 0,$$

and $0 \leq x_{k+1}^m \leq 1$ for all $m \in \mathbf{N}\backslash(i \cup j)$. This results in $v_{k+1} = v_k = 1$. By contradiction one knows that a scrambling A is necessary for $v_{k+1} < v_k$, which completes the proof. □

In order to prove Theorem 4.1, we obtain the following intermediate result.

Proposition 4.2. *For any scrambling matrix $A \in \mathbb{R}^{n\times n}$, the coefficient of ergodicity $\tau(A)$ defined in* (4.2) *satisfies*

$$\tau(A) \leq 1 - \gamma$$

if all the positive elements of A are lower bounded by $\gamma > 0$.

Proof. Consider any two rows of A, denoted by i, j. Define two sets, $\mathbf{i} := \{s : a_{is} > 0\}$ and $\mathbf{j} := \{s : a_{js} > 0\}$. From the scrambling hypothesis, one knows that $\mathbf{i} \cap \mathbf{j} \neq \emptyset$. Thus it holds that

$$\sum_{s=1}^{n} \min(a_{is}, a_{js}) = \sum_{s\in\mathbf{i}\cap\mathbf{j}} \min(a_{is}, a_{js}) \geq \gamma.$$

Then from the definition of $\tau(A)$, it is easy to see

$$\tau(A) = 1 - \min_{i,j} \sum_{s=1}^{n} \min(a_{is}, a_{js}) \leq 1 - \gamma,$$

which completes the proof. □

We are in the position to prove Theorem 4.1 by showing that $v_k \xrightarrow{a.s.} 0$ as $k \to \infty$, where the results obtained in Corollary 3.3 will be used.

Proof of Theorem 4.1. Let $V(x_k) = v_k$ be a finite-step stochastic Lyapunov function candidate for the system dynamics (4.7). It is easy to see $V(x) = 0$ if and only if $x \in \text{span}(\mathbf{1})$. Since all $W(k)$ are stochastic matrices, we observe that

$$\mathbb{E}[V(x_{k+1})|\mathcal{F}_k] - V(x_k) \leq 0$$

from Proposition 4.1, which implies that $V(x_k)$ is exactly a supermartingale with respect to $\mathcal{F}_k$. From Lemma 3.3, we know $V(x_k) \xrightarrow{a.s.} \bar{V}$ for some $\bar{V}$ because $V(x_k) \geq 0$

and $\mathbb{E}V(x_k) < \infty$. From Assumption 4.1, we know that there is an h such that the product $W(k+h,k)$ is scrambling with positive probability for any k. Let $\mathcal{W}_k$ be the set of all possible $W(k+h,k)$ at time k, and n_k the cardinality of $\mathcal{W}_k$. Let n_k^s be the number of scrambling matrices in $\mathcal{W}_k$. We denote each of these scrambling matrices and each of non-scrambling matrices by $S_k^i, i = 1, \dots, n_k^s$ and $\bar{S}_k^j, j = 1, \dots, n_k - n_k^s$, respectively. The probabilities of all the possible $W(k+h,k)$ sum to 1, i.e.,

$$\sum_{i=1}^{n_k^s} \Pr\left[S_k^i\right] + \sum_{j=1}^{n_k - n_k^s} \Pr\left[\bar{S}_k^j\right] = 1. \tag{4.9}$$

Then the conditional expectation of $V(x)$ after finite steps for any k becomes

$$\begin{aligned}
&\mathbb{E}\left[V\left(x_{k+h}\right)\middle|\mathcal{F}_k\right] - V\left(x_k\right) \\
&\qquad = \mathbb{E}\left[V\big(W\left(k+h,k\right)x_k\big)\right] - V\left(x_k\right) \\
&\qquad \le \mathbb{E}\left[\tau\big(W\left(k+h,k\right)\big)\right] V\big(x_k\big) - V\left(x_k\right),
\end{aligned}$$

where $\tau(\cdot)$ is given by (4.2). One can calculate that

$$\begin{aligned}
&\mathbb{E}\left[\tau\Big(W\left(k+h,k\right)\Big)\right] - 1 \\
&= \sum\nolimits_{i=1}^{n_k^s} \Pr\left[S_k^i\right]\tau\left(S_k^i\right) + \sum\nolimits_{j=1}^{n_k-n_k^s} \Pr\left[\bar{S}_k^j\right]\tau\left(\bar{S}_k^j\right) - 1 \\
&\le \sum\nolimits_{i=1}^{n_k^s} \Pr\left[S_k^i\right]\Big(\tau\left(S_k^i\right) - 1\Big),
\end{aligned}$$

where Proposition 4.1 and equation (4.9) have been used. From Assumption 4.1.b), we know that the positive elements of $W(k)$ are lower bounded by α, and thus the positive elements of S_k^i in (4.10) are lower bounded by α^h. Thus $\tau(S_k^i) \le 1 - \alpha^h$ according to Proposition 4.2, and it follows that

$$\begin{aligned}
&\mathbb{E}[V\left(x_{k+h}\right)|\,\mathcal{F}_k] - V\left(x_k\right) \\
&\qquad \le -\sum\nolimits_{i=1}^{n_k^s} \Pr\left[S_k^i\right]\alpha^h \mathbb{E}V(x_k) := \varphi_k\left(x_k\right).
\end{aligned} \tag{4.10}$$

By iterating, one can easily show that

$$\begin{aligned}
\mathbb{E}\left[V\left(x_{nh}\right)\right] - V\left(x_0\right) &\le -\sum\nolimits_{k=0}^{n-1} \varphi_k\left(x_k\right) \\
&= -\sum\nolimits_{k=0}^{n-1}\sum\nolimits_{i=1}^{n_k^s} \Pr\left[S_k^i\right]\alpha^h \mathbb{E}V(x_k).
\end{aligned} \tag{4.11}$$

It then follows that $V\left(x_0\right) - \mathbb{E}\left[V\left(x_{nh}\right)\right] < \infty$ even when $n \to \infty$, since $V(x) \ge 0$. According to the condition (4.6), we know $\sum_{k=0}^{n-1}\sum_{i=1}^{n_k^s} \Pr\left[S_k^i\right] = \infty$. By contradiction, it is easy to infer that $\mathbb{E}V(x_k) \xrightarrow{a.s.} 0$. Since we have already shown that $V(x_k) \xrightarrow{a.s.} \bar{V}$

for some random $\bar{V} \geq 0$, one can conclude that $V(x_k) \xrightarrow{a.s.} 0$. For any given $x_0 \in \mathbb{R}^n$, define the compact set $\mathcal{Q} := \{x : \lceil x \rceil \leq \lceil x_0 \rceil, \lfloor x \rfloor \geq \lfloor x_0 \rfloor$. For any random sequence $\{W(k)\}$, it follows from the system dynamics (4.7) that

$$\begin{aligned} \lceil x_k \rceil \leq \lceil x_{k-1} \rceil \leq \cdots \leq \lceil x_1 \rceil \leq \lceil x_0 \rceil, \\ \lfloor x_k \rfloor \geq \lfloor x_{k-1} \rfloor \geq \cdots \geq \lfloor x_1 \rfloor \geq \lfloor x_0 \rfloor, \end{aligned}$$

and thus x_k will remain within $\mathcal{Q}$. From Corollary 3.3, we know that x_k asymptotically converges to $\{x \in \mathcal{Q} : \varphi_k(x) = 0\}$, or equivalently, $\{x \in \mathcal{Q} : V(x) = 0\}$ almost surely as $k \to \infty$ since $V(x)$ is continuous. In other words, for any $x_0 \in \mathbb{R}^n$, $x_k \xrightarrow{a.s.} \zeta \mathbf{1}$ for some $\zeta \in \mathbb{R}$, which proves Theorem 4.1. □

Compared to the existing results, Theorem 4.1 has provided a quite relaxed condition for the convergence of the backward product (4.1) determined by the random sequence $\{W(k)\}$ to a rank-one matrix: over any time interval of length h, i.e., $[h+k, k]$ for any $k \in \mathbb{N}_0$, the product $W(k+h) \cdots W(k+1)$ has positive probability to be scrambling. The following corollary follows straightforwardly since a Markov matrix is certainly scrambling.

Corollary 4.1. *For a random sequence $\{W(k) : k \in \mathbb{N}\}$, the product* (4.1) *converges to a random matrix $L = \mathbf{1}\xi^\top$ almost surely if there exists an integer h such that for any k the product $W(k+h, k)$ is a Markov matrix with positive probability and*

$$\sum_{i=1}^{\infty} \Pr\left[W\left(k+ih, k+(i-1)\,h\right) \in \mathcal{M}_3\right] = \infty.$$

Next we assume that the sequence $\{W(k)\}$ is driven by an underlying *stationary* process. Then the condition in Theorem 4.1 can be further relaxed. Let us make the following assumption and provide another theorem in this subsection.

Assumption 4.2. *Suppose the random sequence of stochastic matrices $\{W(k) : k \in \mathbb{N}\}$ is driven by a stationary process satisfying the following conditions:*

a) There exists an integer $h > 0$ such that for any $k \in \mathbb{N}_0$, it holds that

$$\Pr\left[W(k+h, k) \in \mathcal{M}_1\right] > 0; \tag{4.12}$$

b) There is a positive number α such that $W_{ij}(k) \geq \alpha$ for any $i, j \in \mathbf{N}, k \in \mathbb{N}_0$ satisfying $W_{ij}(k) > 0$.

In other words, Assumption 4.2 requires that any corresponding matrix product of length h becomes an SIA matrix with positive probability, and the positive elements for any matrix in $\mathcal{M}$ are uniformly lower bounded away from some positive value.

Theorem 4.2. *Under Assumption 4.2, the product of the random sequence of stochastic matrices* $W(k,0)$ *converges to a random matrix* $L = \mathbf{1}\xi^\top$ *almost surely.*

Recall in Section 2.3 that we denote $A_1 \sim A_2$ if these two stochastic matrices are of the same type (have zero elements in the same positions). Obviously, it holds the trivial case $A_1 \sim A_1$. One knows that for any SIA matrix A, there exists an integer l such that A^l is scrambling; it is easy to extend this to the inhomogeneous case, i.e., any product of l stochastic matrices of the same type of A is scrambling if all the matrices' elements are lower bounded away by some positive number. We are now ready to prove Theorem 4.2.

Proof of Theorem 4.2. Since $\{W(k)\}$ is driven by a stationary process, we know that for any $t \in \mathbb{N}_0, h \in \mathbb{N}$, $\{W(t+h), \ldots, W(t+1)\}$ has the same joint distribution as $\{W(t+2h), \ldots, W(t+h+1)\}$. For the h given in Assumption 4.2, there exists an SIA matrix A such that $\Pr[W\big(t+kh+h, t+kh+1\big) = A] > 0$. Thus it follows that $\Pr[W\big(t+kh+2h, t+kh+1\big) = A] > 0$ for any $k \in \mathbb{N}_0$. Thus

$$\Pr\left[\begin{array}{c} W\big(t+(k+2)h, t+(k+1)h\big) \\ \sim W\big(t+(k+1)h, t+kh\big) \end{array} \middle| W(h, t+kh)\right] > 0.$$

When $W(t+h,t) \in \mathcal{M}_1$, which happens with positive probability, we have

$$\begin{aligned} &\Pr\left[W(t+2h,t+h) \sim W(t+h,t), W(t+h,t) \in \mathcal{M}_1\right] \\ &= \Pr\left[\begin{array}{c} W(t+2h,t+h) \\ \sim W(t+h,t) \end{array} \middle| \Pr\left[W(t+h,t) \in \mathcal{M}_1\right]\right] \Pr\left[W(t+h,t) \in \mathcal{M}_1\right] > 0. \end{aligned}$$

By recursion one can conclude that all the m products $W(t+(k+1)h, t+kh), k \in \{0, \ldots, m-1\}$, occur as the same SIA type with positive probability. Since all the products $W(t+(k+1)h, t+kh)$ are of the same type, one can choose m such that $W(t+mh, t)$ is scrambling. This in turn implies that $\Pr\left[W(t+mh,t) \in \mathcal{M}_2\right] > 0$, and the property of stationary process makes sure that (4.6) holds. The conditions in Assumption 4.1 are therefore all satisfied, and then Theorem 4.2 follows from Theorem 4.1. □

Remark 4.1. *Theorems 4.1 and 4.2 have established some sufficient conditions for the convergence of a random sequence of stochastic matrices to a rank-one matrix. A further question is how these results can be applied to controlling distributed computation processes. To answer this question, let us still consider a finite set of stochastic matrices* $\mathcal{M} = \{F_1 \ldots, F_m\}$*, from which each* $W(k)$ *in the random sequence* $\{W(k)\}$ *is sampled. It is defined in [132] that* $\mathcal{M}$ *is a consensus set if the arbitrary product* $\prod_{i=1}^{k} W(i), W(i) \in \mathcal{M}$*, converges to a rank-one matrix. However, it has also been* *shown that to decide whether $\mathcal{M}$ is a consensus set is an NP-hard problem [132, 133]. For a non-consensus set $\mathcal{M}$, it is always not obvious how to find a deterministic sequence that converges, especially when $\mathcal{M}$ has a large number of elements and F_i has zero diagonal entries. However, the convergence can be ensured almost surely by introducing some randomness in the sequence, provided that there is a convergent deterministic sequence intrinsically.*

4.2.2 Estimate of Convergence Rate

In Subsection 4.2.1, we have shown how the product $W(k,0)$ determined by a random process asymptotically converges to a rank-one matrix W a.s. as $k \to \infty$. However, the convergence rate for such a randomized product is not yet clear. It is quite challenging to investigate how fast the process converges, especially when each $W(k)$ may have zero diagonal entries. In this subsection, we address this problem by employing finite-step stochastic Lyapunov functions. Now let us present the main result on convergence rate.

Theorem 4.3. *In addition to Assumption 4.1, if there exists a number p, $0 < p < 1$, such that for any $k \in \mathbb{N}_0$*

$$\Pr\left[W(h,k) \in \mathcal{M}_2\right] \geq p > 0,$$

then the almost sure convergence of the product $W(k,0)$ to a random matrix $L = \mathbf{1}\xi^\top$ is exponential, and the rate is no slower than $\left(1 - p\alpha^h\right)^{1/h}$.

Proof. Choosing $V(x_k) = v_k$ as a finite-step stochastic Lyapunov function candidate, from (4.10) we have

$$\mathbb{E}\left[V(x_{k+h})|\,\mathcal{F}_k\right] - V(x_k) \leq -\sum\nolimits_{i=1}^{n_k^s} \Pr\left[S_k^i\right] \alpha^h V(x_k). \tag{4.13}$$

Furthermore, it is easy to see that

$$\sum\nolimits_{i=1}^{n_k^s} \Pr\left[S_k^i\right] = \Pr\left[W(h,t) \in \mathcal{M}_2\right] \geq p,$$

Substituting it into (4.13) yields

$$\mathbb{E}\left[V(x_{k+h})|\,\mathcal{F}_k\right] \leq \left(1 - p\alpha^h\right) V(x_k).$$

It follows from Corollary 3.3 that $V(x_{k+h}) \xrightarrow{a.s.} 0$, with an convergence rate no slower than $\left(1 - p\alpha^h\right)^{1/h}$. In other words, the agreement is reached exponentially almost surely, which, in turn, completes the proof. $\square$

Theorem 4.3 has established the almost sure exponential convergence rate for the product of $\{W(k)\}$. If any subsequence $\{W(k+1), \ldots, W(k+2), W(k+h)\}$ can result in a scrambling product $W(k+h,k)$ with positive probability and this probability is lower bounded away by some positive number, and then the convergence rate is exponential. *Interestingly, the greater this lower bound is, the faster the convergence becomes.* If we consider a special random sequence which is driven by a stationary ergodic process, the exponential convergence rate follows without any other conditions apart from Assumption 4.2, and an alternative proof is given in Appendix 4.6.

Corollary 4.2. *Suppose the random process governing the evolution of the sequence* $\{W(k) : k \in \mathbb{N}\}$ *is* stationary ergodic*, then the product* $W(k,0)$ *converges to a random rank-one matrix at an exponential rate almost surely under Assumption 4.2.*

4.2.3 Connections to Markov Chains

In this subsection, we show that Theorems 4.2, and 4.3 are the generalizations of some well known results for Markov chains in [68, 71]. A fundamental result on inhomogeneous Markov chains is as follows.

Lemma 4.1 ([71, Th. 4.10], [68])**.** *If the product* $W(k,t)$*, formed from a sequence* $\{W(k)\}$*, satisfies* $W(t+k,t) \in \mathcal{M}_1$ *for any* $k \geq 1, t \geq 0$*, and* $W_{ij}(k) \geq \alpha$ *whenever* $W_{ij}(k) > 0$*, then* $W(k,0)$ *converges to a rank-one matrix as* $k \to \infty$*.*

Let h be the number of distinct types of scrambling matrices of order n. It is known that the product $W(k+h,k)$ is scrambling for any k. In this case, we may take the probability of each product $W(k+h,k)$ being scrambling as $p=1$, and as an immediate consequence of Theorem 4.3, we know that $W(k,0)$ converges to a rank-one matrix at a exponential rate that is no slower than $(1-\alpha^h)^{1/h}$. This convergence rate is consistent with what is estimated in [71, Th. 4.10]. This also applies to the homogeneous case where $W(k) = W$ for any k with W being scrambling. Moreover, it is known that the condition can be relaxed by just requiring W to be SIA to ensure the convergence, which is an immediate consequence of Theorem 4.2.

In next section, we discuss how the results in this section can be further applied to the context of asynchronous computations.

4.3 Agreement Induced by Stochastic Asynchronous Events

In this section, we study the agreement problem of multi-agent systems in networks that are allowed to be periodic (which will be defined later in this section). Periodic networks often lead to oscillating behaviors, but we show that asynchronous updating can induce agreement even the network is periodic. The results on products of random sequences of stochastic matrices obtained in Section 4.2 will be used to construct the proofs.

We take each component x^j in x from (4.7) as the state of agent i in an n-agent system. Define the distributed coordination algorithm

$$x^i(t_{k+1}) = \sum_{j=1}^{n} w_{ij} x^j(t_k), \qquad k \in \mathbb{N}_0, i \in \mathbf{N}, \tag{4.14}$$

where the averaging weights $w_{ij} \geq 0$, $\sum_{j=1}^{n} w_{ij} = 1$, and t_k denote the time instants when updating actions happen. Here we assume the initial state $x(t_0)$ is given. It is always assumed that $T_1 \leq t_{k+1} - t_k \leq T_2$, where $t_0 = 0$ and T_1, T_2 are positive numbers. We say the states of system (4.14) reach *agreement* if $\lim_{k\to\infty} x(t_k) = \mathbf{1}\zeta$, mentioned in Section 4.2. Let $W = [w_{ij}] \in \mathbb{R}^{n\times n}$, and obviously W is a stochastic matrix. The algorithm (4.14) can be rewritten as

$$x(t_{k+1}) = Wx(t_k). \tag{4.15}$$

In fact, the matrix W can be associated with a directed, weighted graph $\mathcal{G}_W = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, 2, \cdots, n\}$ is the vertex set and $\mathcal{E}$ is the edge set for which $(i, j) \in \mathcal{E}$ if $w_{ji} > 0$. The graph $\mathcal{G}_W$ is called a *rooted* one if there exists at least one vertex, called a *root*, from which any other vertex can be reached. It is known that agents are able to reach agreement for all $x(0)$ if W is SIA ([68, 71]). However, the situations when W is not SIA have not been studied before, although they appear often in real systems, such as social networks.

In the context of distributed computation, it is always assumed that each computational unit in the network has access to its own latest state while implementing the iterative update rules [10, 25]. A class of situations that has received considerably less attention in the literature arise when some individuals are not able to obtain their own states, a case which can result from memory loss. Similar phenomena have also been observed in social networks while studying the evolution of opinions. Self-contemptuous people change their opinions solely in response to the opinions of others. The existence of computational units or individuals who are not able to access their own states sometimes might result in the computational failure or opinions' disagreement. As such an example, a periodic matrix W, which must has all zero diagonal entries (no access to their own states for all individuals), always leads the system (4.14) to oscillation. This is because for a periodic W, W^k never converges to a matrix with identical rows as $k \to \infty$. Instead, the positions of W^k that have positive values are periodically changing with k, resulting in a periodically changing value of $W^k x(0)$. We illustrate this point by the following example.

Example 4.1. *For system (4.15), the initial state is given by $x(0) = [1, 2, 3, 4]^T$, and the matrix P is*

$$W = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

By simple computation, one can check that $x(t_1) = [4, 1, 2, 3]^T, x(t_2) = [3, 4, 1, 2]^T, x(t_3) = [2, 3, 4, 1]^T, x(t_4) = [1, 2, 3, 4]^T = x(0)$. It is easy to see that the state equals the initial state after updating for four times. Then the same process will repeat again, which obviously implies a oscillating behavior instead of agreement. △

This motivates us to investigate the particular case where W is periodic. In the following two definitions, we provide the formal definitions of periodic stochastic matrices. We first introduce the definition of periodic irreducible matrices found in [71, Def. 1.6], and then extend this definition to the case when the matrices do not have to be irreducible.

Definition 4.1 ([71, Def. 1.6])**.** *Consider an irreducible stochastic matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$. An index $i \in \{1, 2, \cdots, n\}$ is said to have period $d(i)$ if $d(i)$ is the common divisor of those $m \in \mathbb{N}^+$ for which $a_{ii}^{(m)} > 0$. The matrix A is said to be periodic with period d if $d(i) = d > 1$ for all i.*

Definition 4.2. *Consider a stochastic matrix $A \in \mathbb{R}^{n \times n}$, and let $\mathcal{P} := \{i : \exists m \in \mathbb{N}^+ : a_{ii}^{(m)} > 0\}$. An index $i \in \mathcal{P}$ is said to have period $d(i)$ if $d(i)$ is the common divisor of those m for which $a_{ii}^{(m)} > 0$. The matrix A is said to be periodic if $d(i) > 1$ for any $i \in \mathcal{P}$, and the period d is the common divisor of those m such that $a_{ii}^{(m)} > 0$ for all $i \in \mathcal{P}$.*

Definition 4.2 is a generalization of Definition 4.1. In this definition, a periodic stochastic matrix is not necessarily irreducible. The following example provides some intuition on these two definitions.

Example 4.2. *Consider the following two matrices:*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

One can see that A is irreducible, and B, C are reducible. According to Definition 4.1, it can be calculated that the indices $1, 2$ and 3 of A all have period 3, which means A is periodic with period 3. According to Definition 4.2, $\mathcal{P} = 1, 2$ for B, and the indices $1, 2$ have period 2. Then it is clear that the period of B is 2. Likewise, one can check that the period of C is 6. △

With a slight abuse of terminology, we say the graph $\mathcal{G}_W$ is *periodic* if the associated matrix W is. In this section, we show that agreement can be reached even when W is periodic, *just by introducing asynchronous updating events to the coupled agents.* In fact, perfect synchrony is hard to realize in practice as it is difficult for all agents to have access to a common clock according to which they coordinate their updating actions, while asynchrony is more likely. Researchers have studied how an agreement can be preserved with the existence of asynchrony, see e.g., [13, 14]. Unlike these works, we approach the same problem from a different aspect, where agreement occurs just because of asynchrony.

To proceed, we define a framework of randomly asynchronous updating events. It is usually legitimate to postulate that on occasions more than one, but not all, agents may update. Assume that each agent is equipped with a clock, which need not be synchronized with other clocks. The state of each agent remains unchanged except when an activation event is triggered by its own clock. Denote the set of event times of the ith agent by $\mathcal{T}^i = \{0, t_1^i, \cdots, t_k^i, \cdots\}, k \in \mathbb{N}$. At the event times, agent i updates its state obeying the asynchronous updating rule

$$x_i\left(t_{k+1}^i\right) = \sum\nolimits_{j=1}^{n} w_{ij} x_j\left(t_k^i\right), \tag{4.16}$$

where $i \in \mathbf{N}$. We assume that the clocks which determine the updating events for the agents are driven by an underlying random process. The following assumption is important for the analysis.

Assumption 4.3. *For any agent i, the intervals between two event times, denoted by $h_k^i = t_k^i - t_{k-1}^i$, are such that*

(i) h_k^i are upper bounded with probability 1 for all k and all i;

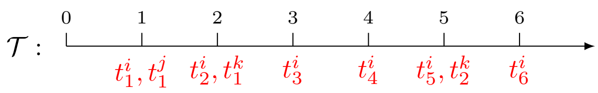


Figure 4.1: Event times of all agents: one (or more) agents can be activated simultaneously.

(ii) $\{h_k^i : k \in \mathbb{N}_0\}$ *is a random sequence, with* $\{h_k^1\}$, $\{h_k^2\}$, …, $\{h_k^n\}$ *being mutually independent.*

Assumption 4.3 ensures that an agent can be activated again within finite time after it is activated at t_{k-1}^i for all $k \in \mathbb{N}$, which implies that all agents will update their states for infinitely many times in the long run. In fact, Assumption 4.3 can be satisfied if the agents are activated by mutually independent Poisson clocks or at rates determined by mutually independent Bernoulli processes ([134, Ch. 6], [124, Ch. 2]).

Let $\mathcal{T} = \{t_0, t_1, t_2, \cdots, t_k, \cdots\}$ denote all event times of all the n agents, in which the event times have been relabeled in a way such that $t_0 = 0$ and $t_\tau < t_{\tau+1}, \tau = \{0, 1, 2, \cdots\}$. This idea has been used in [135] and [10] to study asynchronous iterative algorithms. One situation may occur in which there exists some k such that $t_k \in \mathcal{T}^i$ and $t_k \in \mathcal{T}^j$ for some i, j, which implies more than one agent is activated at some event times. Although this is not likely to happen when the underlying process is some special random ones like Poisson, our analysis and results will not be affected. The arrangement of $\mathcal{T}$ is illustrated clearly by Figure 4.1. For simplicity, we rewrite the set of event times as $\mathcal{T} = \{0, 1, 2, \cdots, k, \cdots\}$. Then the system with asynchronous updating can be treated as one with discrete-time dynamics in which the agents are permitted to update only at certain event times $k, k \in \mathbb{N}$, according to the updating rule (4.16) at each time k. Since each $k \in \mathcal{T}$ can be the event time of any subset of agents, we can associate any set of event times $\{k+1, k+2, \ldots, k+h\}$ with the updating sequence of agents $\{\lambda(k+1), \lambda(k+2), \ldots, \lambda(k+h)\}$ with $\lambda(i) \in \mathcal{V}$. Under Assumption 4.3, one knows that this updating sequence can be *arbitrarily ordered*, and each possible sequence can occur with positive probability, though the particular value is not of concern.

Assume at time k, $m \geq 1$ agents are activated, labeled by $k_1, k_2, \ldots, k_m$, then we define the following matrices

$$W(k) = \left[u_1, \cdots, w_{k_1}^\top, u_{k+1}, \cdots, w_{k_m}^\top, \cdots, u_n\right]^\top, \tag{4.17}$$

where $u_i \in \mathbb{R}^n$ is the ith column of the identity matrix I_n and $w_k \in \mathbb{R}^n$ denotes the kth row of W. We call $W(k)$ the *asynchronous updating matrix* at time k. Then the asynchronous updating rule (4.16) becomes

$$x_k = W(k)x_{k-1}, \qquad k \in \mathbb{N}, \tag{4.18}$$

where $\{W(k)\}$ is a random sequence of asynchronous updating matrices which are stochastic, and $x_0 \in \mathbb{R}^n$ is a given initial state. We say the *asynchronous agreement* is reached if x_k converges to a scaled all-one vector when the agents update asynchronously. It suffices to study the convergence of the product $W(k) \dots W(2)W(1)$ to a rank-one matrix.

In Subsection 4.3.1, we consider the agents are coupled by a strongly connected and periodic network, and show that agreement is reached almost surely if the agents update their states asynchronously under Assumption 4.3. In Subsection 4.3.2, we identify a necessary and sufficient condition on the graph structure for asynchronous agreement, where aperiodicity is not required anymore.

4.3.1 Asynchronous Agreement over Strongly Connected Periodic Networks

In this subsection, we assume that the agents are coupled by a strongly connected and periodic network $\mathcal{G}_W$. Equivalently, the associated stochastic matrix W in the system (4.15) is irreducible and periodic (see Definition 4.1). We show in the following theorem that agreement can be reached if the agents update their states asynchronously.

Theorem 4.4. *Suppose that the agents are coupled by a strongly connected and periodic graph $\mathcal{G}_W$. Then, they can reach agreement almost surely if they update asynchronously under Assumption 4.3.*

We use the results in Corollary 4.1 to construct the proof. Then, it suffices to prove that there is a class of updating sequence of finite length such that the product of the corresponding asynchronous updating matrices, i.e., $W(k)$ in (4.18), is a Markov matrix, and this class of updating sequence appears with positive probability. This is formally stated in the following proposition.

Proposition 4.3. *There exists $T \in \mathbb{N}$ such that the product of the asynchronous updating matrices $W(k+T)W(k+T-1)\cdots W(k+1)$ have a positive probability to be a Markov matrix for any $k \in \mathbb{N}_0$.*

To prove this proposition, we define an operator $\mathcal{N}(\cdot,\cdot)$ for any stochastic matrix and any subset $\mathcal{S} \in \mathcal{V}$

$$\mathcal{N}(A, \mathcal{S}) := \{j : A_{ij} > 0, i \in \mathcal{S}\}, \tag{4.19}$$

and we write $\mathcal{N}(A,\{i\})$ as $\mathcal{N}(A,i)$ for brevity. It is easy to check then for any two stochastic matrices $A_1, A_2 \in \mathbb{R}^{n\times n}$ and for any subset $\mathcal{S} \in \mathcal{V}$, it holds that

$$\mathcal{N}\left(A_2A_1,\mathcal{S}\right)=\mathcal{N}\left(A_1,\mathcal{N}\left(A_2,\mathcal{S}\right)\right). \tag{4.20}$$

Proof of Proposition 4.3. This proposition can be proved by considering a special class of updating sequences, which appears with probability greater than 0. Since the directed graph $\mathcal{G}_W = (\mathcal{V},\mathcal{E})$ considered in this chapter is strongly connected (W is irreducible), for any fixed node $\lambda(1) \in \mathcal{V}$ one can always find some directed paths starting from $\lambda(1)$ and passing through all other nodes with finite lengths. Choose the path with the minimal length $T-1$, denoted by

$$\lambda\left(1\right)\to\lambda\left(2\right)\to\cdots\lambda\left(T-1\right)\to\lambda\left(T\right).$$

Obviously, it satisfies $\bigcup_{i=1}^{T}\lambda\left(i\right)=\mathcal{V}$. Now we assume that the updating sequence of the agents is $\{\lambda\left(1\right),\lambda\left(2\right),\cdots,\lambda(T)\}$, where only one agent updates at the corresponding time. Let $\{W_{\lambda(1)},W_{\lambda(2)},\cdots,W_{\lambda(T)}\}$ denote the sequence of the updating matrices. Let Φ be the backward product of this sequence, and it is given by

$$\Phi=W_{\lambda(T)}W_{\lambda(T-1)}\cdots W_{\lambda(2)}W_{\lambda(1)} \tag{4.21}$$

We next show Φ in (4.21) has at least one positive column. One knows Φ has a positive column if only if all the nodes in the associated graph $\mathcal{G}_\Phi$ share a common neighbor. Then we will prove all the nodes in $\mathcal{G}_\Phi$ share a common neighbor, i.e.,

$$\bigcap_{i=1}^{n}\mathcal{N}\left(\Phi,i\right)\neq\emptyset. \tag{4.22}$$

We first define the following iteration

$$\begin{aligned}&s_m=\{\lambda\left(k_{m-1}\right)\}\cup s_{m-1},\\&k_m=\max\left\{k:\lambda\left(k\right)\notin s_m,1\leq k\leq T\right\}\end{aligned}$$

where $m=2,\cdots,n$. Let $s_1=\emptyset, k_1=T$. Since $\bigcup_{i=1}^{T}\lambda\left(i\right)=\mathcal{V}$, it holds that $\bigcup_{i=1}^{T}\lambda\left(k_i\right)=\mathcal{V}$. For any k_i, it is obvious to see

$$\begin{aligned}&\mathcal{N}\left(W_{\lambda(T)}\cdots W_{\lambda(k_i+1)}W_{\lambda(k_i)}\cdots W_{\lambda(2)}W_{\lambda(1)},\lambda\left(k_i\right)\right)\\&=\mathcal{N}\left(W_{\lambda(k_i)}\cdots W_{\lambda(2)}W_{\lambda(1)},\lambda\left(k_i\right)\right).\end{aligned}$$

As $\lambda(k_i-1)$ is one of the neighbors of $\lambda(k_i)$, i.e.,

$$\lambda\left(k_i-1\right)\in\mathcal{N}\left(W_{\lambda(k_i)},\lambda\left(k_i\right)\right),$$

it follows that

$$\begin{aligned}&\mathcal{N}\left(W_{\lambda(k_i)}W_{\lambda(k_i-1)}\cdots W_{\lambda(2)}W_{\lambda(1)},\lambda\left(k_i\right)\right)\\&\supseteq\mathcal{N}\left(W_{\lambda(k_i-1)}\cdots W_{\lambda(2)}W_{\lambda(1)},\lambda\left(k_i-1\right)\right)\end{aligned} \tag{4.23}$$

where the inequality (4.20) has been used. Also, $\lambda(k_i-2)$ is an neighbor of $\lambda(k_i-1)$, then

$$\begin{aligned}&\mathcal{N}\left(W_{\lambda(k_i-1)}W_{\lambda(k_i-2)}\cdots W_{\lambda(2)}W_{\lambda(1)},\lambda\left(k_i\right)\right)\\&\supseteq\mathcal{N}\left(W_{\lambda(k_i-2)}\cdots W_{\lambda(2)}W_{\lambda(1)},\lambda\left(k_i-2\right)\right)\end{aligned}\tag{4.24}$$

By recurrence one can conclude that

$$\begin{aligned}&\mathcal{N}\left(W_{\lambda(k_i-m)}\cdots W_{\lambda(2)}W_{\lambda(1)},\lambda\left(k_i-m\right)\right)\\&\supseteq\mathcal{N}\left(W_{\lambda(k_i-m-1)}\cdots W_{\lambda(2)}W_{\lambda(1)},\lambda\left(k_i-m-1\right)\right),\end{aligned}$$

where $0\leq m\leq k_i-2$. It is straightforward to see

$$\mathcal{N}\left(\Phi,\lambda(k_i)\right)\supseteq\mathcal{N}\left(W_{\lambda(1)},\lambda\left(1\right)\right)=\mathcal{N}\left(W,\lambda\left(1\right)\right)\tag{4.25}$$

It is worth mentioning that (4.25) holds for any $i=1,2,\cdots,n$, which implies (4.22). Till here we know that all the nodes in the associated graph $\mathcal{G}_\Phi$ have at least one common neighbor which is the neighbor of $\lambda(1)$ in $\mathcal{G}_W$. It is easy to see that Φ has at least one positive column, which implies that it is a Markov matrix.

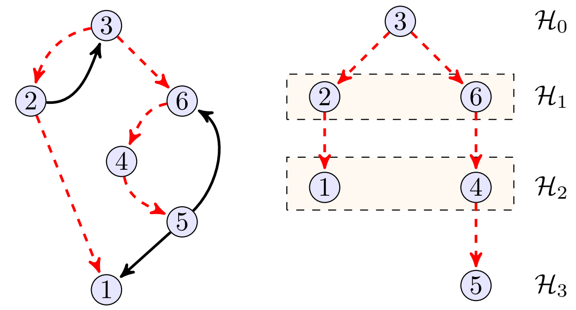
The updating sequence $\{\lambda\left(1\right),\lambda\left(2\right),\cdots,\lambda(T)\}$ can appear with positive probability at every interval of T time steps. This means that the product of the asynchronous updating matrices $W(k+T)W(k+T-1)\cdots W(k+1)$ have a positive probability to be a Markov matrix for any k, which completes the proof. □

4.3.2 A Necessary and Sufficient Condition for Asynchronous Agreement

In the previous subsection, we prove that the agents coupled by a strongly connected and periodic graph can reach an agreement if the agents update asynchronously. It is surprising since it has been believed that agreement through weighted averaging algorithms like (4.16) requires the graph to be aperiodic. In this subsection, we generalize the result in the previous subsection, and obtain a necessary and sufficient condition on the graph structure of $\mathcal{G}_W$ such that asynchronous agreement is ensured. The main result is presented in the following theorem.

Theorem 4.5. *Suppose the agents coupled by a network update asynchronously under Assumption 4.3, then they reach agreement almost surely if and only if the network is rooted, i.e., the matrix W is indecomposable.*

To prove this theorem, we need to introduce some additional concepts and results. It is equivalent to say the associated graph $\mathcal{G}_W$ is rooted if W is indecomposable. Denote the set of all the roots of $\mathcal{G}_W$ by $\mathbf{r}\subseteq\mathcal{V}$. We can partition the vertices of $\mathcal{G}_W$ into some hierarchical subsets as follows. For any $\kappa\in\mathbf{r}$, there must exist at least one directed spanning tree rooted at κ, see e.g., Fig. 4.2 (a). We select any of these



(a) The original graph. (b) Partition of the vertices.

Figure 4.2: An illustration of the graph partition; the hierarchical subsets: $\mathcal{H}_0 = \{3\}$,$\mathcal{H}_1 = \{2, 6\}$,$\mathcal{H}_2 = \{1, 4\}$,$\mathcal{H}_3 = \{5\}$; for example, {3,2,6,1,4,5} is a hierarchical updating vertex sequence.

directed spanning trees, denoted by $\mathcal{G}_W^s$. There exists a directed path from κ to any other vertex $i \in \mathcal{V}\backslash\kappa$, see e.g., Fig. 4.2 (b). Let l_i be the length of the directed path from κ to i, and there exists an integer $L \leq n$ such that $l_i < L$ for all i. Define

$$\mathcal{H}_r := \{i : l_i = r\}, r = 1, \cdots, L-1,$$

and $\mathcal{H}_0 = \{\kappa\}$. From this definition, one can partition the vertices of $\mathcal{G}_W^s$ into L hierarchical subsets, i.e., $\mathcal{H}_0, \mathcal{H}_1, \cdots, \mathcal{H}_{L-1}$, according to the vertices' distances to the root κ. Let n_r be the number of vertices in the subset $\mathcal{H}_r$, $0 \leq r \leq L-1$ (see the example in Fig. 4.2 (b)). Note that given a spanning tree, its corresponding hierarchical subsets H_r's are uniquely determined.

Definition 4.3. *An updating vertex sequence of length n is said to be hierarchical if it can be partitioned into some successive subsequences, denoted by $\{\mathcal{A}_0, \ldots, \mathcal{A}_{L-1}\}$ with $\mathcal{A}_r = \{\lambda_r(1), \lambda_r(2), \cdots, \lambda_r(n_r)\}$, such that $\bigcup_{k=1}^{n_r} \lambda_r(k) = \mathcal{H}_r$ for all $r = 0, \cdots, L-1$, where $\mathcal{H}_r$'s are the hierarchical subsets of some spanning tree $\mathcal{G}_W^s$ in $\mathcal{G}_W$.*

Proposition 4.4. *If agents coupled by $\mathcal{G}_W$ update in a hierarchical sequence $\{a_1, \cdots, a_n\}, a_i \in \mathcal{V}$ for all i, the product of the corresponding asynchronous updating matrices,*

$$\Phi := W_{a_n} \cdots W_{a_2} W_{a_1}$$

is a Markov matrix.

Proof of Proposition 4.4. It suffices to show that all $i \in \mathcal{V}$ share at least one common neighbor in the graph $\mathcal{G}_\Phi$, i.e.,

$$\bigcap_{i=1}^{n} \mathcal{N}(\Phi, i) \neq \emptyset. \tag{4.26}$$

We rewrite the product of asynchronous updating matrices into

$$\Phi = \left\{ W_{\lambda_{L-1}(1)} \cdots W_{\lambda_{L-1}(n_{L-1})} \cdots W_{\lambda_{L-2}(1)} \cdots W_{\lambda_0(1)} \right\}.$$

For any distinct $i, j \in \mathcal{V}$, we know that $\mathcal{N}(W_j, i) = \{i\}$ from the definition of asynchronous updating matrices. Then for any $\lambda_r(t) \in \mathcal{H}_r, t \in \{1, \cdots, n_r\}, r \in \{1, \cdots, L-1\}$, it holds that

$$\begin{aligned}
&\mathcal{N}\left(\Phi, \lambda_r(t)\right) \\
&= \mathcal{N}\left(W_{\lambda_r(t)} W_{\lambda_r(t+1)} \cdots W_{\lambda_r(n_r)} \cdots W_{\lambda_0(1)}, \lambda_r\left(t\right)\right) \\
&= \mathcal{N}\left(W_{\lambda_r(t+1)} \cdots W_{\lambda_r(n_r)} \cdots W_{\lambda_0(1)}, \mathcal{N}\left(W_{\lambda_r(t)}, \lambda_r\left(t\right)\right)\right),
\end{aligned}$$

where the property (4.20) has been used. From Definition 4.3, one knows that there exists at least one vertex $\lambda_{r-1}\left(t_1\right) \in \mathcal{H}_{r-1}$ that can reach $\lambda_r\left(t\right)$ in $\mathcal{G}_W$ and subsequently in $\mathcal{G}_{W_{\lambda_r(t)}}$, which implies

$$\lambda_{r-1}\left(t_1\right) \in \mathcal{N}\left(W_{\lambda_r(t)}, \lambda_r\left(t\right)\right).$$

It then follows

$$\begin{aligned}
&\mathcal{N}\left(W_{\lambda_r(t+1)} \cdots W_{\lambda_r(n_r)} \cdots W_{\lambda_0(1)}, \lambda_{r-1}\left(t_1\right)\right) \\
&\subseteq \mathcal{N}\left(\Phi, \lambda_r(t)\right).
\end{aligned}$$

Similarly, one obtains

$$\begin{aligned}
&\mathcal{N}\left(W_{\lambda_r(t+1)} \cdots W_{\lambda_r(n_r)} \cdots W_{\lambda_0(1)}, \lambda_{r-1}\left(t_1\right)\right) \\
&= \mathcal{N}\left(W_{\lambda_{r-1}(t_1)} \cdots W_{\lambda_r(n_r)} \cdots W_{\lambda_0(1)}, \lambda_{r-1}\left(t_1\right)\right) \\
&= \mathcal{N}\left(W_{\lambda_{r-1}(t_1+1)} \cdots W_{\lambda_0(1)}, \mathcal{N}\left(W_{\lambda_{r-1}(t_1)}, \lambda_{r-1}\left(t_1\right)\right)\right) \\
&\supseteq \mathcal{N}\left(W_{\lambda_{r-1}(t_1+1)} \cdots W_{\lambda_0(1)}, \lambda_{r-2}\left(t_2\right)\right).
\end{aligned}$$

As a recursion, it must be true that

$$\mathcal{N}\left(W_{\lambda_0(1)}, \kappa\right) \subseteq \mathcal{N}\left(\Phi, \lambda_r(t)\right), \tag{4.27}$$

where κ is a root of $\mathcal{G}_W^s$. In fact, it holds that $\lambda_0(1) = \kappa$, and then we know

$$\mathcal{N}\left(W_{\lambda_0(1)}, \kappa\right) = \mathcal{N}\left(W_\kappa, \kappa\right) = \mathcal{N}\left(W, \kappa\right). \tag{4.28}$$

Substituting (4.28) into (4.27) leads to

$$\mathcal{N}\left(W, \kappa\right) \subseteq \mathcal{N}\left(\Phi, \lambda_r(t)\right)$$

for all $\lambda_r(t)$. Since $\bigcup_{r,t} \left\{\lambda_r\left(t\right)\right\} = \mathcal{V}$, we know

$$\mathcal{N}\left(W, \kappa\right) \subseteq \bigcap\nolimits_{r,t} \mathcal{N}\left(\Phi, \lambda_r(t)\right).$$

Straightforwardly, (4.26) follows, which completes the proof. □

Since the hierarchical sequences will appear with positive probability in any sequence of length n, one can easily prove the following proposition by letting $l = n$.

Proposition 4.5. *There exists an integer l such that the product $W(k+l)\cdots W(k+1)$, where $W(k)$ is given in* (4.18)*, is a Markov matrix with positive probability for any $k \in \mathbb{N}$.*

Proof of Theorem 4.5. We prove the necessity by contradiction. Suppose the matrix W is decomposable. Then there are at least two sets of vertices that are isolated from each other. Then agreement will never happen between these two isolated groups if they have different initial states. Let $l = n$, in view of Corollary 4.1, the sufficiency follows directly from Proposition 4.5, which completes the proof. □

Note that the hierarchical sequence is a particular type of updating orders that results in a Markov matrix as the product of the corresponding updating matrices. We have identified another type of updating orders in our earlier work when W is irreducible and periodic in the previous subsection. It is of great interest for future work to look for other updating mechanisms to enable the appearance of Markov matrices or scrambling matrices, which plays a crucial role in giving rise to an asynchronous agreement.

In the next subsection, we demonstrate the obtained results in the two subsections by simulation.

4.3.3 Numerical Examples

In this section, we demonstrate the obtained results by a numerical example. Consider the system (4.15) with the following periodic matrix

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The corresponding graph is given by Fig. 4.3, which is strongly connected and periodic. Let the initial state be $x(0) = [1.1, 4.2, 7.3, 3.4, 4.5, 5.6]^T$. If the agents in the network have a common clock to synchronize the updating actions, the states of the agents cannot reach an agreement, instead, a oscillating behavior takes place, as shown in Fig. 4.4.

However, if individuals update according to their own clocks under Assumption 4.3, the agreement can be reached. To illustrate this, we assume the clocks are driven

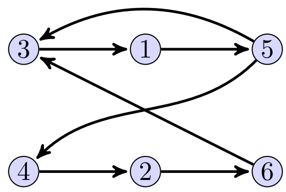


Figure 4.3: Associated graph of P.

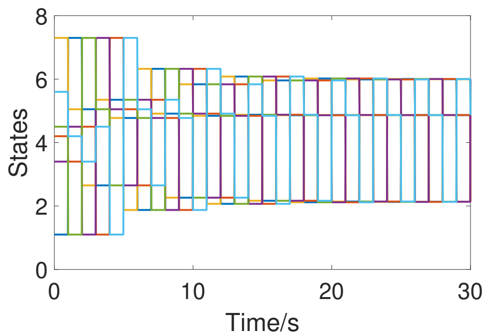


Figure 4.4: Update synchronously: oscillation.

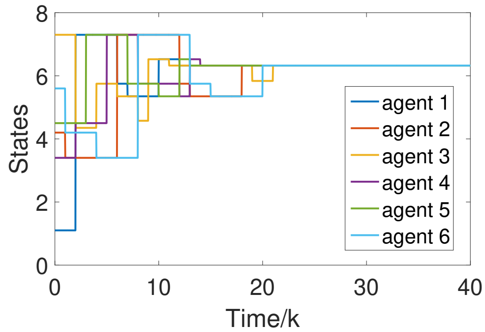


Figure 4.5: Update asynchronously: agreement.

by mutually independent Poisson processes in which the interarrival intervals have the density functions

$$f_i(x) = \lambda_i e^{-\lambda_i x}, \qquad\qquad \text{for } x \geq 0,$$

where $i = 1, 2, \cdots, n$. Let $\lambda_i = 2$ for all i. The evolution of the agents' states is shown in Fig. 4.3.3, which shows that the states converge to a common value instead of an oscillation although the network is periodic. Thus one observes that asynchronous updating events have played a fundamental role in giving rise to agreement.

4.4 A Linear Algebraic Equation Solving Algorithm

In this section, we apply the finite step Lyapunov criteria obtained in Chapter 3 to solving linear algebraic equations distributively.

Researchers have been quite interested in solving a system of linear algebraic equations in the form of $Ax = b$ in a distributed way [29, 30, 113, 114]. In this section we deal with the problem under the assumption that this system of equations has at least one solution. The set of equations is decomposed into smaller sets and distributed to a network of n processors, referred to as agents, to be solved in parallel. Agents can receive information from their neighbors and the neighbor relationships are described by a time-varying n-vertex directed graph $\mathcal{G}(t)$ with self-arcs. When each agent knows only the pair of real-valued matrices $(A_i^{n_i \times m}, b_i^{n_i \times 1})$, the problem of interest is to devise local algorithms such that all n agents can iteratively compute the same solution to the linear equation $Ax = b$, where $A = [A_1^\top, A_2^\top, \ldots, A_n^\top]^\top, b = [b_1^\top, b_2^\top, \ldots, b_n^\top]^\top$ and $\sum_{i=1}^n n_i = m$.

A distributed algorithm to solve the problem is introduced in [77], where the iterative updating rule for each agent i is described by

$$x_{k+1}^i = x_k^i - \frac{1}{d_k^i} P_i \Big(d_k^i x_k^i - \sum_{j \in \mathcal{N}_i(k)} x_k^j \Big), k \in \mathbb{N}, \tag{4.29}$$

where $x_k^i \in \mathbb{R}^m$, d_k^i is the number of neighbors of agent i at time k, $\mathcal{N}_i(k)$ is the collection of i's neighbors, P_i is the orthogonal projection on the kernel of A_i, and the initial value x_1^i is any solution to the equations of $A_i x = b_i$.

The results in [77] have shown that all x_k^i converge to the same solution exponentially fast if the sequence of graphs $\mathcal{G}(t)$ is repeatedly jointly strongly connected. This condition requires that for some integer l, the composition of the sequence of graphs, $\{\mathcal{G}(k), \ldots, \mathcal{G}(k+l-1)\}$, must be strongly connected for any t. It is not so easy to satisfy this condition if the network is changing randomly. Now assume that the evolution of the sequence of graphs $\{\mathcal{G}(1), \ldots, \mathcal{G}(k), \ldots\}$ is driven by a random process. In this case, results in Theorem 3.1 and Corollary 3.1 can be applied to relaxing the condition in [77] to achieve the following more general result.

Theorem 4.6. *Suppose that each agent updates its state x_k^i according to the rule (4.29). All states x_k^i converge to the same solution to $Ax = b$ almost surely if the following two conditions are satisfied:*

a) there exists an integer l such that for any $k \in \mathbb{N}$ the composition of the sequence of randomly changing graphs $\{\mathcal{G}(k), \mathcal{G}(k+1), \ldots, \mathcal{G}(k+l-1)\}$ is strongly connected with positive probability $p(k) > 0$;

b) for any $k \in \mathbb{N}$, it holds that $\sum_{i=0}^{\infty} p(k+il) = \infty$.

To prove the theorem, we define an error system. Let x^* be any solution to $Ax = b$, so $A_i x^* = b_i$ for any i. Then, we define

$$e_k^i = x_k^i - x^*, i \in \mathcal{V}, k \in \mathbb{N},$$

which, as is done in [77], can be simplified into

$$e_{k+1}^i = \frac{1}{d_k^i} P_i \sum\nolimits_{j \in \mathcal{N}_i(k)} P_j e_k^j. \tag{4.30}$$

Let $e_k = [{e_k^1}^\top, \ldots, {e_k^n}^\top]^\top$, $A(k)$ be the adjacency matrix of the graph $\mathcal{G}(k)$, $D(k)$ be the diagonal matrix whose ith diagonal entry is d_k^i, and $W(k) = D^{-1}(k)A^\top(k)$. It is clear that $W(k)$ is a stochastic matrix, and $\{W(k)\}$ is a stochastic process. Now we write equation (4.30) into a compact form

$$e_{k+1} = P(W(k) \otimes I)Pe_k, k \in \mathbb{N}, \tag{4.31}$$

where $\otimes$ denotes the Kronecker product, $P := \text{diag}\{P_1, P_2, \ldots, P_n\}$, and $\{W(k)\}$ is a random process. We will show this error system is globally a.s. asymptotically stable. Define the transition matrix of this error system by

$$\Phi(k+T, k) = P(W(k+T-1) \otimes I)P \cdots P(W(k) \otimes I)P.$$

In order to study the stability of the error system (4.31), we define a mixed-matrix norm for an $n \times n$ block matrix $Q = [Q_{ij}]$ whose ijth entry is a matrix $Q_{ij} \in \mathbb{R}^{m \times m}$, and

$$[\![Q]\!] = |\langle Q \rangle|_\infty,$$

where $\langle Q \rangle$ is the matrix in $\mathbb{R}^{n \times n}$ whose ijth entry is $|Q_{ij}|_2$. Here $\| \cdot \|_2$ and $\| \cdot \|_\infty$ denote the induced 2 norm and infinity norm, respectively. It is easy to show that $[\![\cdot]\!]$ is a norm. Since $\|Ax\|_2 \leq \|A\|_2 \|x\|_2$ for $x \in \mathbb{R}^{nm \times nm}$, it follows straightforwardly that $[\![Ax]\!] \leq [\![A]\!] \, [\![x]\!]$. It has been proven in [77] that $\Phi(k+T, k)$ is non-expansive for any $k > 0, T \geq 0$. In other words, it holds that

$$[\![\Phi(k+T, k)]\!] \leq 1.$$

Moreover, the transition matrix is a contraction, i.e., $[\![\Phi(k+T,k)]\!] < 1$, if there exists a route $j = i_0, i_1, \ldots, i_T = i$ over the sequence $\{\mathcal{G}(k), \ldots, \mathcal{G}(k+T-1)\}$ for any $i, j \in \mathcal{V}$ that satisfies $\bigcup_{k=0}^{T} \{i_k\} = \mathcal{V}$. Now we are ready to prove Theorem 4.6.

Proof of Theorem 4.6. Let $V(e_k) = [\![e_k]\!]$ be a finite-step stochastic Lyapunov function candidate. Let $\{\mathcal{F}_k\}$, where $\mathcal{F}_k = \sigma(\mathcal{G}(1), \cdots, \mathcal{G}(k), \cdots)$, be an increasing sequence of σ-fields. We first show that $V(e_k)$ is a supermartingale with respect to $\mathcal{F}_k$ by observing

$$\mathbb{E}\big[V\big(e_{k+1}\big)\big|\,\mathcal{F}_k\big] = \mathbb{E}\,[\![\Phi_k e_k]\!] \leq \mathbb{E}\,[\![\Phi_k]\!]\ [\![e_k]\!] \leq [\![e_k]\!]\ ,$$

where $\Phi_k = \Phi(k,k) = P(W(k) \otimes I)Pe_k$. The last inequality follows from the fact that $\mathbb{E}\,[\![\Phi_k]\!] \leq 1$ since all the possible Φ_k are non-expansive. Consider the sequence of randomly changing graphs $\{\mathcal{G}(1), \mathcal{G}(2), \cdots, \mathcal{G}(q)\}$, where $q = (n-1)^2 l$. Let $r = n-1$, and partition this sequence into r successive subsequences $\mathcal{G}_1 = \{\mathcal{G}(1), \ldots, \mathcal{G}(rl)\}$, $\mathcal{G}_2 = \{\mathcal{G}(rl+1), \ldots, \mathcal{G}(2rl)\}, \cdots$, $\mathcal{G}_r = \{\mathcal{G}((r-1)l+1), \ldots, \mathcal{G}(r^2 l)\}$. Let $\mathbb{C}_z$ denote the composition of the graphs in the zth subsequence, i.e., $\mathbb{C}_z = \mathcal{G}(zl) \circ \cdots \circ \mathcal{G}((z-1)l+2) \circ \mathcal{G}((z-1)l+1)$, $z = 1, 2, \ldots, r$. Since all the subsequences have the length rl, each can be further partitioned into r successive sub-subsequences of length l. From the condition of Theorem 4.6, one knows that the composition of the graphs in any sub-subsequence has positive probability to be strongly connected. The event that the composition of the graphs in each of the r sub-subsequences in $\mathcal{G}_z$ is strongly connected also has positive probability. This holds for all z. We know that the composition of any r or more strongly connected graphs, within which each vertex has a self-arc, results in a complete graph [9]. It follows straightforwardly that the graphs $\mathbb{C}_1, \cdots, \mathbb{C}_r$ have positive probability to be all complete. Therefore, for any pair $i, j \in \mathcal{V}$, there exists a route from j to i over the graph $\mathbb{C}_z$ for any z. It is easy to check that there exists a route $i_1, i_2, \ldots, i_n$ over the graphs $\mathbb{C}_1, \cdots, \mathbb{C}_r$, where $i_1, i_2, \ldots, i_n$ can be any reordered sequence of $\{1, 2, \ldots, n\}$. Similarly, for any x there must exist a route of length rl, $i_z = i_z^1, i_z^2, \ldots, i_z^{rl} = i_{z+1}$, over $\mathcal{G}_z$. Thus there is a route $i_1^1, i_1^2, \ldots, i_1^{rl}, i_2^2, \ldots, i_2^{rl} \ldots, i_r^{rl}$ over the graph sequence $\{\mathcal{G}(1), \mathcal{G}(2), \cdots, \mathcal{G}(q)\}$ so that $\bigcup_{\delta=1}^{r} \bigcup_{\theta=1}^{rl} \{i_\delta^\theta\} = \mathcal{V}$. This implies that the probability that $\Phi(q,1)$ being a contraction is positive. Since all $\Phi(q,1)$ are non-expansive, there is a number $\rho(1) < 1$ such that $\mathbb{E}\,[\![\Phi(q,1)]\!] = \rho(1)$. Straightforwardly, it also holds $\mathbb{E}\,[\![\Phi(k+q,k)]\!] = \rho(k) < 1$ for all $k < \infty$. Thus there a.s. holds that

$$\begin{aligned}\mathbb{E}\big[V\left(e_{k+q}\right)\big|\,\mathcal{F}_k\big] - V(e_k) &= \mathbb{E}\,[\![\Phi\left(k+q,k\right)e_k]\!] - V(e_k)\\ &\leq \mathbb{E}\,[\![\Phi\left(k+q,k\right)]\!] \cdot [\![e_k]\!] - V(e_k) = (\rho(k)-1)V(e_k).\end{aligned}$$

Similarly as in the proof of Theorem 4.1, the condition b) in Theorem 4.6 ensures that $\sum_{i=1}^{\infty}(1-\rho(k)) = \infty$. It follows that $V\left(e_k\right) \xrightarrow{a.s.} 0$ as $t \to \infty$ since $V(e_0) -$

$\mathbb{E}\big[V(e_{nq})|\mathcal{F}_k\big] < \infty$ for any N. Define the set $\mathcal{Q} := \{e : V(e) \leq V(e_1)\}$ for any initial e_1 corresponding to x_1. For any random sequence $\{\mathcal{G}(k)\}$, it follows from the system dynamics (4.31) that

$$V(e_k) \leq V(e_{k-1}) \cdots \leq V(e_2) \leq V(e_1),$$

and thus e_k will stay within the set $\mathcal{Q}$ with probability 1. From Theorem 3.1 and Corollary 3.1, it follows that e_k asymptotically converges to $\{e : V(e) = 0\}$ almost surely. Moreover, since $V(e)$ is a norm of e, it can be concluded from Corollary 3.1 that the error system (4.31) is globally a.s. asymptotically stable. The proof is complete. □

It is worth mentioning that the error system is globally a.s. exponentially stable under the assumption that the probability of the composition of any sequence of randomly-changing graphs, $\{\mathcal{G}(k), \ldots, \mathcal{G}(k+1), \mathcal{G}(k+l-1)\}$, for any $k \in \mathbb{N}$, being strongly connected is lower bounded by some positive number. This can be proven with the help of Theorem 3.2 and Corollary 3.2.

4.5 Concluding Remarks

In this chapter, we have shown how the finite-step Lyapunov criteria established in the Chapter 3 can be applied to studying several distributed coordination algorithms. As the first application, we look at the product of random sequences of stochastic matrices, including those with zero diagonal entries, and obtain sufficient conditions to ensure that the product almost surely converges to a matrix with identical rows; we also show that the rate of convergence can be exponential under additional conditions. Using these results, we have further investigated how asynchronous updating events can induce agreement among agents coupled by periodic networks. As another application, we have studied a distributed network algorithm for solving linear algebraic equations. We relax the existing conditions on the network structures, while still guaranteeing the equations are solved asymptotically.

4.6 Appendix: An Alternative Proof of Corollary 4.2

For ergodic stationary sequences, the following important property is the key to construct the convergence rate.

Lemma 4.2 (Birkhoff's Ergodic Theorem, see [109, Th. 7.2.1])**.** *For an ergodic sequence* $\{X_k\}, k \in \mathbb{N}_{\geq 0}$*, of random variables, it holds that*

$$\lim_{m\to\infty} \frac{1}{m} \sum\nolimits_{k=0}^{m-1} X_k \xrightarrow{a.s.} E(X_0) \tag{4.32}$$

For the product given in (4.1), we say $W(k,0)$ converges to a rank-one matrix $W = 1\xi^\top$ a.s. as $k \to \infty$ if $\tau(W(k,0)) \to 0$ as $k \to \infty$, where $\tau(\cdot)$ is defined in (4.2). According to Definition 3.1, if there exists $\beta > 1$ such that

$$\beta^k \tau\big(W(k,0)\big) \xrightarrow{a.s.} 0, k \to \infty, \tag{4.33}$$

then the convergence rate is said to be exponential at the rate no slower than β^{-1}. We are now ready to present the proof of Corollary 4.2.

Proof of Corollary 4.2. Let h be the same as that in Assumption 4.2. There is an integer $\theta \in \mathbb{N}$ such that $W(t+\theta h, t)$ is scrambling with positive probability. Let $T = \theta h$. Consider a sufficiently large r, and then $W(r,0)$ can be written as

$$W(r,0)) = \bar{W} \cdot W\big(mT, (m-1)\,T\big) \cdots W\,(T,0)\,,$$

where m is the largest integer such that $mT \leq r$, $W\,(kT+T, kT)\,, k = 0, \cdots, m-1$, are the matrix products defined by (4.1), and $\bar{W} = W(r, mT)$ is the remaining part, which is obviously a stochastic matrix. To study the limiting behavior of $W(r,0)$, we compute its coefficients of ergodicity

$$\begin{aligned}\tau\big(W\,(r,0)\big) &\leq \tau\,(\bar{W}) \prod\nolimits_{k=0}^{m-1} \tau\big(W\,(kT+T, kT)\,\big)\\ &\leq \prod\nolimits_{k=0}^{m-1} \tau\big(W\,(kT+T, kT)\,\big),\end{aligned}$$

where the property (4.4) has been used. The last inequality follows from the property of coefficients of ergodicity, i.e., $\tau(A) \leq 1$ for a stochastic matrix A. Taking logarithms yields that

$$\log \tau\big(W\,(r,0)\big) \leq \sum\nolimits_{k=0}^{m-1} \log \tau\big(W\,(KT+T, kT)\,\big). \tag{4.34}$$

Since the sequence $\{W(k)\}$ is ergodic, it is easy to see that the sequence of products $\{W\,(kT+T, kT)\}$, $k = 0, \cdots, m-1$, over non-overlapping intervals of length T, is also ergodic. It follows in turn that $\{\log \tau\big(W\,(kT+T,kT)\,\big)\}$ is ergodic. From Lemma 4.2, one can further obtain

$$\lim_{m\to\infty}\frac{1}{m}\sum_{k=0}^{m-1}\log \tau\big(W\,(kT+T,kT)\,\big)$$
$$\xrightarrow{a.s.}\mathbb{E}\Big[\log \tau\big(W\,(T,0)\big)\Big]\le \log \mathbb{E}\big[\tau\big(W\,(T,0)\big)\big].$$

The last inequality follows from Jensen's inequality (see [109, Th. 1.5.1]) since $\log(\cdot)$ is concave. According to Assumption 4.1, one knows that $W(t+h,t)$ is scrambling with positive probability, and thus it follows that $0<\mathbb{E}\big[\tau\,(W\,(T,0))\,\big]<1$. Taking a positive number λ satisfying $\lambda<-\log \mathbb{E}\big[\tau\big(W\,(T,0)\big)\big]$, one obtains

$$m\lambda+\sum\nolimits_{k=0}^{m-1}\log \tau\big(W\,(KT+T,kT)\big)\xrightarrow{a.s.}-\infty.$$

Adding $m\lambda$ to both sides of (4.34) yields that

$$m\lambda+\log \tau\big(W\,(r,0)\big)$$
$$\le m\lambda+\sum\nolimits_{k=0}^{m-1}\log \tau\big(W\,(kT+T,kT)\big)\xrightarrow{a.s.}-\infty.$$

It follows straightforwardly that

$$\big(e^{\lambda}\big)^{m}\tau\big(W\,(r,0)\big)\xrightarrow{a.s.}0.$$

Let $\beta=e^{\lambda}$, which apparently satisfies $\beta>1$. From Definition 3.1, one can conclude that the product $W(k,0)$ almost surely converges to a rank-one stochastic matrix exponentially at a rate no slower than β^{-1}, which completes the proof. □

Part II

Partial Synchronization of Kuramoto Oscillators:

Partial Stability Methods

Overview of Part II

Synchronization is a ubiquitous phenomenon that has been observed pervasively in many natural, social and man-made systems [46, 136–138]. Remarkable examples include synchronized flashing of fireflies [4], animal flocking [7], pedestrian footwalk synchrony on London's Millennium Bridge [139], phase synchronization of coupled Josephson junction circuits [140], and synchronous operation of power generators [49].

Global synchronization describes the situation where all units in a network evolve in unison. Strong network coupling plays a fundamental role in the emergence and stability of global synchronization [78]. Recently, another form of synchronization, termed partial synchronization, has attracted a lot of attention [82, 141, 142]. In contrast to global synchronization, partial synchronization characterizes a circumstance in which only some parts of, instead of all, units in a network have similar dynamics. It is believed to be more common [82] in nature, for example in the human brain.

Neuronal synchronization across cortical regions of the human brain, which has been widely detected through recording and analyzing brain waves, is believed to facilitate communication among neuronal ensembles [55]. Only closely correlated oscillating neuronal ensembles can exchange information effectively, because their input and output windows are open at the same time [52]. In healthy human brain, it is frequently observed that only a part of its cortical regions are synchronized [59], and such a phenomenon is commonly referred to as partial phase cohesiveness or partial synchronization of brain neural networks. In contrast, in the pathological brain of an epileptic patient, global synchronization of neural activities are detected to take place across the entire brain [60]. These observations suggest that healthy brain has powerful regulation mechanisms that are not only able to render synchronization, but also capable of preventing unnecessary synchronization among neuronal ensembles. Partly motivated by these experimental studies, researchers are interested in theoretically studying cluster synchronization [82, 85, 142, 143] and chimera states [88], even though analytical results are much more difficult to obtain, while analytical results for global synchronization are ample, e.g., [78, 144, 145].

In this part of the thesis, our objective is to identify some possible underlying mechanisms that could give rise to partial synchronization in complex networks, particularly in human brain networks. The Kuramoto model and its variations [62] will be used to describe the dynamics of oscillators. We first investigate in Chapter 5 how partial synchronization can take place among directly connection regions. We find that strong local or regional coupling is a possible mechanism. Oscillators that are tightly connected can exhibit coordinating behavior, while the rest that are weakly connected to them remain different. In addition, we also study how remote synchronization, a phenomenon also detected in the human brain [92], can take place in star networks. In order to study remote synchronization, we develop some new criteria for partial stability of nonlinear systems in Chapter 6. These new criteria are then used to analytically study remote synchronization in Chapter 7.