

Sampled-Data Consensus Over Random Networks

Junfeng Wu, Ziyang Meng, Tao Yang, Guodong Shi, and Karl Henrik Johansson, *Fellow, IEEE*

Abstract—This paper considers the consensus problem for a network of nodes with random interactions and sampled-data control actions. We first show that consensus in expectation, in mean square, and almost surely are equivalent for a general random network model when the inter-sampling interval and maximum node degree satisfy a simple relation. The three types of consensus are shown to be simultaneously achieved over an independent or a Markovian random network defined on an underlying graph with a directed spanning tree. For both independent and Markovian random network models, necessary and sufficient conditions for mean-square consensus are derived in terms of the spectral radius of the corresponding state transition matrix. These conditions are then interpreted as the existence of critical value on the inter-sampling interval, below which a global mean-square consensus is achieved and above which the system diverges in a mean-square sense for some initial states. Finally, we establish an upper bound on the intersampling interval below which almost sure consensus is reached, and a lower bound on the intersampling interval above which almost sure divergence is reached. Some numerical simulations are given to validate the theoretical results and some discussions on the critical value of the inter-sampling intervals for the mean-square consensus are provided.

Index Terms—Consensus, Markov chain, sampled-data, random networks.

I. INTRODUCTION

IN traditional consensus algorithm, each node exchanges information with a few neighbors, typically given by their relative states, and then updates its own state according to a weighted average. It turns out that with suitable (and rather general) connectivity conditions imposed on the communication graph, all nodes asymptotically reach an agreement in which the nodes' initial values are encoded [1], [2]. Various consensus algorithms have been proposed in the literature. The most common continuous-time consensus algorithm is given by an ordinary differential equation in terms of the relative states of each agent with respect to its neighboring agents [2], [3]. The agent state is driven towards the states of its neighbors, so eventually the algorithm ensures that the

whole network reaches an agreement provided that the network is jointly connected. In [4], [5], the authors developed discrete-time consensus algorithms. In such algorithms, each agent updates its states as a convex combination of the state of itself and that of its neighboring agents. Due to the fact that most algorithms are implemented by a digital device and that the communication channels are unreliable and often subject to limited communication capacity, sampled-data consensus algorithms have also been proposed [6]–[10]. In a sampled-data setting, agent dynamics are continuous and control input is piecewise continuous. The closed-loop system is transformed into discrete-time dynamics and conditions on uniform or nonuniform sample periods are critical to ensure consensus.

Consensus over random networks has drawn much attention since communication networks are naturally random. In [11], [12], the authors studied distributed average consensus in sensor networks with quantized data and independent, identically distributed (i.i.d.) symmetric random topologies. The authors of [13] evaluated the mean-square convergence of consensus algorithms with random asymmetric topologies. Mean-square performance for consensus algorithms over i.i.d. random graphs was studied in [14], and the impact of random packet drops was investigated in [15]. Recently, the i.i.d. assumption was relaxed in [16], [17] to the case where the communication graph is modeled by a finite-state Markov chain. Probabilistic consensus has also been investigated in the literature. It was shown in [18] that for a random network generated by i.i.d. stochastic matrices, almost sure, in probability, and L^p ($p \geq 1$) consensus are equivalent. In [19], the authors showed that almost sure convergence is reached for i.i.d. random graphs and Erdős-Rényi random graphs. The analysis was later extended to directed graphs and more general random graph processes [20], [21]. In [22], the authors showed that for a stochastic linear dynamical system asymptotic almost sure consensus over i.i.d. random networks is reached if and only if the graph contains a directed spanning tree in expectation. The [23] provided a necessary and sufficient condition for consensus over ergodic and stationary graph processes. Divergence in random consensus networks has also been considered, as representing asymptotic disagreement in social networks. Almost sure divergence of consensus algorithms was considered in [24], [25].

In this paper, we consider sampled-data consensus problems over random networks. In the presence of sampled-data control actions, the sampled-data consensus problem is converted into a discrete-time consensus algorithm over directed random networks. Due to the effect of the inter-sampling interval, at sampling instants each node updates its own state not necessarily as a nonnegative-weighted average of the state of itself and that of its neighboring nodes. We analyze the convergence of the consensus algorithm under two random network models. In the first model, each node independently samples its neighbors in a random manner over the underlying graph, while in the second model each node samples its neighbors by following a Markov chain. The impact of sampling intervals on consensus convergence and divergence is studied. We believe that the models considered in this paper are applicable to some applications since they incorporate sampling by digital devices, limited node connections, and random interactions imposed by unreliable networks. Three types of

Manuscript received March 18, 2015; revised February 02, 2016; accepted April 14, 2016. Date of publication May 12, 2016; date of current version July 21, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Amir Asif. This work was supported in part by the Knut and Alice Wallenberg Foundation, the Swedish Research Council, and the NNSF of China under Grant No. 61120106011. (*Corresponding author: Guodong Shi.*)

J. Wu and K. H. Johansson are with the ACCESS Linnaeus Center, School of Electrical Engineering, Royal Institute of Technology, Stockholm 114 28, Sweden (e-mail: junfengw@kth.se; kallej@kth.se).

Z. Meng is with the State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instrument, Tsinghua University, Beijing 100084, China (e-mail: ziyangmeng@mail.tsinghua.edu.cn).

T. Yang is with the Department of Electrical Engineering, University of North Texas, Denton, TX 76203 USA (e-mail: taoyang.work@gmail.com).

G. Shi is with the College of Engineering and Computer Science, The Australian National University, Canberra 0200, Australia (e-mail: guodong.shi@anu.edu.au).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2016.2568168

consensus—consensus in expectation, mean-square and almost sure sensor—are considered. The main contributions of this paper are summarized as follows. For both independent and Markovian random network models, necessary and sufficient conditions for mean-square consensus are derived in terms of the spectral radius of the corresponding state transition matrix. These conditions can be interpreted as critical thresholds on the inter-sampling interval and we show that they can be computed by a generalized eigenvalue problem, which can be further stated as a quasi-convex optimization problem. For each random network model, we obtain an upper bound on the inter-sampling interval below which almost sure convergence is reached, and a lower bound on the inter-sampling interval above which almost sure divergence is reached. To the best of our knowledge, this is the first time that almost sure consensus convergence and divergence are studied for sampled-data systems, and also the first time that almost sure divergence is considered for Markovian random graphs.

The remainder of the paper is organized as follows. Section II provides the problem formulation and introduces the probabilistic consensus notions. Then their relations are discussed. Section III focuses on independent random networks. In this section, we present necessary and/or sufficient conditions for expectation consensus, mean-square consensus, almost sure consensus, and almost sure divergence. The same problems are addressed under a Markovian network in Section IV. In Section V, we illustrate our theoretical results through numerical simulations. Finally, some concluding remarks are drawn in Section VI.

Notation: \mathbb{N} , \mathbb{C} , \mathbb{R} and \mathbb{R}_+ are the sets of nonnegative integers, complex numbers, real numbers and positive real numbers, respectively. For $x, y \in \mathbb{R}$, $x \vee y$ and $x \wedge y$ stand for the maximum and minimum of x and y , respectively. The set of n by n positive semi-definite (positive definite) matrices (that are restricted to be Hermitian) over the field \mathbb{C} is denoted as \mathbb{S}_+^n (\mathbb{S}_{++}^n). For simplicity, we write $X \geq Y$ ($X > Y$), where $X, Y \in \mathbb{S}_+^n$, if $X - Y \in \mathbb{S}_+^n$ ($X - Y \in \mathbb{S}_{++}^n$). For a matrix $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$, $\|X\|$ represents the spectral norm of X ; X^* and X' are the Hermitian conjugate and the transpose of X , respectively. The Kernel of X is defined as $\ker(X) = \{v \in \mathbb{R}^n : Xv = 0\}$. $\text{vec}(X)$ is the vectorization of X , i.e., $\text{vec}(X) := [x'_1, x'_2, \dots, x'_n]' \in \mathbb{R}^{mn}$. \otimes denotes a Kronecker product of two matrices. If $m = n$, $\rho(X)$ and $\text{Tr}(X)$ are the spectral radius and the trace of X , respectively. For vectorization and Kronecker product, the following properties are frequently used in this work: *i)* $\text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$; *ii)* $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, where A, B, C and D are matrices of compatible dimensions. For vectors $x, y \in \mathbb{R}^n$, $x \perp y$ is a short hand for $\langle x, y \rangle = 0$, where $\langle \cdot, \cdot \rangle$ denotes Euclidean inner product. For a set \mathcal{A} , $2^{\mathcal{A}}$ means the power set of \mathcal{A} . The indicator function of a subset $\mathcal{A} \subset \Omega$ is a function $1_{\mathcal{A}} : \Omega \rightarrow \{0, 1\}$, where $1_{\mathcal{A}}(\omega) = 1$ if $\omega \in \mathcal{A}$, and $1_{\mathcal{A}}(\omega) = 0$ if $\omega \notin \mathcal{A}$. The notation $\sigma(\cdot)$ represents the σ -algebra generated by random variables. Depending on the argument, $|\cdot|$ stands for the absolute value of a real number, or the cardinality of a set.

II. PROBLEM FORMULATION

A. Sampling and Random Networks

Consider a network of N nodes indexed in the set $V = \{1, 2, \dots, N\}$. Each node i holds a value $x_i(t) \in \mathbb{R}$ for $t \in [0, \infty)$. The evolution of $x_i(t)$ is described by

$$\dot{x}_i(t) = u_i(t), \quad (1)$$

where $u_i \in \mathbb{R}$ is the control input.

The directed interaction graph $G = (V, E)$ describes underlying information exchange. Here $E \subseteq V \times V$ is an arc set and $(j, i) \in E$ means there is a (possibly unreliable) communication link from node j to node i . The set of neighbors of node i in the underlying graph G is denoted as $\mathcal{N}_i := \{j \in V : (j, i) \in E\}$. The maximum degree of G is defined as $D_{\max} := \max_{i \in V} |\mathcal{N}_i|$.

The Laplacian matrix $L := [l_{ij}] \in \mathbb{R}^{N \times N}$ associated with G is defined as

$$l_{ij} = \begin{cases} -1, & \text{if } i \neq j \text{ and } (j, i) \in E \\ \sum_{m \neq i} 1_{\{(m, i) \in E\}}, & \text{if } i = j. \end{cases}$$

A directed *path* from node i_1 to node i_l is a sequence of nodes $\{i_1, \dots, i_l\}$ such that $(i_j, i_{j+1}) \in E$ for $j = 1, \dots, l-1$. A directed tree is a directed subgraph of $G = (V, E)$ such that every node has exactly one parent, except a single root node with no parents. Therefore, there must exist a directed path from the root to every other node. A directed spanning tree is a directed tree that contains all the nodes of G .

Let \mathcal{G} be a σ -algebra associated with G , which contains all subgraphs of G , and $\{G_k = (V, E_k)\}_{k \in \mathbb{N}}$ be a sequence of random graphs, in which by definition each G_k is a random variable taking values in \mathcal{G} . The Laplacian matrix $L(k) := [l_{ij}(k)] \in \mathbb{R}^{N \times N}$ associated with G_k is defined as

$$l_{ij}(k) = \begin{cases} -1, & \text{if } i \neq j \text{ and } (j, i) \in E_k \\ \sum_{m \neq i} 1_{\{(m, i) \in E_k\}}, & \text{if } i = j. \end{cases}$$

The set of neighbors of node i in denoted as $\mathcal{N}_i(k) := \{j : (j, i) \in E_k\}$. Let the triple $(\mathcal{G}^{\mathbb{N}}, \mathcal{F}, \mathbb{P})$ denote the probability space capturing the randomness contained in the random graph sequence, where \mathcal{F} is the set of all subsets of $\mathcal{G}^{\mathbb{N}}$. Furthermore, we define a filtration $\mathcal{F}_k = \sigma(G_0, \dots, G_k)$ for $k \in \mathbb{N}$.

We define a sequence of node sampling instants as $0 = t_0 < \dots < t_k < t_{k+1} < \dots$ with $\tau_k = t_{k+1} - t_k$ representing the inter-sampling interval. The sampled-data consensus scheme associated with the random graph sequence $\{G_k\}_{k \in \mathbb{N}}$ is given by

$$u_i(t) = \sum_{j \in \mathcal{N}_i(k)} [x_j(t_k) - x_i(t_k)], \quad t \in [t_k, t_{k+1}). \quad (2)$$

The closed-loop system can then be written in the compact form

$$x(t_{k+1}) = [I - \tau_k L(k)] x(t_k) := W(k) x(t_k) \quad (3)$$

with $W(k) := [w_{ij}(k)]$ and $x(t_k) := [x_1(t_k)', \dots, x_N(t_k)']'$. Note that in $W(k)$, τ_k is a positive weight of $L(k)$. Since τ_k can be arbitrarily large, $W(k)$ is not necessarily nonnegative.

Remark 1: In the sampled-data algorithm (3), each node samples its own state at the sampling instants $\{t_k\}_{k=0}^{\infty}$. If each node has continuous access to its own state for all $t \geq 0$, we can introduce the algorithm:

$$u_i(t) = \sum_{j \in \mathcal{N}_i(t_k)} [x_j(t_k) - x_i(t)], \quad t \in [t_k, t_{k+1}), \quad (4)$$

as considered in [26]. The the corresponding closed-loop system is then

$$x(t_{k+1}) = [I - (1 - e^{-\tau_k}) L(k)] x(t_k). \quad (5)$$

By replacing τ_k in (3) with $1 - e^{-\tau_k}$ in (5), all conclusions for (3) throughout the paper can thus be readily translated into those for (4).

B. Consensus Metrics

Define $x_{\max}(t_k) := \max_{i \in V} x_i(t_k)$ and $x_{\min}(t_k) := \min_{i \in V} x_i(t_k)$ and the agreement measure $\mathfrak{X}(k) := x_{\max}(t_k) - x_{\min}(t_k)$. We have the following definitions for consensus convergence and divergence.

Definition 1:

- i) Algorithm (3) achieves (global) consensus in expectation if for any initial state $x(t_0) \in \mathbb{R}^N$ there holds $\lim_{k \rightarrow \infty} \mathbb{E}[\mathfrak{X}(k)] = 0$.
- ii) Algorithm (3) achieves (global) consensus in mean square if for any initial state $x(t_0) \in \mathbb{R}^N$ there holds $\lim_{k \rightarrow \infty} \mathbb{E}[\mathfrak{X}^2(k)] = 0$.
- iii) Algorithm (3) achieves (global) consensus almost surely if for any initial state $x(t_0) \in \mathbb{R}^N$ there holds $\mathbb{P}(\lim_{k \rightarrow \infty} \mathfrak{X}(k) = 0) = 1$.
- iv) Algorithm (3) diverges almost surely if there holds $\mathbb{P}(\limsup_{k \rightarrow \infty} \mathfrak{X}(k) = \infty) = 1$ for any initial state $x(t_0) \in \mathbb{R}^N$ except for $x(t_0) \perp \mathbf{1}$, where $\mathbf{1} := [1, \dots, 1]^T \in \mathbb{R}^N$.

We focus on whether or not Algorithm (3) is able to achieve agreement in terms of various metrics rather than which limiting point the algorithm agrees to, if any. The latter problem is out of the scope of the present paper, and will be pursued in the future.

C. Relations of Consensus Notions

The following lemma suggests that if the inter-sampling interval is small enough, the consensus notations in Definition 1 are equivalent.

Lemma 1: Suppose $\tau_k \in (0, 1/D_{\max}]$ for all k . Then expectation consensus, mean-square consensus, and almost sure consensus are all equivalent for Algorithm (3).

Proof: We begin with the observation that $W(k)$ is a row stochastic matrix for all $k \in \mathbb{N}$ when $\tau_k \in (0, 1/D_{\max}]$, where a row stochastic matrix means a nonnegative square matrix with the sum of each row being 1. Therefore,

$$\begin{aligned} x_{\max}(t_{k+1}) &= \max_{i \in V} \sum_{j=1}^N w_{ij}(k) x_j(t_k) \\ &\leq \max_{i \in V} \sum_{j=1}^N w_{ij}(k) (x_j(t_k) \vee x_{\max}(t_k)) \\ &= x_{\max}(t_k), \end{aligned} \quad (6)$$

implying that $x_{\max}(t_k)$ is non-increasing in k . We show that $x_{\min}(t_k)$ is non-decreasing in k in precisely the same way. The foregoing two observations together suggest that $\mathfrak{X}(k)$ is non-increasing in k . Finally, the conclusion follows by showing the following implications:

- i) Expectation consensus \implies mean-square consensus. Since $\mathfrak{X}(k)$ is non-increasing, we have $\mathbb{E}[\mathfrak{X}^2(k)] \leq \mathfrak{X}(0)\mathbb{E}[\mathfrak{X}(k)]$. By the hypothesis, $\mathbb{E}[\mathfrak{X}^2(k)] \leq \mathfrak{X}(0)\mathbb{E}[\mathfrak{X}(k)] \rightarrow 0$ as $k \rightarrow \infty$.
- ii) Mean-square consensus \implies almost sure consensus. According to Chebyshev's inequality [27], $\mathbb{P}(|\mathfrak{X}(k)| > \epsilon) \leq \frac{\mathbb{E}[\mathfrak{X}^2(k)]}{\epsilon^2}$ holds for any $\epsilon > 0$. If $\lim_{k \rightarrow \infty} \mathbb{E}[\mathfrak{X}^2(k)] = 0$, then $\lim_{k \rightarrow \infty} \mathbb{P}(|\mathfrak{X}(k)| > \epsilon) = 0$. As a result, there exists a subsequence of $\{\mathfrak{X}(k)\}_{k \in \mathbb{N}}$ that converges to 0 almost surely [28]. Since $\{\mathfrak{X}(k)\}_{k \in \mathbb{N}}$ is non-increasing, $\lim_{k \rightarrow \infty} \mathfrak{X}(k) = 0$ almost surely.

- iii) Almost sure consensus \implies expectation consensus. Since the sequence $\{\mathfrak{X}(k)\}_{k \in \mathbb{N}}$ is nonnegative and non-increasing, and $\mathfrak{X}(0)$ is given, by the Monotone Convergence Theorem [28], $\lim_{k \rightarrow \infty} \mathbb{E}[\mathfrak{X}(k)] = 0$. \blacksquare

Remark 2: In [29], the equivalence of L^p consensus, consensus in probability, and almost sure consensus was proved over a random network generated by i.i.d. stochastic matrices. In Lemma 1, we show that this equivalence holds regardless of the type of random process the row stochastic matrices are generated by.

III. INDEPENDENT RANDOM NETWORKS

In this section, we investigate sampled-data consensus when the random graph G_k is obtained by each node independently sampling its neighbors in a random manner over G . Regarding the connectivity of the underlying graph G , we adopt the following assumption:

- (A1) *The underlying graph G has a directed spanning tree.*
We also impose the following assumption.
- (A2) *The random variables $1_{\{(j,i) \in E_k\}}$, $(j,i) \in E$, $k \in \mathbb{N}$, are (temporally and spatially) i.i.d. Bernoulli with mean $q > 0$.*

The techniques developed in this section also apply when $q = q(i)$ is a function of node index i . In order to simplify the notation used in the derivation of the results through this section, we also make the following assumption.

- (A3) Let $\tau_k = \tau_*$ for all k with $\tau_* > 0$.

When each node samples its neighbors as Assumption (A2) describes, $\{L(k)\}_{k \in \mathbb{N}}$ are i.i.d. random variables, whose randomness originates from the primitive random variables $1_{\{(j,i) \in E_k\}}$'s. We denote the sample space of $L(k)$ by $\mathcal{L} := \{L^{(1)}, L^{(2)}, \dots, L^{(M)}\}$, where $M = |\mathcal{G}|$ and $L^{(l)} := [L_{ij}^{(l)}] \in \mathbb{R}^{N \times N}$ is the Laplacian matrix associated with a subgraph $G^{(l)} \in \mathcal{G}$. By counting how many edges are present in G_k and how many are absent from G_k , respectively, the distribution of $L(k)$ is computed by

$$\mathbb{P}\left(L(k) = L^{(i)}\right) = q^{\text{Tr}(L^{(i)})} (1-q)^{\text{Tr}(L-L^{(i)})} := \pi_i \quad (7)$$

for $i = 1, \dots, M$. When $\tau_k = \tau_*$, $W(k)$ inherits the same distribution as $L(k)$ from G_k . We denote $W^{(l)} := I - \tau_* L^{(l)}$.

A. Conjunction of Various Consensus Metrics

When the inter-sampling interval is small enough (to be precise $\tau_* < 1/D_{\max}$, each node updates its state as a convex combination of the previous states of its own and its neighbors. Every update drives nodes' states closer to each other and can be thought of as attraction of the nodes' states. Under the independent random network model, we show in the following theorem that Algorithm (3) achieves consensus, simultaneously in expectation, in mean square, and in almost sure sense, provided that G has a directed spanning tree.

Theorem 1: Let Assumptions (A1), (A2), and (A3) hold. Then expectation consensus, mean-square consensus, and almost sure consensus are achieved under Algorithm (3) if $\tau_* \in (0, 1/D_{\max})$.

Proof: By Lemma 1, it suffices to show that Algorithm (3) achieves consensus in expectation.

Fix a directed spanning tree $G_T := (V, E_T)$ of graph G and a sampling time t_k . Let the root of G_T be $i_1 \in V$, and define a set $\mathcal{M}_1 := \{i_1\}$. Denote

$$\eta := (\tau_*) \wedge (1 - D_{\max} \tau_*).$$

Then, there holds $\eta > 0$ when $\tau_* \in (0, 1/D_{\max})$. We assume $x_{i_1}(t_k) \leq 1/2(x_{\max}(t_k) + x_{\min}(t_k))$ while the other case for $x_{i_1}(t_k) > 1/2(x_{\max}(t_k) + x_{\min}(t_k))$ will be discussed later.

Choose a node $i_2 \in V$ such that $i_2 \notin \mathcal{M}_1$ and $(i_1, i_2) \in E_T$. Define $\mathcal{M}_2 := \mathcal{M}_1 \cup \{i_2\}$. Consider the event $\mathcal{E}_2 := \{(i_1, i_2) \in E_{k+1}\}$. When \mathcal{E}_2 happens, $x_{i_2}(t_{k+1})$ evolves as follows:

$$\begin{aligned} x_{i_2}(t_{k+1}) &= w_{i_2 i_1}(k) x_{i_1}(t_k) + \sum_{j \neq i_1} w_{i_2 j}(k) x_j(t_k) \\ &\leq \frac{1}{2} w_{i_2 i_1}(k) (x_{\min}(t_k) + x_{\max}(t_k)) \\ &\quad + (1 - w_{i_2 i_1}(k)) x_{\max}(t_k) \\ &\leq \frac{1}{2} \eta x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta\right) x_{\max}(t_k), \end{aligned}$$

where the last inequality holds because $\eta \leq w_{i_2 i_1}(k)$. Since $\eta \leq w_{i_1 i_1}(k)$, we show that $x_{i_1}(t_{k+1})$ is bounded by

$$x_{i_1}(t_{k+1}) \leq \frac{1}{2} \eta x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta\right) x_{\max}(t_k).$$

At time t_{k+2} ,

$$\begin{aligned} x_{i_2}(t_{k+2}) &= w_{i_2 i_2}(k+1) x_{i_2}(t_{k+1}) + \sum_{j \neq i_2} w_{i_2 j}(k+1) x_j(t_{k+1}) \\ &\leq w_{i_2 i_2}(k+1) \left[\frac{1}{2} \eta x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta\right) x_{\max}(t_k) \right] \\ &\quad + (1 - w_{i_2 i_2}(k+1)) x_{\max}(t_{k+1}) \\ &\leq \frac{1}{2} \eta^2 x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta^2\right) x_{\max}(t_k), \end{aligned}$$

where the last inequality is due to $x_{\max}(t_{k+1}) \leq x_{\max}(t_k)$ by (6) and $\eta \leq w_{i_2 i_2}(k+1)$. The same is true of node i_1 , i.e., $x_{i_1}(t_{k+2}) \leq \frac{1}{2} \eta^2 x_{\min}(t_k) + (1 - \frac{1}{2} \eta^2) x_{\max}(t_k)$. Recursively, we see that $x_{i_1}(t_{k+n}) \leq \frac{1}{2} \eta^n x_{\min}(t_k) + (1 - \frac{1}{2} \eta^n) x_{\max}(t_k)$ and $x_{i_2}(t_{k+n}) \leq \frac{1}{2} \eta^n x_{\min}(t_k) + (1 - \frac{1}{2} \eta^n) x_{\max}(t_k)$ holds for $n = 1, 2, \dots$

Again, choose a node $i_3 \in V$ such that $i_3 \notin \mathcal{M}_2$ and there exists a node $j \in \mathcal{M}_2$ satisfying $(j, i_3) \in E_T$. Define $\mathcal{M}_3 := \mathcal{M}_2 \cup \{i_3\}$. Consider the event $\mathcal{E}_3 := \{(j, i_3) \in E_{k+2} : (j, i_3) \in E_T, j \in \mathcal{M}_2\}$. If \mathcal{E}_3 happens, we obtain a similar result for node i_3 :

$$\begin{aligned} x_{i_3}(t_{k+2}) &\leq \eta (x_{i_1}(t_{k+1}) \vee x_{i_2}(t_{k+1})) + (1 - \eta) x_{\max}(t_{k+1}) \\ &\leq \frac{1}{2} \eta^2 x_{\min}(t_k) + \left(\eta - \frac{1}{2} \eta^2\right) x_{\max}(t_k) + (1 - \eta) x_{\max}(t_k) \\ &= \frac{1}{2} \eta^2 x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta^2\right) x_{\max}(t_k). \end{aligned}$$

From the same argument as above,

$$x_{i_3}(t_{k+n}) \leq \frac{1}{2} \eta^n x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta^n\right) x_{\max}(t_k)$$

holds for $n = 2, 3, \dots$

We choose nodes i_1, \dots, i_N in sequel and accordingly define $\mathcal{M}_1, \dots, \mathcal{M}_N$ and $\mathcal{E}_2, \dots, \mathcal{E}_N$. Consider $\mathcal{E}_2, \dots, \mathcal{E}_N$ sequentially

happen, then

$$x_{i_m}(t_{k+n}) \leq \frac{1}{2} \eta^n x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta^n\right) x_{\max}(t_k)$$

holds for all $1 \leq m \leq N$ and $n \leq N - 1$, which entails

$$\begin{aligned} x_{\max}(t_{k+N-1}) &= \max_i x_i(t_{k+N-1}) \\ &\leq \frac{1}{2} \eta^{N-1} x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta^{N-1}\right) x_{\max}(t_k). \end{aligned}$$

In this case, the relationship between $\mathfrak{X}(k + N - 1)$ and $\mathfrak{X}(k)$ is given by

$$\begin{aligned} \mathfrak{X}(k + N - 1) &= x_{\max}(t_{k+N-1}) - x_{\min}(t_{k+N-1}) \\ &\leq \frac{1}{2} \eta^{N-1} x_{\min}(t_k) + \left(1 - \frac{1}{2} \eta^{N-1}\right) x_{\max}(t_k) - x_{\min}(t_k) \\ &= \left(1 - \frac{1}{2} \eta^{N-1}\right) \mathfrak{X}(k). \end{aligned}$$

If $x_{i_1}(t_k) > 1/2(x_{\max}(t_k) + x_{\min}(t_k))$ is assumed, a symmetric analysis leads to that, when $\mathcal{E}_2, \dots, \mathcal{E}_N$ sequentially occur, $x_{\min}(t_{k+N-1}) \geq \frac{1}{2} \eta^{N-1} x_{\max}(t_k) + (1 - \frac{1}{2} \eta^{N-1}) x_{\min}(t_k)$. Then $\mathfrak{X}(k + N - 1)$ is bounded by

$$\begin{aligned} \mathfrak{X}(k + N - 1) &= x_{\max}(t_{k+N-1}) - x_{\min}(t_{k+N-1}) \\ &\leq x_{\max}(t_k) - \frac{1}{2} \eta^{N-1} x_{\max}(t_k) - \left(1 - \frac{1}{2} \eta^{N-1}\right) x_{\min}(t_k) \\ &= \left(1 - \frac{1}{2} \eta^{N-1}\right) \mathfrak{X}(k), \end{aligned}$$

exactly the same result as when $x_{i_1}(t_k) \leq 1/2(x_{\max}(t_k) + x_{\min}(t_k))$ is assumed. Therefore, the above inequality holds irrespective of the state of $x_{i_1}(t_k)$.

In addition, we know that probability that the events $\mathcal{E}_2, \dots, \mathcal{E}_N$ sequentially occur is

$$\mathbb{P}\left(1_{\cap_{i=2}^N \mathcal{E}_i} = 1\right) = \prod_{i=2}^N \mathbb{P}(1_{\mathcal{E}_i} = 1) \geq q^{N-1}.$$

Combining all the above analysis,

$$\begin{aligned} \mathbb{E}[\mathfrak{X}(k + N - 1)] &\leq q^{N-1} \left(1 - \frac{1}{2} \eta^{N-1}\right) \mathbb{E}[\mathfrak{X}(k)] + (1 - q^{N-1}) \mathbb{E}[\mathfrak{X}(k)] \\ &= \left(1 - \frac{1}{2} (q\eta)^{N-1}\right) \mathbb{E}[\mathfrak{X}(k)]. \end{aligned} \quad (8)$$

Since $0 < q\eta < 1$, then $\lim_{k \rightarrow \infty} \mathbb{E}[\mathfrak{X}(k)] = 0$, which completes the proof. \blacksquare

When the inter-sampling interval τ_* is too large, then $W(k)$ may have negative entries. Consequently, some nodes may mutually repel, and consensus of Algorithm (3) may not be achieved. When repulsive actions exist, expectation consensus, mean-square consensus, and almost sure consensus are not equivalent in general since the Monotone Convergence Theorem does

not apply. Of course, consensus in mean square still implies expectation consensus as consistent with that convergence of random variables in L^r -norm implies convergence in L^s -norm for $r > s \geq 1$. In the subsequent two subsections, mean-square consensus and almost sure consensus/divergence will be separately analyzed.

B. The Mean-Square Consensus Threshold

In this part, we focus on mean-square consensus. First of all, we give a necessary and sufficient mean-square consensus condition in terms of the spectral radius of a matrix that depends on τ_* , \mathbf{G} and q , by studying the spectral property of a linear system. Note that the analysis is carried out on the spectrum restricted to the smallest invariant subspace containing $I - \frac{1}{N}\mathbf{1}\mathbf{1}'$. The condition is then interpreted as the existence of critical threshold on the inter-sampling intervals, below which Algorithm (3) achieves mean-square consensus and above which $\mathfrak{X}(k)$ diverges in mean-square sense for some initial state $x(t_0)$. This translation relies on the equivalence between the stability of a certain matrix and the feasibility of a linear matrix inequality.

Proposition 1: Let Assumptions (A1), (A2), and (A3) hold. Then the following statements are equivalent:

- i) Algorithm (3) achieves mean-square consensus;
- ii) There holds $\rho(\mathbb{E}[W(0) \otimes W(0)](J \otimes J)) < 1$, where

$$J := I - \frac{1}{N}\mathbf{1}\mathbf{1}'; \quad (9)$$

- iii) There exists a matrix $S > 0$ such that

$$\phi(S) := \sum_{i=1}^M \pi_i JW^{(i)}JSJ(W^{(i)})'J < S, \quad (10)$$

where π_i is defined in (7).

Proof: The proof needs the following lemma.

Lemma 2 (Lemma 2 in [30]): For any $G \in \mathbb{C}^{n \times n}$ there exist $G_i \in \mathbb{S}_+^n$, $i = 1, 2, 3, 4$, such that $G = (G_1 - G_2) + (G_3 - G_4)\mathbf{i}$, where $\mathbf{i} = \sqrt{-1}$.

Define the difference between $x(t_k)$ and its average as

$$d(k) := x(t_k) - \frac{1}{N}\mathbf{1}\mathbf{1}'x(t_k). \quad (11)$$

Evidently, $d(k) = Jx(t_k)$. Since

$$\begin{aligned} \mathfrak{X}(k) &= x_{\max}(k) - \frac{1}{N}\mathbf{1}'x(t_k) - \left[x_{\min}(t_k) - \frac{1}{N}\mathbf{1}'x(t_k) \right] \\ &\leq \left| x_{\max}(t_k) - \frac{1}{N}\mathbf{1}'x(t_k) \right| + \left| x_{\min}(t_k) - \frac{1}{N}\mathbf{1}'x(t_k) \right| \\ &\leq \sqrt{2 \sum_{i=1}^N \left[x_i(t_k) - \frac{1}{N}\mathbf{1}'x(t_k) \right]^2} = \sqrt{2} \|d(k)\| \end{aligned} \quad (12)$$

and

$$\begin{aligned} \mathfrak{X}(k) &= N^{-1/2} \sqrt{N(x_{\max}(t_k) - x_{\min}(t_k))^2} \\ &\geq N^{-1/2} \sqrt{\sum_{i=1}^N \left[x_i(t_k) - \frac{1}{N}\mathbf{1}'x(t_k) \right]^2} \\ &= N^{-1/2} \|d(k)\|, \end{aligned} \quad (13)$$

$\lim_{k \rightarrow \infty} \mathbb{E}[\mathfrak{X}^2(k)] = 0$ is equivalent to $\lim_{k \rightarrow \infty} \mathbb{E} \|d(k)\|^2 = 0$. From the Cauchy-Schwarz inequality, $|\mathbb{E}[d_i(k)d_j(k)^*]| \leq \mathbb{E}[|d_i(k)|^2]^{1/2} \mathbb{E}[|d_j(k)|^2]^{1/2}$ holds for any $1 \leq i, j \leq N$, which furthermore implies the equivalence between $\lim_{k \rightarrow \infty} \mathbb{E} \|d(k)\|^2 = 0$ and $\lim_{k \rightarrow \infty} \mathbb{E}[d(k)d(k)^*] = 0$. Thus, to study the mean-square consensus, we only need to focus on whether $\mathbb{E}[d(k)d(k)^*]$ converges to a zero matrix.

Observe that

$$\begin{aligned} d(k) &= JW(k-1)x(t_{k-1}) \\ &= JW(k-1)x(t_{k-1}) - \frac{1}{N}JW(k-1)\mathbf{1}\mathbf{1}'x(t_{k-1}) \\ &= JW(k-1)d(k-1) \end{aligned} \quad (14)$$

holds for $k = 1, 2, \dots$, where the second equality is due to $JW(k)\mathbf{1} = J\mathbf{1} = 0$. It entails

$$\begin{aligned} \mathbb{E}[d(k)d(k)^*] &= \mathbb{E}[JW(k-1)d(k-1) \\ &\quad d(k-1)^*W(k-1)'] \end{aligned}$$

Taking vectorization on both sides yields

$$\begin{aligned} &\text{vec}(\mathbb{E}[d(k)d(k)^*]) \\ &= \mathbb{E}[(JW(k-1) \otimes (JW(k-1))) \text{vec}(d(k-1)d(k-1)^*)] \\ &= (J \otimes J) \mathbb{E}[W(0) \otimes W(0)] \text{vec}(\mathbb{E}[d(k-1)d(k-1)^*]) \\ &= ((J \otimes J) \mathbb{E}[W(0) \otimes W(0)])^k \text{vec}(d(0)d(0)^*) \\ &= ((J \otimes J) \mathbb{E}[W(0) \otimes W(0)])^k (J \otimes J) \text{vec}(x(t_0)x(t_0)^*) \\ &= (J \otimes J) (\mathbb{E}[W(0) \otimes W(0)](J \otimes J))^k \text{vec}(x(t_0)x(t_0)^*), \end{aligned} \quad (15)$$

where the first equality is based on the property $\text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$ for matrices A, B and C of compatible dimensions, and the separation of expectations in the second equality is due to the independence of the random interconnections.

The implications from one statement to the next are provided as follows.

(i) \Rightarrow (ii). If $\rho(\mathbb{E}[W(0) \otimes W(0)](J \otimes J)) \geq 1$, there exist a number λ with $|\lambda| \geq 1$ and a non-zero vector $v \in \mathbb{C}^{N^2}$ corresponding to λ satisfying $\mathbb{E}[W(0) \otimes W(0)](J \otimes J)v = \lambda v$. Let v_1, \dots, v_l be all the eigenvectors corresponding to the eigenvalue 0 of $J \otimes J$. Since $\mathbb{E}[W(0) \otimes W(0)](J \otimes J)v_i = 0$ for any $i = 1, \dots, l$, there holds $v \neq \sum_{i=1}^l a_i v_i$ for any $a_i \in \mathbb{R}$ and $(J \otimes J)v \neq 0$. Therefore

$$\begin{aligned} \lim_{k \rightarrow \infty} (J \otimes J) (\mathbb{E}[W(0) \otimes W(0)](J \otimes J))^k v &= \lim_{k \rightarrow \infty} \lambda^k (J \otimes J) v \\ &\neq 0. \end{aligned} \quad (16)$$

In order to show that mean-square consensus is not achieved for Algorithm (3), it remains to prove that v can be expressed as a linear combination of different initial states. Note that there exist $G \in \mathbb{C}^{N \times N}$ and $G_1, \dots, G_4 \in \mathbb{S}_+^N$ such that $v = \text{vec}(G)$ and $G = G_2 - G_4 + (G_3 - G_1)\mathbf{i}$ by Lemma 2 (the order of G_1, G_2, G_3 and G_4 is immaterial in this lemma). Since each G_i can be expressed as

$$G_i = \sum_{j=1}^N \lambda_j^{(i)} u_j^{(i)} (u_j^{(i)})^*,$$

where $G_i = U^{(i)} \text{diag}\{\lambda_1^{(i)}, \dots, \lambda_N^{(i)}\} (U^{(i)})^*$ with $\lambda_j^{(i)} \in \sigma(G_i)$ and $U^{(i)} =: [u_1^{(i)}, \dots, u_N^{(i)}]$ unitary. Then, we have

$$v = \sum_{i=1}^4 \sum_{j=1}^N -\lambda_j^{(i)} \mathbf{i} \text{vec} \left(u_j^{(i)} \left(u_j^{(i)} \right)^* \right).$$

We see from (16) that mean-square consensus is not achieved for some $x(t_0) = u_j^{(i)}$. Let $w = w_0 + \mathbf{i}w_1$, where $w_0, w_1 \in \mathbb{R}^N$, be such a $u_j^{(i)}$ that with $x(t_0) = w$ Algorithm (3) achieves mean-square divergence. When $x(t_0) = w$, we have

$$\text{vec} \left(\mathbb{E} [d(k) d(k)^*] \right) = (J \otimes J) \left(\mathbb{E} [W(0) \otimes W(0)] (J \otimes J) \right)^k \cdot (w_0 w_0' + w_1 w_1' + \mathbf{i}w_1 w_0' - \mathbf{i}w_0 w_1').$$

From the Cauchy-Schwarz inequality, for any $1 \leq i, j \leq N$, $|\mathbb{E}[d_i(k) d_j(k)^*]| \leq \mathbb{E}[|d_i(k)|^2]^{1/2} \mathbb{E}[|d_j(k)|^2]^{1/2}$, therefore Algorithm (3) achieves mean-square divergence when $x(t_0) = w_0$ or when $x(t_0) = w_1$.

(ii) \Rightarrow (iii). Denote $R := (J \otimes J) \mathbb{E}[W(0) \otimes W(0)] (J \otimes J)$. From (ii),

$$\begin{aligned} \rho(R) &= \rho \left(\mathbb{E} [W(0) \otimes W(0)] (J \otimes J)^2 \right) \\ &= \rho \left(\mathbb{E} [W(0) \otimes W(0)] (J \otimes J) \right) < 1. \end{aligned}$$

Then, $(I - R)^{-1}$ exists and is nonsingular, $(I - R)^{-1} = \sum_{j=0}^{\infty} R^j$. For any given positive definite matrix $V \in \mathbb{R}^{N \times N}$, there corresponds a unique matrix $S \in \mathbb{R}^{N \times N}$ such that

$$\text{vec}(V) = (I - R) \text{vec}(S). \quad (17)$$

Then,

$$\begin{aligned} \text{vec}(V) &= (I - \mathbb{E} [(JW(0)J) \otimes (JW(0)J)]) \text{vec}(S) \\ &= \text{vec}(S - \phi(S)), \end{aligned}$$

where $\phi(\cdot)$ is defined in (10), which implies $S - \phi(S) > 0$ by the one-to-one correspondence of the vectorization operator. The positive definiteness of S follows from

$$\begin{aligned} \text{vec}(S) &= (I - R)^{-1} \text{vec}(V) \\ &= \sum_{i=0}^{\infty} R^i \text{vec}(V) \\ &= \text{vec} \left(\sum_{i=0}^{\infty} \phi^i(V) \right), \end{aligned}$$

implying $S = \sum_{i=0}^{\infty} \phi^i(V) \geq V > 0$, again by the one-to-one correspondence of the vectorization operator.

(iii) \Rightarrow (i). By the hypothesis, there always exists a $\mu \in (0, 1)$ satisfying $\phi(S) < \mu S$. Fix any given $X \in \mathbb{S}_+^N$ and then choose a $c > 0$ satisfying $X \leq cS$. Then, by the linearity and non-decreasing properties of $\phi(X)$ in X over the positive semi-definite cone,

$$\phi^k(X) \leq c \phi^k(S) < c \phi^{k-1}(\mu S) = c \mu \phi^{k-1}(S) < \dots < c \mu^k S$$

holds for all $k \in \mathbb{N}$. It leads to $\lim_{k \rightarrow \infty} \phi^k(X) = 0$, which means

$$\lim_{k \rightarrow \infty} R^k \text{vec}(X) = 0. \quad (18)$$

In light of Lemma 2, for any $G \in \mathbb{R}^{n \times n}$ there exist $X_1, X_2, X_3, X_4 \in \mathbb{S}_+^n$ such that $G = (X_1 - X_2) + (X_3 - X_4)\mathbf{i}$. Then, we see from (18)

$$\begin{aligned} &\lim_{k \rightarrow \infty} R^k \text{vec}(G) \\ &= \lim_{k \rightarrow \infty} R^k (\text{vec}(X_1) - \text{vec}(X_2) + \text{vec}(X_3)\mathbf{i} - \text{vec}(X_4)\mathbf{i}) \\ &= 0. \end{aligned}$$

Since G is arbitrarily chosen, we have $\rho(\mathbb{E}[W(0) \otimes W(0)](J \otimes J)) = \rho(R) < 1$. Then,

$$\begin{aligned} &\lim_{k \rightarrow \infty} \text{vec}(\mathbb{E}[d(k) d(k)^*]) \\ &= (J \otimes J) \lim_{k \rightarrow \infty} (\mathbb{E}[W(0) \otimes W(0)] (J \otimes J))^k \text{vec}(x(t_0) x(t_0)^*) \\ &= 0 \end{aligned}$$

holds for any $x(t_0) \in \mathbb{R}^N$, which means $\lim_{k \rightarrow \infty} \mathbb{E}[d(k) d(k)^*] = 0$. \blacksquare

The following result holds.

Theorem 2: Let Assumptions (A1), (A2), and (A3) hold. Then Algorithm (3) achieves mean-square consensus if and only if $\tau_* \leq \tau_{\dagger}$, where τ_{\dagger} is given by the following quasi-convex optimization problem:

$$\begin{aligned} &\text{minimize}_{\tau} \quad -\tau \\ &\text{subject to} \quad S - \phi(S) > 0, \\ &\quad \quad \quad S > 0, \end{aligned} \quad (19)$$

where ϕ is defined in (10).

Proof: Consider the following optimization problem:

$$\begin{aligned} &\text{minimize}_{\tau} \quad -\tau \\ &\text{subject to} \quad \Psi > 0 \end{aligned} \quad (20a)$$

$$Y, Z > 0, \quad (20b)$$

$$Y - \tau Z \geq 0, \quad (20c)$$

where Ψ is defined in (20), shown at the bottom of the next page. The problem (20) is a generalized eigenvalue problem, which is quasiconvex [31]. Next we shall show the equivalence between (19) and (20).

Necessity: Suppose that there exists a matrix $S > 0$ such that $\phi(S) < S$ holds. First we shall show $\sum_{i=1}^M \pi_i JW^{(i)} JSJ (W^{(i)})' J < JSJ + \mathbf{1}\mathbf{1}'$. Without loss of generality, choose for (v_1, \dots, v_N) an orthonormal basis of \mathbb{R}^N with $v_1 = \frac{1}{N} \mathbf{1}$. Then, any vector $0 \neq x \in \mathbb{R}^n$ can be expressed as $x = \sum_{i=1}^N a_i v_i$ with coefficients a_1, \dots, a_N not all 0. We have

$$x' \phi(S) x = \left(\sum_{i=2}^N a_i v_i \right)' \phi(S) \left(\sum_{i=2}^N a_i v_i \right)$$

and

$$x' (JSJ + \mathbf{1}\mathbf{1}') x = \left(\sum_{i=2}^N a_i v_i \right)' S \left(\sum_{i=2}^N a_i v_i \right) + a_1^2.$$

Since a_1, \dots, a_N are not all 0 and $\phi(S) < S$, there holds $\sum_{i=1}^M \pi_i JW^{(i)} JSJ (W^{(i)})' J < JSJ + \mathbf{1}\mathbf{1}'$. Finally, let $Z = S$ and $Y = \tau_* S$. By Schur complement lemma, we see that (20a) and (20c) hold.

Sufficiency: Suppose that there exist Y and Z such that (20a), (20b), and (20c) hold. According to Schur complement lemma, (20a) is equivalent to

$$JZJ + \mathbf{1}\mathbf{1}' - \sum_{i=1}^M \pi_i (JZ - JL^{(i)}JY) Z^{-1} (JZ - JL^{(i)}JY)^* > 0,$$

which gives

$$\begin{aligned} & JZJ + \mathbf{1}\mathbf{1}' \\ & > \sum_{i=1}^M \pi_i (JZ - JL^{(i)}JY) Z^{-1} (JZ - JL^{(i)}JY)^* \\ & \geq \sum_{i=1}^M \pi_i \left[\tau_* JL^{(i)}JYJ(L^{(i)})'J - JYJ(L^{(i)})'J - JL^{(i)}JYJ \right] \\ & \quad + JZJ \\ & \geq JZJ - \tau_*^{-1}JYJ + \tau_*^{-1}\phi(Y), \end{aligned} \quad (21)$$

where the second inequality holds by substituting Z^{-1} with τ_*Y^{-1} in accordance with (20c). Therefore, it leads to $JYJ + \tau_*\mathbf{1}\mathbf{1}' > \phi(Y)$. Letting $S = JYJ + \tau_*\mathbf{1}\mathbf{1}'$, we have

$$\phi(Y) = \sum_{i=1}^M \pi_i JW^{(i)}J(JYJ + \tau_*\mathbf{1}\mathbf{1}')J(W^{(i)})'J = \phi(S)$$

and then $S > \phi(S)$. In addition, the positive definiteness of S can be seen from the following lemma.

Lemma 3: There holds $JMJ + \epsilon\mathbf{1}\mathbf{1}' > 0$ for all $M > 0$ and $\epsilon > 0$.

Proof: Choose for (v_1, \dots, v_N) an orthonormal basis with $v_1 = \frac{1}{N}\mathbf{1}$. For any nonzero vector $x = \sum_{i=1}^N a_i v_i$,

$$x'(JMJ + \epsilon\mathbf{1}\mathbf{1}')x = \left(\sum_{i=2}^N a_i v_i \right)' M \left(\sum_{i=2}^N a_i v_i \right) + \epsilon a_1^2.$$

Since a_1, \dots, a_N are not all 0 and $M > 0$, we have $x'(JMJ + \epsilon\mathbf{1}\mathbf{1}')x > 0$. \blacksquare

By Proposition 1, Algorithm (3) achieves mean-square consensus if and only if there exists $S > 0$ such that $\phi(S) < S$, which completes the proof. \blacksquare

The optimization problem (19) can be efficiently solved by interior-point algorithms. Many codes, which are based on interior point methods, are available, such as CSDP, SeDuMi, SDPT3, DSDP, SDPA [32]. The computational complexity of solving (19) is in $O(N^3)$ by using, for instance, the algorithm in [33], which is rather efficient for large-scale graphs.

C. Almost Sure Consensus/Divergence

In this part, we focus on the impact of sampling intervals on almost sure consensus/divergence of Algorithm (3). The following theorem gives the relationships between τ_* and almost sure consensus/divergence: almost sure divergence is achieved when τ_* exceeds an upper bound and almost sure consensus is guaranteed when τ_* is sufficiently small. Also note these two boundaries are not equal in general.

Theorem 3: Let Assumptions (A1), (A2), and (A3) hold.

- i) If $\tau_* \leq \tau_{\dagger}$ with τ_{\dagger} given in Theorem 2, Algorithm (3) achieves almost sure consensus.
- ii) If $\tau_* > \tau_{\ddagger}$, where $\tau_{\ddagger} \in \mathbb{R}_+$ is given by

$$\begin{aligned} \tau_{\ddagger} := & \min \left\{ \tau : \log \frac{2N(\tau - 1)}{N - 1} \geq \frac{(1 - q) \log(2N)}{q_* q} \right\} \\ & \vee \min \left\{ \tau : \lambda_{\min} \left(\tau(L^{(i)})'JL^{(i)} - JL^{(i)} - (L^{(i)})'J \right) \geq 0, \right. \\ & \quad \left. \forall L^{(i)} \in \mathcal{L} \right\} \end{aligned}$$

with $q_* := \min\{(1 - q)^{|\mathcal{A}_i| + |\mathcal{A}_j|} : (j, i) \in E\}$, Algorithm (3) diverges almost surely for any initial state $x(t_0) \in \mathbb{R}^N$ except $x(t_0) \perp \mathbf{1}$.

Proof: We start by presenting supporting lemmas.

Lemma 4 (Lemma (5.6.10) in [34]): Let $A \in \mathbb{C}^{n \times n}$ and $\epsilon > 0$ be given. There is a matrix norm $\|\cdot\|_{\dagger}$ such that $\rho(A) \leq \|A\|_{\dagger} \leq \rho(A) + \epsilon$.

Lemma 5 (Borel-Cantelli Lemma): Let $(\mathcal{S}, \mathcal{S}, \mu)$ be a probability space. Assume that events $\mathcal{A}_i \in \mathcal{S}$ for all $i \in \mathbb{N}$. If $\sum_{i=0}^{\infty} \mu(\mathcal{A}_i) < \infty$, then $\mu(\mathcal{A}_i \text{ i.o.}) = 0$, where “ $\mathcal{A}_i \text{ i.o.}$ ” means \mathcal{A}_i occurs infinitely often. In addition, assuming that events $\mathcal{A}_i, i \in \mathbb{N}$, are independent, then $\sum_{i=0}^{\infty} \mu(\mathcal{A}_i) = \infty$ implies $\mu(\mathcal{A}_i \text{ i.o.}) = 1$.

Proof of (i): Note that

$$\begin{aligned} \mathbb{E} \left[\|d(k)\|^2 \right] &= \text{Tr} \left(\mathbb{E} [d(k) d(k)^*] \right) \\ &\leq N^{1/2} \|\text{vec}(\mathbb{E} [d(k) d(k)^*])\|. \end{aligned}$$

The inequality results from the fact that, for any $X := [x_{ij}] \in \mathbb{S}_+^n$, $\|\text{vec}(X)\|^2 = \sum_{i=1}^n \sum_{j=1}^n x_{ij}^2 \geq \sum_{i=1}^n x_{ii}^2 \geq \frac{1}{n} (\sum_{i=1}^n x_{ii})^2 = \frac{1}{n} (\text{Tr}(X))^2$. If $\tau_* < \tau_{\dagger}$ or equivalently $\rho(\mathbb{E}[W(0) \otimes W(0)](J \otimes J)) < 1$ by Theorem 2, there exists a matrix norm $\|\cdot\|_{\dagger}$ such that $\|\mathbb{E}[W(0) \otimes W(0)](J \otimes J)\|_{\dagger} < \lambda < 1$ by Lemma 4. Moreover, by the equivalence of norms on a finite-dimensional vector space, for the two norms $\|\cdot\|$ and $\|\cdot\|_{\dagger}$, there exists a real number $c \in \mathbb{R}_+$ implying $\|X\| \leq c \|X\|_{\dagger}$ for all $X \in \mathbb{R}^{n \times n}$. From the forgoing observations, (15) and the

$$\Psi := \begin{bmatrix} JZJ + \mathbf{1}\mathbf{1}' & \sqrt{\pi_1} (JZ - JL^{(1)}JY) & \cdots & \sqrt{\pi_M} (JZ - JL^{(M)}JY) \\ * & Z & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & Z \end{bmatrix} \quad (20)$$

with $*$'s standing for entries that are the Hermitian conjugates of entries in the upper triangular part.

submultiplicativity of a matrix norm,

$$\begin{aligned}
& \mathbb{E}[\|d(k)\|^2] \\
& \leq \mathbf{N}^{1/2} \|(J \otimes J)(\mathbb{E}[W(0) \otimes W(0)] \\
& \quad \times (J \otimes J)^k \text{vec}(x(t_0)x(t_0)^*))\| \\
& \leq \mathbf{N}^{1/2} c \left\| (\mathbb{E}[W(0) \otimes W(0)](J \otimes J)^k \right\|_{\dagger} \|\text{vec}(x(t_0)x(t_0)^*)\| \\
& \leq \mathbf{N}^{1/2} c \|\mathbb{E}[W(0) \otimes W(0)](J \otimes J)\|_{\dagger}^k \|\text{vec}(x(t_0)x(t_0)^*)\| \\
& < c\lambda^k \mathbf{N}^{1/2} \|\text{vec}(x(t_0)x(t_0)^*)\|.
\end{aligned}$$

Therefore,

$$\sum_{k=0}^{\infty} \mathbb{E}[\|d(k)\|^2] < c(1-\lambda)^{-1} \mathbf{N}^{1/2} \|\text{vec}(x(t_0)x(t_0)^*)\| < \infty, \quad (22)$$

together with Markov's inequality resulting in that,

$$\sum_{k=0}^{\infty} \mathbb{P}(\|d(k)\| > \delta) \leq (1/\delta^2) \sum_{k=0}^{\infty} \mathbb{E}[\|d(k)\|^2] < \infty$$

holds for any $\delta > 0$. According to Lemma 5, $\lim_{k \rightarrow \infty} \|d(k)\| = 0$ almost surely for any initial state $x(t_0) \in \mathbb{R}^{\mathbf{N}}$. Then, $\mathbb{P}(\lim_{k \rightarrow \infty} \mathfrak{X}(k) = 0) = 1$ follows from (12) and (13).

Proof of (ii): Intuitively, the key point to show almost sure divergence is to show the repulsion effect during state update can beat the attraction effect in a long run. To prove (ii), we partition the rest of the proof into three steps. In the first step, we study a best-case attraction effect for any given sequence of G_k 's. In the second step, we construct a specific event, each sample point of which contributes at least a certain amount of repulsion effect. At the third step, the strong law of large numbers is applied to derive the almost sure divergence from the averages of some random events. To meet the i.i.d. assumption of the strong law of large numbers used in Step 3, we introduce auxiliary random variables in the second step.

Step 1. First of all, observe that for all $k \in \mathbb{N}$ and $\omega \in \mathcal{G}^{\mathbf{N}}$

$$\begin{aligned}
\|d(k+1, \omega)\|^2 &= d(k, \omega)^* W(k, \omega)' J J W(k, \omega) d(k, \omega) \\
&\geq \min_{\substack{\|v\|=1, \\ v \perp \mathbf{1}}} \|v^* W(k, \omega)' J W(k, \omega) v\| \|d(k, \omega)\|^2,
\end{aligned}$$

where the inequality holds because $d(k, \omega) \perp \mathbf{1}$. If $\lambda_{\min}(W(k, \omega)' J W(k, \omega) + \frac{1}{\mathbf{N}} \mathbf{1}\mathbf{1}') \geq 1$ for any $k \in \mathbb{N}$ and $\omega \in \mathcal{G}^{\mathbf{N}}$, then $\min_{\substack{\|v\|=1, \\ v \perp \mathbf{1}}} \|v^* W(k, \omega)' J W(k, \omega) v\| \geq 1$, which together with (12) and (13) implies that

$$\mathbb{P}\left(\mathfrak{X}^2(k) \geq \frac{\mathfrak{X}^2(k-1)}{2\mathbf{N}}\right) = 1 \quad (23)$$

holds for all $k \in \mathbb{N}$. Therefore, $\mathfrak{X}(k) > 0$ for all $k \in \mathbb{N}$ provided that $\mathfrak{X}(0) > 0$. The following random variables are well defined:

$$\xi(k) := \frac{\mathfrak{X}^2(k+1)}{\mathfrak{X}^2(k)}, \quad k \in \mathbb{N}. \quad (24)$$

One condition guaranteeing $\lambda_{\min}(W(k, \omega)' J W(k, \omega) + \frac{1}{\mathbf{N}} \mathbf{1}\mathbf{1}') \geq 1$ is established as follows. Note that for any $L^{(i)} \in \mathcal{L}$,

$$\begin{aligned}
& \lambda_{\min}\left(W(k, \omega)' J W(k, \omega) + \frac{1}{\mathbf{N}} \mathbf{1}\mathbf{1}'\right) \\
&= \lambda_{\min}\left(\tau^2 \left(L^{(i)}\right)' J L^{(i)} - \tau J L^{(i)} - \tau \left(L^{(i)}\right)' J + I\right) \\
&= \tau \lambda_{\min}\left(\tau \left(L^{(i)}\right)' J L^{(i)} - J L^{(i)} - \left(L^{(i)}\right)' J\right) + 1.
\end{aligned}$$

Introduce

$$\begin{aligned}
\tau_{\ddagger} &= \min\left\{\tau : \lambda_{\min}\left(\tau \left(L^{(i)}\right)' J L^{(i)} - J L^{(i)} - \left(L^{(i)}\right)' J\right) \right. \\
&\quad \left. \geq 0, \forall L^{(i)} \in \mathcal{L}\right\}. \quad (25)
\end{aligned}$$

A basic but vital observation is that $\tau_{\ddagger} < \infty$, which makes τ_{\ddagger} well defined. To see this, choose a positive number τ_i for any given $L^{(i)} \in \mathcal{L}$. If $\lambda_{\min}(\tau_i (L^{(i)})' J L^{(i)} - J L^{(i)} - (L^{(i)})' J) \geq 0$ for any $L^{(i)}$, we are done. Otherwise let $v \in \mathbb{C}^{\mathbf{N}}$ with $\|v\| = 1$ be any vector such that

$$v^* \left(\tau_i (L^{(i)})' J L^{(i)} - J L^{(i)} - (L^{(i)})' J\right) v < 0. \quad (26)$$

Using the property $\ker((L^{(i)})' J L^{(i)}) = \ker(J L^{(i)})$, we deduce from (26) $v \notin \ker(J L^{(i)})$ and $v^* (L^{(i)})' J L^{(i)} v > 0$. Let τ_i take a new value satisfying

$$\tau_i > \frac{\max_{\|v\|=1} v^* (J L^{(i)} + (L^{(i)})' J) v}{\min_{\|v\|=1, v \notin \ker(J L^{(i)})} v^* (L^{(i)})' J L^{(i)} v}.$$

Then, $\lambda_{\min}(\tau_i (L^{(i)})' J L^{(i)} - J L^{(i)} - (L^{(i)})' J) \geq 0$. Finally, letting $\tau_0 = \max_{1 \leq i \leq n} \tau_i$, we have $\tau_{\ddagger} \leq \tau_0 < \infty$. According to Weyl Theorem (Theorem 4.3.1 in [34]), $\lambda_{\min}(\tau (L^{(i)})' J L^{(i)} - J L^{(i)} - (L^{(i)})' J) \geq 0$ whenever $\tau > \tau_{\ddagger}$ for each $L^{(i)} \in \mathcal{L}$. Recalling that $L(k, \omega) \in \mathcal{L}$, we see that $\tau > \tau_{\ddagger}$ guarantees $\lambda_{\min}(W(k, \omega)' J W(k, \omega) + \frac{1}{\mathbf{N}} \mathbf{1}\mathbf{1}') \geq 1$ for all $k \in \mathbb{N}$ and $\omega \in \mathcal{G}^{\mathbf{N}}$. *Step 2.* First, we propose the following claim.

Claim. There always exist two (random) nodes $i, j \in \mathbf{V}$ at each time k such that $(j, i) \in \mathbf{E}$ and $|x_i(t_k) - x_j(t_k)| \geq \frac{1}{\mathbf{N}-1} \mathfrak{X}(k)$.

To prove the claim, fix any time instance k . Without loss of generality, index all the nodes in the graph such that $x_{\min}(t_k) = x_{i_1}(t_k) \leq x_{i_2}(t_k) \leq \dots \leq x_{i_{\mathbf{N}}}(t_k) = x_{\max}(t_k)$. Then, there at least exists a node $i_n \in \{i_2, \dots, i_{\mathbf{N}}\}$ satisfying $x_{i_n}(t_k) - x_{i_{n-1}}(t_k) \geq \frac{1}{\mathbf{N}-1} \mathfrak{X}(k)$; otherwise $x_{i_{\mathbf{N}}}(t_k) - x_{i_1}(t_k) = \sum_{l=2}^{\mathbf{N}} (x_{i_l}(t_k) - x_{i_{l-1}}(t_k)) < \mathfrak{X}(k)$, reaching a contradiction. If $|x_i(t_k) - x_j(t_k)| < \frac{1}{\mathbf{N}-1} \mathfrak{X}(k)$ for all $(j, i) \in \mathbf{E}$, then neither $(i, j) \notin \mathbf{E}$ nor $(j, i) \notin \mathbf{E}$ for any $i = i_1, \dots, i_{n-1}$ and $j = i_n, \dots, i_{\mathbf{N}}$, since $x_{i_n}(t_k) - x_{i_{n-1}}(t_k) \geq \frac{1}{\mathbf{N}-1} \mathfrak{X}(k)$, contradicting with the hypothesis that \mathbf{G} has a directed spanning tree.

In view of this claim, for each $\omega \in \mathcal{G}^{\mathbf{N}}$, we choose two nodes $i_k(\omega), j_k(\omega) \in \mathbf{V}$ at time k according to $\mathbf{G}_0, \dots, \mathbf{G}_{k-1}$ such that $(j_k(\omega), i_k(\omega)) \in \mathbf{E}$ and $|x_{i_k(\omega)}(t_k) - x_{j_k(\omega)}(t_k)| \geq \frac{1}{\mathbf{N}-1} \mathfrak{X}(k, \omega)$. The dependence of the node selections on specific sample paths gives rise to a challenge in the analysis—without independence, it is difficult to apply the strong law of numbers at Step 3. To get rid of this, we introduce an additional

sequence of random variables. Let $\{z_k\}_{k \in \mathbb{N}}$ be a sequence of i.i.d. random variables defined on $((0, 1)^{\mathbb{N}}, (\mathcal{B}(0, 1))^{\mathbb{N}}, \mathbb{P})$, where $\mathcal{B}(0, 1)$ denotes the Borel algebra on $(0, 1)$, with $z_k(\zeta) = \zeta_k$ for all $\zeta \in (0, 1)^{\mathbb{N}}$ and each z_k uniformly distributed in $(0, 1)$. Let z_0, z_1, \dots and G_0, G_1, \dots be independent. Formally, we are allowed to define a product probability space $(\mathcal{S}, \mathcal{S}, \mu)$ where $\mathcal{S} = \mathcal{G}^{\mathbb{N}} \times (0, 1)^{\mathbb{N}}$, \mathcal{S} is the σ -algebra generated by $\{\mathcal{A} \times \mathcal{B} : \mathcal{A} \in \mathcal{F}, \mathcal{B} \in (\mathcal{B}(0, 1))^{\mathbb{N}}\}$, and μ is the probability measure satisfying $\mu(\mathcal{A} \times \mathcal{B}) = \mathbb{P}(\mathcal{A}) \prod_{l \in \mathbb{N}} \mathbb{P}(\mathcal{B}_l)$. Define $\mathcal{S}_k = \sigma((G_0, z_0), \dots, (G_k, z_k))$. Introduce a sequence of events associated with $i_k(\omega)$, $j_k(\omega)$ and z_k :

$$\mathcal{D}(k) = \left\{ \bigcup_{\omega \in \mathcal{G}^{\mathbb{N}}} (\omega \times \mathcal{B}_k(\omega)) : \mathcal{N}_{i_k(\omega)}(k, \omega) = \{j_k(\omega)\}, \mathcal{N}_{j_k(\omega)}(k, \omega) = \emptyset \right\}$$

with

$$\mathcal{B}_k(\omega) = \left\{ \zeta \in (0, 1)^{\mathbb{N}} : z_k(\zeta) < q_*/(1-q)^{|\mathcal{N}_{i_k(\omega)}| + |\mathcal{N}_{j_k(\omega)}|} \right\}.$$

Since $i_k(\omega), j_k(\omega) \in \mathcal{F}_{k-1}$, one can verify $\mathcal{D}(k) \in \mathcal{S}_k$. If $\tau_* > 1$, for all $(\omega, \zeta) \in \mathcal{D}(k)$ and $k \in \mathbb{N}$,

$$\begin{aligned} \mathfrak{X}(k+1, \omega) &\geq |x_{i_k(\omega)}(t_{k+1}) - x_{j_k(\omega)}(t_{k+1})| \\ &= (\tau_* - 1) |x_{i_k(\omega)}(t_k) - x_{j_k(\omega)}(t_k)| \\ &\geq \frac{\tau_* - 1}{\mathbf{N} - 1} \mathfrak{X}(k, \omega). \end{aligned} \quad (27)$$

Direct calculation yields

$$\mu((\omega, \zeta) \in \mathcal{D}(k)) = \frac{q_* q}{1 - q}. \quad (28)$$

Step 3. Now we define random variables

$$\mathfrak{M}(k) = \begin{cases} \frac{\tau_* - 1}{\mathbf{N} - 1}, & \text{if } (\omega, \zeta) \in \mathcal{D}(k), \\ \frac{1}{2\mathbf{N}}, & \text{otherwise;} \end{cases} \quad (29)$$

which together with (27) leads to

$$\mu\left(\xi_k = \frac{\mathfrak{X}^2(k+1)}{\mathfrak{X}^2(k)} \geq \mathfrak{M}^2(k)\right) = 1.$$

Therefore, $\mu(\prod_{k=0}^t \xi_k = \frac{\mathfrak{X}^2(t+1)}{\mathfrak{X}^2(0)} \geq \prod_{k=0}^t \mathfrak{M}^2(k)) = 1$, which gives

$$\mu\left(\log \mathfrak{X}(t+1) - \log \mathfrak{X}(0) \geq \sum_{k=0}^t \log \mathfrak{M}(k)\right) = 1. \quad (30)$$

Since each node samples the neighbors independently, where the ‘‘independence’’ is in both spatial and temporal sense (Assumption (A2)), therefore, for any $k \in \mathbb{N}$,

$$\mu((\omega, \zeta) \in \mathcal{D}(k) \mid \mathcal{S}_{k-1}) = \frac{p_* p}{1 - p} = \mu((\omega, \zeta) \in \mathcal{D}(k)),$$

indicating that $\mathfrak{M}(k)$ ’s are independent random variables for $\mathcal{D}(0), \dots, \mathcal{D}(k-1) \in \mathcal{S}_{k-1}$. By induction, we eventually have $\{\mathfrak{M}(k)\}_{k \in \mathbb{N}}$ are i.i.d. with the mean computed as

$$\begin{aligned} \mathbb{E}[\log \mathfrak{M}(k)] &= \frac{q_* q}{1 - q} \log \frac{\tau_* - 1}{\mathbf{N} - 1} + \left(1 - \frac{q_* q}{1 - q}\right) \log \frac{1}{2\mathbf{N}} \\ &:= \mathfrak{m}(\tau_*). \end{aligned} \quad (31)$$

Additionally, since $\mathfrak{M}(k)$ ’s have uniformly bounded covariances, Kolmogorov’s strong law of large numbers [35] shows that

$$\mu\left(\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{k=0}^t \log \mathfrak{M}(k) = \mathfrak{m}(\tau_*)\right) = 1, \quad (32)$$

which together with (30) implies that, when $\mathfrak{m}(\tau_*) > 0$, $\mathbb{P}(\liminf_{k \rightarrow \infty} \mathfrak{X}(k) = \infty) = 1$. Notice that $\mathfrak{m}(\tau_*)$ is increasing in τ_* . Defining $\tau_b = \inf\{\tau : \mathfrak{m}(\tau) > 0\}$ and choosing $\tau_* > \tau_b \vee \tau_b := \tau_b$, the conclusion follows. \blacksquare

IV. MARKOVIAN RANDOM NETWORKS

In this section, we continue to investigate sampled-data consensus when each node samples the neighbors by following a Markov chain. The following assumption is imposed.

(A4) *Independently among $(j, i) \in \mathbb{E}$, the random variables $\mathbf{1}_{\{(j, i) \in \mathbb{E}_k\}}$, $k = 0, 1, \dots$, are a binary Markov chain with the failure rate $\mathbb{P}(\mathbf{1}_{\{(j, i) \in \mathbb{E}_{k+1}\}} = 0 \mid \mathbf{1}_{\{(j, i) \in \mathbb{E}_k\}} = 1) := p$ and the recovery rate $\mathbb{P}(\mathbf{1}_{\{(j, i) \in \mathbb{E}_{k+1}\}} = 1 \mid \mathbf{1}_{\{(j, i) \in \mathbb{E}_k\}} = 0) := q$ positive and strictly less than one.*

Note that the techniques developed in this section also apply when $p(i)$ and $q(i)$ vary depending on the node index i .

Under Assumption (A4), $\{L(k)\}_{k \in \mathbb{N}}$ are a sequence of random variables taking values from \mathcal{L} , governed by a finite-state time-homogeneous Markov chain. Compared with random networks, Markovian networks allow each link to have ‘‘memory’’. The transition probability of $\{L(k)\}_{k \in \mathbb{N}}$ is induced from the transition of edges between the ‘‘on’’ state and the ‘‘off’’ state, which is

$$\begin{aligned} \mathbb{P}(L(k) = L^{(j)} \mid L(k-1) = L^{(i)}) &= p^{s_1} (1-p)^{s_2} q^{s_3} (1-q)^{s_4} \\ &:= \pi_{ij}. \end{aligned} \quad (33)$$

where

$$s_1 = \sum_{(n, m) \in \mathbb{E}} \mathbf{1}_{\{l_{m,n}^{(i)} \neq 0, l_{m,n}^{(j)} = 0\}}, s_2 = \sum_{(n, m) \in \mathbb{E}} \mathbf{1}_{\{l_{m,n}^{(i)} \neq 0, l_{m,n}^{(j)} \neq 0\}},$$

$$s_3 = \sum_{(n, m) \in \mathbb{E}} \mathbf{1}_{\{l_{m,n}^{(i)} = 0, l_{m,n}^{(j)} \neq 0\}}$$

and

$$s_4 = \sum_{(n, m) \in \mathbb{E}} \mathbf{1}_{\{l_{m,n}^{(i)} = 0, l_{m,n}^{(j)} = 0\}}.$$

For convenience, we denote $\Pi := [\pi_{ij}]$ as the *transition probability matrix* of $L(k)$. Again, $W(k)$ inherits the same distribution from $L(k)$. The positiveness of the recovery and failure rates in Assumption (A4) makes $L(k)$ an ergodic Markov chain and Π a positive matrix.

A. Conjunction of Various Consensus Metrics

In this part, we show that an analog of Theorem 1 holds over a Markovian random network. From the probabilistic point of view, the difference between the independent model and the Markovian model can be interpreted using a finite permutation argument as follows. Let \mathfrak{q} be a finite permutation from \mathbb{N} onto \mathbb{N} such that $\mathfrak{q}(i) \neq i$ for finitely many i . For any given $\omega \in \mathcal{G}^{\mathbb{N}}$, we define a finite permutation as $(\mathfrak{q}\omega)_i = \omega_{\mathfrak{q}(i)}$ for all $i \in \mathbb{N}$. In the i.i.d. model, the probability measure is invariant with respect to a finite permutation of the sample path, i.e., $\mathbb{P}(\omega \in \mathcal{F}) = \mathbb{P}(\omega \in \mathfrak{q}\mathcal{F})$; while in the Markovian model this property is absent because

of the Markov property. Nevertheless, if $\tau_k = \tau_* \in (0, 1/D_{\max})$ for all k , any finite permutations does not play any key role in whether or not $\mathfrak{X}(k)$ converges in expectation for Algorithm (3). Moreover, Lemma 1 guarantees mean-square consensus, and almost sure consensus regardless of the random network model.

Theorem 4: Let Assumptions (A1), (A3), and (A4) hold. Then expectation consensus, mean-square consensus, and almost sure consensus are achieved for Algorithm (3) if $\tau_* \in (0, 1/D_{\max})$.

Proof: The proof is similar to that of Theorem 1. Here we only provide a sketch. Fix a directed spanning tree G_T of graph G and a sampling time t_k . We choose i_1, \dots, i_N and define $\mathcal{M}_1, \dots, \mathcal{M}_N$ in sequel by the following iterated algorithm: 1) Set i_1 as the root node of G_T , $\mathcal{M}_1 := \{i_1\}$ and $l = 2$; 2) Choose a node $i_l \in V$ such that there exists a node $j \in \mathcal{M}_{l-1}$ satisfying $(j, i_l) \in G_T$ and $i_l \notin \mathcal{M}_{l-1}$; 3) Update $\mathcal{M}_l := \mathcal{M}_{l-1} \cup \{i_l\}$; 4) If $l \leq N$, set $l = l + 1$ and go to step 2); otherwise stop. Consider a sequence of events $\mathcal{E}_2, \dots, \mathcal{E}_N$ where $\mathcal{E}_l := \{L(k+l-1) \in \{L^{(j)} \in \mathcal{L} : l_{i_l i_{l-1}}^{(j)} \neq 0\}\}$ for $l = 2, 3, \dots, N$. If $\mathcal{E}_2, \dots, \mathcal{E}_N$ sequentially occur, similar to the proof of Theorem 1, we see that

$$\mathfrak{X}(k+N-1) \leq \left(1 - \frac{1}{2}\eta^{N-1}\right) \mathfrak{X}(k) \quad (34)$$

where $\eta := (\tau_*) \wedge (1 - \tau_*/D_{\max}) > 0$. Then, we estimate the probability of the sequential occurrence of $\mathcal{E}_2, \dots, \mathcal{E}_3$ by

$$\begin{aligned} & \mathbb{P}\left(\mathbf{1}_{\Gamma_{i=2}^N \mathcal{E}_i} = 1 \mid L(k-1) \in \mathcal{L}\right) \\ &= \mathbb{P}\left(\mathbf{1}_{\mathcal{E}_N} = 1 \mid \mathbf{1}_{\mathcal{E}_{N-1}} = 1\right) \cdots \mathbb{P}\left(\mathbf{1}_{\mathcal{E}_2} = 1 \mid L(k-1) \in \mathcal{L}\right) \\ &\geq \pi^{N-1}, \end{aligned}$$

where $\pi := \min_{1 \leq i, j \leq M} \pi_{ij} > 0$. Therefore,

$$\begin{aligned} & \mathbb{E}[\mathfrak{X}(k+N-1)] \\ &= \sum_{\gamma=0,1} \mathbb{E}\left[\mathbb{P}\left(\mathbf{1}_{\Gamma_{i=2}^N \mathcal{E}_i} = \gamma \mid L(k-1)\right)\right. \\ &\quad \times \mathbb{E}\left[\mathfrak{X}(k+N-1) \mid \mathbf{1}_{\Gamma_{i=2}^N \mathcal{E}_i} = \gamma, L(k-1)\right]\Big] \\ &\leq (1 - \pi^{N-1}) \mathbb{E}[\mathbb{E}[\mathfrak{X}(k) \mid L(k-1)]] \\ &\quad + \pi^{N-1} \left(1 - \frac{1}{2}\eta^{N-1}\right) \mathbb{E}[\mathbb{E}[\mathfrak{X}(k) \mid L(k-1)]] \\ &= \left(1 - \frac{1}{2}(\pi\eta)^{N-1}\right) \mathbb{E}[\mathfrak{X}(k)], \end{aligned}$$

which implies $\lim_{k \rightarrow \infty} \mathbb{E}[\mathfrak{X}(k)] = 0$ and therefore consensus in expectation is achieved. Finally, the conclusion follows from Lemma 1. \blacksquare

Remark 3: The assumption of a uniform inter-sampling interval τ_k simplifies the notations used in Theorems 1 and 4. It should be emphasized that the techniques used in the proof of Theorems 1 and 4 also apply to the non-uniform inter-sampling interval case. To make the conclusion hold, we require $\lim_{k \rightarrow \infty} \log(\mathbb{E}[\mathfrak{X}(k(N-1))]/\mathbb{E}[\mathfrak{X}(0)]) = -\infty$, which can be guaranteed by $\sum_{k=0}^{\infty} \prod_{j=0}^{N-2} \eta_{k+j} = \infty$ with $\eta_k = (\tau_k) \wedge (1 - (N-1)\tau_k)$ for $k \in \mathbb{N}$. This is seen from (8) and the fact that, for a sequence $\{a_k\}_{k \in \mathbb{N}}$ with $a_k \in [0, 1)$, $\sum_{k=1}^{\infty} a_k = \infty$ if and only if $\prod_{k=1}^{\infty} (1 - a_k) = 0$ [36].

B. The Mean-Square Consensus Threshold

Now, we are interested in establishing a necessary and sufficient condition on τ_* for mean-square consensus of Algorithm (3). We first present an implicit condition in terms of the spectral radius of a certain matrix. Then, this stability condition is translated to a threshold on τ_* . The analysis in this section is based on the techniques using in the proof of Proposition 1 as well as the tools from the theory of Markov jump linear systems.

Proposition 2: Let Assumptions (A1), (A3), and (A4) hold and, for each $(j, i) \in E$, let $\mathbf{1}_{\{(j,i) \in E_0\}}$ start at any initial distribution. Then the following statements are equivalent:

- i) Algorithm (3) achieves mean-square consensus;
- ii) There holds $\rho(\Gamma\Theta) < 1$, where

$$\Gamma := \text{diag}\left(W^{(1)} \otimes W^{(1)}, \dots, W^{(M)} \otimes W^{(M)}\right) \quad (35)$$

and

$$\Theta := \Pi' \otimes (J \otimes J) \quad (36)$$

with J defined in (9) and Π defined in (33);

- iii) There exist matrices $S_1, \dots, S_M > 0$ such that

$$\varphi_j(S) := \sum_{i=1}^M \pi_{ij} J W^{(j)} J S_i J \left(W^{(j)}\right)' J < S_j \quad (37)$$

holds for all $1 \leq j \leq M$, where $S := (S_1, \dots, S_M)$.

Proof: Recall $d(k)$ from (11). Obviously, (12) to (14) still hold. In what follows, we consider a linear space over the complex field \mathbb{C} : $\mathbb{H}^M := \{(M_1, \dots, M_M) : M_i \in \mathbb{C}^{N \times N}, i = 1, \dots, M\}$ and a convex cone in \mathbb{H}^M : $\mathbb{H}_+^M := \{(G_1, \dots, G_M) : G_i \in \mathbb{S}_N^+, i = 1, \dots, M\}$. Define

$$H(k) := \left(d(k) d(k)^* \mathbf{1}_{\{L(k)=L^{(1)}\}}, \dots, d(k) d(k)^* \mathbf{1}_{\{L(k)=L^{(M)}\}}\right).$$

Since

$$\mathbb{E}[d(k) d(k)^*] = \mathbb{E}[H(k)] \underbrace{[I_N, \dots, I_N]'}_{M \text{ times}},$$

it follows from (12) and (13) that $\lim_{k \rightarrow \infty} \mathbb{E}[\mathfrak{X}^2(k)] = 0$ is equivalent to $\lim_{k \rightarrow \infty} \mathbb{E}[H(k)] = 0$. Taking vectorization on both sides of $\mathbb{E}[H(k)]$ gives

$$\begin{aligned} & \text{vec}(\mathbb{E}[H(k)]) \\ &= \Pi' \otimes \begin{bmatrix} (JW^{(1)}) \otimes (JW^{(1)}) & & \\ & \ddots & \\ & & (JW^{(M)}) \otimes (JW^{(M)}) \end{bmatrix} \end{aligned}$$

$$\cdot \begin{bmatrix} \text{vec}\left(\mathbb{E}\left[d(k-1) d(k-1)^* \mathbf{1}_{\{L(k-1)=L^{(1)}\}}\right]\right) \\ \vdots \\ \text{vec}\left(\mathbb{E}\left[d(k-1) d(k-1)^* \mathbf{1}_{\{L(k-1)=L^{(M)}\}}\right]\right) \end{bmatrix}$$

$$= ((\Pi' \otimes I_{N^2}) (I_M \otimes (J \otimes J)) \Gamma) \text{vec}(\mathbb{E}[H(k-1)])$$

$$= ((\Pi' \otimes I_{N^2}) (I_M \otimes (J \otimes J)) \Gamma)^k \text{vec}(\mathbb{E}[H(0)])$$

$$= ((I_M \otimes (J \otimes J)) (\Pi' \otimes I_{N^2}) \Gamma)^k (I_M \otimes (J \otimes J)) \text{vec}(\mathbb{E}[H(0)])$$

$$= (I_M \otimes (J \otimes J)) ((\Pi' \otimes I_{N^2}) \Gamma (I_M \otimes (J \otimes J)))^k$$

$$\cdot \text{vec}(\mathbb{E}[\mathfrak{h}(x(t_0), L(0))]),$$

where

$$\mathfrak{h}(x(t_0), L(0)) := \left(x(t_0) x(t_0)^* \mathbf{1}_{\{L(0)=L^{(1)}\}}, \dots, x(t_0) x(t_0)^* \mathbf{1}_{\{L(0)=L^{(M)}\}} \right).$$

The fourth equality holds because $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ for matrices A, B, C and D of compatible dimensions. In addition, $\rho((P' \otimes I_{N^2})\Gamma(I_M \otimes (J \otimes J))) = \rho(\Gamma\Theta)$.

It follows from Lemma 2 that for any $H \in \mathbb{H}^M$, there exist $H_1, \dots, H_4 \in \mathbb{H}_+^M$ such that $H = (H_1 - H_2) + (H_3 - H_4)\mathbf{i}$. Moreover, for each $H_i = (G_1^{(i)}, \dots, G_M^{(i)})$,

$$H_i = \sum_{m=1}^M \sum_{n=1}^N \lambda_{m,n}^{(i)} \mathfrak{h}(u_{m,n}^{(i)}, L^{(m)})$$

with $U_m^{(i)} =: [u_{m,1}^{(i)}, \dots, u_{m,N}^{(i)}]$ unitary and $G_m^{(i)} = U_m^{(i)} \text{diag}\{\lambda_{m,1}^{(i)}, \dots, \lambda_{m,N}^{(i)}\} (U_m^{(i)})^*$ for $m = 1, \dots, M$, which means that, for any $H \in \mathbb{H}^M$, $\text{vec}(H)$ can be expressed as a linear combination of different initial states.

The rest of the proof follows from the arguments used in the proof of Theorem 1 and the mean-square stability theory of Markov jump linear systems ([37], Theorem 3.9). \square

The following theorem holds, analogous to Theorem 2. The proof is omitted.

Theorem 5: Let Assumptions (A1), (A3), and (A4) hold. Then Algorithm (3) achieves mean-square consensus if and only if $\tau_* \leq \tilde{\tau}_\dagger$, where $\tilde{\tau}_\dagger$ is given by the following quasi-convex optimization problem:

$$\begin{aligned} & \text{minimize}_\tau \quad -\tau \\ & \text{subject to} \quad S_j - \varphi_j(S) > 0, \\ & \quad \quad \quad S_j > 0, j = 1, \dots, M, \end{aligned} \quad (38)$$

where $S := (S_1, \dots, S_M)$ and φ_j is defined in (37).

To solve the optimization problem (38), it has a complexity $O(N^3 2^{|\mathbb{E}|})$ by using interior-point algorithms [33]. In general, solving (38) is computationally intractable for large-scale networks. More efficient approximate algorithms will be left as a future work.

C. Almost Sure Consensus/Divergence

In this part, we explore the almost sure consensus/divergence condition for Algorithm (3) over Markovian random networks. The following theorem exhibits a correlation between τ_* and the asymptotic behavior of every sample path, that is, a small τ_* guarantees almost sure consensus while a large τ_* tends to result in almost sure divergence. In the following theorem, the almost sure divergence analysis is restricted to complete graphs. The assumption of complete graph simplifies the analysis. We believe that the techniques used in developing almost sure divergence result in Theorem 6 can also deal with generic directed graphs. We plan to remove this restriction and consider more general graphs in future work.

Theorem 6: Let Assumptions (A1), (A3), and (A4) hold and, for each $(j, i) \in \mathbb{E}$. Let $\mathbf{1}_{\{(j,i) \in \mathbb{E}_0\}}$ start at any initial distribution.

- i) If $\tau_* \leq \tilde{\tau}_\dagger$, then Algorithm (3) achieves almost sure consensus.

- ii) If G is a complete graph and $\tau_* > \tilde{\tau}_\dagger$, where $\tilde{\tau}_\dagger$ is given by

$$\begin{aligned} \tilde{\tau}_\dagger := & \min \left\{ \tau : \tilde{q}_* \log \frac{2N(\tau-1)}{N-1} + \log \frac{1}{2N} \geq 0 \right\} \\ & \vee \min \left\{ \tau : \lambda_{\min} \left(\tau (L^{(i)})' J L^{(i)} - J L^{(i)} - (L^{(i)})' J \right) \geq 0, \right. \\ & \quad \left. \forall L^{(i)} \in \mathcal{L} \right\} \end{aligned}$$

with $\tilde{q}_* := (1-p) \wedge q$, then Algorithm (3) diverges almost surely for any initial state $x(t_0) \in \mathbb{R}^N$ except $x(t_0) \perp \mathbf{1}$.

Proof: To show (i), note that

$$\begin{aligned} \mathbb{E} \left[\|d(k)\|^2 \right] &= \text{Tr} \left(\mathbb{E} [H(k)] \underbrace{[I_N, \dots, I_N]'}_{M \text{ times}} \right) \\ &\leq (MN)^{1/2} \|\text{vec}(\mathbb{E} [H(k)])\|. \end{aligned}$$

When $\rho(\Gamma\Theta) < 1$, by using the same argument as in (22), we know that $\sum_{k=0}^{\infty} \mathbb{E} [\|d(k)\|^2] < \infty$ holds for any initial state $x(t_0) \in \mathbb{R}^N$ and any initial distribution of $\mathbf{1}_{\{(j,i) \in \mathbb{E}_0\}}$ for each $(j, i) \in \mathbb{E}$. By Markov's inequality and Lemma 5, $\lim_{k \rightarrow \infty} \|d(k)\| = 0$ almost surely.

Next, we shall prove (ii). Similar to the proof of Theorem 3, the analysis is divided into three steps. *Step 1.* Suppose $\tau > \tau_\ddagger$, where $\tau_\ddagger \in \mathbb{R}_+$ is defined in (25). Adopting the analysis used in the proof of Theorem 3, we define $\xi(k)$ as in (24) and conclude that (23) holds for all $k \in \mathbb{N}$.

Step 2. In the first place, for each $\omega \in \mathcal{G}^{\mathbb{N}}$, we choose two (random) nodes $i_k(\omega), j_k(\omega) \in \mathbb{V}$ at time k such that $|x_{i_k(\omega)}(t_k) - x_{j_k(\omega)}(t_k)| = \mathfrak{X}(k, \omega)$. Let $\{z_k\}_{k \in \mathbb{N}}$ be a sequence of i.i.d. random variables defined on $((0, 1)^{\mathbb{N}}, (\mathcal{B}(0, 1))^{\mathbb{N}}, \mathfrak{l})$ with $z_k(\zeta) = \zeta_k$ for all $\zeta \in (0, 1)^{\mathbb{N}}$ and each z_k uniformly distributed in $(0, 1)$, and let z_0, z_1, \dots and G_0, G_1, \dots be independent. Formally, we are allowed to define a product probability space $(\mathcal{S}, \mathcal{S}, \mu)$ with μ the product probability measure satisfying $\mu(\mathcal{A} \times \mathcal{B}) = \mathbb{P}(\mathcal{A})\mathfrak{l}(\mathcal{B})$ for any $\mathcal{A} \in \mathcal{F}$ and $\mathcal{B} \in (\mathcal{B}(0, 1))^{\mathbb{N}}$. Define $\mathcal{S}_k = \sigma((G_0, z_0), \dots, (G_k, z_k))$. Introduce a sequence of events associated with $i_k(\omega), j_k(\omega)$ and z_k :

$$\mathcal{D}(k) = \{\cup_{\omega \in \mathcal{G}^{\mathbb{N}}} (\omega \times \mathcal{B}_k(\omega)) : (j_k(\omega), i_k(\omega)) \in \mathbb{E}_k\}$$

with $\mathcal{B}_k(\omega)$ given by

$$\mathcal{B}_k(\omega) = \begin{cases} \left\{ \zeta \in (0, 1)^{\mathbb{N}} : z_k(\zeta) < ((1-p) \wedge q) / (1-p) \right\}, \\ \text{if } (j_k(\omega), i_k(\omega)) \in \mathbb{E}_{k-1}(\omega); \\ \left\{ \zeta \in (0, 1)^{\mathbb{N}} : z_k(\zeta) < ((1-p) \wedge q) / q \right\}, \\ \text{if } (j_k(\omega), i_k(\omega)) \notin \mathbb{E}_{k-1}(\omega). \end{cases}$$

We have the following claim due to a complete underlying graph G .

Claim. Suppose $\tau_* > 1$. There holds $\mathfrak{X}(k+1, \omega) \geq (\tau_* - 1)\mathfrak{X}(k, \omega)$ for all $(\omega, \zeta) \in \mathcal{D}(k)$ and $k \in \mathbb{N}$.

Step 3. We define random variables

$$\mathfrak{M}(k) = \begin{cases} \frac{\tau_* - 1}{N - 1}, & \text{if } (\omega, \zeta) \in \mathcal{D}(k); \\ \frac{1}{2N}, & \text{otherwise.} \end{cases} \quad (39)$$

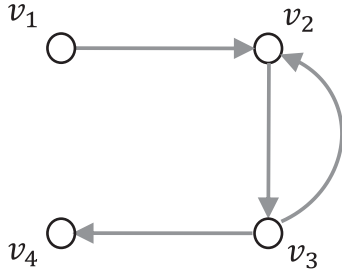


Fig. 1. A underlying graph G consisting of four nodes.

Similar to the proof of Theorem 3, for any $t \in \mathbb{N}$,

$$\mu \left(\log \mathfrak{X}(t+1) - \log \mathfrak{X}(0) \geq \sum_{k=0}^t \log \mathfrak{M}(k) \right) = 1. \quad (40)$$

Since each node independently samples among its neighbors, for any $k \in \mathbb{N}$,

$$\begin{aligned} \mu \left((\omega, \zeta) \in \mathcal{D}(k) \mid \mathcal{S}_{k-1} \right) &= (1-p) \wedge q = \mu((\omega, \zeta) \in \mathcal{D}(k)) \\ &:= \tilde{q}_*, \end{aligned}$$

indicating that $\mathfrak{M}(k)$'s are independent random variables for $\mathcal{D}(0), \dots, \mathcal{D}(k-1) \in \mathcal{S}_{k-1}$. By induction, we eventually have $\{\mathfrak{M}(k)\}_{k \in \mathbb{N}}$ are i.i.d. with the mean computed as

$$\mathbb{E}[\log \mathfrak{M}(k)] \geq \tilde{q}_* \log \frac{\tau_* - 1}{\mathbf{N} - 1} + (1 - \tilde{q}_*) \log \frac{1}{2\mathbf{N}} := \tilde{m}(\tau_*) \quad (41)$$

In addition, since $\mathfrak{M}(k)$'s have uniformly bounded covariances, again by Kolmogorov strong law of large numbers [35],

$$\mu \left(\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{k=0}^t \log \mathfrak{M}(k) = \mathbb{E}[\log \mathfrak{M}(k)] \right) = 1, \quad (42)$$

together with (30) implying that, when $\tilde{m}(\tau_*) > 0$, $\mathbb{P}(\liminf_{k \rightarrow \infty} \mathfrak{X}(k) = \infty) = 1$. Notice that $\tilde{m}(\tau_*)$ is increasing in τ_* . The proof is completed by defining $\tilde{\tau}_b = \inf\{\tau : \tilde{m}(\tau) > 0\}$ and choosing $\tau_* > \tau_b \vee \tilde{\tau}_b := \tilde{\tau}_b$. ■

V. NUMERICAL EXAMPLES

In this section, we provide numerical examples to validate the theoretical results. We first illustrate the existence of the threshold on τ_* , which decides the mean-square convergence or divergence (see Theorems 2 and 5). We then discuss and illustrate how the threshold depends on the number of nodes with cyclic underlying graphs, for i.i.d. and Markovian network models, respectively.

A. Mean-Square Convergence vs. Divergence

We consider a network consisting of $\mathbf{N} = 4$ nodes indexed by $\mathbf{V} = \{v_1, v_2, v_3, v_4\}$. Let $\mathbf{E} = \{(v_1, v_2), (v_2, v_3), (v_3, v_2), (v_3, v_4)\}$. The underlying graph $G = (\mathbf{V}, \mathbf{E})$ is illustrated in Fig. 1. Evidently, G has a directed spanning tree. The random variables $\mathbf{1}_{\{(j,i) \in \mathbf{E}_k\}}$, $(j, i) \in \mathbf{E}$ and $k \in \mathbb{N}$, are i.i.d. Bernoulli ones with $\mathbb{P}((j, i) \in \mathbf{E}_k) = 0.5$. We choose a uniform inter-sampling interval, i.e., $\tau_k = \tau_*$ for all k . Then Algorithm (3) is given by

$$x(t_{k+1}) = [I - \tau_* L(k)] x(t_k). \quad (43)$$

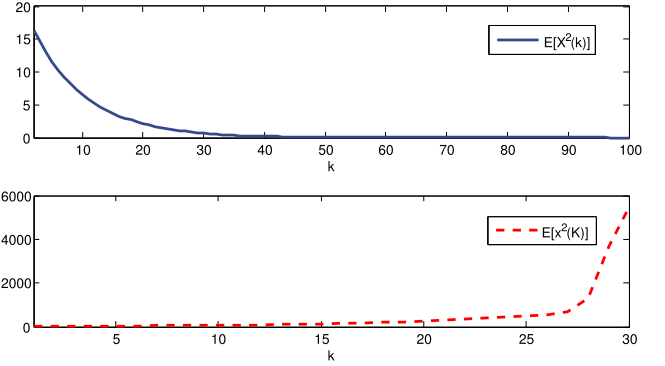


Fig. 2. The evolutions of $\mathbb{E}[\mathfrak{X}^2(k)]$ for different sample periods over an independent random network with $q = 0.5$. In the upper figure, $\mathbb{E}[\mathfrak{X}^2(k)]$ converges to 0 as $k \rightarrow \infty$ when $\tau_* = 1$. In the bottom figure, $\mathbb{E}[\mathfrak{X}^2(k)]$ diverges as $k \rightarrow \infty$ when $\tau_* = 1.14$.

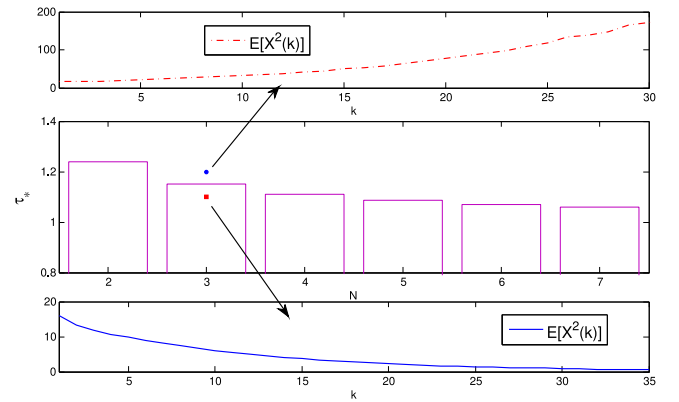


Fig. 3. The relationship between \mathbf{N} and τ_b for a cycle underlying graph over independent random networks ($q = 0.6$). For $\mathbf{N} = 3$, two sample periods, $\tau_* = 1.1$ (the red block) and $\tau_* = 1.2$ (the blue dot), are chosen to illustrate the convergence and divergence behaviors of $\mathbb{E}[\mathfrak{X}^2(k)]$ respectively.

According to Theorem 2, we compute that system (43) achieves consensus in mean square if and only if $\tau_* \leq 1.07$. We next illustrate this conclusion using simulations. Choose $x(t_0) = [5 \ 2 \ 1 \ 1]'$, run 10^6 Monte Carlo simulations, and then use the average as an approximation of $\mathbb{E}[\mathfrak{X}^2(k)]$. Fig. 2 illustrates that $\mathbb{E}[\mathfrak{X}^2(k)]$ converges to 0 as k becomes large when $\tau_* = 1$ and diverges as k increases when $\tau_* = 1.14$, validating the conclusion of Theorem 2.

B. Independent and Markovian Random Graphs

Consider a network of \mathbf{N} nodes connected by a directed cycle graph as the underlying graph. We choose $q = 0.6$ for the i.i.d. model. The relationship between \mathbf{N} and the critical sampling interval τ_b is plotted in Fig. 3. As for the Markovian model, we choose $p = 0.4$ and $q = 0.7$. The relationship between \mathbf{N} and τ_b is plotted in Fig. 4. Note that each $\mathbf{1}_{\{(j,i) \in \mathbf{E}_k\}}$ has a stationary distribution identical to the distribution of $\mathbf{1}_{\{(j,i) \in \mathbf{E}_k\}}$ in the independent model.

VI. CONCLUSION

In this paper, we have considered sampled-data consensus problem over random networks. We first defined three types of

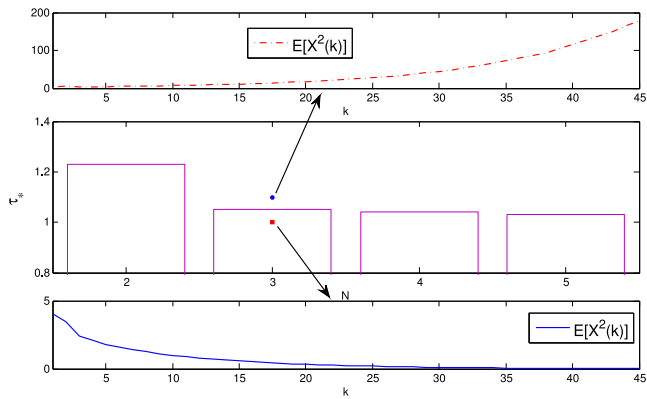


Fig. 4. The relationship between \mathbf{N} and $\bar{\tau}_\dagger$ for a cycle underlying graph over Markovian random networks ($p = 0.6, q = 0.9$). For $\mathbf{N} = 3$, two sample periods, $\tau_* = 1.0$ (the red block) and $\tau_* = 1.1$ (the blue dot), are chosen to illustrate the convergence and divergence behaviors of $\mathbb{E}[\mathfrak{X}^2(k)]$ respectively.

TABLE I
SUMMARY OF THE CONSENSUS RESULTS FOR INDEPENDENT AND MARKOVIAN RANDOM NETWORKS

	Mean-square Consensus	Mean-square Divergence	Almost Sure Consensus	Almost Sure Divergence
Independent Sampling	$\tau_* \leq \tau_\dagger$	$\tau_* > \tau_\dagger$	$\tau_* \leq \tau_\dagger$	$\tau_* > \tau_\dagger$
Markovian Sampling	$\tau_* \leq \bar{\tau}_\dagger$	$\tau_* > \bar{\tau}_\dagger$	$\tau_* \leq \bar{\tau}_\dagger$	$\tau_* > \bar{\tau}_\dagger$

random consensus notions and established the equivalence of these consensus notions provided a sufficient condition in terms of the inter-sampling interval and the maximum degree of the underlying graph. Under this condition, three types of consensus were shown to be simultaneously achieved if the underlying graph contains a directed spanning tree. Both independent and Markovian random networks are then considered. In either network model, necessary and sufficient conditions for mean-square consensus were derived in terms of the inter-sampling interval. Sufficient conditions for almost sure convergence/divergence were also provided, respectively, in terms of the size of inter-sampling interval. The results for the independent and Markovian random networks are summarized in Table I. It is surprising that the phase transition phenomenon of mean-square consensus exists for both types of random networks.

REFERENCES

- [1] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, 2003.
- [2] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, 2004.
- [3] Z. Lin, B. Francis, and M. Maggiore, "State agreement for continuous-time coupled nonlinear systems," *SIAM J. Control Optim.*, vol. 46, no. 1, pp. 288–307, 2007.
- [4] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, pp. 65–78, 2004.
- [5] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, 2005.
- [6] Y. Cao and W. Ren, "Multi-vehicle coordination for double-integrator dynamics under fixed undirected/directed interaction in a sampled-data setting," *International J. Robust Nonlinear Control*, vol. 20, pp. 987–1000, 2010.
- [7] Y. Zhang and Y.-P. Tian, "Consensus of data-sampled multi-agent systems with random communication delay and packet loss," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 939–943, 2010.
- [8] Y. Gao and L. Wang, "Sampled-data based consensus of continuous-time multi-agent systems with time-varying topology," *IEEE Trans. Autom. Control*, vol. 56, no. 5, pp. 1226–1231, 2011.
- [9] W. Yu, W. X. Zheng, G. Chen, W. Ren, and J. Cao, "Second-order consensus in multi-agent dynamical systems with sampled position data," *Automatica*, vol. 47, no. 7, pp. 1496–1503, 2011.
- [10] F. Xiao and T. Chen, "Sampled-data consensus for multiple double integrators with arbitrary sampling," *IEEE Trans. Autom. Control*, vol. 57, no. 12, pp. 3230–3235, 2012.
- [11] S. Kar and J. M. Moura, "Sensor networks with random links: Topology design for distributed consensus," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3315–3326, 2008.
- [12] S. Kar and J. M. Moura, "Distributed consensus algorithms in sensor networks: Quantized data and random link failures," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1383–1400, 2009.
- [13] S. S. Pereira and A. Pagés-Zamora, "Mean square convergence of consensus algorithms in random WSNs," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2866–2874, 2010.
- [14] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 634–649, 2008.
- [15] F. Fagnani and S. Zampieri, "Average consensus with packet drop communication," *SIAM J. Control Optim.*, vol. 48, no. 1, pp. 102–133, 2009.
- [16] A. Tabbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 225–230, 2010.
- [17] I. Matei, J. S. Baras, and C. Somarakis, "Convergence results for the linear consensus problem under markovian random graphs," *SIAM J. Control Optim.*, vol. 51, no. 2, pp. 1574–1591, 2013.
- [18] Q. Song, G. Chen, and D. W. Ho, "On the equivalence and condition of different consensus over a random network generated by i.i.d. stochastic matrices," *IEEE Trans. Autom. Control*, vol. 56, no. 5, pp. 1203–1207, 2011.
- [19] Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1876–1872, 2005.
- [20] C. W. Wu, "Synchronization and convergence of linear dynamics in random directed networks," *IEEE Trans. Autom. Control*, vol. 51, no. 7, pp. 1207–1210, 2006.
- [21] M. Porfiri and D. Stilwell, "Consensus seeking over random weighted directed graphs," *IEEE Trans. Autom. Control*, vol. 51, no. 7, pp. 1767–1773, 2007.
- [22] A. Tabbaz-Salehi and A. Jadbabaie, "A necessary and sufficient condition for consensus over random networks," *IEEE Trans. Autom. Control*, vol. 53, no. 3, pp. 791–795, 2008.
- [23] A. Tabbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 225–230, 2010.
- [24] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Math. Oper. Res.*, vol. 38, no. 1, pp. 1–27, 2013.
- [25] G. Shi, M. Johansson, and K. H. Johansson, "How agreement and disagreement evolve over random dynamic networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 6, pp. 1061–1071, 2013.
- [26] F. Xiao and L. Wang, "Asynchronous consensus in continuous-time multi-agent systems with switching topology and time-varying delays," *IEEE Trans. Autom. Control*, vol. 53, no. 8, pp. 1804–1816, 2008.
- [27] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge, U.K.: Cambridge Univ. Press, 1952.
- [28] R. Durrett, *Probability: Theory and Examples*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [29] Q. Song, G. Chen, and D. Ho, "On the equivalence and condition of different consensus over a random network generated by i.i.d. stochastic matrices," *IEEE Trans. Autom. Control*, vol. 56, no. 5, pp. 1203–1207, 2011.
- [30] O. L. V. Costa and M. Fragoso, "Comments on 'Stochastic stability of jump linear systems'," *IEEE Trans. Autom. Control*, vol. 49, no. 8, pp. 1414–1416, 2004.
- [31] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.

- [32] H. D. Mittelmann, "An independent benchmarking of SDP and SOCP solvers," *Math. Program.*, vol. 95, no. 2, pp. 407–430, 2003.
- [33] P. Gahinet, A. Nemirovskii, A. J. Laub, and M. Chilali, *LMI Control Toolbox User's Guide*. Natick, MA, USA: The Mathworks Partner Series, 1995.
- [34] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
- [35] W. Feller, *An Introduction to Probability Theory and its Applications Vol. I*. New York, NY, USA: Wiley, 1950.
- [36] W. Rudin, *Real and Complex Analysis*. New York, NY, USA: Tata McGraw-Hill Education, 1987.
- [37] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-time Markov Jump Linear Systems*. New York, NY, USA: Springer Science & Business Media, 2006.



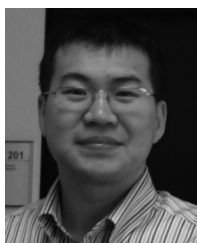
Junfeng Wu received the B.Eng. degree from the Department of Automatic Control, Zhejiang University, Hangzhou, China, in 2009 and the Ph.D. degree in electrical and computer engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2013. From September to December 2013, he was a Research Associate in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong. He is currently a Postdoctoral Researcher at ACCESS (Autonomic Complex Communication nEt-

works, Signals and Systems) Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, Sweden. His research interests include networked control systems, state estimation, and wireless sensor networks, multi-agent systems. He received the Guan Zhao-Zhi Best Paper Award at the 34th Chinese Control Conference in 2015.



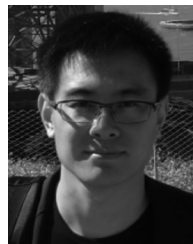
Ziyang Meng received his Bachelor degree with honors from Huazhong University of Science & Technology, Wuhan, China, in 2006, and Ph.D. degree from Tsinghua University, Beijing, China, in 2010. He was an exchange Ph.D. student at Utah State University, Logan, USA from September 2008 to September 2009. From 2010 to 2015, he held postdoc, researcher, and Humboldt research fellow positions at, respectively, Shanghai Jiao Tong University, Shanghai, China, KTH Royal Institute of Technology, Stockholm, Sweden, and Technical University

of Munich, Munich, Germany. He joined Department of Precision Instrument, Tsinghua University, China as an Associate Professor since Sept. 2015. His research interests include multiagent systems, small satellite systems, distributed control and optimization, nonlinear systems and information fusion techniques. He was selected to the national "1000-Youth Talent Program" of China in 2015.



Tao Yang received the M.S. degree with distinction in control engineering from City University, London in 2004, and the Ph.D. degree in electrical engineering from Washington State University in 2012. Between August 2012 and August 2014, he was an ACCESS Post-Doctoral Researcher with the ACCESS Linnaeus Centre, Royal Institute of Technology, Sweden. He is currently an Assistant Professor at the Department of Electrical Engineering, University of North Texas (UNT). Before joining in the UNT, he was a Scientist with Energy & Environmental Di-

rectorate, Pacific Northwest National Laboratory. His research interests include distributed control and optimization in power systems, Cyber Physical Systems, networked control systems, and multi-agent systems. He is a member of the IEEE Control Systems Society Technical Committee on Smart Grids and Technical Committee on Nonlinear Systems and Control.



Guodong Shi received his B.Sc. degree in mathematics and applied mathematics from School of Mathematics, Shandong University, Jinan, China, in July 2005, and his Ph.D. in systems theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in July 2010, respectively. From August 2010 to April 2014 he was a Postdoctoral Researcher at the ACCESS Linnaeus Centre, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. He held a visiting position from Oct. 2013 to Dec.

2013 at the School of Information and Engineering Technology, University of New South Wales, Canberra, Australia. Since May 2014 he has been with the Research School of Engineering, College of Engineering and Computer Science, The Australian National University, Canberra, Australia, as a Lecturer and Future Engineering Research Leadership Fellow. Dr. Shi was selected in the Triennial IFAC Young Author Prize Finalist in 2011, and was a co-recipient of Best Paper Award in Control Theory from the World Congress on Intelligent Control and Automation in 2014 as well as the Guan Zhao-Zhi Best Paper Award from the Chinese Control Conference in 2015. His current research interests include distributed control systems, quantum networking and decisions, and social opinion dynamics.



Karl Henrik Johansson (SM'08–F'12) is Director of the ACCESS Linnaeus Centre and Professor at the School of Electrical Engineering, KTH Royal Institute of Technology, Sweden. He is a Wallenberg Scholar and has held a Senior Researcher Position with the Swedish Research Council. He also heads the Stockholm Strategic Research Area ICT The Next Generation. He received M.Sc. and Ph.D. degrees in electrical engineering from Lund University. He has held visiting positions at UC Berkeley, California Institute of Technology, Nanyang Technological University, and Institute of Advanced Studies Hong Kong University of Science

and Technology. His research interests are in networked control systems, cyber-physical systems, and applications in transportation, energy, and automation systems. He has been a member of the IEEE Control Systems Society Board of Governors and the Chair of the IFAC Technical Committee on Networked Systems. He has been on the Editorial Boards of several journals, including *Automatica*, *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, and *IET Control Theory and Applications*. He is currently a Senior Editor of *IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS* and Associate Editor of *European Journal of Control*. He has been Guest Editor for a special issue of *IEEE TRANSACTIONS ON AUTOMATIC CONTROL* on cyber-physical systems and one of *IEEE Control Systems Magazine* on cyber-physical security. He was the General Chair of the ACM/IEEE Cyber-Physical Systems Week 2010 in Stockholm and IPC Chair of many conferences. He has served on the Executive Committees of several European research projects in the area of networked embedded systems. He received the Best Paper Award of the IEEE International Conference on Mobile Ad-hoc and Sensor Systems in 2009 and the Best Theory Paper Award of the World Congress on Intelligent Control and Automation in 2014. In 2009 he was awarded Wallenberg Scholar, as one of the first ten scholars from all sciences, by the Knut and Alice Wallenberg Foundation. He was awarded Future Research Leader from the Swedish Foundation for Strategic Research in 2005. He received the triennial Young Author Prize from IFAC in 1996 and the Peccei Award from the International Institute of System Analysis, Austria, in 1993. He received Young Researcher Awards from Scania in 1996 and from Ericsson in 1998 and 1999.