Consensus Conditions of Multi-Agent Systems With Time-Varying Topologies and Stochastic Communication Noises

Tao Li, *Member, IEEE*, and Ji-Feng Zhang, *Senior Member, IEEE*

***Abstract*—This paper investigates the average-consensus problem of first-order discrete-time multi-agent networks in uncertain communication environments. Each agent can only use its own and neighbors' information to design its control input. To attenuate the communication noises, a distributed stochastic approximation type protocol is used. By using probability limit theory and algebraic graph theory, consensus conditions for this kind of protocols are obtained: (A) For the case of fixed topologies, a necessary and sufficient condition for mean square average-consensus is given, which is also sufficient for almost sure consensus. (B) For the case of time-varying topologies, sufficient conditions for mean square average-consensus and almost sure consensus are given, respectively. Especially, if the network switches between jointly-containing-spanning-tree, instantaneously balanced graphs, then the designed protocol can guarantee that each individual state converges, both almost surely and in mean square, to a common random variable, whose expectation is right the average of the initial states of the whole system, and whose variance describes the static maximum mean square error between each individual state and the average of the initial states of the whole system.**

***Index Terms*—Average-consensus, distributed coordination, distributed estimation, multi-agent systems, stochastic systems.**

I. Introduction

DISTRIBUTED coordination control of multi-agent networks is a basic problem and attracts a lot of attention from the control community. Among others, consensus control is one of the most fundamental problems in this area, which roughly speaking means to design a network protocol such that as time goes on, all agents asymptotically reach an agreement. In some cases, the agreement is a common value which may be the average of the initial states of the system, and often called average-consensus and has wide application background in the area such as formation control [1], distributed filtering [2], multi-senor data fusion [3] and distributed computation [4]. Olfati-Saber and Murray [5] considered the average-consensus control for first-order integrator networks with fixed and switching topologies. They proved that, if at each time instant, the network is a strongly connected and balanced digraph, then the weighted average-type protocol can ensure average-consensus. Kingston and Beard [6] extended the results of [5] to the discrete-time models and weakened the condition of instantaneous strong connectivity. They proved that if at each time instant the topology graph is balanced and the union of graphs over every bounded time interval is strongly connected, then average-consensus can be achieved. Xiao and Boyd [7] considered first-order discrete-time average-consensus with fixed and undirected topologies. They designed the weighted adjacency matrix to optimize the convergence rate by semi-definite programming. In addition to the above works, some researchers also considered the high-order dynamics [8], [9], the topologies of random graphs [10]–[13] or control design based on individual performance optimization [14]–[17].

Most of the above mentioned works assumed an ideal communication channel between agents, that is, each agent measures its neighbors' states accurately. Obviously, this assumption is only an ideal approximation for real communication channels. Real networks are often interfered by various kinds of noises during the sending, transmission and receiving of information, such as thermal noise, channel fading, quantization effect during encoding and decoding [18], etc. Consensus of dynamic networks with stochastic communication noises is a common problem in distributed systems [19], and has attracted the attention of some researchers [20]–[25]. Ren *et al.* [22]) and Kingston *et al.* [23] introduced time-varying consensus gains and designed consensus protocols based on a Kalman filter structure. They proved that, when there is no communication noise, the designed protocols can ensure consensus to be achieved asymptotically. Xiao *et al.* [24] considered the first-order discrete-time average-consensus control with fixed topologies and additive input noises. They designed the optimal weighted adjacency matrix to minimize the static mean square consensus error. However, since the consensus gain and the adjacency matrix are time-invariant, as time goes on, the state average of the system will diverge with probability one, even if the noises are bounded. Huang and Manton [20] considered the first-order discrete-time consensus control with fixed topologies and communication noises. They introduced decreasing consensus gains $a(k)$ (where k is the discrete time instant) in the protocol to attenuate the noises. They proved that, if $a(k)$ is the same order as $1/k^{\gamma}$, $k \to \infty$, $\gamma \in (0.5, 1]$,

Manuscript received April 19, 2008; April 26, 2009 and November 21, 2009; accepted December 13, 2009. First published February 17, 2010; current version published September 09, 2010. This paper was presented in part at the 27th Chinese Control Conference and the 29th Chinese Control Conference. This work was supported by the National Natural Science Foundation of China (under Grants 60934006 and 60821091) and the Chinese National Laboratory of Space Intelligent Control. Recommended by Associate Editor I.-J. Wang.

The authors are with the Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (e-mail: litao@amss.ac.cn; jif@iss.ac.cn).

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/TAC.2010.2042982

and the network is a strongly connected circulant graph, then the static mean square error between the individual state and the average of the initial states of all agents is in the same order as the variance of the noises; if $a(k)$ satisfy the step rule of classical stochastic approximation, and the network is a connected undirected graph, then the designed protocol can ensure mean square weak consensus. Li and Zhang [25], [26] considered the first-order continuous-time average-consensus control with fixed topologies and communication noises. They used time-varying consensus gains in the protocol and gave a necessary and sufficient condition for asymptotically unbiased mean square average-consensus.

From [26] we can conclude that: (i) if there are communication noises, then the weighted average-type protocol proposed by [5] cannot ensure the stability of the closed-loop system, and even if the noises are bounded, the state average of the system will diverge with probability one. (ii) distributed stochastic approximation type consensus protocols are effective for attenuating communication noises and ensuring mean square consensus. Particularly, Li and Zhang [25], [26] showed that a necessary and sufficient condition for asymptotically unbiased mean square average-consensus is that the consensus gains satisfy the step rule similar to that of classical stochastic approximation.

The existing work on distributed stochastic approximation type consensus protocol is almost focused on the case of fixed topologies. However, there are many kinds of uncertainties for real networks, which are due to not only the existence of communication noises, but also the time-variation of network topologies and parameters. On the one hand, stochastic noises interfere the communication channels among different agents; on the other hand, the network topologies and parameters are time-varying. These two aspects constitute the basic factors of the uncertainties in distributed communication environment. Thus, for distributed coordination of multi-agent systems in uncertain environments, consensus control with time-varying topologies and stochastic communication noises is still a fundamental problem to be solved.

In this paper, we consider the average-consensus control for networks of discrete-time first-order agents with directed topologies. The information available for each agent to design its control input is its local state and the states of its neighbors corrupted by stochastic communication noises. The noises considered here are martingale differences with uniformly bounded second-order moments and finite conditional second-order moments, which include bounded and Gaussian white noises as special cases. Comparing with the existing work [20], [25], [26], here the independency of noises of different communication channels is not required. To attenuate the noises, a distributed stochastic approximation type protocol is used and the convergence properties of the closed-loop system are analyzed by combining probability limit theory and algebraic graph theory.

In summary, the paper is characterized by the following points: i) For the case of fixed topologies, we prove that if the network is a balanced graph containing a spanning tree and the consensus gains satisfy the step rule of classical stochastic approximation, then the closed-loop system will achieve asymptotically unbiased mean square average-consensus. Precisely, the designed protocol can guarantee that each individual state converges in mean square to a common random variable, whose expectation is right the average of the initial states of the whole system. Actually, this condition, in some sense, is necessary and sufficient for asymptotically unbiased mean square average-consensus. ii) For the case of time-varying topologies, the convergence of the closed-loop system is analyzed. The results for the fixed topologies show that to ensure average-consensus with communication noises, it is necessary to introduce decreasing consensus gains. However, the use of decreasing consensus gains leads to the failure of a key assumption, which requires that the nonzero off-diagonal elements of the state matrix are uniformly bounded away from zero. This key assumption has been widely used in the relevant literature for the noise-free cases [27]–[34]. The absence of this assumption together with the time-varying topologies and the communication noises brings essential difficulty to the convergence analysis. We combine martingale convergence theory and algebraic graph theory together. By properly selecting Lyapunov functions, we convert the convergence analysis of matrix products into that of scalar sequences, and show that if the network switches between jointly-containing-spanning tree, instantaneously balanced graphs and the consensus gains decrease with a proper rate, then the closed-loop system achieves asymptotically unbiased mean square average-consensus. iii) The sample path behavior of the distributed consensus protocols is analyzed. It is well known that, for a stochastic system, what can be really observed is always a sample path, so it is more meaningful of designing a protocol to ensure almost sure consensus. By using the nonnegative martingale convergence theorem, sufficient conditions are presented to ensure almost sure consensus for fixed and time-varying topology cases, respectively. Particularly, for the case of fixed topologies, a convergence rate of n step mean consensus error is given in the sense of probability one.

The remainder of this paper is organized as follows. In Section II, some concepts in graph theory is described, and the problem to be investigated is formulated. In Section III, for the case of fixed topologies, a necessary and sufficient condition on the network topology and the consensus gains is given for mean square average-consensus, which is shown to be sufficient for almost sure consensus. In Section IV, for the case of time-varying topologies, sufficient conditions are given to ensure mean square and almost sure consensus, respectively. In Section V, a numerical example is given to illustrate our algorithm. In Section VI, some concluding remarks and future research topics are discussed.

The following notations will be used throughout this paper: We denote a column vector with all ones by $\mathbf{1}$ and the $m \times m$ dimensional identity matrix by I_m. For a given set $\mathcal{S}$, its number of elements is denoted by $|\mathcal{S}|$. For a given vector or matrix A, A^T denotes its transpose, and its 2-norm is denoted by $||A||$; its spectral radius is denoted by $\rho(A)$, and its trace is denoted by $tr(A)$. For a given random variable X, its mathematical expectation is denoted by $E(X)$; and its variance is denoted by $Var(X)$. For a given positive number x, the natural logarithm of x is denoted by $\ln(x)$, the maximum integer which is less than or equal to x is denoted by $\lfloor x \rfloor$, and the minimum integer which is greater than or equal to x is denoted by $\lceil x \rceil$. For a family of random variables (r.v.s) $\{\xi_\lambda, \lambda \in \Lambda\}$, the σ-algebra $\sigma\{\{\xi_\lambda \in B\}, B \in \mathcal{B}, \lambda \in \Lambda\}$ is denoted by $\sigma\{\xi_\lambda, \lambda \in \Lambda\}$, where $\mathcal{B}$ denotes the 1-D Borel set. For a σ-algebras $\mathcal{F}$ and a r.v. ξ, we say ξ is adapted to $\mathcal{F}$, if ξ is $\mathcal{F}$ measurable.

II. Problem Formulation

A. Preliminary in Graph Theory

Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}_\mathcal{G}, \mathcal{A}_\mathcal{G}\}$ be a weighted digraph, where $\mathcal{V} = \{1, 2, \cdots, N\}$ is the set of nodes with node i representing the ith agent, $\mathcal{E}_\mathcal{G}$ is the set of edges and $\mathcal{A}_\mathcal{G} = [a_{ij}] \in \mathbb{R}^{N\times N}$ is the weighted adjacency matrix of $\mathcal{G}$. An edge in $\mathcal{G}$ is denoted by an ordered pair (j, i), and $(j, i) \in \mathcal{E}_\mathcal{G}$ if and only if the jth agent can send information to the ith agent directly. The neighborhood of the ith agent is denoted by $N_i = \{j \in \mathcal{V} \mid (j, i) \in \mathcal{E}_\mathcal{G}\}$. An element of N_i is called a neighbor of i. The ith node is called a source, if it has no neighbors but is a neighbor of another node in $\mathcal{V}$. A node is called an isolated node, if it has no neighbor and it is not a neighbor of any other node.

For any $i, j \in \mathcal{V}$, $a_{ij} \geq 0$, and $a_{ij} > 0 \Leftrightarrow j \in N_i$. The in-degree of i is defined as $deg_{in}(i) = \sum_{j=1}^N a_{ij}$ and the out-degree of i is defined as $deg_{out}(i) = \sum_{j=1}^N a_{ji}$. The Laplacian matrix of $\mathcal{G}$ is defined as $L_\mathcal{G} = \mathcal{D}_\mathcal{G} - \mathcal{A}_\mathcal{G}$, where $\mathcal{D}_\mathcal{G} = diag(deg_{in}(1), \cdots, deg_{in}(N))$.

$\mathcal{G}$ is called a balanced digraph, if $deg_{in}(i) = deg_{out}(i)$, $i = 1, 2, \ldots, N$. $\mathcal{G}$ is called an undirected graph, if $\mathcal{A}_\mathcal{G}$ is a symmetric matrix. It is easily shown that an undirected graph must be a balanced digraph.

For a given positive integer k, the union of k digraphs $\mathcal{G}_1 = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}_1}, \mathcal{A}_{\mathcal{G}_1}\}, \ldots, \mathcal{G}_k = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}_k}, \mathcal{A}_{\mathcal{G}_k}\}$ is denoted by $\sum_{j=1}^k \mathcal{G}_j = \{\mathcal{V}, \cup_{j=1}^k \mathcal{E}_{\mathcal{G}_j}, \sum_{j=1}^k \mathcal{A}_{\mathcal{G}_j}\}$. By the definition of Laplacian matrix, we know that $L_{\sum_{j=1}^k \mathcal{G}_j} = \sum_{j=1}^k L_{\mathcal{G}_j}$.

A sequence $(i_1, i_2), (i_2, i_3), \ldots, (i_{k-1}, i_k)$ of edges is called a directed path from node i_1 to node i_k. $\mathcal{G}$ is called a strongly connected digraph, if for any $i, j \in \mathcal{V}$, there is a directed path from i to j. A strongly connected undirected graph is also called a connected graph. A directed tree is a digraph, where every node except the root has exactly one neighbor and the root is a source. A spanning tree of $\mathcal{G}$ is a directed tree whose node set is $\mathcal{V}$ and whose edge set is a subset of $\mathcal{E}_\mathcal{G}$. For a balanced digraph, containing a spanning tree is equivalent to being strongly connected. We call $\{\mathcal{G}_1, \ldots, \mathcal{G}_k\}$ jointly-containing-spanning-tree, if $\sum_{j=1}^k \mathcal{G}_j$ has a spanning tree. Especially, if $\mathcal{G}_j$, $j = 1, 2, \ldots, k$, are all undirected graphs and $\{\mathcal{G}_1, \ldots, \mathcal{G}_k\}$ jointly contains a spanning tree, then $\sum_{j=1}^k \mathcal{G}_j$ is connected. In this case, $\{\mathcal{G}_1, \ldots, \mathcal{G}_k\}$ is called jointly-connected [33].

Below is a basic theorem on Laplacian matrices:

Theorem 2.1: [5], [35] If $\mathcal{G} = \{\mathcal{V}, \mathcal{E}_\mathcal{G}, \mathcal{A}_\mathcal{G}\}$ is an undirected graph, then $L_\mathcal{G}$ is a symmetric matrix, and has N real eigenvalues, in an ascending order

$$0 = \lambda_1(L_\mathcal{G}) \leq \lambda_2(L_\mathcal{G}) \leq \ldots \leq \lambda_N(L_\mathcal{G})$$

and

$$\min_{x\neq 0, \mathbf{1}^T x = 0} \frac{x^T L_\mathcal{G} x}{\|x\|^2} = \lambda_2(L_\mathcal{G})$$

where $\lambda_2(L_\mathcal{G})$ is called the algebraic connectivity of $\mathcal{G}$. Particularly, if $\mathcal{G}$ is connected, then $\lambda_2(L_\mathcal{G}) > 0$. □

B. Consensus Protocol

In this paper, we consider the average-consensus control for a network of discrete-time first-order agents with the dynamics

$$x_i(t+1) = x_i(t) + u_i(t), \quad t = 0, 1, \ldots, \quad i = 1, \ldots, N \quad (1)$$

where $x_i(t)$ and $u_i(t)$ are the state and control of the ith agent. Here for simplicity, we suppose that $x_i(t)$ and $u_i(t)$ are scalars, and the initial state $x_i(0)$ is deterministic.

The ith agent can receive information from its neighbors

$$y_{ji}(t) = x_j(t) + w_{ji}(t), \quad j \in N_i \quad (2)$$

where $y_{ji}(t)$ denotes the measurement of the jth agent's state $x_j(t)$ by the ith agent; $\{w_{ji}(t), i, j = 1, 2, \ldots, N\}$ are the communication noises. The graph $\mathcal{G}$ shows the structure of the information flow in the system (1), called the information flow graph or network topology graph of the system (1). Denote $X(t) = [x_1(t), \ldots, x_N(t)]^T$, $(\mathcal{G}, X)$ is usually called a dynamic network [5].

We call the group of controls $\mathcal{U} = \{u_i, i = 1, 2, \ldots, N\}$ a measurement-based distributed protocol, if $u_i(t)$ depends only on the state of the ith agent and the measurement of its neighbors' states, that is

$$u_i(t) \in \sigma\left(\cup_{s=0}^t \sigma\left(x_i(s), y_{ji}(s), j \in N_i\right)\right), \quad \forall t = 0, 1, \ldots, \quad i = 1, 2, \ldots, N.$$

The so-called consensus control means to design a measurement-based distributed protocol for the dynamic network $(\mathcal{G}, X)$, such that all agents achieve an agreement on their states in some sense, when $t \to \infty$. The so-called average-consensus control means to design a distributed protocol for the dynamic network $(\mathcal{G}, X)$, such that for any initial value $X(0)$, the states of all the agents converge to $(1/N)\sum_{j=1}^N x_j(0)$ when $t \to \infty$, that is, $(1/N)\sum_{j=1}^N x_j(0)$ can be computed in a distributed way. In this case $(1/N)\sum_{j=1}^N x_j(0)$ is called the group decision value [5]).

Applying the distributed protocol $\mathcal{U}$ to the system (1), (2), generally speaking, will lead to a stochastic closed-loop system, and $x_i(t)$, $i = 1, 2, \ldots, N$ are all stochastic processes. Below we introduce the definition of average-consensus protocol in mean square for stochastic systems.

Definition 2.1: [25] A distributed protocol $\mathcal{U}$ is called an asymptotically unbiased mean square average-consensus protocol if it renders the system (1), (2) has the following properties: for any given $X(0) \in \mathbb{R}^n$, there is a random variable x^*, such that $E(x^*) = (1/N)\sum_{j=1}^N x_j(0)$, $Var(x^*) < \infty$, and

$$\lim_{t\to\infty} E\left[x_i(t) - x^*\right]^2 = 0, \quad i = 1, 2, \ldots, N.$$

□

For the dynamic network $(\mathcal{G}, X)$, we use the distributed protocol

$$u_i(t) = a(t) \sum_{j\in N_i} a_{ij}\left(y_{ji}(t) - x_i(t)\right), \quad t = 0, 1, \ldots \quad (3)$$

where and whereafter $a(t) > 0$, called consensus-gain function, and if $|N_i| = 0$, then the sum $\sum_{j\in N_i}(\cdot)$ is defined as zero.

Remark 1: When $a_{ij} = (1/|N_i|)$, (3) becomes the protocol of [20]. Similar to [20], we call (3) a distributed stochastic approximation type protocol. It intuitively means that each agent updates its state in the direction of the nonnegative gradient of its local Laplacian potential. If the digraph $\mathcal{G}$ is balanced, then $V_{\mathcal{G}} \triangleq x^T L_{\mathcal{G}} x = (1/2)\sum_{i=1}^N \sum_{j\in N_i} a_{ij}(x_j - x_i)^2$ is called the Laplacian potential function associated with $\mathcal{G}$ ([5], which represents the degree of deviation between different agents' states. $(1/2)\sum_{j\in N_i} a_{ij}(x_j - x_i)^2$ is called the local Laplacian potential of the ith agent, and the nonnegative gradient direction with respect to x_i is $\sum_{j\in N_i} a_{ij}(x_j - x_i)$. Due to the communication noises, in the protocol (3), the update direction of the ith agent at time t is $\sum_{j\in N_i} a_{ij}(y_{ji}(t) - x_i(t))$, and $a(t)$ is the step size. When there is no communication noise (i.e. $y_{ji}(t) = x_j(t)$) and $a(t) \equiv 1$, (3) degenerates to the protocol (A.1) of [5].

III. Fixed Topology Case

In this section, we will prove that under mild conditions, the control law (3) is an asymptotically unbiased mean square average-consensus and almost sure strong consensus protocol.

For conciseness of expression, in the sequel we will use the following notations:

$$\mathcal{S} = \left\{\xi \middle| \left\{\xi(t) \in \mathbb{R}^{N^2}, \mathcal{F}^{\xi}(t)\right\} \text{ is a martingale difference}, \right.$$
$$\left. \sigma_{\xi} \triangleq \sup_{t\geq 0} E\|\xi(t)\|^2 < \infty \right\},$$
$$\mathcal{S}' = \left\{\xi \middle| \xi \in \mathcal{S}, \sup_{t\geq 0} E\left(\|\xi(t)\|^2 \middle| \mathcal{F}^{\xi}(t-1)\right) < \infty \ a.s.\right\},$$
$$\widetilde{\mathcal{S}}' = \left\{\xi \middle| \xi \in \mathcal{S}, \sup_{t\geq 0, m\geq 0} E\left(\|\xi(t+m)\|^2 \middle| \mathcal{F}^{\xi}(t)\right) < \infty \ a.s.\right\}$$

where and whereafter $\mathcal{F}^{\xi}(t) = \sigma\{\xi(0), \ldots, \xi(t)\}$.

Remark 2: Obviously, $\widetilde{\mathcal{S}}' \subset \mathcal{S}' \subset \mathcal{S}$. If $\{\xi(t) \in \mathbb{R}^{N^2}, t = 0, 1, \ldots\}$ is a sequence of independent r.v.s with zero mean and uniformly bounded second-order moments, then $\{\xi(t), t = 0, 1, \ldots\} \in \widetilde{\mathcal{S}}'$. So bounded and Gaussian white noises both belong to $\widetilde{\mathcal{S}}'$.

Substituting the protocol (3) into (1) leads to

$$X(t+1) = [I_N - a(t)L_{\mathcal{G}}]X(t) + a(t)D_{\mathcal{G}}W(t), \quad t = 0, 1, \ldots \tag{4}$$

where and whereafter $D_{\mathcal{G}} = diag(\alpha_1^T, \ldots, \alpha_N^T)$ is an $N \times N^2$ dimensional block diagonal matrix with α_i being the ith row of $\mathcal{A}_{\mathcal{G}}$; $W(t) = [w_1^T(t), \ldots, w_N^T(t)]^T$ with $w_i(t) = [w_{1i}(t), \ldots, w_{Ni}(t)]^T$.

We need the following assumptions

A1) $\mathcal{G}$ is a balanced digraph;

A2) $\mathcal{G}$ contains a spanning tree;

A3) $\sum_{t=0}^{\infty} a(t) = \infty$, $\sum_{t=0}^{\infty} a^2(t) < \infty$;

A3a) $\sum_{t=0}^{\infty} a(t) = \infty$, $\lim_{t\to\infty} a(t) = 0$.

Remark 3: Assumption A3a) is weaker than Assumption A3). For example, if $a(k) = (1/(k+1)^{\gamma})$, $k = 0, 1, \ldots$, where $\gamma \in (0, (1/2)]$, then A3a) holds, but A3) fails.

We have the following theorems.

Theorem 3.1: Apply the protocol (3) to the system (1), (2), and suppose that A1)–A2) and A3a) hold. Then for any $W \in \mathcal{S}$

$$\lim_{t\to\infty} E[V(t)] = 0, \quad \forall X(0) \in \mathbb{R}^N. \tag{5}$$

That is, (3) is a mean square weak consensus protocol [20]. Here $V(t) = \|(I_N - (1/N)\mathbf{1}\mathbf{1}^T)X(t)\|^2$ is the energy function of the consensus error.

Proof: Denote $J = (1/N)\mathbf{1}\mathbf{1}^T$, and

$$\delta(t) = (I_N - J)X(t). \tag{6}$$

Then $V(t) = \delta^T(t)\delta(t)$. Thus, from A1) and Theorem 6 of [5] we have $\mathbf{1}^T L_{\mathcal{G}} = 0$ and $JL_{\mathcal{G}} = 0$, which together with (4) leads to

$$\begin{aligned}\delta(t+1) &= X(t) - a(t)L_{\mathcal{G}}X(t) + a(t)D_{\mathcal{G}}W(t) - JX(t) \\ &\quad - a(t)JD_{\mathcal{G}}W(t) \\ &= \delta(t) - a(t)L_{\mathcal{G}}X(t) + a(t)(I_N - J)D_{\mathcal{G}}W(t) \\ &= [I_N - a(t)L_{\mathcal{G}}]\delta(t) + a(t)(I_N - J)D_{\mathcal{G}}W(t) \end{aligned} \tag{7}$$

and

$$\begin{aligned}V(t+1) &= V(t) - 2a(t)\delta^T(t)\widehat{L}_{\mathcal{G}}\delta(t) + a^2(t)\delta^T(t)L_{\mathcal{G}}^T L_{\mathcal{G}}\delta(t) \\ &\quad + 2a(t)\delta^T(t)\left(I_N - a(t)L_{\mathcal{G}}^T\right)(I_N - J)D_{\mathcal{G}}W(t) \\ &\quad + a^2(t)W^T(t)D_{\mathcal{G}}^T(I_N - J)^2 D_{\mathcal{G}}W(t). \end{aligned} \tag{8}$$

From A1) and Theorem 7 of [5], $\widehat{L}_{\mathcal{G}} = ((L_{\mathcal{G}} + L_{\mathcal{G}}^T)/2)$ is the Laplacian matrix of the symmetrized graph[1] $\widehat{\mathcal{G}}$ of $\mathcal{G}$. From A2), noticing that $\widehat{\mathcal{G}}$ is undirected, we know that $\widehat{\mathcal{G}}$ is strongly connected, and hence, from Theorem 2.1, $\lambda_2(\widehat{L}_{\mathcal{G}}) > 0$. Therefore, from $\delta^T(t)\widehat{L}_{\mathcal{G}}\delta(t) \geq \lambda_2(\widehat{L}_{\mathcal{G}})V(t)$ and (8) we have

$$\begin{aligned}V(t+1) &\leq \left(1 - 2\lambda_2(\widehat{L}_{\mathcal{G}})a(t) + a^2(t)\|L_{\mathcal{G}}\|^2\right)V(t) \\ &\quad + 2a(t)\delta^T(t)\left(I_N - a(t)L_{\mathcal{G}}^T\right)(I_N - J)D_{\mathcal{G}}W(t) \\ &\quad + a^2(t)W^T(t)D_{\mathcal{G}}^T(I_N - J)^2 D_{\mathcal{G}}W(t). \end{aligned} \tag{9}$$

Noticing that $\delta(t) \in \mathcal{F}^W(t-1)$ and $W \in \mathcal{S}$, taking mathematical expectation on both sides of the above inequality, we have

$$\begin{aligned}E[V(t+1)] &\leq \left(1 - 2\lambda_2(\widehat{L}_{\mathcal{G}})a(t) + a^2(t)\|L_{\mathcal{G}}\|^2\right)E[V(t)] \\ &\quad + a^2(t)\|D_{\mathcal{G}}\|^2\|I_N - J\|^2\sigma_W. \end{aligned} \tag{10}$$

Noticing that $\lambda_2(\widehat{L}_{\mathcal{G}}) > 0$ and $a(t) \to 0$, $t \to \infty$, we know that there is $t_0 > 0$ such that $a(t)\|L_{\mathcal{G}}\|^2 < \lambda_2(\widehat{L}_{\mathcal{G}})$, and $2a(t)\lambda_2(\widehat{L}_{\mathcal{G}}) \leq 1$, $\forall t \geq t_0$. Thus

$$0 \leq 1 - 2a(t)\lambda_2(\widehat{L}_{\mathcal{G}}) + a^2(t)\|L_{\mathcal{G}}\|^2 < 1, \quad \forall\, t \geq t_0. \tag{11}$$

Then by A3a), we have

$$\sum_{t=t_0}^{\infty}\left(2a(t)\lambda_2(\widehat{L}_{\mathcal{G}}) - a^2(t)\|L_{\mathcal{G}}\|^2\right) \geq \lambda_2(\widehat{L}_{\mathcal{G}})\sum_{t=t_0}^{\infty} a(t) = \infty \tag{12}$$

and

$$\frac{a^2(t)}{2a(t)\lambda_2(\widehat{L}_{\mathcal{G}}) - a^2(t)\|L_{\mathcal{G}}\|^2} \to 0, \quad t \to \infty \tag{13}$$

[1] The definition of the symmetrized graph of a digraph is referred to Definition 2 of [5].

which together with (11), (12) and Lemma A.1 in Appendix A leads to (5). □

Remark 4: In [20], it is proved that if A3) holds, and the network is a connected undirected graph, then the stochastic approximation type protocol can ensure mean square weak consensus. Here, in Theorem 3.1, we give a weaker condition A3a) to ensure mean square weak consensus for balanced digraphs.

In the protocol (3), an agent-independent consensus gain $a(t)$ is used. This requires some coordination of the consensus gain across the agents. It is interesting to investigate the case with agent-dependent consensus gains. For instance, in practical applications, there may be a small error between the actual consensus gain $a_i(t)$ of the ith agent and the designed consensus gain $a(t)$. In this case, the protocol (3) becomes

$$u_i(t) = a_i(t)\sum_{j=1}^{N} a_{ij}\left(y_{ji}(t) - x_i(t)\right), \quad t = 0, 1, \ldots \tag{14}$$

We have the following theorem.

Theorem 3.2: Apply the protocol (14) to the system (1), (2). If Assumptions A1)–A2) hold, and

$$\sum_{t=0}^{\infty} a_j(t) = \infty, \quad j = 1, 2, \ldots, N \tag{15}$$

$$\lim_{t\to\infty} a_j(t) = 0, \quad j = 1, 2, \ldots, N \tag{16}$$

$$\max_{1\le i,j\le N} |a_i(t) - a_j(t)| = o\left(\sum_{j=1}^{N} a_j(t)\right), \quad t \to \infty \tag{17}$$

then for any $W \in \mathcal{S}$

$$\lim_{t\to\infty} E\left[V(t)\right] = 0, \quad \forall\, X(0) \in \mathbb{R}^N. \tag{18}$$

Proof: Denote

$$\bar{a}(t) = \frac{1}{N}\sum_{j=1}^{N} a_j(t),$$
$$\Delta(t) = diag\left(\Delta_1(t), \ldots, \Delta_N(t)\right)$$

where $\Delta_i(t) = \bar{a}(t) - a_i(t)$. Substituting the protocol (14) into the system (1), (2), similar to (8), we have

$$\begin{aligned} V(t+1) &= V(t) - 2\bar{a}(t)\delta^T(t)\widehat{L}_{\mathcal{G}}\delta(t) \\ &+ \bar{a}^2(t)\delta^T(t)L_{\mathcal{G}}^T L_{\mathcal{G}}\delta(t) \\ &+ 2\delta^T(t)(I_N - \bar{a}(t)L_{\mathcal{G}}^T)(I_N - J)\Delta(t)L_{\mathcal{G}}\delta(t) \\ &+ \delta^T(t)L_{\mathcal{G}}^T\Delta(t)(I_N - J)\Delta(t)L_{\mathcal{G}}\delta(t) \\ &+ 2\delta^T(t)L_{\mathcal{G}}^T\Delta(t)(I_N - J) \\ &\times (\bar{a}(t)I_N - \Delta(t))D_{\mathcal{G}}W(t) \\ &+ 2\delta^T(t)(I_N - \bar{a}(t)L_{\mathcal{G}}^T)(I_N - J) \\ &\times (\bar{a}(t)I_N - \Delta(t))D_{\mathcal{G}}W(t) \\ &+ W^T(t)D_{\mathcal{G}}^T(\bar{a}(t)I_N - \Delta(t))(I_N - J) \\ &\times (\bar{a}(t)I_N - \Delta(t))D_{\mathcal{G}}W(t). \end{aligned} \tag{19}$$

From above, noting that $\max_j \sup_{t\ge0} a_j(t) < \infty$, similar to (10), we have

$$E\left[V(t+1)\right] \le (1-q(t))\, E\left[V(t)\right] + N^2\bar{a}^2(t)\|D_{\mathcal{G}}\|^2\|I_N - J\|^2\sigma_W \tag{20}$$

where

$$q(t) = 2\lambda_2(\widehat{L}_{\mathcal{G}})\bar{a}(t) - \bar{a}^2(t)|L_{\mathcal{G}}|^2 - 2\left(1+\alpha_0\|L_{\mathcal{G}}\|\right)\|L_{\mathcal{G}}\|\,\|\Delta(t)\| - \|\Delta(t)\|^2\,\|L_{\mathcal{G}}\|^2$$

and $\alpha_0 = \max_j \sup_{t\ge0} a_j(t)$. By (16), we know that

$$\lim_{t\to\infty} \bar{a}(t) = 0. \tag{21}$$

By (15), we have

$$\sum_{t=0}^{\infty} \bar{a}(t) = \infty. \tag{22}$$

Noting that $\|\Delta(t)\| \le \max_{1\le i,j\le N} |a_i(t) - a_j(t)|$, by (17), we know that

$$\|\Delta(t)\| = o\left(\bar{a}(t)\right), \quad t \to \infty.$$

Then by (20), (21), and (22), similar to (11)–(13), we know that there exists $t_1 > 0$, such that

$$0 < q(t) \le 1, \quad \forall\, t \ge t_1 \tag{23}$$

$$\sum_{t=t_1}^{\infty} q(t) = \infty \tag{24}$$

and

$$\frac{\bar{a}^2(t)}{q(t)} \to 0, \quad t \to \infty. \tag{25}$$

Then by (23), (24), (25) and Lemma A.1 in Appendix A, we have (18). □

For the sufficient conditions to ensure (3) to be an asymptotically unbiased mean square average-consensus protocol, we have the following theorem.

Theorem 3.3: Apply the protocol (3) to the system (1), (2). If Assumptions A1)–A3) hold, then for any $W \in \mathcal{S}$

$$\lim_{t\to\infty} E\left[x_i(t) - x^*\right]^2 = 0, \quad i = 1, 2, \ldots, N, \quad \forall\, X(0) \in \mathbb{R}^N \tag{26}$$

where x^* is a r.v. dependent on W and $X(0)$, satisfying

$$E(x^*) = \frac{1}{N}\sum_{j=1}^{N} x_j(0),$$

$$Var(x^*) \le \frac{\sigma_W^* |\mathcal{E}_{\mathcal{G}}| \sum_{i=1}^{N}\sum_{j\in N_i} a_{ij}^2}{N^2}\sum_{t=0}^{\infty} a^2(t).$$

Especially, if $\{w_{ji}(t), t=0,1,\ldots\}$, $i=1,2,\ldots,N$, $j\in N_i$ are mutually independent, then $V_*\le Var(x^*)\le V^*$, where

$$V^*=\frac{\sigma_W^*|\mathcal{E}_\mathcal{G}|\max_{1\le i<j\le N}a_{ij}^2}{N^2}\sum_{t=0}^{\infty}a^2(t),$$
$$V_*=\frac{\sigma_{W_*}|\mathcal{E}_\mathcal{G}|\min_{1\le i<j\le N}a_{ij}^2}{N^2}\sum_{t=0}^{\infty}a^2(t),$$
$$\sigma_W^*=\max_{(j,i)\in\mathcal{E}_\mathcal{G}}\sup_{t\ge 0}E\left[w_{ji}(t)\right]^2,$$
$$\sigma_{W_*}=\min_{(j,i)\in\mathcal{E}_\mathcal{G}}\inf_{t\ge 0}E\left[w_{ji}(t)\right]^2.$$

That is, (3) is an asymptotically unbiased mean square average-consensus protocol.

Proof: For all $W\in\mathcal{S}$, from (4) and $\mathbf{1}^TL_\mathcal{G}=0$ it follows that:

$$\frac{1}{N}\sum_{j=1}^{N}x_j(t+1)=\frac{1}{N}\sum_{j=1}^{N}x_j(t)+a(t)\frac{1}{N}\mathbf{1}^TD_\mathcal{G}W(t).$$

Taking summation for both sides of the above equations from $t=0$ to $t=n-1$ leads to

$$\frac{1}{N}\sum_{j=1}^{N}x_j(n)=\frac{1}{N}\sum_{j=1}^{N}x_j(0)+\frac{1}{N}\mathbf{1}^TD_\mathcal{G}\sum_{t=0}^{n-1}a(t)W(t). \quad (27)$$

By $W\in\mathcal{S}$ and $\sum_{t=0}^{\infty}a^2(t)<\infty$, we know that $(\sum_{t=0}^{n}a(t)W(t),\mathcal{F}^W(n))$ is a martingale with

$$\sup_{n\ge 0}E\left\|\sum_{t=0}^{n}a(t)W(t)\right\|^2<\infty.$$

Then by Theorem 7.6.10 of [36], it is known that $\sum_{t=0}^{n}a(t)W(t)$ converges in mean square as $n\to\infty$. Denote the limit by $\sum_{t=0}^{\infty}a(t)W(t)$. Then, (26) follows from Theorem 3.1 with

$$x^*=\frac{1}{N}\sum_{j=1}^{N}x_j(0)+\frac{1}{N}\mathbf{1}^TD_\mathcal{G}\sum_{t=0}^{\infty}a(t)W(t).$$

By Corollary 4.2.5 of [37], we have

$$E(x^*)=\frac{1}{N}\sum_{j=1}^{N}x_j(0),$$
$$Var(x^*)=\lim_{n\to\infty}E\left(\frac{1}{N}\mathbf{1}^TD_\mathcal{G}\sum_{t=0}^{n}a(t)W(t)\right)^2$$
$$=\frac{1}{N^2}\sum_{t=0}^{\infty}\left\{a^2(t)E\left(\sum_{i,j}a_{ij}w_{ji}(t)\right)^2\right\}. \quad (28)$$

This together with the Cauchy inequality gives

$$Var(x^*)\le\frac{\sum_{i=1}^{N}|N_i|}{N^2}\lim_{n\to\infty}\sum_{t=0}^{n}\left\{a^2(t)\sum_{i,j}a_{ij}^2E\left(w_{ji}(t)\right)^2\right\}$$
$$\le\frac{\sigma_W^*|\mathcal{E}_\mathcal{G}|\sum_{i,j}a_{ij}^2}{N^2}\sum_{t=0}^{\infty}a^2(t).$$

When $\{w_{ji}(t), t=0,1,\ldots\}$, $i=1,2,\ldots,N$, $j\in N_i$ are independent, by (28) we have

$$Var(x^*)=\frac{1}{N^2}\lim_{n\to\infty}\sum_{t=0}^{n}\left\{a^2(t)\sum_{i,j}a_{ij}^2E\left(w_{ji}(t)\right)^2\right\}$$
$$\le\frac{\sigma_W^*|\mathcal{E}_\mathcal{G}|\max_{1\le i<j\le N}a_{ij}^2}{N^2}\sum_{t=0}^{\infty}a^2(t),$$
$$Var(x^*)\ge\frac{\sigma_{W_*}|\mathcal{E}_\mathcal{G}|\min_{1\le i<j\le N}a_{ij}^2}{N^2}\sum_{t=0}^{\infty}a^2(t).$$

This completes the proof of the theorem. □

Remark 5: Theorems 3.1–3.3 indicate that, for fixed topologies, Assumptions A1)–A3) is a sufficient condition for the protocol (3) to ensure mean square weak consensus and asymptotically unbiased mean square average-consensus. Assumption A2) is to ensure the connectivity of the network to some extent, that is, the algebraic connectivity $\lambda_2(\widehat{L}_\mathcal{G})>0$, such that different agents may asymptotically agree on their states; Assumption A1) is to ensure the state average evolve around $(1/N)\sum_{j=1}^{N}x_j(0)$ such that an average-consensus can be achieved.

Assumption A3) is the step rule of standard stochastic approximation. From the proof of Theorem 3.1, it can be seen that the condition $\sum_{t=0}^{\infty}a(t)=\infty$ is to ensure the consensus error converges to zero with a certain rate. From the proof of Theorem 3.3, one can see the important role played by the condition $\sum_{t=0}^{\infty}a^2(t)<\infty$: when there are communication noises, by (27), the state average of the closed-loop system is not a constant any more, and $\sum_{t=0}^{\infty}a^2(t)<\infty$ ensures convergence of the state average of the closed-loop system.

Remark 6: From Theorem 3.3, it can be seen that, under the control of the protocol (3), there exists static error between the final state of the closed-loop system and the average of the initial states. $Var(x^*)$ describes the static error in the sense of mean square. In fact, it can be shown that, if the conditions of Theorem 3.3 hold, then

$$Var(x^*)=\limsup_{t\to\infty}\max_{1\le i\le N}E\left[x_i(t)-\frac{1}{N}\sum_{j=1}^{N}x_j(0)\right]^2$$

that is, $Var(x^*)$ gives the static maximum mean square error between each individual state and the average of the initial states of the whole system.

Remark 7: In some application of the information fusion of wireless sensor networks, the number N of network nodes is usually very large. Theorem 3.3 gives the analytic expression of the static maximum mean square error between each individual state and the average of the initial states of the whole system, from which, one can see that the impact of N on the accuracy of the information fusion. When the noises of different communication channels are mutually independent, $Var(x^*)$ is proportional to $|\mathcal{E}|/N^2$. Especially, if $|\mathcal{E}|=O(N)$, $\max_{1\le i\le j\le N}a_{ij}=O(1)$, then $Var(x^*)=O(N^{-1})$, $N\to\infty$. This means that the more the network nodes are, the better the effect of the information fusion is. However, a large number of nodes will result in a high cost for running and maintenance of the whole network, so the choice of N is a trade-off between the fusion accuracy and the cost.

On necessity of Assumptions A1)–A3) for asymptotically unbiased mean square average-consensus, we have the following result.

Theorem 3.4: Apply the protocol (3) to the system (1), (2). If (3) is an asymptotically unbiased mean square average-consensus protocol for any $W \in \mathcal{S}$, then Assumptions A1)–A3) hold.

Proof: The proof is provided in Appendix B. □

Remark 8: In [20], sufficient conditions were given to ensure mean square weak consensus for undirected graphs with independent and identically distributed communication noises. In [26], a necessary and sufficient condition was given to ensure continuous-time mean square average-consensus for Gaussian noises. Here, From Theorems 3.3–3.4, it can be seen that A1)–A3) is a necessary and sufficient condition to ensure that (3) is a mean square average-consensus protocol for any communication noises which are martingale differences with bounded second order moments.

For the special case with no communication noise, a sufficient condition for the protocol (3) to ensure average-consensus is given by the following theorem.

Theorem 3.5: Apply the protocol (3) to the system (1), (2) with $W(t)=0, t=0,1,\ldots$. If A1)–A2) hold, $\sum_{t=0}^{\infty} a(t)=\infty$ and

$$\limsup_{t\to\infty} a(t) < \mu \tag{29}$$

where $\mu=\min\{2\lambda_2(\widehat{L}_{\mathcal{G}})/\|L_{\mathcal{G}}\|^2, 1/2\lambda_2(\widehat{L}_{\mathcal{G}})\}$, then

$$\lim_{t\to\infty}\|X(t)-JX(0)\|=0, \quad \forall\, X(0)\in\mathbb{R}^N$$

Proof: Noticing that $W(t)=0$, $t=0,1,\ldots$, by A1) and A2), similar to (10), we have

$$\begin{aligned} V(t+1) \le &\left(1-2\lambda_2(\widehat{L}_{\mathcal{G}})a(t)+a^2(t)\|L_{\mathcal{G}}\|^2\right)V(t)\\ &+a^2(t)\|D_{\mathcal{G}}\|^2\|I_N-J\|^2\sigma_W. \end{aligned}$$

Take a constant $\epsilon_0 \in (0, \mu-\limsup_{t\to\infty} a(t))$. Then by (29), we know that there is $t_0>0$, such that

$$a(t) \le \frac{2\lambda_2(\widehat{L}_{\mathcal{G}})}{\|L_{\mathcal{G}}\|^2}-\epsilon_0, \quad \forall\, t\ge t_0 \tag{30}$$

and

$$a(t) \le \frac{1}{2\lambda_2(\widehat{L}_{\mathcal{G}})}-\epsilon_0, \quad \forall\, t\ge t_0. \tag{31}$$

By (30), we have

$$\begin{aligned} 1-2\lambda_2(\widehat{L}_{\mathcal{G}})a(t)+a^2(t)\|L_{\mathcal{G}}\|^2 &\le 1-\epsilon_0\|L_{\mathcal{G}}\|^2a(t)\\ &<1, \quad \forall\, t\ge t_0. \end{aligned} \tag{32}$$

By (31), we have

$$\begin{aligned} &1-2\lambda_2(\widehat{L}_{\mathcal{G}})a(t)+a^2(t)\|L_{\mathcal{G}}\|^2\\ &\quad\ge 1-2\lambda_2(\widehat{L}_{\mathcal{G}})a(t)\\ &\quad\ge 2\epsilon_0\lambda_2(\widehat{L}_{\mathcal{G}})\ge 0, \quad \forall\, t\ge t_0. \end{aligned} \tag{33}$$

From (30) and $\sum_{t=0}^{\infty} a(t)=\infty$, one gets

$$\sum_{t=t_0}^{\infty}\left(2\lambda_2(\widehat{L}_{\mathcal{G}})a(t)-a^2(t)\|L_{\mathcal{G}}\|^2\right)\ge\epsilon_0\|L_{\mathcal{G}}\|^2\sum_{t=t_0}^{\infty}a(t)=\infty$$

which together with (32), (33) and Lemma A.1 leads to

$$\lim_{t\to\infty} V(t)=0. \tag{34}$$

Similar to (27), noticing that $W(t)=0$, $t=0,1,\ldots$, we have

$$\frac{1}{N}\sum_{j=1}^{N}x_j(t)=\frac{1}{N}\sum_{j=1}^{N}x_j(0), \quad t=1,2,\ldots$$

This together with (34) leads to the conclusion of the theorem. □

Remark 9: Comparing Theorem 3.1 of [26] with Theorem 3.5 here, one can see the difference between continuous-time and discrete-time protocols. In the noise-free case, for the continuous-time protocol in [26] to ensure average-consensus, it is only required that the consensus gain $a(t)$ satisfies $\int_0^\infty a(t)=\infty$; while for the discrete-time protocol (3), generally speaking, only $\sum_0^\infty a(t)=\infty$ is not sufficient, since overlarge consensus gains will make the eigenvalues of the closed-loop state matrix go out of the unit circle of the complex plane. This also reflects the essential difference between continuous-time and discrete-time systems.

From the following theorem, it can be seen that under Assumptions A1)–A3), for a class of communication noises, the protocol (3) can ensure almost sure consensus as well.

Theorem 3.6: Apply the protocol (3) to the system (1), (2). If Assumptions A1)–A3) hold, then for any $W\in\mathcal{S}'$

$$\lim_{t\to\infty} x_i(t)=x^* \quad a.s., \quad i=1,2,\ldots,N, \quad \forall\, X(0)\in\mathbb{R}^N. \tag{35}$$

That is, (3) is an almost sure strong consensus protocol [20]. Here, x^* is given by Theorem 3.3. Furthermore, if $a(t)\downarrow 0$, $t\to\infty$, then

$$\frac{1}{n}\sum_{t=0}^{n}\|\delta(t)\|=o\left(\frac{1}{\sqrt{a(n)n}}\right), \quad n\to\infty\ a.s. \tag{36}$$

where $\delta(t)$ is given by (6).

Proof: For all $W\in\mathcal{S}'$, from $\delta(t)\in\mathcal{F}^W(t-1)$ and (9) it follows that:

$$\begin{aligned} &E\left(V(t+1)|\mathcal{F}^W(t-1)\right)\\ &\quad\le\left(1+a^2(t)\|L_{\mathcal{G}}\|^2\right)V(t)-2\lambda_2(\widehat{L}_{\mathcal{G}})a(t)V(t)\\ &\qquad+a^2(t)tr\left(D_{\mathcal{G}}^T(I-J)^2D_{\mathcal{G}}\right)\\ &\qquad\times E\left(\|W(t)\|^2|\mathcal{F}^W(t-1)\right)\ a.s. \end{aligned} \tag{37}$$

Noticing that $\sup_{t\ge0}E(\|W(t)\|^2|\mathcal{F}^W(t-1))<\infty$ a.s. and $\sum_{t=0}^{\infty}a^2(t)<\infty$, by nonnegative supermartingale convergence theorem [38], [39] we know that $V(t)$ converges almost surely as $t\to\infty$, and

$$\sum_{t=0}^{\infty}a(t)V(t)<\infty\ a.s. \tag{38}$$

Furthermore, by Theorem 3.1 and $\mathcal{S}' \subset \mathcal{S}$ (see Remark 3)

$$\lim_{t\to\infty} V(t) = 0 \; a.s. \tag{39}$$

By $W \in \mathcal{S}'$ and $\sum_{t=0}^{\infty} a^2(t) < \infty$, we know that $\{\sum_{t=0}^{n} a(t)W(t), \mathcal{F}^W(n)\}$ is a martingale with

$$\sup_{n\geq 0} E\left\|\sum_{t=0}^{n} a(t)W(t)\right\|^2 < \infty.$$

Then by Theorem 7.6.10 of [36], it is known that $\sum_{t=0}^{n} a(t)W(t)$ converges both in mean square and almost surely as $n \to \infty$. Thus, by (27) and (39), one gets $x_i(t)$ converges almost surely as $t \to \infty$, $i = 1, 2, \ldots, N$. This together with Theorem 3.3 gives (35).

If $a(t) \downarrow 0$, $t \to \infty$, then by Kronecker lemma [37] and (38) we have

$$\lim_{n\to\infty} a(n)\sum_{t=0}^{n} V(t) = 0 \; a.s.$$

which together with Cauchy inequality results in

$$\frac{1}{n}\sum_{t=0}^{n}\|\delta(t)\| \leq \left(\frac{1}{n}\sum_{t=0}^{n} V(t)\right)^{1/2} = o\left(\frac{1}{\sqrt{a(n)n}}\right) \; a.s.$$

□

Remark 10: Theorem 3.6 implies that, under the same conditions, the states of different agents converge asymptotically to a common random variable with probability one. Note that this random variable may not be precisely the average of the initial states, although its sample mean is. In this case, (36) gives a rough estimate for the convergence rate of n step mean consensus error.

IV. Time-Varying Topology Case

In this section, we consider the case of time-varying topologies. In this case, the distributed protocol is running over a flow of topology graphs $\{\mathcal{G}(t), t = 0, 1, \ldots\}$, where $\mathcal{G}(t) = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}(t)}, \mathcal{A}_{\mathcal{G}(t)}\}$, $t = 0, 1, \ldots$, is a sequence of digraphs with the same vertex set. The edge sets and weighted adjacency matrices are time-varying.

The networks with time-varying topologies can be found in many engineering, biological, social and economic systems, such as the creation and failure of communication links, the loss of data packages, the variation of the channel parameters, and the evolvement and reconfiguration of formations in swarm and flocking [40]. In many cases, fixed topologies are only ideal models, even if the protocol is designed for a fixed topology, it is necessary to consider the robustness of the protocol with respect to the time-variation of the topology. For the stability and consensus of time-varying networked systems without communication noises, the readers are referred to [27], [34].

Here we will consider two kinds of typical topology graph flows

$$\Gamma_1 = \Big\{\{\mathcal{H}(t), t = 0, 1, \ldots\} \,|\, \mathcal{H}(t) \text{ is a balanced digraph}, \quad \forall\, t \geq 0, \quad \sup_{t\geq 0}\left\|\mathcal{A}_{\mathcal{H}(t)}\right\| < \infty\Big\}$$

which is a family of all sequences of balanced graphs with bounded weighted adjacent matrix. It can be seen that a sequence of undirected graphs with bounded weighted adjacent matrix belongs to Γ_1

$$\Gamma_2 = \{\{\mathcal{H}(t), t = 0, 1, \ldots\} \,|\, \mathcal{H}(t) \text{ is a balanced digraph}, \quad \forall t \geq 0, |\{\mathcal{H}(t), t = 0, 1, \ldots\}| < \infty\}$$

which is a family of sequences of switching balanced graphs. Obviously, $\Gamma_2 \subset \Gamma_1$. If $\{\mathcal{H}(t), t = 0, 1, \ldots\} \in \Gamma_2$, then the set $\{\mathcal{H}(t), t = 0, 1, \ldots\}$ has only finite elements. The most common sequence of switching balanced graphs is the sequence of undirected graphs $\{\mathcal{H}(t) = \{\mathcal{V}, \mathcal{E}_{\mathcal{H}(t)}, \mathcal{A}_{\mathcal{H}(t)}\}, t = 0, 1, \ldots\}$ with weighted adjacent matrices $\mathcal{A}_{\mathcal{H}(t)} = [a_{ij}(t)]_{N\times N}$, whose elements take only two kinds of values: when i and j are mutually neighbors, $a_{ij}(t) = a_{ji}(t) = a_{ij} > 0$, otherwise, $a_{ij}(t) = a_{ji}(t) = 0$, $i, j \in \mathcal{V}$. This kind of sequences of switching undirected graphs are widely involved in the synchronization of Vicsek models [33], [41].

For $\{\mathcal{G}(t), t = 0, 1, \ldots\}$, the distributed network protocol is given by

$$u_i(t) = a(t)\sum_{j\in N_i(t)} a_{ij}(t)\left(y_{ji}(t) - x_i(t)\right), \quad \forall\, t = 0, 1, \ldots \tag{40}$$

where $a_{ij}(t)$ is the element of ith row and jth column of $\mathcal{A}_{\mathcal{G}(t)}$, which is the weighted adjacency matrix at time t. $a_{ij}(t) > 0 \Leftrightarrow (j, i) \in \mathcal{E}_{\mathcal{G}(t)}$, and $N_i(t) = \{j \in \mathcal{V} |\, a_{ij}(t) > 0\}$. Here

$$y_{ji}(t) = x_j(t) + w_{ji}(t), \quad j \in N_i(t), \quad t = 0, 1, \ldots \tag{41}$$

Since $u_i(t)$ is adapted to $\sigma(x_i(t), y_{ji}(t), j \in N_i(t))$, $t = 0, 1, \ldots, i = 1, 2, \ldots, N$, $\mathcal{U} = \{u_1, \ldots, u_N\}$ is a distributed protocol.

Substituting the protocol (40) into (1) gives

$$X(t+1) = \left[I_N - a(t)L_{\mathcal{G}(t)}\right]X(t) + a(t)D_{\mathcal{G}(t)}W(t), \quad t = 0, 1, \ldots \tag{42}$$

In this section, we need the following assumption

A4) $a(t+1) \leq a(t)$, $t = 0, 1, \ldots$, and

$$\limsup_{t\to\infty} \frac{a(t)}{a(t+1)} < \infty.$$

Remark 11: If $a(t) = 1/(t+1)$ or $a(t) = \ln(t+2)/(t+1)$, then both A3) and A4) hold. In fact, if there is $\beta_1 \leq 1$, $\beta_2 > -0.5$, $\gamma_1 \leq 1$ and $\gamma_2 > 0.5$, $c_1 > 0$, $c_2 > 0$, such that for sufficiently large t, $c_1/t^{\gamma_1}[\ln(t)]^{\beta_1} \leq a(t) \leq c_2[\ln(t)]^{\beta_2}/t^{\gamma_2}$, then A3) holds. If $a(t)$ decreases monotonically, and there is $\gamma \in (0.5, 1]$ and $\beta \geq -1$, $c_3 > 0$, $c_4 > 0$, such that for sufficiently large t, $c_3[\ln(t)]^{\beta}/t^{\gamma} \leq a(t) \leq c_4[\ln(t)]^{\beta}/t^{\gamma}$, then both A3) and A4) hold.

For convenience of citation, below we denote $\lambda_k^h = \lambda_2(L_{\widehat{\mathcal{G}_k^h}})$, where $\mathcal{G}_k^h = \sum_{i=k}^{k+h-1}\mathcal{G}(i)$. The main results of this section are summarized in the following theorems.

Theorem 4.1: Apply the protocol (40) to the system (1), (41). For any given $\{\mathcal{G}(t), t = 0, 1, \ldots\} \in \Gamma_1$, if there is an integer

$h > 0$ such that $\inf_{m\geq 0} \lambda_{mh}^h > 0$ and Assumptions A3a)–A4) hold, then for any $W \in \mathcal{S}$

$$\lim_{t\to\infty} E[V(t)] = 0, \quad \forall X(0) \in \mathbb{R}^N. \tag{43}$$

That is, (40) is a mean square weak consensus protocol.

Proof: Noticing that $\mathcal{G}(t)$ is a balanced graph, similar to (7), by (42) we have

$$\delta(t+1) = \left[I_N - a(t)L_{\mathcal{G}(t)}\right]\delta(t) + a(t)(I_N - J)D_{\mathcal{G}(t)}W(t) \tag{44}$$

and hence

$$\delta[(m+1)h] = \Phi((m+1)h, mh)\,\delta(mh) + \overline{W}_{mh}^h \tag{45}$$

where $\Phi(n+1,i) = (I_N - a(n)L_{\mathcal{G}(n)})\Phi(n,i)$, $\Phi(i,i) = I_N$, $i = 0,1,\ldots,n$, $n = 1,2,\ldots$; $\overline{W}_k^h = \sum_{j=k}^{k+h-1}\Phi(k+h-1,j)a(j)(I_N - J)D_{\mathcal{G}(j)}W(j)$.

By Assumption A4), we know that there exits constant $C_h > 0$ and positive integer m_0, such that $a(mh) \leq C_h a[(m+1)h]$ and $a(mh) \leq 1, \forall m \geq m_0$. Then by $\sup_{t\geq 0}\|\mathcal{A}_{\mathcal{G}(t)}\| < \infty$, noting that $a(t) \downarrow 0$, we have

$$\begin{aligned}&\Big\|\Phi^T((m+1)h,mh)\,\Phi((m+1)h,mh) - I_N \\ &\quad - \sum_{i=mh}^{(m+1)h-1} a(i)\left(L_{\mathcal{G}(i)} + L_{\mathcal{G}^T(i)}\right)\Big\| \\ &\leq a^2(mh)\sum_{l=2}^{2h}\left(P_h\left(\max\left\{\sup_{t\geq 0}\left\|L_{\mathcal{G}(t)}\right\|, 1\right\}\right)^{P_h}\right) \\ &\leq a^2[(m+1)h]\,M_h, \quad \forall m \geq m_0\end{aligned}$$

where $P_h = 2^{2h} - 2h - 1$

$$M_h = C_h^2 P_h(2h-1)\left(\max\left\{\sup_{t\geq 0}\left\|L_{\mathcal{G}(t)}\right\|, 1\right\}\right)^{P_h}. \tag{46}$$

Thus, by the definition of $V(t)$ and (45), we have

$$\begin{aligned}&V[(m+1)h] \\ &\leq V(mh) - 2\delta^T(mh)\left[\sum_{i=mh}^{(m+1)h-1} a(i)\hat{L}(i)\right]\delta(mh) \\ &\quad + a^2((m+1)h)\,M_h V(mh) + \left(\overline{W}_{mh}^h\right)^T\overline{W}_{mh}^h \\ &\quad + 2\delta^T(mh)\left(\Phi((m+1)h,mh)\right)^T\overline{W}_{mh}^h \\ &\leq V(mh) - 2a((m+1)h)\,\delta^T(mh) \\ &\quad \times\left[\sum_{i=mh}^{(m+1)h-1}\hat{L}(i)\right]\delta(mh) + a^2((m+1)h)\,M_h V(mh) \\ &\quad + \left(\overline{W}_{mh}^h\right)^T\overline{W}_{mh}^h + 2\delta^T(mh) \\ &\quad \times\left(\Phi((m+1)h,mh)\right)^T\overline{W}_{mh}^h, \quad \forall m \geq m_0\end{aligned} \tag{47}$$

where $\hat{L}(i) = ((L_{\mathcal{G}(i)} + L_{\mathcal{G}(i)}^T)/2)$. From $W \in \mathcal{S}$, $\delta(mh) \in \mathcal{F}^W(mh-1)$ and the definition of $\overline{W}_{mh}^h$ it follows that:

$$E\left[\delta^T(mh)(\Phi((m+1)h,mh))^T\,\overline{W}_{mh}^h|\mathcal{F}^W(mh-1)\right] = 0\ a.s.$$

which implies

$$E\left[\delta^T(mh)\left(\Phi((m+1)h,mh)\right)^T\overline{W}_{mh}^h\right] = 0. \tag{48}$$

Further, by $\sup_{t\geq 0}\|D_{\mathcal{G}(t)}\| < \infty$, there exists a constant $N_h > 0$ such that

$$E\left[\overline{W}(mh)^T\overline{W}_{mh}^h\right] \leq N_h\sum_{i=mh}^{(m+1)h-1} a^2(i). \tag{49}$$

Since $\mathcal{G}(i)$, $i = 0,1,\ldots$, is a balanced digraph, by Theorem 7 of [5], $\hat{L}_i = L_{\widehat{\mathcal{G}}(i)}$, where $\widehat{\mathcal{G}} = \{\mathcal{V}, \mathcal{E}_{\widehat{\mathcal{G}}}, \mathcal{A}_{\widehat{\mathcal{G}}}\}$ denotes the symmetrized graph of $\mathcal{G} = \{\mathcal{V}, \mathcal{E}_{\mathcal{G}}, \mathcal{A}_{\mathcal{G}}\}$. By the definition of the union graph of symmetrized graphs we have $\sum_{i=mh}^{(m+1)h-1}\widehat{\mathcal{G}}(i) = \widehat{\mathcal{G}}_{mh}^h$, which in turn gives

$$\sum_{i=mh}^{(m+1)h-1}\hat{L}(i) = \sum_{i=mh}^{(m+1)h-1} L_{\widehat{\mathcal{G}}(i)} = L_{\sum_{i=mh}^{(m+1)h-1}\widehat{\mathcal{G}}(i)} = L_{\widehat{\mathcal{G}}_{mh}^h}.$$

Thus, by (47), (48), (49) and Theorem 2.1 we have

$$\begin{aligned}E(V[(m+1)h]) &\leq \big(1 - 2\lambda_{mh}^h a((m+1)h) \\ &\quad + a^2((m+1)h)\,M_h\big)\,E[V(mh)] \\ &\quad + N_h\sum_{i=mh}^{(m+1)h-1} a^2(i) \\ &= \left(1 - 2\left(\inf_{m\geq 0}\lambda_{mh}^h\right)a((m+1)h)\right. \\ &\quad \left. + a^2((m+1)h)\,M_h\right)E[V(mh)] \\ &\quad + N_h\sum_{i=mh}^{(m+1)h-1} a^2(i), \quad \forall m \geq m_0.\end{aligned} \tag{50}$$

Noticing that $\inf_{m\geq 0}\lambda_{mh}^h > 0$

$$\sum_{m=0}^{\infty} a(mh) \geq \frac{1}{h}\sum_{m=0}^{\infty}\sum_{i=mh}^{(m+1)h-1} a(i) = \sum_{t=0}^{\infty} a(t) = \infty$$

and

$$\sum_{i=mh}^{(m+1)h-1} a^2(i) \to 0, \quad m \to \infty$$

similar to (11), (12) and (13), by (50) and Lemma A.1, we get $E[V(mh)] \to 0$, $m \to \infty$.

Therefore, for any given $\epsilon > 0$, there is an $m_1 > 0$ such that

$$E[V(mh)] \leq \epsilon, \quad \forall m \geq m_1 \tag{51}$$

and

$$a^2(t) < \epsilon, \quad \forall t \geq m_1 h. \tag{52}$$

Let $m_t = \lfloor t/h \rfloor$. Then, for any given $t \geq m_1 h$, we have $m_t \geq m_1$ and

$$0 \leq t - m_t h \leq h. \tag{53}$$

From (44) and the definition of $V(t)$, we have

$$E[V(n+1)] \le \widetilde{\phi}(n+1,k)E[V(k)] + K_h \sum_{i=k}^{n} \widetilde{\phi}(n,i)a^2(i), \quad \forall\, k \ge 0 \quad (54)$$

where $K_h = \sup_{t\ge 0} \|D_{\mathcal{G}(t)}\|^2 \|I_N - J\|^2 \sigma_W$, $\widetilde{\phi}(n,i) = \prod_{j=i}^{n-1} (1 - 2\lambda_2(\widehat{L}_{\mathcal{G}(j)})a(j) + a^2(j)\|L_{\mathcal{G}(j)}\|^2)$, $i = 0,1,\ldots,n-1$, $n = 1,2,\ldots$; $\widetilde{\phi}(i,i) = 1$, $i = 0,1,\ldots$. Thus, there exists a $\gamma \ge 1$ such that $|\widetilde{\phi}(n,i)| \le \gamma^{n-i}$, $\forall n \ge i \ge 0$. This together with (53), (51), (52) and (54) gives

$$\begin{aligned} E[V(t+1)] &\le \widetilde{\phi}(t+1,m_t h)E[V(m_t h)] + K_h \sum_{i=m_t h}^{t} \widetilde{\phi}(t,i)a^2(i) \\ &\le \gamma^h \epsilon + \gamma^h K_h \sum_{i=m_t h}^{t} a^2(i) \\ &\le \gamma^h (1 + K_h(h+1))\epsilon, \quad \forall\, t \ge m_1 h. \end{aligned}$$

Hence, (43) follows from the arbitrariness of ϵ. □

Theorem 4.2: Apply the protocol (40) to the system (1), (41). For any given $\{\mathcal{G}(t), t = 0,1,\ldots\} \in \Gamma_1$, if there is an integer $h > 0$ such that $\inf_{m\ge 0} \lambda_{mh}^h > 0$ and Assumptions A3)–A4) hold, then for any $W \in \mathcal{S}$

$$\lim_{t\to\infty} E[x_i(t) - \widetilde{x}_*]^2 = 0, \quad i = 1,2,\ldots,N, \quad \forall\, X(0) \in \mathbb{R}^N \quad (55)$$

where $\widetilde{x}_*$ is a r.v. dependent on W, $X(0)$ and $\{\mathcal{G}(t), t = 0,1,\ldots\}$, satisfying $E(\widetilde{x}_*) = (1/N)\sum_{j=1}^N x_j(0)$ and $Var(\widetilde{x}_*) < \infty$. That is, (40) is an asymptotically unbiased mean square average-consensus protocol.

Proof: By (42) and similar to (27) we have

$$\frac{1}{N}\sum_{j=1}^N x_j(n) = \frac{1}{N}\sum_{j=1}^N x_j(0) + \frac{1}{N}\mathbf{1}^T \sum_{t=0}^{n-1} a(t) D_{\mathcal{G}(t)} W(t). \quad (56)$$

By $W \in \mathcal{S}$, $\sup_{t\ge 0} \|D_{\mathcal{G}(t)}\|^2 < \infty$ and $\sum_{t=0}^\infty a^2(t) < \infty$, $\sum_{t=0}^n a(t) D_{\mathcal{G}(t)} W(t)$ is convergent in mean square. Hence, by Theorem 4.1, similar to the proof of Theorem 3.3, we have (55). □

Remark 12: Different from the randomly time-varying communication link failures considered in [42], here, the network topology may change continuously and to ensure mean square average-consensus, we do not need additional distribution conditions on the events of the link failures and creations.

Theorem 4.3: Apply the protocol (40) to the system (1), (41). For any given $\{\mathcal{G}(t), t = 0,1,\ldots\} \in \Gamma_1$, if there is an integer $h > 0$ such that $\inf_{m\ge 0} \lambda_{mh}^h > 0$ and Assumptions A3)–A4) hold, then for any $W \in \widetilde{\mathcal{S}}'$

$$\lim_{t\to\infty} x_i(t) = \widetilde{x}_* \; a.s. \; i = 1,2,\ldots,N, \; \forall\, X(0) \in \mathbb{R}^N \quad (57)$$

where $\widetilde{x}_*$ is given by Theorem 4.2. That is, (40) is an almost sure strong consensus protocol. □

The proof of Theorem 4.3 needs the following two lemmas.

Lemma 4.1: For a sequence of digraphs $\{G(t), t = 0,1,\ldots\}$, the following three statements are equivalent.

i) There is an integer $h > 0$ such that $\inf_{m\ge 0} \lambda_{mh}^h > 0$.
ii) There is an integer $h > 0$ such that $\inf_{k\ge 0} \lambda_k^h > 0$.
iii) There are integers $h > 0$ and $k_0 > 0$ such that $\inf_{m\ge 0} \lambda_{k_0+mh}^h > 0$.

Proof: $(ii) \Rightarrow (i)$ and $(ii) \Rightarrow (iii)$ are straightforward. It suffices to show that $(i) \Rightarrow (ii)$ and $(iii) \Rightarrow (i)$.

$(i) \Rightarrow (ii)$. Suppose that $\inf_{m\ge 0} \lambda_{mh_0}^{h_0} > 0$ holds for some integer $h_0 > 0$. For any given $k \ge 0$, set $n_k = \lceil k/h_0 \rceil$. Then, $L_{\sum_{j=k}^{k+2h_0-1} \widehat{G}(j)} - L_{\sum_{i=n_k h_0}^{(n_k+1)h_0-1} \widehat{G}(i)}$ is the Laplacian matrix of the union of graphs: $\sum_{j=k}^{n_k h_0 - 1} \widehat{G}(j)$ and $\sum_{i=(n_k+1)h_0}^{k+2h_0-1} \widehat{G}(i)$. Thus, $L_{\sum_{j=k}^{k+2h_0-1} \widehat{G}(j)} - L_{\sum_{i=n_k h_0}^{(n_k+1)h_0-1} \widehat{G}(i)}$ is positive semidefinite, which together with Theorem 2.1 gives

$$\lambda_k^{2h_0} \ge \lambda_{n_k h_0}^{h_0} \ge \inf_{m\ge 0} \lambda_{mh}^h > 0, \quad k = 0,1,\ldots$$

Thus $\inf_{k\ge 0} \lambda_k^{2h_0} > 0$.

$(iii) \Rightarrow (i)$. Suppose that $\inf_{m\ge 0} \lambda_{k_0+mh_0}^{h_0} > 0$ holds for some integers $k_0 > 0$ and $h_0 > 0$. Let $\overline{h} = k_0 + 2h_0$, $n_m = \lceil (m\overline{h} - k_0)/h_0 \rceil$, $m = 1,2,\ldots$. Then, $L_{\sum_{i=m\overline{h}}^{(m+1)\overline{h}-1} \widehat{G}(i)} - L_{\sum_{j=k_0+n_m h_0}^{k_0+(n_m+1)h_0-1} \widehat{G}(j)}$ is positive semidefinite, which together with Theorem 2.1 gives

$$\lambda_{m\overline{h}}^{\overline{h}} \ge \lambda_{k_0+n_m h_0}^{h_0} \ge \inf_{m\ge 0} \lambda_{k_0+mh_0}^{h_0} > 0, \quad m = 1,2,\ldots$$

Noticing that $\lambda_0^{\overline{h}} \ge \lambda_{k_0}^{h_0} \ge \inf_{m\ge 0} \lambda_{k_0+mh_0}^{h_0} > 0$, we have

$$\inf_{m\ge 0} \lambda_{m\overline{h}}^{\overline{h}} \ge \inf_{m\ge 0} \lambda_{k_0+mh_0}^{h_0} > 0.$$

□

Lemma 4.2: Apply the protocol (40) to the system (1), (41). For any given $\{\mathcal{G}(t), t = 0,1,\ldots\} \in \Gamma_1$, if there are integers $h > 0$ and $k_0 \ge 0$ such that $\inf_{m\ge 0} \lambda_{k_0+mh}^h > 0$ and Assumptions A3)–A4) hold, then for any $W \in \mathcal{S}'$

$$\lim_{m\to\infty} V(k_0 + mh) = 0 \; a.s.$$

Proof: First, by Lemma 4.1 and Theorem 4.1 we have

$$\lim_{t\to\infty} E[V(t)] = 0. \quad (58)$$

By $W \in \widetilde{\mathcal{S}}'$, there is a constant $\overline{N}_h > 0$ such that

$$\begin{aligned} \sup_{m\ge 0} E\left[\left\|\overline{W}_{k_0+mh}^h\right\|^2 |\mathcal{F}^W(k_0+mh-1)\right] &\le \overline{N}_h \sup_{t\ge 0, m\ge 0} E\left[\|W(t+m)\|^2 |\mathcal{F}^W(t)\right] \\ &\quad \times \sum_{i=k_0+mh}^{k_0+(m+1)h-1} a^2(i). \end{aligned} \quad (59)$$

By (44), similar to (45) and (47), we have

$$\delta[k_0 + (m+1)h] = \Phi(k_0 + (m+1)h, k_0 + mh) \times \delta(k_0 + mh) + \overline{W}_{k_0+mh}^h$$

and

$$\begin{aligned} V[k_0+(m+1)h] &\le \big(1-2\lambda_{k_0+mh}^h a(k_0+(m+1)h) \\ &\quad + a^2(k_0+(m+1)h)M_h\big)V(k_0+mh) \\ &\quad + \left(\overline{W}_{k_0+mh}^h\right)^T \overline{W}_{k_0+mh}^h + 2\delta^T(k_0+mh)\Phi^T \\ &\quad \times (k_0+(m+1)h, k_0+mh)\overline{W}_{k_0+mh}^h. \end{aligned} \tag{60}$$

Notice that $\{(V(k_0+mh), \mathcal{F}^W(k_0+mh-1)), m=0,1,\ldots\}$ is an adapted sequence. Then, from (60) and (59) it follows that

$$\begin{aligned} E\big(V[k_0+(m+1)h]\,|\mathcal{F}^W(k_0+mh-1)\big) &\le \big(1+a^2(k_0+(m+1)h)M_h\big)V(k_0+mh) \\ &\quad + \overline{N}_h \sup_{t\ge 0, m\ge 0} E\left[\|W(t+m)\|^2\,|\mathcal{F}^W(t)\right] \\ &\quad \times \sum_{i=k_0+mh}^{k_0+(m+1)h-1} a^2(i)\ a.s. \end{aligned} \tag{61}$$

where M_h is given by (46). Since $\sum_{m=1}^\infty a^2(k_0+mh)<\infty$, and $\sum_{m=0}^\infty \sum_{i=k_0+mh}^{k_0+(m+1)h-1} a^2(i)<\infty$, by (59), (61) and the nonnegative supermartingale convergence theorem we know that $V(k_0+mh)$ converges a.s. as $m\to\infty$. Furthermore, by (58) we have

$$\lim_{m\to\infty} V(k_0+mh)=0\ a.s.$$

□

Proof of Theorem 4.3: By Lemma 4.1, there exists $\widetilde{h}>0$ such that $\inf_{m\ge 0}\lambda_{l+m\widetilde{h}}^{\widetilde{h}} \ge \inf_{k\ge 0}\lambda_k^{\widetilde{h}}>0$, $l=0,1,\ldots,\widetilde{h}-1$. Thus, it follows from Lemma 4.2 that

$$\lim_{m\to\infty} V(l+m\widetilde{h})=0\ a.s.,\ l=0,1,\ldots\widetilde{h}-1.$$

This implies

$$\lim_{t\to\infty} V(t)=0\ a.s. \tag{62}$$

Since $\{\sum_{t=0}^n a(t)D_{\mathcal{G}(t)}W(t), \mathcal{F}^W(n)\}$ is martingale with $\sup_{n\ge 0} E\|\sum_{t=0}^n a(t)D_{\mathcal{G}(t)}W(t)\|^2<\infty$, by Theorem 7.6.10 of [36], we know that $\sum_{t=0}^n a(t)D_{\mathcal{G}(t)}W(t)$ converges almost surely as $n\to\infty$. This together with (62) and (56) implies that $x_i(t)$, $i=1,2,\ldots,N$, converges almost surely as $t\to\infty$. Thus, by Theorem 4.2 we get (57). □

Corollary 4.1: Apply the protocol (40) to the system (1), (41). For any given $\{\mathcal{G}(t), t=0,1,\ldots\}\in\Gamma_2$, if there is an integer $h>0$ such that for any $m\ge 0$, $\sum_{i=mh}^{(m+1)h-1}\mathcal{G}(i)$ contains a spanning tree, and Assumptions A3)–A4) hold, then for any $W\in\mathcal{S}$

$$\lim_{t\to\infty} E[x_i(t)-\widetilde{x}_*]^2=0, \quad i=1,2,\ldots,N, \quad \forall X(0)\in\mathbb{R}^N.$$

Proof: Since $\sum_{i=mh}^{(m+1)h-1}\mathcal{G}(i)$, $m=0,1,\ldots$, has a spanning tree, $\widehat{\mathcal{G}}_{mh}^h$, $m=0,1,\ldots$, is strongly connected, which together with Theorem 2.1 implies $\lambda_{mh}^h>0$, $m=0,1,\ldots$. Furthermore, by $\{\mathcal{G}(t), t=0,1,\ldots\}\in\Gamma_2$, $|\{\lambda_{mh}^h, m=0,1,\ldots\}|<\infty$, and hence, $\inf_{m\ge 0}\lambda_{mh}^h=\min_{m\ge 0}\lambda_{mh}^h>0$. This together with Theorem 4.2 and $\Gamma_2\subset\Gamma_1$ completes the proof. □

Corollary 4.2: Apply the protocol (40) to the system (1), (41). For any given $\{\mathcal{G}(t), t=0,1,\ldots\}\in\Gamma_2$, if there is an integer $h>0$ such that for any $m\ge 0$, $\sum_{i=mh}^{(m+1)h-1}\mathcal{G}(i)$ contains a spanning tree, and A3)–A4) hold, then for any $W\in\widetilde{\mathcal{S}}'$

$$\lim_{t\to\infty} x_i(t)=\widetilde{x}_*\ a.s.\ i=1,2,\ldots,N,\ \forall X(0)\in\mathbb{R}^N$$

where $\widetilde{x}_*$ is given by Theorem 4.2. That is, (40) is an almost sure strong consensus protocol.

Proof: Similar to Corollary 4.1, we can get $\inf_{m\ge 0}\lambda_{mh}^h>0$. This together with Theorem 4.3 and $\Gamma_2\subset\Gamma_1$ leads to the desired conclusion. □

Remark 13: Theorems 4.1–4.3 are for the case of time-varying graph flows, while Corollaries 4.1–4.2 are for the special cases of switching graph flows, where the network switches among a finite number of digraphs and the condition that there is $h>0$ such that $\inf_{m\ge 0}\lambda_{mh}^h>0$ is equivalent to that there is $h>0$ such that for any $m\ge 0$, $\sum_{i=mh}^{(m+1)h-1}\mathcal{G}(i)$ contains a spanning tree, that is, $\{\mathcal{G}(i), i=mh, mh+1,\ldots,(m+1)h-1\}$, $m=0,1,\ldots$, are all jointly-containing-spanning-tree.

V. Numerical Example

Example 1: Consider a dynamic network of three agents with switching topologies $\mathcal{G}(t)=\{\mathcal{V}=\{1,2,3\},\mathcal{E}(t),\mathcal{A}(t)\}$. When $t=2k+1, k=0,1,\ldots,\mathcal{E}(t)=\{(1,2),(2,1)\},\mathcal{A}(t)=[\alpha_{ij}]_{3\times 3}$, where $\alpha_{12}=\alpha_{21}=0.8, \alpha_{ij}=0, (i,j)\notin\mathcal{E}(t)$. When $t=2k, k=0,1,\ldots,\mathcal{E}(t)=\{(2,3),(3,2)\},\mathcal{A}(t)=[\beta_{ij}]_{3\times 3}$, where $\beta_{23}=\beta_{32}=0.8$, $\beta_{ij}=0$, $(i,j)\notin\mathcal{E}(t)$. It can be seen that $\mathcal{G}_j$ is balanced, $j=0,1,\ldots$ and $\sum_{j=2k}^{2k+1}\mathcal{G}(j)$ has a spanning tree, $k=0,1,\ldots$. The initial states of agents are given by $x_1(0)=0.5$, $x_2(0)=0.3$ and $x_3(0)=-0.8$. The communication noises $w_{12}(t)$, $w_{21}(t)$, $w_{23}(t)$ and $w_{32}(t)$ are independent white noises with uniform distribution on [−0.3, 0.3]. If the protocol (A.1) of [5] is used, due to the communication noises, the actual control input of each agent is given by

$$u_1(t)=\begin{cases} 0, & t=2k, \\ 0.8\,(x_2(t)-x_1(t)+w_{21}(t)), & t=2k+1; \end{cases} \tag{63}$$

$$u_2(t)=\begin{cases} 0.8\,(x_3(t)-x_2(t)+w_{32}(t)), & t=2k, \\ 0.8\,(x_1(t)-x_2(t)+w_{12}(t)), & t=2k+1; \end{cases} \tag{64}$$

$$u_3(t)=\begin{cases} 0, & t=2k+1, \\ 0.8\,(x_2(t)-x_3(t)+w_{23}(t)), & t=2k \end{cases} \tag{65}$$

where $k=0,1,\ldots$ According to [6], if the network switches between jointly-containing-spanning-tree, instantaneously balanced graphs, then the weighted average type protocol can ensure average-consensus. However, due to the communication noises, we can see that the actual closed-loop system is divergent as shown in Fig. 1, where and whereafter, the curves of states are sample path plots for one-time implementation of the communication noises.

If we take the consensus gain $a(t)=\ln(t+2)/(t+2)$, $t=0,1,\ldots$, then the control input of each agent is given by

$$u_1(t)=\begin{cases} 0, & t=2k, \\ a(t)\,(x_2(t)-x_1(t)+w_{21}(t)), & t=2k+1; \end{cases} \tag{66}$$

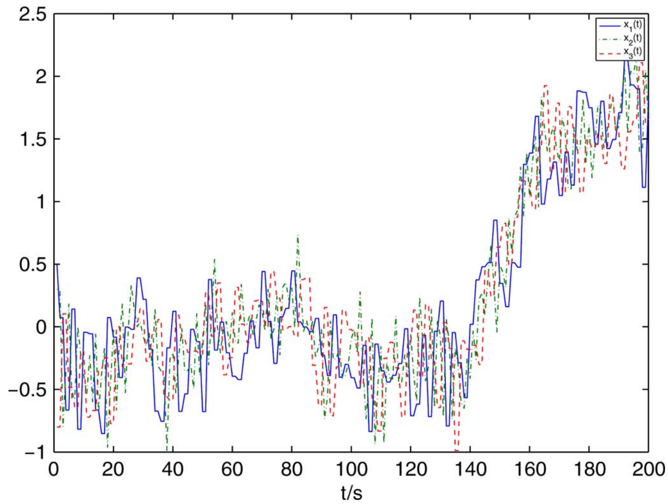


Fig. 1. Curves of states under the protocol (63)–(65) with bounded noises.

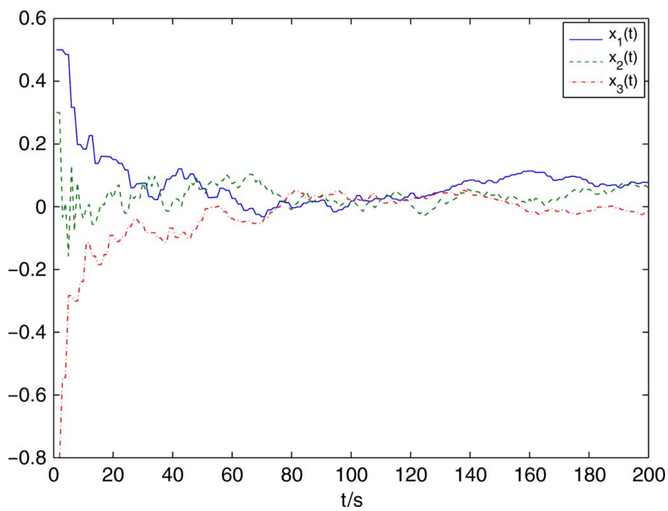


Fig. 2. Curves of states under the protocol (66)–(68) with bounded noises.

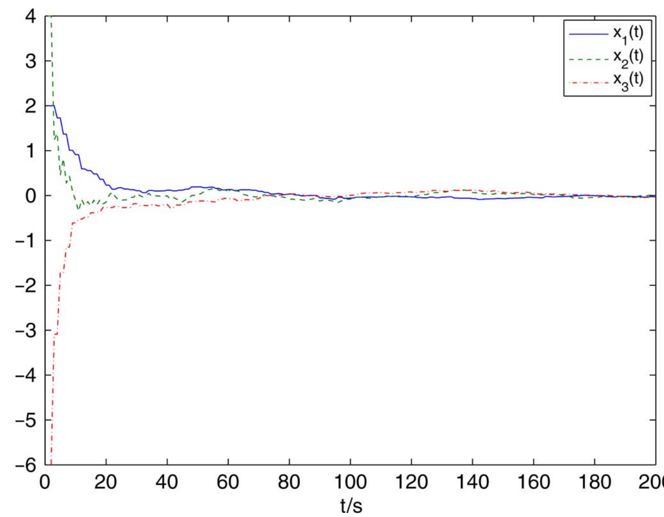


Fig. 3. Curves of states under the protocol (66)–(68) with Gaussian noises.

$$u_2(t) = \begin{cases} a(t)\left(x_3(t) - x_2(t) + w_{32}(t)\right), & t = 2k, \\ a(t)\left(x_1(t) - x_2(t) + w_{12}(t)\right), & t = 2k+1; \end{cases} \quad (67)$$

$$u_3(t) = \begin{cases} 0, & t = 2k+1, \\ a(t)\left(x_2(t) - x_3(t) + w_{23}(t)\right) & t = 2k \end{cases} \quad (68)$$

where $k = 0, 1, \ldots$ In this case, the conditions of Corollary 4.1–4.2 hold. The states of the closed-loop system are shown in Fig. 2. It is shown that average-consensus is achieved asymptotically.

If the communication noises are Gaussian white noises with distribution $N(0, 0.25)$, and the initial states are given by $x_1(0) = 2$, $x_2(0) = 4$ and $x_3(0) = -6$, then under the protocol (66)–(68), the states of the closed-loop system are shown in Fig. 3. It is shown that average-consensus is achieved asymptotically.

VI. Conclusion

In this paper, average-consensus control has been considered for networks of discrete-time first-order agents with fixed and time-varying topologies. The control input of each agent can only use its local state and the states of its neighbors corrupted by stochastic communication noises. Due to the communication noises, the stability of the closed-loop system cannot be ensured by using only the weighted average-type protocol proposed by [5]. To solve this problem, a distributed stochastic approximation type protocol is adopted to reduce the impact of the noises. By the probability limit theory and algebraic graph theory, consensus conditions for this kind of protocols are obtained. For the case of fixed topologies, a necessary and sufficient condition for mean square average-consensus is given, which is also sufficient for almost sure consensus. For the case of time-varying topologies, sufficient conditions for mean square average-consensus and almost sure consensus are given, respectively. Our research shows that distributed stochastic approximation type consensus protocol is strongly robust against both the time-variation of topologies and the communication noises. If the network topology is a balanced digraph jointly containing a spanning tree, then the designed protocol can guarantee that each individual state converges in mean square to a common random variable. For the future research, continuous-time cases, more complex agent dynamics may be considered. In addition, it is interesting to investigate the case with both communication noises and time-delays.

Appendix A

Lemma A.1: [43] Let $\{u(k), k = 0, 1, \cdots\}$, $\{\alpha(k), k = 0, 1, \cdots\}$ and $\{q(k), k = 0, 1, \cdots\}$ be real sequences, satisfying $0 < q(k) \leq 1$, $\alpha(k) \geq 0$, $k = 0, 1, \cdots$, $\sum_{k=0}^{\infty} q(k) = \infty$, $\alpha(k)/q(k) \to 0$, $k \to \infty$, and

$$u(k+1) \leq (1 - q(k))\, u(k) + \alpha(k).$$

Then $\limsup_{k\to\infty} u(k) \leq 0$. In particular, if $u(k) \geq 0$, $k = 0, 1, \cdots$, then $u(k) \to 0$, $k \to \infty$.

Appendix B
The Proof of Theorem 3.4

Lemma B.1: Apply the protocol (3) to the system (1), (2). If $W(t) = 0$, $t = 0, 1, \ldots$, then

$$\lim_{t\to\infty} \|X(t) - JX(0)\| = 0, \quad \forall\, X(0) \in \mathbb{R}^N \qquad \text{(B.1)}$$

only if A1)–A2) hold.

Proof: It suffices to show that if $\mathcal{G}$ is a non-balanced graph or $\mathcal{G}$ does not contain a spanning tree, then (B.1) does not hold.

Step 1: Consider the case where $\mathcal{G}$ is a non-balanced graph. In this case, since $L_\mathcal{G}$ is the Laplacian matrix, $L_\mathcal{G}$ has a zero eigenvalue. And hence, there exists an N-dimensional vector α, $\alpha^T\mathbf{1} = 1$, such that $\alpha^T L_\mathcal{G} = 0$. Furthermore, by Theorem 6 of [5], $\alpha \neq (1/N)\mathbf{1}$. This together with (4) and $W(t) \equiv 0$ implies that $\alpha^T X(t+1) = \alpha^T X(t), t = 0, 1, \ldots$. Thus

$$\alpha^T X(t) \equiv \alpha^T X(0), \quad \forall X(0) \in \mathbb{R}^n. \tag{B.2}$$

When (B.1) holds, so does (5), and hence

$$\lim_{t\to\infty} \alpha^T X(t) = \alpha^T J X(0) = \frac{1}{N}\mathbf{1}^T X(0), \quad \forall X(0) \in \mathbb{R}^n$$

which together with (B.2) leads to $\alpha = (1/N)\mathbf{1}$. This contradicts $\alpha \neq (1/N)\mathbf{1}$. Thus, (B.1) does not hold.

Step 2: Consider the case where $\mathcal{G}$ does not contain a spanning tree. In this case, there are only three possibilities [29]:

I) $\mathcal{G}$ has at least one isolated node i_0. Applying protocol (3) results in

$$\begin{cases} x_{i_o}(t+1) = x_{i_o}(t), \\ \widetilde{X}(t+1) = \left(I_{N-1} - a(t)\widetilde{\mathcal{L}}\right)\widetilde{X}(t),\ t = 0, 1, \ldots \end{cases}$$

where $\widetilde{X}(t) = [x_1(t), \cdots, x_{i_o-1}(t), x_{i_o+1}(t), \cdots, x_N(t)]^T$; $\widetilde{\mathcal{L}}$ is the Laplacian matrix of the graph removing the isolated node i_0. Take $x_{i_o}(0) = 0, x_j(0) = 1, \forall\, j \neq i_0$. Then, $x_{i_o}(0) = 0$ implies $x_{i_o}(t) \equiv 0$. By $\widetilde{\mathcal{L}}\mathbf{1} = 0$ we have $x_j(t) \equiv 1, j \neq i_0$. Thus, (B.1) does not hold.

II) $\mathcal{G}$ does not have any isolated node, but has at least two source nodes i_1, i_2. Take $x_{i_1}(0) = 0, x_{i_2}(0) = 1$. Then, applying protocol (3), similar to I), we have $x_{i_1}(t) \equiv 0 \neq 1 \equiv x_{i_2}(t)$. Thus, (B.1) does not hold.

III) $\mathcal{G}$ does not contain isolated and contain at most one source node, but can be divided into two subgraphs $\mathcal{G}_1 = \{\mathcal{V}_1, \mathcal{E}_1, \mathcal{A}_1\}$ and $\mathcal{G}_2 = \{\mathcal{V}_2, \mathcal{E}_2, \mathcal{A}_2\}$, satisfying $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{V}_1 \cap \mathcal{V}_2 = \Phi, \mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2, \mathcal{E}_1 \cap \mathcal{E}_2 = \Phi$. Without loss of generality, suppose that $\mathcal{V}_1 = \{1, 2, \ldots, |\mathcal{V}_1|\}$, $\mathcal{V}_2 = \{|\mathcal{V}_1| + 1, \ldots, |\mathcal{V}_1| + |\mathcal{V}_2|\}$, $\mathcal{A}_\mathcal{G} = diag(\mathcal{A}_1, \mathcal{A}_2)$ is a diagonal block matrix. Then, applying protocol (3) leads to

$$\begin{cases} X_1(t+1) = \left(I_{|\mathcal{V}_1|} - a(t)\widetilde{\mathcal{L}}_1\right) X_1(t), \\ X_2(t+1) = \left(I_{|\mathcal{V}_2|} - a(t)\widetilde{\mathcal{L}}_2\right) X_2(t) \end{cases}$$

where $X_1(t), X_2(t)$ are the states of the nodes in $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively; $\widetilde{\mathcal{L}}_1, \widetilde{\mathcal{L}}_2$ are the Laplacian matrices of $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively. Take $x_i(0) = 0, i \in \mathcal{V}_1, x_j(0) = 1$, $j \in \mathcal{V}_2$. Then, similar to I), we have $x_i(t) \equiv 0 \neq 1 \equiv x_j(t), i \in \mathcal{V}_1, j \in \mathcal{V}_2$. Thus, (B.1) does not hold. □

Lemma B.2: Apply the protocol (3) to the system (1), (2). If for any $M > 0$, there is $t_0 \geq M$ such that $P\{\|(I_N - J)D_\mathcal{G} W(t_0)\| > 0\} > 0$, then for any given $K \geq 0$, there is $t_1 \geq K$ such that $E[V(t_1)] > 0$.

Proof: By contradiction, suppose there was a $K_0 > 0$ such that for all $t \geq K_0$, $E[V(t)] = 0$. Then, there would be $\delta(t) = 0$ a.s., $\forall\, t \geq K_0$, which together with (7) implies that $P\{\|(I_N - J)D_\mathcal{G} W(t)\| = 0\} = 1, \forall\, t \geq K_0$. This contradicts the condition of the lemma. Thus, the lemma is true. □

Proof of Theorem 3.4: We need only to show that none of the following four cases is true:

I) Assumption A1) does not hold.
II) Assumption A2) does not hold.
III) Under Assumption A1), $\sum_{t=0}^{\infty} a^2(t) = \infty$.
IV) Under Assumption A1)

$$\sum_{t=0}^{\infty} a^2(t) < \infty, \quad \sum_{t=0}^{\infty} a(t) < \infty.$$

By $W = \{W(t) = 0, t = 0, 1, \ldots\} \in \mathcal{S}$ and Lemma B.1, it is clear that neither I) nor II) is true. So, it suffices to show that neither III) nor IV) is true.

Step 1: Prove that III) is not true.

Suppose that there is at least a node which is not a source node, that is, there is $i_0 > 0$, $i_0 \neq j_0 > 0$, such that $a_{i_0 j_0} > 0$. Without loss of generality, suppose $i_0 = 1$, $j_0 = 2$. Let $W = \{[0, \widetilde{w}_{21}(t), \ldots, 0, \ldots, 0]^T, t = 0, 1, \ldots\}$, where $\{\widetilde{w}_{21}(t), t = 0, 1, \ldots\}$ is a standard white noise sequence. Then, one can see that $W \in \mathcal{S}$.

If $\sum_{t=0}^{\infty} a^2(t) = \infty$, then by (27) and the condition that $x_i(t)$, $i = 1, 2, \ldots, N$, converges in mean square to a common random variable with finite second-order moment, we would have that $(1/N)\mathbf{1}^T D_\mathcal{G} \sum_{t=0}^{n-1} a(t)W(t)$ converges in mean square to a random variable with finite second-order moment x_w, as $t \to \infty$. Furthermore, by Corollary 4.2.5 of [37], we get

$$\lim_{n\to\infty} E\left(\frac{1}{N}\mathbf{1}^T D_\mathcal{G} \sum_{t=0}^{n-1} a(t)W(t)\right)^2 = E(x_w)^2 < \infty. \tag{B.3}$$

On the other hand

$$\lim_{n\to\infty} E\left(\frac{1}{N}\mathbf{1}^T D_\mathcal{G} \sum_{t=0}^{n-1} a(t)W(t)\right)^2 = \frac{a_{12}^2}{N^2}\sum_{t=0}^{\infty} a^2(t) = \infty.$$

This contradicts (B.3). Thus, III) is not true.

Step 2: Prove that IV) is not true.

Similar to Step 1, suppose that $a_{12} > 0$, and that W is the same as in Step 1. Since $\sum_{t=0}^{\infty} a^2(t) < \infty, a(t) \to 0$ as $t \to \infty$. Notice that $E[\|(I_N - J)D_\mathcal{G} W(t)\|^2] = ((N-1)/N)a_{12}^2 > 0$, $\forall\, t \geq 0$. Then, by Lemma B.2, there is $t_0 > 0$ such that

$$E\left[V(t_0)\right] > 0 \tag{B.4}$$

$$0 \leq 2a(t)\lambda_{max}(\widehat{L}_\mathcal{G}) < \frac{\ln 2}{2}, \quad \forall t \geq t_0. \tag{B.5}$$

By (8), similar to (10) we have

$$E\left[V(t+1)\right] \geq \left(1 - 2\lambda_{max}(\widehat{L}_\mathcal{G})a(t)\right) E\left[V(t)\right], \quad \forall t \geq t_0.$$

This together with (B.5) and $1 - x \geq e^{-2x}$, $x \in [0, (\ln 2/2))$, implies

$$E\left[V(n)\right] \geq \exp\left\{-4\lambda_{max}(\widehat{L}_\mathcal{G})\sum_{t=t_0}^{n-1} a(t)\right\} E\left[V(t_0)\right], \ \forall n > t_0.$$

Thus, from (B.4) and $\sum_{t=0}^{\infty} a(t) < \infty$ we have

$$\liminf_{t\to\infty} E\left[V(t)\right] \geq \exp\left\{-4\lambda_{max}(\widehat{L}_\mathcal{G})\sum_{t=t_0}^{\infty} a(t)\right\} E[V(t_0)] > 0.$$

This contradicts the fact that $x_i(t)$, $i = 1, 2, \ldots, N$, converges in mean square to a common random variable. Thus, IV) is not true. $\square$

REFERENCES

[1] A. Sinha and D. Ghose, "Generalization of linear cyclic pursuit with application to rendezvous of multiple autonomous agents," *IEEE Trans. Autom. Control*, vol. 51, no. 11, pp. 1819–1824, Nov. 2006.

[2] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. 44th IEEE Conf. Decision Control Eur. Control Conf.*, Seville, Spain, Dec. 2005, pp. 8179–8184.

[3] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4th Int. Symp. Inform. Processing Sensor Networks*, 2005, pp. 63–70.

[4] N. Lynch, *Distributed Algorithms*. San Matero, CA: Morgan Kaufmann, 1996.

[5] R. Olfati-Saber and R. M. Murray, "Consensus problem in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[6] D. B. Kingston and R. W. Beard, "Discrete-time average-consensus under switching network topologies," in *Proc. Amer. Control Conf.*, Minneapolis, MN, Jun. 2006, pp. 3551–3556.

[7] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, 2004.

[8] G. M. Xie and L. Wang, "Consensus control for a class of networks of dynamic agents," *Int. J. Robust Nonlin. Control*, vol. 17, pp. 941–959, 2007.

[9] W. Ren and E. Atkins, "Distributed multi-vehicle coordinated control via local information exchange," *Int. J. Robust Nonlin. Control*, vol. 17, p. 1002C1033, 2007.

[10] Y. Hatano and M. Mesbahi, "Agreement over random network," *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1867–1872, Nov. 2005.

[11] C. W. Wu, "Synchronization and convergence of linear dynamics in random directed networks," *IEEE Trans. Autom. Control*, vol. 51, no. 7, pp. 1207–1210, Jul. 2006.

[12] F. Fagnani and S. Zampieri, "Average consensus with packet drop communication," in *Proc. 45th IEEE Conf. Decision Control*, San Diego, CA, Dec. 2006, pp. 1007–1012.

[13] A. T. 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, Mar. 2008.

[14] D. Bauso, L. Giarré, and R. Pesenti, "Non-linear protocols for optimal distributed consensus in networks of dynamics agents," *Syst. Control Lett.*, vol. 55, no. 11, pp. 918–928, 2006.

[15] T. Li and J. F. Zhang, "Decentralized tracking-Type games for multi-agent systems with coupled ARX models: Asymptotic nash equilibria," *Automatica*, vol. 44, no. 3, pp. 713–725, 2008.

[16] M. Huang, P. E. Caines, and R. P. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ-Nash equilibria," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1560–1571, Sep. 2007.

[17] T. Li and J. F. Zhang, "Asymptotically optimal decentralized control for large population stochastic multi-agent systems," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1643–1660, Jun. 2008.

[18] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192–1203, 2007.

[19] W. Ren, R. W. Beard, and E. M. Atkins, "A survy of consensus problems in multi-agent coordination," in *Proc. Amer. Control Conf.*, Portland, OR, Jun. 2005, pp. 1859–1864.

[20] M. Huang and J. H. Manton, "Coordination and consensus of networked agents with noisy measurement: Stochastic algorithms and asymptotic behavior," *SIAM J. Control Optim.*, vol. 48, no. 1, pp. 134–161, 2009.

[21] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri, "Communication constraints in coordinated consensus problems," in *Proc. Amer. Control Conf.*, Minneapolis, MN, Jun. 2006, pp. 4189–4194.

[22] W. Ren, R. W. Beard, and D. B. Kingston, "Multi-agent Kalman consensus with relative uncertainty," in *Proc. Amer. Control Conf.*, Portland, OR, Jun. 2005, pp. 1865–1870.

[23] D. B. Kingston, W. Ren, and R. W. Beard, "Consensus algorithm are input-to-state stable," in *Proc. Amer. Control Conf.*, Portland, OR, Jun. 2005, pp. 1686–1690.

[24] L. Xiao, S. Boyd, and S. J. Kim, "Distributed average consensus with least-mean square deviation," *J. Paralell Distrib. Comp.*, vol. 67, no. 1, pp. 33–46, 2007.

[25] T. Li, "Asymptotically unbiased average consensus under measurement noises and fixed topologies," in *Proc. 17th IFAC World Congress*, Seoul, Korea, Jul. 2008, pp. 2867–2873.

[26] T. Li and J. F. Zhang, "Mean square average consensus under measurement noises and fixed topologies: Necessary and sufficient conditions," *Automatica*, vol. 45, no. 8, pp. 1929–1936, 2009.

[27] L. Moreau, "Stability of multi-agent systems with dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169–182, Feb. 2005.

[28] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *Proc. 44th IEEE Conf. Decision Control Eur. Control Conf.*, Seville, Spain, Dec. 2005, pp. 2996–3000.

[29] 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, May 2005.

[30] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. AC-31, no. 9, pp. 803–812, Sep. 1986.

[31] A. Olshevsky and J. N. Tsitsiklis, "Convergence rates in distributed consensus and averaging," in *Proc. 45th IEEE Conf. Decision Control*, San Diego, CA, 2006, pp. 3387–3392.

[32] D. P. Bertsekas and J. N. Tsitsiklis, "Comments on 'coordination of groups of mobile autonomous agents using nearest neighbor rules'," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 968–969, May 2007.

[33] A. Jadbabaie, J. Lin, and S. M. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003.

[34] L. Moreau and S. G. Belgium, "Stability of continuous-time distributed consensus algorithms," in *Proc. 43rd IEEE Conf. Decision Control*, Atlantis, Paradise Island, Bahamas, Dec. 2004, pp. 3998–4003.

[35] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York: Springer, 2001.

[36] R. B. Ash, *Real Analysis and Probability*. New York: Academic Press, 1972.

[37] Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, 3rd ed. New York: Springer-Verlag, 1997.

[38] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

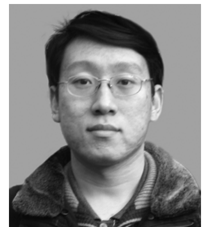
[39] J. Neveu, *Discrete-Parameter Martingales*. Amsterdam, The Netherlands: North-Holland, 1975.

[40] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[41] M. Cao, D. A. Spielman, and A. S. Morse, "A lower bound on convergence of a distributed network consensus algorithm," in *Proc. 44th IEEE Conf. Decision Control, Eur. Control Conf.*, Seville, Spain, Dec. 2005, pp. 2356–2361.

[42] M. Huang and J. H. Manton, "Stochastic consensus seeking with measurement noise: Convergence and asymptotic normality," in *Proc. Amer. Control Conf.*, Seattle, WA, Jun. 2008, pp. 1337–1342.

[43] B. T. Polyak, *Introduction to Optimization*. New York: Optimization Software, Inc., 1987.



Tao Li (M'09) was born in Tianjin, China, in 1981. He received the B.S. degree in automation from Nankai University, Tianjin, China, in 2004, and the Ph.D. degree in systems theory from the Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences (CAS), Beijing, China, in 2009.

From February 2008 to January 2009, he was with the School of Electrical and Electronic Engineering, Nanyang Technology University, Singapore, as a Research Assistant. Since July 2009, he has been with AMSS, CAS, where he is a Research Associate. His current research interests include system modeling, stochastic systems, and multi-agent systems.

Dr. Li received the Best Paper Award of the 7th Asian Control Conference in 2009.

Ji-Feng Zhang (M'92–SM'97) was born in Shandong, China, in 1963. He received the B.S. degree in mathematics from Shandong University, Shandong, in 1985, and the Ph.D. degree from the Institute of Systems Science (ISS), Chinese Academy of Sciences (CAS), Beijing, in 1991.

Since 1985, he has been with ISS, CAS, where he is now a Professor of Academy of Mathematics and Systems Science, the Vice-Director of the ISS. He now serves as Deputy Editor-in-Chief of *Acta Automatica Sinica*, *Journal of Systems Science and Mathematical Sciences*, and *Control Theory & Applications*, Managing Editor of the *Journal of Systems Science and Complexity*, and Associate Editor of the *SIAM Journal on Control and Optimization* among others. His current research interests include system modeling and identification, adaptive control, stochastic systems, and multi-agent systems.

Dr. Zhang received the Distinguished Young Scholar Fund from National Natural Science Foundation of China in 1997, the First Prize of the Young Scientist Award of CAS in 1995, the Outstanding Advisor Award of CAS in 2007, 2008, and 2009, respectively, and serves as the Director of the Technical Committee on Control Theory, Chinese Association of Automation. He was an Associate Editor with the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.