

Distributed Consensus With Limited Communication Data Rate

Tao Li, *Member, IEEE*, Minyue Fu, *Fellow, IEEE*, Lihua Xie, *Fellow, IEEE*, and Ji-Feng Zhang, *Senior Member, IEEE*

Abstract—Communication data rate and energy constraints are important factors which have to be considered when investigating distributed coordination of multi-agent networks. Although many proposed average-consensus protocols are available, a fundamental theoretic problem remains open, namely, how many bits of information are necessary for each pair of adjacent agents to exchange at each time step to ensure average consensus? In this paper, we consider average-consensus control of undirected networks of discrete-time first-order agents under communication constraints. Each agent has a real-valued state but can only exchange symbolic data with its neighbors. A distributed protocol is proposed based on dynamic encoding and decoding. It is proved that under the protocol designed, for a connected network, average consensus can be achieved with an exponential convergence rate based on merely one bit information exchange between each pair of adjacent agents at each time step. An explicit form of the asymptotic convergence rate is given. It is shown that as the number of agents increases, the asymptotic convergence rate is related to the scale of the network, the number of quantization levels and the ratio of the second smallest eigenvalue to the largest eigenvalue of the Laplacian of the communication graph. We also give a performance index to characterize the total communication energy to achieve average consensus and show that the minimization of the communication energy leads to a tradeoff between the convergence rate and the number of quantization levels.

Index Terms—Average-consensus, communication energy, data rate, distributed consensus, distributed coordination, distributed estimation, multi-agent systems, quantization, sensor network.

I. INTRODUCTION

DISTRIBUTED consensus and average-consensus problems have attracted great interests in recent years. Many distributed control and estimation strategies are designed based

Manuscript received November 14, 2008; revised September 24, 2009, October 09, 2009, and April 15, 2010; accepted April 15, 2010. Date of publication June 07, 2010; date of current version February 09, 2011. This work was supported by Singapore Millennium Foundation, Agency for Science, Technology and Research under Grant SERC 052 101 0037, National Natural Science Foundation of China (NSFC) under Grants 60828006, 60821091, 60934006, 61004029, NSFC-Guangdong Joint Foundation (U0735003), Major State Basic Research Development Program of China under Grant 2009cb320600 and the Chinese National Laboratory of Space Intelligent Control. Recommended by Associate Editor M. Prandini.

T. Li is 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. He is also with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: litao@amss.ac.cn).

M. Fu is with the School of EE&CS, University of Newcastle, Newcastle, Australia (e-mail: minyue.fu@newcastle.edu.au).

L. Xie is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: elhxie@ntu.edu.sg).

J. F. Zhang is 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: 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.2010.2052384

on consensus algorithms ([1], [2]). The study of distributed consensus is closely related to an essential issue of the complexity science: how can local communications and cooperations among individuals lead to certain desirable global behaviors?

The average-consensus problem originated from distributed computation and decision-making, which usually means to design a network protocol such that, as time goes on, the states of all agents asymptotically reach the average of the initial states. In the 1980s, Tsitsiklis [3] and Tsitsiklis *et al.* [4] studied the consensus problem and proposed weighted-average protocols. Thereafter, Jadbabaie [5] studied similar protocols motivated by biological group behaviors and stirred the excitement of the research on distributed cooperative control in the control community. From then on, many valuable results ([6]–[10]) have appeared. However, most of the works in the literature assume that each agent can obtain the exact state information of its neighbors through local communications. When the states of agents are real-valued, this assumption is equivalent to the requirement that the communication channels between agents have unlimited capacity (bandwidth). It is well known that in real digital networks, communication channels have a finite channel capacity, thus, at each time step, agents can only transmit a finite amount of information to their neighbors. The communication between different agents can be viewed as such a process: At each time step, the sender encodes the quantized state and sends out the code. When the neighbors receive the code, they use a decoding algorithm to obtain an estimate of the sender's state. Thus, quantization plays an important role in information exchange among agents. Therefore, consensus problems based on quantized communication become interesting and more meaningful.

Kashyap *et al.* ([11], [12]) and Nedic *et al.* ([13]) designed average-consensus algorithms based on the assumption that the states of agents are integer-valued. These algorithms can drive each agent to some integer approximation of the average of the initial states. Carli *et al.* and Frasca *et al.* ([14]–[17]) and Kar and Moura ([18]) studied the average-consensus problem with real-valued states based on quantized communication. Carli *et al.* ([14]–[16]) proposed algorithms with a uniform quantizer of infinite levels to ensure the boundness of the consensus error and gave an error bound in terms of the system parameters. Furthermore, an algorithm based on dynamic quantization to ensure average consensus was developed in Carli *et al.* ([17]). The number of quantization levels, however, will increase as the number of agents increases no matter how the control parameters are chosen. Kar and Moura ([18]) added a random dither to the sensor state before quantization to make the quantization error a “white” noise. By a distributed stochastic approximation method ([19]–[23]) and pathwise analysis, it was proved that, for an infinite-level quantizer, consensus can be achieved with

probability 1, while for a finite-level quantizer, the states of all agents will eventually enter into a small neighborhood of the average of the initial states with high probability. More relevant results with dithered quantizers can be found in [24] and [25].

It is well known that for feedback control of single-agent systems with communication constraints, Tatikonda and Mitter ([26]) and Nair and Evans ([27]) gave the minimal bit rate (channel capacity) for stabilizing a linear time-invariant system and the case with a logarithmic quantizer was considered by Elia and Mitter ([28]) and Fu and Xie ([29]). For distributed cooperative control problems of multi-agent systems under communication constraints, though various kinds of algorithms have been proposed as mentioned above, some fundamental theoretic problems remain open. For example, to achieve consensus of the whole network, how many bits of information does each pair of adjacent agents need to exchange at each time step?

In this paper, we consider the average-consensus control for discrete-time first-order undirected networks under communication constraints. Each agent has a real-valued state but can only exchange symbolic data with its neighbors. The communication between agents is based on dynamic encoding and decoding with finite-level quantization. We design a distributed protocol with error compensation. The protocol is characterized by three parameters: the control gain, the scaling function, and the number of quantization levels.

We show that if the network is connected, then for any given uniform quantizer with finite levels, the control gain and the scaling function can be chosen properly such that under the protocol designed, the average state of the whole network is preserved and the deviation between the state of each agent and the average state of the whole network converges to zero exponentially. In particular, it is shown that the control parameters can be properly chosen such that the average consensus can be achieved by using a single-bit quantizer. This indicates that no matter how large a network is, as long as it is connected, we can always design a distributed protocol to ensure asymptotic average consensus with an exponential convergence rate using merely one bit information exchange between each pair of adjacent agents at each time step.

We also analyze the relationship between the convergence rate and the number of quantization levels, and show that under the protocol designed, faster convergence requires more bits for quantization. We give a performance index to characterize the total communication energy to achieve average consensus and show that the minimization of the communication energy leads to a tradeoff between the convergence rate and the number of quantization levels.

Our proposed consensus algorithm has potential applications in the distributed estimation over sensor networks, where the number of nodes is often large. This motivates us to investigate asymptotic properties as the number of nodes N approaches infinity. We show that in some sense, the asymptotic convergence rate is $O(\exp\{-(KQ_N^2/(2\sqrt{N}))t\})$ when using a $(2K + 1)$ -level quantizer, where Q_N is the ratio of the second smallest eigenvalue (algebraic connectivity) to the largest eigenvalue (spectral radius) of the Laplacian matrix of the topology graph. The ratio Q_N is an important physical

quantity reflecting the synchronizability of a network ([30]). Our result shows that the convergence rate of distributed consensus depends not only on the connectivity but also on the synchronizability of the communication graph for a given finite communication data rate.

The remainder of this paper is organized as follows. In Section II, we present the model of the network, propose a distributed protocol, and formulate the problem to be investigated. In Section III, we prove that under the protocol designed and some mild conditions, average consensus can be achieved with an exponential convergence rate. Then we analyze the asymptotic performance as $N \rightarrow \infty$ and give an explicit form of the asymptotic convergence rate. In Section IV, we give a performance index to characterize the total communication energy cost to achieve average consensus. In Section V, we consider how to select the control parameters for an important class of complex networks called expander networks. In Section VI, we give some concluding remarks and future research topics.

The following notation will be used throughout this paper: $\mathbf{1}$ denotes a column vector with all ones. I denotes the identity matrix with an appropriate size. For a given set \mathcal{S} , the number of its elements is denoted by $|\mathcal{S}|$. For a given vector or matrix A , its transpose is denoted by A^T , its ∞ -norm is denoted by $\|A\|_\infty$, and its Euclidean norm is denoted by $\|A\|_2$, its trace is denoted by $\text{tr}(A)$. For a given positive number x , the natural logarithm of x is denoted by $\ln(x)$, the logarithm of x with base 2 is denoted by $\log_2(x)$, the maximum integer less than or equal to x is denoted by $\lfloor x \rfloor$; the minimum integer greater than or equal to x is denoted by $\lceil x \rceil$. For a given random variable X , the mathematical expectation of X is denoted by $E[X]$.

II. PROBLEM FORMULATION

A. Average-Consensus Problem

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

$$x_i(t+1) = x_i(t) + hu_i(t), \quad t = 0, 1, \dots, \quad i = 1, 2, \dots, N \quad (1)$$

where $x_i(t) \in \mathbb{R}$ and $u_i(t) \in \mathbb{R}$ are, respectively, the state and input of the i th agent, and h is the control gain. The communications between different agents are modeled as an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$, where $\mathcal{V} = \{1, 2, \dots, N\}$ is the set of nodes with i representing the i th agent, \mathcal{E} is the set of edges and $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ is the weighted adjacency matrix of \mathcal{G} . Note that \mathcal{A} is a symmetric matrix. An edge denoted by the pair (j, i) represents a communication channel from j to i and $(j, i) \in \mathcal{E}$ if and only if $(i, j) \in \mathcal{E}$. The neighborhood of the i th agent is denoted by $N_i = \{j \in \mathcal{V} | (i, j) \in \mathcal{E}\}$. For any $i, j \in \mathcal{V}$, $a_{ij} = a_{ji} \geq 0$, and $a_{ij} > 0$ if and only if $j \in N_i$. Also, $\text{deg}_i = \sum_{j=1}^N a_{ij}$ is called the degree of i , and $d^* = \max_i \text{deg}_i$ is called the degree of \mathcal{G} . A graph \mathcal{G} is called d -regular, if $|N_i| = d$, $i = 1, 2, \dots, N$, where $d > 0$ is a constant and $a_{ij} = 1$, $(i, j) \in \mathcal{E}$. For a d -regular graph, the degree of all nodes is d . The Laplacian matrix of \mathcal{G} is defined as $\mathcal{L} = \mathcal{D} - \mathcal{A}$, where $\mathcal{D} = \text{diag}(\text{deg}_1, \dots, \text{deg}_N)$. The Laplacian matrix \mathcal{L} is a symmetric positive semidefinite matrix and its eigenvalues in an ascending order are denoted by

$0 = \lambda_1(\mathcal{L}) \leq \lambda_2(\mathcal{L}) \leq \dots \leq \lambda_N(\mathcal{L})$, where $\lambda_N(\mathcal{L})$ is the spectral radius of \mathcal{L} and $\lambda_2(\mathcal{L})$ is called the algebraic connectivity of \mathcal{G} ([31]). A sequence of edges $(i_1, i_2), (i_2, i_3), \dots, (i_{k-1}, i_k)$ is called a path from node i_1 to node i_k . The graph \mathcal{G} is called a connected graph if for any $i, j \in \mathcal{V}$, there is a path from i to j .

The dynamic system (1) together with the communication graph \mathcal{G} is usually called a dynamic network ([6]). A group of controls $\mathcal{U} = \{u_i, i = 1, 2, \dots, N\}$ is called a distributed protocol if for all i , $u_i(t)$ only depends on $x_i(s)$ and $x_j(s)$, $j \in N_i$, $s \leq t$. The so-called average consensus control means designing a distributed protocol for the dynamic network, such that for any initial values $x_1(0), \dots, x_N(0)$, the states of all agents converge to $(1/N) \sum_{j=1}^N x_j(0)$ as $t \rightarrow \infty$. That is, $(1/N) \sum_{j=1}^N x_j(0)$ can be computed asymptotically in a distributed way.

B. Protocol Design

In [6], a weighted-average protocol was proposed

$$u_i(t) = \sum_{j \in N_i} a_{ij} (x_j(t) - x_i(t)), \quad t = 0, 1, \dots, \\ i = 1, 2, \dots, N. \quad (2)$$

In (2), the i th agent needs the exact state information of its neighbors. In this paper, we assume that the exact state information is not available, but only symbolic data can be exchanged between agents, and the communication channels are modeled as noiseless digital channels each with a pair of encoder and decoder. The encoder Φ_j of the j th agent is given by

$$\begin{cases} \xi_j(0) = 0, \\ \xi_j(t) = g(t-1)\Delta_j(t) + \xi_j(t-1), \\ \Delta_j(t) = q\left(\frac{1}{g(t-1)}(x_j(t) - \xi_j(t-1))\right), \end{cases} \quad t = 1, 2, \dots \quad (3)$$

where $\xi_j(t)$ is the internal state of Φ_j , and $\Delta_j(t)$, which is the output of Φ_j , is sent to the neighbors of the j th agent. Here, $q(\cdot)$ is a finite-level uniform quantizer, and $g(t) > 0$ is a scaling function.

The quantizer $q(\cdot): \mathbb{R} \rightarrow \Gamma$ is a map from \mathbb{R} to the set of quantized levels Γ . In this paper, we consider a finite-level uniform symmetric quantizer with

$$\Gamma = \{0, \pm i, i = 1, 2, \dots, K\}.$$

The number of quantization levels is $2K + 1$. The associated quantizer $q(\cdot)$ is given by

$$q(y) = \begin{cases} 0, & -1/2 < y < 1/2, \\ i, & \frac{2i-1}{2} \leq y < \frac{2i+1}{2}, \quad i = 1, 2, \dots, K-1, \\ K, & y \geq \frac{2K-1}{2}, \\ -q(-y), & y \leq -1/2. \end{cases} \quad (4)$$

Remark 1: The encoder Φ_j is a difference encoder with scaling, and $\xi_j(t)$ is a one-step predictor. In this difference coding algorithm, what is quantized at each time step is a ‘‘prediction error’’, $x_j(t) - \xi_j(t-1)$, rather than the state $x_j(t)$. Intuitively speaking, the amplitude of the prediction error is smaller than that of the state itself, so it can be represented by fewer bits. \square

Remark 2: If consensus is achieved asymptotically, then the prediction error $x_j(t) - \xi_j(t-1)$ tends to zero as $t \rightarrow \infty$. Therefore, intuitively the scaling function $g(t)$ should have the following properties. On one hand, $g(t)$ should decay gradually to make the quantizer persistently excited, such that the agents receive the information from their neighbors continuously. On the other hand, $g(t)$ should be large enough such that the quantizer will not be saturated. \square

Remark 3: The quantizer is implemented using the following strategy: when the output $\Delta_j(t)$ of the quantizer is zero, the j th agent does not send any information, so for a $(2K + 1)$ -level quantizer $q(\cdot)$, the communication channel (j, i) , $i \in N_j$ is required to be capable of transmitting $\lceil \log_2(2K) \rceil$ bits without error at each time step. In particular, the quantizer $q(\cdot)$ given by

$$q(y) = \begin{cases} 0, & -1/2 < y < 1/2, \\ 1, & y \geq 1/2, \\ -1, & y \leq -1/2 \end{cases} \quad (5)$$

is a one-bit quantizer. \square

For each communication channel $(j, i) \in \mathcal{E}$, the i th agent receives $\Delta_j(t)$, and then uses the following decoder Ψ_{ji} to estimate $x_j(t)$:

$$\begin{cases} \hat{x}_{ji}(0) = 0, \\ \hat{x}_{ji}(t) = g(t-1)\Delta_j(t) + \hat{x}_{ji}(t-1), \end{cases} \quad t = 1, 2, \dots \quad (6)$$

where $\hat{x}_{ji}(t)$ is the output of Ψ_{ji} .

Remark 4: The proposed encoder (3) and decoder (6) both require memory. Some related work on the exploitation of memory to speed up the convergence of distributed consensus can be found in [33]. \square

We propose a distributed protocol as

$$u_i(t) = \sum_{j \in N_i} a_{ij} (\hat{x}_{ji}(t) - \xi_i(t)), \quad t = 0, 1, \dots, \quad i = 1, 2, \dots, N. \quad (7)$$

Denote

$$\begin{aligned} X(t) &= [x_1(t), \dots, x_N(t)]^T, \\ \hat{X}(t) &= [\hat{\xi}_1(t), \dots, \hat{\xi}_N(t)]^T, \\ e(t) &= X(t) - \hat{X}(t), \\ \delta(t) &= X(t) - J_N X(t) \end{aligned} \quad (8)$$

where $J_N = (1/N)\mathbf{1}\mathbf{1}^T$.

Remark 5: If $\xi_i(t)$ is replaced by $x_i(t)$, then the protocol (7) becomes

$$u_i(t) = \sum_{j \in N_i} a_{ij} (\hat{x}_{ji}(t) - x_i(t)), \quad t = 0, 1, \dots, \\ i = 1, \dots, N. \quad (9)$$

The protocol (9) is a natural extension of the protocol (2) to the case with quantized communications and it has some computational advantage over the protocol (7). One may wonder why we use the protocol (7) rather than the protocol (9). We give some explanations below.

From (3) and (6), it follows that:

$$\hat{x}_{ji}(t) = \xi_j(t), \quad t = 0, 1, \dots, \quad i \in N_j, \quad j = 1, 2, \dots, N. \quad (10)$$

Thus, the internal state $\xi_j(t)$ of encoder Φ_j is equal to the estimates of $x_j(t)$ by its neighbors. By the symmetry of \mathcal{A} and (10), the protocol (7) can be rewritten as

$$\begin{aligned} u_i(t) &= \sum_{j \in N_i} a_{ij} [x_j(t) - x_i(t) \\ &\quad - (x_j(t) - \hat{x}_{ji}(t)) + (x_i(t) - \xi_i(t))] \\ &= \sum_{j \in N_i} a_{ij} [x_j(t) - x_i(t)] - \sum_{j \in N_i} a_{ij} [x_j(t) - \hat{x}_{ji}(t)] \\ &\quad + \sum_{j \in N_i} a_{ji} [x_i(t) - \hat{x}_{ij}(t)]. \end{aligned} \quad (11)$$

It can be seen that, in our protocol (7), the control input of the i th agent consists of three terms. The first term, $\sum_{j \in N_i} a_{ij} [x_j(t) - x_i(t)]$, which is just the control input of the protocol (2), plays the main role. The second term, $-\sum_{j \in N_i} a_{ij} [x_j(t) - \hat{x}_{ji}(t)]$, represents the weighted sum of estimation errors for the neighbors' states $x_j(t)$, $j \in N_i$. The last term $\sum_{j \in N_i} a_{ji} [x_i(t) - \hat{x}_{ij}(t)]$ is the weighted sum of estimation errors for $x_i(t)$ by the neighbors.

The last term in (11), which we call an error-compensation term, plays an important role in our protocol. Substituting the protocol (3), (6) and (7) into the system (1) leads to

$$\begin{cases} X(t+1) = (I - h\mathcal{L})X(t) + h\mathcal{L}e(t), \\ \hat{X}(t+1) = g(t)Q\left(\frac{X(t+1) - \hat{X}(t)}{g(t)}\right) + \hat{X}(t) \end{cases} \quad (12)$$

where $Q([y_1, \dots, y_N]^T) = [q(y_1), \dots, q(y_N)]^T$. From the above, noting that $J_N \mathcal{L} = 0$, we have

$$\frac{1}{N} \sum_{j=1}^N x_j(t+1) = \frac{1}{N} \sum_{j=1}^N x_j(t), \quad t = 0, 1, \dots \quad (13)$$

It can be seen that the closed-loop system preserves the average state under the protocol (7). If the error-compensation term is removed, then by (11), the protocol (7) reduces to the protocol (9), and the closed-loop system becomes

$$X(t+1) = (I - h\mathcal{L})X(t) - h\mathcal{A}e(t). \quad (14)$$

Generally speaking, the closed-loop system (14) does not preserve the average state, and worse still, it can be shown that the closed-loop system (14) may be divergent if $e(t)$ is a bounded white noise ([34]). That is why we use the protocol (7) rather than the protocol (9). Indeed, a similar error compensation approach has been considered in [14] and [15]. \square

III. FINITE-LEVEL QUANTIZED CONSENSUS

For the protocol designed and the resulting closed-loop system (12), the following questions are naturally put forward:

Questions: Can the whole network achieve consensus with finite-level quantized communication? If so, how many bits are necessary for each pair of adjacent agents to exchange at each time step? What is the relationship between the convergence rate and the control gain, the scaling function and the number of quantization levels?

In this section, we will answer the above questions.

A. Convergence Analysis

We make the following assumptions:

A1) \mathcal{G} is connected.

A2) $\max_i |x_i(0)| \leq C_x$, $\max_i |\delta_i(0)| \leq C_\delta$, where C_x and C_δ are known nonnegative constants.

To get the main results (Theorems 3.1 and 3.3), we need the following lemma.

Lemma 3.1: If Assumption A1) holds and $h < 2/\lambda_N(\mathcal{L})$, then $\rho_h < 1$, where

$$\rho_h = \max_{2 \leq i \leq N} |1 - h\lambda_i(\mathcal{L})|. \quad (15)$$

Furthermore, if $h < 2/(\lambda_2(\mathcal{L}) + \lambda_N(\mathcal{L}))$, then $\rho_h = 1 - h\lambda_2(\mathcal{L})$.

Proof: It follows directly from the properties of graph Laplacian ([32]), and the details are omitted. \square

Theorem 3.1: Suppose Assumptions A1)–A2) hold. For any given $h \in (0, 2/\lambda_N(\mathcal{L}))$ and $\gamma \in (\rho_h, 1)$, let

$$K_1(h, \gamma) = \left\lfloor M_1(h, \gamma) - \frac{1}{2} \right\rfloor + 1 \quad (16)$$

$$M_1(h, \gamma) = \frac{\sqrt{N}h^2\lambda_N^2(\mathcal{L})}{2\gamma(\gamma - \rho_h)} + \frac{1 + 2hd^*}{2\gamma} \quad (17)$$

and for any given $K \geq K_1(h, \gamma)$, let

$$g_0 > \max \left\{ \frac{C_x}{K + \frac{1}{2}}, \frac{2(\gamma - \rho_h)(C_\delta\gamma + hC_x\lambda_N(\mathcal{L}))}{h\lambda_N(\mathcal{L})} \right\}. \quad (18)$$

Then under the protocol given by (3), (6) and (7) with the $(2K + 1)$ -level uniform quantizer (4) and the scaling function $g(t) = g_0\gamma^t$, the closed-loop system (12) satisfies

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{N} \sum_{j=1}^N x_j(0), \quad i = 1, 2, \dots, N \quad (19)$$

and

$$\limsup_{t \rightarrow \infty} \frac{\|\delta(t)\|_2}{\gamma^t} \leq \frac{\sqrt{N}g_0h\lambda_N(\mathcal{L})}{2\gamma(\gamma - \rho_h)} \quad (20)$$

where $\delta(t)$ is the consensus error defined by (8). Furthermore, $r_{\text{asym}} \leq \gamma$, where

$$r_{\text{asym}} = \sup_{X(0) \neq J_N X(0)} \lim_{t \rightarrow \infty} \left(\frac{\|X(t) - J_N X(0)\|_2}{\|X(0) - J_N X(0)\|_2} \right)^{1/t} \quad (21)$$

which is defined as *convergence rate of average consensus* ([35]).

Proof: Let g_0 be chosen as in (18). By (12), noting that $\mathcal{L}J_N = J_N \mathcal{L} = 0$ and

$$X(t+1) - \hat{X}(t) = (I + h\mathcal{L})e(t) - h\mathcal{L}\delta(t)$$

we have

$$\begin{cases} \delta(t+1) = (I - h\mathcal{L})\delta(t) + h\mathcal{L}e(t), \\ e(t+1) = (I + h\mathcal{L})e(t) - h\mathcal{L}\delta(t) \\ \quad - g(t)Q\left(\frac{1}{g(t)}((I + h\mathcal{L})e(t) - h\mathcal{L}\delta(t))\right). \end{cases} \quad (22)$$

Let

$$w(t) = \frac{1}{g(t)}\delta(t) \quad (23)$$

$$z(t) = \frac{1}{g(t)}e(t). \quad (24)$$

Then from (22) and $g(t) = g_0\gamma^t$, we get

$$\begin{aligned}\gamma w(t+1) &= (I - h\mathcal{L})w(t) + h\mathcal{L}z(t), \\ \gamma z(t+1) &= (I + h\mathcal{L})z(t) - h\mathcal{L}w(t) \\ &\quad - Q((I + h\mathcal{L})z(t) - h\mathcal{L}w(t))\end{aligned}\quad (25)$$

which can be written as

$$\begin{aligned}w(t+1) &= \gamma^{-1}(I - h\mathcal{L})w(t) + \gamma^{-1}h\mathcal{L}z(t), \\ z(t+1) &= \gamma^{-1}\Delta(t)\end{aligned}\quad (26)$$

where

$$\Delta(t) = (I + h\mathcal{L})z(t) - h\mathcal{L}w(t) - Q((I + h\mathcal{L})z(t) - h\mathcal{L}w(t)).\quad (27)$$

Now we claim that if a $(2K+1)$ -level uniform quantizer (4) with $K \geq K_1(h, \gamma)$ is applied, then the quantizer will never be saturated. Indeed, by (3), (6) and (10), we know that $\hat{X}(0) = 0$. Using Assumption A2) and (18), we have

$$\|z(0)\|_\infty \leq \frac{C_x}{g_0}\quad (28)$$

and

$$\|(I + h\mathcal{L})z(0) - h\mathcal{L}w(0)\|_\infty = \left\| \frac{X(0)}{g_0} \right\|_\infty \leq \frac{C_x}{g_0} < K + \frac{1}{2}.$$

Hence, when $t = 0$, the quantizer is unsaturated. For any given nonnegative integer k , suppose that when $t = 0, 1, \dots, k$, the quantizer is not saturated. Then, we have

$$\sup_{0 \leq t \leq k} \|\Delta(t)\|_\infty \leq \frac{1}{2}.$$

This together with (26) gives

$$\sup_{1 \leq t \leq k+1} \|z(t)\|_\infty \leq \frac{1}{2\gamma}.\quad (29)$$

Since \mathcal{L} is symmetric, we can take the unitary matrix $T = [(1/\sqrt{N}), \phi_2, \dots, \phi_N]$ defined by $\phi_i^T \mathcal{L} = \lambda_i \phi_i^T$, $i = 2, \dots, N$. Let $\tilde{w}(t) = T^{-1}w(t) = T^T w(t)$ and decompose $\tilde{w}(t) = [\tilde{w}_1(t), \tilde{w}_2^T(t)]^T$ with a scalar $\tilde{w}_1(t)$. Then $\tilde{w}_1(t) = 0$

$$\tilde{w}_2(t+1) = \text{diag}\left(\frac{1 - h\lambda_2(\mathcal{L})}{\gamma}, \dots, \frac{1 - h\lambda_N(\mathcal{L})}{\gamma}\right) \tilde{w}_2(t) + \gamma^{-1}h\phi^T \mathcal{L}z(t), \quad t = 0, 1, \dots\quad (30)$$

where $\phi = [\phi_2, \dots, \phi_N]$.

Denote $\tilde{P}_{\gamma, h} = \text{diag}((1 - h\lambda_2(\mathcal{L}))/\gamma, \dots, (1 - h\lambda_N(\mathcal{L}))/\gamma)$. From (30), it follows that:

$$\begin{aligned}\tilde{w}_2(t+1) &= [\tilde{P}_{\gamma, h}]^{t+1} \tilde{w}_2(0) + \gamma^{-1}h[\tilde{P}_{\gamma, h}]^t \phi^T \mathcal{L}z(0) \\ &\quad + \gamma^{-1}h \sum_{i=0}^{t-1} [\tilde{P}_{\gamma, h}]^i \phi^T \mathcal{L}z(t-i), \quad t = 0, 1, \dots\end{aligned}\quad (31)$$

where $\sum_{i=0}^l (\cdot)$ is defined as 0, when $l < 0$. Thus, for the time instant $t = k+1$, noting that $\tilde{w}_2(t) = \phi^T w(t)$ and $w(t) = \phi \tilde{w}_2(t)$, we have

$$\begin{aligned}w(k+1) &= \phi [\tilde{P}_{\gamma, h}]^{k+1} \phi^T w(0) + \gamma^{-1}h\phi [\tilde{P}_{\gamma, h}]^k \phi^T \mathcal{L}z(0) \\ &\quad + \gamma^{-1}h\phi \sum_{i=0}^{k-1} [\tilde{P}_{\gamma, h}]^i \phi^T \mathcal{L}z(k-i).\end{aligned}\quad (32)$$

Now we estimate the three terms on the right-hand side of the above equation, separately. For the first term, noting that

$\|\phi\|_2 = 1$ and $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{N}\|x\|_\infty$, for any N dimensional vector x , we have

$$\begin{aligned}\left\| \phi [\tilde{P}_{\gamma, h}]^{k+1} \phi^T w(0) \right\|_2 &\leq \left\| \phi [\tilde{P}_{\gamma, h}]^{k+1} \phi^T \right\|_2 \|w(0)\|_2 \\ &\leq \frac{\sqrt{N} \|\delta(0)\|_\infty}{g_0} \left(\frac{\rho_h}{\gamma} \right)^{k+1} \\ &\leq \frac{\sqrt{N} \|\delta(0)\|_\infty}{g_0} \left(\frac{\rho_h}{\gamma} \right)^k.\end{aligned}\quad (33)$$

For the second term in (32), using (28), $\gamma \in (\rho_h, 1)$ and $\|\mathcal{L}\|_2 = \lambda_N(\mathcal{L})$, we have

$$\left\| \gamma^{-1}h\phi [\tilde{P}_{\gamma, h}]^k \phi^T \mathcal{L}z(0) \right\|_2 \leq \frac{\sqrt{N}hC_x\lambda_N(\mathcal{L})}{g_0\gamma} \left(\frac{\rho_h}{\gamma} \right)^k.\quad (34)$$

Similarly, for the last term in (32), by (29) and noting that

$$\left\| \sum_{i=0}^{k-1} [\tilde{P}_{\gamma, h}]^i \right\|_2 \leq \sum_{i=0}^{k-1} \|\tilde{P}_{\gamma, h}\|_2^i = \frac{1 - \left(\frac{\rho_h}{\gamma} \right)^k}{1 - \frac{\rho_h}{\gamma}}$$

we have

$$\begin{aligned}\left\| \gamma^{-1}h\phi \sum_{i=0}^{k-1} [\tilde{P}_{\gamma, h}]^i \phi^T \mathcal{L}z(k-i) \right\|_2 \\ \leq \frac{\sqrt{N}h\lambda_N(\mathcal{L})}{2\gamma(\gamma - \rho_h)} \left[1 - \left(\frac{\rho_h}{\gamma} \right)^k \right].\end{aligned}\quad (35)$$

Then by Assumption A1), $\gamma \in (\rho_h, 1)$ and (32)–(35), we have

$$\|w(k+1)\|_2 \leq \max \left\{ \frac{\sqrt{N}C_\delta\gamma + h\sqrt{N}C_x\lambda_N(\mathcal{L})}{g_0\gamma}, \frac{h\sqrt{N}\lambda_N(\mathcal{L})}{2\gamma(\gamma - \rho_h)} \right\}.\quad (36)$$

This together with (29) and (18) leads to

$$\begin{aligned}\|(I + h\mathcal{L})z(k+1) - h\mathcal{L}w(k+1)\|_\infty \\ \leq \|(I + h\mathcal{L})z(k+1)\|_\infty + h\|\mathcal{L}\|_2 \|w(k+1)\|_2 \\ \leq \|(I + h\mathcal{L})\|_\infty \|z(k+1)\|_\infty + h\|\mathcal{L}\|_2 \|w(k+1)\|_2 \\ \leq (1 + 2hd^*) \|z(k+1)\|_\infty + h\|\mathcal{L}\|_2 \|w(k+1)\|_2 \\ \leq \frac{1 + 2hd^*}{2\gamma} + h\lambda_N(\mathcal{L}) \\ \times \left[\max \left\{ \frac{\sqrt{N}C_\delta\gamma + h\sqrt{N}C_x\lambda_N(\mathcal{L})}{g_0\gamma}, \frac{\sqrt{N}h\lambda_N(\mathcal{L})}{2\gamma(\gamma - \rho_h)} \right\} \right] \\ = M_1(h, \gamma) < \left[M_1(h, \gamma) - \frac{1}{2} \right] + \frac{3}{2} \\ = K_1(h, \gamma) + \frac{1}{2} \leq K + \frac{1}{2}.\end{aligned}\quad (37)$$

So when $t = k+1$, the quantizer is also unsaturated. Therefore, by induction, we conclude that if a $(2K+1)$ -level uniform quantizer (4) with $K \geq K_1$ is applied, then it will never be saturated.

Noting that $\|w(0)\|_\infty \leq C_\delta/g_0$, by (36) and (18), we have

$$\sup_{t \geq 0} \|w(t)\|_\infty \leq \max \left\{ \frac{C_\delta}{g_0}, \frac{h\sqrt{N}\lambda_N(\mathcal{L})}{2\gamma(\gamma - \rho_h)} \right\} < \infty.\quad (38)$$

Thus, by the definition of $w(t)$ and $0 < \gamma < 1$, we get

$$\lim_{t \rightarrow \infty} \|\delta(t)\|_\infty = 0.$$

This together with (13) gives (19).

From (32)–(35), noting that $\delta(t) = g_0 \gamma^t w(t)$, we have

$$\begin{aligned} \|\delta(k+1)\|_2 &\leq \|\delta(0)\|_2 (\rho_h)^k + \sqrt{N} h \gamma^{-1} C_x \lambda_N(\mathcal{L}) (\rho_h)^k \\ &+ \frac{\sqrt{N} g_0 h \lambda_N(\mathcal{L})}{2\gamma(\gamma - \rho_h)} \gamma^k. \end{aligned}$$

This together with $\gamma > \rho_h$ leads to (20).

Similar to (32)–(35), we have

$$\begin{aligned} \frac{\|\delta(k+1)\|_2}{\|\delta(0)\|_2} &\leq \gamma (\rho_h)^k + \frac{h\sqrt{N} C_x \lambda_N(\mathcal{L})}{\|\delta(0)\|_2} (\rho_h)^k \\ &+ \frac{hg_0 \lambda_N(\mathcal{L}) \sqrt{N}}{2\|\delta(0)\|_2 (\gamma - \rho_h)} \gamma^k, \quad \forall \delta(0) \neq 0 \end{aligned}$$

which together with $\gamma > \rho_h$ leads to

$$\begin{aligned} \ln \left\{ \frac{\|\delta(k+1)\|_2}{\|\delta(0)\|_2} \right\} &\leq \ln \left\{ \frac{hg_0 \lambda_N(\mathcal{L}) \sqrt{N}}{2\|\delta(0)\|_2 (\gamma - \rho_h)} \gamma^k \right\} \\ &+ \ln(1 + o(1)), \quad k \rightarrow \infty \\ &= k \ln(\gamma) + O(1), \quad k \rightarrow \infty. \end{aligned}$$

Thus

$$\begin{aligned} &\lim_{k \rightarrow \infty} \left(\frac{\|\delta(k+1)\|_2}{\|\delta(0)\|_2} \right)^{\frac{1}{k+1}} \\ &= \lim_{k \rightarrow \infty} \exp \left\{ \frac{1}{k+1} \ln \left(\frac{\|\delta(k+1)\|_2}{\|\delta(0)\|_2} \right) \right\} \\ &\leq \exp \left\{ \lim_{k \rightarrow \infty} \frac{1}{k+1} (k \ln(\gamma) + O(1)) \right\} \\ &= \gamma, \quad \forall \delta(0) \neq 0. \end{aligned}$$

By this and the definition of r_{asym} , we get $r_{\text{asym}} \leq \gamma$. \square

Remark 6: Assumption A2) says that the upper bound of the initial states and the initial consensus error are both known. If the bound C_x of the initial states is known, then $\max_i |\delta_i(0)| \leq 2C_x$. Thus Assumption A2) holds if the following assumption holds:

A2') $\max_i |x_i(0)| \leq C_x$, where C_x is a known nonnegative constant.

But we use Assumption A2) rather than A2'), due to the following considerations: 1) In many cases, the bound of the initial consensus error may be much smaller than that of the initial states. 2) From the proof of Theorem 3.1, it is shown that the upper bound of the initial states and the initial consensus error have different impacts on the dynamic evolution of the scaled consensus error $w(t)$. If the upper bound of the initial consensus error is known *a priori*, then the associated estimates can be obtained more accurately. \square

Remark 7: Here, the scaling function $g(t)$ is designed offline. By Lemma A.2 and the property of Laplacian matrix, we have $d^* \leq \lambda_N(\mathcal{L}) \leq 2d^*$, then by $0 < \gamma - \rho_h < 1$, we know that (18) holds if

$$g_0 > \max \left\{ \frac{C_x}{K + \frac{1}{2}}, \frac{2(C_\delta + 2hd^*C_x)}{hd^*} \right\}.$$

By the above inequality, the choice of g_0 requires the knowledge of the upper bound of the initial states C_x and the degree of the communication graph. This is a conservative selection, and in practice, smaller g_0 may be available. \square

Remark 8: Theorem 3.1 says that by using a scaling function decaying exponentially and a $\lceil \log_2(2K_1(h, \gamma)) \rceil$ -bit uniform quantizer, the protocol (3), (6) and (7) can ensure average consensus to be achieved asymptotically. It is worth pointing out that for any given h and γ , the bit number $\lceil \log_2(2K_1(h, \gamma)) \rceil$ is a conservative estimate, and in practice, fewer bits may be required. However, the number $K_1(h, \gamma)$ gives us some intuitive clues on the relationship between the number of bits required and the control gain h and the scaling factor γ . \square

Remark 9: Theorem 3.1 gives an estimate for the convergence rate of the consensus. The smaller the γ , the faster the consensus error converges to zero. Note that γ can be made arbitrarily close to ρ_h , which is the convergence rate for the case with perfect communication ([35]). From Theorem 3.1, it is shown that a smaller γ , namely a faster convergence rate, requires more bits to be communicated, and when $\gamma \rightarrow \rho_h$, the required number of bits goes to infinity. \square

B. Probabilistic Analysis

From the proof of Theorem 3.1, it can be seen that the quantizers of all agents are unsaturated, then it is shown that the convergence of the consensus error $\delta(t)$ is no slower than $O(\gamma^t)$. This result is based on the worst case analysis, since the upper bound of the quantization error $\|\Delta(t)\|$, which is $1/2$, is used. If the quantization errors are modeled as stochastic noises uniformly distributed on $[-1/2, 1/2]$ as in [15] and [16], then we can show that the mean-square-root consensus error $\sqrt{E\|\delta(t)\|_2^2}$ converges to zero asymptotically, and the convergence rate is just $O(\gamma^t)$.

Denote

$$\begin{aligned} e_q(t) &= \frac{1}{g(t)} ((I + h\mathcal{L})e(t) - h\mathcal{L}\delta(t)) \\ &\quad - Q \left(\frac{1}{g(t)} ((I + h\mathcal{L})e(t) - h\mathcal{L}\delta(t)) \right) \end{aligned}$$

which is the vector with the i th component as the quantization error of agent i . Then the closed-loop system (22) can be written as

$$\begin{aligned} \delta(t+1) &= (I - h\mathcal{L})\delta(t) + h\mathcal{L}e(t), \\ e(t+1) &= g(t)e_q(t). \end{aligned} \quad (39)$$

For a probabilistic analysis, we introduce a stochastic model of (39)

$$\begin{aligned} \delta_s(t+1) &= (I - h\mathcal{L})\delta_s(t) + h\mathcal{L}e_s(t), \\ e_s(t+1) &= g(t)e_{q_s}(t) \end{aligned} \quad (40)$$

where $\{e_{q_s}(t), t = 0, 1, \dots\}$ are stochastic noises.

Theorem 3.2: For (40), let $\{e_{q_s}(t), t = 0, 1, \dots\}$ be an N dimensional random vector sequence with $E[e_{q_s}(t)] = 0$ and $E[e_{q_s}(t)e_{q_s}(t)^T] = (1/12)I$, $E[e_{q_s}(t)e_{q_s}(k)^T] = 0$, $t \neq k$. If \mathcal{L} is the Laplacian matrix of a connected graph,

$h \in (0, 2/\lambda_N(\mathcal{L}))$ and $g(t) = g_0\gamma^t$ with $\gamma \in (\rho_h, 1)$ and $g_0 > 0$, then

$$\liminf_{t \rightarrow \infty} \frac{\sqrt{E \left[\|\delta_s(t)\|_2^2 \right]}}{\gamma^t} \geq \frac{g_0 h \lambda_2(\mathcal{L})}{2\sqrt{3}\gamma\sqrt{\gamma^2 - \rho_h^2}} \quad (41)$$

$$\limsup_{t \rightarrow \infty} \frac{\sqrt{E \left[\|\delta_s(t)\|_2^2 \right]}}{\gamma^t} \leq \frac{\sqrt{N-1}g_0 h \lambda_N(\mathcal{L})}{2\sqrt{3}\gamma\sqrt{\gamma^2 - \rho_h^2}} \quad (42)$$

and $\bar{r}_{\text{asym}} = \gamma$, where

$$\bar{r}_{\text{asym}} = \sup_{X(0) \neq J_N X(0)} \lim_{t \rightarrow \infty} \left(\frac{\sqrt{E \left[\|X(t) - J_N X(0)\|_2^2 \right]}}{\|X(0) - J_N X(0)\|_2} \right)^{1/t}.$$

Proof: From (40), similar to (32), we have

$$\begin{aligned} \delta_s(k+1) &= \gamma^{k+1} \phi[\tilde{P}_{\gamma,h}]^{k+1} \phi^T \delta_s(0) \\ &\quad + g_0 \gamma^k h \phi[\tilde{P}_{\gamma,h}]^k \phi^T \mathcal{L} z_s(0) \\ &\quad + g_0 \gamma^{k-1} h \phi \sum_{i=0}^{k-1} (\tilde{P}_{\gamma,h})^i \phi^T \mathcal{L} e_{q_s}(k-1-i) \end{aligned}$$

where $z_s(t) = e_s(t)/g(t)$ and $\delta_s(0) = \delta(0)$.

In view of the properties of $e_{q_s}(t)$, it follows that:

$$\begin{aligned} E \left[\delta_s(k+1) \delta_s(k+1)^T \right] &= \gamma^{2(k+1)} \phi[\tilde{P}_{\gamma,h}]^{k+1} \phi^T \delta_s(0) \delta_s(0)^T \phi[\tilde{P}_{\gamma,h}]^{k+1} \phi^T \\ &\quad + g_0^2 \gamma^{2k} h^2 \phi[\tilde{P}_{\gamma,h}]^k \phi^T \mathcal{L} z_s(0) z_s(0)^T \mathcal{L} \phi[\tilde{P}_{\gamma,h}]^k \phi^T \\ &\quad + g_0 \gamma^{2k+1} h \phi[\tilde{P}_{\gamma,h}]^k \phi^T \mathcal{L} z_s(0) \delta_s(0)^T \phi[\tilde{P}_{\gamma,h}]^{k+1} \phi^T \\ &\quad + \frac{1}{12} g_0^2 \gamma^{2(k-1)} h^2 \\ &\quad \times \sum_{i=0}^{k-1} \phi[\tilde{P}_{\gamma,h}]^i \phi^T \mathcal{L}^2 \phi[\tilde{P}_{\gamma,h}]^i \phi^T. \end{aligned} \quad (43)$$

Noting that $\phi^T \phi = I$, we have

$$\begin{aligned} &\text{tr} \left[\phi[\tilde{P}_{\gamma,h}]^i \phi^T \mathcal{L}^2 \phi[\tilde{P}_{\gamma,h}]^i \phi^T \right] \\ &= \text{tr} \left[\mathcal{L} \phi[\tilde{P}_{\gamma,h}]^i \phi^T \phi[\tilde{P}_{\gamma,h}]^i \phi^T \mathcal{L} \right] \\ &= \text{tr} \left[\mathcal{L} \phi[\tilde{P}_{\gamma,h}]^i (\tilde{P}_{\gamma,h})^i \phi^T \mathcal{L} \right] \\ &= \text{tr} \left[(\tilde{P}_{\gamma,h})^i \phi^T \mathcal{L}^2 \phi[\tilde{P}_{\gamma,h}]^i \right] \\ &\geq \lambda_2^2(\mathcal{L}) \text{tr} \left[(\tilde{P}_{\gamma,h})^{2i} \right] \\ &\geq \lambda_2^2(\mathcal{L}) \left(\frac{\rho_h}{\gamma} \right)^{2i}. \end{aligned} \quad (44)$$

This together with (43) and $\|\tilde{P}_{\gamma,h}\|_2 = \rho_h/\gamma < 1$ leads to

$$\begin{aligned} \liminf_{t \rightarrow \infty} \frac{E \left[\|\delta_s(t+1)\|_2^2 \right]}{\gamma^{2(t+1)}} &\geq \frac{\lambda_2^2(\mathcal{L}) g_0^2 \gamma^{-4} h^2}{12} \sum_{i=0}^{\infty} \left(\frac{\rho_h}{\gamma} \right)^{2i} \\ &= \frac{g_0^2 h^2 \lambda_2^2(\mathcal{L})}{12\gamma^2(\gamma^2 - \rho_h^2)} \end{aligned}$$

which results in (41). Similar to (44), we have

$$\begin{aligned} &\text{tr} \left[\phi[\tilde{P}_{\gamma,h}]^i \phi^T \mathcal{L}^2 \phi[\tilde{P}_{\gamma,h}]^i \phi^T \right] \\ &\leq \lambda_N^2(\mathcal{L}) \text{tr} \left[\phi[\tilde{P}_{\gamma,h}]^i (\tilde{P}_{\gamma,h})^i \phi^T \right] \\ &= \lambda_N^2(\mathcal{L}) \text{tr} \left[(\tilde{P}_{\gamma,h})^{2i} \right] \\ &\leq (N-1) \lambda_N^2(\mathcal{L}) \left(\frac{\rho_h}{\gamma} \right)^{2i}. \end{aligned}$$

This together with (43) leads to

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{E \left[\|\delta_s(t+1)\|_2^2 \right]}{\gamma^{2(t+1)}} &\leq \frac{(N-1) \lambda_N^2(\mathcal{L}) g_0^2 \gamma^{-4} h^2}{12} \\ &\quad \times \sum_{i=0}^{\infty} \left(\frac{\rho_h}{\gamma} \right)^{2i} \\ &= \frac{(N-1) g_0^2 h^2 \lambda_N^2(\mathcal{L})}{12\gamma^2(\gamma^2 - \rho_h^2)} \end{aligned}$$

which gives (42). Further, from (41), (42) and the definition of \bar{r}_{asym} , we get $\bar{r}_{\text{asym}} = \gamma$. \square

C. Any-Rate Convergence

From Theorem 3.1, it can be seen that if the convergence rate γ is fixed (i.e., independent of the number of agents), then the number of quantization levels, $2K_1(h, \gamma) + 1$, will tend to infinity as $N \rightarrow \infty$. However, in practical applications, the bit rate is usually limited. To satisfy this requirement, we can use a fixed number of quantization levels at the cost of slower convergence. We have the following result.

Theorem 3.3: Suppose Assumptions A1–A2 hold. For any given $K \geq 1$, let

$$\begin{aligned} \Omega_K &= \left\{ (\alpha, \beta) \mid \alpha \in \left(0, \frac{2}{\lambda_2(\mathcal{L}) + \lambda_N(\mathcal{L})} \right), \right. \\ &\quad \left. \beta \in (\rho_\alpha, 1), \quad M_1(\alpha, \beta) < K + \frac{1}{2} \right\} \end{aligned} \quad (45)$$

where ρ_α is defined by (15) and $M_1(\alpha, \beta)$ is defined by (17). Then, (i) Ω_K is nonempty. (ii) For any $(h, \gamma) \in \Omega_K$, under the protocol given by (3), (6) and (7) with $g(t) = g_0\gamma^t$ and the $(2K+1)$ -level uniform quantizer (4), the closed-loop system (12) satisfies

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{N} \sum_{j=1}^N x_j(0), \quad i = 1, 2, \dots, N$$

where g_0 is a constant satisfying (18).

Proof:

(i) Noting that

$$\lim_{\alpha \rightarrow 0} \left[\frac{\sqrt{N} \alpha \lambda_N^2(\mathcal{L})}{2\lambda_2(\mathcal{L})} + \frac{1 + 2\alpha d^*}{2} \right] = \frac{1}{2}$$

we know that for any given $K \geq 1$, there exists $\alpha^* \in (0, 2/\lambda_2(\mathcal{L}) + \lambda_N(\mathcal{L}))$ such that

$$\frac{\sqrt{N} \alpha^* \lambda_N^2(\mathcal{L})}{2\lambda_2(\mathcal{L})} + \frac{1 + 2\alpha^* d^*}{2} < K + \frac{1}{2}. \quad (46)$$

By Lemma 3.1, it is known that $\rho_{\alpha^*} = 1 - \alpha^* \lambda_2(\mathcal{L}) < 1$. By this and (17), we get

$$\lim_{\gamma \rightarrow 1} M_1(\alpha^*, \gamma) = \frac{\sqrt{N} \alpha^* \lambda_N^2(\mathcal{L})}{2\lambda_2(\mathcal{L})} + \frac{1 + 2\alpha^* d^*}{2}.$$

Then by (46), we know that there exists $\gamma^* \in (\rho_{\alpha^*}, 1)$, such that

$$M_1(\alpha^*, \gamma^*) < K + \frac{1}{2}.$$

Therefore $(\alpha^*, \gamma^*) \in \Omega_K$, that is, Ω_K is nonempty.

(ii) For any $(h, \gamma) \in \Omega_K$, by (45), we know that $h \in (0, 2/\lambda_N(\mathcal{L}))$, $\gamma \in (\rho_h, 1)$, and

$$\frac{1}{2} < M_1(h, \gamma) < K + \frac{1}{2}.$$

Thus, by (16), one gets $K_1(h, \gamma) \leq K$, which together with Theorem 3.1 leads to the conclusion of the theorem. \square

Remark 10: From Theorems 3.1 and 3.3, it is shown that as long as the network is connected, we can always design a distributed protocol to ensure exponentially fast average consensus with each agent sending merely one bit of information to its neighbors at each time step. \square

The set Ω_K is a plane point set described by three nonlinear inequalities. Generally speaking, it is difficult to get an explicit solution of these inequalities. However, by introducing a free parameter $\epsilon_0 \in (0, 1)$, we can get a simple algorithm to choose (h, γ) from Ω_K for any given $K \geq 1$.

Algorithm 1

- (i) Choose a constant $\epsilon_0 \in (0, 1)$.
- (ii) Choose the control gain $h \in (0, h_K^*(\epsilon_0))$, where

$$h_K^*(\epsilon_0) = \min \left\{ \frac{2}{\lambda_2(\mathcal{L}) + \lambda_N(\mathcal{L})}, 2K\epsilon_0\lambda_2(\mathcal{L}) \left(\sqrt{N}\lambda_N^2(\mathcal{L}) + 2\lambda_2(\mathcal{L})d^*\epsilon_0 + (2K+1)\lambda_2^2(\mathcal{L})\epsilon_0(1-\epsilon_0) \right)^{-1} \right\} \quad (47)$$

- (iii) Let $\gamma = 1 - (1 - \epsilon_0)h\lambda_2(\mathcal{L})$.

The result below shows that any pair (h, γ) generated by Algorithm 1 belongs to Ω_K and any point in Ω_K can be generated by Algorithm 1.

Lemma 3.2: For any given $K \geq 1$, and $\epsilon_0 \in (0, 1)$, let

$$\Omega_{K, \epsilon_0} = \{(\alpha, \beta) | \alpha \in (0, h_K^*(\epsilon_0)), \beta = 1 - (1 - \epsilon_0)\alpha\lambda_2(\mathcal{L})\}. \quad (48)$$

Then we have

$$\Omega_K = \bigcup_{\epsilon_0 \in (0, 1)} \Omega_{K, \epsilon_0}. \quad (49)$$

The proof of Lemma 3.2 is given in Appendix.

Next we use an example to demonstrate the validity of the proposed consensus protocol and Theorem 3.3.

Example 1: We consider a network with 10 nodes and 0–1 weights, which means that $a_{ij} = 1$, if $(i, j) \in \mathcal{E}$, otherwise,

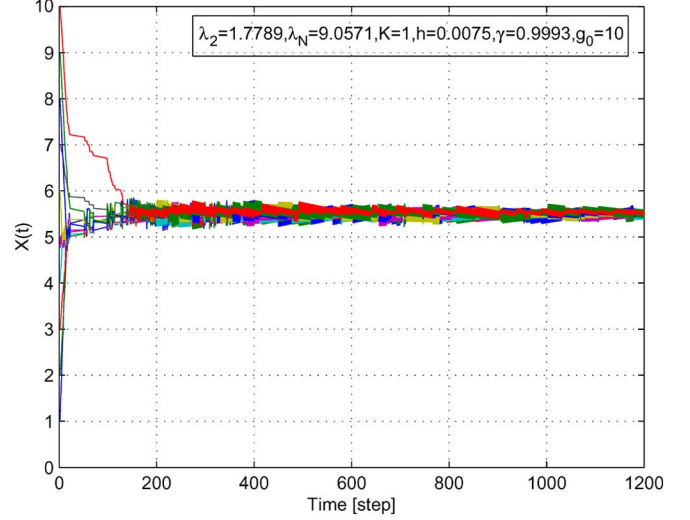


Fig. 1. Curves of states of Example 1 with $h = 0.0075$ and $\gamma = 0.9993$.

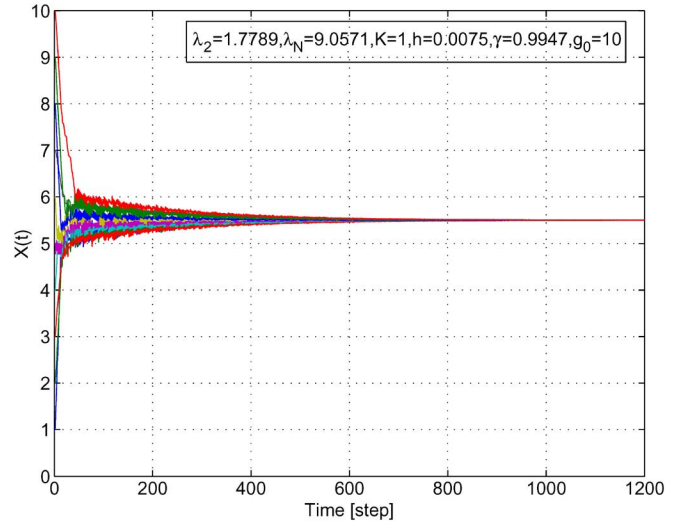


Fig. 2. Curves of states of Example 1 with $h = 0.0075$ and $\gamma = 0.9947$.

$a_{ij} = 0$. The edges of the graph are randomly generated according to $P\{(i, j) \in \mathcal{E}\} = 0.5$, for any unordered pair (i, j) . Here, $\lambda_2(\mathcal{L}) = 1.7789$ and $\lambda_{10}(\mathcal{L}) = 9.0571$. The initial states are chosen as $x_i(0) = i$, $i = 1, \dots, 10$. The control gain h and the scaling factor γ are taken as 0.0075 and 0.9993 ($\epsilon_0 = 0.95$), respectively. From (16), we get $K_1 = 1$. The one-bit quantizer is used. The evolution of the states is shown in Fig. 1. It can be seen that average consensus is achieved asymptotically with an exponential convergence rate. Then we take $\gamma = 0.9947$ ($\epsilon_0 = 0.6$). The evolution of the states is shown in Fig. 2. Comparing Fig. 1 with Fig. 2, we can see that a smaller γ gives faster convergence.

D. Asymptotic Convergence Rate

In many cases, the number N of the network nodes is large and we are concerned about asymptotic properties as N approaches infinity. In this section, we investigate the asymptotic performance of the closed-loop system. It can be seen that the asymptotic value of γ has a very compendious expression. Instead of a given network topology graph \mathcal{G} , what we consider

is a sequence of connected graphs $\mathcal{G}_N = \{\mathcal{V}_N, \mathcal{E}_N, \mathcal{A}_N\}$ with $|\mathcal{V}_N| = N$, $N = 1, 2, \dots$. However, for succinctness, the subscript N is omitted in the following.

Theorem 3.4: Suppose Assumption A1) holds. Then for any given $K \geq 1$

$$\lim_{N \rightarrow \infty} \frac{\inf_{(h, \gamma) \in \Omega_K} \gamma}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}} = 1 \quad (50)$$

where $Q_N = \lambda_2(\mathcal{L})/\lambda_N(\mathcal{L})$.

The proof of Theorem 3.4 is given in Appendix.

Remark 11: Q_N is an important physical factor. It is shown that a network exhibits better synchronizability if Q_N is large ([30]). Theorem 3.4 shows that in some sense, the asymptotic convergence rate we can achieve using the proposed protocol is $O(\exp\{- (KQ_N^2/(2\sqrt{N}))t\})$. Therefore, the asymptotic convergence rate is closely related to the number of the quantization levels, the scale and the synchronizability of the network.

Note that we do not claim that the proposed protocol given by (3), (6) and (7) provides the optimal convergence rate among all possible coding and decoding schemes. Theorem 3.4 only gives the asymptotic highest convergence rate under the proposed protocol. It would be an interesting future topic to find the fundamental upper bound for convergence rate among all possible coding and decoding schemes. \square

IV. MINIMIZATION OF COMMUNICATION ENERGY COST

Average-consensus protocols can be viewed as distributed least-mean-square estimation algorithms in sensor networks ([36]). Limited power is perhaps the most critical constraint for applications of wireless sensor networks. Simulations show that for a large-scale sensor network, communications between nodes consume far more power than computation ([37]). For power saving, reducing the communication load for distributed estimation is a critical issue. This motivates us to consider how to minimize the total communication energy cost for average consensus. Define the *convergence time constant* as $\tau_{\text{asym}} = (\ln(1/r_{\text{asym}}))^{-1}$, which gives the asymptotic number of steps for the consensus error to decrease by $1/e$ ([35]). For a $(2K + 1)$ -level quantizer, at each time instant, $\lceil \log_2(2K) \rceil$ bits are sent by each sending node and are received by each receiving node. For simplicity, we assume that the transmission power and receiving energy cost for each node are the same. Therefore, if a $(2K_1(h, \gamma) + 1)$ -level quantizer is used, then the communication energy of the whole sensor network to achieve consensus is given by

$$\Phi(\gamma) = B_h(\gamma)(c_1 N + 2c_2 |\mathcal{E}|)$$

where $B_h(\gamma) = \lceil \log_2(2K_1(h, \gamma)) \rceil \tau_{\text{asym}}$, and c_1 and c_2 are respectively the energy costs for transmitting and receiving each bit. This model follows from [38].

It can be seen that minimizing $\Phi(\gamma)$ is equivalent to minimizing $B_h(\gamma)$. By Theorem 3.1, we know that

$$B_h(\gamma) \leq B_h^*(\gamma)$$

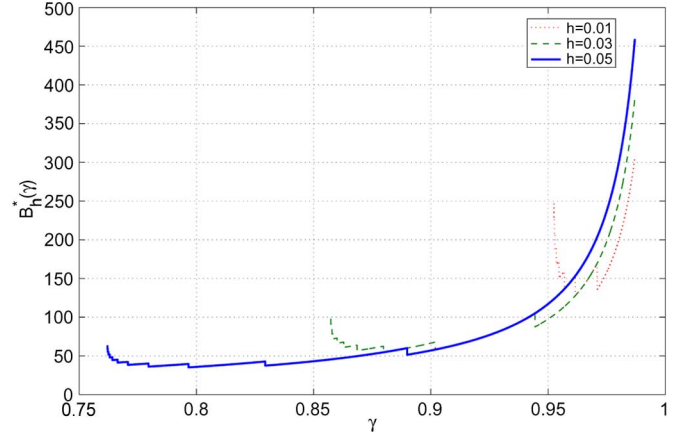


Fig. 3. Curves of B_h^* with respect to γ when $h = 0.01$, $h = 0.03$ and $h = 0.05$.

where

$$B_h^*(\gamma) = \lceil \log_2(2K_1(h, \gamma)) \rceil (\ln(1/\gamma))^{-1}$$

which is a function of γ for the given network topology \mathcal{G} and the control gain h . The smaller the γ , the smaller the $(\ln(1/\gamma))^{-1}$, but the larger the $\lceil \log_2(2K_1(h, \gamma)) \rceil$. Minimizing $B_h^*(\gamma)$ with respect to γ subject to $\gamma \in (\rho_h, 1)$, we can get a good solution for γ .

Remark 12: Here, we only get a sub-optimal solution of γ , since what we minimize is the upper bound of $B_h(\gamma)$. If the stochastic model of quantization errors is used as in Theorem 3.2, and the convergence time constant is defined as $\tau_{\text{asym}} = (\ln(1/\bar{r}_{\text{asym}}))^{-1}$, then by Theorem 3.2, we have $B_h(\gamma) = B_h^*(\gamma)$ and minimizing $B_h^*(\gamma)$ gives the optimal solution of γ .¹ \square

Example 2: We consider a network with 30 nodes and 0–1 weights, where $\lambda_2(\mathcal{L}) = 4.7622$ and $\lambda_{30}(\mathcal{L}) = 22.6239$. The curves of $B_h^*(\gamma)$ when $h = 0.01$, $h = 0.03$ and $h = 0.05$ are shown in Fig. 3. The discontinuity of the curves is due to the rounding up and down operators. It is shown that the optimized value of γ decreases as h increases.

V. QUANTIZED CONSENSUS FOR EXPANDER NETWORKS

From Theorem 3.3, to ensure average consensus by using any given $\lceil \log_2(2K) \rceil$ -bit quantizer, it is sufficient that the control gain h and the scaling factor γ be chosen from

$$\bar{\Omega}_K = \left\{ (h, \gamma) \mid h \in \left(0, \frac{2}{\lambda_2(\mathcal{L}) + \lambda_N(\mathcal{L})} \right), \right. \\ \left. \gamma \in (1 - h\lambda_*, 1), \quad \bar{M}_1(h, \gamma) < K + \frac{1}{2} \right\}$$

where

$$\bar{M}_1(h, \gamma) = \frac{2\sqrt{N}h^2(d^*)^2}{\gamma[\gamma - (1 - h\lambda_*)]} + \frac{1 + 2hd^*}{2\gamma}$$

and λ_* is a positive lower bound of $\lambda_2(\mathcal{L})$. Therefore, the choice of h and γ depends on a positive lower bound of the algebraic connectivity. Generally speaking, the positive lower bound of

¹In this case, $\Phi(\gamma)$ is defined as the communication energy of the whole network to achieve consensus in the stochastic average sense.

the algebraic connectivity of a complex network is not known, however, for an important class of regular networks, it can be worked out by using other physical parameters.

Definition 5.1: ([39]) The *isoperimetric constant* or *expander constant* of an equally weighted graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$ on N vertices is

$$i_c(\mathcal{G}) = \min \left\{ \frac{|\partial\mathcal{F}|}{|\mathcal{F}|} : \mathcal{F} \subseteq \mathcal{V}, \quad 0 < |\mathcal{F}| \leq \frac{N}{2} \right\}$$

where the boundary $\partial\mathcal{F}$ of \mathcal{F} is the set of edges with one extremity in \mathcal{F} and the other in $\mathcal{V} - \mathcal{F}$.

Definition 5.2: ([39]): Let c be a positive constant. A d -regular graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$ is called a c -expander, if $i_c(\mathcal{G}) \geq c$.

Expander graphs are highly connected sparse graphs, which play an important role in designing efficient communication networks in computer science ([30], [40]). From Corollary 2.3 of [41], a key property of a c -expander is

$$\lambda_2(\mathcal{L}) \geq d - \sqrt{d^2 - c^2}. \quad (51)$$

For expander networks, we have the following result.

Theorem 5.1: Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$ be a d -regular c -expander. Suppose Assumption A2) holds. For any given $h \in (0, (2d)^{-1})$ and $\gamma \in (1 - h(d - \sqrt{d^2 - c^2}), 1)$, let

$$K_2(h, \gamma) = \left\lfloor M_2(h, \gamma) - \frac{1}{2} \right\rfloor + 1 \quad (52)$$

$$M_2(h, \gamma) = \frac{2\sqrt{N}h^2d^2}{\gamma[\gamma - 1 + h(d - \sqrt{d^2 - c^2})]} + \frac{1 + 2hd}{2\gamma} \quad (53)$$

and for any given $K \geq K_2(h, \gamma)$ let

$$g_0 > \max \left\{ \frac{C_x}{K + \frac{1}{2}}, \frac{4hd(d^2 + \sqrt{d^2 - c^2})(C_\delta + C_x)}{hc^2} \right\}. \quad (54)$$

Then under the protocol given by (3), (6) and (7) with the $(2K + 1)$ -level uniform quantizer (4) and the scaling function $g(t) = g_0\gamma^t$, the closed-loop system (12) satisfies

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{N} \sum_{j=1}^N x_j(0), \quad i = 1, 2, \dots, N. \quad (55)$$

Proof: Note that for a c -expander, Assumption A1) holds. From the property of graph Laplacian ([32]), we know that $\lambda_N(\mathcal{L}) \leq 2d$, then by Lemma 3.1 and $h \in (0, (2d)^{-1})$, we get

$$\rho_h = 1 - h\lambda_2(\mathcal{L}). \quad (56)$$

This together with (51), (52), (53), (16) and (17) leads to

$$K_2(h, \gamma) \geq K_1(h, \gamma). \quad (57)$$

By (56), we have

$$\gamma - \rho_{h,\mathcal{G}} \leq 1 - (1 - h\lambda_2(\mathcal{L})) \leq 2hd$$

which together with $h\lambda_N(\mathcal{L}) < 1$, and $\lambda_N(\mathcal{L}) \geq \lambda_2(\mathcal{L}) \geq d - \sqrt{d^2 - c^2}$ gives

$$\begin{aligned} & \frac{2(\gamma - \rho_{h,\mathcal{G}})(C_\delta\gamma + hC_x\lambda_N(\mathcal{L}))}{h\lambda_N(\mathcal{L})} \\ & \leq \frac{4hd(d + \sqrt{d^2 - c^2})(C_\delta + C_x)}{hc^2}. \end{aligned}$$

By this, (54), (18), (57) and Theorem 3.1, we have (55). \square

Similar to Theorem 3.3, we have the following theorem.

Theorem 5.2: Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$ be a d -regular c -expander. Suppose Assumption A2) holds. For any given $K \geq 1$, let

$$\Omega_K^c = \left\{ (\alpha, \beta) \mid \alpha \in \left(0, \frac{1}{2d}\right), \beta \in \left(1 - \alpha(d - \sqrt{d^2 - c^2}), 1\right), \right. \\ \left. M_2(\alpha, \beta) < K + \frac{1}{2} \right\}$$

where $M_2(\alpha, \beta)$ is given by (53). Then, (i) Ω_K^c is nonempty. (ii) For any given $(h, \gamma) \in \Omega_K^c$, under the protocol given by (3), (6) and (7) with $g(t) = g_0\gamma^t$ and the $(2K + 1)$ -level uniform quantizer (4), the closed-loop system (12) satisfies

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{N} \sum_{j=1}^N x_j(0), \quad i = 1, 2, \dots, N$$

where g_0 is a constant satisfying (54).

The proof of Theorem 5.2 is similar to that of Theorem 3.3 and is omitted here.

To choose (h, γ) from Ω_K^c for any given $K \geq 1$, we have the following algorithm.

Algorithm 2

- (i) Choose a constant $\epsilon_0 \in (0, 1)$.
- (ii) Choose the control gain $h \in (0, \tilde{h}_c(\epsilon_0))$, where

$$\tilde{h}_c(\epsilon_0) = \min \left\{ \frac{1}{2d}, \frac{2Kd^{-1}\epsilon_0}{\frac{\sqrt{Nd}(d + \sqrt{d^2 - c^2})}{c^2} + 2\epsilon_0 + (4K + 2)\epsilon_0(1 - \epsilon_0)} \right\}$$

- (iii) Let $\gamma = 1 - (1 - \epsilon_0)h(d + \sqrt{d^2 - c^2})$.

Remark 13: A Ramanujan graph \mathcal{G} is defined as a d -regular graph with $\lambda_2(\mathcal{L}) \geq d - 2\sqrt{d - 1}$. Ramanujan graphs are expanders which achieve the largest possible algebraic connectivity ([39]–[42]). Recently, it was found that in fact, Ramanujan graphs are close to optimal graphs with the best synchronizability ([30], [43]). For a Ramanujan graph \mathcal{G} with $d \geq 3$, $\lambda_2(\mathcal{L}) \geq 3 - 2\sqrt{2}$. This property is of particular importance, since $\lambda_2(\mathcal{L})$ is lower bounded, independent of the graph \mathcal{G} . So by Theorem 5.2, for the case with a d -regular Ramanujan graph with $d \geq 3$, we can choose h and γ , such that $h \in (0, \tilde{h}_c(\epsilon_0))$ and $\gamma = 1 - (1 - \epsilon_0)(3 - 2\sqrt{2})h$, where

$$\tilde{h}_c(\epsilon_0) = \min \left\{ \frac{1}{2d}, \frac{2K\epsilon_0}{\frac{\sqrt{Nd}d^2}{3-2\sqrt{2}} + 2d\epsilon_0 + 2(2K+1)d\epsilon_0(1-\epsilon_0)} \right\}$$

for any given $K \geq 1$. \square

Remark 14: Since $Q_N \leq 1$, the asymptotic convergence rate is slower than $O(\exp\{-(K/(2\sqrt{N}))t\})$. In [30], it is shown that entangled networks, which have largest Q_N will typically be Ramanujan graphs. In [43], it is also shown that Ramanujan graphs asymptotically optimize Q_N for a large class of regular graphs. For a family of d -regular Ramanujan graphs, $Q_N \geq ((d - 2\sqrt{d-1})/(2d))$. Therefore, the asymptotic convergence rate we can achieve is higher than $O(\exp\{-(K(d - 2\sqrt{d-1}))/((4d\sqrt{N}))t\})$. \square

VI. CONCLUSION

In this paper, the average-consensus control problem has been considered for undirected networks of discrete-time first-order agents under finite bit-rate communication. Based on uniform quantization with scaling, a dynamic difference encoding and decoding scheme is used for the communication between each pair of agents. A distributed protocol has been proposed, where the control input of each agent is a weighted sum of the difference between the estimate of its neighbor's state and the internal state of its own encoder. This type of protocol is equivalent to adding an error compensation term to the original weighted average type protocol. It is shown that for a connected undirected dynamic network with first-order agents, no matter how many agents there are, we can always design a distributed protocol to ensure that average consensus is achieved asymptotically with as few as one bit information exchange between each pair of adjacent agents at each time step. It is shown that the asymptotic convergence rate is related to the number of network nodes, the number of quantization levels and the synchronizability of the network.

In this paper, we assume that all agents are synchronized and the communication channels are noiseless. For future research, the case with asynchronous protocols, robustness with respect to packet-loss, link failures and time-delay may be considered, and the case with noisy digital communication channel is also interesting.

APPENDIX

Proof of Lemma 3.2: First, we prove that

$$\bigcup_{\epsilon_0 \in (0,1)} \Omega_{K,\epsilon_0} \subseteq \Omega_K. \quad (\text{A1})$$

For any $(\alpha_0, \beta_0) \in \bigcup_{\epsilon_0 \in (0,1)} \Omega_{K,\epsilon_0} \subseteq \Omega_K$, there exists $\epsilon_0^* \in (0, 1)$ such that

$$\alpha_0 \in (0, h_K^*(\epsilon_0^*)) \quad (\text{A2})$$

and

$$\beta_0 = 1 - (1 - \epsilon_0^*) \alpha_0 \lambda_2(\mathcal{L}). \quad (\text{A3})$$

This together with (47) leads to

$$\alpha_0 \in \left(0, \frac{2}{\lambda_2(\mathcal{L}) + \lambda_N(\mathcal{L})}\right) \quad (\text{A4})$$

and

$$\beta \in (\rho_{\alpha_0}, 1). \quad (\text{A5})$$

From (17) and (A3), it follows that $M_1(\alpha_0, \beta_0) = S(\alpha_0, \epsilon_0^*)$, where:

$$\begin{aligned} S(\alpha_0, \epsilon_0^*) &= \frac{1 + 2\alpha_0 d^*}{2[1 - (1 - \epsilon_0^*) \alpha_0 \lambda_2(\mathcal{L})]} \\ &\quad + \frac{\sqrt{N} \alpha_0 \lambda_N^2(\mathcal{L})}{2[1 - (1 - \epsilon_0^*) \alpha_0 \lambda_2(\mathcal{L})] \epsilon_0^* \lambda_2(\mathcal{L})} \\ &= \frac{\epsilon_0^* \lambda_2(\mathcal{L})(1 + 2\alpha_0 d^*) + \sqrt{N} \alpha_0 \lambda_N^2(\mathcal{L})}{2[1 - (1 - \epsilon_0^*) \alpha_0 \lambda_2(\mathcal{L})] \epsilon_0^* \lambda_2(\mathcal{L})}. \end{aligned} \quad (\text{A6})$$

Note that $S(\alpha_0, \epsilon_0^*) < K + 1/2$ is equivalent to $\alpha_0 < \alpha_K^*$, where

$$\begin{aligned} \alpha_K^* &= (2K \epsilon_0^* \lambda_2(\mathcal{L})) \left(\sqrt{N} \lambda_N^2(\mathcal{L}) + 2\lambda_2(\mathcal{L}) d^* \epsilon_0^* \right. \\ &\quad \left. + (2K + 1) \lambda_2^2(\mathcal{L}) \epsilon_0^* (1 - \epsilon_0^*) \right)^{-1}. \end{aligned}$$

Then by (A2) and (47), we have $M_1(\alpha_0, \beta_0) < K + 1/2$. This together with (A4), (A3), (A5) and (45) leads to (A1).

Now we prove that

$$\Omega_K \subseteq \bigcup_{\epsilon_0 \in (0,1)} \Omega_{K,\epsilon_0}. \quad (\text{A7})$$

For any $(\alpha_0, \beta_0) \in \Omega_K$, by (45), we know that

$$0 < \alpha_0 < \frac{2}{\lambda_2(\mathcal{L}) + \lambda_N(\mathcal{L})}. \quad (\text{A8})$$

This together with Lemma 3.1 gives

$$\rho_{\alpha_0} = 1 - \alpha_0 \lambda_2(\mathcal{L}). \quad (\text{A9})$$

Denote $\epsilon_0' = 1 - ((1 - \beta_0)/\alpha_0 \lambda_2(\mathcal{L}))$. Then

$$\beta_0 = 1 - (1 - \epsilon_0') \alpha_0 \lambda_2(\mathcal{L}). \quad (\text{A10})$$

By (A9) and $\beta_0 \in (\rho_{\alpha_0}, 1)$, we have $\epsilon_0' \in (0, 1)$. From (A10) and (17), it follows that $M_1(\alpha_0, \beta_0) = S(\alpha_0, \epsilon_0')$, where $S(\alpha_0, \epsilon_0')$ is given by (A6) with ϵ_0^* replaced by ϵ_0' . This together with $M_1(\alpha_0, \beta_0) < K + 1/2$ leads to $S(\alpha_0, \epsilon_0') < K + 1/2$. Note that $S(\alpha_0, \epsilon_0') < K + 1/2$ is equivalent to

$$\alpha_0 < \alpha_K'$$

where

$$\begin{aligned} \alpha_K' &= (2K \epsilon_0' \lambda_2(\mathcal{L})) \left(\sqrt{N} \lambda_N^2(\mathcal{L}) + 2\lambda_2(\mathcal{L}) d^* \epsilon_0' \right. \\ &\quad \left. + (2K + 1) \lambda_2^2(\mathcal{L}) \epsilon_0' (1 - \epsilon_0') \right)^{-1}. \end{aligned}$$

Then by (A8) and (47), we have $\alpha_0 \in (0, h_K^*(\epsilon_0'))$. This together with (A10) leads to $(\alpha_0, \beta_0) \in \Omega_{K,\epsilon_0'}$, that is, (A7) holds. \square

The proof of Theorem 3.4 needs the following lemmas.

Lemma A.1: ([31]): Suppose \mathcal{G} is a weighted graph with N nodes, and \mathcal{L} is the Laplacian matrix. Then $\lambda_2(\mathcal{L}) \leq (N/(N-1)) \min_i \deg_i$.

Lemma A.2: Suppose $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A} = [a_{ij}]_{N \times N}\}$ is a weighted graph with $|\mathcal{V}| = N \geq 2$, \mathcal{L} is the Laplacian matrix, and d^* is the maximum degree. Then $\lambda_N(\mathcal{L}) \geq (N/(N-1))d^*$.

Proof: Denote $a^* = \max_{i,j} a_{ij}$, and $\bar{\mathcal{L}} = a^*NI - a^*\mathbf{1}\mathbf{1}^T - \mathcal{L}$. Then it can be seen that $\bar{\mathcal{L}}$ is a Laplacian matrix.

By Lemma A.1, the properties of Laplacian matrices and the complementary graph method ([31]), we have

$$\begin{aligned} \lambda_N(\mathcal{L}) &= \max_{\|\delta\|_2=1, \delta^T \mathbf{1}=0} \delta^T \mathcal{L} \delta \\ &= \max_{\|\delta\|_2=1, \delta^T \mathbf{1}=0} \delta^T (a^*NI - a^*\mathbf{1}\mathbf{1}^T - \bar{\mathcal{L}}) \delta \\ &= a^*N - \min_{\|\delta\|_2=1, \delta^T \mathbf{1}=0} \delta^T \bar{\mathcal{L}} \delta \\ &= a^*N - \lambda_2(\bar{\mathcal{L}}) \\ &\geq a^*N - \frac{N}{N-1} \min_i (a^*(N-1) - \deg_i) \\ &= a^*N - \frac{N}{N-1} (a^*(N-1) - \max_i \deg_i) \\ &= \frac{N}{N-1} d^* \end{aligned} \quad \square$$

Lemma A.3: Suppose Assumption A1) holds. For any given $K \geq 1$, and $\epsilon_0 \in (0, 1)$, let $h_K^*(\epsilon_0)$ be given by (47). Define

$$\Gamma_K = \{\theta | \theta = 1 - (1 - \epsilon_0)h\lambda_2(\mathcal{L}), \epsilon_0 \in (0, 1), h \in (0, h_K^*(\epsilon_0))\}.$$

Then

$$\inf_{\gamma \in \Gamma_K} \gamma \geq 1 - \frac{K\lambda_2^2(\mathcal{L})}{2\sqrt{N}\lambda_N^2(\mathcal{L})}.$$

Proof: From (47), we have

$$h < \frac{2K\epsilon_0\lambda_2(\mathcal{L})}{\sqrt{N}\lambda_N^2(\mathcal{L})}, \quad \forall h \in (0, h_K^*(\epsilon_0)).$$

Then for any $\gamma \in \Gamma_K$ with $\epsilon_0 \in (0, 1)$ and $h \in (0, h_K^*(\epsilon_0))$, noting that $\epsilon_0(1 - \epsilon_0) \leq (1/4)$, we get

$$\gamma > 1 - \frac{2K(1 - \epsilon_0)\epsilon_0\lambda_2^2(\mathcal{L})}{\sqrt{N}\lambda_N^2(\mathcal{L})} \geq 1 - \frac{K\lambda_2^2(\mathcal{L})}{2\sqrt{N}\lambda_N^2(\mathcal{L})}.$$

This leads to the conclusion of this lemma.

Lemma A.4: Suppose Assumption A1) holds. For any given $K \geq 1$, and $\epsilon_0 \in (0, 1)$, let $h_K^*(\epsilon_0)$ be given by (47) and $\gamma = 1 - (1 - \epsilon_0)h\lambda_2(\mathcal{L})$. Then

$$\inf_{\gamma \in \Gamma_K} \gamma \leq 1 - \frac{K\lambda_2^2(\mathcal{L})}{2(\sqrt{N} + 2K + 3)\lambda_N^2(\mathcal{L})}. \quad (\text{A11})$$

Proof: From (47) and Lemma A.2, we have

$$\begin{aligned} h_K^*(\epsilon_0) &\geq \min \left\{ \frac{1}{\lambda_N(\mathcal{L})}, \right. \\ &\quad \left. \frac{2K\epsilon_0\lambda_2(\mathcal{L})}{\sqrt{N}\lambda_N^2(\mathcal{L}) + 2\lambda_N^2(\mathcal{L}) + (2K + 1)\lambda_N^2(\mathcal{L})} \right\} \\ &= \min \left\{ \frac{1}{\lambda_N(\mathcal{L})}, \frac{2K\epsilon_0\lambda_2(\mathcal{L})}{(\sqrt{N} + 2K + 3)\lambda_N^2(\mathcal{L})} \right\} \\ &= \frac{2K\epsilon_0\lambda_2(\mathcal{L})}{(\sqrt{N} + 2K + 3)\lambda_N^2(\mathcal{L})} \end{aligned}$$

which together with $\inf_{h \in (0, h_K^*(\epsilon_0))} \gamma = 1 - (1 - \epsilon_0)h_K^*(\epsilon_0)\lambda_2(\mathcal{L})$ gives

$$\inf_{h \in (0, h_K^*(\epsilon_0))} \gamma \leq 1 - \frac{2(1 - \epsilon_0)\epsilon_0 K \lambda_2^2(\mathcal{L})}{(\sqrt{N} + 2K + 3)\lambda_N^2(\mathcal{L})}.$$

From this, it follows that:

$$\begin{aligned} \inf_{\gamma \in \Gamma_K} \gamma &\leq 1 - \max_{\epsilon_0 \in (0, 1)} \frac{2(1 - \epsilon_0)\epsilon_0 K \lambda_2^2(\mathcal{L})}{(\sqrt{N} + 2K + 3)\lambda_N^2(\mathcal{L})} \\ &= 1 - \frac{K\lambda_2^2(\mathcal{L})}{2(\sqrt{N} + 2K + 3)\lambda_N^2(\mathcal{L})} \end{aligned}$$

that is, (A11) holds. \square

Proof of Theorem 3.4: By Lemma A.3, we have

$$\frac{\inf_{\gamma \in \Gamma_K} \gamma}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}} \geq \frac{1 - \frac{KQ_N^2}{2\sqrt{N}}}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}}, \quad \forall N \geq 1$$

where $h_K^*(\epsilon_0)$ is given by (47), which together with $KQ_N^2/\sqrt{N} \rightarrow 0, N \rightarrow \infty$ gives

$$\liminf_{N \rightarrow \infty} \frac{\inf_{\gamma \in \Gamma_K} \gamma}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}} \geq 1. \quad (\text{A12})$$

By Lemma A.4, we have

$$\frac{\inf_{\gamma \in \Gamma_K} \gamma}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}} \leq \frac{\left(1 - \frac{KQ_N^2}{2\sqrt{N}} \frac{\sqrt{N}}{\sqrt{N} + 2K + 3}\right)}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}}, \quad \forall N \geq 1$$

which together with $KQ_N^2/\sqrt{N} \rightarrow 0, N \rightarrow \infty$ gives

$$\limsup_{N \rightarrow \infty} \frac{\inf_{\gamma \in \Gamma_K} \gamma}{\exp\left\{-\frac{KQ_N^2}{2\sqrt{N}}\right\}} \leq 1. \quad (\text{A13})$$

By Lemma 3.2 and (48), we know that

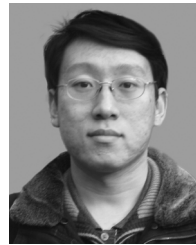
$$\inf_{(h, \gamma) \in \Omega_K} \gamma = \inf_{\epsilon_0 \in (0, 1)} \inf_{(h, \gamma) \in \Omega_{K, \epsilon_0}} \gamma = \inf_{\gamma \in \Gamma_K} \gamma.$$

By this, (A12) and (A13), we get (50). \square

REFERENCES

- [1] W. Ren and R. W. Beard, *Distributed Consensus in Multi-Vehicle Cooperative Control*. London, U.K.: Springer, 2008.
- [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] J. N. Tsitsiklis, "Problems in Decision Making and Computation," Ph.D. dissertation, Dept. Elect. Eng. Comp. Sci., MIT, Cambridge, MA, 1984.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] L. Moreau, "Stability of multi-agent systems with dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169–182, Feb. 2005.

- [8] 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.
- [9] W. Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proc. Amer. Control Conf.*, Portland, OR, Jun. 2005, pp. 1859–1864.
- [10] 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, Jan. 2007.
- [11] A. Kashyap, T. Basar, and R. Srikant, "Consensus with quantized information updates," in *Proc. 45th IEEE Conf. Decision Control*, San Diego, CA, Dec. 13–15, 2006, pp. 2728–2733.
- [12] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192–1203, 2007.
- [13] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On Distributed Averaging Algorithms and Quantization Effects" *Arxiv preprint arXiv:0711.4179*, 2007.
- [14] R. Carli, F. Fagnani, P. Frasca, T. Taylor, and S. Zampieri, "Average consensus on networks with transmission noise or quantization," in *Proc. Eur. Control Conf.*, 2007, pp. 1852–1857.
- [15] P. Frasca, R. Carli, F. Fagnani, and S. Zampieri, "Average consensus on networks with quantized communication," *Int. J. Nonlin. Robust Control*, vol. 19, no. 16, pp. 1787–1816, 2009.
- [16] R. Carli, F. Fagnani, P. Frasca, and S. Zampieri, "A probabilistic analysis of the average consensus algorithm with quantized communication," in *Proc. 17th IFAC World Congress*, Seoul, Korea, Jul. 6–11, 2008.
- [17] R. Carli, F. Fagnani, P. Frasca, and S. Zampieri, "Efficient quantized techniques for consensus algorithms," in *Proc. NeCST Workshop*, Nancy, France, 2007 [Online]. Available: http://calvino.polito.it/~frasca/docs/CFFZ_necst07.pdf
- [18] S. Kar and J. M. F. Moura, "Distributed Consensus Algorithms in Sensor Networks: Quantized Data and Random Link Failures" *Arxiv preprint arXiv:0712.1609*, 2007.
- [19] M. Huang and J. H. Manton, "Coordination and consensus of networked agents with noisy measurement: Stochastic algorithms and asymptotic behavior," *SIAM J. Control Optim.: Special Issue Control Optim. Cooperative Networks*, vol. 48, no. 1, pp. 134–161, 2009.
- [20] M. Huang and J. H. Manton, "Stochastic approximation for consensus seeking: Mean square and almost sure convergence," in *Proc. 46th IEEE Conf. Decision Control*, New Orleans, LA, Dec. 12–14, 2007, pp. 306–311.
- [21] 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.
- [22] T. Li and J. F. Zhang, "Sampled-data based average consensus with measurement noises: Convergence analysis and uncertainty principle," *Sci. China Series F: Inform. Sci.*, vol. 52, no. 11, pp. 2089–2103, 2009.
- [23] T. Li and J. F. Zhang, "Consensus conditions of multiagent systems with time-varying topologies and stochastic communication noises," *IEEE Trans. Autom. Control*, to be published.
- [24] T. C. Aysal, M. J. Coates, and M. G. Rabbat, "Distributed average consensus with dithered quantization," *IEEE Trans. Signal Processing*, vol. 56, no. 10, pp. 4905–4918, Sep. 2008.
- [25] R. Carli, F. Fagnani, P. Frasca, and S. Zampieri, "Average consensus by gossip algorithms with quantized communication," in *Proc. 47th IEEE Conf. Decision Control*, Dec. 9–11, 2008, pp. 4831–4836.
- [26] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.
- [27] G. N. Nair and R. J. Evans, "Exponential stability of finite-dimensional linear systems with limited data rates," *Automatica*, vol. 39, no. 4, pp. 585–593, 2003.
- [28] N. Elia and S. K. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. Autom. Control*, vol. 46, no. 9, pp. 1384–1400, Sep. 2001.
- [29] M. Fu and L. Xie, "The sector bound approach to quantized feedback control," *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1698–1711, Nov. 2001.
- [30] L. Donetti, P. I. Hurtado, and M. A. Muñoz, "Entangled networks, synchronization, and optimal network topology," *Phys. Rev. Lett.*, vol. 95, no. 18, 2005 [Online]. Available: <http://prl.aps.org/abstract/PRL/v95/i18/e188701>
- [31] M. F. Praha, "Algebraic connectivity of graphs," *Czechoslovak Math. J.*, vol. 23, no. 98, pp. 298–305, 1973.
- [32] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York: Springer, 2001.
- [33] E. Kokiopoulou and P. Frossard, "Polynomial filtering for fast convergence in distributed consensus," *IEEE Trans. Signal Processing*, vol. 57, no. 1, pp. 342–354, Jan. 2009.
- [34] L. Xiao, S. Boyd, and S. J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel Distrib. Comp.*, vol. 67, no. 1, pp. 33–46, 2007.
- [35] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, 2004.
- [36] 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.
- [37] V. Shnayder, M. Hempstead, B. Chen, G. W. Allen, and M. Welsh, "Simulating the power consumption of large-scale sensor network applications," in *Proc. 2nd Int. Conf. Embedded Networked Sensor Syst.*, 2004, pp. 188–200.
- [38] J. Lin, W. Xiao, F. Lewis, and L. Xie, "Energy efficient distributed adaptive multi-sensor scheduling for target tracking in wireless sensor networks," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 6, pp. 1886–1896, Dec. 2009.
- [39] G. P. Davidoff, P. Sarnak, and A. Valette, *Elementary Number Theory, Group Theory and Ramanujan Graphs*. London, U.K.: Cambridge University Press, 2003.
- [40] F. Bien, "Construction of telephone networks by group representations," *Notices Amer. Math. Soc.*, vol. 36, no. 1, pp. 5–22, 1989.
- [41] A. Berman and X. D. Zhang, "Lower bounds for the eigenvalues of Laplacian matrices," *Linear Algebra Appl.*, vol. 316, no. 1, pp. 13–20, 2000.
- [42] R. Olfati-Saber, "Algebraic connectivity ratio of Ramanujan graphs," in *Proc. Amer. Control Conf.*, New York, NY, Jul. 9–13, 2007, pp. 4619–4624.
- [43] S. Kar, S. Aldosari, and J. M. F. Moura, "Topology for distributed inference on graphs," *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2609–2613, Dec. 2008.



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 Technological University, Singapore, as a Research Assistant. Since July, 2009, he has been with AMSS, CAS, where he is now a Research Associate. His current research interests include system modeling, stochastic systems and multi-agent systems.

Dr. Li received the Singapore Millennium Foundation Postdoctoral Research Fellowship in 2009.



Minyue Fu (F'04) received the B.S. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 1982, and the M.S. and Ph.D. degrees in electrical engineering from the University of Wisconsin-Madison, in 1983 and 1987, respectively.

From 1983 to 1987, he held a Teaching Assistantship and a Research Assistantship at the University of Wisconsin-Madison. He was a Computer Engineering Consultant at Nicolet Instruments, Inc., Madison, WI, during 1987. From 1987 to 1989, he served as an Assistant Professor in the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI. He joined the Department of Electrical and Computer Engineering, the University of Newcastle, Australia, in 1989. Currently, he is a Chair Professor in Electrical Engineering. In addition, he was a Visiting Associate Professor at University of Iowa, Ames, from 1995 to 1996, and a Senior Fellow/Visiting Professor at Nanyang Technological University, Singapore, in 2002. He holds a Changjiang Visiting Professorship at Shandong University and visiting positions at South China University of Technology and Zhejiang University in China. He was an Associate Editor for *Automatica* and the *Journal of Optimization and Engineering*. His main research interests include control systems, signal processing and communications.

Dr. Fu was an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.



Lihua Xie (F'07) received the B.E. and M.E. degrees in electrical engineering from Nanjing University of Science and Technology, Nanjing, China, in 1983 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Newcastle, Newcastle, Australia, in 1992.

He was with the Department of Automatic Control, Nanjing University of Science and Technology, from 1986 to 1989. Since 1992, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore,

where he is currently a Professor, Deputy Head of Control and Instrumentation Division, and the Director of Centre for Intelligent Machines. He is also a Changjiang visiting Professor with South China University of Technology. He is an Associate Editor at Large of the *Journal of Control Theory and Applications*. He served as an Associate Editor of *Automatica*. His current research interests include robust control and estimation, networked control, sensor networks, time delay systems, and control of disk drive systems. In these areas, he has published four books, two patents and many papers.

Dr. Xie is a Fellow of the Institution of Engineers, Singapore. He is currently an Associate Editor of the IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY. He served as an Associate Editor of a number of archival journals including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and the IEEE TRANSACTIONS ON CIRCUIT AND SYSTEMS-II. He was the General Chairman of the 9th International Conference on Control, Automation, Robotics and Vision and the 7th IEEE International Conference on Control and Automation.



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, China, in 1985 and the Ph.D. degree from Institute of Systems Science (ISS), Chinese Academy of Sciences (CAS), Beijing, China, 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 is a Managing Editor of the *Journal of Systems Science and Complexity*; Deputy Editor-in-Chief of the following three journals: *Acta Automatica Sinica*, *Journal of Systems Science and Mathematical Sciences*, *Control Theory & Applications*; and Associate Editor of several other journals, including the *SIAM Journal on Control and Optimization*. 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. He was an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He serves as the Director of the Technical Committee on Control Theory, Chinese Association of Automation.

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. He was an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He serves as the Director of the Technical Committee on Control Theory, Chinese Association of Automation.