An Adaptive Quantization Scheme for Distributed Consensus

Jun Fang
Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, NJ 07030
Email: jun.fang@stevens.edu

Hongbin Li
Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, NJ 07030
Email: hongbin.li@stevens.edu

***Abstract*— The problem of distributed average consensus with quantized data is considered in this paper. We firstly propose a simple modification to the classical consensus protocol. Under a condition that the quantization noise variance converges to zero, the proposed protocol achieves a consensus in a mean squared sense and the consensus value is equal to the average of the initial state. Based on this result, we develop an adaptive quantization scheme which can adaptively adjust its quantization threshold and step-size by learning from previous runs, in a way such that the quantization noise variance at each sensor decreases to zero. Simulation results are presented to illustrate the effectiveness of the proposed algorithm.**

Index Terms – **Distributed average consensus, quantized data, wireless sensor network (WSN).**

I. Introduction

Distributed average consensus has attracted much attention over the past few years. It is a fundamental problem arising from the wide range of applications in wireless sensor networks (WSNs), such as distributed function computation and distributed parameter estimation. In contrast to a centralized approach, a distributed average consensus scheme does not need a fusion center. Instead, sensors exchange data with their respective neighbors and carry out local computations to eventually reach a consensus, which is usually the average value of all sensor measurements. This characteristic makes the distributed average consensus scheme suitable for WSN applications where power and bandwidth are severely constrained.

A multitude of studies on distributed average consensus have appeared recently. Among them, a major research direction focuses on the computation of the optimal weights [1], [2] or development of schemes [3], [4] to accelerate the convergence rate of the distributed consensus algorithms. In these work, they usually assume that the real data are exchanged among neighboring sensors without distortion. This assumption, however, is undermined in practice due to the link noise and quantization errors. Unfortunately, it is shown [5] that in the presence of noise, the classical consensus protocol adopted in [1], [2] does not achieve a consensus and the asymptotic mean squared error is unbounded. Many recent efforts [6]–[10] have been made to address this issue. Specifically, when only quantization errors are considered, the authors [6] proposed two coding schemes by exploiting the temporal correlation among successive states. A consensus can be achieved under the condition that the quantization noise variance converges to zero. In addition, [7] proposed a dithered quantization scheme to reach a consensus, in which a set of weights decaying to zero and satisfying a persistence condition are used. Despite these efforts, we are still away from our objective because these algorithms either cannot reach a consensus [9], [10] or converge to a final consensus value which is random and not the desired average of the initial states [6], [7].

In this paper, we present a modified version of the classical consensus algorithm. The proposed consensus protocol is shown to converges in the mean squared sense to a desired state under a condition that the quantization noise variance decreases to zero. We propose an adaptive quantization (AQ) scheme which can adaptively adjust its quantization threshold and step-size by learning from previous states, in a way such that the quantization noise variance at each sensor decreases to zero.

II. Consensus Protocols

We model the WSN as an undirected graph $G = (V, E)$ whose vertices $V = \{1, 2, \ldots, N\}$ correspond to the sensors and whose edges $E = \{(i, j)|i, j \in V\}$ represent available communication links among sensors. An edge between i and j exists if sensor i can communicate directly with sensor j. We focus our study on the connected graph, i.e. there exists a multihop communication path connecting every pair of vertices. The structure of the graph can be described by an $N \times N$ symmetric affinity matrix $\mathbf{A}$

$$a_{i,j} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $a_{i,j}$ denotes the (i, j)th entry of $\mathbf{A}$. The Laplacian matrix of the graph G is defined as

$$\mathbf{L} \triangleq \mathbf{D} - \mathbf{A} \tag{2}$$

where $\mathbf{D} \triangleq \text{diag}(\mathbf{A1})$ is the degree matrix and $\mathbf{1}$ denotes a column vector with all unity elements. $\mathbf{L}$ is a positive semidefinite matrix with only one null eigenvalue associated with the eigenvector $\frac{1}{\sqrt{N}}\mathbf{1}$ [11]. Assuming ideal links and no quantization, the classical distributed average consensus algorithm [1] updates its state as

$$\mathbf{x}(t+1) = (\mathbf{I} - \alpha\mathbf{L})\mathbf{x}(t) \triangleq \mathbf{W}\mathbf{x}(t) \quad (3)$$

where $\mathbf{x}(t) \triangleq [x_1(t) \ x_2(t) \ \ldots \ x_N(t)]^T$, $x_n(t)$ denotes the state value of sensor n at iteration t. It can be easily verified that for $0 < \alpha < \frac{2}{\lambda_{\max}(\mathbf{L})}$, the above system converges to the average of the initial state, i.e. $\frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{x}(0)$, where $\lambda_{\max}(\mathbf{L})$ denotes the largest eigenvalue of $\mathbf{L}$. However, when quantization error[1] is present, a direct application of the consensus algorithm to the quantized data: $\mathbf{x}(t+1) = \mathbf{W}Q(\mathbf{x}(t))$, results in an error propagation, where $Q(\mathbf{x}(t)) \triangleq [Q(x_1(t)) \ \ldots \ Q(x_N(t))]^T$, $Q(x_n(t))$ denotes the quantized data of $x_n(t)$ using certain quantization scheme. To see this, we rewrite the above update equation as

$$\mathbf{x}(t+1) = \mathbf{W}\left(\mathbf{x}(t) + \mathbf{v}(t)\right) \quad (4)$$

where $\mathbf{v}(t) \triangleq [v_1(t) \ v_2(t) \ \ldots \ v_N(t)]^T$, $v_n(t) = Q(x_n(t)) - x_n(t)$ denotes the quantization error of sensor n introduced at iteration t. The state $\mathbf{x}(t)$ therefore can be expressed as

$$\mathbf{x}(t) = \mathbf{W}^t\mathbf{x}(0) + \sum_{i=0}^{t-1}\mathbf{W}^{t-i}\mathbf{v}(i). \quad (5)$$

Noting that $\mathbf{W}^t \to \frac{1}{N}\mathbf{1}\mathbf{1}^T$ as $t \to \infty$ [1], we have

$$\begin{aligned}\lim_{t\to\infty}\mathbf{x}(t) &= \lim_{t\to\infty}\left(\mathbf{W}^t\mathbf{x}(0) + \sum_{i=0}^{k}\mathbf{W}^{t-i}\mathbf{v}(i)\right. \\ &\quad \left. + \sum_{i=k+1}^{t-1}\mathbf{W}^{t-i}\mathbf{v}(i)\right) \\ &= \frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{x}(0) + \sum_{i=0}^{k}\frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{v}(i) \\ &\quad + \lim_{t\to\infty}\sum_{i=k+1}^{t-1}\mathbf{W}^{t-i}\mathbf{v}(i) \end{aligned} \quad (6)$$

where k can be any finite positive integer. From (6), we see that the quantization errors incurred at each iteration are preserved throughout the process and eventually contributes to the final state. As a result, the system usually does not converge to the desired initial average, even if a consensus can be achieved; in more severe cases, error accumulation may lead to an unbounded state divergent from the average of the initial state [5].

In [6], [7], the authors considered another consensus protocol in which each sensor, say sensor n, uses its local unquantized data, $x_n(t)$, instead of the quantized version, $Q(x_n(t))$, in updating its state, which leads to:

$$\begin{aligned}\mathbf{x}(t+1) &= (\mathbf{I} - \alpha\mathbf{D})\,\mathbf{x}(t) + \alpha\mathbf{A}Q(\mathbf{x}(t)) \\ &= \mathbf{W}\mathbf{x}(t) + \alpha\mathbf{A}\mathbf{v}(t). \end{aligned} \quad (7)$$

Since $\frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{A} = \frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{D} \neq \mathbf{0}$, by following a similar derivation, it can be readily shown that (7), as (4), also suffers from error propagation. This is the main reason that the algorithms [6], [7], although reaching a consensus under certain conditions, fail to converge to the average of the initial state.

To address this issue, we propose the following consensus protocol

$$\begin{aligned}\mathbf{x}(t+1) &= \mathbf{x}(t) - \alpha\mathbf{L}Q(\mathbf{x}(t)) \\ &= \mathbf{W}\mathbf{x}(t) - \alpha\mathbf{L}\mathbf{v}(t) \end{aligned} \quad (8)$$

where for each sensor, both the real data $x_n(t)$ and its quantized version $Q(x_n(t))$ are used in the update. This protocol, albeit simple, can suppress the error propagation to a certain extent. We have

$$\begin{aligned}\lim_{t\to\infty}\mathbf{x}(t) &= \lim_{t\to\infty}\left(\mathbf{W}^t\mathbf{x}(0) - \alpha\sum_{i=0}^{t-1}\mathbf{W}^{t-1-i}\mathbf{L}\mathbf{v}(i)\right) \\ &= \frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{x}(0) - \alpha\sum_{i=0}^{k}\frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{L}\mathbf{v}(i) \\ &\quad - \lim_{t\to\infty}\alpha\sum_{i=k+1}^{t-1}\mathbf{W}^{t-1-i}\mathbf{L}\mathbf{v}(i) \\ &\overset{(a)}{=} \frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{x}(0) - \lim_{t\to\infty}\alpha\sum_{i=k+1}^{t-1}\mathbf{W}^{t-1-i}\mathbf{L}\mathbf{v}(i) \end{aligned} \quad (9)$$

where (a) comes by noting that $\mathbf{1}^T\mathbf{L} = \mathbf{0}$. We observe that for a specific i, the quantization error component $\mathbf{v}(i)$ introduced at iteration i will eventually vanish as the system evolves over time. This observation implies that a convergence to the desired state may be achieved if the sequence $\{\mathbf{v}(t)\}$ satisfies a certain condition. Specifically, as [6], we model the quantization errors as spatially and temporally uncorrelated random variables[2]. We then have the following important result.

Proposition 1: Suppose that the quantization errors $\{v_n(t)\}$ have zero mean and their variance converge to zero, i.e. $E[v_n^2(t)] \to 0$ as $t \to \infty \ \forall n$, then the above iteration (8) achieves a consensus in a mean-squared sense. The consensus value is equal to the average of the initial state, i.e. $\frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{x}(0)$.

Proof: See Appendix A. ■

This result inspires us to propose an adaptive quantization (AQ) scheme which can adaptively adjust its quantization threshold and step-size by learning from previous states, in a way such that the quantization noise variance at each sensor decreases to zero.

III. Proposed Adaptive Quantization Scheme

In this section, we will first introduce the one-bit AQ scheme and then extend it to the case of multiple bits.

[1] We only consider quantization error in this paper and assume the quantized data are exchanged without error.

[2] In [12], Sripad *et al.* proved that the quantization noise can be modeled as a uniformly distributed random variable uncorrelated with the input message under a certain mild condition.

A. One-bit AQ

The AQ scheme involves encoding and decoding process. The received quantized (encoded) data has to be decoded before it is applied to the recursive update. Hence the quantized data $Q(x_n(t))$ we discussed in the consensus protocols correspond to the data decoded at the receiver, but not the encoded data at the transmitter.

Let us consider encoding first. For each sensor, say sensor n, it, firstly, uses two globally specified parameters: an initial threshold τ, and an initial quantization step-size Δ, to generate its one-bit encoded data (encoded data are also quantized data, we refer to them as encoded data in order to differentiate them from the quantized data we mentioned in the consensus protocols) of the first three iterations:

$$\begin{aligned} b_n(0) =& \mathrm{sgn}(x_n(0) - \tau) \\ b_n(1) =& \mathrm{sgn}(x_n(1) - \tau_n(1)) \\ b_n(2) =& \mathrm{sgn}(x_n(2) - \tau_n(2)) \end{aligned} \tag{10}$$

where

$$\begin{cases} \mathrm{sgn}\{x\} = -1, & \text{if } x \leq 0, \\ \mathrm{sgn}\{x\} = 1, & \text{if } x > 0. \end{cases}$$

$\tau_n(1) = \tau + b_n(0)\Delta$ and $\tau_n(2) = \tau_n(1) + b_n(1)\Delta$, $b_n(t)$ denotes the encoded data of sensor n at iteration t, $\tau_n(t)$ is the corresponding threshold used for quantization. At iteration $t \geq 2$, sensor n computes its threshold by performing accumulation of the previous bits, weighted by a variable step-size $\Delta_n(t)$:

$$\tau_n(t+1) = \tau_n(t) + b_n(t)\Delta_n(t), \tag{11}$$

where $\Delta_n(t)$ evolves using the following dynamic model:

$$\Delta_n(t) = \Delta_n(t-1)K^{b_n(t-1)b_n(t-2)} \qquad t = 2, 3, \ldots, \tag{12}$$

where $K > 1$ is a constant, and $\Delta_n(1) = \Delta$. Then sensor n uses $\tau_n(t)$ as a threshold to generate its current encoded data:

$$b_n(t) = \mathrm{sgn}(x_n(t) - \tau_n(t)) \tag{13}$$

The decoding of the AQ scheme is simple and described as follows. Suppose sensor m is one of the neighboring sensors of sensor n. After receiving the encoded data $b_n(t)$ from sensor n, sensor m recovers $\tau_n(t+1)$ and uses it as the decoded output data for the recursive update, i.e. $Q(x_n(t)) = \tau_n(t+1)$. The reconstruction of $\tau_n(t+1)$ can be inferred from the received encoded data $\{b_n(i)\}_{i=0}^{t}$, i.e. $\tau_n(1) = \tau + b_n(0)\Delta$, $\tau_n(2) = \tau_n(1) + b_n(1)\Delta$ and (11) for $t \geq 2$.

We can immediately recognize that the above process is reminiscent of the Delta modulation (DM) with variable step-size (VS), but is implemented in a distributed fashion to solve a consensus problem. As shown in [13], the VS-DM is able to adaptively adjust its step-size and encode a waveform with decreasing granular noise, which is exactly the property we desired. An example of VS-DM encoding is illustrated in Fig. 1. We see that the threshold, i.e. the quantized data, gets closer and closer to its true value as time evolves.

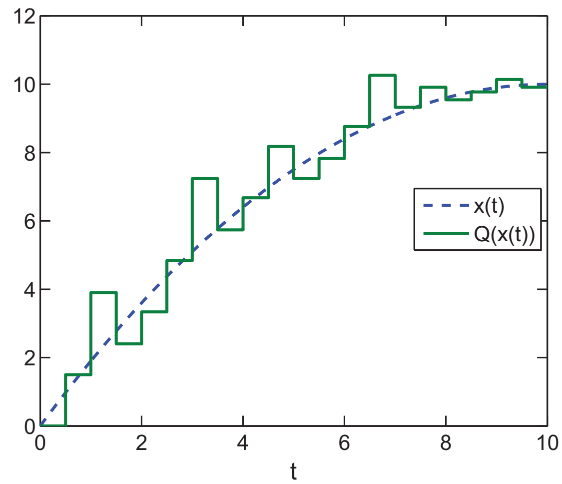

Fig. 1. An example of VS-DM encoding.

The AQ scheme can be easily implemented in a distributed fashion since it only involves very simple algebraic calculations. Also, unlike [6], no knowledge of the global topology is required at each sensor. On the other hand, for both encoding and decoding, not only current data but also information from prior runs are required. Hence, for each sensor, a moderate amount of memory resource has to be allocated to store the prior information for encoding and decoding. Note that each sensor needs to encode its own data and, at the same time, to decode the data received from its neighboring sensors.

B. Multiple-bit AQ

The extension of one-bit AQ to multiple bits is straightforward. At the encoding side, we replace the one-bit encoder: $b_n(t) = \mathrm{sgn}(x_n(t) - \tau_n(t))$ with a multiple-bit encoder. Let $m_n(t)$ denote the multiple-bit encoded data with bit length q. Using the current interval length, i.e. quantization step size $\Delta_n(t)$, we construct a set of data points: $S = \{\pm 2^{k-1}\Delta_n(t)\}_{k=1}^{q}$. The encoded data $m_n(t)$ is obtained by rounding the message $y_n(t) \triangleq x_n(t) - \tau_n(t)$ to its nearest data point in S, i.e.

$$m_n(t) = \frac{s_i}{\Delta_n(t)}$$
$$\text{where } s_i \in S \text{ and } |s_i - y_n(t)| = \min_j\{|s_j - y_n(t)|\} \tag{14}$$

The threshold $\tau_n(t+1)$ is updated as

$$\tau_n(t+1) = \tau_n(t) + m_n(t)\Delta_n(t) \tag{15}$$

where, as before, the quantization step size is set to be Δ for $t = 0, 1$ and evolves using the same dynamic model (12) when $t \geq 2$, with $b_n(t)$ denoting the sign of the encoded data $m_n(t)$, i.e. $b_n(t) = \mathrm{sgn}(m_n(t))$.

The decoding procedure is the same as that of the one-bit AQ. At each sensor, say sensor m, it reconstructs $Q(x_n(t)) = \tau_n(t+1)$ using the encoded information $\{m_n(i)\}_{i=0}^{t}$ received

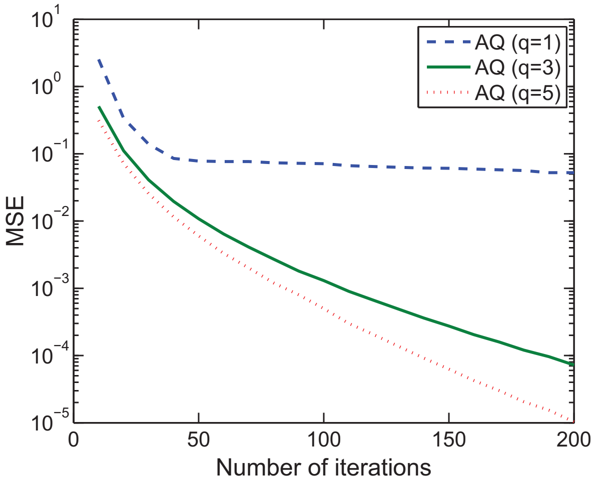

Fig. 2. MSE vs. the number of iterations, with the number of quantization bits q equal to 1, 3, 5, respectively.

from sensor n for the recursive update. We have $\tau_n(1) = \tau + m_n(0)\Delta$, $\tau_n(2) = \tau + m_n(1)\Delta$, and (15) for $t \geq 2$.

IV. Simulation Results

We present simulation results to illustrate the performance of our proposed AQ scheme. The sensor network is constructed using a random geographic graph model, in which $N = 25$ sensors are placed uniformly at random on a two-dimensional unit area and communicate with their neighbors within a radius r. The transmission radius r is set to be $\sqrt{\log N/N}$ to ensure that the graph is connected with a high probability. The initial values of the sensors, $\{x_i(0)\}$, are generated according to a Gaussian distribution with zero mean and unit variance. The performance is measured by an empirical mean-square error $\|\mathbf{x}(t) - \frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{x}(0)\|_2^2$, which is averaged over 1000 Monte Carlo runs, with the graph[3] and the initial state independently generated for each run. Fig. 2 shows the mean-square errors (MSEs) of our proposed algorithm as a function of the number of iterations N_{itr}, where we set $K = 1.2$ in (12) for quantization step size update, and the number of quantization bits, q, varies from 1 to 5. From Fig. 2, we see that, even with a small number of quantization bits, the proposed AQ scheme achieves a superior estimation performance within a moderate number of iterations. This shows the effectiveness of our proposed AQ scheme in solving the distributed consensus problem.

Appendix A
Proof of Proposition 1

Our objective is to show that

$$E[(\mathbf{x}(t) - \bar{\mathbf{x}}_0)^T(\mathbf{x}(t) - \bar{\mathbf{x}}_0)] \rightarrow 0 \quad \text{as } t \rightarrow \infty. \tag{16}$$

[3]Only connected graphs are counted in.

where $\bar{\mathbf{x}}_0 \triangleq \frac{1}{N}\mathbf{1}\mathbf{1}^T\mathbf{x}(0)$ and

$$\begin{aligned}\mathbf{x}(t) =& \mathbf{W}^t\mathbf{x}(0) - \alpha\sum_{i=0}^{t-1}\mathbf{W}^{t-1-i}\mathbf{L}\mathbf{v}(i)\\ \triangleq& \mathbf{f}_1 - \mathbf{f}_2\end{aligned} \tag{17}$$

Since $\mathbf{f}_1$ is deterministic and $\mathbf{f}_2$ is a matrix-weighted combination of the zero-mean random vectors $\{\mathbf{v}(i)\}$, we can decompose the mean-square error as

$$\begin{aligned}E[(\mathbf{x}(t) - \bar{\mathbf{x}}_0)^T(\mathbf{x}(t) - \bar{\mathbf{x}}_0)] =& (\mathbf{f}_1 - \bar{\mathbf{x}}_0)^T(\mathbf{f}_1 - \bar{\mathbf{x}}_0) + E\left[\mathbf{f}_2^T\mathbf{f}_2\right]\\ \triangleq& \epsilon_1 + \epsilon_2.\end{aligned} \tag{18}$$

Due to the manuscript length limit, we will not carry out a detailed analysis for computing ϵ_1 and ϵ_2. It turns out that we have the mean-square deviation upper-bounded by

$$\begin{aligned}&E[(\mathbf{x}(t) - \bar{\mathbf{x}}_0)^T(\mathbf{x}(t) - \bar{\mathbf{x}}_0)]\\ <&\rho^{2t}\|\mathbf{x}(0)\|_2^2 + \alpha^2 N\sum_{i=0}^{t-1}\lambda_{\max}(\mathbf{C}_{v,i})\lambda_{\max}^2(\mathbf{L})\rho^{2(t-1-i)}\end{aligned} \tag{19}$$

where ρ denotes the spectral radius of $(\mathbf{W} - \frac{1}{N}\mathbf{1}\mathbf{1}^T)$ and we can readily verify that $\rho < 1$; $\lambda_{\max}(\mathbf{L})$ and $\lambda_{\max}(\mathbf{C}_{v,i})$ denote the largest eigenvalue of $\mathbf{L}$ and $\mathbf{C}_{v,i}$, respectively. It can be shown that both terms on the right-hand side of (19) will vanish as $t \rightarrow \infty$ and the quantization noise variance at each sensor converges to zero, i.e. there exists t_0 such that for any arbitrary small $\epsilon > 0$, $\lambda_{\max}(\mathbf{C}_{v,t}) < \epsilon$ holds for $t > t_0$.

References

[1] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems and Control Letters*, no. 53, pp. 65–78, Feb. 2004.

[2] D. S. Scherber and H. C. Papadopoulos, "Distributed computation of averages over ad hoc networks," *IEEE Journal on Sel. Areas In Comm.*, vol. 23, pp. 776–787, April 2005.

[3] E. Kokiopoulou and P. Frossard, "Accelerating distributed consensus using extrapolation," *IEEE Signal Processing Letters*, vol. 14, pp. 665–668, October 2007.

[4] C. C. Moallemi and B. V. Roy, "Consensus propagation," *IEEE Trans. Inform. Theory*, no. 11, pp. 4753–4766, Nov. 2006.

[5] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *Proceedings of the 17th international symposium on mathematical theory of networks and systems*, 2006.

[6] M. E. Yildiz and A. Scaglione, "Coding with side information for rate constrained consensus," *IEEE Trans. Signal Processing*, to appear, Available: http://people.ece.cornell.edu/scaglione/.

[7] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks: quantized data," *submitted to IEEE Tran. Signal Processing*, 2007, Available: http://arxiv.org/PS_cache/arxiv/pdf/0712/0712.1609v1.pdf.

[8] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in ad hoc WSNs with noisy links - Part I: distributed estimation of deterministic signals," *IEEE Trans. Signal Processing*, no. 1, pp. 350–364, Jan. 2008.

[9] S. Barbarossa and G. Scutari, "Decentralized maximum-likelihood estimation for sensor networks composed of nonlinearly coupled dynamical systems," *IEEE Trans. Signal Processing*, no. 7, pp. 3456–3470, July 2007.

[10] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus," *Automatica*, pp. 1192–1203, 2007.

[11] F. R. K. Chung, *Spectral Graph Theory*. CBMS Regional Conference Series in Mathematics, No. 92, 1997.

[12] A. B. Sripad and D. L. Snyder, "A necessary and sufficient condition for quantization errors to be uniform and white," *IEEE Trans. Acoust., Speech, Signal Processing*, no. 5, pp. 442–448, Oct. 1977.

[13] J. G. Proakis, *Digital Communications, 3rd ed.* Mcgraw-Hill College, 1995.