Distributed Average Consensus with Dithered Quantization

Tuncer C. Aysal, Mark J. Coates, and Michael G. Rabbat
Telecommunications and Signal Processing–Computer Networks Lab.
Department of Electrical and Computer Engineering
McGill University, Montreal, QC
{tuncer.aysal,mark.coates,michael.rabbat}@mcgill.ca

Abstract

In this paper, we develop algorithms for distributed computation of averages of the node data over networks with bandwidth/power constraints or large volumes of data. Distributed averaging algorithms fail to achieve consensus when deterministic uniform quantization is adopted. We propose a distributed algorithm in which the nodes utilize probabilistically quantized information, *i.e.*, dithered quantization, to communicate with each other. The algorithm we develop is a dynamical system that generates sequences achieving a consensus at one of the quantization values almost surely. In addition, we show that the expected value of the consensus is equal to the average of the original sensor data. We derive an upper bound on the mean square error performance of the probabilistically quantized distributed averaging (PQDA). Moreover, we show that the convergence of the PQDA is monotonic by studying the evolution of the minimum-length interval containing the node values. We reveal that the length of this interval is a monotonically non–increasing function with limit zero. We also demonstrate that all the node values, in the worst case, converge to the final two quantization bins at the same rate as standard unquantized consensus. Finally, we report the results of simulations conducted to evaluate the behavior and the effectiveness of the proposed algorithm in various scenarios.

Index Terms

Distributed algorithms, average consensus, sensor networks, probabilistic quantization, dithering.

Portions of this paper were presented at the IEEE Statistical Signal Processing Workshop, Madison, WI, Aug. 2007 [1] and the Allerton Conference on Communication, Control, and Computing, Monticello, IL, Sep. 2007 [2].

I. Introduction

Ad hoc networks of autonomous sensors and actuators are attractive solutions for a broad range of applications. Such networks find use in civilian and military applications, including target tracking and surveillance for robot navigation, source localization, weather forecasting, medical monitoring and imaging. In general, the networks envisioned for many of these applications involve large numbers of possibly randomly distributed inexpensive sensors, with limited sensing, processing and communication power on board. In many of the applications, limitations in bandwidth, sensor battery power and computing resources place tight constraints in the rate and form of information that can be exchanged [3]–[5]. Other applications such as camera networks and distributed tracking demand communication of large volumes of data. When the power and bandwidth constraints, or large volume data sets are considered, communication with unquantized values is impractical.

Distributed average consensus – the task of ensuring that all nodes in a network are aware of the average of a set of network-wide measurements – is a fundamental problem in ad hoc network applications, including distributed agreement and synchronization problems [6], distributed coordination of mobile autonomous agents [7], [8], and distributed data fusion in sensor networks [3], [9], [10]. It is also a central topic for load balancing (with divisible tasks) in parallel computers [11], [12]. Our previous work has illustrated how distributed average consensus can be used for two distributed signal processing tasks: source localization [13], and data compression [14]. Decentralized data compression, in particular, requires the computation of many consensus values in parallel (one for each compression coefficient). By appropriately quantizing each coefficient, multiple coefficients can be transmitted in a single packet, leading to a significantly more efficient implementation.

Distributed averaging algorithms are extremely attractive for applications in wirelessly networked systems because nodes only exchange information and maintain state for their immediate neighbors. Consequently, there is no need to establish or maintain complicated routing structures. Also, there is no single bottleneck link (as in a tree) where the result of in-network computation can be compromised or lost or jammed by an adversary. Finally, consensus algorithms have the attractive property that, at termination, the computed value is available throughout the network, so a network user can query any node and immediately receive a response, rather than waiting for the query and response to propagate to and from a fusion center.

In both wireless sensor and peer–to–peer networks, there is interest in simple protocols for computing aggregate statistics [15]–[18]. In this paper we focus on a particular class of iterative algorithms for average consensus. Each node updates its state with a weighted sum of values from neighboring nodes, *i.e.*,

$$x_i(t+1) = W_{ii}x_i(t) + \sum_{j\in\mathcal{N}_i} W_{ij}x_j(t) \tag{1}$$

for $i = 1, 2, \ldots, N$ and $t = 0, 1, \ldots$. Here W_{ij} is a weight associated with the edge $\{i, j\}$ and N is the total number

of nodes. These weights are algorithm parameters [3], [9]. Furthermore, $\mathcal{N}_i$ denotes the set of nodes that have a direct (bidirectional) communication link with node i. The state at each node in the iteration consists of a single real number, which overwrites the previous value. The algorithm parameters $\mathbf{W}$ are time–independent, *i.e.*, do not depend on t. Under easily-verified conditions on $\mathbf{W}$ it is easy to show that the value $x_i(t)$ at each node converges to $1/N \sum_{i=1}^{N} x_i(0)$ asymptotically as $t \to \infty$.

Xiao, Boyd and Kim extended the distributed consensus algorithm to admit noisy communication links where each node updates its local variable with a weighted average of its neighbors values, and each new value is corrupted by an additive noise with zero mean [19]:

$$x_i(t+1) = W_{ii} x_i(t) + \sum_{j \in \mathcal{N}_i} W_{ij} x_j(t) + w_i(t) \tag{2}$$

where $w_i(t)$ is the additive zero–mean noise with fixed variance. They pose and solve the problem of designing weights $W_{i,j}$ that lead to optimal steady-state behavior, based on the assumption that the noise terms $w_i(t)$ are independent.

A. Related Work

While there exists a substantial body of work on average consensus protocols with infinite precision and noise–free peer–to–peer communications, little research has been done introducing distortions in the message exchange. Recently, Yildiz and Scaglione, in [20], explored the impact of quantization noise through modification of the consensus algorithm proposed by Xiao, Boyd and Kim [19]. They note that the noise component in (2) can be considered as the quantization noise and they develop an algorithm for predicting neighbors' unquantized values in order to correct errors introduced by quantization [20]. Simulation studies for small N indicate that if the increasing correlation among the node states is taken into account, the variance of the quantization noise diminishes and nodes converge to a consensus.

Kashyap *et al.* examine the effects of quantization in consensus algorithms from a different point of view [21]. They require that the network average, $\bar{x} = 1/N \sum_{i=1}^{N} x_i(t)$, be preserved at every iteration. To do this using quantized transmissions, nodes must carefully account for round-off errors. Suppose we have a network of N nodes and let Δ denote the "quantization resolution" or distance between two quantization lattice points. If $\bar{x}$ is not a multiple of $N\Delta$, then it is not possible for the network to reach a strict consensus (*i.e.*, $\lim_{t\to\infty} \max_{i,j} |x_i(t) - x_j(t)| = 0$) while also preserving the network average, $\bar{x}$, since nodes only ever exchange units of Δ. Instead, Kashyap *et. al* define the notion of a "quantized consensus" to be such that all $x_i(t)$ take on one of two neighboring quantization values while preserving the network average; *i.e.*, $x_i(t) \in \{k\Delta, (k+1)\Delta\}$ for all i and some k, and $\sum_i x_i(T) = N\bar{x}$. They show that, under reasonable conditions, their algorithm will converge to a quantized

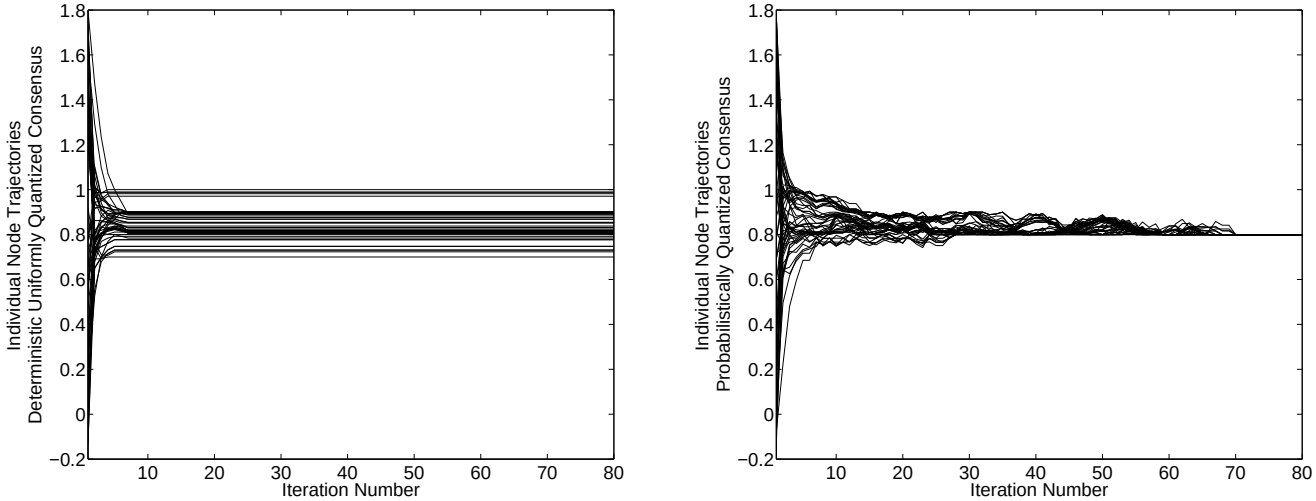
consensus. However, the quantized consensus is clearly not a strict consensus, *i.e.*, all nodes do not have the same value. Even when the algorithm has converged, as many as half the nodes in the network may have different values. If nodes are strategizing and/or performing actions based on these values (*e.g.*, flight formation), then differing values may lead to inconsistent behavior.

Of note is that both related works discussed above utilize standard deterministic uniform quantization schemes to quantize the data. In contrast to [20], where quantization noise terms are modeled as independent zero-mean random variables, we explicitly introduce randomization in our quantization procedure, *i.e.*, "dithering". Careful analysis of this randomization allows us to provide concrete theoretical rates of convergence in addition to empirical results. Moreover, the algorithm proposed in this paper converges to a strict consensus, as opposed to the approximate "quantized consensus" achieved in [21]. In addition to proving that our algorithm converges, we show that the network average is preserved in expectation, and we characterize the limiting mean squared error between the consensus value and the network average.

B. Summary of Contributions

Constraints on sensor cost, bandwidth, and energy budget dictate that information transmitted between nodes has to be quantized in practice [3], [4]. In this paper, we propose a simple distributed and iterative scheme to compute the average at each sensor node utilizing only quantized information communication. Standard, deterministic uniform quantization does not lead to the desired result. Although the standard distributed averaging algorithm converges to a fixed point when deterministic uniform quantization is used, it fails to converge to a consensus as illustrated in Fig. 1(a). Instead, we adopt the probabilistic quantization (PQ) scheme described in [4]. PQ has been shown to be very effective for estimation with quantized data since the noise introduced by PQ is zero-mean [4]. This makes PQ suitable for average–based algorithms. As shown in Section II, the PQ algorithm is a form dithered quantization method. Dithering has been widely recognized as a method to render the quantization noise independent of the quantized data, reducing some artifacts created by deterministic quantization and there is a vast literature on the topic, see [22] and reference therein.

In the scheme considered here, each node exchanges quantized state information with its neighbors in a simple and bidirectional manner. This scheme does not involve routing messages in the network; instead, it diffuses information across network by updating each node's data with a weighted average of its neighbors' quantized ones. We do not burden the nodes with extensive computations, and we provide theoretical results, *i.e.*, we show here that the distributed average computation utilizing probabilistic consensus indeed achieves a consensus almost surely (Fig. 1), and the consensus is one of the quantization levels. Furthermore, the expected value of the achieved consensus is equal to the desired value, *i.e.*, the average of the initial analog node measurements. We also give an



(a) (b)

Fig. 1. Individual node trajectories (*i.e.*, $x_i(t), \forall i$) taken by the distributed average consensus using (a:) deterministic uniform quantization and (b:) probabilistic quantization. The number of nodes is $N = 50$, the nodes' initial average is $\overline{\mathbf{x}}(0) = 0.85$, and the quantization resolution is set to $\Delta = 0.1$. The consensus value, in this case, is 0.8.

upper bound on the mean square error performance of the probabilistically quantized distributed averaging (PQDA) algorithm.

We also investigate the evolution with time of the interval occupied by the node values. Specifically, we show that the size of this interval is a monotonically non–increasing function, with limit zero. These results indicate that the convergence of the PQDA algorithm is monotonic in the sense that the global trend of the node values is towards the consensus. Moreover, we show here that all the node values, in the worst case, arrive in the final two quantization bins at the same rate as standard unquantized consensus. Of note is that there is always a non–zero probability of achieving consensus when all the node values are in the final two bins. Finally, we present simulation results evaluating the proposed algorithm in varying scenarios and showing the effectiveness of the PQDA algorithm.

C. Paper Organization

The remainder of this paper is organized as follows. Section II reviews the graph theoretical concepts, introduces the distributed average consensus problem along with the probabilistic quantization scheme, and, reveals the connections between probabilistic quantization and dithering theory. The proposed algorithm, along with its properties, is detailed in Section III. Section IV presents results regarding the convergence characteristics of the proposed PQDA algorithm. Numerical examples evaluating the performance of the proposed algorithm in varying scenarios are provided in Section V. Some extensions of the proposed algorithm along with additional practical considerations are detailed in Section VI. Finally, we conclude with Section VII.

II. Preliminaries

In the following, the distributed average consensus problem is formulated utilizing the probabilistic quantization concept. We first review some basic concepts from graph theory and then formulate the consensus problem in which the nodes communicate with quantized data. Finally, we provide a brief review of probabilistic quantization and reveal the connections between probabilistic quantization and dithering theory.

A. Review of Graph Theoretical Concepts

Let $G = (\mathcal{V}, \mathcal{E})$ be a graph consisting of a set of vertices, $\mathcal{V}$, and a set of edges, $\mathcal{E}$. Let $N = |\mathcal{V}|$ denote the number of vertices, where $|\cdot|$ denotes the cardinality. We denote an edge between vertices i and j as an unordered pair $(i, j) \in \mathcal{E}$. The presence of an edge between two vertices indicates that they can establish bidirectional noise–free communication with each other. We assume that transmissions are always successful and that the topology is fixed. We assume connected network topologies and the connectivity pattern of the graph is given by the $N \times N$ adjacency matrix $\mathbf{A} = [A_{ij}]$, where

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} . \tag{3}$$

Moreover, we denote the neighborhood of the node i by $\mathcal{N}_i \triangleq \{j \in \mathcal{V} : i \neq j, (i,j) \in \mathcal{E}\}$. Also, the degree of the node i is given by $\deg_i = |\mathcal{N}_i|$.

B. Distributed Average Consensus

We consider a set of nodes of a network (vertices of the graph), each with an initial real valued scalar $y_i \in [-U, U]$, where $i = 1, 2, \ldots, N$. Let $\mathbf{1}$ denote the vector of ones. Our goal is to develop a distributed iterative algorithm that computes the value $\overline{\mathbf{y}} \triangleq (N)^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{y}$ at every node in the network, while using quantized communication. We hence aim to design a system such that the states at all nodes converge to a consensus and the expectation of the consensus achieved, in the limit, is the average of the initial states.

The average of node measurements is a sufficient statistic for many problems of interest. The following two remarks briefly discusses two examples.

Remark 1. *When the observations follow $y_i = \theta + n_i$ where $i = 1, 2, \ldots, N$ and θ is the scalar to be estimated, and the noises, $\{n_i : i = 1, 2, \ldots, N\}$, are independent and identically distributed (i.i.d.) zero–mean Gaussian with variance σ^2, the maximum–likelihood estimate is given by the average, $\hat{\theta} = (N)^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{y}$ with the associated mean square error σ^2/N.*

Remark 2. *Suppose node measurements $\{y_i : i = 1, 2, \ldots, N\}$ are i.i.d. conditioned on some hypothesis H_j, with $j \in \{0, 1\}$, where $\Pr\{H_0\} = \alpha$. Let, $z_i \triangleq \log\{\Lambda(y_i)\}$, where $\Lambda(y_i) \triangleq \Pr\{y_i|H_1\}/\Pr\{y_i|H_0\}$. Then, the optimal*

decision is to perform the following detection rule: $\overline{\mathbf{z}} \gtrless_{H_0}^{H_1} \alpha/(1-\alpha)$ *where* $\overline{\mathbf{z}} \triangleq (N)^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{z}$.

C. Probabilistic Quantization and Dithering

In the following, we present a brief review of the quantization scheme adopted in this paper. Suppose that the scalar value $x_i \in \mathbb{R}$ is bounded to a finite interval $[-U, U]$. Furthermore, suppose that we wish to obtain a quantized message q_i with length l bits, where l is application dependent. We therefore have $L = 2^l$ quantization points given by the set $\boldsymbol{\tau} = \{\tau_1, \tau_2, \ldots, \tau_L\}$ where $\tau_1 = -U$ and $\tau_L = U$. The points are uniformly spaced such that $\Delta = \tau_{j+1} - \tau_j$ for $j \in \{1, 2, \ldots, L-1\}$. It follows that $\Delta = 2U/(2^l - 1)$. Now suppose $x_i \in [\tau_j, \tau_{j+1})$ and let $q_i \triangleq \mathcal{Q}(x_i)$ where $\mathcal{Q}(\cdot)$ denotes the PQ operation. Then x_i is quantized in a probabilistic manner:

$$\Pr\{q_i = \tau_{j+1}\} = r, \text{ and, } \Pr\{q_i = \tau_j\} = 1 - r \tag{4}$$

where $r = (x_i - \tau_j)/\Delta$. Of note is that when the variable to quantize is exactly equal to a quantization centroid, there is zero probability of choosing another centroid. The following lemma, adopted from [4], discusses two important properties of PQ.

Lemma 1. *[4] Suppose $x_i \in [\tau_j, \tau_{j+1})$ and let q_i be an l–bit quantization of $x_i \in [-U, U]$. The message q_i is an unbiased representation of x_i,* i.e.,

$$\mathbb{E}\{q_i\} = x_i, \textit{ and, } \mathbb{E}\{(q_i - x_i)^2\} \leq \frac{U^2}{(2^l - 1)^2} \equiv \frac{\Delta^2}{4}. \tag{5}$$

As noted in the following lemma, a careful observation shows that probabilistic quantization is equivalent to a "dithered quantization" method.

Lemma 2. *Suppose $x_i \in [\tau_j, \tau_{j+1})$ and let $q_i \triangleq \mathcal{Q}(x_i)$. Probabilistic quantization is equivalent to the following dithered quantization scheme:*

$$q_i = \min_j |\tau_j - (x_i + u)| \tag{6}$$

where u is a uniform random variable with support on $[-\Delta/2, \Delta/2]$.

Proof: Without loss of generality, suppose $|\tau_j - x_i| < |\tau_{j+1} - x_i|$. Moreover, suppose we are utilizing a deterministic uniform quantizer. Then,

$$\Pr\{q_i = \tau_j\} = \Pr\left\{\tau_j - \frac{\Delta}{2} \leq x_i + u \leq \tau_j - \frac{\Delta}{2}\right\} \tag{7}$$

$$= \Pr\left\{u \leq \tau_j + \frac{\Delta}{2} - x_i\right\} \tag{8}$$

$$= \frac{\tau_{j+1} - x_i}{\Delta}. \tag{9}$$

Note that the last line is equivalent to $1-r$, so the proof is complete. ■

Thus, before we perform any quantization, we add uniform random variable u with support defined on $[-\Delta/2, \Delta/2]$ and we form $x_i' = u + x_i$. Now, performing standard deterministic uniform quantization, *i.e.*, letting $q_i = \min_j |\tau_j - x_i'|$, yields quantized values, q_i's that are statistically identical to the ones of the probabilistic quantization. Thus, probabilistic quantization is a form of dithering where one, before performing standard deterministic uniform quantization, adds a uniform random variable with support equal to the quantization bin size. This is a substractively dithered system [22]. It has been shown by Schuchman that the substractive dithering process utilizing uniform random variable with support on $[-\Delta/2, \Delta/2]$ yields error signal values that are statistically independent from each other and the input [23].

III. Distributed Average Consensus with Probabilistically Quantized Communication

In the following, we propose a quantized distributed average consensus algorithm and incorporate PQ into the consensus framework for networks. Furthermore, we analyze the effect of PQ on the consensus algorithm. Specifically, we present theorems revealing the limiting consensus, expectation and mean square error of the proposed PQDA algorithm.

At $t = 0$ (after all sensors have taken the measurement), each node initializes its state as $x_i(0) = y_i$, *i.e.*, $\mathbf{x}(0) = \mathbf{y}$ where $\mathbf{x}(0)$ denotes the initial states at the nodes. It then quantizes its state to generate $q_i(0) = \mathcal{Q}(x_i(0))$. At each following step, each node updates its state with a linear combination of its own quantized state and the quantized states at its neighbors

$$x_i(t+1) = W_{ii} q_i(t) + \sum_{j \in \mathcal{N}_i} W_{ij} q_j(t) \tag{10}$$

for $i = 1, 2, \ldots, N$, where $q_j(t) = \mathcal{Q}(x_j(t))$, and t denotes the time step. Also, W_{ij} is the weight on $x_j(t)$ at node i. Moreover, setting $W_{ij} = 0$ whenever $\Phi_{ij} = 0$, the distributed iterative process reduces to the following recursion

$$\mathbf{x}(t+1) = \mathbf{W}\mathbf{q}(t) \tag{11}$$

where $\mathbf{q}(t)$ denotes the quantized state vector, followed by

$$\mathbf{q}(t+1) = \mathcal{Q}(\mathbf{x}(t+1)). \tag{12}$$

The PQDA algorithm hence refers to the iterative algorithm defined by (11) and (12). In the sequel, we assume that $\mathbf{W}$, the weight matrix, is symmetric, non-negative and satisfies the conditions required for asymptotic average consensus without quantization [19]:

$$\mathbf{W}\mathbf{1} = \mathbf{1},\ \mathbf{1}^{\mathrm{T}}\mathbf{W} = \mathbf{1}^{\mathrm{T}}, \text{ and, } \rho(\mathbf{W} - \mathbf{J}) < 1, \tag{13}$$

where $\rho(\mathbf{U})$ denotes the spectral radius of a matrix $\mathbf{U}$ (*i.e.*, the largest eigenvalue of $\mathbf{U}$ in absolute value), and $\mathbf{J} \triangleq (N)^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}}$, where $\mathbf{J}\mathbf{x}$ projects $\mathbf{x}$ onto the N–dimensional "diagonal" subspace (*i.e.*, the set of vectors in $\mathbb{R}^N$ corresponding to a strict consensus). Weight matrices satisfying the required convergence conditions are easy to find if the underlying graph is connected and non–bipartite, *e.g.*, Maximum–degree and Metropolis weights [19].

The following theorem considers the convergence of the probabilistically quantized distributed average computation.

Theorem 1. *The probabilistically quantized distributed iterative process achieves a consensus, almost surely,* i.e.,

$$\Pr\left\{\lim_{t\to\infty}\mathbf{x}(t) = c\mathbf{1}\right\} = 1 \tag{14}$$

where $c \in \boldsymbol{\tau}$.

Proof: Without loss of generality, we focus on integer quantization in the range $[1, m]$. Define $\mathcal{M}$ as the discrete Markov chain with initial state $\mathbf{q}(0)$ and transition matrix defined by the combination of the deterministic transformation $\mathbf{x}(t+1) = \mathbf{W}\mathbf{q}(t)$ and the probabilistic quantizer $\mathbf{q}(t+1) \sim \Pr\{\mathbf{q}(t+1)|\mathbf{x}(t+1)\}$.

Let $\mathcal{S}_0$ be the set of quantization points that can be represented in the form $q\mathbf{1}$ for some integer q and denote by $\mathcal{S}_k$ the set of quantization points with Manhattan distance k from $\mathcal{S}_0$. Moreover, let $\mathcal{C}(\mathbf{q})$ be the open hypercube centered at $\mathbf{q}$ and defined as $(q_1 - 1, q_1 + 1) \times (q_2 - 1, q_2 + 1) \times ... \times (q_N - 1, q_N + 1)$. Here q_k denotes the k–th coefficient of $\mathbf{q}$. Note that any point in $\mathcal{C}(\mathbf{q})$ has a non-zero probability of being quantized to $\mathbf{q}$. Let

$$\mathcal{A}_k = \bigcup_{\mathbf{q}\in\mathcal{S}_k} \mathcal{C}(\mathbf{q}). \tag{15}$$

The consensus operator has the important property that $|\mathbf{W}\mathbf{q} - \mu(\mathbf{W}\mathbf{q})| < |\mathbf{q} - \mu(\mathbf{q})|$ for $|\mathbf{q} - \mu(\mathbf{q})| > 0$, where $\mu(\cdot)$ denotes the projection of its argument onto the $\mathbf{1}$–vector. Moreover, $c\mathbf{W}\mathbf{1} = c\mathbf{1}$. The latter property implies that $\mathbf{q} \in \mathcal{S}_0$ is an absorbing state, since $\mathcal{Q}(\mathbf{x}(t+1)) = \mathcal{Q}(\mathbf{W}\mathbf{q}(t)) = \mathcal{Q}(\mathbf{q}(t)) = \mathbf{q}(t)$. The former property implies that there are no other absorbing states, since $\mathbf{x}(t+1)$ cannot equal $\mathbf{q}(t)$ (it must be closer to the $\mathbf{1}$–vector). This implies, from the properties of the quantizer $\mathcal{Q}$, that there is a non-zero probability that $\mathbf{q}(t+1) \neq \mathbf{q}(t)$.

In order to prove that $\mathcal{M}$ is an absorbing Markov chain, it remains to show that it is possible to reach an absorbing state from any other state. We prove this by induction, demonstrating first that

$$\Pr\{\mathbf{q}(t+1) \in \mathcal{S}_0 | \mathbf{q}(t) \in \mathcal{S}_1\} > 0 \tag{16}$$

and subsequently that

$$\Pr\{\mathbf{q}(t+1) \in \bigcup_{i=0}^{k-1} S_i | \mathbf{q}(t) \in \mathcal{S}_k\} > 0. \tag{17}$$

Define the open set $\mathcal{V}_k$ as

$$\mathcal{V}_k = \left\{ \mathbf{x} : |\mathbf{x} - \mu(\mathbf{x})| < k\sqrt{N-1}/\sqrt{N} \right\}. \tag{18}$$

To commence, observe that $\mathcal{V}_1 \subset \mathcal{A}_0$. The distance $|\mathbf{q} - \mu(\mathbf{q})| = \sqrt{N-1}/\sqrt{N}$ for $\mathbf{q} \in \mathcal{S}_1$. Hence, if $\mathbf{q}(t) \in \mathcal{S}_1$, $\mathbf{x}(t+1) = \mathbf{W}\mathbf{q}(t) \in \mathcal{V}_1 \in \mathcal{A}_0$ and $\Pr\{\mathbf{q}(t+1) \in \mathcal{S}_0\} > 0$. Similarly, the set

$$\mathcal{V}_k \subset \bigcup_{i=0}^{k-1} \mathcal{A}_i \tag{19}$$

is contained in the union of the first k hypercubes, $\mathcal{A}_i$, $i = 0, 1, \ldots, k-1$. The maximum distance $|\mathbf{q} - \mu(\mathbf{q})|$ for any point $\mathbf{q} \in \mathcal{S}_k$ is $k\sqrt{N-1}/\sqrt{N}$. This implies that

$$\mathbf{x}(t+1) = \mathbf{W}\mathbf{q}(t) \in \mathcal{V}_k \in \bigcup_{i=0}^{k-1} \mathcal{A}_i. \tag{20}$$

There is thus some $i < k$ and some $\mathbf{q} \in \mathcal{S}_i$ such that $\Pr\{\mathcal{Q}(\mathbf{x}(t+1)) = \mathbf{q}\} > 0$. This argument implies that for any starting state $\mathbf{q}(0)$ such that $\mathbf{q}(0) \in \mathcal{S}_k$ for some k, there exists a sequence of transitions with non–zero probability whose application results in absorption. ■

The theorem reveals that the probabilistically quantized distributed process indeed achieves a strict consensus at one of the quantization values. It is of interest to note that the stationary points of the PQDA algorithm are in the form of $c\mathbf{1}$ where $c \in \boldsymbol{\tau}$. We, hence, construct an absorbing Markov chain where the absorbing states are given by the stationarity points and show that for any starting state, there exists a sequence of transitions with non–zero probability whose application results in absorption. The following theorem discusses the expectation of the limiting random vector, *i.e.*, the expected value of $\mathbf{x}(t)$ as t tends to infinity.

Theorem 2. *The expectation of the limiting random vector is given by*

$$\mathbb{E}\left\{ \lim_{t\to\infty} \mathbf{x}(t) \right\} = (N)^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}}\mathbf{x}(0). \tag{21}$$

Proof: Note that $||\mathbf{x}(t)|| \leq \sqrt{N}U$, for $t \geq 0$, and, $\{x_i(t) : i = 1, 2, \ldots, N\}$ is bounded for all t. Moreover, from Theorem 1, we know that the random vector sequence $\mathbf{x}(t)$ converges in the limit, *i.e.*, $\lim_{t\to\infty} \mathbf{x}(t) = c\mathbf{1}$ for some $c \in \boldsymbol{\tau}$. Thus, by the Lebesgue dominated convergence theorem [24], we have

$$\mathbb{E}\left\{ \lim_{t\to\infty} \mathbf{x}(t) \right\} = \lim_{t\to\infty} \mathbb{E}\{\mathbf{x}(t)\}. \tag{22}$$

In the following, we derive $\lim_{t\to\infty} \mathbb{E}\{\mathbf{x}(t)\}$ and utilize the above relationship to arrive at the desired result.

In terms of quantization noise $\mathbf{v}(t)$, we can write $\mathbf{q}(t) = \mathbf{x}(t) + \mathbf{v}(t)$. The distributed iterative process reduces

to the following recursion $\mathbf{x}(t+1) = \mathbf{W}\mathbf{x}(t) + \mathbf{W}\mathbf{v}(t)$. Repeatedly utilizing the state recursion gives

$$\mathbf{x}(t) = \mathbf{W}^t\mathbf{x}(0) + \sum_{j=0}^{t-1} \mathbf{W}^{t-j}\mathbf{v}(j). \tag{23}$$

Taking the statistical expectation of $\mathbf{x}(t)$ as $t \to \infty$ and noting that the only random variables are $\mathbf{v}(j)$ for $j = 0, 1, \ldots, t-1$, yields

$$\lim_{t\to\infty} \mathbb{E}\{\mathbf{x}(t)\} = \lim_{t\to\infty} \mathbf{W}^t\mathbf{x}(0) + \sum_{j=0}^{t-1} \mathbf{W}^{t-j}\mathbb{E}\{\mathbf{v}(j)\} \tag{24}$$

$$= \lim_{t\to\infty} \mathbf{W}^t\mathbf{x}(0) \tag{25}$$

since $\mathbb{E}\{\mathbf{v}(j)\} = \mathbf{0}$ for $j = 0, 1, \ldots, t-1$; a corollary of Lemma 1. Furthermore, noting that $\lim_{t\to\infty} \mathbf{W}^t = (N)^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}}$ gives

$$\lim_{t\to\infty} \mathbb{E}\{\mathbf{x}(t)\} = (N)^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}}\mathbf{x}(0). \tag{26}$$

Recalling (22) gives the desired result. ■

This result indicates that the expectation of the limiting random vector is indeed equal to the initial analog node measurements' average. Furthermore, this theorem, combined with the previous one, indicates that the consensus value, c, is a discrete random variable with support defined by $\boldsymbol{\tau}$, and whose expectation is equal to the average of the initial states.

After establishing that the consensus value is a random variable with the desired expectation in the limit, the next natural quantity of interest is the limiting mean squared error, *i.e.*, the limiting average squared distance of the consensus random variable from the desired initial states' average value. The following theorem, thus, considers the expectation of the error norm of probabilistically quantized consensus as t and N tend to infinity.

Theorem 3. *Let us define* $\overline{\mathbf{x}}(0) \triangleq \mathbf{J}\mathbf{x}(0)$. *The expectation of the error norm of the probabilistically quantized distributed average consensus is asymptotically bounded by*

$$\lim_{t\to\infty} \lim_{N\to\infty} (\sqrt{N})^{-1}\mathbb{E}\{||\mathbf{q}(t) - \overline{\mathbf{x}}(0)||\} \leq \frac{\Delta}{2}\frac{1}{1-\rho(\mathbf{W}-\mathbf{J})} \tag{27}$$

where $\rho(\cdot)$ *denotes the spectral radius of its argument.*

Proof: See Appendix A. ■

The proof exploits averaging characteristics of $\mathbf{W}$, properties of norm operators, and uses a Law of Large Numbers argument to bound the error contributed by quantization noise.

Note that the upper bound decreases with decreasing spectral radius of $(\mathbf{W}-\mathbf{J})$, where a smaller (larger) $\rho(\mathbf{W}-\mathbf{J})$ can be, in a loose manner, interpreted as better (worse) "averaging ability" of the weight matrix. Furthermore, as

expected, the upper bound on the error norm increases with decreasing quantization resolution (*i.e.*, increasing Δ).

IV. Convergence Characteristics of Probabilistically Quantized Distributed Average Consensus

The convergence characteristics of the PQDA are essential for further understanding of the algorithm. In the following, we consider the evolution of the intervals occupied by the quantized and unquantized state values. Interestingly, we reveal that the length of the smallest interval containing all of the quantized state values, (*i.e.*, the range of the quantized state values) is non–increasing with a limit of zero as the time step tends to infinity. Moreover, we show that size of the minimum length interval, with boundaries constrained to the quantization points, that contains all of the unquantized node state values, is also non–increasing. This also has limit zero as the time step tends to infinity.

Let us denote the smallest and largest order statistics of any vector $\mathbf{u} \in \mathbb{R}^N$ as $u_{(1)} \triangleq \min_i\{u_i\}$ and $u_{(N)} \triangleq \max_i\{u_i\}$, respectively. Furthermore, let $\mathrm{I}[\mathbf{x}(t)]$ denote the interval of the node state values at time t, *i.e.*, the interval in which $\{x_i(t) : i = 1, 2, \ldots, N\}$ values lie,

$$\mathrm{I}[\mathbf{x}(t)] \triangleq [x_{(1)}(t), x_{(N)}(t)] \tag{28}$$

and $\mathrm{I}[\mathbf{q}(t)]$ denote the domain of the quantized node state values at time t, *i.e.*,

$$\mathrm{I}[\mathbf{q}(t)] \triangleq [q_{(1)}(t), q_{(N)}(t)]. \tag{29}$$

Moreover, let

$$\tau_L(t) \triangleq \max_j\{\tau_j : x_i(t) \geq \tau_j, \forall i\} \tag{30}$$

and

$$\tau_U(t) \triangleq \min_j\{\tau_j : x_i(t) \leq \tau_j, \forall i\} \tag{31}$$

along with $\mathrm{I}[\tau(t)] \triangleq [\tau_L(t), \tau_U(t)]$.

The following theorem discusses the evolution of the interval of the quantized node state values, and the minimum range quantization bin that encloses the node state values. The theorem reveals that both intervals are non–expanding.

Theorem 4. *For some $t \geq 0$, suppose that $q_i(t) \in \mathrm{I}[\mathbf{q}(t)]$, and, $x_i(t) \in \mathrm{I}[\mathbf{x}(t)]$, for $i = 1, 2, \ldots, N$. By construction*

$$\mathrm{I}[\mathbf{x}(t)] \subseteq \mathrm{I}[\tau(t)]. \tag{32}$$

Then, for $k \geq 1$, the followings hold:

(i) *The interval of the quantized state vector is non–expanding,* i.e.,

$$\mathrm{I}[\mathbf{q}(t+k)] \subseteq \mathrm{I}[\mathbf{q}(t)]. \tag{33}$$

(ii) *The minimum length interval with boundaries defined by quantization points that encloses the state vector values is non–expanding,* i.e.,

$$\mathrm{I}[\mathbf{x}(t+k)] \subseteq \mathrm{I}[\tau(t+k)] \subseteq \mathrm{I}[\tau(t)]. \tag{34}$$

Proof: Consider (i) first. Suppose that $q_i(t) \in \mathrm{I}[\mathbf{q}(t)]$, for $i = 1, 2, \ldots, N$, and recall that the state recursion follows as $\mathbf{x}(t+1) = \mathbf{W}\mathbf{q}(t)$. Let $\mathbf{w}^i$ denote the row vector formed as the i–th row of the weight matrix $\mathbf{W}$. Now, we can write the node specific update equation as

$$x_i(t+1) = \mathbf{w}^i \mathbf{q}(t). \tag{35}$$

Note that $x_i(t+1)$ is a linear combination of quantized local node values and $w_j^i \geq 0$ for $j = 1, 2, \ldots, N$, where w_j^i denotes the j–th entry of $\mathbf{w}^i$. Moreover, $\mathbf{w}^i \mathbf{1} = 1$, since $\mathbf{W}\mathbf{1} = \mathbf{1}$. Thus, $x_i(t+1)$ is a convex combination of the quantized node state values and its own quantized state. The node state value $x_i(t+1)$ is then in the convex hull of quantized state values $\{q_i(t) : i = 1, 2, \ldots, N\}$. The convex hull of the quantized state values at time t is given by $\mathrm{I}[\mathbf{q}(t)]$, indicating that

$$x_i(t+1) \in \mathrm{I}[\mathbf{q}(t)] \tag{36}$$

for $i = 1, 2, \ldots, N$, and subsequently,

$$\mathrm{I}[\mathbf{x}(t+1)] \subseteq \mathrm{I}[\mathbf{q}(t)]. \tag{37}$$

Hence, we see that

$$q_{(N)}(t+1) = \mathcal{Q}(x_i(t+1)) \leq q_{(N)}(t) \tag{38}$$

for some $i \in \{1, 2, \ldots, N\}$ and

$$q_{(1)}(t+1) = \mathcal{Q}(x_j(t+1)) \geq q_{(1)}(t) \tag{39}$$

for some $j \in \{1, 2, \ldots, N\}$ and $j \neq i$. It follows that

$$\mathrm{I}[\mathbf{q}(t+1)] \subseteq \mathrm{I}[\mathbf{q}(t)]. \tag{40}$$

Repeatedly utilizing the above steps completes the proof.

Now consider (ii). Suppose that $x_i(t) \in \mathrm{I}[\mathbf{x}(t)]$ for $i = 1, 2, \ldots, N$. Then, by construction, $\mathrm{I}[\mathbf{x}(t)] \subseteq \mathrm{I}[\tau(t)]$.

Furthermore, since

$$\tau_L(t) \leq q_{(1)}(t) \tag{41}$$

and

$$\tau_U(t) \geq q_{(N)}(t) \tag{42}$$

it follows that

$$q_i(t) = \mathcal{Q}(x_i(t)) \in \mathrm{I}[\mathbf{q}(t)] \subseteq \mathrm{I}[\tau(t)] \tag{43}$$

for $i = 1, 2, \ldots, N$. The convex combination property, similar to the previous case, indicates that, $x_i(t+1) \in \mathrm{I}[\mathbf{q}(t)]$ for $i = 1, 2, \ldots, N$, and subsequently, $\mathrm{I}[\mathbf{x}(t+1)] \subseteq \mathrm{I}[\mathbf{q}(t)]$. Moreover, since, $\mathrm{I}[\mathbf{q}(t)] \subseteq \mathrm{I}[\tau(t)]$, it follows that

$$\mathrm{I}[\mathbf{x}(t+1)] \subseteq \mathrm{I}[\tau(t)]. \tag{44}$$

Finally combining all the results indicating that

$$\mathrm{I}[\mathbf{x}(t+1)] \subseteq \mathrm{I}[\mathbf{q}(t+1)] \subseteq \mathrm{I}[\tau(t+1)] \subseteq \mathrm{I}[\tau(t)] \tag{45}$$

and repeatedly utilizing the above steps completes the proof. ■

The proof of this theorem indicates that each iteration is indeed a convex combination of the previous set of quantized node state values, and uses of the properties of convex functions to arrive at the stated results.

Let us define $r_{\mathbf{q}}(t)$ as the Lebesgue measure of the domain of the quantized state vector at time step t, *i.e.*, the range of $\mathbf{q}(t) \in \mathbb{R}^N$,

$$r_{\mathbf{q}}(t) \triangleq q_{(N)}(t) - q_{(1)}(t) \tag{46}$$

where $r_{\mathbf{q}}(t) \in \{0, \Delta, \ldots, (L-1)\Delta\}$. Similar to the quantized state vector case, we define $r_\tau(t) \triangleq \tau_U(t) - \tau_L(t)$ as the length of the interval $\mathrm{I}[\tau(t)]$.

The following corollary (the proof of which is omitted since it follows directly from Theorem 1 and Theorem 4), compiled from Theorem 1 and Theorem 4, discusses properties of interest of $r_{\mathbf{q}}(t)$ and $r_\tau(t)$.

Corollary 1. *The functions $r_{\mathbf{q}}(t)$ and $r_\tau(t)$, with initial conditions $r_{\mathbf{q}}(0) \geq 0$ and $r_\tau(0) \geq 0$, tend to zero as t tends to infinity,* i.e.,

$$\lim_{t\to\infty} r_{\mathbf{q}}(t) = 0 \tag{47}$$

and

$$\lim_{t\to\infty} r_\tau(t) = 0. \tag{48}$$

Moreover, $r_{\mathbf{q}}(t)$ *and* $r_{\tau}(t)$ *are monotonically non–increasing functions.*

The presented theorem and corollary indicate that the convergence of the PQDA is monotonic in the sense that the global trend of both the quantized and unquantized node state values is towards the consensus and that the minimum-length intervals containing all values do not expand, and in fact, converge to zero-length monotonically. The following theorem investigates the rate of convergence of the PQDA to a state where there is a first time non–zero probability of converging to the consensus (all values are contained within two quantization bins).

Theorem 5. *Let* $r_{\mathbf{q}}(t) \triangleq q_{(N)}(t) - q_{(1)}(t)$ *and* $r_{\mathbf{x}}(t) \triangleq x_{(N)}(t) - x_{(1)}(t)$ *denote the range of the quantized and unquantized node state values at time step* t*, with the initial values* $r_{\mathbf{q}}(0)$ *and* $r_{\mathbf{x}}(0)$*, respectively. Then,*

$$\mathbb{E}\{r_{\mathbf{q}}(t)\} \leq \sqrt{\frac{N-1}{N}}\rho^t(\mathbf{W}-\mathbf{J})r_{\mathbf{x}}(0) + 2\Delta \tag{49}$$

where $\rho(\cdot)$ *denotes the spectral radius of its argument.*

Proof: See Appendix B. ■

In the appendix, we compile an upper and lower bound on the largest and smallest order statistics of the quantized node state vector using results from [25]. Then, the task reduces to deriving a bound on the convergence rate of the normed difference of any row i and j with time, and combining this bound with the bounds on the order statistics gives the desired result.

Theorem 5 reveals that the PQDA converges to the final two bins with the same rate as standard consensus. Theorem 5 also relates the convergence of the quantized node state values range to the range of initial node measurements.

After all the node state values are in the final two bins, there is always a non–zero probability to immediately converge to consensus. Note that, in the absence of knowledge of the norm of the initial node states or the initial state range, the bound given above reduces to

$$\mathbb{E}\{r_{\mathbf{q}}(t)\} \leq 2\sqrt{\frac{N-1}{N}}\rho^t(\mathbf{W}-\mathbf{J})U + 2\Delta \tag{50}$$

where we used the facts that $\max_i\{x_i(0)\} \leq U$ and $\min_i\{x_i(0)\} \geq -U$.

To understand the convergence of the PQDA algorithm after all the quantized states converged to the final two bins, first, let us discuss the behavior of the PQDA algorithm in the final bin, *i.e.*, $\max_i\{q_i(t)\} - \min_i\{q_i(t)\} = \Delta$. Suppose $x_i(t) \in [\tau_j, \tau_{j+1}]$, for some j. In this case, all the nodes state values need to be quantized to τ_j or τ_{j+1} to achieve a consensus at time step t. Hence, the effect of the weight matrix on the convergence rate significantly decreases and the convergence rate is mainly dominated by the probabilistic quantization. Moreover, we hence believe that the time interval, where all the node state values are in τ_j and $\tau_j + 2\Delta$, is a transition period between

the dominating effect of the weight matrix, *i.e.*, the spectral radius of $\mathbf{W} - \mathbf{J}$, and the dominating effect of probabilistic quantization. Obtaining analytical expressions of convergence rate for these transition and final bin regions appears to be a challenging task. Although our current research efforts focus on this challenge, we assess the convergence performance of the PQDA algorithm with extensive simulations in the following section.

V. Numerical Examples

This section details numerical examples evaluating the performance of the distributed average computation using probabilistic quantization. Throughout the simulations we utilized the Metropolis weight matrix defined for a graph $G = (\mathcal{V}, \mathcal{E})$ [3]. The Metropolis weights on a graph are defined as follows:

$$W_{ij} = \begin{cases} (1 + \max\{\deg_i, \deg_j\})^{-1}, & i \neq j, \text{ and } (i,j) \in \mathcal{E} \\ 1 - \sum_{k \in \mathcal{N}_i} W_{ik}, & i = j \\ 0, & \text{otherwise} \end{cases} . \tag{51}$$

This method of choosing weights is adapted from the Metropolis algorithm in the literature of Markov chain Monte Carlo [3], [19]. The Metropolis weights are very simple to compute and are well suited for distributed implementation. In particular, each node only needs to know the degrees of its neighbors to determine the weights on its adjacent edges. Furthermore, the nodes do not need any global knowledge of the communication graph or even the total number of nodes.

We simulate a network with $N = 50$ nodes randomly dispersed on the unit square $[0,1] \times [0,1]$, connecting two nodes by an edge if the distance between them is less than the connectivity radius, *i.e.*, $d = \sqrt{\log(N)/N}$. Thus, a link exists between any two nodes that are at a range less than d. Throughout this section, the initial states are drawn i.i.d. from a uniform distribution as following: $\mathbf{y} = \overline{\mathbf{x}}(0) + \mathbf{n}$, where $\mathbf{n}$ is i.i.d. uniformly distributed with support in the $[-1, 1]$ interval. The initial states are then regularized such that $\overline{\mathbf{y}} = \overline{\mathbf{x}}(0) = 0.85$. The quantization resolution is taken as $\Delta = 0.1$. Plotted in Fig. 2 is $q_{(N)}$ and $q_{(1)}$ at every time step t (corresponding to node trajectories given in Fig. 1). The figure indicates that the proposed algorithm does indeed achieve consensus as the interval in which the quantized state vector converges to zero and is monotonically non–expanding, corroborating the theoretical results. In this case, the consensus is $\lim_{t\to\infty} \mathbf{x}(t) = (0.8)\mathbf{1}$, which is in agreement with the theoretical results indicating that the consensus is at one of the quantization levels.

We next investigate the effect of the quantization resolution and the location of the initial state average on the consensus standard deviation. Figure 3 plots the error norm of the consensus for varying $\Delta \in [0.05, 0.5]$ when $\overline{\mathbf{x}}(0) = 0.85$ and for varying $\overline{\mathbf{x}}(0) \in [0.5, 1.5]$ when $\Delta = 0.25$. Also plotted is the derived upper-bound on the PQDA. Note that each data point in the plots is an ensemble average of 1000 trials. The variance, as expected, tends to increase as Δ increases and exhibits a harmonic behavior as the location of the average changes. This is

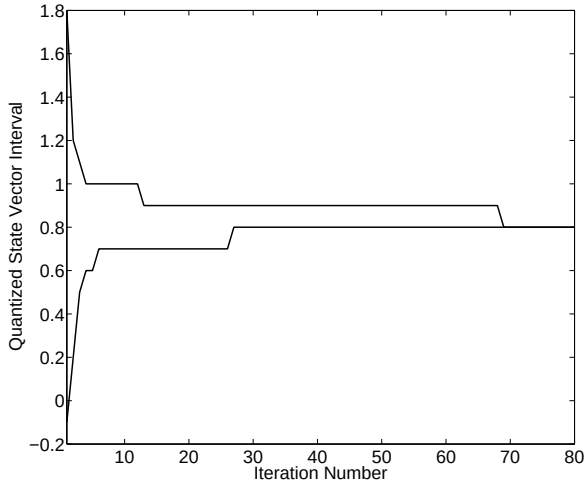
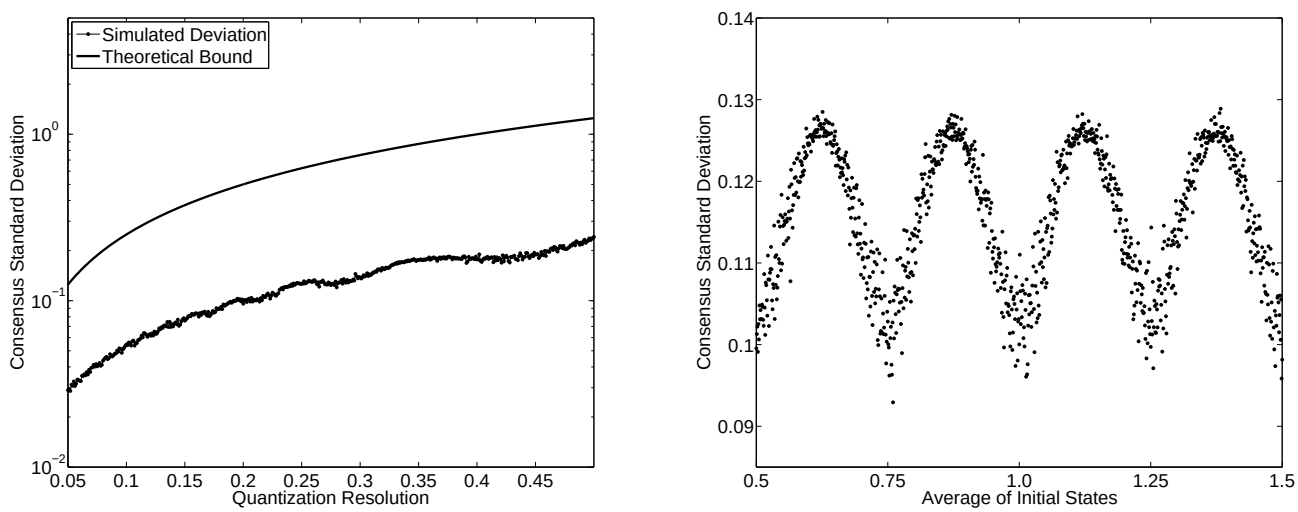


Fig. 2. The plotted are is the interval in which the quantized state vector is. The number of nodes is $N = 50$, the nodes' initial average is $\overline{\mathbf{x}}(0) = 0.85$, and the quantization resolution is set to $\Delta = 0.1$.



(a) (b)

Fig. 3. The error norm of the consensus with respect to (a) the quantization resolution, *i.e.*, $\Delta \in [0.05, 0.5]$ with $\overline{\mathbf{x}}(0) = 0.85$ and (b) the initial state average with $\Delta = 0.25$. The network parameter are: $N = 50$ and $d = \sqrt{\log(N)/N}$.

due to the effect induced by the distance of the average to the quantization levels.

Figure 4 shows the behavior of the average mean square error (MSE) per iteration defined as:

$$\mathrm{MSE}(t) = \frac{1}{N} \sum_{i=1}^{N} \left(x_i(t) - \overline{\mathbf{x}}(0)\right)^2 \tag{52}$$

for $\Delta \in \{0.05, 0.1, 0.15, 0.2\}$. In other words, $\mathrm{MSE}(t)$ is the average mean squared distance of the states at iteration t from the initial mean. Each curve is an ensemble average of 1000 experiments and the network parameters are: $N = 50$, $\overline{\mathbf{x}}(0) = 0.85$ and $d = \sqrt{\log(N)/N}$. The plots suggest that smaller quantization bins yield a smaller

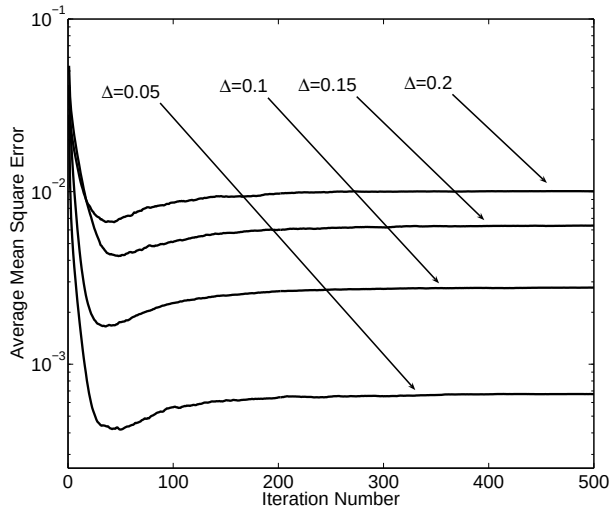


Fig. 4. The average MSE of the probabilistically quantized distributed average consensus for varying quantization resolution where $\Delta \in \{0.05, 0.1, 0.15, 0.2\}$. The remaining network parameters are: $N = 50$, $\overline{\mathbf{x}}(0) = 0.85$ and $d = \sqrt{\log(N)/N}$.

steady state MSE and that as quantization bin size increases, the number of iterations taken by PQDA to reach the final quantization bin decreases. The quasi-convex shape of the MSE curves are due to the fact that the algorithm, after all the state values converge into a quantization range $[\tau_i, \tau_{i+1})$ for some $i \in \{1, 2, \ldots, L-1\}$, drifts to a quantization value.

Considered next is the consensus value of the PQDA algorithm. Figure 5 plots the histograms of the consensus value for varying initial state average, *i.e.*, $\overline{\mathbf{x}}(0) \in \{0.80, 0.825, \ldots, 1.00\}$ for $\Delta = 0.2$. The number of nodes in the network is $N = 50$. Note that the consensus values shift as the initial average value shifts from 0.80 to 1.00. This is directly related to the fact that the consensus, in expectation, is equal to the average of initial states as provided by the theoretical results.

We investigate the average convergence time of the distributed average consensus using probabilistic quantization for varying $\Delta \in \{0.05, 0.1, 0.2\}$ against the number of nodes in the network, Fig. 6(a) and (b). We also show the average number of iterations taken to achieve the final quantization bin. Moreover, Fig. 6(c) and (d) plot the average normalized distance to the closest absorbing state at the first time step when all the quantized node state values are in the final quantization bin. The initial state averages are $\overline{\mathbf{x}}(0) = 0.85$ and $\overline{\mathbf{x}}(0) = 0.90$, and the connectivity radius is $d = \sqrt{4\log(N)/N}$. Each data point is an ensemble average of 10000 trials. Note that the convergence time increases with the number of nodes in the network. The plots suggest that the number of iterations taken by the PQDA algorithm to converge to final quantization bin decreases as Δ increases. This can be seen by noting that the algorithm has to go through less "averaging" (multiplication with the weight matrix) before arriving at the final bin. It is hence clear that the algorithm needs to run for a smaller number of iterations to arrive at a larger final bin size.

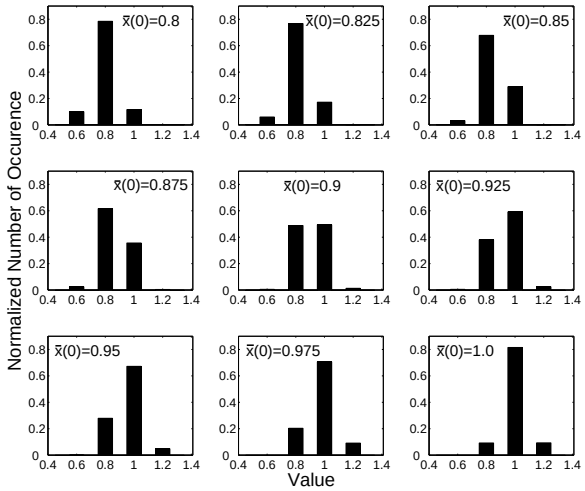


Fig. 5. Histograms of the consensus value achieved by the probabilistically quantized consensus for varying initial state average where $\overline{\mathbf{x}}(0) \in \{0.80, 0.825, \ldots, 1.00\}$ and $\Delta = 0.2$. The number of nodes in the network is $N = 50$.

On the other hand, as discussed in more detail below, the expected number of iterations taken to achieve consensus is dominated by the number of iterations taken to converge to an absorbing state after all the node values are in the final bin. The probabilistic quantization is the dominant effect in the final bin. The time taken to converge to an absorbing state is heavily dependent on the distance to that absorbing state at the first time step when all values enter the final bin. This distance is affected by two factors. First, if more averaging operations occur prior to the entry step, then there is more uniformity in the values, decreasing the distance. Second, if the initial data average is close to a quantization value, then, on average, the entry point will be closer to an absorbing state (note that $\mathbb{E}\{\mathbf{1}^{\mathrm{T}}\mathbf{q}(t)\} = \mathbf{1}^{\mathrm{T}}\mathbf{x}(0)$). These observations explain the results of Fig. 6. Note that the convergence time order for $\overline{\mathbf{x}}(0) = 0.85$ and $\overline{\mathbf{x}}(0) = 0.90$ cases flip for $\Delta = 0.2$ and $\Delta = 0.1$. That is due to the fact that the average distance to an absorbing when, at the first time step, all the node values enter the final bin is smaller for $\overline{\mathbf{x}}(0) = 0.85$ when $\Delta = 0.2$ compared to $\Delta = 0.1$, and is smaller for $\overline{\mathbf{x}}(0) = 0.90$ when $\Delta = 0.1$ compared to $\Delta = 0.2$. Moreover, note that $\Delta = 0.05$ yields the smallest distance to an absorbing state for both initial conditions. Although, it takes more iterations to converge to final bin, in both cases, PQDA algorithm with $\Delta = 0.05$ yields the smallest average distance to an absorbing state when all the node values enter to the final bin for the first time step, hence, the smallest average number of iterations to achieve the consensus.

We consider next the effect of the connectivity radius on the average number of iterations taken to achieve the consensus. Figure 7 depicts the average number of iterations to achieve the consensus for the cases where the initial state average is $\overline{\mathbf{x}}(0) = 0.85$ and $\overline{\mathbf{x}}(0) = 0.90$. As expected, the average number of iterations taken to achieve consensus decreases with increasing connectivity radius. This is related to the fact that higher connectivity radius, implies a lower second largest eigenvalue for the weight matrix. Moreover, as in the previous case, the convergence

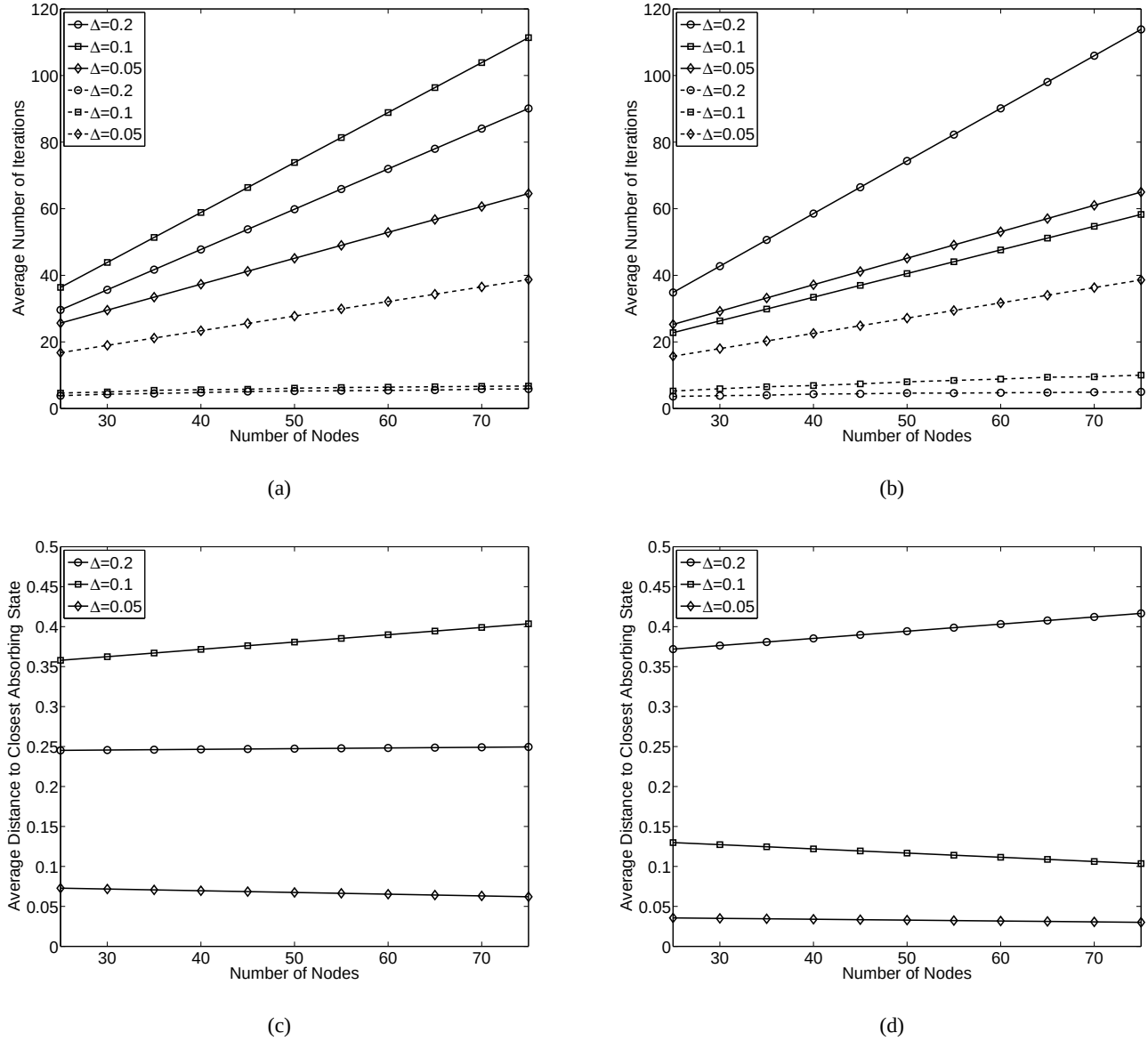


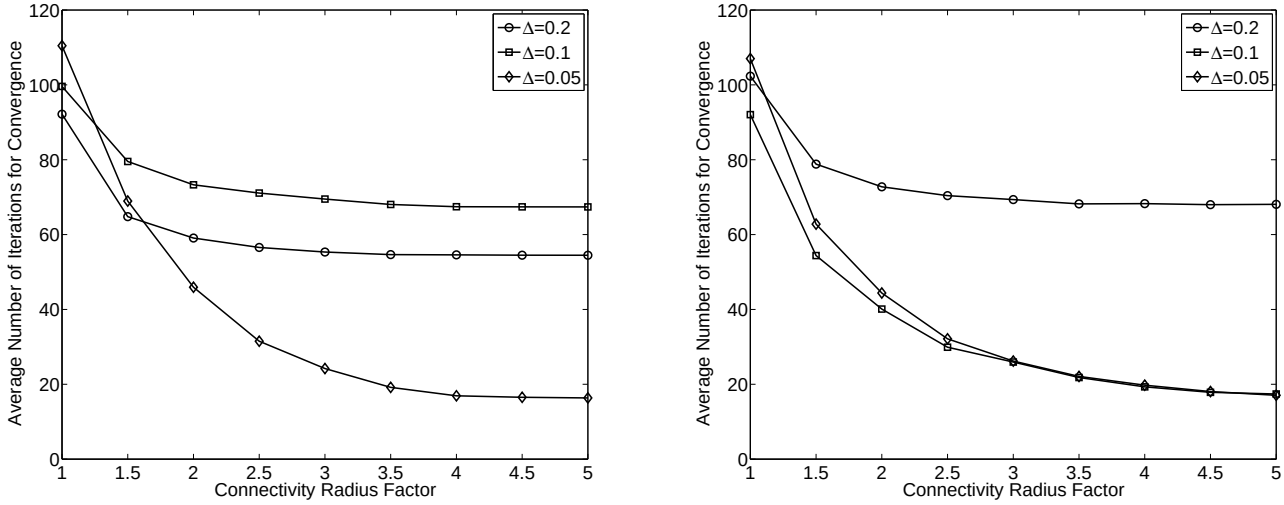
Fig. 6. Average number of iterations taken by the probabilistically quantized distribute average computation to achieve final quantization bin (dashed) and consensus (solid) for $\Delta \in \{0.05, 0.1, 0.2\}$ and for varying N with (a:) $\overline{\mathbf{x}}(0) = 0.85$ and (b:) $\overline{\mathbf{x}}(0) = 0.90$, along with the corresponding average distance to the closest absorbing state at the first time step when all the quantized node state value are in the final quantization bin for (c:) $\overline{\mathbf{x}}(0) = 0.85$ and (d:) $\overline{\mathbf{x}}(0) = 0.90$.

time is related to the distance of the initial state average to a quantization value for a given quantization resolution. Of note is that 0.85 is a quantization point for $\Delta = 0.05$, and 0.90 is a quantization point for both $\Delta = 0.1$ and $\Delta = 0.05$. The combined results of the presented experiments indicate that the expected number of iterations required to reach a consensus depends on the following factors:

(1) Quantization resolution;

(2) Initial node measurements;

(3) Number of nodes in the network; and



(a) (b)

Fig. 7. Average number of iterations taken by the probabilistically quantized distribute average computation for $\Delta \in \{0.05, 0.1, 0.2\}$ with (a:) $\overline{\mathbf{x}}(0) = 0.85$ and (b:) $\overline{\mathbf{x}}(0) = 0.90$. The number of nodes in the network is $N = 50$. Connectivity radius factor is defined as the modulating constant k, in the expression for the connectivity radius $d = k\sqrt{\log(N)/N}$.

(4) Connectivity radius.

Note that (1), (3) and (4) are system design choices, affected by energy budgets and bandwidth constraints, but (2) is data-dependent. This implies that the quantization resolution, given the bandwidth and power constraints of the application, should be chosen to minimize the expected (or worst-case) convergence time over the range of possible initial averages.

VI. Further Considerations

The analysis presented in this paper makes two main simplifying assumptions: 1) the network topology does not change over time, and 2) communication between neighboring nodes is always successful. The simplifying assumptions essentially allow us to focus on the case where the weight matrix, $\mathbf{W}$, does not change with time. However, time-varying topologies and unreliable communications are important practical issues which have been addressed for un-quantized consensus algorithms (see, *e.g.*, [3], [26], [27]). Since $\mathbf{W}$ has the same support as the adjacency matrix of the underlying communication graph, when the topology changes with time, the averaging weights must also vary. Likewise, an unsuccessful transmission between two nodes is equivalent to the link between those nodes vanishing for one iteration. In either case, we can now think of $\mathbf{W}(t)$ as random process. Typical results for this scenario roughly state that average consensus is still accomplished when the weight matrix varies with time, so long as the expected weight matrix, $\mathbb{E}[\mathbf{W}(t)]$, is connected. This condition ensures that there is always non-zero probability that information will diffuse throughout the network. We expect that the same techniques employed in [3], [26], [27] can be used to show convergence of our average consensus with probabilistic quantization with

time-varying $\mathbf{W}$.

In this paper we also restricted ourselves to the scenario where the quantization step size Δ remains fixed over all time. Recall that when the algorithm has converged to a consensus, all $q_i(t)$ are at the same quantization point, so $\|\mathbf{q}(t) - \mathbf{J}\mathbf{q}(t)\| = 0$. Letting $D(t) \triangleq \|\mathbf{q}(t) - \mathbf{J}\mathbf{q}(t)\|$ denote Euclidean distance to convergence, we know that when the algorithm is far from converging (*i.e.*, $D(t)$ large), quantization errors have less of an effect on convergence of the algorithm. This is because the averaging effects of $\mathbf{W}$ are multiplicative and thus have a stronger influence when $D(t)$ is large, whereas the quantization error is bounded by a constant which only depends on Δ and not on $D(t)$. When $D(t)$ is of the same order as the quantization noise variance, quantization essentially wipes away the effects of averaging and hampers the time to convergence. A natural extension of the algorithm proposed in this paper involves shrinking the quantization step size, Δ, over time, *e.g.*, setting $\Delta^{\text{new}} = \Delta^{\text{old}}/2$ once $D(t)$ is established to be below the threshold where quantization effects outweigh averaging. We expect that this modification should improve the rate at which $D(t)$ tends to zero without affecting statistical properties of the limiting consensus values (*i.e.*, unbiased w.r.t. to $\overline{\mathbf{x}}(0)$, and no increase in the limiting variance). Solidifying this procedure is a topic of current investigation.

VII. Concluding Remarks

We have described probabilistically quantized distributed averaging (PQDA), a framework for distributed computation of averages of the node data over networks with bandwidth/power constraints or large volumes of data. The proposed method unites the distributed average consensus algorithm and probabilistic quantization, which is a form of "dithered quantization". The proposed PQDA algorithm achieves a consensus, and the consensus is a discrete random variable whose support is the quantization values and expectation is equal to the average of the initial states. We have derived an upper bound on the mean square error performance of the PQDA algorithm. Our analysis demonstrates that the minimum-length intervals (with boundaries constrained to quantization points) containing the quantized and unquantized state values are non-expanding. Moreover, the lengths of these intervals are non-increasing functions with limit zero, indicating that convergence is monotonic. In addition, we have shown that, all the node state values, in the worst case, arrive in the final two quantization bins at the same rate as standard, unquantized consensus algorithms. Finally, we have provided numerical examples illustrating the effectiveness of the proposed algorithm and highlighting the factors that impact the convergence rate.

APPENDIX A

PROOF OF THEOREM 3–ERROR NORM OF PROBABILISTICALLY QUANTIZED DISTRIBUTED AVERAGE CONSENSUS

Consider the following set of equalities

$$||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)|| = ||\mathbf{x}(t) + \mathbf{v}(t) - \mathbf{J}\mathbf{x}(0)|| \tag{53}$$

$$= ||\mathbf{W}^t\mathbf{x}(0) + \sum_{j=0}^{t-1}\mathbf{W}^{t-j}\mathbf{v}(j) - \mathbf{J}\mathbf{x}(0) + \mathbf{v}(t)|| \tag{54}$$

$$= ||(\mathbf{W} - \mathbf{J})^t\mathbf{x}(0) + \sum_{j=0}^{t-1}\mathbf{W}^{t-j}\mathbf{v}(j) + \mathbf{v}(t)|| \tag{55}$$

where we use the facts that $(\mathbf{W}^t - \mathbf{J}) = (\mathbf{W} - \mathbf{J})^t$, for $t \geq 1$, and $\mathbf{J}^t = \mathbf{J}$, for $t \geq 1$. Now the eigendecomposition of $\mathbf{W}$ yields

$$\mathbf{W} = \sum_{k=1}^{N}\lambda_k\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}} \tag{56}$$

where $\mathbf{u}_k$ denotes the eigenvector associated with the eigenvalues λ_k. Eigendecomposition further indicates that

$$\mathbf{W}^{t-j} = \sum_{k=1}^{N}\lambda_k^{t-j}\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}}. \tag{57}$$

Since $\mathbf{W}\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^{\mathrm{T}}\mathbf{W} = \mathbf{1}^{\mathrm{T}}$, the eigenvector associated with the eigenvalue $\lambda_1 = 1$ is given by $\mathbf{u}_1 = (\sqrt{N})^{-1}\mathbf{1}$. Substituting this information into the error norm equation gives

$$||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)|| = ||(\mathbf{W} - \mathbf{J})^t\mathbf{x}(0) + \sum_{j=0}^{t-1}\sum_{k=1}^{N}\lambda_k^{t-j}\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}}\mathbf{v}(j) + \mathbf{v}(t)|| \tag{58}$$

$$= ||(\mathbf{W} - \mathbf{J})^t\mathbf{x}(0) + \sum_{j=0}^{t-1}\left(\frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}} + \sum_{k=2}^{N}\lambda_k^{t-j}\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}}\right)\mathbf{v}(j) + \mathbf{v}(t)|| \tag{59}$$

$$= ||(\mathbf{W} - \mathbf{J})^t\mathbf{x}(0) + \sum_{j=0}^{t-1}\overline{\mathbf{v}}(j)\mathbf{1} + \sum_{j=0}^{t-1}\sum_{k=2}^{N}\lambda_k^{t-j}\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}}\mathbf{v}(j) + \mathbf{v}(t)||. \tag{60}$$

Moreover, applying the Triangle inequality and using the facts that $||\overline{\mathbf{v}}(j)\mathbf{1}|| = |\overline{\mathbf{v}}(j)|||\mathbf{1}|| = |\overline{\mathbf{v}}(j)|\sqrt{N}$ and that

$$\sum_{k=2}^{N}\lambda_k^{t-j}\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}} = (\mathbf{W} - \mathbf{J})^{t-j} \tag{61}$$

after multiplying both sides with $(\sqrt{N})^{-1}$ gives

$$(\sqrt{N})^{-1}||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)|| \leq (\sqrt{N})^{-1}||(\mathbf{W} - \mathbf{J})^t\mathbf{x}(0)|| + \sum_{j=0}^{t-1}|\overline{\mathbf{v}}(j)| + \sum_{j=0}^{t-1}(\sqrt{N})^{-1}||(\mathbf{W} - \mathbf{J})^{t-j}\mathbf{v}(j)|| + (\sqrt{N})^{-1}||\mathbf{v}(t)||. \tag{62}$$

We need to following lemma to continue with the proof.

Lemma 3. *The sum of quantization noise terms, at all time steps, converges in probability,* i.e.,

$$\lim_{N\to\infty} \Pr\{|\overline{\mathbf{v}}(j)| \geq \epsilon\} = 0 \tag{63}$$

for $\epsilon > 0$ *and all* j*. Thus,* $\lim_{N\to\infty} \mathbb{E}\{|\overline{\mathbf{v}}(j)|\} = 0$*.*

Proof: Recall that $\mathbb{E}\{\overline{\mathbf{v}}(j)\} = 0$ and

$$\begin{aligned}
\mathrm{Var}\{\overline{\mathbf{v}}(j)\} &= \mathrm{Var}\left\{\frac{1}{N}\sum_{i=1}^{N} v_i(j)\right\} && (64)\\
&= \frac{1}{N^2}\sum_{i=1}^{N} \mathrm{Var}\{v_i(j)\} && (65)\\
&\leq \frac{1}{N^2}\sum_{i=1}^{N} \frac{\Delta^2}{4} && (66)\\
&= \frac{1}{N}\frac{\Delta^2}{4}. && (67)
\end{aligned}$$

Now using Chebyshev's Inequality, we obtain

$$\begin{aligned}
\Pr\{|\overline{\mathbf{v}}(j)| \geq \epsilon\} &\leq \frac{\mathrm{Var}\{|\overline{\mathbf{v}}(j)|\}}{\epsilon^2} && (68)\\
&\leq \frac{1}{N\epsilon^2}\frac{\Delta^2}{4}. && (69)
\end{aligned}$$

Now, take the limit of $N \to \infty$, we see that the RHS of the above goes to zero for all $\epsilon > 0$. Thus, the probability $|\overline{\mathbf{v}}(j)|$ being greater than zero is equal to zero for $N \to \infty$ and the implication of $\lim_{N\to\infty} \mathbb{E}\{|\overline{\mathbf{v}}(j)|\} = 0$. ■

The error norm equation, after taking the expectation and limit as $N \to \infty$, since the limit of each $\mathbb{E}\{|\overline{\mathbf{v}}(j)|\}$ exists and equals zero (from Lemma 3), reduces to

$$\begin{aligned}
&\lim_{N\to\infty} (\sqrt{N})^{-1}\mathbb{E}\{||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)||\}\\
&\qquad \leq \lim_{N\to\infty} (\sqrt{N})^{-1}\mathbb{E}\{||(\mathbf{W}-\mathbf{J})^t\mathbf{x}(0)||\} + \sum_{j=0}^{t-1}(\sqrt{N})^{-1}\mathbb{E}\{||(\mathbf{W}-\mathbf{J})^{t-j}\mathbf{v}(j)||\} + (\sqrt{N})^{-1}\mathbb{E}\{||\mathbf{v}(t)||\}.
\end{aligned} \tag{70}$$

Furthermore, utilizing the Norm inequality gives

$$\lim_{N\to\infty} (\sqrt{N})^{-1}\mathbb{E}\{||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)||\} \leq \lim_{N\to\infty} \rho^t(\mathbf{W}-\mathbf{J})(\sqrt{N})^{-1}||\mathbf{x}(0)|| + \sum_{j=0}^{t} \rho^{t-j}(\mathbf{W}-\mathbf{J})(\sqrt{N})^{-1}\mathbb{E}\{||\mathbf{v}(j)||\}. \tag{71}$$

In the following, we derive an upper-bound of $\mathbb{E}\{||\mathbf{v}(j)||\}$ for $j = 0, 1, \ldots, t$ to bound $\mathbb{E}\{||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)||\}$.

Consider the expectation of the quantization noise,

$$\mathbb{E}\{||\mathbf{v}(j)||\} = \mathbb{E}\left\{\sqrt{\sum_{i=1}^{N} v_i^2(j)}\right\}. \tag{72}$$

Note that $f(u) = \sqrt{u}$ is a concave function. The concavity indicates that utilizing Jensen's inequality gives

$$\mathbb{E}\left\{\sqrt{\sum_{i=1}^{N} v_i^2(j)}\right\} \leq \sqrt{\sum_{i=1}^{N} \mathbb{E}\{v_i^2(j)\}}. \tag{73}$$

Now using the upper-bound for the expectation of the quantization noise variance term, *i.e.*, Lemma 1, indicates that the expectation of the quantization noise norm is bounded by

$$\mathbb{E}\{||\mathbf{v}(j)||\} \leq \sqrt{\sum_{i=1}^{N} \frac{\Delta^2}{4}} = \sqrt{N}\frac{\Delta}{2}. \tag{74}$$

Now, substituting this result into the error norm equation, after some manipulations, gives

$$\lim_{N\to\infty} (\sqrt{N})^{-1}\mathbb{E}\{||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)||\} \leq \lim_{N\to\infty} \rho^t(\mathbf{W}-\mathbf{J})(\sqrt{N})^{-1}||\mathbf{x}(0))|| + \frac{\Delta}{2}\sum_{j=0}^{t} \rho^{t-j}(\mathbf{W}-\mathbf{J}). \tag{75}$$

Recall that $\rho(\mathbf{W}-\mathbf{J}) < 1$, hence, applying the Geometric Series equality, *i.e.*,

$$\sum_{j=0}^{t} \rho^j(\mathbf{W}-\mathbf{J}) = \frac{1-\rho^{t+1}(\mathbf{W}-\mathbf{J})}{1-\rho(\mathbf{W}-\mathbf{J})} \tag{76}$$

further yields

$$\lim_{N\to\infty} (\sqrt{N})^{-1}\mathbb{E}\{||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)||\} \leq \lim_{N\to\infty} \rho^t(\mathbf{W}-\mathbf{J})(\sqrt{N})^{-1}||\mathbf{x}(0))|| + \frac{\Delta}{2}\frac{1-\rho^{t+1}(\mathbf{W}-\mathbf{J})}{1-\rho(\mathbf{W}-\mathbf{J})}. \tag{77}$$

Now, taking the limit as t tends to infinity yields

$$\lim_{t\to\infty}\lim_{N\to\infty} (\sqrt{N})^{-1}\mathbb{E}\{||\mathbf{q}(t) - \mathbf{J}\mathbf{x}(0)||\} \leq \lim_{t\to\infty}\lim_{N\to\infty} \rho^t(\mathbf{W}-\mathbf{J})(\sqrt{N})^{-1}||\mathbf{x}(0)|| + \frac{\Delta}{2}\frac{1-\rho^{t+1}(\mathbf{W}-\mathbf{J})}{1-\rho(\mathbf{W}-\mathbf{J})}. \tag{78}$$

Note that the limit of each term exists. Also consider the following:

$$\lim_{t\to\infty} \rho^t(\mathbf{W}-\mathbf{J}) = 0 \tag{79}$$

since $\rho(\mathbf{W}-\mathbf{J}) < 1$, and, subsequently,

$$\lim_{t\to\infty} \frac{1-\rho^{t+1}(\mathbf{W}-\mathbf{J})}{1-\rho(\mathbf{W}-\mathbf{J})} = \frac{1}{1-\rho(\mathbf{W}-\mathbf{J})}. \tag{80}$$

Combining these findings and substituting them into (78) yields the desired result.

APPENDIX B

PROOF OF THEOREM 5–CONVERGENCE RATE TO THE FINAL TWO BINS

Note that $r_{\mathbf{q}}(t) = \max_i\{q_i(t)\} - \min_i\{q_i(t)\}$. In order to bound the expected range, we will upper and lower bound the largest and smallest order statistics of the quantized node state values at time step t. To prove the proposed theorem, we make use of the following bounds for the maximum and minimum order statistics of (possibly dependent) $\{q_i\}_{i=1}^N$ samples [25]:

$$\mathbb{E}\{\max_i\{q_i\}\} \leq \max_i\{\mathbb{E}\{q_i\}\} + \mathbb{E}\{\max_i\{q_i - \mathbb{E}\{q_i\}\}\} \tag{81}$$

and

$$\mathbb{E}\{\min_i\{q_i\}\} \geq \min_i\{\mathbb{E}\{q_i\}\} + \mathbb{E}\{\min_i\{q_i - \mathbb{E}\{q_i\}\}\} \tag{82}$$

respectively. Using these bounds, in our setup, for the largest order statistics gives:

$$\mathbb{E}\{\max_i\{q_i(t)\}\} \leq \max_i \mathbf{w}_i^t\mathbf{x}(0) + \mathbb{E}\{\max_i\{v_i(t)\}\} \tag{83}$$

$$\leq \mathbf{w}_{i^*}^t\mathbf{x}(0) + \Delta \tag{84}$$

where we define $\mathbf{w}_i^t$ to be i–th row of the weight matrix taken to the power t and $i^* = i : \mathbf{w}_{i^*}^t\mathbf{x}(0) \geq \mathbf{w}_i^t\mathbf{x}(0)$ for $i = 1, 2, \ldots, N$ and used the properties of probabilistic quantization and the fact that $|v_i(t)| \leq \Delta$ almost surely. Similarly, we have shown that

$$\mathbb{E}\{\min_i\{q_i(t)\}\} \geq \mathbf{w}_{i_*}^t\mathbf{x}(0) - \Delta \tag{85}$$

where $i_* = i : \mathbf{w}_{i_*}^t\mathbf{x}(0) \leq \mathbf{w}_i^t\mathbf{x}(0)$ for $i = 1, 2, \ldots, N$, yielding

$$\mathbb{E}\{r_{\mathbf{q}}(t)\} \leq (\mathbf{w}_{i^*}^t - \mathbf{w}_{i_*}^t)\mathbf{x}(0) + 2\Delta. \tag{86}$$

Utilizing the Cauchy–Schwartz inequality reduces the above expression to:

$$\mathbb{E}\{r_{\mathbf{q}}(t)\} \leq ||\mathbf{w}_{i^*}^t - \mathbf{w}_{i_*}^t||||\mathbf{x}(0) - \mathbf{J}\mathbf{x}(0)|| + 2\Delta. \tag{87}$$

Clearly, to upper bound $\mathbb{E}\{r_{\mathbf{q}}(t)\}$, we need to upper bound $||\mathbf{w}_{i^*}^t - \mathbf{w}_{i_*}^t||$. Hence, we derive an upper bound for $||\mathbf{w}_{i_1}^t - \mathbf{w}_{i_2}^t||$ for any (i_1, i_2) pair such that $i_1 \neq i_2$, in the following:

$$||\mathbf{w}_{i_1}^t - \mathbf{w}_{i_2}^t|| = ||\mathbf{w}_{i_1}^t - \frac{1}{N}\mathbf{1}^{\mathrm{T}} - (\mathbf{w}_{i_2}^t - \frac{1}{N}\mathbf{1}^{\mathrm{T}})|| \tag{88}$$

$$\leq^{(a)} ||\mathbf{w}_{i_1}^t - \frac{1}{N}\mathbf{1}^{\mathrm{T}}|| + ||\mathbf{w}_{i_2}^t - \frac{1}{N}\mathbf{1}^{\mathrm{T}}|| \tag{89}$$

$$\leq^{(b)} ||\mathbf{W}^t - \mathbf{J}|| + ||\mathbf{W}^t - \mathbf{J}|| \tag{90}$$

$$=^{(c)} ||(\mathbf{W} - \mathbf{J})^t|| + ||(\mathbf{W} - \mathbf{J})^t|| \tag{91}$$

$$\leq^{(d)} ||(\mathbf{W} - \mathbf{J})||^t + ||(\mathbf{W} - \mathbf{J})||^t \tag{92}$$

$$=^{(e)} 2\rho^t(\mathbf{W} - \mathbf{J}) \tag{93}$$

where (a) follows from the Triangle inequality, (b) from the fact that the norm of any row of a matrix is smaller that the norm of the matrix, (c) using the properties of the weight matrix, (d) by the Norm inequality, and (e) due to the symmetric assumption on the weight matrix. Finally, substituting (93) into (87) yields:

$$\mathbb{E}\{r_{\mathbf{q}}(t)\} \leq 2\rho^t(\mathbf{W} - \mathbf{J})||\mathbf{x}(0) - \mathbf{J}\mathbf{x}(0)|| + 2\Delta. \tag{94}$$

Moreover, using Thomson's sharp bound relating order statistics and sample standard deviation (for N even, but a very similar result exists for odd N) [28]:

$$2||\mathbf{x}(0) - \mathbf{J}\mathbf{x}(0)|| \leq \sqrt{\frac{N-1}{N}}\left(\max_i\{x_i(0)\} - \min_i\{x_i(0)\}\right) \triangleq \sqrt{\frac{N-1}{N}}\mathbf{r}_{\mathbf{x}}(0) \tag{95}$$

one can relate the bound given on the quantized node state values range to the initial states' range, *i.e.*, the result stated in the theorem.

REFERENCES

[1] T. C. Aysal, M. J. Coates, and M. G. Rabbat, "Distributed average consensus using probabilistic qantization," in *Proceedings of the IEEE Statistical Signal Processing Workshop*, Madison, WI, Aug. 2007.

[2] ——, "Rates of convergence of distributed average consensus with probabilistic qantization," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep. 2007.

[3] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proceedings of the IEEE/ACM International Symposium on Information Processing in Sensor Networks*, Los Angeles, CA, Apr. 2005.

[4] J.-J. Xiao and Z.-Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3564–3575, Oct. 2005.

[5] H. Papadopoulos, G. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 978–1002, Mar. 2001.

[6] N. Lynch, *Distributed Algorithms*. Morgan Kaufmann Publishers, Inc., San Francisco, CA, 1996.

[7] W. Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[8] D. S. Scherber and H. C. Papadopoulos, "Locally constructed algorithms for distributed computations in ad-hoc networks," in *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*, Berkeley, CA, Apr. 2004.

[9] C. C. Moallemi and B. V. Roy, "Consensus propagation," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 4753–4766, Nov. 2006.

[10] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Distributed sensor fusion using dynamic consensus," in *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, Jul. 2005.

[11] C.-Z. Xu and F. Lau, *Load balancing in parallel computers: theory and practice*. Kluwer, Dordrecht, 1997.

[12] Y. Rabani, A. Sinclair, and R. Wanka, "Local divergence of markov chains and the analysis of iterative load-balancing schemes," in *Proceedings of the IEEE Symposium on Foundations of Computer Science*, Palo Alto, CA, Nov. 1998.

[13] M. Rabbat, R. Nowak, and J. Bucklew, "Robust decentralized source localization via averaging," in *Proceedings of the IEEE ICASSP*, Philadelphia, PA, Mar. 2005.

[14] M. Rabbat, J. Haupt, A. Singh, and R. Nowak, "Decentralized compression and predistribution via randomized gossiping," in *Proceedings of the Information Processing in Sensor Networks*, Nashville, TN, Apr. 2006.

[15] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," in *Proceedings of the ACM/IEEE Internation Conference on Mobile Computing and Networking*, Boston, MA, Aug. 2000.

[16] J. Zhao, R. Govindan, and D. Estrin, "Computing aggregates for monitoring wireless sensor networks," in *Proceedings of the International Workshop on Sensor Network Protocols and Applications*, Anchorage, AL, May 2003.

[17] S. R. Madden, R. Szewczyk, M. J. Franklin, and D. Culler, "Supporting aggregate queries over ad-hoc wireless sensor networks," in *Proceedings of the Workshop on Mobile Computing Systems and Applications*, Callicoon, NY, Jun. 2002.

[18] A. Montresor, M. Jelasity, and O. Babaoglu, "Robust aggregation protocols for large-scale overlay networks," in *Proceedings of the International Conference on Dependable Systems and Networks*, Florence, Italy, Jun. 2004.

[19] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least–mean–square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33–46, Jan. 2007.

[20] M. E. Yildiz and A. Scaglione, "Differential nested lattice encoding for consensus problems," in *Proceedings of the Information Processing in Sensor Networks*, Cambridge, MA, Apr. 2007.

[21] A. Kashyap, T. Basar, and R.Srikant, "Quantized consensus," *Automatica*, vol. 43, pp. 1192–1203, Jul. 2007.

[22] R. A. Wannamaker, S. P. Lipshitz, J. Vanderkooy, and J. N. Wright, "A theory of nonsubtractive dither," *IEEE Transactions on Signal Processing*, vol. 8, no. 2, pp. 499–516, Feb. 2000.

[23] L. Schuchman, "A theory of nonsubtractive dither," *IEEE Transactions on Communication Technology*, vol. COMM–12, pp. 162–165, Dec. 1964.

[24] O. Kallenberg, *Foundations of Modern Probability*. Springer–Verlag, Second Ed., 2002.

[25] L. P. Devroye, "Inequalities for the completion times of PERT networks," *Mathematics of Operations Research*, vol. 4, no. 4, pp. 441–447, Nov. 1979.

[26] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, Sept. 2004.

[27] M. G. Rabbat, R. D. Nowak, and J. A. Bucklew, "Generalized consensus algorithms in networked systems with erasure links," in *Proceedings IEEE Workshop on Signal Processing Advances in Wireless Communications*, New York, NY, June 2005.

[28] G. W. Thomson, "Bounds for the ratio of range to standard deviation," *Biometrika*, vol. 42, no. 1/2, pp. 268–269, Jun. 1955.