Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

On Distributed Averaging Algorithms and Quantization Effects

Angelia Nedić, Alex Olshevsky, Asuman Ozdaglar, John N. Tsitsiklis

***Abstract*— We consider distributed iterative algorithms for the averaging problem over time-varying topologies. Our focus is on the convergence time of such algorithms when complete (unquantized) information is available, and on the degradation of performance when only quantized information is available. We study a large and natural class of averaging algorithms, which includes the vast majority of algorithms proposed to date, and provide tight polynomial bounds on their convergence time. We then propose and analyze distributed averaging algorithms under the additional constraint that agents can only store and communicate quantized information. We show that these algorithms converge to the average of the initial values of the agents within some error. We establish bounds on the error and tight bounds on the convergence time, as a function of the number of quantization levels.**

I. Introduction

There has been much recent interest in distributed control and coordination of networks consisting of multiple, potentially mobile, agents. This is motivated mainly by the emergence of large scale networks, characterized by the lack of centralized access to information and time-varying connectivity. Control and optimization algorithms deployed in such networks should be completely distributed, relying only on local observations and information, and robust against unexpected changes in topology such as link or node failures.

A canonical problem in distributed control is the *consensus problem*. The objective in the consensus problem is to develop distributed algorithms that can be used by a group of agents in order to reach agreement (consensus) on a common decision (represented by a scalar or a vector value). The agents start with some different initial decisions and communicate them locally under some constraints on connectivity and inter-agent information exchange. The consensus problem arises in a number of applications including coordination of UAVs (e.g., aligning the agents' directions of motion), information processing in sensor networks, and distributed optimization (e.g., agreeing on the estimates of some unknown parameters). The *averaging problem* is a special case in which the goal is to compute the exact average of the initial values of the agents. A natural and widely studied consensus algorithm, proposed and analyzed in [14] and [15], involves, at each time step, every agent taking a weighted average of its own value with values received from some of the other agents. Similar algorithms have been studied in the load-balancing literature (see for example [6]). Motivated by observed group behavior in biological and dynamical systems, the recent literature in cooperative control has studied similar algorithms and proved convergence results under various assumptions on agent connectivity and information exchange (see [16], [7], [12], [11],[10]).

In this paper, our goal is to provide tight bounds on the convergence time (defined as the number of iterations required to reduce a suitable Lyapunov function by a constant factor) of a general class of consensus algorithms, as a function of the number n of agents. We focus on algorithms that are designed to solve the averaging problem. We consider both problems where agents have access to exact values and problems where agents only have access to quantized values of the other agents. Our contributions can be summarized as follows.

In the first part of the paper, Sections II and III, we consider the case where agents can exchange and store continuous values, which is a widely adopted assumption in the previous literature. We consider a large class of averaging algorithms defined by the condition that the weight matrix is a possibly nonsymmetric, doubly stochastic matrix. For this class of algorithms, we prove that the convergence time is $O(n^2/\eta)$, where n is the number of agents and η is a lower bound on the nonzero weights used in the algorithm. To the best of our knowledge, *this is the best polynomial-time bound on the convergence time of such algorithms*. We also show that this bound is tight. We remark that it is possible to choose the coefficients in a distributed manner in order to improve this convergence time to $O(n^2)$. This matches the best currently available convergence time guarantee for the much simpler case of static connectivity graphs [13].

A. Nedić is with the Industrial and Enterprise Systems Engineering Department, University of Illinois at Urbana-Champaign, Urbana IL 61801 (e-mail:angelia@uiuc.edu)

A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis are with the Laboratory for Information and Decision Systems, Electrical Engineering and Computer Science Department, Massachusetts Institute of Technology, Cambridge MA, 02139 (e-mails: alex_o@mit.edu, asuman@mit.edu, jnt@mit.edu)

This research was partially supported by the National Science Foundation under grants ECCS- 0701623, DMI-0545910 and CMMI 07-42538, and by the DARPA ITMANET program.

In the second part of the paper, Section IV, we impose the additional constraint that agents can only store and transmit quantized values. This model provides a good approximation for communication networks that are subject to communication bandwidth or storage constraints. We focus on a particular quantization rule, which rounds down the values to the nearest quantization level. We propose a distributed algorithm that uses quantized values and we prove that all agents have the same value after $O((n^2/\eta)\log(nQ))$ time steps, where Q is the number of quantization levels per unit value. Due to the rounding-down feature of the quantizer, this algorithm does not preserve the average of the values at each iteration. However, we provide bounds on the error between the final consensus value and the initial average, as a function of the number Q of available quantization levels. In particular, we show that the error goes to 0 at a rate of $(\log Q)/Q$, as the number Q of quantization levels increases to infinity.

Other than the papers cited above, our work is also related to [8], [3], and [4], which study the effects of quantization on the performance of averaging algorithms. In [8], Kashyap *et al.* proposed randomized *gossip-type* quantized averaging algorithms under the assumption that each agent value is an integer. They showed that these algorithms converge to approximate consensus and provided bounds on the convergence time for specific static topologies. In the recent work [3], Carli *et al.* proposed a distributed algorithm that uses quantized values and showed favorable convergence properties using simulations on some static topologies. Our results on quantized averaging algorithms differ from these works in that *we study the more general case of time-varying topologies, and provide tight polynomial bounds on both the convergence time and the discrepancy from the initial average, in terms of the number of quantization levels.*

II. Preliminaries on Distributed Averaging

We consider a set $N = \{1, 2, \ldots, n\}$ of agents, which will henceforth be referred to as “nodes.” Each node i starts with a scalar value $x_i(0)$. At each nonnegative integer time k, node i receives from some of the other nodes j a message with the value of $x_j(k)$, and updates its value according to

$$x_i(k+1) = \sum_{j=1}^{n} a_{ij}(k)x_j(k), \qquad (1)$$

where the $a_{ij}(k)$ are nonnegative weights with the property that $a_{ij}(k) > 0$ only if node i receives information from node j at time k. We use the notation $A(k)$ to denote the *weight matrix* $[a_{ij}(k)]_{i,j=1,\ldots,n}$. Given a matrix A, we use $\mathcal{E}(A)$ to denote the set of directed edges (j, i), including self-edges (i, i), such that $a_{ij} > 0$. At each time k, the nodes' connectivity can be represented by the directed graph $G(k) = (N, \mathcal{E}(A(k)))$.

Our goal is to study the convergence of the iterates $x_i(k)$ to the average of the initial values, $(1/n)\sum_{i=1}^{n} x_i(0)$, as k approaches infinity. In order to establish such convergence, we impose some assumptions on the weights $a_{ij}(k)$ and the graph sequence $G(k)$.

Assumption 1: For all $k \geq 0$, the weight matrix $A(k)$ is a doubly stochastic matrix with positive diagonal. Additionally, there exists a constant $\eta > 0$ such that if $a_{ij}(k) > 0$, then $a_{ij}(k) \geq \eta$.

The doubly stochasticity assumption on the weight matrix guarantees that the average of the node values remains the same at each iteration. The second part of this assumption states that each node gives significant weight to its values and to the values of its neighbors at each time k.

Our next assumption ensures that the graph sequence $G(k)$ is sufficiently connected for the nodes to repeatedly influence each other's values.

Assumption 2: There exists an integer $B \geq 1$ such that the directed graph

$$\left(N, \mathcal{E}(A(kB)) \bigcup \cdots \bigcup \mathcal{E}(A((k+1)B-1))\right)$$

is strongly connected for all $k \geq 0$.

Any algorithm of the form given in Eq. (1) with the sequence of weights $a_{ij}(k)$ satisfying Assumptions 1 and 2 solves the averaging problem. This is formalized in the following theorem.

Theorem 1: Let Assumptions 1 and 2 hold. Let $\{x(k)\}$ be generated by the iteration (1). Then, for all i, we have

$$\lim_{k\to\infty} x_i(k) = \frac{1}{n}\sum_{j=1}^{n} x_j(0).$$

This theorem is a minor modification of known results in [14], [15], [7], [2], where the convergence of each $x_i(k)$ to the same value is established under weaker versions of Assumptions 1 and 2. The fact that the limit is the average of the entries of the vector $x(0)$ follows from the fact that multiplication of a vector by a doubly stochastic matrix preserves the average of the vector's components.

Recent research has focused on methods of choosing weights $a_{ij}(k)$ that satisfy Assumptions 1 and 2, and minimize the convergence time of the resulting averaging algorithm (see [17] for the case of static graphs, see [12] and [1] for the case of symmetric weights, i.e., weights satisfying $a_{ij}(k) = a_{ji}(k)$, and also see [5]). For static graphs, some recent results on optimal time-invariant algorithms may be found in [13].

III. Convergence time

In this section, we give an analysis of the convergence time of averaging algorithms of the form (1). Our goal is to obtain tight estimates on the convergence time, under Assumptions 1 and 2. As a convergence measure, we use the "sample variance" of a vector $x \in \mathbb{R}^n$, defined as

$$V(x) = \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

where $\bar{x}$ is the average of the entries of x.

Let $x(k)$ denote the vector of node values at time k [i.e., the vector of iterates generated by algorithm (1) at time k]. We are interested in providing an upper bound on the number of iterations it takes for the "sample variance" $V(x(k))$ to decrease to a small fraction of its initial value $V(x(0))$. We omit the proofs of some of the results due to space restrictions and refer the interested reader to the technical report [**?**].

A. Preliminaries on Doubly Stochastic Matrices

We begin by analyzing how the sample variance $V(x)$ changes when the vector x is multiplied by a doubly stochastic matrix A. The next lemma shows that $V(Ax) \leq V(x)$. Thus, under Assumption 1, the sample variance $V(x(k))$ is nonincreasing in k, and $V(x(k))$ can be used as a Lyapunov function.

Lemma 1: Let A be a doubly stochastic matrix. Then, for all $x \in \mathbb{R}^n$,

$$V(Ax) = V(x) - \sum_{i<j} w_{ij}(x_i - x_j)^2,$$

where w_{ij} is the (i,j)-th entry of the matrix $A^T A$.

Because the weight matrix $A(k)$ has nonnegative entries, the entries $w_{ij}(k)$ of $A(k)^T A(k)$ are nonnegative. In view of this, Lemma 1 implies that

$$V(x(k+1)) \leq V(x(k)) \qquad \text{for all } k,$$

and the amount of variance decrease is given by

$$V(x(k)) - V(x(k+1)) = \sum_{i<j} w_{ij}(k)(x_i(k) - x_j(k))^2.$$

We will use this result to provide a lower bound on the amount of decrease of the sample variance $V(x(k))$ in between iterations.

Since every positive entry of $A(k)$ is at least η, it follows that every positive entry of $A(k)^T A(k)$ is at least η^2. Therefore, it is immediate that

$$\text{if} \quad w_{ij}(k) > 0, \quad \text{then} \quad w_{ij}(k) \geq \eta^2.$$

In our next lemma, we establish a stronger lower bound. In particular, instead of focusing on an individual w_{ij}, we consider all w_{ij} associated with edges (i,j) that cross a particular cut in the graph $(N, \mathcal{E}(A^T A))$, which allows us to provide a lower bound linear in η.

Lemma 2: Let A be a stochastic matrix with positive diagonal, and assume that its smallest positive entry is at least η. If (S^-, S^+) is a partition of the set $N = \{1, \ldots, n\}$ into two disjoint sets with

$$\sum_{i \in S^-,\ j \in S^+} w_{ij} > 0,$$

then

$$\sum_{i \in S^-,\ j \in S^+} w_{ij} \geq \frac{\eta}{2}.$$

B. A Bound on Convergence Time

With the preliminaries on doubly stochastic matrices in place, we can now proceed to derive bounds on the decrease of our Lyapunov function $V(x(k))$ during the interval $[kB, (k+1)B-1]$. In what follows, we denote by $V(k)$ the sample variance $V(x(k))$ at time k.

Lemma 3: Let Assumptions 1 and 2 hold. Let $\{x(k)\}$ be generated by the update rule (1). Suppose that the components $x_i(kB)$ of the vector $x(kB)$ have been ordered from largest to smallest, with ties broken arbitrarily. Then, we have

$$V(kB) - V((k+1)B) \geq \frac{\eta}{2} \sum_{i=1}^{n-1} (x_i(kB) - x_{i+1}(kB))^2.$$

Proof: By Lemma 1, we have for all t,

$$V(t) - V(t+1) = \sum_{i<j} w_{ij}(t)(x_i(t) - x_j(t))^2, \quad (2)$$

where $w_{ij}(t)$ is the (i,j)-th entry of $A(t)^T A(t)$. Summing up the variance differences $V(t) - V(t+1)$ over different values of t, we obtain

$$V(kB) - V((k+1)B) = \sum_{t=kB}^{(k+1)B-1} \sum_{i<j} w_{ij}(t)(x_i(t) - x_j(t))^2. \quad (3)$$

We next introduce some notation.

(a) For all $d \in \{1, \ldots, n-1\}$, let t_d be the first time larger than or equal to kB (if it exists) at which there is a communication between two nodes belonging to the two sets $\{1, \ldots, d\}$ and $\{d+1, \ldots, n\}$, to be referred to as a communication across the cut d.

(b) For all $t \in \{kB, \ldots, (k+1)B-1\}$, let $D(t) = \{d \mid t_d = t\}$, i.e., $D(t)$ consists of "cuts" $d \in \{1, \ldots, n-1\}$ such that time t is the first communication time larger than or equal to kB between nodes in the sets $\{1,\ldots,d\}$ and $\{d+1,\ldots,n\}$. Because of Assumption 2, the union of the sets $D(t)$ includes all indices $1,\ldots,n-1$.

(c) For all $d \in \{1,\ldots,n-1\}$, let $C_d = \{(i,j) \mid i \leq d,\ d+1 \leq j\}$.

(d) For all $t \in \{kB,\ldots,(k+1)B-1\}$, let $F_{ij}(t) = \{d \in D(t) \mid (i,j) \in C_d\}$, i.e., $F_{ij}(t)$ consists of all cuts d such that the edge (i,j) at time t is the first communication across the cut at a time larger than or equal to kB.

(e) To simplify notation, let $y_i = x_i(kB)$. By assumption, we have $y_1 \geq \cdots \geq y_n$.

We make two observations:

(1) Suppose that $d \in D(t)$. Then, for some $(i,j) \in C_d$, we have either $a_{ij}(t) > 0$ or $a_{ji}(t) > 0$. In either case, we obtain $w_{ij}(t) > 0$. By Lemma 2, we obtain
$$\sum_{(i,j)\in C_d} w_{ij}(t) \geq \frac{\eta}{2}. \tag{4}$$

(2) Fix some (i,j), with $i < j$, and time $t \in \{kB,\ldots,(k+1)B-1\}$, and suppose that $F_{ij}(t)$ is nonempty. Let $F_{ij}(t) = \{d_1,\ldots,d_k\}$, where the d_j are arranged in increasing order. Since $d_1 \in F_{ij}(t)$, we have $d_1 \in D(t)$ and therefore $t_{d_1} = t$. By the definition of t_{d_1}, this implies that there has been no communication between a node in $\{1,\ldots,d_1\}$ and a node in $\{d_1+1,\ldots,n\}$ during the time interval $[kB, t-1]$. It follows that $x_i(t) \geq y_{d_1}$. By a symmetrical argument, we also have
$$x_j(t) \leq y_{d_k+1}. \tag{5}$$

These relations imply that
$$\begin{aligned} x_i(t) - x_j(t) &\geq y_{d_1} - y_{d_k+1} \\ &= \sum_{h=1}^{k-1}(y_{d_h} - y_{d_{h+1}}) \\ &\quad +(y_{d_k} - y_{d_k+1}) \\ &\geq \sum_{d\in F_{ij}(t)} (y_d - y_{d+1}), \end{aligned}$$
where the last inequality follows because we have $y_{d_i} - y_{d_{i+1}} \geq y_{d_i} - y_{d_i+1}$ for all $i = 1,\ldots,k-1$. Since the components of y are sorted in nonincreasing order, we have $y_d - y_{d+1} \geq 0$, for every $d \in F_{ij}(t)$. For any nonnegative numbers z_i, we have
$$(z_1 + \cdots + z_k)^2 \geq z_1^2 + \cdots + z_k^2,$$
which implies that
$$(x_i(t) - x_j(t))^2 \geq \sum_{d\in F_{ij}(t)} (y_d - y_{d+1})^2. \tag{6}$$

We now use these two observations to provide a lower bound on the expression on the right-hand side of Eq. (2) at time t. We use Eq. (6) and then Eq. (4), to obtain
$$\begin{aligned} \sum_{i<j} w_{ij}(t)(x_i(t) &- x_j(t))^2 \\ &\geq \sum_{i<j} w_{ij}(t) \sum_{d\in F_{ij}(t)} (y_d - y_{d+1})^2 \\ &= \sum_{d\in D(t)} \sum_{(i,j)\in C_d} w_{ij}(t)(y_d - y_{d+1})^2 \\ &\geq \frac{\eta}{2} \sum_{d\in D(t)} (y_d - y_{d+1})^2. \end{aligned}$$

We sum both sides of the above inequality for different values of t, and use Eq. (3), to obtain
$$\begin{aligned} V(kB) &- V((k+1)B) \\ &= \sum_{t=kB}^{(k+1)B-1} \sum_{i<j} w_{ij}(t)(x_i(t) - x_j(t))^2 \\ &\geq \frac{\eta}{2} \sum_{t=kB}^{(k+1)B-1} \sum_{d\in D(t)} (y_d - y_{d+1})^2 \\ &= \frac{\eta}{2} \sum_{d=1}^{n-1} (y_d - y_{d+1})^2, \end{aligned}$$
where the last inequality follows from the fact that the union of the sets $D(t)$ is only missing those d for which $y_d = y_{d+1}$. ■

We next establish a bound on the variance decrease.

Lemma 4: Let Assumptions 1 and 2 hold, and suppose that $V(kB) > 0$. Then,
$$\frac{V(kB) - V((k+1)B)}{V(kB)} \geq \frac{\eta}{2n^2} \qquad \text{for all } k.$$

Proof: Without loss of generality, we assume that the components of $x(kB)$ have been sorted in nonincreasing order. We introduce the notation $\Delta(kB) = V(kB) - V((k+1)B)$. By Lemma 3, we have
$$\Delta(kB) \geq \frac{\eta}{2} \sum_{i=1}^{n-1} (x_i(kB) - x_{i+1}(kB))^2.$$
This implies that
$$\frac{\Delta(kB)}{V(kB)} \geq \frac{\eta}{2} \frac{\sum_{i=1}^{n-1} (x_i(kB) - x_{i+1}(kB))^2}{\sum_{i=1}^{n} (x_i(kB) - \bar{x}(kB))^2}.$$
Observe that the right-hand side does not change when we add a constant to every $x_i(kB)$. We can therefore assume, without loss of generality, that $\bar{x}(kB) = 0$, so that
$$\frac{\Delta(kB)}{V(kB)} \geq \frac{\eta}{2} \min_{\substack{x_1 \geq x_2 \geq \cdots \geq x_n \\ \sum_i x_i = 0}} \frac{\sum_{i=1}^{n-1} (x_i - x_{i+1})^2}{\sum_{i=1}^{n} x_i^2}.$$

Note that the right-hand side is unchanged if we multiply each x_i by the same constant. Therefore, we can assume, without loss of generality, that $\sum_{i=1}^n x_i^2 = 1$, so that

$$\frac{\Delta(kB)}{V(kB)} \geq \frac{\eta}{2} \min_{\substack{x_1 \geq x_2 \geq \cdots \geq x_n \\ \sum_i x_i = 0,\ \sum_i x_i^2 = 1}} \sum_{i=1}^{n-1} (x_i - x_{i+1})^2. \quad (7)$$

The requirement $\sum_i x_i^2 = 1$ implies that the average value of x_i^2 is $1/n$, which implies that there exists some j such that $|x_j| \geq 1/\sqrt{n}$. Without loss of generality, let us suppose that this x_j is positive.[1]

The rest of the proof relies on a technique from [9] to provide a lower bound on the right-hand side of Eq. (7). Let

$$z_i = x_i - x_{i+1} \quad \text{for } i < n, \quad \text{and} \quad z_n = 0.$$

Note that $z_i \geq 0$ for all i and

$$\sum_{i=1}^n z_i = x_1 - x_n.$$

Since $x_j \geq 1/\sqrt{n}$ for some j, we have that $x_1 \geq 1/\sqrt{n}$; since $\sum_{i=1}^n x_i = 0$, it follows that at least one x_i is negative, and therefore $x_n < 0$. This gives us

$$\sum_{i=1}^n z_i \geq \frac{1}{\sqrt{n}}.$$

Combining with Eq. (7), we obtain

$$\frac{\Delta(kB)}{V(kB)} \geq \frac{\eta}{2} \min_{z_i \geq 0,\ \sum_i z_i \geq 1/\sqrt{n}} \sum_{i=1}^n z_i^2.$$

The minimization problem on the right-hand side is a symmetric convex optimization problem, and therefore has a symmetric optimal solution, namely $z_i = 1/n^{1.5}$ for all i. This results in an optimal value of $1/n^2$. Therefore, $\Delta(kB)/V(kB) \geq \eta/(2n^2)$, which is the desired result. ∎

Our main result follows immediately from Lemma 4.

Theorem 2: Let Assumptions 1 and 2 hold. Then, there exists an absolute constant[2] c such that we have

$$V(k) \leq \epsilon V(0) \qquad \text{for all } k \geq c(n^2/\eta) B \log(1/\epsilon).$$

The next theorem shows that the convergence time bound of the preceding theorem is tight.

[1] Otherwise, we can replace x with $-x$ and subsequently reorder to maintain the property that the components of x are in descending order. It can be seen that these operations do not affect the objective value.

[2] We say c is an absolute constant when it does not depend on any of the parameters in the problem, in this case n, B, η, ϵ.

Theorem 3: There exist constants c and n_0 with the following property. For any $n \geq n_0$, nonnegative integer B, $\eta < 1/2$, and $\epsilon < 1$, there exist a sequence of weight matrices $A(k)$ satisfying Assumptions 1 and 2, and an initial value $x(0)$ such that if $V(k)/V(0) \leq \epsilon$, then

$$k \geq c \frac{n^2}{\eta} B \log \frac{1}{\epsilon}.$$

It is possible to synthesize the coefficients $a_{ij}(k)$ in a distributed manner using the 3-hop neighborhood information for each node. This yields an averaging algorithm with the improved $O(n^2 B)$ convergence time. The details of this algorithm can be found in the technical report [**?**].

IV. Quantization Effects

In this section, we consider a quantized version of the update rule (1). This model is a good approximation for a network of nodes communicating through finite bandwidth channels, so that at each time instant, only a finite number of bits can be transmitted. We incorporate this constraint in our algorithm by assuming that each node, upon receiving the values of its neighbors, computes the convex combination $\sum_{j=1}^n a_{ij}(k) x_j(k)$ and quantizes it. This update rule also captures a constraint that each node can only store quantized values.

Unfortunately, under Assumptions 1 and 2, if the output of Eq. (1) is rounded to the nearest integer, the sequence $x(k)$ is not guaranteed to converge to consensus; see [8]. We therefore choose a quantization rule that rounds the values down, according to

$$x_i(k+1) = \left\lfloor \sum_{j=1}^n a_{ij}(k) x_j(k) \right\rfloor, \qquad (8)$$

where $\lfloor \cdot \rfloor$ represents rounding *down* to the nearest multiple of $1/Q$, and where Q is some positive integer. For convenience we define $U = \max_i x_i(0)$ and $L = \min_i x_i(0)$.

We adopt a slightly different measure of convergence for the analysis of the quantized consensus algorithm. For any $x \in \mathbb{R}^n$, we define $m(x) = \min_i x_i$ and

$$\underline{V}(x) = \sum_{i=1}^n (x_i - m(x))^2.$$

We will also use the simpler notation $m(k)$ and $\underline{V}(k)$ to denote $m(x(k))$ and $\underline{V}(x(k))$, respectively, where it is more convenient to do so. The function $\underline{V}$ will be our Lyapunov function for the analysis of the quantized consensus algorithm. The reason for not using our earlier Lyapunov function, V, is that for the quantized algorithm, V is not guaranteed to be monotonically nonincreasing in time, because the operation of quantizing down can increase V. On the other hand, $\underline{V}$ is only decreased when each value is quantized down. It can be verified that for any x, we have $V(x) \leq \underline{V}(x) \leq nV(x)$. As a consequence, any convergence time bounds expressed in terms of $\underline{V}$ translate to essentially the same bounds expressed in terms of V, up to a logarithmic factor. The next theorem contains our main result on the convergence time of the quantized algorithm.

Theorem 4: Let Assumptions 1 and 2 hold. Let $\{x(k)\}$ be generated by the update rule (8). Then, there exists an absolute constant c such that for all $\epsilon > 0$,

$$\underline{V}(k) \leq \epsilon \underline{V}(0) \qquad \text{for all } k \geq c\,(n^2/\eta)B\log(1/\epsilon).$$

Similar to Theorem 3, we can establish that the bound in this theorem is tight.

Despite favorable convergence properties of our quantized averaging algorithm (8), the update rule does not preserve the average of the values at each iteration. Therefore, the common limit of the sequences $x_i(k)$, denoted by x_f, need not be equal to the exact average of the initial values. We next provide an upper bound on the error between x_f and the initial average, as a function of the number of quantization levels.

Theorem 5: Let Assumptions 1 and 2 hold. Then, there is an absolute constant c such that for the common limit x_f of the values $x_i(k)$ generated by the quantized algorithm (8), we have

$$\left| x_f - \frac{1}{n}\sum_{i=1}^{n} x_i(0) \right| \leq \frac{c}{Q}\;\frac{n^2}{\eta}\,B\log(Qn(U-L)).$$

Let us assume that the parameters B, η, and $U-L$ are fixed. Theorem 5 implies that as n increases, the number of bits used for each communication, which is proportional to $\log Q$, needs to grow only as $O(\log n)$ to make the error negligible. Furthermore, this is true even if the parameters B, $1/\eta$, and $U-L$ grow polynomially in n.

V. Conclusions

We studied distributed algorithms for the averaging problem over networks with time-varying topology, with a focus on tight bounds on the convergence time of a general class of averaging algorithms. We first considered algorithms for the case where agents can exchange and store continuous values, and established tight convergence time bounds. We next studied averaging algorithms under the additional constraint that agents can only store and send quantized values. We showed that these algorithms guarantee convergence of the agent values to consensus within some error from the average of the initial values and provided a bound on the error that highlights the dependence on the number of quantization levels. Future work includes investigation of other quantization schemes and their impact on convergence time.

References

[1] P.A. Bliman and G. Ferrari-Trecate, *Average consensus problems in networks of agents with delayed communications*, Proceedings of the Joint 44th IEEE Conference on Decision and Control and European Control Conference, 2005.

[2] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis, *Convergence in multiagent coordination, consensus, and flocking*, Proceedings of the Joint 44th IEEE Conference on Decision and Control and European Control Conference, 2005.

[3] R. Carli, F. Fagnani, P. Frasca, T. Taylor, and S. Zampieri, *Average consensus on networks with transmission noise or quantization*, Proceedings of European Control Conference, 2007.

[4] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri, *Communication constraints in coordinated consensus problems*, Proceedings of IEEE ACC, 2006.

[5] J. Cortes, *Analysis and design of distributed algorithms for chi-consensus*, Proceedings of the 45th IEEE Conference on Decision and Control, 2006.

[6] G. Cybenko, *Dynamic load balancing for distributed memory multiprocessors*, Journal of Parallel and Distributed Computing **7** (1989), no. 2, 279–301.

[7] A. Jadbabaie, J. Lin, and A.S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Transactions on Automatic Control **48** (2003), no. 3, 988–1001.

[8] A. Kashyap, T. Basar, and R. Srikant, *Quantized consensus*, Proceedings of the 45th IEEE Conference on Decision and Control, 2006.

[9] H.J. Landau and A.M. Odlyzko, *Bounds for the eigenvalues of certain stochastic matrices*, Linear Algebra and its Applications **38** (1981), 5–15.

[10] Q. Li and D. Rus, *Global clock synchronization in sensor networks*, IEEE Transactions on Computers **55** (2006), no. 2, 214–224.

[11] L. Moreau, *Stability of multiagent systems with time-dependent communication links*, IEEE Transactions on Automatic Control **50** (2005), no. 2, 169–182.

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

[13] A. Olshevsky and J.N. Tsitsiklis, *Convergence speed in distributed consensus and averaging*, Preprint, 2006.

[14] J.N. Tsitsiklis, *Problems in decentralized decision making and computation*, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, MIT, 1984.

[15] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Transactions on Automatic Control **31** (1986), no. 9, 803–812.

[16] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Schochet, *Novel type of phase transitions in a system of self-driven particles*, Physical Review Letters **75** (1995), no. 6, 1226–1229.

[17] L. Xiao and S. Boyd, *Fast linear iterations for distributed averaging*, Systems and Control Letters **53** (2004), 65–78.