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

Average consensus by gossip algorithms with quantized communication

Paolo Frasca, Ruggero Carli, Fabio Fagnani, and Sandro Zampieri

Abstract— **This work studies how the randomized gossip algorithm can solve the average consensus problem on networks with quantized communications. The algorithm is proved to converge to the average value, up to the size of the quantization bins, whenever the the graph is connected. Moreover, its speed of convergence is estimated.**

I. INTRODUCTION

In the latest years, an increasing interest in the control community has been devoted to distributed systems, and much research has been focused on the so called average consensus problem [13], [3]. Suppose we have a graph $\mathcal{G}$ with set of nodes $\mathcal{V} = \{1, \dots, N\}$ and a real quantity x_i for every node $i \in \mathcal{V}$. The average consensus problem consists in computing the average $x_{ave} = N^{-1} \sum_i x_i$ in an iterative and distributed way, exchanging information among nodes exclusively along the available edges in $\mathcal{G}$. Most of the literature on the consensus problem assumes that the communication channel between the nodes allows to transfer real numbers with no errors: however, from the implementation point of view, it is more natural to assume the agents to communicate by finite capacity digital channels. This clearly forces a quantization on the real numbers that agents have to transmit. This crucial issue has attracted attention only very recently, in the works [14], [4], [1], [7], [9]. Other quantization effects have also been considered, namely [10] studies the convergence when the agents have quantized states, and [12] when quantization is in the updating rule.

The use of quantized communication has been noticed since [14] to complicate the convergence of the algorithms used for the average consensus problem. This happens mainly because the initial average of states is not preserved. The correction term introduced in [4] allows to avoid this shortcoming and to preserve the average at each step of the algorithm, every time the evolution matrix is doubly stochastic.

This mild assumption covers the case of the gossip algorithm, introduced in [2] and further discussed in [6], in which at each time step only a randomly chosen pair of neighboring nodes exchanges information and performs an adjournment of their states. In this work we prove that, under the constraint of quantized communication, an adaptation of such algorithm converges to the average consensus, up to the size of the quantization bin, for any connected undirected graph. This is shown in Section III, while in Section IV we give estimates of its speed of convergence. In Section V, we present significant simulations and draw our conclusions.

Fabio Fagnani and Paolo Frasca are with the Dipartimento di Matematica del Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy. {fabio.fagnani,paolo.frasca}@polito.it

Ruggero Carli is with the Department of Mechanical Engineering, University of California at Santa Barbara, CA 93106-5070, USA. carlirug@engineering.ucsb.edu

Sandro Zampieri is with the DEI, University of Padova, Via Gradenigo 6/a, 35131 Padova, Italy. zampi@dei.unipd.it

II. STATEMENT

Assume we are given an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Let $\mathcal{V} = \{1, \dots, N\}$ be the set of vertices (nodes) of the graph, and let $\mathcal{E}$ be the set of edges, which is a subset of $\{\{i, j\} : i, j \in \mathcal{V}, i \neq j\}$. Each vertex corresponds to an agent, and each edge to an available link connecting two agents. Two agents connected by a link are said to be *neighbors*. We assume that the graph is *connected*, that is, for any given pair of vertices $\{i, j\}$ there exists a path which connects i to j. A *path* in $\mathcal{G}$ consists in a (ordered) sequence of vertices $(i = i_1, i_2, \dots, i_r = j)$ such that $\{i_j, i_{j+1}\} \in \mathcal{E}$ for every $j \in \{1, \dots, r-1\}$. A graph is said to be *fully connected* or *complete* if $\mathcal{E} = \{\{i, j\} : i, j \in \mathcal{V}, i \neq j\}$.

Each agent $i \in \mathcal{V}$ is given a time-dependent state $x_i(t) \in \mathbb{R}$ and the goal is to design an adjournment algorithm such that in the limit each agent's state tends to the average of initial states.

In this work, we assume that the agents cannot access the values of their neighbors' states, but only an *integer approximation* of them. This is intended to model a system in which only digital communications are allowed between agents. With this constraint, it is clear that the agents' states can not converge to an exact consensus. However, algorithms can be designed to drive the system to a weaker condition of quantized average consensus. Let $x(t)$ be the vector of states $[x_1(t), \dots, x_N(t)]^*$ where * denotes the transpose.

Definition 1: A *quantized average consensus state* is a state $\bar{x} \in \mathbb{R}^N$ such that $\left|\bar{x}_i - N^{-1} \sum_{j=1}^N x_j(0)\right| < 1$ for all $i \in \mathcal{V}$. The algorithm is said to have reached the quantized average consensus if T_{con} exists such that, for all $t > T_{con}$, $x(t)$ is a quantized average consensus state.

We define a quantized gossip algorithm as follows. At each time step, one edge $\{i, j\}$ is randomly selected in $\mathcal{E}$ with probability $P_{\{i,j\}}$ such that $\sum_{\{i,j\} \in \mathcal{E}} P_{\{i,j\}} = 1$. Let $P \in \mathbb{R}^{N \times N}$ be

$$P_{ij} = P_{ji} = \begin{cases} P_{\{i,j\}} & \text{if} \quad \{i,j\} \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Remark that, if we denote as $\mathbf{1}$ the vector of length N whose component are all equal to 1, then $\mathbf{1}^* P \mathbf{1} = 2$. The agents insisting on the selected edge average their states following

$$\begin{aligned} x_i(t+1) &= x_i(t) - \alpha q[x_i(t)] + \alpha q[x_j(t)] \\ x_j(t+1) &= x_j(t) - \alpha q[x_j(t)] + \alpha q[x_i(t)], \end{aligned} \tag{2}$$

where $q[z]$ is the nearest integer to the real number z, and $\alpha \in (0, 1)$ is a parameter of the method. We assume, with no loss of generality, that every edge is selected with positive probability.

It is easy to check that the average of states is preserved, that is, defining $x_{ave}(t) = N^{-1} \sum_{k=1}^{N} x_k(t)$, $x_{ave}(t+1) = x_{ave}(t)$.

Of special interest is the case $\alpha = 1/2$, which is the most natural choice. Moreover, in this case we are able to prove convergence results. From now on, unless otherwise stated, we assume $\alpha = 1/2$.

III. Convergence

In this section we give conditions assuring that with probability one the system converges to a quantized consensus state *in finite time*. To this aim, we can take advantage of a symbolic dynamics which lies under the real states dynamics. The idea of its construction comes from [7], and in its analysis we can adapt results in [10].

To construct such symbolic dynamics we need the following technical lemma, proved in the appendix. Let $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor and ceiling operators from $\mathbb{R}$ to $\mathbb{Z}$.

Lemma 2: Given $a, b \in \mathbb{N}$ and $x \in \mathbb{R}$, it holds

$$\lfloor x \rfloor = \left\lfloor \frac{\lfloor ax \rfloor}{a} \right\rfloor \tag{3}$$

$$q[x] = \lfloor x + 1/2 \rfloor = \left\lceil \frac{1}{2} \left\lfloor \frac{\lfloor 2bx \rfloor}{b} \right\rfloor \right\rceil \tag{4}$$

We define $n_i(t) = \lfloor 2x_i(t) \rfloor$ for all $i \in \mathcal{V}$. Simple properties of floor and ceiling operators, together with the above lemma, allow us to remark that $q[x_i(t)] = \left\lceil \frac{n_i(t)}{2} \right\rceil$ and that

$$\begin{aligned} x_i(t+1) &= x_i(t) - \frac{1}{2} q[x_i(t)] + \frac{1}{2} q[x_j(t)] \\ \lfloor 2x_i(t+1) \rfloor &= \lfloor 2x_i(t) \rfloor - q[x_i(t)] + q[x_j(t)], \end{aligned}$$

from which we can obtain that

$$\begin{aligned} n_i(t+1) &= n_i(t) - \left\lceil \frac{n_i(t)}{2} \right\rceil + \left\lceil \frac{n_j(t)}{2} \right\rceil \\ &= \left\lfloor \frac{n_i(t)}{2} \right\rfloor + \left\lceil \frac{n_j(t)}{2} \right\rceil . \end{aligned}$$

We have thus found an iterative system involving only the symbolic signals $n_i(t)$. When the edge $\{i, j\}$ is selected, i and j adjourn their states following the pair dynamics

$$(n_i(t+1), n_j(t+1)) = g(n_i(t), n_j(t)) \tag{5}$$

where

$$g(h, k) = \left(\left\lfloor \frac{h}{2} \right\rfloor + \left\lceil \frac{k}{2} \right\rceil , \left\lfloor \frac{k}{2} \right\rfloor + \left\lceil \frac{h}{2} \right\rceil \right).$$

It is clear that g is symmetric in the arguments, in the sense that if $g(h, k) = (\eta, \chi)$, then $g(k, h) = (\chi, \eta)$.

The analysis of the evolution of (5) will then allow us to obtain information about the asymptotics of $x_i(t)$, since $n_i(t) = \lfloor 2x_i(t) \rfloor$.

We now start the analysis of system (5), which we based on a slight extension of results in [10]. We define the following quantities

$$m(t) = \min_{1 \leq i \leq N} n_i(t) \qquad M(t) = \max_{1 \leq i \leq N} n_i(t),$$

and

$$D(t) = M(t) - m(t).$$

From the form of (5) one can easily remark that $m(t)$ can not decrease and $M(t)$ can not increase. Hence $D(t)$ is not increasing. A much stronger result about the monotonicity of $D(t)$ is the content of the following lemma.

Lemma 3: If $D(t) \geq 2$, then there exists $\tau \in \mathbb{N}$ such that $\mathbb{P}[D(t + \tau) < D(t)] > 0$.

Proof: Let $\mathcal{I}(t) = \{j \in \mathcal{V} \text{ s.t. } n_j(t) = m(t)\}$. We start by proving that $|\mathcal{I}(t)|$, i.e., the cardinality of $\mathcal{I}(t)$, does not increase and that, if $D(t) \geq 2$, then there is a positive probability that it decreases within a finite number of time steps. Notice first that, for $h, k \in \mathbb{Z}$, $g(h+2, k+2) = g(h, k) + 2$. Hence, by an appropriate translation of the initial condition, we can always restrict ourselves to the case $m(t) \in \{0, 1\}$, which is easier to handle.

Case $m(t) = 0$. In this case it is possible for a nonzero state to decrease to 0, but only in the case of a swap between 0 and 1. This assures that $|\mathcal{I}(t)|$ is nonincreasing. Let $\mathcal{S}(t)$ denote the set of nodes which have value $m(t) + 2$ or more. Since $D(t) \geq 2$ then $\mathcal{S}(t)$ is non empty at time t. Now let $(v_1, v_2, \ldots, v_{p-1}, v_p)$ be a shortest path between $\mathcal{I}(t)$ and $\mathcal{S}(t)$. Such a path exists since $\mathcal{G}$ is connected. Note that $v_1 \in \mathcal{I}(t)$ and $v_p \in \mathcal{S}(t)$ and that $\{v_2, \ldots, v_{p-1}\}$ could be an empty set; in this case a shortest path between $\mathcal{I}(t)$ and $\mathcal{S}(t)$ has length 1. Moreover note also that all the nodes in the path except v_1 and v_p have value 1 at time t, otherwise $(v_1, v_2, \ldots, v_{p-1}, v_p)$ is not a shortest path. Since each edge of the communication graph has a positive probability of being selected in any time, there is also a positive probability that in the $p-1$ time units following t the edges of this path are selected sequentially, starting with the edge $\{v_1, v_2\}$. At the last step of this sequence we have that the values of v_{p-1} and v_p are updated. By observing again, that the pair of value $(0, 1)$ is transformed by (5) into the pair $(1, 0)$ we have that the value of v_{p-1}, when the edge $\{v_{p-1}, v_p\}$ is selected, is equal to 0. This update, for the form of (5), will cause the value of both nodes to be strictly greater than 0. Therefore, this proves that $|\mathcal{I}(t+p-1)| < |\mathcal{I}|$ with positive probability. Clearly, if $|\mathcal{I}(t)| = 1$ then we have also that $D(t+p-1) < D(t)$ with positive probability.

Case $m(t) = 1$. In this case no state can decrease to 1, and then $|\mathcal{I}(t)|$ is not increasing. Let $\mathcal{I}(t)$, $\mathcal{S}(t)$ and $(v_1, v_2, \ldots, v_{p-1}, v_p)$ be defined as in the previous case. Obviously in this case all the nodes $v_2, \ldots, v_{p-1}$ in the path have value equal to 2. Moreover observe that also the sequence of edges $\{v_{p-1}, v_p\}$, $\{v_{p-2}, v_{p-1}\}$,…,$\{v_2, v_3\}$, $\{v_1, v_2\}$ has positive probability of being selected in the $p-1$ time units following t. At the last step of this sequence of edges, the values of v_1 and v_2 are updated. Clearly the value of v_1 is equal to 1. Since the value of v_p at time t is greater or equal to 3, and since the pair $(2, 3)$ is transformed by (5) into $(3, 2)$, we have that the value of v_2 when the edge v_1, v_2 is selected, is greater or equal to 3. This update, for (5), will cause the value of both nodes to be strictly greater than 1. Hence $|\mathcal{I}(t+p-1)| < |\mathcal{I}|$ with positive probability. Again, if $|\mathcal{I}(t)| = 1$ then we have also that $D(t+p-1) < D(t)$ with positive probability.

Consider now the following sequence of times $t_0 = t, t_1, t_2, \ldots$. For each $i \geq 0$, if $|\mathcal{I}(t)| > 1$, then we let t_{i+1} to be the first time for which there is a positive probability that $|\mathcal{I}(t_{i+1})| < |\mathcal{I}(t_i)|$. Let now $k \in \mathbb{N}$ be such that $|\mathcal{I}(t_k)| = 1$. Then we have that $D(t_{k+1}) < D(t_k)$. This concludes the proof. ∎

Before stating the main result regarding the convergence properties of (5), we provide two notational definitions. Let

$$\omega = \left\lfloor \frac{1}{N} \sum_{i=1}^{N} n_i(t) \right\rfloor = \left\lfloor \frac{1}{N} \sum_{i=1}^{N} n_i(0) \right\rfloor,$$

and

$$\mathcal{R} = \left\{ r \in \mathbb{N}^N : r_i \in \{\omega, \omega+1\} \quad \forall\, i \in \mathcal{V} \right\}.$$

Theorem 4: Almost surely there exists $T_{con} \in \mathbb{N}$ such that $D(t) < 2$, for all $t > T_{con}$. Consequently $n(t) \in \mathcal{R}$ for all $t > T_{con}$.

Proof:

Notice first that, fixed the initial condition $n(0)$, there exists a fixed finite set $\mathcal{F}$ such that $n(t) \in \mathcal{F}$ for all t. It follows from a repeated application of Lemma 3 that for every $\bar{n} \in \mathcal{F}$, there exists $t_{\bar{n}}$ such that

$$\mathbb{P}(n(t_{\bar{n}} + s) \in \mathcal{R} | n(s) = \bar{n}) = p_{\bar{n}} > 0\,.$$

Put

$$T = \max_{\bar{n} \in \mathcal{F}} t_{\bar{n}}\,, \quad p = \min_{\bar{n} \in \mathcal{F}} p_{\bar{n}}$$

Using the fact that $\mathcal{R}$ is invariant by the dynamics of $n(t)$ ($n(t) \in \mathcal{R}$ yields $n(t+1) \in \mathcal{R}$), we have that,

$$\mathbb{P}(n(s+T) \in \mathcal{R} | n(s) \in \mathcal{F}) \geq p > 0\,.$$

Now

$$\begin{aligned}\mathbb{P}\left[n(tT) \notin \mathcal{R}\right] &= \prod_{r=2}^{t} \mathbb{P}\left[n(rT) \notin \mathcal{R} | n((r-1)T) \notin \mathcal{R}\right] \cdot \\ &\quad \cdot \mathbb{P}\left[n(T) \notin \mathcal{R} | n(0) \in \mathcal{F}\right] \leq (1-p)^t\,.\end{aligned}$$

Hence, $\mathbb{P}(n(tT) \notin \mathcal{R}\, \forall t) = 0$. This proves the thesis. ∎

We can go back to the original system, and prove the following result. Let $v \in \mathbb{R}^N$ and denote $\|v\|_\infty = \max_i v_i$ and $\|v\|_2 = \left(\sum_i v_i^2\right)^2$.

Corollary 5: Consider the algorithm (2). Let $\alpha = 1/2$. Then, almost surely, there exists $T_{con} \in \mathbb{N}$ such that

$$|x_i(t) - x_j(t)| \leq 1 \quad \forall\, i, j \quad \forall\, t > T_{con}, \tag{6}$$

and hence $\|x(t) - x_{ave}\|_\infty \leq 1$ and $N^{-1/2}\|x(t) - x_{ave}\|_2 \leq 1/2$.

Proof: The proof is an immediate consequence of Theorem 4 and of the relation $n_i(t) = \lfloor 2x_i(t) \rfloor$. ∎

Remark 1: It is worth noting that Lemma 3 and Theorem 4 are an extension of respectively, Lemma 3 and Theorem 1 in [10]. In [10] the authors introduced the so-called class of *quantized gossip algorithms*. According to their definition, a quantized gossip algorithm is as follows. Assume that $\{i, j\}$ is the edge selected at time t and that $n_i(t)$ and $n_j(t)$ are respectively the values of node i and of node j at time t. If $n_i(t) = n_j(t)$ then $n_i(t+1) = n_i(t)$ and $n_j(t+1) = n_j(t)$. Otherwise, defined $D_{ij} = |n_i(t) - n_j(t)|$, if $D_{ij} \geq 1$ the method used to update the values has to satisfy the following three properties:

(P1) $n_i(t+1) + n_j(t+1) = n_i(t) + n_j(t)$,

(P2) if $D_{ij}(t) > 1$ then $D_{ij}(t+1) < D_{ij}(t)$, and

(P3) if $D_{ij}(t) = 1$ and (without loss of generality) $n_i(t) < n_j(t)$, then $n_i(t+1) = n_j(t)$ and $n_j(t+1) = n_i(t)$. This update is called *s*wap.

Now we substitute the property $(P3)$ either with the property

(P3') if $D_{ij}(t) = 1$ and (without loss of generality) $n_i(t) < n_j(t)$, then, if $n_i(t)$ is odd then $n_i(t+1) = n_j(t)$ and $n_j(t+1) = n_i(t)$, otherwise if $n_i(t)$ is even then $n_i(t+1) = n_i(t)$ and $n_j(t+1) = n_j(t)$

or with the property

(P3") if $D_{ij}(t) = 1$ and (without loss of generality) $n_i(t) < n_j(t)$, then, if $n_i(t)$ is even then $n_i(t+1) = n_j(t)$ and $n_j(t+1) = n_i(t)$, otherwise if $n_i(t)$ is odd then $n_i(t+1) = n_i(t)$ and $n_j(t+1) = n_j(t)$.

We call the class of algorithms satisfying (P1), (P2), (P3') or satisfying (P1), (P2), (P3"), *extended quantized gossip algorithms*. It is possible to prove that Lemma 3 and Theorem 1 stated in [10] are true also for this class (the proof are analogous to the proofs of Lemma 3 and Theorem 4 provided in this paper). Moreover it is easy to see that the algorithm (5) satisfies the properties (P1), (P2), (P3'). This represents an alternative way to prove Theorem 4.

A. General case

We conjecture that Corollary 5 can be extended, still using a symbolic dynamics, to cover all cases in which α is a rational number in $(0, 1/2]$.

Conjecture 1: For a gossip algorithm with adjournment step (2), $\alpha = \frac{h}{k}$, and $2h < k$, almost surely there exists $T_{con} \in \mathbb{N}$ such that

$$|x_i(t) - x_j(t)| \leq 1 \quad \forall\, i, j \quad \forall\, t > T, \tag{7}$$

and hence $\|x(t) - x_{ave}\|_\infty \leq 1$ and $N^{-1/2}\|x(t) - x_{ave}\|_2 \leq 1/2$.

Computer simulations (Figure 5), which intrinsically use rational numbers, confirm this convergence property. We expect that similar behavior should also show up in those cases when $\alpha \in (0, 1/2]$ is not necessarily rational but to prove this we would need other ideas: a possibility would be to use suitable rational approximations and a stronger version of Corollary 1 where some uniform estimation of T_{con} is obtained. Finally, we do not expect to have a similar result for $\alpha > 1/2$. Indeed, computer simulations show in this case that there is convergence to a bounded interval, but the length of the interval is bigger than 1, and grows as α approaches 1.

IV. SPEED OF CONVERGENCE

In the above section we just proved convergence for the system (2), with no attention to estimate how fast this convergence occurs. In this section, we get some results in that direction, and we focus on two examples, the complete graph and the ring graph. The latter is the graph with edges set $\mathcal{E} = \{\{i, i+1\} : i = 1, \dots, N-1\} \cup \{(1, N)\}$.

A first approach can be to perform a probabilistic analysis of the symbolic dynamics, and derive bounds on $\mathbb{E}[T_{con}]$. This can be done on the lines of [10], Section 6, using some theory on the hitting times of Markov chains [11]. This provides bounds which for $N \to \infty$ are of order $O(N^3)$ on the complete graph and $O(N^4)$ on the ring graph.

Are these bounds tight? Simulations (Figure 4) suggest that this approach can be too conservative. Then a different approach should be developed.

We recall that the not-quantized gossip algorithm has been proved to converge exponentially fast, and its $\epsilon-$convergence time

$$T_\epsilon = \sup_{x(0)} \inf \left\{ t : \mathbb{P}\left(\frac{||x(t) - x_{ave}\mathbf{1}||_2}{||x(0)||_2} \geq \epsilon \right) \leq \epsilon \right\}$$

has been estimated in [2] and [5] [1].

It comes out that, assuming every edge is selected with equal probability,

- $T_\epsilon = \Theta(N)$ for $N \to \infty$, for the complete graph;
- $T_\epsilon = \Theta(N^3)$ for $N \to \infty$, for the ring graph.

Moreover, in simulations it is evident that the quantized communication algorithm converges almost as fast as the not quantized version, until it approaches its inherent precision limit, the quantization step size. At that moment, it slows down until it reaches its best achievable performance.

Then the idea is to prove that, as long as the distance from the agreement is much larger than the quantization step, the speed of convergence is almost the same as the not-quantized algorithm, until the algorithm starts to slow down.

Consider the following function of $x(t)$

$$V(x(t)) = x^*(t)\Omega x(t) = ||x(t) - x_{ave}||_2^2$$

where $\Omega = I - N^{-1}\mathbf{1}\mathbf{1}^*$. First observe that

$$\begin{aligned} &\mathbb{E}[V(x(t+1))|V(x(t))] = \\ &= \sum_{\{i,j\}\in\mathcal{E}} P_{\{i,j\}} \mathbb{E}[V(x(t+1))|V(x(t)), e(t) = \{i,j\}] \end{aligned}$$

where $e(t)$ denotes the edge selected at time t. Observe now that

$$\begin{aligned} &\mathbb{E}[V(x(t+1))|V(x(t)), e(t) = \{i,j\}] - V(x(t)) = \\ &= (x_i(t+1) - x_{ave})^2 + (x_j(t+1) + x_{ave})^2 - \\ &(x_i(t) - x_{ave})^2 - (x_j(t) - x_{ave})^2 \end{aligned}$$

where $x_{ave} = N^{-1}\sum_i x_i(t)$.

[1] In the latter, the algorithm we are discussing is called *symmetric* gossip.

Then, we substitute (2) and we obtain

$$\begin{aligned} &(x_i(t+1) - x_{ave})^2 + (x_j(t+1) + x_{ave})^2 + \\ &\quad - (x_i(t) - x_{ave})^2 - (x_j(t) - x_{ave})^2 = \\ &\frac{(q[x_i(t)] - q[x_j(t)])^2}{2} - (q[x_i(t)] - q[x_j(t)])\,(x_i(t) - x_j(t)) \\ &= (q[x_i(t)] - q[x_j(t)])\,\frac{q[x_i(t)] - q[x_j(t)]}{2}\,(x_i(t) - x_j(t)). \end{aligned}$$

Since $|q[z] - z| \leq 1/2$ for all $z \in \mathbb{R}$, then

$$x_i(t) - x_j(t) - 1 \leq q[x_i(t)] - q[x_j(t)] \leq x_i(t) - x_j(t) + 1.$$

The latter inequality implies

$$\begin{aligned} &(q[x_i(t)] - q[x_j(t)])\,(\frac{q[x_i(t)] - q[x_j(t)]}{2} - (x_i(t) - x_j(t))) \\ &\leq -\frac{1}{2}(x_i(t) - x_j(t))^2 + \frac{1}{2}, \end{aligned}$$

and then

$$\begin{aligned} &\mathbb{E}[V(x(t+1))|V(x(t)), e(t) = \{i,j\}] - V(x(t)) \leq \\ &\leq -\frac{1}{2}(x_i(t) - x_j(t))^2 + \frac{1}{2}. \end{aligned}$$

Hence we have that

$$\begin{aligned} &\mathbb{E}[V(x(t+1))|V(x(t))] \\ &\leq \sum_{\{i,j\}\in\mathcal{E}} P_{\{i,j\}} \left(V(x(t)) - \frac{1}{2}(x_i(t) - x_j(t))^2 + \frac{1}{2} \right) = \\ &= \frac{1}{2} - \frac{1}{2} \sum_{\{i,j\}\in\mathcal{E}} P_{\{i,j\}} (x_i(t) - x_j(t))^2 + V(x(t)). \end{aligned}$$

Notice now that

$$\sum_{\{i,j\}\in\mathcal{E}} P_{\{i,j\}} (x_i(t) - x_j(t))^2 = 2\,x^*(t)(\operatorname{diag}(P\mathbf{1}) - P)x(t)$$

and moreover that $\operatorname{diag}(P\mathbf{1}) - P \geq \lambda\Omega$ where λ is the smallest eigenvalue of $\operatorname{diag}(P\mathbf{1}) - P$ which is different from zero. From these facts we argue that

$$\begin{aligned} &\mathbb{E}[V(x(t+1))] = \mathbb{E}[\mathbb{E}[V(x(t+1))|V(x(t))]] \leq \\ &\mathbb{E}[\frac{1}{2} - \frac{1}{2}2\lambda x(t)^*\Omega x(t) + V(x(t))] = \\ &\mathbb{E}[(1-\lambda)V(x(t)) + \frac{1}{2}] = \\ &= (1-\lambda)\mathbb{E}[V(x(t))] + \frac{1}{2}. \end{aligned}$$

From the recurrence inequality

$$\mathbb{E}[V(x(t+1))] \leq (1-\lambda)\mathbb{E}[V(x(t))] + \frac{1}{2}$$

we can argue that

$$\mathbb{E}[V(x(t))] \leq (1-\lambda)^t \mathbb{E}[V(x(0))] + \frac{1 - (1-\lambda)^t}{2\lambda}$$

and then

$$\mathbb{E}[V(x(t))] \leq (1-\lambda)^t V(x(0)) + \frac{1}{2\lambda} \tag{8}$$

This shows that $\mathbb{E}[V(x(t))]$ tends initially to decrease exponentially fast at the rate $1 - \lambda$, and then to saturate to a constant $\frac{1}{2\lambda}$. It is important to remark that λ is a function of P, and then of the topology of the graph. This will be clear in the following examples.

Given a set A, we denote as $|A|$ its cardinality. We assume, for simplicity, that $P_{\{i,j\}} = \frac{1}{|\mathcal{E}|}$ for all $\{i,j\} \in \mathcal{E}$. This implies that

$$\text{diag}(P\mathbf{1}) - P = \frac{1}{|\mathcal{E}|} L_{\mathcal{G}},$$

where $L_{\mathcal{G}}$ is the Laplacian matrix[2] of $\mathcal{G}$. Let us consider two particular cases.

Example 1: Assume that $\mathcal{G}$ is the complete graph. Note that $|\mathcal{E}| = \frac{N(N-1)}{2}$. Then $P_{\{i,j\}} = \frac{2}{N(N-1)}$. In this case we have that

$$\text{diag}(P\mathbf{1}) - P = \frac{2}{N-1} I - \frac{2}{N(N-1)} \mathbf{1}\mathbf{1}^*,$$

from which it follows that $\lambda = \frac{2}{N-1}$.

Example 2: Assume that $\mathcal{G}$ is the ring graph. Note that $|\mathcal{E}| = N$. Assume that each edge is chosen with the same probability $1/N$. Hence it turns out that $\text{diag}(P\mathbf{1}) - P =$

$$= 2 \begin{pmatrix} \frac{1}{N} & -\frac{1}{2N} & 0 & 0 & \cdots & 0 & -\frac{1}{2N} \\ -\frac{1}{2N} & \frac{1}{N} & \frac{1}{2N} & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{2N} & \frac{1}{N} & \frac{1}{2N} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ -\frac{1}{2N} & 0 & 0 & 0 & \cdots & -\frac{1}{2N} & \frac{1}{N} \end{pmatrix}.$$

In this case, it is possible to prove, using results in [3], that $\lambda = \frac{2}{N}\left(1 - \cos\frac{2\pi}{N}\right)$, and then $\lambda = \frac{4\pi^2}{N^3} + o\left(1/N^3\right)$ for $N \to \infty$.

In these examples we see that, in the early steps, *the convergence rate is of the same order as the not quantized case* recalled above, but as time goes on the expected V saturates to a constant level, which also depends on N. Since $\frac{1}{2\lambda}$ increases drastically as N increases, for huge N this approach is less informative.

V. SIMULATIONS AND FINAL REMARKS

To illustrate and motivate our analytical results, we show some simulative results for the complete and ring graphs, and also for the random geometric graphs, which are a very common model in the analysis of wireless networks [8]. The random geometric graphs are constructed by randomly placing N nodes in the unit square, and joining them with edges whenever their distance is below a threshold $R = \Theta(\sqrt{\log N/N})$ for $N \to \infty$. In this case, it is known [2] that $T_\epsilon = \Theta(N^2/\log N)$ for $N \to \infty$.

As already mentioned, simulations point out the interesting question of the relationship between the gossip algorithm with or without quantization. How far are they apart in terms of performance, namely, asymptotic distance from the average consensus and speed of convergence?

We discussed the former issue in Section III, while the latter has been partially answered in Section IV. Any refinement would be of great interest: the bound (8) does not seem to be tight, since the dependence on N of the saturation level (see Figure 3) is not expected from the results of Section III, and does not appear in simulations. Moreover, we want to recall that an analysis developed for t going to infinity is not, at least in principle, appropriate for a system which is governed by a symbolic dynamics, converging in finite time. Our contribution has been to show that, as long as the distance from the agreement is much larger than the quantization step, the speed of convergence is almost the same as the not-quantized algorithm. Instead, when we are near to the consensus, the granularity effects have to come out, so that a full understanding of the algorithm has to be based on both the not quantized approximation and the analysis of some integer dynamics. Finally, the issue of generalizing the results to different values of the local averaging parameter α has been addressed in Section III-A: we believe this is a relevant point because it is well known [2], [5] that the 'best' value of α from the point of view of speed of convergence is not, in general, $1/2$.

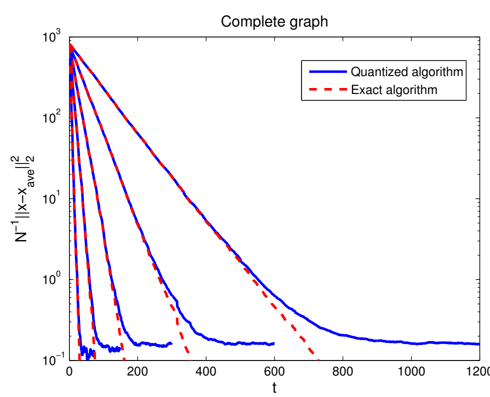

Fig. 1. Time evolution of the squared euclidean distance from agreement, with normalization, for the complete graph on $N = 5, 10, 20, 40, 80$ nodes, from left to rigth. Algorithm with quantized (solid lines) and with non-quantized communication (dashed). Average of 50 trials, from random initial conditions from a uniform distribution. Remark that the committed error is in facts smaller than predicted by theory.

REFERENCES

[1] Tuncer C. Aysal, Mark Coates, and Michael Rabbat. Distributed average consensus using probabilistic quantization. *Statistical Signal Processing, 2007. SSP '07. IEEE/SP 14th Workshop on*, pages 640–644, 26-29 Aug. 2007.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[3] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri. Communication constraints in the average consensus problem. *Automatica*, 44(3):671–684, March 2008.

[4] Ruggero Carli, Fabio Fagnani, Paolo Frasca, Thomas Taylor, and Sandro Zampieri. Average consensus on networks with transmission noise or quantization. In *ECC '07, Kos*, pages 1852–1857, 2007.

[5] F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. In *Proceedings Information Theory and Application Workshop*, 2007.

[6] Fabio Fagnani and Sandro Zampieri. Randomized consensus algorithms over large scale networks. *IEEE Journal on Selected Areas in Communications*, 26(4):634–649, May 2008.

[2]The Laplacian matrix of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is defined as $L_{\mathcal{G}} = A_{\mathcal{G}}\mathbf{1} - A_{\mathcal{G}}$, where $A_{\mathcal{G}}$ is the adjacency matrix of the graph. The latter is such that $(A_{\mathcal{G}})_{ij} = (A_{\mathcal{G}})_{ji} = 1$ if $\{i,j\} \in \mathcal{E}$ and 0 otherwise.

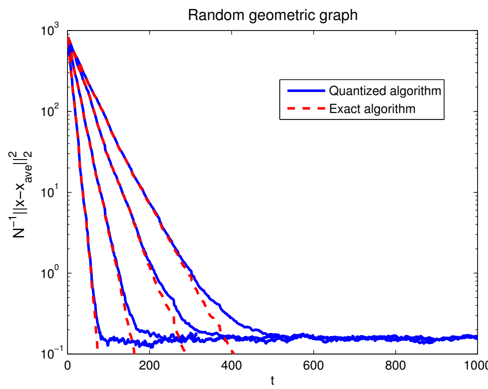

Fig. 2. Time evolution of the squared euclidean distance from agreement, with normalization, for the random geometric graph on $N = 10, 20, 30, 40$ nodes. Algorithm with quantized (solid lines) and with non-quantized communication (dashed). Average of 40 trials.

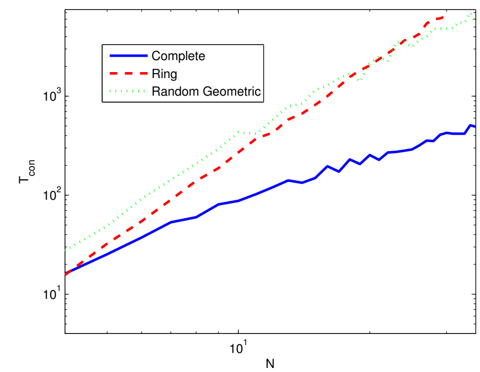

Fig. 4. Dependence of T_{con} on N in the examples: complete, ring, and random geometric. We plot the average of 40, 40, 100 trials respectively. Remark that the order of growth is approximately N, N^3, and N^2, respectively.

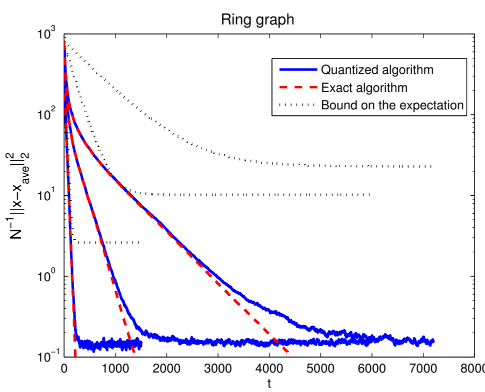

Fig. 3. Time evolution of the squared euclidean distance from agreement, with normalization, for the ring graph (Example 2) on $N = 10, 20, 30$ nodes. The algorithm performance, with quantized and with non-quantized communication, is compared with the bound (8). Average of 40 trials.

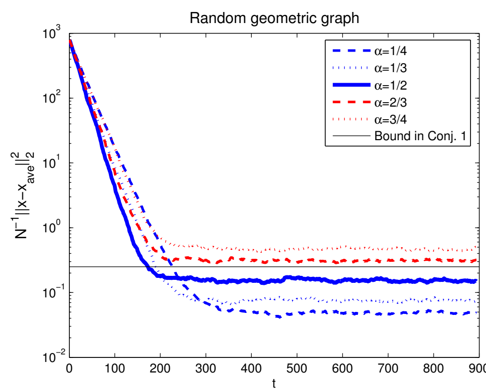

Fig. 5. Time evolution of the squared euclidean distance from agreement, with normalization, for the random geometric graph on $N = 20$ nodes, for different values of α (from bottom to top). Average of 40 trials. The evidence is supporting Conjecture 1.

[7] Paolo Frasca, Ruggero Carli, Fabio Fagnani, and Sandro Zampieri. Average consensus on networks with quantized communication. *International Journal of Robust and Nonlinear Control*, 2008. to appear.

[8] P. Gupta and P.R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, Mar 2000.

[9] Soummya Kar and Jose M. F. Moura. Distributed consensus algorithms in sensor networks: Quantized data. submitted, Dec 2007.

[10] A. Kashyap, T. Basar, and R Srikant. Quantized consensus. *Automatica*, 43:1192–1203, 2007.

[11] D.A Levin, Y. Peres, and E. L. Wilmer. Markov chains and mixing times. Available on line, 2006.

[12] Angelia Nedic, Alex Olshevsky, Asuman Ozdaglar, and John N. Tsitsiklis. On distributed averaging algorithms and quantization effects. Technical Report 2778, LIDS-MIT, Nov. 2007.

[13] 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(9):1520–1533, 2004.

[14] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *IPSN'05, Los Angeles, pp. 63-70*, 2005.

APPENDIX I
PROOF OF LEMMA 2

We first prove (3). Let $m = \lfloor x \rfloor$. So

$$am \leq ax < am + a\,.$$

Hence, we can find $s \in \mathbb{N}, 0 \leq s \leq a-1$ such that $am+s \leq ax < am+s+1$. This yields $\lfloor ax \rfloor = am+s$ and $\left\lfloor \frac{\lfloor ax \rfloor}{a} \right\rfloor = m$. We now prove equation (4). The equality $q[x] = \lfloor x+1/2 \rfloor$ is clear from the definition of $q[x]$. To prove the second equality, let $h = \lfloor 2x \rfloor$. Then $h \leq 2x < h+1$, from which follows that

$$\frac{h}{2} + \frac{1}{2} = \frac{h+1}{2} \leq x + 1/2 < \frac{h+2}{2} = \frac{h}{2} + 1.$$

From this inequality it follows that $\lfloor x+1/2 \rfloor = \lceil \frac{h}{2} \rceil$. This, with (3), implies (4).