Gossip consensus algorithms via quantized communication ⋆

Ruggero Carli [a], Fabio Fagnani [b], Paolo Frasca [b], Sandro Zampieri [c],

[a] *Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA 93106, USA*

[b] *Dipartimento di Matematica, Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129 Torino, Italy*

[c] *DEI, Università di Padova, Via Gradenigo 6/a, 35131 Padova, Italy*

Abstract

This paper considers the average consensus problem on a network of digital links, and proposes algorithms based on pairwise "gossip" communications and updates. Through both analytical results and simulations, we investigate two design questions, arising from the literature: whether the agents should use a deterministic or a probabilistic quantizer, and whether they should use, or not, exact information regarding their own states in the update.

Key words: Average consensus; quantization; Markov chain.

1 Introduction

In the latest years, algorithms to solve consensus problems have attracted a lot of interest. In a consensus problem a group of agents has to agree about a certain quantity, starting from different initial estimates. A special interest is devoted to average consensus, where the agents are requested to agree on the average of their initial estimates. Among the vast literature, we refer the reader to [14] [4] and references therein. The difficulty of the problem resides in the communication constraints which are given to the agents. Such communication constraints are usually represented by a graph: nodes are agents and edges are available communication links. Moreover, the communication across the links can be assumed to be perfect, or rather be digital and possibly subject to bandwidth constraints, interferences, erasures, packet losses, noise, delays. The constraint of quantization, due to the use of digital channels or to computing and memory constraints, has been recently investigated in several papers [18,3,8,9,1,7,12,19]. Among the many algorithms for consensus proposed in the literature, particularly interesting is the so called (symmetric) gossip algorithm: at every time instant a randomly chosen pair of agents communicate and average their states. Such algorithm, studied in detail in [2], has many appealing features: it reduces the number of communications with respect to deterministic algorithms and avoids data collision.

The goal of the present paper is to analyze the effects of quantization on the gossip algorithm: the agents states are assumed to be real numbers, while the sent message are integer numbers. We consider both a *deterministic* uniform quantizer and a *probabilistic* uniform quantizer, which are defined rigorously Section 2. To perform the states update, we introduce two alternative strategies, the *partially quantized* strategy and the *globally quantized* strategy, depending on whether the systems use exact information regarding their own state, or not, to update their states. We analyze these strategies, both with the deterministic quantizer, and with the probabilistic quantizer.

⋆ This paper was not presented at any IFAC meeting. Corresponding author P. Frasca Tel. +39-011-5647552.
Email addresses: `carlirug@engineering.ucsb.edu` (Ruggero Carli), `fabio.fagnani@polito.it.` (Fabio Fagnani), `paolo.frasca@polito.it` (Paolo Frasca), `zampi@unipd.it` (Sandro Zampieri).

Our results are obtained by two different techniques. For the partially quantized strategy with probabilistic quantization, we give a mean squared error analysis and convergence is proved for time going to infinity. In the three other cases, we study a Markov chain symbolic dynamics, obtaining results of convergence in finite time. Such a fact is remarkable, since it underlines the discrete nature of the problem, in spite of the state space being continuous. We show that the globally quantized strategy, both using the deterministic quantizer and the probabilistic quantizer, ensures that, almost surely, the consensus is reached in a finite time. The drawback of this strategy is that it does not preserve the average of the initial conditions. On the other hand, the partially quantized strategy preserves the initial average at each iteration of the algorithm, but does not guarantee that an exact consensus is reached. However, we prove that the states get as close to consensus as the size of quantization steps.

The paper is organized as follows. In Section 2 we formulate the problem. In particular we introduce the *partially* quantized strategy and the *globally* quantized strategy. In Section 3 and in Section 4, we analyze these two strategies assuming, respectively, that the systems quantize the information by means of deterministic quantizers and by means of probabilistic quantizers. Finally in Section 5 we gather out our conclusions.

2 Quantized gossip algorithms

For the sake of the clarity, we start by briefly reviewing the gossip consensus algorithm, where the systems communicate each other the exact value of their states; we follow the treatment in [2].
Assume we are given an undirected graph $\mathcal{G} = (V, \mathcal{E})$, $\mathcal{E} \subset \{(i,j) : i,j \in V\}$. At each time step, one edge (i,j) is randomly selected in $\mathcal{E}$ with probability $W^{(i,j)}$ such that $\sum_{(i,j)\in\mathcal{E}} W^{(i,j)} = 1$. Let W be the matrix with entries $W_{ij} = W^{(i,j)}$. The two agents connected by that edge average their states according to

$$\begin{aligned} x_i(t+1) &= \frac{1}{2}x_i(t) + \frac{1}{2}x_j(t) \\ x_j(t+1) &= \frac{1}{2}x_j(t) + \frac{1}{2}x_i(t) \end{aligned} \tag{1}$$

while

$$x_h(t+1) = x_h(t) \qquad \text{if } h \neq i,j. \tag{2}$$

Let $E_{ij} = (e_i - e_j)(e_i - e_j)^*$ and

$$P(t) = I - \frac{1}{2}E_{ij}$$

where $e_i = [0,\ldots,0,1,0,\ldots,0]^*$ is a $N \times 1$ unit vector with the i-th component equal to 1, then (1) and (2) can be written in a vector form as

$$x(t+1) = P(t)x(t) \tag{3}$$

where $x(t) = [x_1(t),\ldots,x_N(t)]^*$ denotes the state of the overall system. Note that $P(t)$ is a doubly stochastic matrix. It is well known [5,16] that, if the graph $\mathcal{G}$ is connected and each edge $(i,j) \in \mathcal{E}$ can be selected with a strictly positive probability $W^{(i,j)}$, then (3) reaches, almost surely, the *average consensus*, namely

$$\lim_{t\to\infty} x(t) = x_{ave}\mathbf{1},$$

where $x_{ave} = \frac{1}{N}\mathbf{1}^*x(0)$. In the sequel, we make the following assumption.

Assumption 1 *The graph $\mathcal{G} = (V,\mathcal{E})$ is a undirected connected graph and, at every time instant $t \geq 0$, each edge $(i,j) \in \mathcal{E}$ can be selected with a strictly positive probability $W^{(i,j)}$.*

Note that the algorithm (3) relies upon a crucial assumption: each agent transmits to its neighboring agents the precise value of its state. This implies the exchange of perfect information through the communication network. In this paper, we consider a more realistic case, i.e., we assume that the communication network is constituted of rate-constrained digital links. This prevents the agents from having a precise knowledge about the state of the other agents. In fact, through a digital channel, the i-th agent can only send to its neighbors symbolic data in a finite alphabet: using only this data, the neighbors of the i-th agent can build an estimate of the i-th agent's state. We denote this estimate by $\hat{x}_i(t)$, and let $\hat{x}(t) = [\hat{x}_1(t), \ldots, \hat{x}_N(t)]^*$. In this work, the estimate is simply the received symbol.

We proceed now by illustrating two types of quantizers which have been introduced in the literature in order to transmit information through a digital channel, for consensus purposes. In [3,7], the authors analyze the case in which

$$\hat{x}_i(t) = q_d(x_i(t)), \tag{4}$$

where, given a real number z, $q_d : \mathbb{R} \to \mathbb{Z}$ is the mapping sending z to its nearest integer, namely,

$$q_d(z) = n \in \mathbb{Z} \quad \Leftrightarrow \quad \begin{array}{ll} z \in [n - 1/2, n + 1/2[, & \text{if } z \geq 0 \\ z \in]n - 1/2, n + 1/2], & \text{if } z < 0. \end{array} \tag{5}$$

We refer to this quantizer as the *deterministic quantizer*. Instead in [1,17], the so-called *probabilistic quantizer* is introduced. This quantizer is defined as follows. Let $x \in \mathbb{R}$ and let $q_p(\cdot)$ denote the *probabilistic quantizer*. As for the *deterministic quantizer* above described, the set of quantization levels is the integer numbers, and $q_p(x)$ is the binary random variable defined as

$$q_p(x) = \begin{cases} \lfloor x \rfloor \text{ with probability } \lceil x \rceil - x \\ \lceil x \rceil \text{ with probability } x - \lfloor x \rfloor, \end{cases} \tag{6}$$

where we let $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor and ceiling operators from $\mathbb{R}$ to $\mathbb{Z}$. The following straightforward lemma states two important properties of the *probabilistic quantizer*.

Lemma 2.1 *Let $q_p(x)$ be a probabilistic quantization of $x \in \mathbb{R}$. Then $q_p(x)$ is an unbiased representation of x, i.e.,*

$$\mathbb{E}[q_p(x)] = x. \tag{7}$$

Moreover

$$\mathbb{E}\left[(x - q_p(x))^2\right] \leq \frac{1}{4}. \tag{8}$$

From now on, with a slight abuse of notation, given a vector $x \in \mathbb{R}^N$, we use the notation $q_d(x) \in \mathbb{R}^N$ (respectively $q_p(x) \in \mathbb{R}^N$) to denote the vector such that $q_d(x) = [q_d(x_1), \ldots, q_d(x_N)]^*$ (respectively $q_p(x) = [q_p(x_1), \ldots, q_p(x_N)]^*$).

In this paper, we introduce two updating rules of the state using quantized information. In the first strategy, if (i, j) is the edge selected at the t-th iteration, i and j, in order to update its state, use only the estimates of their states, as follows,

$$\begin{aligned} x_i(t+1) &= \frac{1}{2}\hat{x}_i(t) + \frac{1}{2}\hat{x}_j(t) \\ x_j(t+1) &= \frac{1}{2}\hat{x}_j(t) + \frac{1}{2}\hat{x}_i(t), \end{aligned} \tag{9}$$

or, equivalently in vector form, by recalling the definition of $P(t)$,

$$x(t+1) = P(t)\hat{x}(t). \tag{10}$$

To define the second strategy, we remark that (1) can be written as

$$\begin{aligned} x_i(t+1) &= x_i(t) - \frac{1}{2}x_i(t) + \frac{1}{2}x_j(t) \\ x_j(t+1) &= x_j(t) - \frac{1}{2}x_j(t) + \frac{1}{2}x_i(t). \end{aligned}$$

We then propose the following updating rule, where the agents use also perfect information regarding their own states,

$$x_i(t+1) = x_i(t) - \frac{1}{2}\hat{x}_i(t) + \frac{1}{2}\hat{x}_j(t)$$
$$x_j(t+1) = x_j(t) - \frac{1}{2}\hat{x}_j(t) + \frac{1}{2}\hat{x}_i(t), \tag{11}$$

or, equivalently in vector form,

$$x(t+1) = x(t) + (P(t) - I)\hat{x}(t). \tag{12}$$

We call the law (9) *globally quantized* and the law (11) *partially quantized.* It is easy to see that the *partially quantized* law (11), as the law (1), maintains the initial state average. Formally, defining $x_{ave}(t) = \frac{1}{N}\mathbf{1}^* x(t)$, we have that the *globally quantized* law (9) satisfies $x_{ave}(t) = x_{ave}(0)$, for all $t \geq 0$. Indeed, it is immediate to verify that $\mathbf{1}^* x(t+1) = \mathbf{1}^* x(t) + \mathbf{1}^*(P(t) - I)\hat{x}(t) = \mathbf{1}^* x(t)$, where the last equality follows from the fact that, since $P(t)$ is doubly stochastic for all $t \geq 0$, then $\mathbf{1}^*(P(t) - I) = 0$ for all $t \geq 0$.
We proceed with our analysis of these two rules by assuming first that $\hat{x}_i(t) = q_d(x_i(t))$, i.e., the information transmitted is quantized by means of deterministic quantizer, and then by assuming that $\hat{x}_i(t) = q_p(x_i(t))$, i.e., the information transmitted is quantized by means of probabilistic quantizer.

Remark 2.2 *In this paper we consider quantizers having quantization step equal to* 1*. More general quantizers, with quantization step a generic positive real number* ϵ*, can be obtained from* q_d *and* q_p *by defining* $q_d^{(\epsilon)}(x) = \epsilon q_d(x/\epsilon)$ *and* $q_p^{(\epsilon)}(x) = \epsilon q_p(x/\epsilon)$*. Hence, the general case can be simply recovered by a suitable scaling.*

3 Quantized gossip algorithms via deterministic quantizers

In this section we assume that the information exchanged between the agents is quantized by means of the deterministic quantizer q_d described in (5), namely $\hat{x}_i(t) = q_d(x_i(t))$. In this context, we separately analyze the partially and globally quantized strategies, starting from the first one.

3.1 Partially quantized strategy

Consider the partially quantized strategy

$$x_i(t+1) = x_i(t) - \frac{1}{2}q_d(x_i(t)) + \frac{1}{2}q_d(x_j(t))$$
$$x_j(t+1) = x_j(t) - \frac{1}{2}q_d(x_j(t)) + \frac{1}{2}q_d(x_i(t)). \tag{13}$$

Define

$$y(t) = \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^*\right)x(t) = x(t) - \frac{1}{N}\mathbf{1}\mathbf{1}^* x(0)\,, \tag{14}$$

and

$$d(t) = \frac{1}{\sqrt{N}}\|y(t)\|_2. \tag{15}$$

Such quantity represents the distance of the state $x(t)$ from the average of the states.

As an example we report in Figure 1 the result of simulations relative to a connected random geometric graph. Such graph has been drawn placing $N = 50$ nodes uniformly at random inside the unit square and connecting two nodes whenever the distance between them is less that $R = 0.3$. The initial condition $x_i(0)$ is randomly chosen inside the interval $[-100, 100]$ for all $1 \leq i \leq N$. Note that $d(t)$ does not converge to 0, meaning that the average consensus is not reached. However its value gets very close to 0, implying that the values of the state get very close to the initial average.

In the following we will give a general formal proof of this fact, quantifying the distance from consensus the states of the agents asymptotically achieve. This will be done exploiting a natural *symbolic dynamics* interpretation of the states dynamics and adapting to it the results presented in [9]. To start, we need the following technical lemma, whose proof can be found in [7].

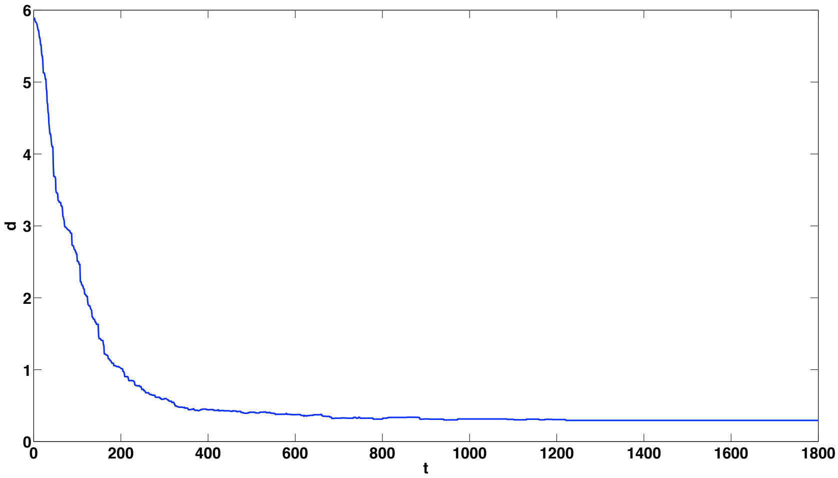

Fig. 1. Behavior of d for a connected random graph with $N = 50$ in case of deterministic quantizers and of partially quantized strategy.

Lemma 3.1 *Given $\alpha, \beta \in \mathbb{N}$ and $x \in \mathbb{R}$, it holds*

$$\lfloor x \rfloor = \left\lfloor \frac{\lfloor \alpha x \rfloor}{\alpha} \right\rfloor \tag{16}$$

$$q_d(x) = \lfloor x + 1/2 \rfloor = \left\lceil \frac{1}{2} \left\lfloor \frac{\lfloor 2\beta x \rfloor}{\beta} \right\rfloor \right\rceil. \tag{17}$$

Let us start our analysis. We define $n_i(t) = \lfloor 2x_i(t) \rfloor$ for all $i \in V$, and let $n(t) = [n_1(t), \dots, n_N(t)]^*$. Simple properties of floor and ceiling operators, together with the Lemma 3.1, allow us to remark that $q_d(x_i(t)) = \left\lceil \frac{n_i(t)}{2} \right\rceil$ and that

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

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

$$(n_i(t+1), n_j(t+1)) = g(n_i(t), n_j(t)) \tag{18}$$

where $g : \mathbb{Z} \times Z \to \mathbb{Z} \times \mathbb{Z}$ is

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

Notice 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 (18) then allows us to obtain information about the asymptotics of $x_i(t)$, since $n_i(t) = \lfloor 2x_i(t) \rfloor$. Before stating the main result regarding the convergence properties of (18), we define the following set

$$\mathcal{R} = \left\{ r \in \mathbb{Z}^N : \exists \alpha \in \mathbb{Z} \text{ s. t. } r - \alpha \mathbf{1} \in \{0, 1\}^N \right\}. \tag{19}$$

We have the following result.

Theorem 3.2 *Let $n(t)$ evolve according to (18). For every fixed initial condition $n(0) \in \mathbb{Z}$, almost surely there exists $T_{con} \in \mathbb{N}$ such that $n(t) \in \mathcal{R}$ for all $t \geq T_{con}$.*

Proof: The proof is based on verifying the following three facts:

(i) the set $\mathcal{R}$, defined in (19), is an invariant subset for the evolution described by (18);
(ii) $n(t)$ is a Markov process on a finite number of states;
(iii) there is a positive probability for $n(t)$ to reach a state in $\mathcal{R}$ in a finite number of steps.

Standard results on Markov chains [13] ensure that, if the above three facts yield true, the thesis is proven. Let us now check them in order.

(i) Let $h \in \mathbb{Z}$. Observe that

$$g(h, h+1) = \begin{cases} (h+1, h) & \text{if } h \text{ is even} \\ (h, h+1) & \text{if } h \text{ is odd} \end{cases}$$

This implies that $\mathcal{R}$ is an invariant subset for the dynamics described by (18).

(ii) Markovianity immediately follows from the fact that subsequent random choices of the edges are independent. We prove now that the states are finite. To this aim let $(h', k') = g(h, k)$. By the structure of g, it is easy to see that

$$\max\{h', k'\} \leq \max\{h, k\} \qquad \min\{h, k\} \leq \min\{h', k'\}.$$

Therefore we have that $m(n(t)) \geq m(n(0))$ and $M(n(t)) \leq M(n(0))$. This implies that the cardinality of the set of the states is upper bounded by $(M(n(0)) - m(n(0)) + 1)^N$.

(iii) First define

$$m(t) = \min_{1 \leq i \leq N} n_i(t) \tag{20}$$

$$M(t) = \max_{1 \leq i \leq N} n_i(t), \tag{21}$$

and,

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

The proof of (iii) is based on the following strong result about the monotonicity of $D(t)$: if $D(t) \geq 2$, then there exists $\tau \in \mathbb{N}$ such that

$$\mathbb{P}[D(t+\tau) < D(t)] > 0. \tag{22}$$

Now we prove (22).

Let $\mathcal{I}(t) = \{j \in V \text{ s.t. } n_j(t) = m(t)\}$. 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 of course 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 larger. 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 $(v_1, v_2, \ldots, v_{p-1}, v_p)$ would not be 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 (18) 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 (18), causes the value of both nodes to be strictly greater than 0. Therefore, this proves that $|\mathcal{I}(t+p-1)| < |\mathcal{I}(t)|$ 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 thus $|\mathcal{I}(t)|$ 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 (18) 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 (18), causes the value of both nodes to be strictly greater than 1. Hence $|\mathcal{I}(t+p-1)| < |\mathcal{I}(t)|$ 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 ensures the validity of (22).

The proof of the fact (iii) follows directly from (22). Indeed, let $\bar{n} \notin \mathcal{R}$, then, from a repeated application of (22) it follows that, there exists a path connecting $\bar{n}$ to a state $\bar{n}' = [\bar{n}'_1, \ldots, \bar{n}'_N]$, such that $\max\{\bar{n}'_1, \ldots, \bar{n}'_N\} - \min\{\bar{n}'_1, \ldots, \bar{n}'_N\} < 2$, that is, $\bar{n}' \in \mathcal{R}$.

This proves the thesis. We can go back to the original system, and prove the following result.

Corollary 3.3 *Consider the algorithm (13). Then, almost surely, there exists $T_{con} \in \mathbb{N}$ such that*

$$|x_i(t) - x_j(t)| \leq 1 \quad \forall\, i, j \quad \forall\, t \geq T_{con}, \tag{23}$$

and hence,

$$\|x(t) - x_{ave}\mathbf{1}\|_\infty \leq 1.$$

Proof: The proof is an immediate consequence of Theorem 3.2 and of the relation $n_i(t) = \lfloor 2x_i(t) \rfloor$, which assure that the states belong to two consecutive quantization bins.

Remark 3.4 *It is worth noting that Theorem 3.2 is an extension of Lemma 3 and Theorem 1 in [9]. In [9] the authors introduced a class of quantized gossip algorithms, satisfying the following assumptions. Let (i, j) be the edge selected at time t and let $n_i(t)$ and $n_j(t)$ the values at time t of node i and of node j respectively. 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 $D_{ij} = |n_i(t) - n_j(t)|$, 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* $n_j(t+1) = n_i(t)$. *Such update is called* swap.

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* $n_j(t+1) = n_j(t)$.

If we consider the class of algorithms satisfying (P1), (P2), (P3') or satisfying (P1), (P2), (P3"), it is possible to prove that Lemma 3 and Theorem 1 stated in [9] hold true also for this class. The proofs are analogous to that of Theorem 3.2 provided in this paper. Moreover it is easy to see that the algorithm (18) satisfies the properties (P1), (P2), (P3'). This represents an alternative way to prove Theorem 3.2.

Remark 3.5 *Another feature of the algorithms proposed in [9] is that the state of each node is always an integer. On one hand this represents a clear advantage from a computational point of view. On the other hand, the nodes, at the initial step, have to quantize the initial conditions that could be any arbitrary real number. In general the average of the quantized states will be different from the average of the initial states, thus introducing an error that will be propagated along the iterations of the algorithm. This fact is illustrated in Figure 2 where we provide a comparison between the partially quantized strategy via deterministic quantizers and the algorithm proposed in [9], that we denote by KBS. Precisely, we plotted the behavior of $d(t)$ for both strategies on the same graph considered in Figure 1.*

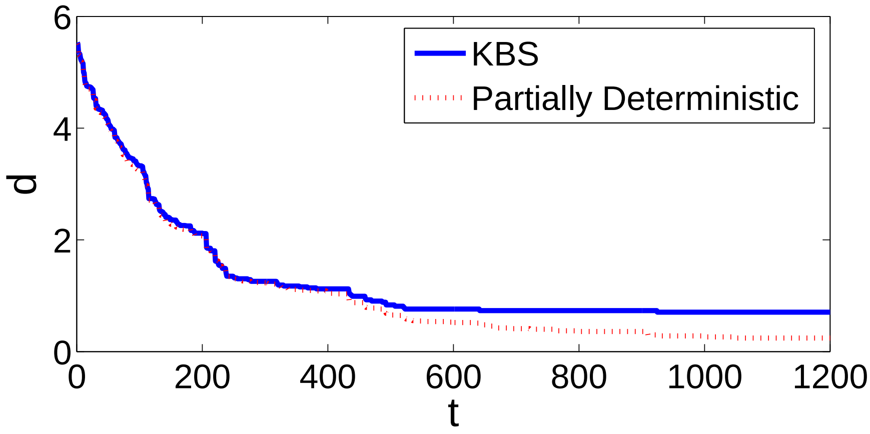

Fig. 2. Plot of $d(t)$, as in (15), for the partially quantized strategy with deterministic quantization, and for the algorithm *KBS* proposed in [9].

3.2 Globally quantized strategy

In this subsection we consider the globally quantized strategy

$$\begin{aligned} x_i(t+1) &= \frac{1}{2}q_d(x_i(t)) + \frac{1}{2}q_d(x_j(t)) \\ x_j(t+1) &= \frac{1}{2}q_d(x_j(t)) + \frac{1}{2}q_d(x_i(t)). \end{aligned} \tag{24}$$

We underline immediately that the fact that (24) uses only quantized information and not perfect information combined with quantized information as in (13) makes the analysis of (13) slightly easier than the analysis of (24). Remarkably, we show in this subsection that the law (13) drives, almost surely, the systems to exact consensus at an integer value. Unfortunately, the initial average of states is not preserved in general. Again, the analysis of this algorithm can be performed efficiently by means of the *symbolic dynamics* previously introduced.
Let again $n_i(t) = \lfloor 2x_i(t) \rfloor$ for all $i \in V$. From (24) and the fact that $q_d(x_i(t)) = \left\lceil \frac{n_i(t)}{2} \right\rceil$ we obtain

$$(n_i(t+1), n_j(t+1)) = (g_1(n_i(t), n_j(t)), g_1(n_i(t), n_j(t))) \tag{25}$$

where $g_1 : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ is defined as

$$g_1(h, k) = \left\lceil \frac{h}{2} \right\rceil + \left\lceil \frac{k}{2} \right\rceil .$$

Define

$$\mathcal{A} = \{y \in \mathbb{Z}^N : \exists \alpha \in \mathbb{Z} \text{ such that } y = 2\alpha \mathbf{1}\}. \tag{26}$$

We have the following result.

Theorem 3.6 *Let $n(t)$ evolve according to (25). For every fixed initial condition $n(0)$, almost surely there exists $T_{con} \in \mathbb{N}$ such that $n(t) \in \mathcal{A}$ for all $t \geq T_{con}$.*

Proof: As for the proof of Theorem 3.2, it will be sufficient to verify the following three facts:

(i) each element in the set $\mathcal{A}$ is invariant for the evolution described by (25);
(ii) $n(t)$ is a Markov process on a finite number of states;
(iii) there is a positive probability for $n(t)$ to reach a state in $\mathcal{A}$ in a finite number of steps.

Let us now check them in order.

(i) is trivial.

(ii) Markovianity immediately follows from the fact that subsequent random choices of the edges are independent. To prove that the states are finite, define $m(t)$ and $M(t)$ as in (20) and (21). Let $p, q \in \mathbb{Z}$ with $p \leq q$. Then, from the structure of g_1 we have that $p \leq g_1(p,q) \leq q + r_q$ where r_q denotes the remainder in the Euclidean division of q by 2. It follows that

$$m(0) \leq n_i(t) \leq M(0) + r_{M(0)} \ \forall i \in V \ \ \forall t \geq 0\,. \tag{27}$$

This yields (ii).

(iii) Let us fix $t = t_0$, and assume that $n(t_0) \notin \mathcal{A}$. We prove that there exists $\tau \in \mathbb{N}$ such that $\mathbb{P}[n(t_0+\tau) \in \mathcal{A}] > 0$. We start by observing that, from the assumption of having a connected graph, there exists $(h,k) \in \mathcal{E}$ such that $n_h(t_0) = m(t_0)$, $n_k(t_0) = q$ and $g(m(t_0), q) > m(t_0)$. Indeed, two cases are given when $n(t_0) \notin \mathcal{A}$.
- If $m(t_0) < M(t_0)$, then it suffices to consider an edge (h,k) such that $n_h(t_0) = m(t_0)$ and $n_k(t_0) = q > m(t_0)$, which gives $g_1(m(t_0), p) > m(t_0)$. Note that such an edge exists from the hypothesis of having a connected graph;
- if $m(t_0) = M(t_0)$, necessarily we have that $m(t_0)$ and $M(t_0)$ are odd; then $g(m(t_0), m(t_0)) > m(t_0)$.

We define now $\mathcal{I}_a(t) = \{i \in \mathcal{V} : n_i(t) = a\}$. The above discussion implies that $|\mathcal{I}_{m(t_0)}(t_0+1)| < |\mathcal{I}_{m(t_0)}(t_0)|$ with the positive probability of choosing the edge (h,k) and hence that there is also a positive probability that at some finite time $t' > t_0$, $|\mathcal{I}_{m(t_0)}(t')| = 0$, that is $m(t') > m(t_0)$. Iterating this argument and recalling that (see (27)) $M(t) \leq M(t_0) + r_{M(t_0)}$ for all $t \geq t_0$, it follows that there exists $\tau \in \mathbb{N}$ such that $\mathbb{P}[n(t_0+\tau) \in \mathcal{A}] > 0$.

This proves the thesis. We can now go back to the original system. The following corollary follows immediately from the definition of $n(t)$.

Corollary 3.7 *Let $x(t)$ evolve according to (24). Then almost surely there exists $T_{con} \in \mathbb{N}$ and $\alpha \in \mathbb{Z}$ such that $x_i(t) = \alpha$ for all $i \in \mathcal{V}$ and for all $t \geq T_{con}$.*

We have already underlined the fact that this strategy does not preserve the initial average, in general. Providing some probabilistic estimation of the distance of the consensus point from the initial average is a challenging problem: we limits our analysis to the following simulation. In Figure 3 we plot the variable z that is defined as follows. In the globally quantized strategy we have that, almost surely $\lim_{t\to\infty} = \alpha\mathbf{1}$ for some random integer α. Let $z = |\alpha - 1/N\mathbf{1}^* x(0)|$. In words, z represents the distance from the consensus point to which the globally quantized strategy leads the systems and the average of the initial condition. We have depicted the value of z for a family of random geometric graphs [15] of increasing size from $N = 10$ up to $N = 80$. The initial condition $x_i(0)$ is chosen randomly inside the interval $[-100, 100]$ for all $1 \leq i \leq N$. Moreover for each N, z is computed as the mean of 100

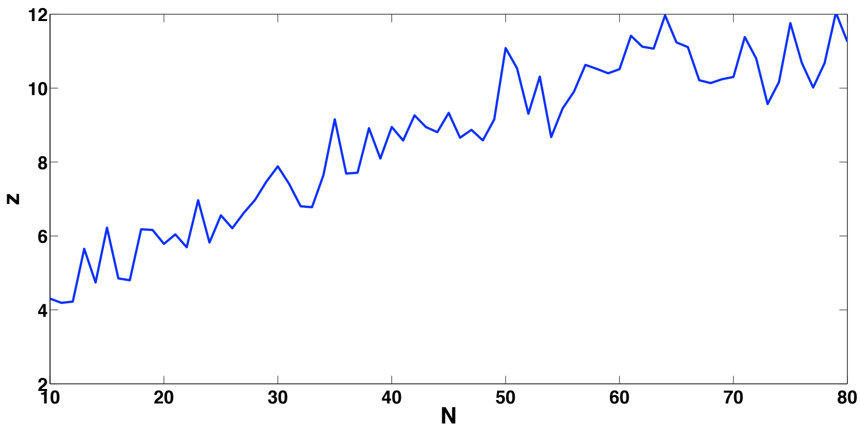

Fig. 3. Behavior of z for a family of random geometric graphs in case of deterministic quantizers and of globally quantized strategy.

trials. We can see that the value of z is increasing in N and assumes values that are not negligible with respect to the quantization step size.

3.3 Speed of convergence

Providing insights on the speed of convergence of (13) and of (24) is quite hard in general. In Figure 4 and Figure 5 we report, respectively, a comparison between the partially quantized strategy (13) and the gossip algorithm with exchange of perfect information (1) and between the globally quantized strategy (24) and again the gossip algorithm with exchange of perfect information (1). The simulations are made on the same random geometric graphs considered in Figure 1, and the initial conditions are randomly chosen inside the interval $[-100, 100]$.

For both strategies we plotted the behavior of the variable $d(t)$ defined in (15).

From the Figure 4 and Figure 5 we can infer that the speed of convergence toward the steady state of the quantized strategies (24) and (13) is similar to the one of the gossip algorithm with perfect exchange of information. This numerical evidence is not completely understood yet, but some interesting preliminary results appear in [6].

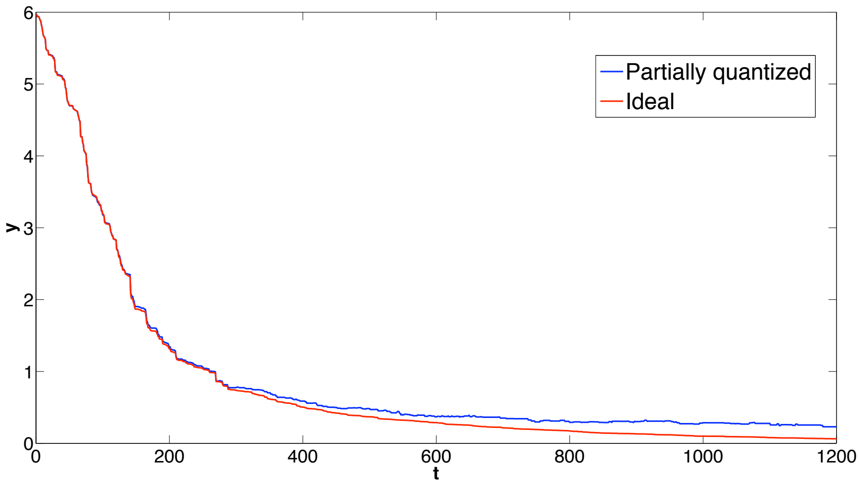

Fig. 4. Behavior of d, when using the *partially quantized* strategy, for a connected random geometric graph with $N = 50$. Note that since the *partially quantized* strategy does not converge to a consensus, $d(t)$ does not go to 0.

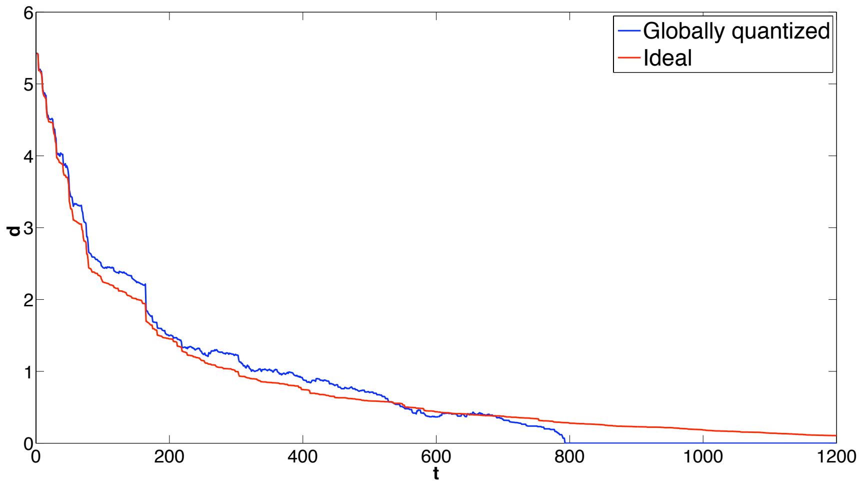

Fig. 5. Behavior of d, when using the *globally quantized* strategy, for a connected random geometric graph with $N = 50$. In this case, accordingly to the theoretical result stated in Corollary 3.7, $d(t)$ tends to 0.

Remark 3.8 *If, depending on the application, one can not relax the convergence requirement, we could suggest the* *following heuristic solution to the consensus problem, which combines the positive features of both strategies,*

$$x(t+1) = Pq_d(x(t)) + \epsilon(t)(x(t) - q_d(x(t))),$$

where $\epsilon(t),\ t \geq 0$, *is a nonnegative sequence such that* $\epsilon(t) \leq 1,\ \forall\, t \geq 0$ *and* $\lim_{t\to\infty} \epsilon(t) = 0$.

4 Quantized gossip algorithms via probabilistic quantizers

In this section we assume that the information exchanged between the systems is quantized by means of the probabilistic quantizer q_p described in (6), namely $\hat{x}_i(t) = q_p(x_i(t))$. We recall the statistics of q_p, as illustrated in Lemma 2.1. Moreover, we make the following natural assumption

Assumption 2 *Given the values* $x_i(t)$ *for all* $i \in V$, *the random variables* $q_p(x_i(t))$, *as* i *varies, form an independent set. Moreover, for every* $i \neq j$, *given* $x_i(t)$, $q_p(x_i(t))$ *is independent from* $x_j(t)$.

As before, we will now separately analyze the partially and globally quantized strategies.

4.1 Partially quantized strategy

The algorithm for partially quantized strategy, when the edge (i, j) is chosen, can be written as

$$\begin{aligned} x_i(t+1) &= x_i(t) - \frac{1}{2} q_p(x_i(t)) + \frac{1}{2} q_p(x_j(t)) \\ x_j(t+1) &= x_j(t) - \frac{1}{2} q_p(x_j(t)) + \frac{1}{2} q_p(x_i(t)). \end{aligned} \tag{28}$$

Similarly to the partially quantized strategy via deterministic quantizers (13), also (28) does not reach the consensus in general. Again we report a simulation showing this fact. In Figure 6 the behavior of the quantity $d(t)$, defined in (15), is depicted for the same connected random geometric graph considered in Figure 1. Note that the quantity

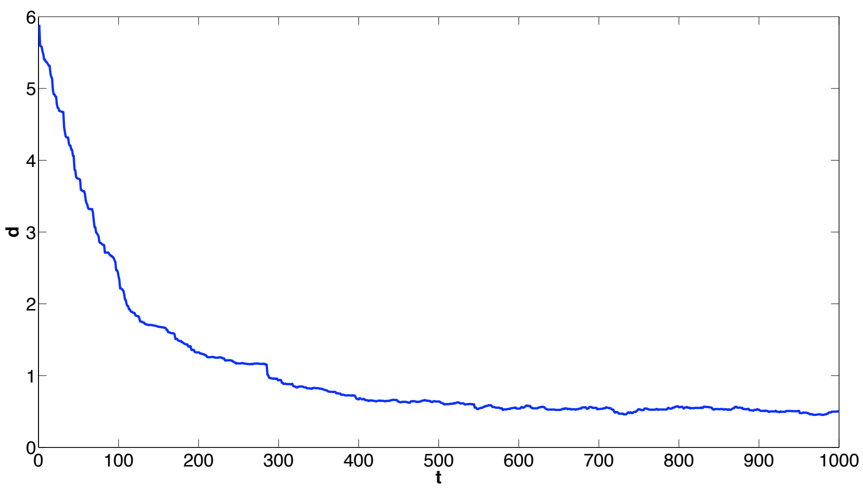

Fig. 6. Behavior of d for a connected random geometric graph with $N = 50$.

$d(t)$ stays visibly away from 0, meaning that the average consensus is not reached.

The analysis of (28) is more complicate than for the corresponding law (13). This is mainly due to the lack of convexity properties which were used in the analysis of (13). The following example shows this type of difficulty.

Example 4.1 Consider (13) and assume that the edge (i, j) has been selected at time t. Without loss of generality assume that $x_i(t) \leq x_j(t)$. Then, by convexity arguments, we have that $\lfloor x_i(t) \rfloor \leq x_i(t+1), x_j(t+1) \leq \lceil x_j(t) \rceil$. This is no longer true for (28). As a numerical example assume that $x_i(t) = 3.4$ and $x_j(t) = 3.6$. Then with probability $1/4$ we will have that $q_p(x_i(t)) = 4$ and $q_p(x_j(t)) = 3$. In this case, by (28), we have that $x_i(t+1) = 2.9$ and that $x_j(t+1) = 4.1$. Hence, $x_i(t+1), x_j(t+1)$ do not belong to the interval $[\lfloor x_i(t) \rfloor, \lceil x_j(t) \rceil]$.

For this reason, we do not develop a symbolic analysis for this algorithm, and we do not prove convergence in finite time. By simulations we can see that (28) does not drive the states of the systems inside the same bin of quantization, as the corresponding strategy (13) using deterministic quantizers. In Figure 7, we depict the behavior of the quantity

$$s(t) = \max_{1 \leq i,j \leq N} |x_i(t) - x_j(t)|.$$

for the same random geometric graph considered in Figure 6. In this simulation we assume that the initial condition $x_i(0)$ is randomly chosen inside the interval $[-10, 10]$. Note that s asymptotically oscillates around 2. Interesting

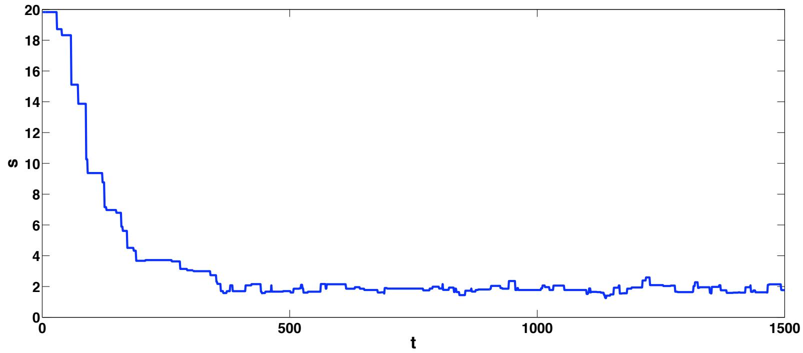

Fig. 7. Behavior of s for a connected random geometric graph with $N = 50$.

results on (28), in terms of both the asymptotic distance from the initial average and the speed of convergence, can be provided by a mean-square analysis. In the sequel of this subsection, we assume that the initial condition $x(0)$ satisfies the following condition.

Assumption 3 *The initial condition $x(0)$ is a random variable such that $\mathbb{E}[x(0)] = 0$ and $\mathbb{E}[x(0)x^*(0)] = \sigma_0^2 I$ for some $\sigma_0^2 > 0$.*

We start by observing that (28) can be rewritten as

$$x(t+1) = P(t)x(t) + (P(t) - I)(q_p(x(t)) - x(t)) \tag{29}$$

Define

$$e(t) = q_p(x(t)) - x(t),$$

the quantization error and recall the definition of $y(t)$ given in (14). From (29), using the fact that $P(t)$ is symmetric and stochastic, we easily obtain the following recursive relation in terms of the variables $e(t)$ and $y(t)$

$$y(t+1) = P(t)y(t) + (P(t) - I)e(t). \tag{30}$$

In order to perform an asymptotic analysis of (30) it is convenient to introduce the following matrices. Let

$$\Sigma_{yy}(t) = \mathbb{E}[y(t)y^*(t)], \quad \Sigma_{ee}(t) = \mathbb{E}[e(t)e(t)^*], \quad \Sigma_{ye}(t) = \mathbb{E}[y(t)e(t)^*].$$

Equation (30) leads to the following recursive equation in terms of the above matrices

$$\begin{aligned} \Sigma_{yy}(t+1) = {} & \mathbb{E}[P(t)\Sigma_{yy}(t)P(t)] + \mathbb{E}[P(t)\Sigma_{ye}(t)(P(t) - I)] + \\ & + \mathbb{E}\left[(P(t) - I)\Sigma_{ye}^* P(t)\right] + (P(t) - I)\Sigma_{ee}(t)(P(t) - I). \end{aligned} \tag{31}$$

From the fact that $x(0)$ is a random variable satisfying Assumption 3, it immediately follows that

$$\Sigma_{yy}(0) = \sigma_0^2 \left(I - N^{-1}\, \mathbf{1}\mathbf{1}^* \right). \tag{32}$$

The following proposition states some correlation properties of the variables y and e.

Proposition 4.2 *Consider the variables $y(t)$ and $e(t)$ above defined. Then*

$$\mathbb{E}\left[e(t)\right] = 0 \qquad \textit{and} \qquad \Sigma_{ee}(t) = \text{diag}\left\{\sigma_1^2(t), \ldots, \sigma_N^2(t)\right\} \tag{33}$$

where the right-hand-side is a diagonal matrix whose diagonal elements $\sigma_i^2(t) = \mathbb{E}\left[e_i^2(t)\right]$ are such that $\sigma_i^2(t) \le 1/4$ for all $1 \le i \le N$ and for all $t \ge 0$.
Moreover

$$\Sigma_{ye}(t) = 0, \tag{34}$$

for all $t \ge 0$.

Proof: Using Lemma 2.1, we have that

$$\begin{aligned}
\mathbb{E}\left[e_i(t)\right] &= \mathbb{E}\left[\,\mathbb{E}\left[q_p(x_i(t)) - x_i(t) | x_i(t)\right]\right] \\
&= \mathbb{E}\left[\,\mathbb{E}\left[q_p(x_i(t)) | x_i(t)\right] - x_i(t)\right] \\
&= \mathbb{E}\left[x_i(t) - x_i(t)\right] \\
&= 0.
\end{aligned} \tag{35}$$

Moreover, for $i \neq j$, using Assumption 2,

$$\begin{aligned}
\mathbb{E}\left[e_i(t)e_j(t)\right] &= \mathbb{E}\left[\mathbb{E}\left[e_i(t)e_j(t) \,|\, x_i(t),\, x_j(t)\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[e_i(t) \,|\, x_i(t),\, x_j(t)\right] \mathbb{E}\left[e_j(t) \,|\, x_i(t),\, x_j(t)\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[e_i(t) \,|\, x_i(t)\right] \mathbb{E}\left[e_j(t) \,|\, x_j(t)\right]\right] \\
&= 0
\end{aligned} \tag{36}$$

If $i = j$, using again Lemma 2.1, we have that

$$\begin{aligned}
\mathbb{E}\left[e_i^2(t)\right] &= \mathbb{E}\left[\left(q_p(x_i(t)) - x_i(t)\right)^2\right] \\
&= \mathbb{E}\left[\,\mathbb{E}\left[\left(q_p(x_i(t)) - x_i(t)\right)^2 | x_i(t)\right]\right] \\
&\le \mathbb{E}\left[\frac{1}{4}\right] \\
&= \frac{1}{4}
\end{aligned} \tag{37}$$

An argument similar than the one above used to prove that $\mathbb{E}\left[e_i(t)e_j(t)\right] = 0$ allows to prove that $\mathbb{E}\left[x_i(t)e_j(t)\right] = 0$ for any $i \neq j$. This easily yields (34).

From the above properties we have that (31) can be rewritten as

$$\Sigma_{yy}(t+1) = \mathbb{E}\left[P(t)\Sigma_{yy}(t)P(t)\right] + \mathbb{E}\left[\left(P(t) - I\right)\Sigma_{ee}(t)\left(P(t) - I\right)\right]. \tag{38}$$

To estimate the asymptotic distance from the initial average, we introduce the cost function

$$J(W) = \limsup_{t \to \infty} \sqrt{\frac{1}{N}\mathbb{E}[\|y(t)|^2]}. \tag{39}$$

The cost depends on the selection probabilities W, and, thanks to the above definitions, can be computed as

$$J(W) = \limsup_{t \to \infty} \sqrt{\frac{1}{N}\text{tr } \{\Sigma_{yy}(t)\}}. \tag{40}$$

We can rewrite the evolution law (38) as

$$\Sigma_{yy}(t+1) = \mathcal{A}(\Sigma_{yy}(t)) + \mathcal{B}(\Sigma_{ee}(t)),$$

where $\mathcal{A}$ and $\mathcal{B}$ are linear operators from $\mathbb{R}^{N\times N}$ to itself. Namely, given a matrix M, $\mathcal{A}(M) = \mathbb{E}\left[P(t)MP(t)\right]$ and $\mathcal{B}(M) = \mathbb{E}\left[(P(t)-I)\,M\,(P(t)-I)\right].$

It is useful to remark that $\mathcal{A}$ is actually the evolution on Σ_{yy} for the gossip algorithm [2], in the absence of quantization error, while $\mathcal{B}$ can be regarded as a disturbance due to the quantization error. From [5], we know that in the case of no quantization the system converges almost surely to consensus. This implies that $\mathcal{A}$ is an asymptotically stable operator when restricted to the subspace $\mathcal{S} = \{M \in \mathbb{R}^{N\times N} : \mathbf{1}^*M\mathbf{1} = 0\}$. Since $\mathbf{1}^*\mathcal{B}(M)\mathbf{1} = 0$ for any matrix M and $\Sigma_{yy}(0) \in \mathcal{S}$, we have that $\Sigma_{yy}(t) \in \mathcal{S}$ for all $t \geq 0$. As a consequence $\Sigma_{yy}(t)$ converges for $t \to +\infty$ and in the definition of $J(W)$ in (40) $\limsup$ can be replaced by $\lim$.

It is actually a general fact that, in most cases, systems with quantization can be regarded as disturbed versions of non-quantized systems. This approach has been taken in [6], to show that the speed of convergence of gossip consensus algorithms is essentially the same with or without quantization, as long as the states are far from consensus.

Providing an expression for $J(W)$ is quite hard in general. We then try to simplify the problem by introducing the following auxiliary system

$$\bar{\Sigma}(t+1) = \mathbb{E}\left[P(t)\bar{\Sigma}(t)P(t)\right] + \frac{1}{4}\mathbb{E}\left[(P(t)-I)^2\right], \qquad (41)$$

where $\bar{\Sigma}(0) = \Sigma_{yy}(0)$, and the following cost function

$$\bar{J} = \limsup_{t\to\infty} \sqrt{\frac{1}{N}\mathrm{tr}\ \{\bar{\Sigma}(t)\}}.$$

In principle, $\bar{J}$ should depend on W, too. However, we are going to prove that this is not the case. We have the following comparison result.

Proposition 4.3 *Consider the cost functions $J(W)$ and $\bar{J}$. We have that*

$$J(W) \leq \bar{J}.$$

Proof: To prove the statement we show, by induction on t, that $\bar{\Sigma}(t) \geq \Sigma_{yy}(t)$ for all $t \geq 0$, where the inequality is meant in matricial sense, that is, $\bar{\Sigma}(t) - \Sigma_{yy}(t)$ is a semidefinite positive matrix.

Since $\bar{\Sigma}(0) = \Sigma_{yy}(0)$ the assertion is true for $t = 0$. Assume now that $\bar{\Sigma}(t) \geq \Sigma_{yy}(t)$ is true for a generic t. We have that

$$\begin{aligned}\bar{\Sigma}(t+1) - \Sigma_{yy}(t+1) &= \mathbb{E}\left[P(t)\bar{\Sigma}(t)P(t)\right] + \frac{1}{4}\mathbb{E}\left[(P(t)-I)^2\right] \\ &\quad - \left(\mathbb{E}\left[P(t)\Sigma_{yy}(t)P(t)\right] + \mathbb{E}\left[(P(t)-I)\,\Sigma_{ee}(t)\,(P(t)-I)\right]\right) \\ &= \mathbb{E}\left[P(t)(\bar{\Sigma}(t) - \Sigma_{yy}(t))P(t)\right] + \mathbb{E}\left[(P(t)-I)\left(\frac{1}{4}I - \Sigma_{ee}(t)\right)(P(t)-I)\right].\end{aligned}$$

Since by inductive hypothesis $\bar{\Sigma}(t) \geq \Sigma_{yy}(t)$ and since by Proposition 4.2 we know that $\Sigma_{ee}(t) \leq \frac{1}{4}I$ for all $t \geq 0$, we have that $\bar{\Sigma}(t+1) - \Sigma_{yy}(t+1) \geq 0$.

Observe now that, since $P(t)^2 = P(t)$ we obtain that $\mathbb{E}[(I - P(t))^2] = I - \mathbb{E}[P(t)]$. From this fact we obtain the following result.

Proposition 4.4 *Given the above definitions and (38),*

$$\lim_{t\to\infty} \bar{\Sigma}(t) = \frac{1}{4}\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^*\right).$$

Proof: Define the matrix $\bar{\mathcal{B}} = \mathbb{E}[(I - P(t))^2]$ Since $\bar{\Sigma}_{yy}(0) \in S$, and $\mathcal{A}$ is asymptotically stable if restricted to the subspace $\mathcal{S}$, then

$$\lim_{t\to\infty} \bar{\Sigma}(t) = \sum_{t=0}^{+\infty} \mathcal{A}^{(t)}(\bar{\mathcal{B}}).$$

This is the only fixed point of the iteration law (41). Thus we are left to prove that $\Sigma^* = \frac{1}{4}\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^*\right)$ is a fixed point, that is $\Sigma^* = \mathcal{A}(\Sigma^*) + \bar{\mathcal{B}}$. This is true, because

$$\begin{aligned}\mathcal{A}(\Sigma^*) + \bar{\mathcal{B}} &= \frac{1}{4}\mathbb{E}\left[P(t)\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^*\right)P(t)\right] + \frac{1}{4}(I - \mathbb{E}[P(t)]) \\ &= \frac{1}{4}\left\{\mathbb{E}\left[P(t)^2\right] - \frac{1}{N}\mathbf{1}\mathbf{1}^* + I - \mathbb{E}[P(t)]\right\} = \frac{1}{4}\left\{I - \frac{1}{N}\mathbf{1}\mathbf{1}^*\right\}.\end{aligned}$$

Corollary 4.5 *For all probability matrix W we have that $J(W) \leq \frac{1}{2}$.*

Proof: From the above proposition we can argue that $\bar{J} = \frac{1}{2}\sqrt{\frac{N-1}{N}}$, and since $J(W) \leq \bar{J}$, we can conclude.

From these theorems we draw a strong conclusion about the convergence of the algorithm. In spite of missing consensus in the strict sense, the asymptotical mean squared error of the algorithm is smaller than the size of the quantization bin, and has a bound which does not depend on the number of the agents, nor on the topology of the graph, nor on the probability of the edges selection.

4.2 Globally quantized strategy

The algorithm for the globally quantized strategy, when the edge (i, j) is chosen, can be written as

$$\begin{aligned} x_i(t+1) &= \frac{1}{2}q_p(x_i(t)) + \frac{1}{2}q_p(x_j(t)) \\ x_j(t+1) &= \frac{1}{2}q_p(x_j(t)) + \frac{1}{2}q_p(x_i(t)). \end{aligned} \tag{42}$$

Below we prove that the law (42), as the law (24), drives almost surely the systems to exact consensus at an integer value. Moreover, we show by simulations, that the consensus point, even if (42) does not preserve the average of the state, is rather close to the average of the initial condition. This represents a significant improvement with respect to the strategy (24), that, as seen in Figure 3, leads to a consensus point whose distance from the average of the initial condition, is not negligible in general.

With the globally quantized strategy (42), like with (28), we have to deal with two sorts of randomness, since the interacting pair is randomly selected, and the quantization map is itself random. This makes the analysis of (42) more complicate than the analysis of (24). However, again, we are able to prove the convergence by a *symbolic dynamics* approach.

Let again $n_i(t) = \lfloor 2x_i(t) \rfloor$ for all $i \in V$ and let $n(t) = [n_1(t), \dots, n_N(t)]^*$. Before finding a recursive equation for $n(t)$, we need to introduce the following random variable. Let

$$T_{all} = \inf\{t : \text{at time t every node in } V \text{ has been selected at least once}\}$$

T_{all} is an integer random variable which is almost surely finite, because nodes are selected with positive probability. Note that, from (42), $x_i(t) \in \{a, a + 1/2\}$ for some integer number a, for all $t \geq T_{all}$. This allows us to disregard the evolution before T_{all} and to analyze, for $t > T_{all}$, the symbolic dynamics as follows. For $t \geq T_{all}$, by recalling how the probabilistic quantizer works, we have that

$$q_p(x_i(t)) = \begin{cases} \frac{n_i(t)}{2} & \text{if } n_i(t) \text{ is even} \\ \lceil \frac{n_i(t)}{2} \rceil \text{ with probability } 1/2 & \text{if } n_i(t) \text{ is odd} \\ \lfloor \frac{n_i(t)}{2} \rfloor \text{ with probability } 1/2 & \end{cases}$$

Let ξ_1 and ξ_2 be two independent Bernoulli random variables with parameter 1/2 and define $g_2 : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ by

$$g_2(h,k) = \left\lceil \frac{h}{2} \right\rceil + \left\lceil \frac{k}{2} \right\rceil - \xi_1 r_h - \xi_2 r_k,$$

where r_h denotes the remainder of the division of h by 2. If, at time instant t, the edge (i,j) is selected, then

$$(n_i(t+1), n_j(t+1)) = (g_2(n_i(t), n_j(t)), g_2(n_i(t), n_j(t))) \,. \tag{43}$$

The following result characterizes the convergence properties of (43). Recall the definition of the set $\mathcal{A}$ in (26).

Theorem 4.6 *Let $n(t)$ evolve according to (43). For every fixed initial condition $n(0)$, almost surely there exists $T_{con} \in \mathbb{N}$ such that $n(t) \in \mathcal{A}$ for all $t \geq T_{con}$.*

Proof: The proof is quite similar to the proof of Theorem 3.6 and Theorem 3.2, and it is based on proving the following three facts:

(i) each element in the set $\mathcal{A}$ is invariant for the evolution described by (43);
(ii) $n(t)$ is a Markov process on a finite number of states;
(iii) there is a positive probability for $n(t)$ to reach a state in $\mathcal{A}$ in a finite number of steps.

Let us now check them in order.

(i) is trivial.
(ii) Markovianity immediately follows from the fact that subsequent random choices of the edges are independent and from (43). To prove that the states are finite, define $m(t)$ and $M(t)$ as in (20) and (21). Let $h \in \mathbb{Z}$. Then, from the structure of g_2 we have that
   - $g_2(h,h) = h$ if h is even;
   - $h - 1 \leq g_2(h,h) \leq h + 1$ if h is odd.

   The above two properties imply that $m(0) - r_{m(0)} \leq n_i(t) \leq M(0) + r_{M(0)}$ for all $i \in V$ and for all $t \geq 0$, This yields (ii).
(iii) Observe that
   $$g_2(h,k) = g_1(h,k) - \xi_1 r_h - \xi_2 r_k,$$
   where g_1 is the map defining the evolution of (25). Hence
   $$\mathbb{P}\left[g_2(h,k) = g_1(h,k)\right] \geq \frac{1}{4}.$$
   This fact, combined with the fact (iii) proved along the proof of Theorem 3.6, ensures that, also for (43), there is a positive probability of reaching a state in $\mathcal{A}$ in a finite time.

The above theorem and the previous remarks about T_{all} lead to the following claim about the original system.

Corollary 4.7 *Let $x(t)$ evolve following (42). Then almost surely there exists $T_{con} \in \mathbb{N}$ and $\alpha \in \mathbb{Z}$ such that $x_i(t) = \alpha$ for all $i \in V$ and for all $t \geq T_{con}$.*

As for (24), it is an open problem to provide a theoretical estimation of the distance between the consensus point to which (42) leads the systems, and the average of the initial condition. We limit our analysis to the following simulations. In Figure 8 we plot the variable z as previously defined for the *globally quantized* strategy using deterministic quantizers, i.e., $z = |c - 1/N \mathbf{1}^* x(0)|$ where c is such that $\lim_{t\to\infty} x(t) = c\mathbf{1}$. The variable z represents the distance between the consensus point to which the globally quantized strategy leads the systems and the average of the initial condition. We plot the value of z for a family of random geometric graphs of increasing size from $N = 10$ up to $N = 80$. The initial condition $x_i(0)$ is chosen randomly inside the interval $[-100, 100]$ for all $1 \leq i \leq N$. Moreover for each N, z is calculated as the mean of 100 trials. In Figure 9 we provide a comparison between (24) and (42). Surprisingly, the globally quantized strategy using probabilistic quantizers, differently from the globally quantized strategy using deterministic quantizers, seems to reach the consensus very close to the average of the initial condition.

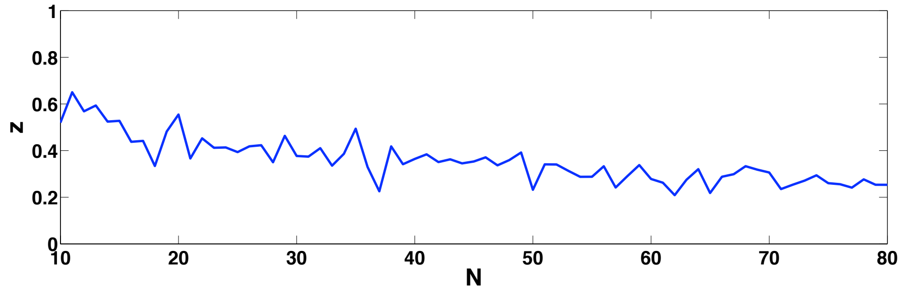

Fig. 8. Behavior of z for a family of random geometric graphs when considering the globally quantized strategy using probabilistic quantizers.

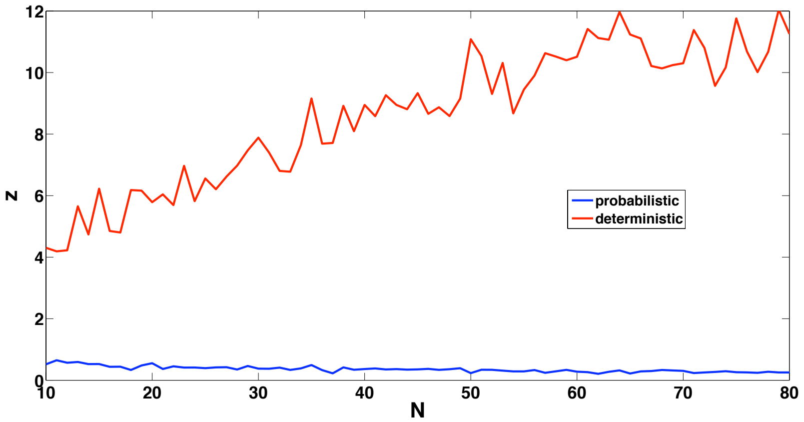

Fig. 9. Comparison in terms of z between the "deterministic" and the "probabilistic" strategy, for a family of random geometric graphs.

5 Conclusion and open questions

In this paper we studied the gossip algorithm for the consensus problem with quantized communication. In order to face the effects due to the quantization (both deterministic and probabilistic) we proposed here two updating rules: the globally quantized strategy and the partially quantized strategy. In the former the nodes use only quantized information in order to update their state. In the latter they have access also to exact information regarding their own state. We summarize our results in the following table.

	Globally Quant.	Partially Quant.
Deterministic	Finite time conv. to consensus Larger averaging error	Finite time conv. to $N^{-1/2}\|x - x_{ave}\mathbf{1}\|_2 \leq 1/2$ Average preserved
Probabilistic	Finite time conv. to consensus Smaller averaging error	Asympt. conv. to $N^{-1/2}\sqrt{\mathbb{E}[\|x - x_{ave}\mathbf{1}\|_2^2]} \leq 1/2$ Average preserved

We have seen that the partially quantized strategy, with both the quantizers, deterministic and probabilistic, does not reach the consensus in general, but maintains the average of the state at each iteration and drives all the states very close to the average of the initial condition. On the other hand, we have shown that the globally strategy leads almost surely to a consensus which, however, does not coincide with the average of the initial condition. We have provided some simulations characterizing the distance between the consensus point and the initial average. While using the deterministic quantizer this distance turns out to be not negligible, with the probabilistic quantizer the consensus is reached surprisingly very close to the average of the initial condition. Providing some theoretical insights on this fact will be the object of future research.

A second issue which deserves attention is the speed of convergence of the presented algorithms. Indeed, the non-quantized gossip algorithm [2] is known to asymptotically converge, in a mean squared sense, at exponential speed, with a rate which depends on the matrix W. It is thus natural to conjecture that the convergence of the quantized version will be roughly exponential, as long as the differences states are much larger than the quantization step. Preliminary results in this sense are in [6] and in Section 4. However, the granularity effects eventually comes out in the convergence, making the systems converge in finite time to some limit point: estimates on such a time are sought in [9] and in the pair of recent papers [10,11]. Giving a rigorous and complete clarification of this question is an interesting open problem.

References

[1] T. C. Aysal, M. Coates, and M. Rabbat. Distributed average consensus using probabilistic quantization. In *IEEE Workshop on Statistical Signal Processing*, pages 640–644, Maddison, Wisconsin, August 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, P. Frasca, T. Taylor, and S. Zampieri. Average consensus on networks with transmission noise or quantization. In *ECC '07, Kos*, pages 1852–1857, 2007.

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

[5] F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. *IEEE J. on Selected Areas of Communications*, 26(4):634–649, 2008.

[6] P. Frasca, R. Carli, F. Fagnani, and S. Zampieri. Average consensus by gossip algorithms with quantized communication. In *IEEE Conf. on Decision and Control*, pages 4831–4836, Cancun, Mexico, December 2008.

[7] P. Frasca, R. Carli, F. Fagnani, and S. Zampieri. Average consensus on networks with quantized communication. *International Journal on Robust and Nonlinear Control*, 2008. in press.

[8] S. Kar and J. M. F. Moura. Distributed consensus algorithms in sensor networks: Quantized data. Dec 2007. submitted.

[9] A. Kashyap, T. Başar, and R. Srikant. Quantized consensus. *Automatica*, 43(7):1192–1203, 2007.

[10] J. Lavaei and R. M. Murray. On quantized consensus by means of gossip algorithm – part i: Convergence proof. *Submitted to American Control Conference.*

[11] J. Lavaei and R. M. Murray. On quantized consensus by means of gossip algorithm – part ii: Convergence time. *Submitted to American Control Conference.*

[12] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis. On distributed averaging algorithms and quantization effects. In *IEEE Conf. on Decision and Control*, pages 4825–4830, Cancun, Mexico, December 2008.

[13] J. R. Norris. *Markov Chains.* Cambridge University Press, Cambridge, UK, 1997.

[14] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *IEEE Proceedings*, 95(1):215–233, 2007.

[15] M. Penrose. *Random Geometric Graphs.* Oxford Studies in Probability. Oxford University Press, 2003.

[16] A. Tahbaz-Salehi and A. Jadbabaie. Consensus over random networks. *IEEE Transactions on Automatic Control*, 53:791 – 795, 2008.

[17] J.-J. Xiao and Z.-Q. Lou. Decentralized estimation in an inhomogeneous sensing enviromnent. *IEEE Transactions on Information Theory*, 51(10):3564–3575, 2005.

[18] L. Xiao, S. Boyd, and S. Lall. A scheme for asynchronous distributed sensor fusion based on average consensus. In *International Conference on Information Processing in Sensor Networks (IPSN'05)*, pages 63–70, Los Angeles, CA, April 2005.

[19] M. Zhu and S. Martinez. On the convergence time of distributed quantized averaging algorithms. In *IEEE Conf. on Decision and Control*, pages 3971–3976, Cancun, Mexico, December 2008.