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

Quantized average consensus via dynamic coding/decoding schemes

Ruggero Carli Francesco Bullo Sandro Zampieri

***Abstract*— In the average consensus a set of linear systems has to be driven to the same final state which corresponds to the average of their initial states. This contribution presents a consensus strategy in which the systems can exchange information among themselves according to a fixed connected digital communication network. Beside the decentralized computational aspects induced by the choice of the communication network, we here have also to face the quantization effects due to the digital links. We here present and discuss two different encoding/decoding strategies with theoretical and simulation results on their performance.**

I. INTRODUCTION

A basic aspect in the analysis and in the design of cooperative agents systems is related to the effect of the agents information exchange on the coordination performance. A coordination task which is widely treated in the literature is the so called average consensus. This is the problem of driving states of a set of dynamic systems to a final common state which corresponds to the average of initial states of each system.

The way in which the information flow on the network influences the consensus performance has been already considered in the literature [1], [2], where the communication cost is modeled simply by the number of active links in the network which admit the transmission of real numbers. However, this model can be too rough when the network links represent actual digital communication channels. Indeed the transmission over a finite alphabet requires the design of efficient ways to translates real numbers into digital information, namely smart quantization techniques. The investigation of consensus under quantized communication started with [3] in which the authors study systems having (and transmitting) integer-valued states and propose a class of gossip algorithms which preserve the average of states and are guaranteed to converge up to one quantization bin. In [4] the quantization error is seen as a zero-mean additive noise and by simulations, it is shown for small N that, if the increasing correlation among the node states is taken into account, the variance of the quantization noise diminishes and nodes converge to a consensus. In [5] the authors propose a distributed algorithm that uses quantized values and preserves the average at each iteration. Even if the consensus is not reached, they showed favorable convergence properties using simulations on some static topologies. The authors in [6] adopt the probabilistic quantization (PQ) scheme to quantize the information before transmitting to the neighboring sensors. They show that the node states reach consensus to a quantized level; only in expectation do they converge to the desired average.

The main contribution of this paper is to introduce a novel quantized strategy that permits both to maintain the initial average and to reach it asymptotically. More precisely we adapt coding/decoding strategies, that were proposed for centralized control and communication problems, to the distributed consensus problem. In particular, we present two coding/decoding strategies, one based on the exchange of logarithmically quantized information, the other on a zoom in - zoom out strategy (this latter involves the use of uniform quantizers). We provide analytical and simulative results illustrating the convergence properties of these strategies. In particular we show that the convergence factors depend smoothly on the accuracy parameter of the quantizers used and that, remarkably, that the critical quantizer accuracy sufficient to guarantee convergence is independent from the network dimension.

The paper is organized as follows. Section II briefly reviews the standard average consensus algorithm. In Section III we present two strategies of coding/decoding of the data throughout reliable digital channels: one based on logarithmic quantizers, the other on uniform quantizers. We analyze the former from a theoretical point in Section IV and Section V. We provide simulations results for the latter in Section VI. Finally, we gather our conclusions in Section VII.

Mathematical Preliminaries

Before proceeding, we collect some definitions and notations which are used through the paper.

In this paper $\mathcal{G} = (V, E)$ denotes a *undirected graph* where $V = \{1, \dots, N\}$ is the set of vertices and E is the set of (directed) edges, i.e., a subset of $V \times V$. Clearly, if $(i, j) \in E$ also $(j, i) \in E$ and this means that i can transmit information about its state to j and vice-versa. Any $(i, i) \in E$ is called a *self loop*. A *path* in $\mathcal{G}$ consists in a sequence of vertices $(i_1, i_2, \dots, i_r)$ such that $(i_j, i_{j+1}) \in E$ for every $j \in \{1, \dots, r-1\}$. A graph is said to be *connected* if for any given pair of vertices (i, j) there exists a path connecting i to j. A matrix M is said to be *stochastic* if $M_{ij} \geq 0$ for all i and j and the sums along each row are equal to 1. Moreover a matrix M is said to be *doubly stochastic* if it is stochastic and also the sums along each column are equal to 1. Given a nonnegative matrix $M \in \mathbb{R}^{N \times N}$, we can define an induced graph $\mathcal{G}_M$ by taking N nodes and putting an edge (j, i) in E if $M_{ij} > 0$. Given a graph $\mathcal{G}$ on V, M is said to be *adapted* or *compatible* with $\mathcal{G}$ if $\mathcal{G}_M \subseteq \mathcal{G}$. Given a vector $v \in \mathbb{R}^N$ and a matrix $M \in \mathbb{R}^{N \times N}$, we denote with v^T and M^T the transpose of v and of M. Then, let $\sigma(M)$ denote the set of eigenvalues of M. If M is symmetric and stochastic we assume that $\sigma(M) = \{1, \lambda_1(M), \dots, \lambda_{N-1}(M)\}$, where

This material is based upon work supported in part by AFOSR MURI Award FA9550-07-1-0528 and ONR N00014-07-1-0721.

R. Carli and F. Bullo are with the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA 93106, USA, {carlirug|bullo}@engr.ucsb.edu.
S. Zampieri is with the Department of Information Engineering, Università di Padova, Via Gradenigo 6/a, 35131 Padova, Italy zampi@dei.unipd.it.

$1, \lambda_1(M), \ldots, \lambda_{N-1}(M)$ are the eigenvalues of M and are such that $\lambda_1(M) \geq \lambda_2(M) \geq \ldots \geq \lambda_{N-1}(M)$. We define

$$\lambda_{\max}(M) = \lambda_1(M), \qquad \lambda_{\min}(M) = \lambda_{N-1}(M).$$

With the symbols $\mathbf{1}$ and $\mathbf{0}$ we denote the N-dimensional vectors having respectively all the components equal to 1 and equal to 0. Given $v = [v_1, ..., v_N]^T \in \mathbb{R}^N$, $\operatorname{diag}\{v\}$ or $\operatorname{diag}\{v_1, \ldots, v_N\}$ mean a diagonal matrix having the components of v as diagonal elements. Moreover, $\|v\|$ and $< v >$ denote the Euclidean norm of v and the subspace generated by v, respectively. Finally, for $f, g : \mathbb{N} \to \mathbb{R}$, we say that $f \in o(g)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.

II. Problem Formulation

We start this section by briefly describing the standard discrete-time consensus algorithm. Assume that we have a set of agents V and a graph $\mathcal{G}$ on V describing the feasible communications among the agents. For each agent $i \in V$ we denote by $x_i(t)$ the estimate of the average of agent i at time t. Standard consensus algorithm are constructed by choosing a doubly stochastic matrix $P \in \mathbb{R}^{N\times N}$ compatible with $\mathcal{G}$ and assuming that at every times t agent i updates its estimate according to

$$x_i(t+1) = \sum_{j=1}^{N} P_{ij} x_j(t). \tag{1}$$

More compactly we can write

$$x(t+1) = Px(t), \tag{2}$$

where $x(t)$ is the column vector entries $x_i(t)$ represent the agents states. In our treatment we will restrict to the case in which P is symmetric, i.e., $P^T = P$. Note that a stochastic symmetric matrix P is automatically doubly stochastic.

It is well known in the literature [7] that, if P is a symmetric stochastic matrix with positive diagonal entries and such that $\mathcal{G}_P$ is connected, then the algorithm (2) solves the *average consensus problem*, namely

$$\lim_{t\to+\infty} x(t) = x_a(0)\mathbf{1},$$

where $x_a(0) := \frac{1}{N}\mathbf{1}^T x(0)$. From now on we will assume the following.

Assumption 1: P is a symmetric stochastic matrix such that $P_{ii} > 0$, for $i \in \{1, \ldots, N\}$, and $\mathcal{G}_P$ is connected.

Note that the algorithm (2) 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 what follows, we consider a more realistic case, i.e., we assume that the communication network is constituted only of rate-constrained digital links. Accordingly, the main objectives of this paper are to understand (i) how the standard consensus algorithm may be modified to overcome the forced quantization effects due to the digital channel and (ii) how much does its performance degrade.

We note that the presence of a rate constraint 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 the j-th agent symbolic data in a finite or countable alphabet; using only this data, the j-th agent can build at most an estimate of the i-th agent's state. To tackle this problem we take a two step approach. First, we introduce a coding/decoding scheme; each agent uses this scheme to estimate the positions of its neighbors. Second, we consider the standard consensus algorithm where, in place of the exact knowledge of the states of the systems, we substitute estimates calculated according to the proposed coding/decoding scheme.

III. Coder/decoder pairs for digital channels

In this section we discuss a general and two specific coder/decoder models for reliable digital channels; we follow the survey [8]. We will later adopt this coder/decoder scheme to define communication protocols in the robotic network.

Suppose a source is communicating to a receiver some time-varying data $x : \mathbb{N} \to \mathbb{R}$ via repeated transmissions at time instants in $\mathbb{N}$. Each transmission takes place through a digital channel, i.e., messages can only be symbols in a finite or countable set. The channel is assumed to be reliable, i.e., the transmitted symbol is received without error.

A coder/decoder pair for a digital channel is given by the sets:

(i) a set Ξ, serving as *state space* for the coder/decoder; a fixed $\xi_0 \in \Xi$ is the *initial coder/decoder state*;

(ii) a finite or countable set $\mathcal{A}$, serving as *transmission alphabet*; elements $\alpha \in \mathcal{A}$ are called message;

and by the maps:

(i) a map $F : \Xi \times \mathcal{A} \to \Xi$, called the *coder/decoder dynamics*;

(ii) a map $Q : \Xi \times \mathbb{R} \to \mathcal{A}$, being the *quantizer function*;

(iii) a map $H : \Xi \times \mathcal{A} \to \mathbb{R}$, called the *decoder function*.

The coder computes the symbols to be transmitted according to, for $t \in \mathbb{N}$,

$$\xi(t+1) = F(\xi(t), \alpha(t)), \quad \alpha(t) = Q(\xi(t), x(t)).$$

Correspondingly, the decoder implements, for $t \in \mathbb{N}$,

$$\xi(t+1) = F(\xi(t), \alpha(t)), \quad \hat{x}(t) = H(\xi(t), \alpha(t)).$$

Coder and decoder are jointly initialized at $\xi(0) = \xi_0$. Note that an equivalent representation for the coder is $\xi(t+1) = F(\xi(t), Q(\xi(t), x(t)))$, and $\alpha(t) = Q(\xi(t), x(t))$. In summary, the coder/decoder dynamics is given by

$$\begin{aligned} \xi(t+1) &= F(\xi(t), \alpha(t)), \\ \alpha(t) &= Q(\xi(t), x(t)), \\ \hat{x}(t) &= H(\xi(t), \alpha(t)). \end{aligned} \tag{3}$$

In what follows we present two interesting coder/decoder pairs: the "zoom in - zoom out" uniform quantizer strategy and the logarithmic quantizer.

A. Zoom in - zoom out uniform coder

In this strategy the information transmitted is quantized by a scalar uniform quantizer which can be described as follows. For $L \in \mathbb{N}$, define the *uniform set of quantization levels*

$$S_L = \Big\{ -1 + \frac{2\ell - 1}{L} \mid \ell \in \{1, \ldots, L\} \Big\} \cup \{-1\} \cup \{1\}$$

and the corresponding *uniform quantizer* (see Figure 1) $\mathrm{unq}_L : \mathbb{R} \to S_L$ by

$$\mathrm{unq}_L(x) = -1 + \frac{2\ell - 1}{L}, \quad \text{for } \ell \in \{1, \dots, L\} \text{ s.t. } \frac{2(\ell - 1)}{L} \le x + 1 \le +\frac{2\ell}{L},$$

and otherwise $\mathrm{unq}_L(x) = 1$ if $x > 1$ or $\mathrm{unq}_L(x) = -1$ if $x < -1$. Note that larger values of the parameter L correspond to more accurate uniform quantizers unq_L. Moreover note that, if we define m to be the number of quantization levels we have that $m = L + 2$.

For $L \in \mathbb{N}$, $k_{in} \in]0,1[$, and $k_{out} \in]1, +\infty[$, the *zoom in - zoom out uniform coder/decoder* has the state space $\Xi = \mathbb{R} \times \mathbb{R}_{>0}$, the initial state $\xi_0 = (0, 1)$, and the alphabet $\mathcal{A} = S_L$. The coder/decoder state is written as $\xi = (\hat{x}_{-1}, f)$ and the coder/decoder dynamics are

$$\hat{x}_{-1}(t+1) = \hat{x}_{-1}(t) + f(t)\alpha(t),$$
$$f(t+1) = \begin{cases} k_{\text{in}}\, f(t), & \text{if } |\alpha(t)| < 1, \\ k_{\text{out}}\, f(t), & \text{if } |\alpha(t)| = 1. \end{cases}$$

The quantizer and decoder functions are, respectively,

$$\alpha(t) = \mathrm{unq}_L\left(\frac{x(t) - \hat{x}_{-1}(t)}{f(t)}\right), \ \hat{x}(t) = \hat{x}_{-1}(t) + f(t)\alpha(t).$$

The coder/decoder pair is analyzed as follows. One can observe that $\hat{x}_{-1}(t+1) = \hat{x}(t)$, i.e., the first component of the coder/decoder state contains the estimate of x. The transmitted messages contain a quantized version of the estimate error $x - \hat{x}_{-1}$ scaled by factor f. Accordingly, the second component of the coder/decoder state f is referred to as the *scaling factor*: it grows when $|x - \hat{x}_{-1}| \ge f$ ("zoom out step") and it decreases when $|x - \hat{x}_{-1}| < f$ ("zoom in step").

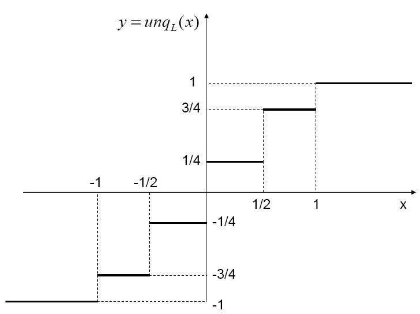

Fig. 1. The uniform quantizer ($m = 6$).

B. Logarithmic coder

This strategy is presented for example in [9]. Given an *accuracy parameter* $\delta \in]0, 1[$, define the *logarithmic set of quantization levels*

$$S_\delta = \left\{\left(\frac{1+\delta}{1-\delta}\right)^\ell\right\}_{\ell \in \mathbb{Z}} \cup \{0\} \cup \left\{-\left(\frac{1+\delta}{1-\delta}\right)^\ell\right\}_{\ell \in \mathbb{Z}}, \quad (4)$$

and the corresponding *logarithmic quantizer* (see Figure 2) $\mathrm{lgq}_\delta : \mathbb{R} \to S_\delta$ by

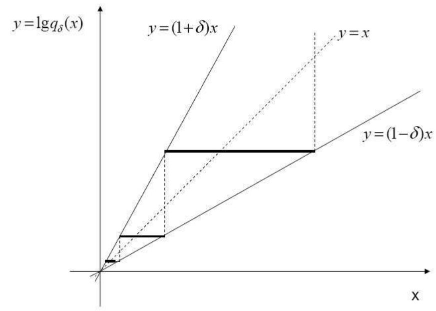

Fig. 2. The logarithmic quantizer.

$$\mathrm{lgq}_\delta(x) = \left(\frac{1+\delta}{1-\delta}\right)^\ell, \quad \text{for } \ell \in \mathbb{Z} \text{ s.t. } \frac{(1+\delta)^{\ell-1}}{(1-\delta)^\ell} \le x \le \frac{(1+\delta)^\ell}{(1-\delta)^{\ell+1}},$$

otherwise $\mathrm{lgq}_\delta(x) = 0$ if $x = 0$ or $\mathrm{lgq}_\delta(x) = -\mathrm{lgq}_\delta(-x)$ if $x < 0$.
Note that smaller values of the parameter δ correspond to more accurate logarithmic quantizers lgq_δ. For $\delta \in]0, 1[$, the *logarithmic coder/decoder* is defined by the state space $\Xi = R$, initial state $\xi_0 = 0$, the alphabet $\mathcal{A} = S_\delta$, and by the maps

$$\xi(t+1) = \xi(t) + \alpha(t)$$
$$\alpha(t) = \mathrm{lgq}_\delta(x(t) - \xi(t))$$
$$\hat{x}(t) = \xi(t) + \alpha(t).$$

The coder/decoder pair is analyzed as follows. One can observe that $\xi(t+1) = \hat{x}(t)$ for $t \in \mathbb{N}$, that is, the coder/decoder state contains the estimate of the data x. The transmitted messages contain a quantized version of the estimate error $x - \xi$. The estimate $\hat{x} : N \to \mathbb{R}$ satisfies the recursive relation

$$\hat{x}(t+1) = \hat{x}(t) + \mathrm{lgq}_\delta\left(x(t+1) - \hat{x}(t)\right),$$

with initial condition $\hat{x}(0) = \mathrm{lgq}_\delta\left(x(0)\right)$ determined by $\xi(0) = 0$. Finally, define the function $r : \mathbb{R} \to \mathbb{R}$ by $r(y) = \frac{\mathrm{lgq}_\delta(y) - y}{y}$ for $y \neq 0$ and $r(0) = 0$. Some elementary calculations show that $|r(y)| \le \delta$ for all $y \in \mathbb{R}$. Accordingly, if we define the trajectory $\omega : \mathbb{N} \to [-\delta, +\delta]$ by $\omega(t) = r(x(t+1) - \hat{x}(t))$, then we obtain that

$$\hat{x}(t+1) = \hat{x}(t) + (1 + \omega(t))\big(x(t+1) - \hat{x}(t)\big). \quad (5)$$

IV. CONSENSUS ALGORITHM WITH EXCHANGE OF QUANTIZED INFORMATION

We consider now the average consensus algorithm with the assumption that the agents can communicate only via digital channels. Here, we adopt the logarithmic coder/decoder scheme (3) described in Section III-B; we analyze the zoom in - zoom out strategy via simulations in Section VI.

Here is an informal description of our proposed scheme. We envision that along each communication edge we implement a logarithmic coder/decoder; in other words, each agent transmits via a dynamic encoding scheme to all its neighbors the quantized information regarding its position. Once state estimates of all node's neighbors are available, each node will then implement the average consensus algorithm.

Next, we provide a formal description of the proposed algorithm. Let $P \in \mathbb{R}^{N\times N}$ be a stochastic symmetric matrix with positive diagonal elements and with connected induced graph $\mathcal{G}_{\mathcal{P}}$. Assume there are digital channels along all edges of $\mathcal{G}_{\mathcal{P}}$ capable of carrying a countable number of symbols. Pick an accuracy parameter $\delta \in \,]0,1[$. The *consensus algorithm with dynamic coder/decoder* is defined as follows:

Processor states: For each $i \in \{1,\dots,N\}$, node i has a state variable $x_i \in \mathbb{R}$ and state estimates $\hat{x}_j \in \mathbb{R}$ of the states of all neighbors j of i in $\mathcal{G}_{\mathcal{P}}$. Furthermore, node i maintains a copy of $\hat{x}_i$.

Initialization: The state $x(0) = [x_1(0),\dots,x_N(0)]^T \in \mathbb{R}^N$ is given as part of the problem. All estimates $\hat{x}_j(0)$, for $j \in \{1,\dots,N\}$, are initialized to 0.

State iteration: At time $t \in \mathbb{N}$, for each i, node i performs three actions in the following order:

(1) Node i updates its own state by

$$x_i(t) = x_i(t-1) + \sum_{j=1}^{N} P_{ij}\left(\hat{x}_j(t-1) - \hat{x}_i(t-1)\right). \tag{6}$$

(2) Node i transmits to all its neighbors the symbol

$$\alpha_i(t) = \mathrm{lgq}_\delta(x_i(t) - \hat{x}_i(t-1)).$$

(3) Node i updates its estimates

$$\hat{x}_j(t) = \hat{x}_j(t-1) + \alpha_j(t), \tag{7}$$

for j being equal to all neighbors of i and to i itself.

Remark 2: Robot i and all its neighbors j maintain in memory an estimate $\hat{x}_i$ of the state x_i. We denote all these estimates by the same symbol because they are all identical: they are initialized in the same manner and they are updated through the same equation with the same information. On the other hand, it would be possible to adopt distinct quantizer accuracies δ_{ij} for each communication channel (i,j). In such a case then we would have to introduce variables $\hat{x}_{ij}$ that node i and j would maintain for the estimate of x_i.

We now analyze the algorithm. First, we write the closed-loop system in matrix form. Equation (6) is written as

$$x(t+1) = x(t) + (P-I)\hat{x}(t). \tag{8}$$

The N-dimensional vector of state estimates $\hat{x} = [\hat{x}_1,\dots,\hat{x}_N]^T$ is updated according to the multiplicative-noise model in equation (5). In other words, there exist $\omega_j : \mathbb{N} \to [-\delta, +\delta]$, for $j \in \{1,\dots,N\}$, such that

$$\hat{x}_j(t+1) = \hat{x}_j(t) + (1+\omega_j(t))\big(x_j(t+1) - \hat{x}_j(t)\big),$$

and, for $\Omega(t) := \mathrm{diag}\{\omega_1(t),\dots,\omega_N(t)\}$,

$$\hat{x}(t+1) = \hat{x}(t) + (I+\Omega(t))\big(x(t+1) - \hat{x}(t)\big). \tag{9}$$

Equations (8) and (9) determine the closed-loop system. Next, we define the estimate error $e = \hat{x} - x \in \mathbb{R}^N$. By straightforward calculations we can rewrite the close-loop system in terms of the quantities x and e, for $t \in \mathbb{Z}_{\geq 0}$, as

$$\begin{bmatrix} x(t+1) \\ e(t+1) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \Omega(t) \end{bmatrix} \begin{bmatrix} P & P-I \\ P-I & P-2I \end{bmatrix} \begin{bmatrix} x(t) \\ e(t) \end{bmatrix}. \tag{10}$$

Initial conditions are $x(0)$ and $e(0) = -x(0)$. We state now the main properties of our quantized consensus algorithm.

Theorem 3: Assume $P \in \mathbb{R}^{N\times N}$ satisfies Assumption 1 and define $\bar{\delta} \in \mathbb{R}$ by

$$\bar{\delta} := (1+\lambda_{\min}(P))\,/\,(3-\lambda_{\min}(P)). \tag{11}$$

The solution $t \mapsto (x(t), e(t))$ of the consensus algorithm with dynamic coder/decoder satisfies:

(i) the state average is preserved by the algorithm, that is, $\frac{1}{N}\sum_{i=1}^N x_i(t) = \frac{1}{N}\sum_{i=1}^N x_i(0)$ for all $t \in \mathbb{N}$;

(ii) if $0 < \delta < \bar{\delta}$, then the state variables converge to their average value and the estimate error vanishes, that is,

$$\lim_{t\to\infty} x(t) = x_a(0)\mathbf{1}, \qquad \lim_{t\to\infty} e(t) = 0.$$

where $x_a(0) = \frac{1}{N}\mathbf{1}^T x(0)$.

Remark 4: Consider the sequence of circulant matrices $P_N \in \mathbb{R}^{N\times N}$ defined by

$$P_N = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & 0 & \cdots & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & \cdots & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \frac{1}{3} & 0 & 0 & 0 & \cdots & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}. \tag{12}$$

For this sequence of symmetric stochastic matrices we have that $\lambda_{\min}(P_N) = \frac{1}{3} - \frac{2}{3}\cos\left(\frac{2\pi}{N}\left\lfloor\frac{N}{2}\right\rfloor\right)$. Hence $\lambda_{\min}(P_N) \geq -\frac{1}{3}$, implying therefore that $\bar{\delta} \geq \frac{1}{5}$ for all N. This shows that $\bar{\delta}$ is uniformly bounded away from 0. This is a remarkable property of scalability on the dimension of the network. However the fact that the critical accuracy sufficient to guarantee convergence is independent on the network dimension is more general than what seen in the above example. Indeed, assume that $P_N \in \mathbb{R}^{N\times N}$ is a sequence of matrices of increasing size, where each P_N satisfies Assumption 1 and where each P_N has all the diagonal elements greater than a positive real number $\bar{p}$. Then, by Gershgorin's Theorem [10], we have that $\lambda_{\min}(P_N) \geq -1 + 2\bar{p}$ and hence $\bar{\delta} \geq \frac{\bar{p}}{2-\bar{p}}$ for all N. It follows that the critical accuracy sufficient to guarantee convergence is bounded away from zero uniformly on the dimension of the network. □

V. Exponential convergence

The objective of this section is to understand, by means of a Lyapunov analysis, how much the quantization affects the performance of the consensus algorithm. We start by introducing some definitions. A function $f : \mathbb{N} \to \mathbb{R}$ converges to 0 *exponentially fast* if there exist a constant $C > 0$ and another constant $\xi \in [0,1[$ such that $|f(t)| \leq C\xi^t$, for all t; the infimum among all numbers $\xi \in [0,1[$ satisfying the exponential convergence property is called the *exponential convergence factor* of f. In other words, the exponential convergence factor of f is given by

$$\limsup_{t\to\infty} |f(t)|^{\frac{1}{t}}.$$

Consider first the system (2). To quantify the speed of convergence of (2) toward consensus, we introduce the variable

$$\bar{x}(t) := x(t) - x_a(0)\mathbf{1}.$$

Clearly, $\lim_{t\to\infty} x(t) = x_a(0)\mathbf{1}$ if and only if $\lim_{t\to\infty} \bar{x}(t) = 0$. It is easy to see that the variable $\bar{x}$ satisfies the same recursive equation of the variable x, i.e.,

$$\bar{x}(t+1) = P\bar{x}(t). \tag{13}$$

Moreover note that $\mathbf{1}^T\bar{x}(t) = 0$, for all $t \geq 0$. We define the exponential convergence factor of $\bar{x}(t)$, for a given initial condition $\bar{x}_0 \in <\mathbf{1}>^\perp$, to be $\rho(P, \bar{x}_0) := \limsup_{t\to\infty} ||\bar{x}(t)||^{\frac{1}{t}}$. We can get rid of the initial condition and define the *exponential convergence factor* of the system (2) as follows

$$\rho(P) := \sup_{\bar{x}_0 \in <\mathbf{1}>^\perp} \rho(P, \bar{x}_0) \tag{14}$$

Consider now the positive semidefinite matrix $I - P$. Notice that

$$\rho(P, \bar{x}_0) = \limsup_{t\to\infty} (\bar{x}(t)^T(I-P)\bar{x}(t))^{\frac{1}{2t}}$$

and so we can characterize the speed of convergence to 0 of the variable $\bar{x}$ by studying the exponential convergence factor of the Lyapunov function $\bar{x}(t)^T(I-P)\bar{x}(t)$.

Theorem 5: Consider (13) with $P \in \mathbb{R}^{N\times N}$ satisfing Assumption 1. Then the function $t \mapsto (\bar{x}(t)^T(I-P)\bar{x}(t))^{1/2}$, defined along any trajectory $t \mapsto \bar{x}(t)$, converges exponentially fast to 0. Moreover, the factor $\rho(P)$, defined in equation (14), satisfies

$$\rho(P) = \max\{\lambda_{\max}(P), -\lambda_{\min}(P)\}.$$

This concludes the analysis of the algorithm (2). In the sequel of this section, we provide a similar analysis of the system (10). For the sake of the notational simplicity, let $z(t) := \left[x^T(t) \;\; e^T(t)\right]^T$ and

$$\mathcal{A}(t) := \begin{bmatrix} I & 0 \\ 0 & \Omega(t) \end{bmatrix} \begin{bmatrix} P & P-I \\ P-I & P-2I \end{bmatrix}.$$

Clearly $z(0) = \left[x(0)^T \;\; e(0)^T\right]$. To perform a Lyapunov analysis of (10), it is convenient to introduce the variable

$$\bar{z}(t) = \begin{bmatrix} I - \frac{1}{N}\mathbf{1}\mathbf{1}^T & 0 \\ 0 & I \end{bmatrix} z(t).$$

Clearly, condition (ii) of Theorem 3 holds true if and only if $\lim_{t\to\infty} \bar{z}(t) = 0$. It is straightforward to see that $\bar{z}$ satisfies the same recursive equation of $z(t)$, i.e.,

$$\bar{z}(t+1) = \mathcal{A}(t)\bar{z}(t) \tag{15}$$

and that $\left[\mathbf{1}^T \;\; \mathbf{0}^T\right]^T \bar{z}(t) = 0$ for all $t \geq 0$. Consider now the matrix $L \in \mathbb{R}^{2N\times 2N}$ defined as

$$L = \begin{bmatrix} I-P & 0 \\ 0 & \gamma I \end{bmatrix}.$$

For each $\gamma > 0$ define $\tilde{\rho}\left(P, \delta, \gamma; \bar{z}_0, \{\mathcal{A}(t)\}_{t=0}^{\infty}\right) := \limsup_{t\to\infty} (\bar{z}(t)^T L \bar{z}(t))^{\frac{1}{2t}}$. We can get rid of the initial conditions $\bar{z}_0$ and the sequences $\{\mathcal{A}(t)\}_{t=0}^{\infty}$ by considering

$$\tilde{\rho}(P, \delta, \gamma) := \sup_{\bar{z}_0, \{\mathcal{A}(t)\}_{t=0}^{\infty}} \tilde{\rho}\left(P, \delta, \gamma; \bar{z}_0, \{\mathcal{A}(t)\}_{t=0}^{\infty}\right) \tag{16}$$

It can be shown that $\tilde{\rho}(P, \delta, \gamma)$ is independent of γ and for this reason we denote it as $\tilde{\rho}(P, \delta)$.

We characterize now $\tilde{\rho}(P, \delta, \gamma)$. To this aim, consider the following semidefinite programming problem

$$\bar{\beta}(P, \delta, \gamma) := \begin{array}{rc} & \max \;\; \beta \\ \text{such that} & R_1^T L R_1 - L \leq -\beta L \end{array} \tag{17}$$

We have the following result.

Theorem 6: Consider (15) with the matrix P satisfying Assumption 1. Let $\bar{\delta}$ be as in (11) and let $\delta \in \mathbb{R}$ be such that $0 \leq \delta < \bar{\delta}$. Moreover let $\gamma \in \mathbb{R}$ be such that $\gamma > 0$, and let $\bar{\beta}(P, \delta, \gamma)$ be as in (17). Then, the function $t \to (\bar{z}(t)^T L \bar{z}(t))^{1/2}$, defined along any trajectory $t \to \bar{z}(t)$ converges exponentially fast to 0 and the factor $\tilde{\rho}(P, \delta)$, defined in equation (16), satisfies

$$\tilde{\rho}(P, \delta) \leq \left(1 - \bar{\beta}(P, \delta, \gamma)\right)^{1/2}.$$

In general, assigned P and the value of the accuracy parameter δ, one could be interested in determining the maximum value of $\bar{\beta}$, as function of γ. Clearly, the best bound on $\tilde{\rho}(P, \delta)$ corresponds to to the maximum value of $\bar{\beta}$, i.e.,

$$\tilde{\rho}(P, \delta) \leq \left(1 - \bar{\beta}_{opt}(P, \delta)\right)^{1/2}$$

where $\bar{\beta}_{\text{opt}}(P, \delta) := \max_{\gamma>0} \bar{\beta}(P, \delta, \gamma)$. We illustrate this discussion in the following example.

Example 7: In this example we consider a connected random geometric graph generated by choosing $N = 30$ points at random in the unit square, and then placing an edge between each pair of pints at distance less than 0.4. The matrix P is built using the Metropolis weights [11]. In Figure 3, we depict the behavior of $\left(1 - \bar{\beta}_{\text{opt}}(P, \delta)\right)^{1/2}$ as a function of δ. The dotted line represents the value of $\rho(P)$, that is, the convergence factor of the ideal algorithm (13). Notice that the convergence factor $\left(1 - \bar{\beta}_{\text{opt}}(P, \delta)\right)^{1/2}$ depends smoothly on the accuracy parameter δ and satisfies

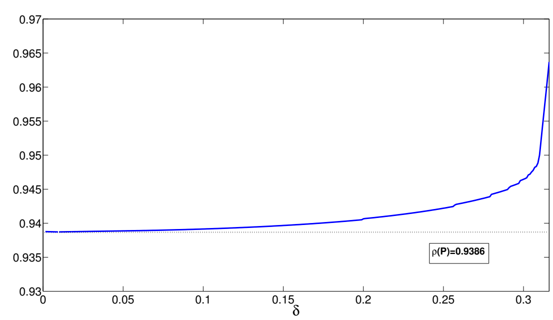

Fig. 3. Behavior of $\left(1 - \bar{\beta}_{\text{opt}}(P, \delta)\right)^{1/2}$.

$$\lim_{\delta\to 0^+} \left(1 - \bar{\beta}_{\text{opt}}(P, \delta)\right)^{1/2} = \rho(P).$$

An interesting characterization of $\tilde{\rho}$ can be provided when considering a family of matrices $\{P_N\}$ of increasing size whose maximum eigenvalue converges to 1. It is worth noting that this situation is encountered in many practical situations [12], [2], [13]. We formalize this situation as follows.

Assumption 8 (Vanishing spectral gap): Assume we have a sequence of symmetric stochastic matrices $\{P_N\} \subset \mathbb{R}^{N\times N}$ satisfying Assumption 1 and the following conditions

(i) $\lambda_{\min}(P_N) > c$ for some $c \in \,]-1,1[$ and for all $N \in \mathbb{N}$;

(ii) $\lambda_{\max}(P_N) = 1 - \epsilon(N) + o(\epsilon(N))$ as $N \to \infty$, where $\epsilon : \mathbb{N} \to \mathbb{R}$ is a positive function such that $\lim_{N\to\infty} \epsilon(N) = 0$.

According to Theorem 5, as $N \to \infty$, we have that $\rho(P_N) = 1 - \epsilon(N) + o(\epsilon(N))$. In considering the quantized version of the consensus algorithm, together with the sequence $\{P_N\}$, we have also to fix the sequence $\{\delta_N\}$. For simplicity, in the following we will assume that, $\{\delta_N\}$ is a constant sequence, i.e., $\delta_N = \delta$ with suitable δ such that $\delta < \frac{1+c}{3-c}$ which ensures the stability for all N.

Theorem 9: Let $\{P_N\} \subset \mathbb{R}^{N\times N}$ be a family of matrices of increasing size satisfying Assumptions 1 and 8. Let $\delta \in \mathbb{R}$ be such that $\delta < \frac{1+c}{3-c}$. Then, as $N \to \infty$, we have that

$$\tilde{\rho}(P_N, \delta) \leq 1 - \left(1 - \frac{1 + c + \delta^2(c-3)}{4(1-\delta^2)}\right)\epsilon(N) + o(\epsilon(N)).$$

Notice that the coefficient in from of $\epsilon(N)$ is negative. Indeed, it can be seen that that coefficient is negative if and only if $\delta^2 < (3-c)/(1+c)$ and this is true since we have chosen $\delta < (1+c)/(3-c)$ and since $\delta < 1$.

VI. Numerical simulations

In this section we provide some numerical results illustrating the performance of the Zoom in -Zoom out strategy. We consider the same connected random geometric graph of Example 7. We assume that the initial conditions has been randomly generated inside the interval $[-100, 100]$. For all the experiments, we set the parameters k_{in} and k_{out} to the values $1/2$ and 2 respectively, and initialized the scaling factor f of each agent to the value 50. Moreover we run simulations for two different values of m, $m = 5$ and $m = 10$. The results obtained are reported in Figure 4. The variable plotted is the normalized Euclidean norm of the vector $\bar{x}(t) := x(t) - x_a(0)\mathbf{1}$, that is,

$$s(t) = \frac{1}{N}\|\bar{x}(t)\|^{1/2}.$$

Note that, as depicted in Figure 4, also the zoom in- zoom out uniform coder- decoder strategy seems to be very efficient in achieving the consensus. In particular it is remarkable that this strategy works well even if the uniform quantizer has few quantization levels ($m = 5$). Finally it is worth observing that, as seen in Example 7, also in this case the performance degrades smoothly as the quantization becomes coarser.

VII. Conclusions

In this paper we presented a new approach solving the average consensus problem in presence of only quantized exchanges of information. In particular we considered two strategies, one based on logarithmic quantizers, and the other one based on a zooming in-zooming out strategy. We studied them with theoretical and experimental results proving that using these schemes the quantized average consensus problem can be efficiently solved. Additionally, we show that the convergence factors depend smoothly on the accuracy parameter of the quantizer and that, remarkably,

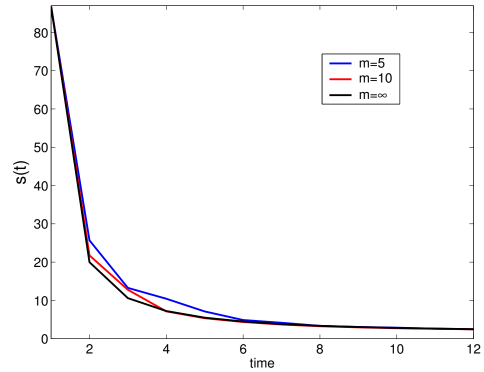

Fig. 4. Zoom in- zoom out strategy

that the critical quantizer accuracy sufficient to guarantee convergence is independent from the network dimension. A field of future research will be to look for encoding and decoding methods which are able to solve the average problem also with noisy digital channels.

References

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

[2] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri, "Communication constraints in the average consensus problem," *Automatica*, vol. 44, no. 3, pp. 671–684, 2008.

[3] A. Kashyap, T. Başar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192–1203, 2007.

[4] M. E. Yildiz and A. Scaglione, "Differential nested lattice encoding for consensus problems," in *Symposium on Information Processing of Sensor Networks (IPSN)*, (Cambridge, MA), pp. 89–98, Apr. 2007.

[5] R. Carli, F. Fagnani, P. Frasca, T. Taylor, and S. Zampieri, "Average consensus on networks with transmission noise or quantization," in *European Control Conference*, (Kos, Greece), pp. 1852–1857, June 2007.

[6] T. C. Aysal, M. Coates, and M. Rabbat, "Distributed average consensus using probabilistic quantization," in *IEEE Workshop on Statistical Signal Processing*, (Maddison, Wisconsin), pp. 640–644, Aug. 2007.

[7] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *IEEE Proceedings*, vol. 95, no. 1, pp. 215–233, 2007.

[8] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *IEEE Proceedings*, vol. 95, no. 1, pp. 108–137, 2007.

[9] N. Elia and S. K. Mitter, "Stabilization of linear systems with limited information," *IEEE Transactions on Automatic Control*, vol. 46, no. 9, pp. 1384–1400, 2001.

[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.

[11] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Symposium on Information Processing of Sensor Networks (IPSN)*, (Los Angeles, CA), pp. 63–70, Apr. 2005.

[12] S. Martínez, F. Bullo, J. Cortés, and E. Frazzoli, "On synchronous robotic networks – Part I: Models, tasks and complexity," *IEEE Transactions on Automatic Control*, vol. 52, no. 12, pp. 2199–2213, 2007.

[13] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Gossip algorithms: Design, analysis, and application," in *IEEE Conference on Computer Communications (INFOCOM)*, pp. 1653–1664, Mar. 2005.