On Quantized Consensus by Means of Gossip Algorithm – Part I: Convergence Proof

Javad Lavaei and Richard M. Murray

***Abstract*— This paper is concerned with the distributed averaging problem subject to a quantization constraint. Given a group of agents associated with scalar numbers, it is assumed that each pair of agents can communicate with each other with a prescribed probability, and that the data being exchanged between them is quantized. In this part of the paper, it is proved that the stochastic gossip algorithm proposed in a recent paper leads to reaching the quantized consensus. Some important properties of the system in the steady-state (after reaching the consensus) are also derived. The results developed here hold true for any arbitrary quantization, provided the tuning parameter of the gossip algorithm is chosen properly. The expected value of the convergence time is lower and upper bounded in the second part of the paper.**

I. Introduction

Consider a group of agents, each of which is associated with some data such as a real number or an image. The problem of contriving a strategy by means of which all agents can update themselves so that they agree upon some universal shared data is called the *consensus* or *state agreement* problem [1], [2]. Consensus has a long history in computer science, particularly in distributed computation where a program is divided into parts that run simultaneously on multiple computers communicating over a network [3], [4].

There are many important real-world problems whose treatment is contingent upon the notion of consensus. In the load-balancing problem, the tasks of disparate processors are to be equalized in order to refrain from overloading any processor [5], [6]. In the synchronization of coupled oscillators appearing in systems biology, the frequencies of oscillation of all agents are desired to become equal [7], [8]. In multi-agent coordination and flocking, there are a number of applications in which the state-agreement problem shows up [9], [10]. For instance, the heading angles of different mobile agents may be required to be aligned [11]. In a sensor network comprising a set of sensors measuring the same quantity in a noisy environment, the state estimations of different agents must be averaged [12]. A more complete survey on these topics is given in the recent paper [2].

Consider the distributed average consensus in which the values owned by the agents are to be averaged in a distributed fashion. Since it may turn out in some applications that all agents cannot update their numbers synchronously, the gossip algorithm has been widely exploited by researchers to handle the averaging problem asynchronously [13], [14]. This type of algorithm selects a pair of agents at each time, and updates their values based on some averaging policy. The consensus problem in the context of gossip algorithm has been thoroughly investigated in the literature [15], [16], [17], [18].

In light of communication constraints, the data being exchanged between any pair of agents is normally to be quantized. This has given rise to the emergence of quantized gossip algorithms. The notion of quantized consensus is defined in [17] for the case when quantized values (integers) are to be averaged over a connected network with digital communication channels. This paper shows that the quantized gossip algorithm leads to reaching the quantized consensus. This result is extended in [18] to the case when the quantization is uniform, and the initial numbers owned by the agents are reals (as opposed to being integers). The paper [18] shows that the quantized gossip algorithm works for a particular choice of the updating parameter, although it conjunctures that this result is true for a wide range of updating parameters. A related paper on quantized consensus gives a synchronous algorithm in order to reach a consensus with arbitrary precision, at the cost of not preserving the average of the initial numbers [19].

Part I of the current work starts with proving the above-mentioned property of quantized consensus. More precisely, a weighted connected graph is considered together with a set of scalars sitting on its vertices. The weight of each edge represents the probability of establishing a communication between its corresponding vertices through the updating procedure. It is shown that the quantized consensus is reached under the stochastic gossip algorithm proposed in [18], for a range of updating parameters. The results hold true for any arbitrary quantizer, including uniform and logarithmic ones. Some elegant properties of the system in the steady state (after reaching the consensus) are subsequently derived. The second part of the paper deals with the expect value of the time at which the consensus is reached [20]. This quantity (in the worst case) is upper and lower bounded in terms of the weighted Laplacian of the graph. A convex optimization is then proposed to investigate what set of weighes on the edges results in the smallest convergence time.

This paper is organized as follows. Some preliminaries are presented in Section II, and the problem is formulated accordingly. The convergence proof is provided in Section III for uniform quantizers, and is generalized to arbitrary quantizers in Section IV. The results are illustrated in Section V through simulations. Some concluding remarks are finally

This work has been supported by AFOSR and Air Force MURI.

The authors are with the Department of Control and Dynamical Systems, California Institute of Technology, Pasadena, USA (emails: lavaei@cds.caltech.edu; murray@cds.caltech.edu).

drawn in Section VI.

II. Problem formulation

Consider a connected weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$, where:

- $\mathcal{V} := \{v_1, v_2, ..., v_\nu\}$ is the set of vertices of $\mathcal{G}$;
- $\mathcal{E}$ is the set of edges of $\mathcal{G}$;
- $\mathcal{P} := \{p_{ij}\}_{i,j}$ is the set of weights assigned to the edges of $\mathcal{G}$.

Assume that:

- $\sum_{i,j \in \boldsymbol{\nu}} p_{ij}$ is equal to 2, where $\boldsymbol{\nu} := \{1, 2, ..., \nu\}$.
- p_{ij}, $i, j \in \boldsymbol{\nu}$, is equal to zero if $(i, j) \notin \mathcal{E}$; otherwise, it is strictly positive. In particular, $p_{11}, p_{22}, ..., p_{\nu\nu}$ are all equal to zero.

The scalar p_{ij} associated with the edge (i, j) represents the probability of choosing the edge (i, j) when an edge of $\mathcal{G}$ is to be picked at random. Suppose that a real number x_i has been assigned to the vertex v_i, for any $i \in \boldsymbol{\nu}$. Let $q(x) : \Re \to \Re$ be a given quantization operator, which can be, for instance, a logarithmic or constant quantizer. In what follows, a quantized gossip algorithm is presented [18].

Algorithm 1:

Step 1: Given a positive real ε, set $k = 0$. Define $x_i[0] := x_i$, for any $i \in \boldsymbol{\nu}$.

Step 2: Pick an edge of $\mathcal{G}$ at random.

Step 3: Suppose that the ending vertices of the edge selected in step 2 possess the values $x_i[k]$ and $x_j[k]$. Perform the following updates:

$$\begin{aligned} x_i[k+1] &= x_i[k] + \varepsilon \times \big(q(x_j[k]) - q(x_i[k])\big), \\ x_j[k+1] &= x_j[k] + \varepsilon \times \big(q(x_i[k]) - q(x_j[k])\big), \\ x_q[k+1] &= x_q[k], \quad \forall q \in \boldsymbol{\nu} \backslash \{i, j\} \end{aligned} \tag{1}$$

Step 4: Increase k by 1 and jump to step 2.

Let the short-hand notation:

$$\mathbf{x}[k] = \left[\begin{array}{cccc} x_1[k] & x_2[k] & \cdots & x_\nu[k] \end{array} \right], \quad k \in \mathsf{Z} \tag{2}$$

be used hereafter.

The next definition is extracted from [18].

Definition 1: Given a quantization-based protocol acting on $\mathcal{G}(\mathcal{V}, \mathcal{E})$ (e.g. the deterministic gossip algorithm), assume that the vector $\mathbf{x}[k]$ denotes the values on the vertices of $\mathcal{G}$ at time k, obtained using this protocol. It is said that the *quantized consensus* is reached for the graph $\mathcal{G}$ under the protocol $\mathcal{C}$ if for any arbitrary initial state $\mathbf{x}_i[0] \in \Re^\nu$, there exists a natural number k_0 such that:

$$|x_i[k] - x_{\text{ave}}| < 1, \quad \forall k \geq k_0, \; \forall i \in \boldsymbol{\nu} \tag{3}$$

where $x_{\text{ave}} := \frac{x_1[0] + x_2[0] + \cdots + x_\nu[0]}{\nu}$.

In line with the above definition, if the protocol $\mathcal{C}$ is stochastic (e.g. Algorithm 1), one would say that the quantized consensus is reached *almost surely* if there exists a number $k_0 \in \mathsf{N}$, almost surely, for which the inequality (3) holds. In the rest of the paper, the short name *consensus* will be used for *quantized consensus*.

It is shown in [18] that if the quantizer $q(x)$ is uniform, the consensus is reached almost surely for the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ under the stochastic gossip algorithm described above (Algorithm 1), provided $\varepsilon = 0.5$. This paper also conjectures that the same result holds true for any positive $\varepsilon < 0.5$, while it may not be necessarily true for $\varepsilon > 0.5$ (as simulation confirms). The objective of this part of the paper is twofold. This conjecture is to be proved first. The results are then to be extended to general quantizers.

III. Main results

In the remainder of the paper, assume that $\varepsilon \in (0, 0.5]$ (unless otherwise stated). Let x_{max} and x_{min} be defined as:

$$x_{max} := \max_{i \in \boldsymbol{\nu}} \lceil x_i \rceil, \quad x_{min} := \min_{i \in \boldsymbol{\nu}} \lfloor x_i \rfloor \tag{4}$$

where $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ stand for the ceiling and floor operators, respectively.

Definition 2: Define $\mathcal{S}$ to be the set of all ν-tuple $(\alpha_1, \alpha_2, ..., \alpha_\nu)$ such that $\alpha_i \in [x_{min}, x_{max}]$ and, in addition, $(\alpha_i - x_i)$ is an integer multiple of ε, for any $i \in \boldsymbol{\nu}$.

Algorithm 1 is stochastic in the sense that an edge must be chosen *at random* at each time update. The deterministic version of this algorithm can be obtained by replacing its step 2 with the following:

Step 2: Pick an edge of $\mathcal{G}$ arbitrarily (at the discretion of the designer).

Let the deterministic version of Algorithm 1 be referred to as "Algorithm 2".

Theorem 1: Assume that step 2 of Algorithm 2 (i.e. picking an edge arbitrarily) can be taken in such a way that the consensus is reached for the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ under Algorithm 2. Then, the consensus is reached almost surely for the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ under Algorithm 1.

Proof: Apply Algorithm 1 to the graph $\mathcal{G}$ with the initial state $\mathbf{x}[0]$. Using induction, one can easily conclude from the equation (1) that:

i) $x_i[k]$ is always in the interval $[x_{min}, x_{max}]$, for any $k \in \mathsf{N} \cup \{0\}$ and $i \in \boldsymbol{\nu}$.
ii) The vector $\mathbf{x}[k]$ belongs to the set $\mathcal{S}$, for any $k \in \mathsf{N} \cup \{0\}$.

It can be inferred from property (ii) that $\mathbf{x}[k]$ always belongs to a finite set, while the argument k changes from 0 to infinity. Therefore, there exists a ν-tuple $\boldsymbol{\alpha} \in \mathcal{S}$ and an infinite set $\mathcal{N}_0 \subset \mathsf{N}$ such that:

$$\mathbf{x}[k] = \boldsymbol{\alpha}, \quad \forall k \in \mathcal{N}_0 \tag{5}$$

By assumption, the consensus is reached for the graph $\mathcal{G}$ under Algorithm 2, by commencing from the initial state $\boldsymbol{\alpha}$. Assume that the consensus is reached at time μ, and that the edge e_j is chosen at time j in step 2 of Algorithm 2, for $j = 1, 2, ..., \mu$. One can observe that:

- If μ edges of $\mathcal{G}$ are to be chosen successively at random, the probability of ending up with the sequence $(e_1, e_2, ..., e_\mu)$ is positive.
- The vector $\boldsymbol{\alpha}$ shows up on the vertices of $\mathcal{G}$ infinitely many times if Algorithm 1 is applied to the graph with the initial state $\mathbf{x}[0]$.

The two observations pointed out above substantiate that almost surely there exists an η in $\mathcal{N}_0$ such that the edge e_j is selected in step 2 of Algorithm 1 at time $\eta + j$, for any $j \in \{1, 2, ..., \mu\}$. This implies that the consensus is reached at the time $k_0 := \eta + \mu$. ■

Theorem 1 states that in order to prove the convergence of the stochastic gossip algorithm, it suffices to show that of its deterministic version. Hence, this converse statement will be proved in the sequel. Throughout the rest of this section, assume that $q(x)$ is a uniform quantizer, i.e. it rounds any real x to its nearest integer (by convention, assume that $q(r + 0.5) = r,\ \forall r \in \mathsf{Z}$). The results will be extended to a general quantizer $q(x)$ in the subsequent section.

Definition 3: Define the following quantities:

$$\begin{aligned} \eta_1 &:= \max\left\{\frac{2k+1}{2}\Big| k \in \mathsf{Z},\ \frac{2k+1}{2} \leq x_{ave}\right\}, \\ \eta_2 &:= \min\left\{\frac{2k+1}{2}\Big| k \in \mathsf{Z},\ \frac{2k+1}{2} \geq x_{ave}\right\} \end{aligned} \tag{6}$$

Definition 4: Let $\mathcal{S}_o$ and $\mathcal{S}_o(\mu)$, $\mu \in \Re$, be defined as follows:

$$\mathcal{S}_o := \left\{(\alpha_1, \alpha_2, ..., \alpha_\nu) \in \mathcal{S} \middle|\ \alpha_i \in (\eta_1, \eta_2],\quad \forall i \in \boldsymbol{\nu}\right\}$$

and:

$$\mathcal{S}_o(\mu) := \left\{(\alpha_1, ..., \alpha_\nu) \in \mathcal{S} \middle|\ \alpha_i \in (\mu - \varepsilon, \mu + \varepsilon],\ \forall i \in \boldsymbol{\nu}\right\}$$

Definition 5: Define the distance function $d_\varepsilon(\cdot, \mathcal{S}_o) : \mathcal{S} \to \mathsf{Z}$ as:

$$d_\varepsilon(\boldsymbol{\alpha}, \mathcal{S}_o) := \min_{\boldsymbol{\beta} \in \mathcal{S}_o} \frac{|\boldsymbol{\alpha} - \boldsymbol{\beta}|_1}{\varepsilon}, \quad \forall \boldsymbol{\alpha} \in \mathcal{S} \tag{7}$$

where $|\cdot|_1$ denotes the L_1 norm, and Z is the set of integers. In the same way, define $d_\varepsilon(\boldsymbol{\alpha}, \mathcal{S}_o(\mu))$ for any real μ.

Note that $d_\varepsilon(\boldsymbol{\alpha}, \mathcal{S}_o)$ is equal to zero if $\boldsymbol{\alpha} \in \mathcal{S}_o$.

Lemma 1: Apply Algorithm 1 to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ with the initial state $\mathbf{x}[0]$.

- Suppose that $\mathbf{x}[k]$ belongs to the set $\mathcal{S}_o$ for some nonnegative integer k. The equality $\mathbf{x}[k+1] = \mathbf{x}[k]$ holds. In other words, each element of $\mathcal{S}_o$ is an equilibrium point of the discrete-time system.
- Assume that $\mathbf{x}[k]$ belongs to the set $\mathcal{S}_o(r + 0.5)$, for some integers k and r. The state $\mathbf{x}[k+1]$ is in the set $\mathcal{S}_o(r+0.5)$ as well. In other words, this set is invariant under the underlying algorithm.

Proof: The proof is straightforward, and is omitted for brevity. ■

Lemma 2: Apply Algorithm 1 to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ with the initial state $\mathbf{x}[0]$. Given $r \in \mathsf{Z}$, the following inequality holds for any nonnegative integer k:

$$d_\varepsilon(\mathbf{x}[k+1], \mathcal{S}_o(r + 0.5)) \leq d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(r + 0.5)) \tag{8}$$

Proof: Assume that the edge (i, j) is chosen at the $(k+1)^{\text{th}}$ time update, and that $x_i[k] \leq x_j[k]$. There are a number of possibilities as follows:

- $x_i[k] - r - 0.5 > 0$ *and* $x_j[k] - r - 0.5 > 0$*:* It can be easily shown that:

$$x_i[k], x_j[k], x_i[k+1], x_j[k+1] > r + 0.5 \tag{9}$$

The above inequalities together with the equality

$$x_i[k] + x_j[k] = x_i[k+1] + x_j[k+1] \tag{10}$$

allow us to conclude that

$$d_\varepsilon(\mathbf{x}[k+1], \mathcal{S}_o(r+0.5)) = d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(r+0.5)). \tag{11}$$

- $x_i[k] - r - 0.5 \leq 0$ *and* $x_j[k] - r - 0.5 \leq 0$*:* It is easy to observe that:

$$x_i[k], x_j[k], x_i[k+1], x_j[k+1] \leq r + 0.5 \tag{12}$$

This leads to the equality (11) (as before).

- $x_i[k] - r - 0.5 \leq$ *and* $x_j[k] - r - 0.5 > 0$*:* Similar to the previous cases, it can be shown that the inequality (8) holds. ■

Lemma 3: Given $r \in \mathsf{Z}$, apply Algorithm 2 to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with the initial state $\mathbf{x}[0]$. At each time update $k \in \mathsf{N}$, select an edge of the graph (in step 2 of the algorithm) such that the function $d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(r + 0.5))$ is minimized. There exists a natural number k_0 for which either of the following cases occurs:

i) $\mathbf{x}[k]$ is in the invariant set $\mathcal{S}_o(r + 0.5)$, for any $k \geq k_0$.
ii) $x_1[k] - r - 0.5, x_2[k] - r - 0.5, ..., x_\nu[k] - r - 0.5$ are either all *negative* or all *strictly positive* for all $k \geq k_0$.

Proof: Since $d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(r+0.5))$ is a nonnegative integer-valued decreasing function (by Lemma 2), there exists a number k_0 with the property:

$$d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(r + 0.5)) = d_\varepsilon(\mathbf{x}[k_0], \mathcal{S}_o(r + 0.5)), \quad \forall k \geq k_0 \tag{13}$$

If $d_\varepsilon(\mathbf{x}[k_0], \mathcal{S}_o(r + 0.5)) = 0$, then case (i) explained in the statement of the lemma definitely occurs. It is desired to prove that if $d_\varepsilon(\mathbf{x}[k_0], \mathcal{S}_o(r + 0.5)) \neq 0$, then case (ii) takes place. To this end, notice that if $x_i[k] - r - 0.5$, $\forall i \in \boldsymbol{\nu}$, are negative (strictly positive) for some proper k, so are $x_i[k + 1] - r - 0.5$, $\forall i \in \boldsymbol{\nu}$. This implies that it suffices to prove case (ii) only for $k = k_0$.

To prove by contradiction, assume that there exist two integers $i, j \in \boldsymbol{\nu}$ such that:

$$x_i[k_0] > r + 0.5, \quad x_j[k_0] \leq r + 0.5 \tag{14}$$

Since the graph $\mathcal{G}$ is connected, the above two inequalities yield that there are two integers $\mu_1, \mu_2 \in \boldsymbol{\nu}$ subject to:

- (μ_1, μ_2) is an edge of the graph $\mathcal{G}$.
- $x_{\mu_1}[k_0] > r + 0.5$ and $x_{\mu_2}[k_0] \leq r + 0.5$.

If $x_{\mu_1}[k_0] > r + 0.5 + \varepsilon$ or $x_{\mu_2}[k_0] \leq r + 0.5 - \varepsilon$, then following the proof of Lemma 2, one can conclude that choosing the edge (μ_1, μ_2) at time $k_0 + 1$ in step 2 of Algorithm 2 results in the reduction of $d_\varepsilon(\mathbf{x}[k_0], \mathcal{S}_o(r+0.5))$, i.e.:

$$d_\varepsilon(\mathbf{x}[k_0 + 1], \mathcal{S}_o(r + 0.5)) < d_\varepsilon(\mathbf{x}[k_0], \mathcal{S}_o(r + 0.5)) \tag{15}$$

which is impossible in light of the equality (13). Thus:

$$r + 0.5 < x_{\mu_1}[k] \leq r + 0.5 + \varepsilon, \tag{16a}$$

$$r + 0.5 \geq x_{\mu_2}[k] > r + 0.5 - \varepsilon \tag{16b}$$

Consider an arbitrary vertex connected to v_{μ_2}, and denote it with v_{μ_3} (if such a vertex does not exist, find a vertex connected to v_{μ_1} instead). It is desired to prove that:

$$r + 0.5 - \varepsilon < x_{\mu_3}[k_0] \leq r + 0.5 + \varepsilon \qquad (17)$$

To this end, consider the following scenarios:

- $x_{\mu_3}[k_0]$ *is greater than* $r + 0.5$: It results from (16b) that if the inequality $x_{\mu_3}[k_0] \leq r + 0.5 + \varepsilon$ does not hold, then choosing the edge (μ_2, μ_3) at time $k_0 + 1$ through Algorithm 2 will reduce the storage function $d_\varepsilon(\mathbf{x}[k_0], \mathcal{S}_o(r + 0.5))$, which is impossible by (13).
- $x_{\mu_3}[k_0]$ *is less than or equal to* $r + 0.5$: If the relation $r + 0.5 - \varepsilon < x_{\mu_3}[k_0]$ does not hold, run step 2 of Algorithm 2 at times $k_0 + 1$ and $k_0 + 2$ as follows:
  - At time $k_0 + 1$, choose the edge (μ_1, μ_2) which gives the updates (in light of (16)):

    $$\begin{aligned} x_{\mu_1}[k_0 + 1] &= x_{\mu_1}[k_0] - \varepsilon, \\ x_{\mu_2}[k_0 + 1] &= x_{\mu_2}[k_0] + \varepsilon \end{aligned} \qquad (18)$$

    Therefore:

    $$d_\varepsilon(\mathbf{x}[k_0+1], \mathcal{S}_o(r+0.5)) = d_\varepsilon(\mathbf{x}[k_0], \mathcal{S}_o(r+0.5)) \qquad (19)$$
  - At time $k_0 + 2$, choose the edge (μ_2, μ_3). The equation (18) leads to:

    $$\begin{aligned} r + 0.5 &< x_{\mu_2}[k_0 + 1] \leq r + 0.5 + \varepsilon, \\ x_{\mu_3}[k_0 + 1] &= x_{\mu_3}[k_0] \leq r + 0.5 - \varepsilon \end{aligned} \qquad (20)$$

    Thus (after some manipulations):

    $$\begin{aligned} & d_\varepsilon(\mathbf{x}[k_0 + 2], \mathcal{S}_o(r + 0.5)) \leq \\ & d_\varepsilon(\mathbf{x}[k_0 + 1], \mathcal{S}_o(r + 0.5)) - 1 \end{aligned} \qquad (21)$$

    which is impossible by the equation (13)

  This shows the validity of the inequality (17). Since the graph is connected, there is a path from v_{μ_2} to any other vertex in $\mathcal{V}$. One can continue the argument made above (on the vertex v_{μ_3}) for the vertices of this path successively to conclude that:

  $$r + 0.5 - \varepsilon < x_i[k_0] \leq r + 0.5 + \varepsilon, \quad \forall i \in \boldsymbol{\nu} \qquad (22)$$

The above inequality signifies that $d_\varepsilon(\mathbf{x}[k_0], \mathcal{S}_o(r + 0.5))$ is equal to zero, while it was earlier assumed to be nonzero. This contradiction completes the proof. ■

Theorem 2: Apply Algorithm 2 to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with the initial state $\mathbf{x}[0]$. Step 2 of this algorithm (i.e. selecting an edge arbitrarily) can be taken in such a way that there exists a positive number k_1 for which one of the following cases takes place:

i) $\mathbf{x}[k]$ belongs to the set $\mathcal{S}_o$, for any $k \geq k_1$.
ii) $\mathbf{x}[k]$ belongs to the set $\mathcal{S}_o(\eta_1)$, for any $k \geq k_1$.
iii) $\mathbf{x}[k]$ belongs to the set $\mathcal{S}_o(\eta_2)$, for any $k \geq k_1$.

Proof: Define the storage functions:

$$\begin{aligned} V_1[k] &:= d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(\eta_1)), \\ V_2[k] &:= d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(\eta_2)) \end{aligned} \qquad (23)$$

In the course of taking step 2 of Algorithm 2, i.e. choosing an edge at one's discretion, select an edge at each time update k so that the function $V_1[k]$ is minimized (as explained in the statement of Lemma 3). Halt this algorithm at a time k_0, where $V_1[k]$ reaches its minimum and remains constant. By the preceding lemma, one of the following cases happens:

- $\mathbf{x}[k]$ *is in the invariant set* $\mathcal{S}_o(\eta_1)$*, for any* $k \geq k_0$*:* If this is the case, the proof is complete.
- $x_1[k]-\eta_1, x_2[k]-\eta_1, ..., x_\nu[k]-\eta_1$ *are all negative, for any* $k \geq k_0$*:* Since x_{ave} is greater or equal to η_1, and it is also the average of the numbers $x_1[k], ..., x_\nu[k]$, this case is ruled out, unless $x_1[k] = x_2[k] = \cdots = x_\nu[k] = x_{ave} = \eta$. Nevertheless, this implies that $\mathbf{x}[k] \in \mathcal{S}_o(\eta_1)$.
- $x_1[k] - \eta_1, x_2[k] - \eta_1, ..., x_\nu[k] - \eta_1$ *are all strictly positive, for any* $k \geq k_0$*:* At time $k = k_0$, ignore the mission of minimizing $V_1[k]$, and after this time take step 2 of Algorithm 2 so that the Lyapunov function $V_2[k]$ is minimized at each time update. Notice that since all entries of $\mathbf{x}[k_0]$ are greater than η_1, they can never go beyond this limit at a future time. Using Lemma 3, it can be argued that there exists a natural number $k_1 > k_0$ for which one of the following cases occurs:
  - $\mathbf{x}[k]$ *is in the invariant set* $\mathcal{S}_o(\eta_2)$*, for any* $k \geq k_0$*:* If this is the case, the proof is complete.
  - $x_1[k]-\eta_2, x_2[k]-\eta_2, ..., x_\nu[k]-\eta_2$ *are all negative, for any* $k \geq k_0$*:* It follows from this case that $\mathbf{x}[k]$ belongs to the set $\mathcal{S}_o$, for any $k \geq k_1$.
  - $x_1[k]-\eta_2, x_2[k]-\eta_2, ..., x_\nu[k]-\eta_2$ *are all strictly positive, for any* $k \geq k_0$*:* This case can be simply ruled out, by adopting an argument similar to the one made above. ■

Theorems 1 and 2 give rise to the conclusion that the consensus is reached almost surely for the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ under the quantized stochastic gossip algorithm (i.e. Algorithm 1), for any $\varepsilon \in (0, 0.5]$.

Given a set $\mathcal{M} \in \Re^\nu$, define the diameter of $\mathcal{M}$ to be the supremum of the infinity norm of the distance between every two points in $\mathcal{M}$.

Remark 1: Definition 1 states that if the consensus is reached at time k_0, the state $\mathbf{x}[k]$ belongs to the box $[x_{\text{ave}} - 1, x_{\text{ave}} + 1]^\nu$, for any $i \in \boldsymbol{\nu}$ and $k \geq k_0$. In contrast, it can be deduced from Theorems 1 and 2 that there exists a positive integer $k_1 \geq k_0$ such that $\mathbf{x}[k]$, $\forall k \geq k_1$, belongs to one of the sets $\mathcal{S}_o$, $\mathcal{S}_o(\eta_1)$, or $\mathcal{S}_o(\eta_2)$. In this regard, two points can be made as follows:

- The diameter of the set given by Definition 1 is equal to 2, whereas that of each of the sets $\mathcal{S}_o$, $\mathcal{S}_o(\eta_1)$ and $\mathcal{S}_o(\eta_2)$ is at most 1. This implies that a better definition of consensus can be presented in terms of these sets.
- If $\mathbf{x}[k]$ in the steady state (for large enough k's) is not constant (almost surely) and can oscillate, it should then belong to either of $\mathcal{S}_o(\eta_1)$ or $\mathcal{S}_o(\eta_2)$, which are both of diameter 2ε (see Lemma 1). Note that the diameter of these sets can become arbitrarily small by rendering an appropriate ε. This implies that running the gossip

algorithm for a small ε either makes the steady state constant or permits it to oscillate while belonging to a set of a small diameter (2ε). In the latter case, $x_i[k]$ can oscillate between only two numbers of difference ε (due to the definition of $\mathcal{S}_o(\mu)$, $\mu \in \Re$).

To clarify Remark 1, consider the nominal values $x_{ave} = 10.6$ and $\varepsilon = 0.2$. The definition of consensus borrowed from [18] states that there exists a positive integer k_0 such that:

$$9.6 < x_1[k], ..., x_\nu[k] < 11.6, \quad \forall k \geq k_0 \tag{24}$$

In contrast, Theorem 2 asserts that there exists a number k_1 so that:

$$10.4 < x_1[k], ..., x_\nu[k] \leq 10.8, \quad \forall k \geq k_1 \tag{25}$$

or:

$$10.5 < x_1[k], ..., x_\nu[k] \leq 11.5, \quad \forall k \geq k_1 \tag{26}$$

(note that case (iii) in Theorem 2 is ruled out in this example, as the average of the entries of $\mathbf{x}[k]$ cannot be smaller that all entries of $\mathbf{x}[k]$). Comparing (24) with (25) and (26), one can simply observe that a precise description of the steady state values on the vertices of $\mathcal{G}$ is delineated by (25) and (26). Besides, notice that if $9.6 < x_1[k], ..., x_\nu[k] < 11.6$ for some integer k, it may not be true that $9.6 < x_1[k+1], ..., x_\nu[k+1] < 11.6$ (because this region does not correspond to an invariant set in general, whereas $\mathcal{S}_o(10.5)$ and $\mathcal{S}_o$ are both invariant).

IV. Generalization to arbitrary quantizers

Let $q(x) : \Re \rightarrow \Re$ be a general quantization operator characterized as follows:

$$q(x) = \begin{cases} L_i & \text{if} \quad x \in [L_i, \bar{L}_i] \\ L_{i+1} & \text{if} \quad x \in (\bar{L}_i, L_{i+1}) \end{cases} \qquad \forall i \in \mathsf{Z} \tag{27}$$

where $\{L_i\}_{-\infty}^{\infty}$ is a monotonically increasing sequence of integers representing the quantization levels, and:

$$\bar{L}_i := \frac{L_i + L_{i+1}}{2}, \quad \forall i \in \mathsf{Z} \tag{28}$$

The scalar quantities L_i and $\bar{L}_i$ will be referred to as *level* and *splitting level*, respectively. The results presented in the preceding section can be readily extended, provided Definitions 1, 3 and 4 are expressed in the general case. This is carried out in the following.

Revised Definition 1: Given a quantization-based protocol acting on $\mathcal{G}(\mathcal{V}, \mathcal{E})$, denote with $\mathbf{x}[k]$, $k \in \mathsf{N} \cup \{0\}$, the vector of values on the vertices of $\mathcal{G}$ at time k, obtained using this protocol. It is said that the (quantized) consensus is reached for the graph $\mathcal{G}$ under the protocol $\mathcal{C}$ if for any arbitrary initial state $\mathbf{x}[0] \in \Re^\nu$, there exist a natural number k_0 and an integer μ such that either of the following sets of relations holds:

$$\begin{cases} \sum_{i=1}^{\nu} x_i[k] = \sum_{i=1}^{\nu} x_i[0] \\ x_j[k] \in [L_\mu, L_{\mu+1}] \end{cases} \qquad \forall k \geq k_0, \ \forall j \in \boldsymbol{\nu} \tag{29}$$

or:

$$\begin{cases} \sum_{i=1}^{\nu} x_i[k] = \sum_{i=1}^{\nu} x_i[0] \\ x_j[k] \in (\bar{L}_\mu, \bar{L}_{\mu+1}] \end{cases} \qquad \forall k \geq k_0, \ \forall j \in \boldsymbol{\nu} \tag{30}$$

Note that the above definition presents a more comprehensive description of consensus, compared to Definition 1 (see the discussion given in Remark 1). Roughly speaking, the revised version of Definition 1 states that the consensus is reached if the numbers on the vertices of the graph eventually lie between two consecutive levels or two consecutive splitting levels.

Revised Definition 3: Define η_1 and η_2 to be:

$$\begin{aligned} \eta_1 &= \max_{i \in \mathsf{Z}} \bar{L}_i \quad \text{s.t.} \quad \bar{L}_i \leq x_{\text{ave}}, \\ \eta_2 &= \min_{j \in \mathsf{Z}} \bar{L}_j \quad \text{s.t.} \quad \bar{L}_j \geq x_{\text{ave}} \end{aligned} \tag{31}$$

Revised Definition 4: Let $\mathcal{S}_o(\bar{L}_i)$, $i \in \mathsf{Z}$ be defined as the set of all ν-tuple $(\alpha_1, \alpha_2, ..., \alpha_\nu) \in \mathcal{S}$ such that:

$$\alpha_j \in \left(\bar{L}_i - \varepsilon(L_{i+1} - L_i), \bar{L}_i + \varepsilon(L_{i+1} - L_i)\right], \quad \forall j \in \boldsymbol{\nu} \tag{32}$$

It is noteworthy that other definitions presented in the previous section carry over to the general case, such as the definitions of $\mathcal{S}$, $\mathcal{S}_o$ and $d_\varepsilon(\cdot, S_o(\mu))$. Moreover, the assumption $\varepsilon \in (0, 0.5]$ remains unchanged.

One can adopt an approach similar to the one proposed in the preceding section to prove all lemmas and theorems (presented earlier) in the general case. This leads to the conclusions that the consensus is reached almost surely for the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ under Algorithm 1, and that $\mathbf{x}[k]$ belongs to one of the invariant sets $\mathcal{S}_o$, $\mathcal{S}_o(\eta_1)$ or $\mathcal{S}_o(\eta_2)$, for large enough k's .

V. Simulation results

Consider a complete graph $\mathcal{G}$ with $\nu = 40$ and, for simplicity, assume that all edges possess the same weight equal to $\frac{2}{\nu(\nu-1)}$. Let the initial values sitting on the vertices of $\mathcal{G}$ be uniformly distributed in the box $[0, 100]^\nu$. We wish to observe how these values evolve under the quantized stochastic gossip algorithm (i.e. Algorithm 1). For this purpose, assume that the quantization is uniform, and that $\varepsilon = 0.2$. Two sets of initial states have been randomly generated, which are spelled out below:

- As the first trial, the initial values randomly generated are depicted in Figure 1. Note that the x-axis of this plot shows the index i changing from 1 to 40, and the y-axis shows the corresponding value of $x_i[0]$. The time k_1 introduced in Theorem 2 turns out to be equal to 770. The final values at this time are plotted in Figure 2. Since these numbers are spread in the interval $[52.5, 53.5]$, the point $\mathbf{x}[k_1]$ belongs to the set $\mathcal{S}_o$ (see Theorem 2). This implies that the steady-state of the vector $\mathbf{x}[k]$ is fixed, i.e. $\mathbf{x}[k] = \mathbf{x}[k_1]$, for any $k \geq k_1$. The storage function $d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o)$ is sketched in Figure 3 to illustrate how it attenuates to zero in a (non-strictly) decreasing way. This is in accordance with Lemma 2.

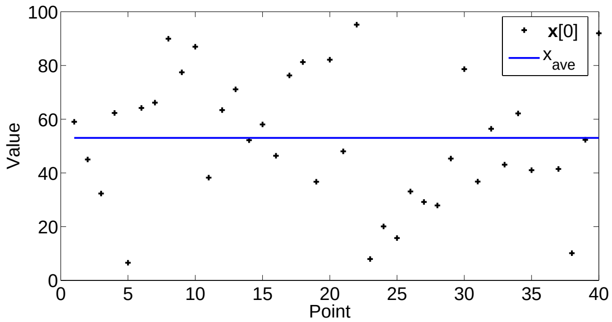

Fig. 1. The initial values on the vertices of the graph $\mathcal{G}$ for the first trial.

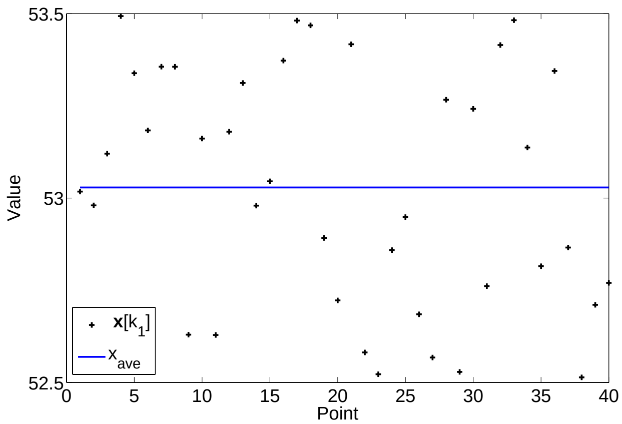

Fig. 2. The final values on the vertices of the graph $\mathcal{G}$ (at the time k_1) for the first trial.

- As the second trial, the initial values randomly generated are shown in Figure 4. The corresponding final values at the time $k_1 = 555$ are depicted in Figure 5. This figure demonstrates that $\mathbf{x}[k_1]$ belongs to the set $\mathcal{S}_o(\eta_1)$, rather than $\mathcal{S}_o$. This confirms the results of Theorem 2. Therefore, the steady-state of the vector $\mathbf{x}[k]$ is not fixed, and this vector can oscillate. However, $x_i[k]$, $i \in \boldsymbol{\nu}$, can take only two possible values with the difference $\varepsilon = 0.2$, in light of the definition of $\mathcal{S}_o(\mu)$. The storage function $d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(\eta_1))$ is plotted in Figure 6 to illustrate the convergence rate of Algorithm 1.

VI. Conclusions

This paper deals with the distributed averaging problem over a connected weighted graph. At each time update, an edge of the graph is to be chosen with the probability equal to its weight, and the values on the ending vertices are to be updated in terms of the quantized data of each other. A quantized stochastic gossip algorithm was proposed in a recent paper, which was shown to work in a particular case. In this part of the current paper, it is proved that the quantized consensus is reached in the general case using this algorithm. The quantizer can be, for instance, constant or logarithmic. Some properties of the steady-state numbers sitting on the vertices of the graph are obtained, which elegantly describe the steady-state behavior of the system. In the second part of the paper, the expected value of the time at which the consensus is reached will be lower and upper bounded in term of the topology of the graph, particularly its weighted Laplacian matrix.

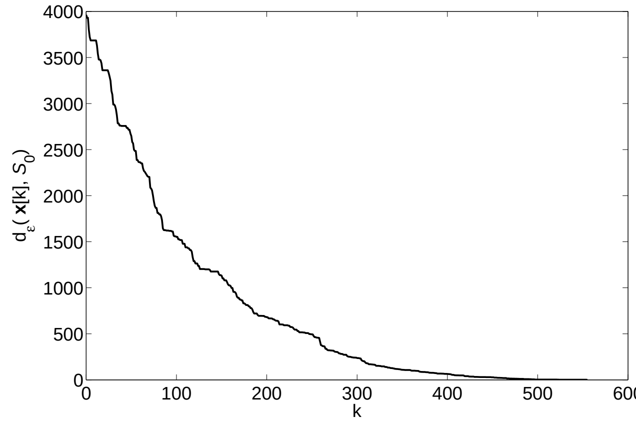

Fig. 3. The storage function $d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o)$ for the first trial.

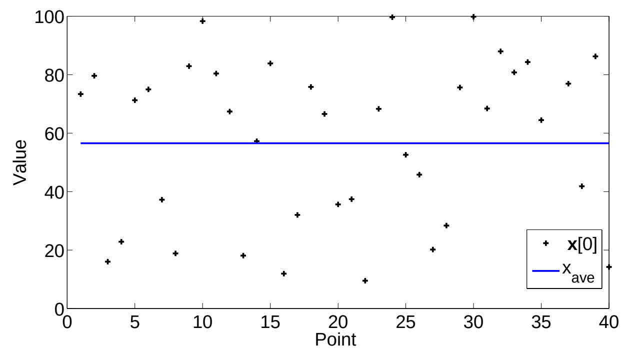

Fig. 4. The initial values on the vertices of the graph $\mathcal{G}$ for the second trial.

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

[3] G. Tel, "Introduction to distributed algorithms," Cambridge University Press, 2000.

[4] N. A. Lynch, "Distributed algorithms," Morgan Kaufmann Publishers, Inc., San Francisco, CA, 1996.

[5] D. P. Bertsekas and J. N. Tsitsiklis, "Parallel and distributed computation: Numerical methods," Belmont, MA: Athena Scientific, 1997.

[6] Y. Rabani, A. Sinclair and R. Wanka,"Local divergence of Markov chains and the analysis of iterative load-balancing schemes," in *Proceedings of IEEE Conference on Foundations of Computer Science*, 1998.

[7] Y. Kuramoto, "Chemical oscillators, waves, and turbulance," Springer-Verlag, Berlin, 1984.

[8] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, pp. 268-276, 2001.

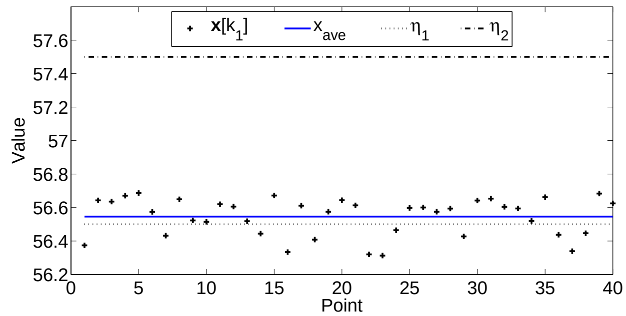

Fig. 5. The final values on the vertices of the graph $\mathcal{G}$ (at the time k_1) for the second trial.

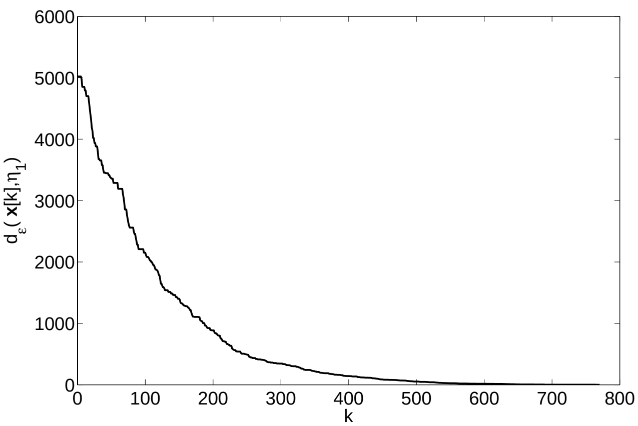

Fig. 6. The storage function $d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(\eta_1))$ for the second trial.

[9] A. V. Savkin, "Coordinated collective motion of groups of autonomous mobile robots: Analysis of Vicsekś model," *IEEE Transactions on Automatic Control*, vol. 49, no. 6, pp. 981982, 2004.

[10] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401420, 2006.

[11] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 9881001, 2003.

[12] A. Speranzon, C. Fischione and K.H. Johansson, "Distributed and collaborative estimation over wireless sensor networks," in *Proceedings of the 45th IEEE Conference on Decision and Control*, 2006.

[13] J. Tsitsiklis, "Problems in decentralized decision making and computation," PhD thesis, Dept. of Electrical Engineering and Computer Science, M.I.T., Boston, MA, 1984.

[14] S. Boyd, A. Ghosh, B. Prabhakar and D. Shah, "Analysis and optimization of randomized gossip algorithms," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2004.

[15] S. Boyd, A. Ghosh, B. Prabhakar and D. Shah , "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508-2530, 2006.

[16] F. Benezit, A. G. Dimakis, P. Thiran and M. Vetterli, "Gossip along the way: Order-optimal consensus through randomized path averaging," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, 2007.

[17] A. Kashyap, T. Basara and R. Srikanta, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192-1203, 2007.

[18] P. Frasca, R. Carli, F. Fagnani and S. Zampieri, "Average consensus by gossip algorithms with quantized communication," to appear in *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008 (available online at http://calvino.polito.it/∼frasca/publ.html).

[19] A. Censi and Richard M. Murray, "A biologically inspired approach to real-valued average consensus over quantized channels with arbitrary deterministic accuracy," submitted to *Proceedings of the 2009 American Control Conference*, 2009.

[20] J. Lavaei and R. M. Murray, "On quantized consensus by means of gossip algorithm – Part II: convergence time," submitted to *Proceedings of the 2009 American Control Conference*, 2009 (available online at www.cds.caltech.edu/∼lavaei/publication/consensus2.pdf).