

On Quantized Consensus by Means of Gossip Algorithm – Part II: Convergence Time

Javad Lavaei and Richard M. Murray

Abstract—This paper deals with the distributed averaging problem over a connected network of agents, subject to a quantization constraint. It is assumed that at each time update, only a pair of agents can update their own numbers in terms of the quantized data being exchanged. The agents are also required to communicate with one another in a stochastic fashion. In the first part of the paper, it was shown that the quantized consensus is reached by means of a stochastic gossip algorithm proposed in a recent paper, for any arbitrary quantization. The current part of the paper considers the expected value of the time at which the quantized consensus is reached. This quantity (corresponding to the worst case) is upper and lower bounded in terms of the topology of the graph, for uniform quantization. In particular, it is shown that these bounds are related to the principal minors of the weighted Laplacian matrix. A convex optimization is also proposed to determine the set of probabilities (used to pick a pair of agents) which leads to the fast convergence of the gossip algorithm.

I. INTRODUCTION

During the past few decades, there has been a particular interest in distributed computations, which aim to compute some quantity over a network of processors in a decentralized fashion [1], [2]. The distributed averaging problem, as a particular case, is concerned with computing the average of numbers owned by the agents of a group [3], [4]. This problem has been investigated through the notion of consensus in several papers, motivated by different applications [5], [6], [7], [8], [9], [10]. For instance, the synchronization of coupled oscillators, appearing in biophysics, neurobiology, and systems biology, is studied in [5] and [6] to explore how to reach a consensus on the frequencies of all agents. Moreover, the problem of aligning the heading angles of a group of mobile agents (e.g. a flock of birds) is treated in [11]. Given a sensor network comprising a set of sensors measuring the same quantity in a noisy environment, the problem of consensus on state estimations is discussed in [12]. The consensus problem for networks of dynamic agents with fixed and switching topologies is tackled in [3], where it is shown that the convergence rate is related to the algebraic connectivity of the network. The work [13] elaborates on the relationship between the amount of information exchanged by the agents and the rate of convergence to the consensus. A more complete survey on this topic is given in the recent paper [4].

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 [14], [15]. 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 [16], [17], [18], [19]. For instance, the work [16] studies the convergence of a general randomized gossip algorithm, and derives conditions under which the algorithm converges. This paper also shows that the averaging time of a gossip algorithm depends on the second largest eigenvalue of a doubly stochastic matrix characterizing the algorithm.

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 [18] 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 [19] 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 [19] shows that the quantized gossip algorithm works for a particular choice of the updating parameter, although it conjectures 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 [20].

In this paper, 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 was shown in Part I of this work that the quantized consensus is reached under the stochastic gossip algorithm proposed in [19], for a wide range of updating parameters and any arbitrary quantizer including uniform and logarithmic ones [21]. The current part of the paper is concerned with the convergence time of the gossip algorithm. More precisely, consider the expected value of the time at which the consensus is reached, and take its maximum over all possible initial states belonging to a given hypercube. Upper and lower bounds on this quantity are provided for a uniform quantizer, which turn out to be related to the Laplacian of the weighted graph. Since

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).

these bounds are both constant multiples of some expression depending on the edge weights, the lower and upper bounds can be simultaneously minimized in order to obtain the best weights (resulting in a small convergence time). To do so, a convex optimization is proposed, which can be solved by a semidefinite program.

This paper is organized as follows. Some preliminaries are presented in Section II, and the problem is formulated accordingly. The main results on the convergence time are provided in Section III. The results are illustrated in Section IV with a numerical example. Finally, some concluding remarks are given in Section V.

II. PROBLEM FORMULATION AND PRELIMINARIES

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

- $\mathcal{V} := \{v_1, v_2, \dots, v_\nu\}$ 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:

- $\sum_{i,j \in \nu} p_{ij}$ is equal to 2, where $\nu := \{1, 2, \dots, \nu\}$.
- p_{ij} , $i, j \in \nu$, is equal to zero if $(i, j) \notin \mathcal{E}$; otherwise, it is strictly positive. In particular, $p_{11}, p_{22}, \dots, p_{\nu\nu}$ 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 $i \in \nu$. Let $q(x) : \mathbb{R} \rightarrow \mathbb{R}$ be a general quantization operator characterized as follows:

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

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

$$\bar{L}_i := \frac{L_i + L_{i+1}}{2}, \quad \forall i \in \mathbb{Z} \quad (2)$$

Note that \mathbb{Z} denotes the set of integers. In what follows, a quantized stochastic gossip algorithm is presented [19].

Algorithm 1:

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

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:

$$\begin{aligned} x_i[k+1] &= x_i[k] + \varepsilon \times (q(x_j[k]) - q(x_i[k])), \\ x_j[k+1] &= x_j[k] + \varepsilon \times (q(x_i[k]) - q(x_j[k])), \\ x_q[k+1] &= x_q[k], \quad \forall q \in \nu \setminus \{i, j\} \end{aligned} \quad (3)$$

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

Let the short-hand notation:

$$\mathbf{x}[k] = [x_1[k] \quad x_2[k] \quad \dots \quad x_\nu[k]], \quad k \in \mathbb{Z} \quad (4)$$

be used hereafter.

Definition 1: Given a quantization-based protocol acting on $\mathcal{G}(\mathcal{V}, \mathcal{E})$, denote with $\mathbf{x}[k]$, $k \in \mathbb{N} \cup \{0\}$, 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 $\mathbf{x}[0] \in \mathbb{R}^\nu$, there exist a natural number k_0 and an integer μ such that either of the following sets of relations holds:

$$\begin{cases} \sum_{i=1}^{\nu} x_i[k] = \sum_{i=1}^{\nu} x_i[0] \\ x_j[k] \in [L_\mu, L_{\mu+1}] \end{cases} \quad \forall k \geq k_0, \quad \forall j \in \nu \quad (5)$$

or:

$$\begin{cases} \sum_{i=1}^{\nu} x_i[k] = \sum_{i=1}^{\nu} x_i[0] \\ x_j[k] \in (\bar{L}_\mu, \bar{L}_{\mu+1}] \end{cases} \quad \forall k \geq k_0, \quad \forall j \in \nu \quad (6)$$

In line with the above definition, if the protocol \mathcal{C} is stochastic, one would say that the quantized consensus is reached *almost surely* if there exists a number $k_0 \in \mathbb{N}$, almost surely, for which either of the relations (5) or (6) holds. For simplicity, the short name *consensus* is used hereafter in order to refer to *quantized consensus*. A few definitions and notations will be introduced in the sequel.

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

$$x_{max} := \max_{i \in \nu} [x_i], \quad x_{min} := \min_{i \in \nu} [x_i] \quad (7)$$

Definition 3: Let η_1 and η_2 be:

$$\begin{aligned} \eta_1 &= \max_{i \in \mathbb{Z}} \bar{L}_i \quad \text{s.t.} \quad \bar{L}_i \leq x_{ave}, \\ \eta_2 &= \min_{j \in \mathbb{Z}} \bar{L}_j \quad \text{s.t.} \quad \bar{L}_j \geq x_{ave} \end{aligned} \quad (8)$$

where $x_{ave} := \frac{x_1 + x_2 + \dots + x_\nu}{\nu}$.

Definition 4: Define:

$$\mathcal{S}_o := \{(\alpha_1, \alpha_2, \dots, \alpha_\nu) \in \mathcal{S} \mid \alpha_i \in (\eta_1, \eta_2], \quad \forall i \in \nu\}$$

Furthermore, let $\mathcal{S}_o(\bar{L}_i)$, $i \in \mathbb{Z}$, be defined as the set of all ν -tuple $(\alpha_1, \alpha_2, \dots, \alpha_\nu) \in \mathcal{S}$ such that:

$$\alpha_j \in (\bar{L}_i - \varepsilon(L_{i+1} - L_i), \bar{L}_i + \varepsilon(L_{i+1} - L_i)], \quad \forall j \in \nu \quad (9)$$

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

$$d_\varepsilon(\alpha, \mathcal{S}_o) := \min_{\beta \in \mathcal{S}_o} \frac{|\alpha - \beta|_1}{\varepsilon}, \quad \forall \alpha \in \mathcal{S} \quad (10)$$

where $|\cdot|_1$ denotes the L_1 norm. Define also the distance function $d_\varepsilon(\cdot, \mathcal{S}_o(\mu))$ in the same vein.

The next result was proved in Part I of the paper [21].

Theorem 1: Given $\varepsilon \in (0, 0.5]$, apply Algorithm 1 to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ with the initial state $\mathbf{x}[0]$. There, almost surely, exists a positive number k_1 for which one of the following cases takes place:

- $\mathbf{x}[k]$ 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 $k \geq k_1$.

It is noteworthy that the above theorem proves reaching the consensus and, besides, describes the behavior of the system in the steady state. In order to study the convergence time of Algorithm 1, it is desired to find lower and upper bounds on the quantity k_1 introduced in Theorem 1. The objective of this part is twofold. First, these lower and upper bounds are to be obtained. Then, it is desired to investigate what probability distribution P results in the fastest convergence.

III. CONVERGENCE TIME

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’’. Throughout the remainder of the paper, assume that $\varepsilon \in (0, 0.5]$, and that $q(x)$ is a uniform quantizer, which rounds any real x to its nearest integer (by convention, let $q(r + 0.5) = r$, $\forall r \in \mathbb{Z}$). The results can be extended to a general quantizer in the same way.

Definition 6: Since Algorithm 1 is stochastic, the quantity k_1 is a random variable. Thus, define $t_c(\varepsilon)$ to be equal to $\max \mathcal{E}\{k_1\}$, where the maximum is taking over every initial state $\mathbf{x}[0]$ in the bounding box $[x_{\min}, x_{\max}]^\nu$ which belongs to none of the steady-state sets \mathcal{S}_o , $\mathcal{S}_o(\eta_1)$ and $\mathcal{S}_o(\eta_2)$ (note that $\mathcal{E}\{\cdot\}$ denotes the expectation operator).

The term $t_c(\varepsilon)$ indeed quantifies the expected value of the convergence time in the worst case. This section aims to characterize $t_c(\varepsilon)$ in terms of x_{\min} , x_{\max} , ε , and the topology of the graph (together with the probabilities assigned to its edges).

Definition 7: Given an integer r , apply Algorithm 1 (or Algorithm 2) to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ with the initial state $\mathbf{x}[0]$. The process of choosing an edge at time k ($k \in \mathbb{N}$) in step 2 of Algorithm 1 (or Algorithm 2) is said to be a *positive action* with respect to (w.r.t.) $d_\varepsilon(\cdot, \mathcal{S}_o(r + 0.5))$ if the inequality:

$$d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(r + 0.5)) < d_\varepsilon(\mathbf{x}[k - 1], \mathcal{S}_o(r + 0.5)) \quad (11)$$

holds; otherwise, it is called a *trivial action*, meaning the following (see Lemma 2 in [21]):

$$d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(r + 0.5)) = d_\varepsilon(\mathbf{x}[k - 1], \mathcal{S}_o(r + 0.5)) \quad (12)$$

Remark 1: Regarding Definition 7, one can observe that there is a reduction in the Lyapunov function $d_\varepsilon(\cdot, \mathcal{S}_o(r + 0.5))$ by at least 1 during any positive action. Moreover, having assumed that the vertices v_i and v_j are chosen at time $k + 1$, where $x_i[k] > x_j[k]$, it is straightforward to show that a positive action occurs at this time if and only if any of the following sets of relations holds:

- $x_i[k] > r + 0.5 + \varepsilon$ and $x_j[k] \leq r + 0.5$; or
- $x_i[k] > r + 0.5$ and $x_j[k] \leq r + 0.5 - \varepsilon$.

Definition 8: Given $\alpha \in \mathbb{R}^\nu$ and $r \in \mathbb{Z}$, define $T_{r,\varepsilon}(\alpha)$ to be the time at which the first positive action is taken w.r.t. $d_\varepsilon(\cdot, \mathcal{S}_o(r + 0.5))$, provided Algorithm 1 is applied to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{P})$ with the initial state α . Notice that since Algorithm 1 is stochastic, $T_{r,\varepsilon}(\alpha)$ is a random variable.

Definition 9: Given $\alpha \in \mathbb{R}^\nu$, let \mathcal{H} denote an infinite sequence whose elements all belong to \mathcal{E} (i.e. \mathcal{H} is an infinite sequence of edges). Define $T_{r,\varepsilon}(\alpha|\mathcal{H})$ to be equal to the time when the first positive action is taken w.r.t. $d_\varepsilon(\cdot, \mathcal{S}_o(r + 0.5))$, provided that Algorithm 2 is applied to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with the initial state α , where the edge selected at time k in step 2 of this algorithm is indeed the k^{th} element of \mathcal{H} , for any $k \in \mathbb{N}$.

Lemma 1: Given $r \in \mathbb{Z}$ and $\alpha := [\alpha_1 \ \dots \ \alpha_\nu]^T \notin \mathcal{S}_o(r + 0.5)$, the following hold for any infinite sequence of edges \mathcal{H} :

- i) If $\alpha_1 > r + 0.5 + \varepsilon$, then:

$$T_{r,\varepsilon}(\alpha|\mathcal{H}) \leq T_{r,\varepsilon}(\alpha_1 - \varepsilon, \alpha_2, \dots, \alpha_\nu|\mathcal{H}) \quad (13)$$

- ii) If $\alpha_1 \leq r + 0.5 - \varepsilon$, then:

$$T_{r,\varepsilon}(\alpha|\mathcal{H}) \leq T_{r,\varepsilon}(\alpha_1 + \varepsilon, \alpha_2, \dots, \alpha_\nu|\mathcal{H}) \quad (14)$$

Proof: Let case (i) be proved here, as the other case is analogous. Apply Algorithm 2 to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ so that its step 2 selects edges from the sequence \mathcal{H} in order. For the initial states $(\alpha_1, \alpha_2, \dots, \alpha_\nu)$ and $(\alpha_1 - \varepsilon, \alpha_2, \dots, \alpha_\nu)$, denote the resulting numbers on the vertices of \mathcal{G} at time k with $\mathbf{w}[k] := (w_1[k], w_2[k], \dots, w_\nu[k])$ and $\bar{\mathbf{w}}[k] := (\bar{w}_1[k], \bar{w}_2[k], \dots, \bar{w}_\nu[k])$, respectively, for any $k \in \mathbb{N} \cup \{0\}$. Furthermore, for notational simplicity, define:

$$m := T_{r,\varepsilon}(\alpha_1 - \varepsilon, \alpha_2, \dots, \alpha_\nu|\mathcal{H}) \quad (15)$$

In order to proceed with the proof by contradiction, assume that there is no positive action w.r.t. $d_\varepsilon(\cdot, \mathcal{S}_o(r + 0.5))$ at time instants $1, 2, \dots, m$, by starting from the initial state α . In other words:

$$d_\varepsilon(\mathbf{w}[0], \mathcal{S}_o(r + 0.5)) = d_\varepsilon(\mathbf{w}[k], \mathcal{S}_o(r + 0.5)) \quad (16)$$

for any $k \in \{0, 1, 2, \dots, m\}$. Now, one can draw a number of conclusions as follows:

- i) $w_1[k]$ is always greater than or equal to $\bar{w}_1[k]$ for any k satisfying the inequality $1 \leq k \leq m$.
- ii) It is an easy consequence of property (i) that $w_i[k] \geq \bar{w}_i[k]$ for any $i \in \nu$ and $k \in \{1, \dots, m\}$.
- iii) The relation $w_i[k] = \bar{w}_i[k]$ holds if $w_i[k] \leq r + 0.5$ or $\bar{w}_i[k] \leq r + 0.5$, for any $i \in \nu$ and $k \in \{0, 1, \dots, m\}$. This result can be easily proven by induction on k , taking property (ii) into account, and using the equality (16).

Assume that the m^{th} element of \mathcal{H} is the edge (i, j) . With no loss of generality, suppose that $\bar{w}_i[m - 1] \geq \bar{w}_j[m - 1]$. Since a positive action occurs at time m w.r.t. $d_\varepsilon(\cdot, \mathcal{S}_o(r + 0.5))$ by starting from the initial state $\bar{\mathbf{w}}[0]$, it follows from Remark 1 that either of the cases pointed out below occurs:

- a) $\bar{w}_i[m - 1] > r + 0.5 + \varepsilon$ and $\bar{w}_j[m - 1] \leq r + 0.5$; or
- b) $\bar{w}_i[m - 1] > r + 0.5$ and $\bar{w}_j[m - 1] \leq r + 0.5 - \varepsilon$.

Assume that case (a) happens (the other case is similar). Properties (ii) and (iii) mentioned above yield that:

$$\begin{aligned} w_i[m-1] &\geq \bar{w}_i[m-1] > r + 0.5 + \varepsilon, \\ w_j[m-1] &= \bar{w}_j[m-1] \leq r + 0.5 \end{aligned} \quad (17)$$

Since the edge (i, j) is to be chosen at time m in step 2 of Algorithm 2, the above inequalities and Remark 1 signify that a positive action occurs at time m for the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with the initial state $\alpha = \mathbf{w}[0]$. In other words, $T_{r,\varepsilon}(\alpha|\mathcal{H})$ must be equal to m , while it has been already assumed that this quantity is greater than m . This contradiction completes the proof. ■

Definition 10: For any integer r and infinite sequence of edges \mathcal{H} , define:

$$\zeta(\varepsilon, r, \mathcal{H}) := \max T_{r,\varepsilon}(\beta|\mathcal{H}) \quad (18)$$

where the maximum is taken over all ν -tuple $\beta := [\beta_1 \ \cdots \ \beta_\nu]$ with the following properties:

- $\beta \notin \mathcal{S}_o(r + 0.5)$;
- There exist $i, j \in \nu$ such that:
 - $\beta_i > r + 0.5 + \varepsilon$ and $\beta_j \leq r + 0.5$; or
 - $\beta_i > r + 0.5$ and $\beta_j \leq r + 0.5 - \varepsilon$.

Proposition 1: Given an integer r and an infinite sequence of edges \mathcal{H} , there exist an integer $\kappa \in \nu$ and a vector $\alpha = [\alpha_1 \ \cdots \ \alpha_\nu]$ subject to:

$$\zeta(\varepsilon, r, \mathcal{H}) = T_{r,\varepsilon}(\alpha|\mathcal{H}), \quad (19a)$$

$$r + 0.5 - \varepsilon < \alpha_i \leq r + 0.5 + \varepsilon, \quad \forall i \in \nu \setminus \{\kappa\} \quad (19b)$$

In addition, α_κ satisfies one of the following relations:

$$r + 0.5 + \varepsilon < \alpha_\kappa \leq r + 0.5 + 2\varepsilon \quad (20)$$

or:

$$r + 0.5 - 2\varepsilon < \alpha_\kappa \leq r + 0.5 - \varepsilon \quad (21)$$

Proof: The proof follows from Lemma 1. The details are omitted here for brevity. Note that the reason why the inequality (19) is not satisfied for every i in ν (i.e. there is a κ for which this inequality does not hold) is that α should not belong to $\mathcal{S}_o(r + 0.5)$, in light of Definition 10. ■

Lemma 2: Consider a vector $\alpha = [\alpha_1 \ \cdots \ \alpha_\nu]$ and an integer $r \in \mathbb{Z}$ satisfying the relations:

$$\begin{aligned} r + 0.5 + \varepsilon &< \alpha_1 \leq r + 0.5 + 2\varepsilon, \\ r + 0.5 - \varepsilon &< \alpha_2 \leq r + 0.5, \\ r + 0.5 - \varepsilon &< \alpha_i \leq r + 0.5 + \varepsilon, \quad \forall i \in \nu \setminus \{1, 2\} \end{aligned} \quad (22)$$

For any infinite sequence of edges \mathcal{H} , the inequality given below holds:

$$T_{r,\varepsilon}(\alpha|\mathcal{H}) \leq T_{r,\varepsilon}(\alpha_1, \alpha_2 + \varepsilon, \dots, \alpha_\nu|\mathcal{H}) \quad (23)$$

Proof: Apply Algorithm 2 to the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ and select edges in its step 2 from the sequence \mathcal{H} successively. For the initial states $(\alpha_1, \alpha_2, \dots, \alpha_\nu)$ and $(\alpha_1, \alpha_2 + \varepsilon, \dots, \alpha_\nu)$, denote the resulting numbers on the vertices of \mathcal{G} at time k with $\mathbf{u}[k] := (u_1[k], u_2[k], \dots, u_\nu[k])$ and $\bar{\mathbf{u}}[k] := (\bar{u}_1[k], \bar{u}_2[k], \dots, \bar{u}_\nu[k])$, respectively, for any $k \in \mathbb{N} \cup \{0\}$. Define also:

$$g := T_{r,\varepsilon}(\alpha_1, \alpha_2 + \varepsilon, \dots, \alpha_\nu|\mathcal{H}) \quad (24)$$

For a proof by contradiction, assume that $T_{r,\varepsilon}(\alpha|\mathcal{H}) > g$. Observe that:

- i) Since $q(\gamma) = r + 1$, for any $\gamma \in (r + 0.5, r + 0.5 + 2\varepsilon]$, it can be verified that:

$$\begin{aligned} u_1[k] &= \bar{u}_1[k] = \alpha_1, \\ u_i[k], \bar{u}_i[k] &\in (r + 0.5 - \varepsilon, r + 0.5 + \varepsilon) \end{aligned} \quad (25)$$

for any $k \in \{0, 1, \dots, g-1\}$ and $i \in \nu \setminus \{1\}$.

- ii) Using property (i) and by means of induction on k , one can show that if $\bar{u}_j[k] \leq r + 0.5$ for some $j \in \nu$ and $k \in \{0, 1, \dots, g-1\}$, then $u_j[k] = \bar{u}_j[k]$.

Let the g^{th} element of \mathcal{H} be the edge (i, j) , where $i < j$. It results from the definition of g and property (i) that $i = 1$ and $\bar{u}_j[m-1] \leq r + 0.5$ (this is the only way to generate a positive action). Therefore, by properties (i) and (ii), one can write:

$$\begin{aligned} u_j[m-1] &= \bar{u}_j[m-1] \leq r + 0.5, \\ u_1[m-1] &> r + 0.5 + \varepsilon \end{aligned} \quad (26)$$

As a result, Remark 1 leads to the conclusion that selecting the edge (i, j) at time m results in a positive action for the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with the initial state α ; i.e. $T_{r,\varepsilon}(\alpha|\mathcal{H}) = g$. This contradicts the aforementioned assumption. ■

Proposition 2: Consider the objects r , \mathcal{H} and $\zeta(\varepsilon, r, \mathcal{H})$ introduced in Proposition 1 and Definition 10. There exists a vector $\alpha := [\alpha_1 \ \cdots \ \alpha_\nu]$ such that:

$$\zeta(\varepsilon, r, \mathcal{H}) = T_{r,\varepsilon}(\alpha|\mathcal{H}), \quad (27a)$$

$$\begin{aligned} \{\alpha_1, \alpha_2, \dots, \alpha_\nu\} &= \left\{ r + 0.5 - \frac{\varepsilon}{2}, r + 0.5 + \frac{\varepsilon}{2}, \dots \right. \\ &\quad \left. , r + 0.5 + \frac{\varepsilon}{2}, r + 0.5 + 3\frac{\varepsilon}{2} \right\} \end{aligned} \quad (27b)$$

(note that the term $r + 0.5 + \frac{\varepsilon}{2}$ appears $\nu - 2$ times in the above set).

Proof: It follows from Proposition 1 and Lemma 2 that there exist a vector $\alpha = [\alpha_1 \ \cdots \ \alpha_\nu]$ and integers $\mu_1, \mu_2 \in \nu$ with the properties:

- $\zeta(\varepsilon, r, \mathcal{H}) = T_{r,\varepsilon}(\alpha|\mathcal{H})$
- The set of inequalities:

$$\begin{aligned} r + 0.5 + \varepsilon &< \alpha_{\mu_1} \leq r + 0.5 + 2\varepsilon, \\ r + 0.5 - \varepsilon &< \alpha_{\mu_2} \leq r + 0.5, \\ r + 0.5 &< \alpha_i \leq r + 0.5 + \varepsilon, \quad \forall i \in \nu \setminus \{\mu_1, \mu_2\} \end{aligned} \quad (28)$$

or:

$$\begin{aligned} r + 0.5 - 2\varepsilon &< \alpha_{\mu_1} \leq r + 0.5 - \varepsilon, \\ r + 0.5 &< \alpha_{\mu_2} \leq r + 0.5 + \varepsilon, \\ r + 0.5 - \varepsilon &< \alpha_i \leq r + 0.5, \quad \forall i \in \nu \setminus \{\mu_1, \mu_2\} \end{aligned} \quad (29)$$

holds.

Due to the symmetry, one can assume with no loss of generality that the set of inequalities given in (28) holds. It is straightforward to show (using the above properties) that $T_{r,\varepsilon}(\alpha|\mathcal{H})$ is unchanged if the following replacements are made:

- α_{μ_1} with $r + 0.5 + \frac{3\varepsilon}{2}$;
- α_{μ_2} with $r + 0.5 - \frac{\varepsilon}{2}$;

- α_i with $r + 0.5 + \frac{\varepsilon}{2}$, for any $i \in \nu \setminus \{\mu_1, \mu_2\}$.

This completes the proof. \blacksquare

Corollary 1: Given an integer r and an infinite sequence of edges \mathcal{H} , there exists a vector $\alpha := [\alpha_1 \ \cdots \ \alpha_\nu]$ such that:

$$\zeta(\varepsilon, r, \mathcal{H}) = T_{0,0.5}(\alpha|\mathcal{H}), \quad (30a)$$

$$\{\alpha_1, \alpha_2, \dots, \alpha_\nu\} = \{0.25, 0.75, \dots, 0.75, 1.25\} \quad (30b)$$

where that the number 0.75 appears $\nu - 2$ times in the set given above

Proof: The proof is straightforward (using Proposition 2). The details are omitted for brevity. \blacksquare

Corollary 1 states that the quantity $\zeta(\varepsilon, r, \mathcal{H})$, introduced in Definition 10, is independent of r and ε . Instead, it is contingent upon only \mathcal{H} and \mathcal{G} .

Define Φ as follows:

$$\Phi := \max_{\alpha} \mathcal{E}\{T_{0,0.5}(\alpha)\} \quad (31)$$

where the maximum is taken over all ν -tuple $\alpha = [\alpha_1 \ \cdots \ \alpha_\nu]$ satisfying the relation $\{\alpha_1, \alpha_2, \dots, \alpha_\nu\} = \{0.25, 0.75, \dots, 0.75, 1.25\}$ in which the value 0.75 appears $\nu - 2$ times.

Theorem 2: Given a real $\varepsilon \in (0, 0.5]$, the quantity $t_c(\varepsilon)$ can be upper and lower bounded as follows:

$$\Phi \leq t_c(\varepsilon) \leq \frac{\nu(x_{\max} - x_{\min} + 2)\Phi}{\varepsilon} \quad (32)$$

Proof: With no loss of generality, assume that:

$$x_{\text{ave}} - x_{\min} \geq x_{\max} - x_{\text{ave}} \quad (33)$$

Recall that the proof of Theorem 2 in Part I of the paper proposes two storage functions as follows [21]:

$$\begin{aligned} V_1[k] &:= d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(\eta_1)), \\ V_2[k] &:= d_\varepsilon(\mathbf{x}[k], \mathcal{S}_o(\eta_2)) \end{aligned} \quad (34)$$

It also suggests to minimize $V_1[k]$ until it reaches a constant level, and subsequently minimize $V_2[k]$ (in other words, the edge being selected in step 2 of the algorithm at each update is to be chosen in such a way that the storage function is minimized). It follows from the inequality (33) and the relation $|x_{\text{ave}} - \eta_1| \leq 1$ that:

$$V_1[0] \leq \sum_{i=1}^{\nu} \frac{|x_i - \eta_1|}{\varepsilon} \leq \nu \frac{x_{\text{ave}} - x_{\min} + 1}{\varepsilon} \quad (35)$$

At time $k = k_0$ where the minimum of $V_1[k]$ is reached, two possibilities can happen according to Lemma 3 in Part I of the paper. The first one is that $\mathbf{x}[k_0] \in \mathcal{S}_o(\eta_1)$, which implies that the consensus is reached, and there is no need to minimize $V_2[k]$ any longer. The second scenario is that the relation $x_i[k_0] > \eta_1$ holds for any $i \in \nu$. Assume that the latter one is the case. It is easy to verify that:

$$V_2[k_0] \leq \sum_{i=1}^{\nu} \frac{|x_i[k_0] - \eta_2|}{\varepsilon} \leq \nu \frac{x_{\max} - x_{\text{ave}} + 1}{\varepsilon} \quad (36)$$

As a result:

$$V_1[0] + V_2[k_0] \leq \frac{\nu(x_{\max} - x_{\min} + 2)}{\varepsilon} \quad (37)$$

On the other hand, the aforementioned discussion indicates that at most $V_1[0] + V_2[k_0]$ positive actions are required to reach the consensus. Moreover, it can be inferred from Corollary 1 that the expected value of the time between two consecutive positive actions is at most equal to Φ (until the consensus is reached). These facts along with the inequality (37) complete the proof. \blacksquare

Theorem 2 states that $t_c(\varepsilon)$ can be upper bounded by a term which is proportional to the inverse of ε .

A question arises as how to compute Φ systematically. This is addressed in the sequel. The following definitions/notations will be convenient to proceed with the development of the paper.

- Let P be the weighted Laplacian of the graph \mathcal{G} . In other words, P is a $\nu \times \nu$ matrix whose (i, j) off-diagonal entry, $i, j \in \nu, i \neq j$, is equal to $-p_{ij}$ and its (i, i) diagonal entry, $i \in \nu$, is equal to $\sum_{k=1}^{\nu} p_{ki}$.
- For any $i \in \nu$ and $M \in \mathbb{R}^{\nu \times \nu}$, define $M_{\sim i}$ to be the matrix obtained from M by removing its i^{th} row and i^{th} column.

Theorem 3: The quantity Φ can be obtained as follows:

$$\Phi = \max_{i \in \nu} |(P_{\sim i})^{-1}|_{\infty} \quad (38)$$

where $|\cdot|_{\infty}$ stands for the infinity norm.

Proof: For any $i, j \in \nu, i \neq j$, let β_{ij} denote a ν -dimensional vector whose elements are all equal to 0.75, except for the i^{th} and j^{th} entries which are 0.25 and 1.25, respectively. In addition, denote the quantity $\mathcal{E}\{T_{0,0.5}(\beta_{ij})\}$ with $\bar{t}_c(i, j)$. We wish to contrive a recursive equation characterizing $\bar{t}_c(i, j)$. To this end, consider the graph \mathcal{G} with the initial state β_{ij} . The expected value of the time at which the first positive action is taken (under Algorithm 1) is, by definition, equal to $\bar{t}_c(i, j)$. To count this number in another way, run the algorithm one iteration. Assume that the edge e_μ is chosen in this iteration. Consider the possibilities pointed out below:

- e_μ is equal to the edge (i, k) , for some $k \in \nu \setminus \{j\}$: In this case, due to the equality $\mathbf{x}[0] = \beta_{ij}$, the vector $\mathbf{x}[1]$ is obtained as β_{kj} . Hence, it is expected to take the first positive action after $\bar{t}_c(k, j)$ time updates (in addition to the first time update taken at the beginning).
- e_μ is equal to the edge (i, j) : This means that a positive action is already taken at the first time update.
- e_μ is equal to the edge (k, l) , for some $k, l \in \nu \setminus \{i\}$: In this case, it is easy to show that $\mathbf{x}[1] = \mathbf{x}[0] = \beta_{ij}$. This implies that it is expected to take the first positive action after $\bar{t}_c(i, j)$ time updates (other than the first one already taken).

The above reasoning leads to the recursive equation:

$$\begin{aligned} \bar{t}_c(i, j) &= 1 + \sum_{k \in \nu \setminus \{j\}} p_{ik} \bar{t}_c(k, j) \\ &+ \left(1 - \sum_{k=1}^{\nu} p_{ik}\right) \bar{t}_c(i, j), \quad \forall i \in \nu \setminus \{j\} \end{aligned} \quad (39)$$

Stack the scalars $\bar{t}_c(1, j), \dots, \bar{t}_c(j-1, j), \bar{t}_c(j+1, j), \dots, \bar{t}_c(\nu, j)$ in a column and denote the resulting vector with $\bar{\mathbf{t}}_c(j) \in \mathbb{R}^{\nu-1}$. The equation (39) can be re-arranged as:

$$P_{\sim j} \bar{\mathbf{t}}_c(j) = E, \quad \forall j \in \nu \quad (40)$$

where $E \in \mathbb{R}^{\nu-1}$ is a vector of 1's. It follows from the equality $|P_{\sim j}^{-1}E|_\infty = |P_{\sim j}^{-1}|_\infty$ and the above equation that:

$$\begin{aligned} \Phi &= \max_{i, j \in \nu, i \neq j} \bar{t}_c(i, j) = \max_{j \in \nu} |\bar{\mathbf{t}}_c(j)|_\infty \\ &= \max_{j \in \nu} |(P_{\sim j})^{-1}E|_\infty = \max_{j \in \nu} |(P_{\sim j})^{-1}|_\infty \end{aligned} \quad (41)$$

This completes the proof. \blacksquare

The results of Theorems 2 and 3 can be combined to explicitly bound the quantity $t_c(\varepsilon)$. This is carried out in the sequel.

Theorem 4: The scalar $t_c(\varepsilon)$ satisfies the following inequalities:

$$\frac{1}{\sqrt{\nu-1}} \left(\max_{i \in \nu} \frac{1}{\lambda_{\min}\{P_{\sim i}\}} \right) \leq t_c(\varepsilon) \quad (42)$$

and

$$t_c(\varepsilon) \leq \frac{\nu \sqrt{\nu-1} (x_{\max} - x_{\min} + 2)}{\varepsilon} \left(\max_{i \in \nu} \frac{1}{\lambda_{\min}\{P_{\sim i}\}} \right) \quad (43)$$

where $\lambda_{\min}(\cdot)$ represents the smallest eigenvalue of a matrix.

Proof: One can write:

$$|(P_{\sim j})^{-1}|_\infty \leq \sqrt{\nu-1} |(P_{\sim j})^{-1}|_2 \quad (44)$$

where $|\cdot|_2$ stands for the 2-norm. Since the weighted Laplacian matrix P is positive semi-definite (PSD), its principal minor $P_{\sim j}$ is PSD too. As a result, it can be deduced from the above inequality that:

$$|(P_{\sim j})^{-1}|_\infty \leq \sqrt{\nu-1} \frac{1}{\lambda_{\min}\{P_{\sim j}\}} \quad (45)$$

Similarly:

$$|(P_{\sim j})^{-1}|_\infty \geq \frac{1}{\sqrt{\nu-1}} |(P_{\sim j})^{-1}|_2 = \frac{1}{\lambda_{\min}\{P_{\sim j}\} \sqrt{\nu-1}} \quad (46)$$

The proof is completed by the inequalities (44) and (46), on noting Theorems 2 and 3. \blacksquare

Remark 2: Theorem 4 states that the convergence time $t_c(\varepsilon)$ is related to the $(\nu-1)$ th order minors of the Laplacian of the graph (i.e. $P_{\sim i}, i \in \nu$), rather than the Laplacian itself. Let $\lambda_2(P)$ denote the second smallest eigenvalue of P . Since the graph \mathcal{G} is connected, $\lambda_2(P)$ is strictly positive. Now, the interlacing theorem can be exploited to argue that:

$$0 < \lambda_{\min}\{P_{\sim j}\} \leq \lambda_2(P) \quad (47)$$

This means that unlike the unquantized consensus whose convergence mainly depends on $\lambda_2(P)$, a more subtle dependency on $\lambda_2(P)$ is governed for the quantized case (in fact, Φ is not directly related to $\lambda_2(P)$).

A. Special graphs

This subsection aims to obtain lower and upper bounds on the quantity $t_c(\varepsilon)$ for both complete and path graphs in the case when all edges have the same weight. In this regard, assume that each edge is associated with the same weight p .

Corollary 2: For a complete graph \mathcal{G} with equally weighted edges, the following inequality holds:

$$\frac{\nu(\nu-1)}{2} \leq t_c(\varepsilon) \leq \frac{\nu^2(\nu-1)(x_{\max} - x_{\min} + 2)}{2\varepsilon} \quad (48)$$

Proof: The weight p for this graph is equal to $\frac{2}{\nu(\nu-1)}$. Using this fact and by means of Theorems 2 and 3, it is straightforward to show the validity of the inequality (48). The details are omitted here. \blacksquare

Corollary 3: Let \mathcal{G} be a path graph with equally weighted edges so that v_i is connected to v_{i+1} , for any $i \in \{1, 2, \dots, \nu-1\}$ (these are the only edges of the graph). The inequality given below holds:

$$\frac{\nu(\nu-1)^2}{2} \leq t_c(\varepsilon) \leq \frac{\nu^2(\nu-1)^2(x_{\max} - x_{\min} + 2)}{2\varepsilon} \quad (49)$$

Proof: Since the graph \mathcal{G} has $\nu-1$ edges, the weight p is equal to $\frac{1}{\nu-1}$. On the other hand, it is easy to show that $\Phi = \bar{t}_c(1, \nu)$. This leads to the following recursive equations (in light of (39)):

$$p\bar{t}_c(1, \nu) - p\bar{t}_c(2, \nu) = 1 \quad (50a)$$

$$-p\bar{t}_c(i-1, \nu) + 2p\bar{t}_c(i, \nu) - p\bar{t}_c(i+1, \nu) = 1 \quad (50b)$$

$$-p\bar{t}_c(\nu-2, \nu) + 2p\bar{t}_c(\nu-1, \nu) = 1 \quad (50c)$$

where the argument i in the equation (50b) belongs to the set $\{2, 3, \dots, \nu-2\}$. Adding up these equalities results in the relation:

$$p\bar{t}_c(\nu-1, \nu) = \nu-1 \quad (51)$$

The (backward) recursive equation (50b) can be solved using conventional techniques to conclude that there exist two constants a and b such that:

$$p\bar{t}_c(i, \nu) = a + bi - \frac{i^2}{2}, \quad i = \nu-1, \nu-2, \dots, 1 \quad (52)$$

One can employ the final conditions given by (50c) and (51) to arrive at:

$$a = \frac{\nu^2 - \nu}{2}, \quad b = \frac{1}{2} \quad (53)$$

This implies that:

$$\Phi = \bar{t}_c(1, \nu) = \frac{\nu^2 - \nu}{2p} = \frac{\nu(\nu-1)^2}{2} \quad (54)$$

The proof follows immediately from the above equation and Theorem 2. \blacksquare

B. Optimal edge weights

In this subsection, it is desired to find out what probabilities the edges of \mathcal{G} should possess so that the consensus is reached fast. For this purpose, observe that the quantity $t_c(\varepsilon)$ is bounded from below and above in (42) and (43). Having

fixed x_{\min} , x_{\max} and ε , these lower and upper bounds are constant multiples of the term:

$$\max_{i \in \nu} \frac{1}{\lambda_{\min}\{P_{\sim i}\}} \quad (55)$$

Therefore, the lower and upper bounds provided in Theorem 4 can be *simultaneously* minimized by merely minimizing the above expression. This interesting fact leads to finding a sub-optimal edge-selection probability distribution by minimizing the function (55) over all possible (discrete) probability distributions captured by P . In what follows, this is accomplished.

Problem 1: Minimize the scalar variable $-\mu$ subject to the constraints:

$$\lambda_{\min}\{P_{\sim i}\} \geq \mu, \quad i = 1, 2, \dots, \nu \quad (56)$$

where P is a matrix variable representing the Laplacian of the weighted graph \mathcal{G} . Denote the global minimizer of this optimization with (μ^*, P^*) (note that there are some implicit constraints stating that the weights on the edges are positive and sum up to 1).

Since the operator $\lambda_{\min}(\cdot)$ is concave in terms of its symmetric argument, it is easy to show that Problem 1 is convex. More precisely, the constraint $\lambda_{\min}\{P_{\sim i}\} \geq \mu$ can be expressed as $P_{\sim i} \geq \mu I$, which is a semidefinite constraint. Hence, the solution P^* can be found efficiently. On the other hand, one can verify that:

$$\mu^* = \max_P \min_i \lambda_{\min}\{P_{\sim i}\} \quad (57)$$

or equivalently:

$$\frac{1}{\mu^*} = \min_P \max_i \frac{1}{\lambda_{\min}\{P_{\sim i}\}} \quad (58)$$

This implies that the solution P^* gives a sub-optimal edge-selection probability distribution (resulting in the fast convergence of Algorithm 1), because of minimizing the term given in (55).

IV. NUMERICAL EXAMPLE

Consider the graph \mathcal{G} drawn in Figure 1. The objective is to find out what probabilities should be assigned to the edges of \mathcal{G} so that the consensus is reached fast under Algorithm 1. To this end, let the convex optimization provided in Problem 1 be solved. This yields the following probability distribution:

$$\begin{aligned} p_{12} &= p_{15} = 0.2087, \\ p_{23} &= p_{24} = p_{45} = p_{35} = 0.1146, \\ p_{34} &= 0.1241 \end{aligned} \quad (59)$$

The quantity Φ corresponding to this set of edge-selection probabilities turns out to be equal to 14.1770. One can make a comparison with two heuristic methods for designing the probability set \mathcal{P} , which are spelled out below:

- The most naive approach is to assume that the edges of the graph are equally weighted. This leads to the probability $p = \frac{1}{7}$ on each edge. The associated quantity Φ is obtained as 17.5.

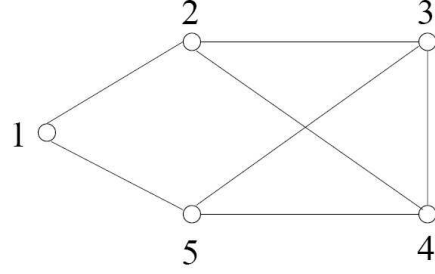


Fig. 1. The graph \mathcal{G} rendered in the numerical example

- Another technique is to devise the probability distribution \mathcal{P} in such a way that all vertices have the same probability of being chosen at each time update, i.e.:

$$\begin{aligned} p_{12} + p_{15} &= p_{21} + p_{23} + p_{24} \\ &= p_{32} + p_{34} + p_{35} \\ &= p_{42} + p_{43} + p_{45} \\ &= p_{51} + p_{53} + p_{54} \end{aligned} \quad (60)$$

Note that $p_{ij} = p_{ji}$, $\forall i, j \in \nu$. The above set of equations has a unique symmetric solution (complying with the symmetry of the graph \mathcal{G}) as follows:

$$\begin{aligned} p_{12} &= p_{15} = 0.2, \\ p_{23} &= p_{24} = p_{45} = p_{35} = 0.1, \\ p_{34} &= 0.2 \end{aligned} \quad (61)$$

The corresponding Φ is equal to 15.

Hence, the value of Φ corresponding to the sub-optimal solution is better from the ones obtained from these two rudimentary techniques.

Let an interesting fact about the edge selection be revealed now. Assume that the graph \mathcal{G} does not have the edge $(1, 5)$ (i.e. remove this edge). In this case, Problem 1 leads to the following solution:

$$\begin{aligned} p_{12} &= 0.3781, \quad p_{23} = p_{24} = 0.1757, \\ p_{34} &= 0, \quad p_{53} = p_{54} = 0.1352 \end{aligned} \quad (62)$$

associated with $\Phi = 23.1292$. Notice that $p_{34} = 0$, which signifies that although a complete graph has the best convergence, if some edges do not exist (e.g. the edge $(1, 5)$), it might be better to ignore some other edges too (e.g. the edge $(3, 4)$). This is fascinating as it reveals the fact that some communications are redundant.

To compare the value obtained for Φ with other possible values, consider the case when all edges have the same weight. This results in the equality $\Phi = 36$. Therefore, there is a noticeable improvement in the value of Φ via the solution of Problem 1.

For the purpose of simulation, the following points have been randomly generated in the interval $[0, 100]$:

$$\begin{aligned} x_1[0] &= 20.1185, \quad x_2[0] = 13.6221, \quad x_3[0] = 97.8356, \\ x_4[0] &= 45.5033, \quad x_5[0] = 45.9224 \end{aligned} \quad (63)$$

The stochastic gossip algorithm was run 1000 times and the average of the scalar k_1 was calculated accordingly (note that k_1 is a random variable). This value for the probability distribution (62) was obtained as 48.5710, while that for the identical probability distribution (equal edge weights) turned out to be 65.3580. This demonstrates that one could save significantly in the convergence time if the solution of Problem 1 is employed, which also obviates the usage of the edge (3, 4).

A. Comparison with the work [16]

The work [16] proposes a SDP optimization in order to obtain optimal weights for the unquantized consensus. We wish to show that the set of optimal weights for the unquantized consensus (using a gossip algorithm) might be by no means a sub-optimal solution for the quantized consensus. To this end, consider again the graph \mathcal{G} with its edge (1, 5) removed. Solving the optimization problem given in [16] arrives at:

$$\begin{aligned} p_{12} &= 0.3077, & p_{23} &= p_{24} = 0.2115, \\ p_{34} &= 0, & p_{53} &= p_{54} = 0.1346 \end{aligned} \quad (64)$$

Note that the gossip algorithm proposed in [19] is somewhat different from the one described here. More specifically, it does not select an edge directly; instead, it first chooses a vertex and then picks one of its edges. However, the corresponding edge and vertex probabilities can be interpreted as the edge probabilities given above. The quantity Φ for this set of weights is equal to 23.8066, which is fairly close to the sub-optimal one obtained earlier. However, the average of k_1 over 1000 runs was obtained as 91.6410, which is much larger than 48.5710 corresponding to the sub-optimal case, as computed above. This demonstrates that optimizing the convergence time under an unquantized gossip algorithm may degrade the convergence time under its corresponding quantized algorithm.

V. CONCLUSIONS

This paper tackles the average consensus problem over a connected weighted graph subject to a quantization constraint. It is assumed that each pair of vertices can be chosen with a certain probability in order to update their numbers in term of the quantized data being exchanged. In the first part of the paper, it was shown that the quantized consensus is reached under the stochastic gossip algorithm given in a recent paper. This part of the paper deals with the time at which the consensus is reached. Lower and upper bounds on the expected value of this quantity in the worst case are provided, which depend on the principal minors of the weighted Laplacian matrix of the graph. These bounds are explicitly computed for equally weighted complete and path graphs. Finally, a convex optimization is provided with the aim of finding out what weights on the edges result in the fast convergence of the gossip algorithm.

REFERENCES

- [1] G. Tel, "Introduction to distributed algorithms," Cambridge University Press, 2000.
- [2] N. A. Lynch, "Distributed algorithms," Morgan Kaufmann Publishers, Inc., San Francisco, CA, 1996.
- [3] 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.
- [4] 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.
- [5] Y. Kuramoto, "Chemical oscillators, waves, and turbulence," Springer-Verlag, Berlin, 1984.
- [6] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, pp. 268-276, 2001.
- [7] D. P. Bertsekas and J. N. Tsitsiklis, "Parallel and distributed computation: Numerical methods," Belmont, MA: Athena Scientific, 1997.
- [8] 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.
- [9] A. V. Savkin, "Coordinated collective motion of groups of autonomous mobile robots: Analysis of Vicsek's model," *IEEE Transactions on Automatic Control*, vol. 49, no. 6, pp. 981-982, 2004.
- [10] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401-420, 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. 988-1001, 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] 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.
- [14] J. Tsitsiklis, "Problems in decentralized decision making and computation," PhD thesis, Dept. of Electrical Engineering and Computer Science, M.I.T., Boston, MA, 1984.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] A. Kashyap, T. Basara and R. Srikanta, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192-1203, 2007.
- [19] 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>).
- [20] 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.
- [21] J. Lavaei and R. M. Murray, "On quantized consensus by means of gossip algorithm – Part I: convergence proof," submitted to *Proceedings of the 2009 American Control Conference*, 2009 (available online at www.cds.caltech.edu/~lavaei/publication/consensus1.pdf).