

Average Consensus and Gossip Algorithms in Networks with Stochastic Asymmetric Communications

Duarte Antunes, Daniel Silvestre, Carlos Silvestre

Abstract— We consider that a set of distributed agents desire to reach consensus on the average of their initial state values, while communicating with neighboring agents through a shared medium. This communication medium allows only one agent to transmit unidirectionally at a given time, which is true, e.g., in wireless networks. We address scenarios where the choice of agents that transmit and receive messages at each transmission time follows a stochastic characterization, and we model the topology of allowable transmissions with asymmetric graphs. In particular, we consider: (i) randomized gossip algorithms in wireless networks, where each agent becomes active at randomly chosen times, transmitting its data to a single neighbor; (ii) broadcast wireless networks, where each agent transmits to all the other agents, and access to the network occurs with the same probability for every node. We propose a solution in terms of a linear distributed algorithm based on a state augmentation technique, and prove that this solution achieves average consensus in a stochastic sense, for the special cases (i) and (ii). Expressions for absolute time convergence rates at which average consensus is achieved are also given.

I. INTRODUCTION

The average consensus problem is a distributed control problem in which a set of agents aims to agree on the average of their initial state values by exchanging messages dictated by a given communication topology. Several multidisciplinary applications of average consensus algorithms have been reported in the literature. These include distributed optimization [1], [2]; motion coordination tasks, such as flocking, leader following [3], and rendezvous problems [4]; and resource allocation in computer networks [5].

An elegant theory is now available in the literature to solve consensus problems using linear distributed algorithms, in which each agent computes a weighted average between its state value and the state values of the agents to which it can communicate (see, e.g. [6], [7]). Many variations of this problem have been addressed in the literature considering, e.g., stochastic packet drops and link failures [8], [9], quantized data transmissions [10], and time-varying communication connectivity [6].

On the other hand, randomized gossip algorithms have been proposed, e.g., in [11], as a decentralized solution to the average consensus problem that can deal with several features such as the absence of a centralized entity, and the possibly varying topology of the network caused by agents that join and leave the network. The premise of gossip

algorithms is that each agent communicates with no more than one neighbor at each transmission time. In [11], it is assumed that at each communication time, the sender and the receiver set their state to the average of their current values, which subsumes that the communication is bidirectional.

In the present work, we consider the average consensus problem in scenarios where communication is unidirectional at each time slot, i.e., at each transmission time a single agent transmits data to one or several agents, but does not receive data. Note that at a different time slot receiver and sender agents may invert their roles, i.e., the word unidirectional refers only to communication at a given transmission time. For concreteness, we consider the two following scenarios: (i) randomized gossip algorithms in wireless networks, where each agent becomes active at randomly chosen times, transmitting its data to a single neighbor; (ii) broadcast wireless networks, where each agent transmits to all the other agents, access to the network occurs with the same probability for every agent, and the intervals between transmissions are independent and identically distributed. As we shall see, the unidirectionality communication constraint precludes in general the existence of a linear distributed algorithm where associated to each agent there is a single scalar state, updated based on the state of the other agents, as in related problems where the communication topology of the network is also time-varying, but satisfies different assumptions (see [6], [7]). We assume a symmetric communication topology, meaning that if an agent a can communicate with an agent b then the agent b can communicate with the agent a , although this does not take place at the same transmission time, i.e., at each transmission time the graphs modeling communications are in general asymmetric. Note that this is typically the case in wireless networks, and therefore this assumption is reasonable to assume in both scenarios (i) and (ii).

The main contribution of the present paper, is to propose a linear distributed algorithm using a state augmentation technique to achieve average consensus when communication constraints impose time-varying unidirectional transmissions, and to prove convergence in expectation to average consensus for scenarios (i) and (ii). Moreover, the stochasticity of the times between transmissions, in both gossip and broadcast scenarios considered herein, render non-trivial the computation of the rates of convergence of the linear distributed algorithms to consensus in absolute time, as opposed to discrete-time, i.e., in terms of the number of transmissions. We provide expressions for the absolute time rates in both scenarios.

A recent work [12] addresses a similar problem, consid-

D. Antunes is with the Department of Mechanical Engineering, Eindhoven University of Technology, the Netherlands, D.Antunes@tue.nl

D. Silvestre, C. Silvestre are with the Dep. of Electrical Eng. and Computer Science, Instituto Superior Técnico, ISR, 1046-001 Lisboa, Portugal. {dsilvestre,cjs}@isr.ist.utl.pt

ering gossip algorithms also with asynchronous communication between the agents. In [12], a method is proposed to achieve average consensus, also using a state augmentation technique, and this method is proved to converge almost surely to consensus. The communication topologies considered in [12] encompass general directed graphs. Hence, [12] does not need to assume a symmetric communication topology, which is crucial to obtain our convergence results. Note, however, that in the method proposed in [12], the state updates depend in a nonlinear way of the current state, while the algorithm that we propose is a linear distributed algorithm, proving a solution that parallels existing algorithms for the standard gossip [11] and linear distributed algorithms [7].

The remainder of the paper is organized as follows. We provide some preliminaries and state the average consensus problem in Section II. The proposed solution is given in Section III and our main results are stated in Section IV. Concluding remarks and directions for future work are given in Section VI.

Notation : The transpose and the spectral radius of a matrix A are denoted by A^\top and $r_\sigma(A)$, respectively. For vectors a_i , $(a_1, \dots, a_n) := [a_1^\top \dots a_n^\top]^\top$. We let $\mathbf{1}_n := [1 \dots 1]^\top$ and $\mathbf{0}_n := [0 \dots 0]^\top$ indicate n -dimension vector of ones and zeros, and I_n denotes the identity matrix of dimension n . Dimensions are omitted when no confusion arises. The vector e_i denotes the canonical vector whose components equal zero, except component i that equals one. The notation $\text{diag}([A_1 \dots A_n])$ indicates a block diagonal matrix with blocks A_i . The Kronecker product is denoted by \otimes .

II. PRELIMINARIES AND PROBLEM STATEMENT

We consider that a set of m agents with scalar state $x_i(t)$, $1 \leq i \leq m$, desire to obtain the average of their initial states, i.e.,

$$\lim_{t \rightarrow \infty} x_i(t) = x_{\text{av}} := \frac{1}{m} \sum_{i=1}^m x_i(0). \quad (1)$$

We refer to this problem as the *average consensus problem*. The communication topology is modeled by a directed graph $G = (V, E)$, where V is the set of m agents, also denoted by nodes, and $E \subseteq V \times V$ is the set of communication links. The node i can send a message to the node j , if $(j, i) \in E$. If there exists at least one $i \in V$ such that $(i, i) \in E$ we say that the graph has self-loops. A standing assumption will be that G is strongly connected and aperiodic [7]. We associate to the graph G an *adjacency matrix* N with entries

$$N_{ij} := \begin{cases} 1, & \text{if } (i, j) \in E \\ 0, & \text{otherwise} \end{cases}. \quad (2)$$

More generally, we can consider a *weighted adjacency matrix* W associated with G with entries $[W]_{ij} = w_{ij} \in \mathbb{R}$, if $(i, j) \in E$ and $[W]_{ij} = 0$, otherwise.

A. Previous Work

A *linear distributed algorithm* is defined by the iteration

$$x(t_{k+1}) = W_k x(t_k), \quad (3)$$

where $\{W_k, k \geq 0\}$ is a set of weighted adjacency matrices associated with the graph G , and t_k are the times at which a transmission occurs, with $t_0 := 0$. It is subsumed throughout the paper that between transmission times, the state variables do not change their values, i.e., e.g., $\dot{x}(t) = 0$, $t \in [t_k, t_{k+1})$.¹ If $W_k = W_c$ is constant for every $k \geq 0$, where W_c is a stochastic matrix ($W_c \mathbf{1} = \mathbf{1}$), and the graph is strongly connected and aperiodic, the linear distributed algorithm (3) converges to consensus (cf. [7]), meaning that $\lim_{k \rightarrow \infty} x_k = \mathbf{1}c$, for a given constant c . See [7] for conditions under which consensus is achieved in the case where W_k are time-varying. Moreover, if all the matrices $\{W_k, k \geq 0\}$ are doubly stochastic ($W_k \mathbf{1} = \mathbf{1}$ and $\mathbf{1}^\top W_k = \mathbf{1}^\top$) and (3) converges to consensus then (3) converges to average consensus, i.e., (1) is achieved.

A randomized gossip algorithm is a special case of a linear distributed algorithms. In gossip algorithms, each of the m nodes has a clock which ticks at a rate 1 Poisson process, i.e., inter-tick times are rate 1 exponential distributed. It is clear that at a given tick time, each node has the same probability of being the node whose clock ticked. Such node, denoted by i , chooses a random out-neighbor j (corresponding to the non-zero entries of column i of the adjacency matrix N) according to a given distribution p_{ij} , which denotes the probability that node $j \in N_{\text{out}}(i)$ is chosen, where $N_{\text{out}}(i)$ denotes the set of out-neighbors of node i and $\sum_{j \in N_{\text{out}}(i)} p_{ij} = 1$. Both nodes update their current state values according to the averages of their values, while the remaining nodes keep their state intact, i.e., if node i ticks at time t_k and chooses node ℓ to communicate with, then $x(t_{k+1}) = W_k x(t_k)$, where $W_k = M_{i\ell}$,

$$M_{i\ell} := I - \frac{(e_i - e_\ell)(e_i - e_\ell)^\top}{2}$$

B. Stability and convergence rate

As in gossip algorithms, there are several works in the literature (see, e.g., [8], [9], and references therein), where the matrices W_k in (3) are assumed to be stochastic matrices, i.e., $\text{Prob}[W_k = M_j] = p_j$, for some matrices M_j , $1 \leq j \leq n_M$, and probabilities p_j , $\sum_{i=1}^{n_M} p_j = 1$. This may be due to several sources of randomness, such as packet drops or link failures. In such cases the definition of achieving average consensus (cf. (1)) must be adapted to a stochastic setting.

Definition 1: We say that a linear distributed algorithm taking the form (3), where $\{W_k, k \geq 0\}$ are stochastic matrices:

- (i) converges almost surely to average consensus if

$$\lim_{t \rightarrow \infty} x_i(t) = x_{\text{av}} := \frac{1}{m} \sum_{i=1}^m x_i(0), \quad \forall i \in \{1, \dots, m\}$$

almost surely.

¹We use this formulation, instead of simply considering discrete-time variables $z_k = x(t_k)$ indexed by the transmission number k , because we are interested in determining convergence rates to consensus in absolute time t .

(ii) converges in expectation to average consensus if

$$\lim_{t \rightarrow \infty} \mathbb{E}[x_i(t)] = x_{av} \quad , \quad \forall i \in \{1, \dots, m\}.$$

(iii) converges in second moment to average consensus if

$$\lim_{t \rightarrow \infty} \mathbb{E}[(x_i(t) - x_{av})^2] \rightarrow 0 \quad , \quad \forall i \in \{1, \dots, m\}.$$

In applications it is typically easier to prove convergence in expectation and in second moment, which are the two stability definitions we shall focus in the present paper. It is clear that convergence in second moment to consensus is stronger than convergence in expectation and also stronger than almost sure convergence. As argued in [11] it is sometimes possible to prove that convergence in expectation implies almost sure convergence. However, we make no such assertion here.

Provided that consensus is achieved one may ask at which rate is it achieved. In the present paper we shall focus on determining the following rate corresponding to the convergence rate to zero of the second moment of the error between state and average consensus.

Definition 2: Suppose that the following condition holds for some $\gamma > 0$ and $c > 0$ and for every $x(0) \in \mathbb{R}^m$,

$$\mathbb{E}[(x_i(t) - x_{av})^2] \leq ce^{-\gamma t} (x_i(0) - x_{av})^2, \forall i \in \{1, \dots, m\}. \quad (4)$$

Then, the second moment convergence rate is defined as

$$\alpha := \sup\{\gamma : (4) \text{ holds for some } c > 0 \text{ and every } x(0) \in \mathbb{R}^m\}.$$

C. Problem Statement

We consider the following communication constraint, inspired by wireless communication constraints:

Communication Constraint 3: At each transmission time t_k , a single agent i sends data to one or various of its out-neighbors, and cannot receive data from its in-neighbors.

For concreteness, we consider the two following scenarios.

Gossip algorithms: We adapt the set-up proposed in [11] and described in Section II-A to accommodate the communication constraint 3. In other words, we consider that each of the m agents has a clock which ticks at a rate 1 Poisson process, randomly choosing one of its out-neighbors according to the distribution p_{ij} associated with node i . Node j can change its state (not necessary a scalar) at this transmission time based on the data received from node i . However, node i does not receive data at this transmission time. We allow p_{ii} to be different from zero, meaning that at some clock ticks, node i can update its state without communicating to any other node.

Broadcast wireless networks: At each transmission time an agent sends data to every other agent in the network. We consider that nodes have the same probability of accessing the networks at each transmission time and assume that the times between transmissions are independent and identically distributed as it is typical to consider in such networks (cf. [13]).

The problems we are interested in are the following.

- (i) In the setup of scenarios 1) and 2), design a linear distributed algorithm that achieves average consensus according to one of the stability notions of Definition 1.
- (ii) Compute the second moment convergence rate to average consensus of this linear distributed algorithm.

Although we shall provide a method to obtain average consensus that can be applied to any communication topology described by (2), our stability results given for scenarios 1) and 2) assume that N is symmetric, which is reasonable to assume in wireless networks, in which if a node a can communicate to a node b then typically b can communicate with a .

The next example illustrates why, in general, it is not possible to achieve average consensus by using the standard solution described in Section II-A, where each node updates a scalar state according to the recursion (3).

Example 4: Suppose that $m = 2$ and the adjacency matrix is given by $N = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. When node $i \in \{1, 2\}$ transmits, node $j \in \{1, 2\}$, $j \neq i$, updates its information based on the received state of x_i and its own state x_j , and node i can only update its state based on its own state information. Thus, a linear distributed algorithm using a single scalar state per node, is described by (3), where $W_k \in \{Z_1, Z_2\}$, and $Z_1 = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 0 & \alpha_3 \end{bmatrix}$, $Z_2 = \begin{bmatrix} \beta_1 & 0 \\ \beta_2 & \beta_3 \end{bmatrix}$, where we assume that $\alpha_2 \neq 0$, and $\beta_2 \neq 0$, since otherwise there would be no communication between the nodes. It is clear that W_k cannot therefore be doubly stochastic matrices. \square

In the next section, we propose a linear distributed algorithm that solves this problem by using a state augmentation technique.

III. PROPOSED SOLUTION

We augment the original state $x(t)$ with an auxiliary vector $y(t) \in \mathbb{R}^{n_y}$, and define

$$z = (x, y). \quad (5)$$

We consider a linear distributed algorithm taking the form

$$z(t_{k+1}) = U_k z(t_k), \quad (6)$$

where $z(0) = (x(0), y(0))$, $y(0) = 0$. Intuitively, the purpose of y it to assure that at each iteration the total state average is kept constant, i.e., that

$$\frac{\sum_1^m x_i(t_{k+1}) + \sum_1^{n_y} y_i(t_{k+1})}{m + n_y} = \frac{\sum_1^m x_i(t_k) + \sum_1^{n_y} y_i(t_k)}{m + n_y}. \quad (7)$$

If we initialize y to zero and guarantee that $y(t)$ goes to zero then average consensus is achieved. More specifically, the proposed algorithm can be described as follows.

Set $n_y = m$. Each component $y_i \in \mathbb{R}$ is kept in memory and updated by node i along with its state x_i . At time t_k , a given node i sends a message containing $x_i(t_k)$ and $y_i(t_k)$ to one or more out-neighbors (corresponding to the non-zero

entries of column i of the adjacency matrix N). The node i does not change its state, i.e.,

$$x_i(t_{k+1}) = x_i(t_k) \quad (8)$$

and resets the auxiliary state to zero

$$y_i(t_{k+1}) = 0. \quad (9)$$

Let $n_{\text{out}}(i, k)$ be the number of out-neighbors of node i to which a message is sent at time t_k . A node j receiving this message, updates its state $x_j(t)$ according to

$$x_j(t_{k+1}) = (1 - \alpha)x_j(t_k) + \alpha x_i(t_k) + \beta y_j(t_k) + \gamma y_i(t_k) \quad (10)$$

and updates its variable $y_j(t_k)$ according to

$$y_j(t_{k+1}) = \frac{y_i(t_k)}{n_{\text{out}}(i, k)} + y_j(t_k) + x_j(t_k) - x_j(t_{k+1}) \quad (11)$$

so that the total state average is kept constant, i.e., (7) holds.

The average consensus problem in the two scenarios considered in the present paper, i.e., gossip and broadcast algorithms, can be modeled by a linear distributed algorithm taking the form (6). In fact, the following matrices U_k can model the algorithm (8)-(11), when applied to the communication assumptions for the gossip and broadcast algorithm described in Section II-C.

Gossip algorithm \mathcal{G} : The matrices U_k are taken from the set $\{Q_{ij}, 1 \leq i, j \leq m, i \neq j\}$, where each Q_{ij} corresponds to a transmission from node i to an out-neighbor node j , and these matrices are described as follows. Let $\Lambda_i := \text{diag}(\mathbf{e}_i)$ and $\Omega_{ij} := I - (\Lambda_i + \Lambda_j)$. Then

$$Q_{ij} = \begin{bmatrix} A_{ij} & B_{ij} \\ C_{ij} & D_{ij} \end{bmatrix} \quad (12)$$

where

$$\begin{aligned} A_{ij} &:= I - \alpha\Lambda_j + \alpha\mathbf{e}_j\mathbf{e}_i^\top \\ B_{ij} &:= \beta\Lambda_j + \gamma\mathbf{e}_j\mathbf{e}_i^\top \\ C_{ij} &:= \Lambda_j(I - A_{ij}) \\ D_{ij} &:= \Omega_{ij} + \Lambda_j(I + \mathbf{e}_j\mathbf{e}_i^\top - B_{ij}). \end{aligned} \quad (13)$$

The matrices defined in (12) also model the case where a node i picks itself when there is a clock tick (with probability p_{ii}), in which case we arbitrate that the state update is described by

$$x_i(t_{k+1}) = x_i(t_k) + (\alpha + \beta)y_i(t_k)$$

and

$$y_i(t_{k+1}) = (1 - (\alpha + \beta))y_i(t_k),$$

for which (7) is met. The matrices U_k are by construction independent and identically distributed, and satisfy

$$\text{Prob}[U_k = Q_{ij}] = \frac{1}{m}p_{ij},$$

($\frac{1}{m}$ is the probability that node i is the one whose clock ticks at t_k and p_{ij} the probability that i picks its out-neighbor node j). The times between transmissions $t_{k+1} - t_k$ are independent and identically distributed with exponential distribution with mean $\frac{1}{m}$ (since m nodes can trigger according to a rate 1 Poisson process).

Broadcast Algorithm \mathcal{B} : The matrices U_k are taken from the set $\{R_i, 1 \leq i \leq m\}$, where each R_i corresponds to a transmission from node i to every other node. Let $\Lambda_i := \text{diag}(\mathbf{e}_i)$, $\Omega_i = (I - \Lambda_i)$. Then

$$R_i = \begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} \quad (14)$$

$$\begin{aligned} A_i &= (1 - \alpha)I + \alpha\mathbf{1}_m\mathbf{e}_i^\top \\ B_i &= \Omega_i(\beta I + \gamma\mathbf{1}_m\mathbf{e}_i^\top) \\ C_i &= \Omega_i(I - A_i) \\ D_i &= \Omega_i(I + \frac{\mathbf{1}_m\mathbf{e}_i^\top}{m-1} - B_i). \end{aligned} \quad (15)$$

The matrices U_k are independent and identically distributed due to our assumption that nodes access the network with the same probability, i.e.

$$\text{Prob}[U_k = R_i] = \frac{1}{m}.$$

The times between transmissions $t_{k+1} - t_k$ are independent and identically distributed with an arbitrary distribution imposed by the network protocol.

Hereafter, we denote by *gossip algorithm \mathcal{G}* the linear distributed algorithm modeled by (6) and (12), and denote by *broadcast algorithm \mathcal{B}* the linear distributed algorithm modeled by (6) and (14). Note that, by construction, for both gossip and broadcast algorithms the matrices $\{U_k, k \geq 0\}$ are such that

$$[\mathbf{1}_m^\top \quad \mathbf{1}_m^\top] U_k = [\mathbf{1}_m^\top \quad \mathbf{1}_m^\top], \quad (16)$$

which means that the total average is preserved at each iteration, i.e., $\mathbf{1}_{2m}^\top z(t_{k+1}) = \mathbf{1}_{2m}^\top z(t_k)$, and

$$U_k \begin{bmatrix} \mathbf{1}_m \\ \mathbf{0}_m \end{bmatrix} = \begin{bmatrix} \mathbf{1}_m \\ \mathbf{0}_m \end{bmatrix} \quad (17)$$

which means that if consensus is achieved at iteration k , i.e., if $x(t_k) = c\mathbf{1}_m$ and $y(t_k) = \mathbf{0}_m$, the state remains unchanged at iteration $k+1$, i.e., $x(t_{k+1}) = c\mathbf{1}_m$ and $y(t_{k+1}) = \mathbf{0}_m$.

Example 5: In the same setup of Example 4 and for $\alpha = 1/2$, $\beta = 1/2$, and $\gamma = 1/2$, the matrices (12), for the proposed gossip algorithm are given by

$$Q_{12} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 \\ -1/2 & 1/2 & 1/2 & 1/2 \end{bmatrix}, \quad Q_{21} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 1/2 & 1/2 \end{bmatrix}$$

and the matrices (14) for the broadcast algorithm are given by $R_1 = Q_{12}$ and $R_2 = Q_{21}$. Suppose that node i transmits with probability $p_1 = 1/2$ and node 2 transmits with probability $p_2 = 1/2$. We will show in the next section (cf. Theorem 6) that the necessary and sufficient conditions for convergence in expectation and for convergence in variance are

$$r_1 := r_\sigma(p_1 A_1 + p_2 A_1 - \frac{1}{2} \begin{bmatrix} \mathbf{1}_2 \\ \mathbf{0}_2 \end{bmatrix} [\mathbf{1}_2^\top \quad \mathbf{1}_2^\top]) < 1,$$

$$r_2 := r_\sigma(p_1 A_1 \otimes A_1 + p_2 A_2 \otimes A_2 - S) < 1,$$

$S := \frac{1}{4}(\begin{bmatrix} \mathbf{1}_2 \\ \mathbf{0}_2 \end{bmatrix} \otimes \begin{bmatrix} \mathbf{1}_2 \\ \mathbf{0}_2 \end{bmatrix})([\mathbf{1}_2^\top \quad \mathbf{1}_2^\top] \otimes [\mathbf{1}_2^\top \quad \mathbf{1}_2^\top])$, respectively, where (A_1, A_2) should be replaced by (Q_{12}, Q_{21}) for the gossip algorithm \mathcal{G} and (R_1, R_2) for the broadcast algorithm

B. Computing r_1 and r_2 yields, $r_1 = 0.5$ and $r_2 = 0.542$ and therefore we conclude that consensus is achieved in expectation and second moment (and a fortiori almost surely), which means that this linear distributed algorithm meets the desired requirement of the problem of Example 4, i.e., it achieves average consensus.

IV. MAIN RESULTS

A. Stability

We start by presenting a general result for the convergence analysis of the stochastic linear distributed algorithm (6).

Theorem 6: Consider a linear distributed algorithm (6) where $\{U_k, k \geq 0\}$ are characterized by (16), (17), and are randomly chosen from a set $\mathcal{M} := \{B_i, 1 \leq i \leq n_p\}$, according to

$$\text{Prob}[U_k = B_i] = p_i, \sum_{i=1}^{n_p} p_i = 1. \quad (18)$$

Then, the linear distributed algorithm converges in expectation to average consensus if and only if

$$r_\sigma \left(\sum_{i=1}^{n_p} p_i B_i - \frac{1}{m} \begin{bmatrix} \mathbf{1}_m \\ \mathbf{0}_m \end{bmatrix} \begin{bmatrix} \mathbf{1}_m^\top & \mathbf{1}_m^\top \end{bmatrix} \right) < 1 \quad (19)$$

and converges in second moment to average consensus if and only if

$$r_\sigma \left(\sum_{i=1}^{n_p} p_i B_i \otimes B_i - S \right) < 1, \quad (20)$$

where

$$S := \frac{1}{m^2} \left(\begin{bmatrix} \mathbf{1}_m \\ \mathbf{0}_m \end{bmatrix} \otimes \begin{bmatrix} \mathbf{1}_m \\ \mathbf{0}_m \end{bmatrix} \right) \left(\begin{bmatrix} \mathbf{1}_m^\top & \mathbf{1}_m^\top \end{bmatrix} \otimes \begin{bmatrix} \mathbf{1}_m^\top & \mathbf{1}_m^\top \end{bmatrix} \right). \quad (21)$$

□

Note that both gossip and broadcast algorithm proposed in Section III comply with the framework of Theorem 6, and therefore we can use the conditions (19) and (20) to assert convergence to average consensus of these algorithms.

The next Theorem establishes that the gossip algorithm \mathcal{G} converges in expectation to consensus when the communication topology is symmetric. The proof is postponed to Subsection IV-C.

Theorem 7: For a graph with a symmetric adjacency matrix N , described by (2), there always exists a symmetric doubly stochastic weighted adjacency matrix P (possibly containing self-loops). Moreover, if we set $p_{ij} = [P]_{ij}$, $\alpha = \beta = \gamma = 1/2$, the gossip algorithm \mathcal{G} converges in expectation to consensus.

□

Note that the probabilities p_{ij} can be any as long as $P : [P]_{ij} = p_{ij}$ is a symmetric matrix. An interesting direction for future work is to investigate how to choose the p_{ij} for example, to optimize the convergence rate to average consensus, in a similar way to the distributed algorithm proposed in [11].

For broadcast wireless networks, we have the following result:

Theorem 8: The broadcast algorithm \mathcal{G} converges in expectation to consensus, when $\alpha = \beta = \gamma = 1/2$.

□

B. Convergence rates

The following theorem provides the second moment convergence rate according to Definition 2 of the general stochastic iteration (6), (18) when the time intervals $t_{k+1} - t_k$ are independent and identically distributed.

Theorem 9: Consider the stochastic iteration (6), (18) and suppose that the intervals $t_{k+1} - t_k$ are independent and identically distributed following a distribution ρ . Then, provided that the linear algorithm converges in second moment to average consensus, the second moment convergence rate is given by the unique α that satisfies

$$\int_0^\infty e^{\alpha t} \rho(dt) = \frac{1}{r_1} \quad (22)$$

where $r_1 := r_\sigma \left(\sum_{i=1}^{n_p} p_i B_i \otimes B_i - S \right)$.

□

The proof uses similar arguments to the ones provided for results for a class of system known as impulsive renewal systems [13], and is omitted due to space limitations.

It is possible to prove that r_1 is the convergence rate at which the discrete variable $z^k := x(t_k)$ converges in second moment to consensus, i.e., $(z_i^k - x_{av}) \leq r_1^k (z_i^0 - x_{av})^i$, for every component i of z . Thus, (22) relates the convergence rate in absolute-time t with the convergence rate in discrete-time k which denotes the number of transmissions.

If we specialize the Theorem 9 to the two scenarios that we consider in the present paper, we obtain the following results.

Corollary 10: For the gossip algorithm \mathcal{G} the second moment convergence rate is given by

$$\alpha = \frac{1 + r_1}{m}$$

where $r_1 = r_\sigma \left(\sum_{i=1}^m \sum_{j \in N_{\text{out}}(i)} \frac{p_{ij}}{m} Q_{ij} \otimes Q_{ij} - S \right)$.

□

Corollary 11: For the broadcast algorithm \mathcal{B} , the second moment convergence rate is given by the unique $\alpha > 0$ such that

$$\int_0^\infty e^{\alpha t} \mu(dt) = \frac{1}{r_1}$$

where $r_1 = r_\sigma \left(\frac{1}{m} \sum_{i=1}^m R_i \otimes R_i - S \right)$.

□

C. Proof of Theorem 7

Let $R := \mathbb{E}[U_k] = \sum_{i=1}^m \sum_{j \in N_{\text{out}}(i)} p_{ij} Q_{ij}$. Since $\mathbb{E}[z(t_{k+1})] = R \mathbb{E}[z(t_k)]$, we have that

$$\mathbb{E}[z(t_{k+1})] = \mathbb{E} \left[\begin{bmatrix} x(t_{k+1}) \\ y(t_{k+1}) \end{bmatrix} \right] = R^k z(0) = R^k \begin{bmatrix} x(0) \\ 0 \end{bmatrix}$$

and therefore it suffices to prove that

$$\lim_{k \rightarrow \infty} R^k = \frac{1}{m} \begin{bmatrix} \mathbf{1}_m \\ \mathbf{0}_m \end{bmatrix} \begin{bmatrix} \mathbf{1}_m^\top & \mathbf{1}_m^\top \end{bmatrix} \quad (23)$$

from which we conclude that $\lim_{k \rightarrow \infty} \mathbb{E}[x(t_{k+1})] = 1_m x_{av}$, $x_{av} = 1_m^\top x(0)$. From (12), (13) we notice that we can partition R into blocks $R = \begin{bmatrix} R_1 & R_2 \\ R_3 & R_4 \end{bmatrix}$ where each block is a linear combination of the following three matrices

$$\begin{aligned} X &= \sum_{i=1}^m \sum_{j \in N_{out}(i)} p_{ij} \Lambda_j, & Y &= \sum_{i=1}^m \sum_{j \in N_{out}(i)} p_{ij} \Lambda_i, \\ Z &= \sum_{i=1}^m \sum_{j \in N_{out}(i)} p_{ij} \mathbf{e}_j \mathbf{e}_i^\top. \end{aligned} \quad (24)$$

It is easy to see that $Z = P^\top = P$ (since we assume that P is symmetric) and $Y = I$. Moreover,

$$X = \sum_{j=1}^m \sum_{i \in N_{in}(j)} p_{ij} \Lambda_j = \sum_{j=1}^m \Lambda_j = I,$$

where we used the fact that $\sum_{i \in N_{in}(j)} p_{ij} = 1$, due to the key assumption that $P : P_{ij} = p_{ij}$ is a doubly stochastic matrix. Therefore, each R_i is a linear combination of the matrices P and I and we can write

$$R = P_1 \otimes I_m + P_2 \otimes P.$$

where for $\alpha = \beta = \gamma = \frac{1}{2}$,

$$P_1 = \begin{bmatrix} 1 - \frac{1}{2m} & \frac{1}{2m} \\ \frac{1}{2m} & 1 - \frac{1}{2m} \end{bmatrix}, \quad P_2 = \begin{bmatrix} \frac{1}{2m} & \frac{1}{2m} \\ -\frac{1}{2m} & \frac{1}{2m} \end{bmatrix}. \quad (25)$$

We denote an eigenvalue of a matrix A by $\lambda_i(A)$ and the set of eigenvalues by $\{\lambda_i(A)\}$. Let $P_S(\delta) := P_1 + \delta P_2$. Then one can obtain that

$$\lambda_i(P_S(\delta)) = 1 + \frac{\delta - 2 \pm \sqrt{2 - \delta^2}}{2m}, \quad i \in \{1, 2\}. \quad (26)$$

Let w_{P_i} be the two eigenvector of $P_S(\delta)$, and v_{P_j} denote the m eigenvectors of P (note that P is symmetric and therefore it has m eigenvectors). Then R has $2m$ eigenvectors $w_{P_i} \otimes v_{P_j}$, since one can show that

$$R(w_{P_i} \otimes v_{P_j}) = \lambda_\ell(R) w_{P_i} \otimes v_{P_j}$$

where the set of eigenvalues of R is given by

$$\begin{aligned} \{\lambda_\ell(R), 1 \leq \ell \leq 2m\} &= \{\lambda_i(P_S(\eta_j)) : \eta_j \in \{\lambda_j(P)\}, \\ &1 \leq i \leq 2, 1 \leq j \leq m\} \end{aligned} \quad (27)$$

Since P is symmetric and doubly stochastic, and it is a weighted adjacency matrix of a strongly connected and aperiodic graph, the eigenvalues of P are real, P has a simple eigenvalue at 1, and all the remaining eigenvalues belong to the set $(-1, 1)$. Corresponding to the simple eigenvalue 1 of P , R has two eigenvalues at $\{\lambda_i(P_1 + P_2)\} = \{1, 1 - 1/m\}$. Corresponding to the eigenvalues of P that belong to the set $(-1, 1)$, the eigenvalues of R are inside the unit circle. This can be shown by noticing that (26) is a strictly increasing function when $-1 < \delta < 1$ for each i and, using this fact, it is easy to conclude that $r_\sigma(P_1 + \delta P_2) < 1$ for $-1 < \delta < 1$. Thus R has a single eigenvalue at 1, all the remaining eigenvalues are inside the unit disk, and the

vectors $v_R := \begin{bmatrix} 1_m \\ 0_m \end{bmatrix}$ and $w_R := [1_m^\top \quad 1_m^\top]$ are left and right eigenvalues of R , respectively, associated with this eigenvalue 1. This implies that $\lim_{k \rightarrow \infty} R^k = \frac{1}{w_R v_R} v_R w_R$, which is (23). ■

V. CONCLUSIONS AND FUTURE WORK

We addressed average consensus and randomized gossip algorithms with the constraint that only one node transmits data at a given time. Our main result is to provide a linear distributed gossip algorithm, that provably converges to consensus in expectation. Directions for future work include: (i) extending the results to the stronger stability notion of convergence in second moment; (ii) considering general digraphs, i.e., not requiring the communication topology to be symmetric; (iii) optimizing the probability parameters of the gossip algorithms with respect to some quantity of interest as the speed of convergence to consensus.

VI. ACKNOWLEDGMENTS

The first author is grateful to João Almeida, for valuable comments and discussions.

REFERENCES

- [1] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *Automatic Control, IEEE Transactions on*, vol. 31, no. 9, pp. 803 – 812, Sep. 1986.
- [2] B. Johansson, T. Keviczky, M. Johansson, and K. Johansson, "Sub-gradient methods and consensus algorithms for solving convex optimization problems," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, 2008, pp. 4185 –4190.
- [3] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," in *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, vol. 3, 2002, pp. 2953 – 2958 vol.3.
- [4] J. Cortes, S. Martinez, and F. Bullo, "Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions," *Automatic Control, IEEE Transactions on*, vol. 51, no. 8, pp. 1289 –1298, 2006.
- [5] M. Chiang, S. Low, A. Calderbank, and J. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255 –312, 2007.
- [6] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1520 – 1533, 2004.
- [7] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009, electronically available at <http://coordinationbook.info>.
- [8] S. Patterson, B. Bamieh, and A. El Abbadi, "Convergence rates of distributed average consensus with stochastic link failures," *Automatic Control, IEEE Transactions on*, vol. 55, no. 4, pp. 880 –892, 2010.
- [9] F. Fagnani and S. Zampieri, "Average consensus with packet drop communication," *SIAM J. Control Optim.*, no. 48, pp. 102–133, 2009.
- [10] R. Carli, F. Bullo, and S. Zampieri, "Quantized average consensus via dynamic coding/decoding schemes," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 2, pp. 156–175, 2010. [Online]. Available: <http://dx.doi.org/10.1002/rnc.1463>
- [11] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *Information Theory, IEEE Transactions on*, vol. 52, no. 6, pp. 2508 – 2530, 2006.
- [12] K. Cai and H. Ishii, "Gossip consensus and averaging algorithms with quantization," in *2010 American Control Conference*, 2010, pp. 6306 –6311.
- [13] D. Antunes, J. P. Hespanha, and C. Silvestre, "Volterra integral approach to impulsive renewal systems: Application to networked control," Nov. 2010, submitted to journal publication.