Abstract—Motivated by applications to wireless sensor, peer-to-peer, and ad hoc networks, we study distributed broadcasting algorithms for exchanging information and for computing in an arbitrarily connected network of nodes. Specifically, we propose a broadcasting-based gossiping algorithm to compute the (possibly weighted) average of the initial measurements of the nodes at every node in the network. We show that the broadcast gossip algorithms almost surely converge to a consensus. In addition, the random consensus value is, in expectation, equal to the desired value, i.e., the average of initial node measurements.

However, the broadcast gossip algorithms do not converge to the initial average in absolute sense because of the fact that the sum is not preserved at every iteration. We provide theoretical results on the mean square error performance of the broadcast gossip algorithms. The results indicate that the mean square error strictly decreases through iterations until the consensus is achieved. Finally, we assess and compare the communication cost of the broadcast gossip algorithms to achieve a given distance to consensus through numerical simulations.

Index Terms—Distributed average consensus, broadcasting, sensor networks, gossip algorithms.

I. INTRODUCTION

A fundamental problem in decentralized networked systems is that of having nodes reach a state of agreement [1]–[3]. For example, the nodes in a wireless sensor network must be synchronized in order to communicate using a TDMA scheme or to use time-difference-of-arrival measurements for localization and tracking. Similarly, one would like a host of unmanned aerial vehicles to make coordinated decisions on a surveillance strategy. This paper focuses on a prototypical example of agreement in asynchronous networked systems, namely, the randomized average consensus problem: each node initially has a scalar value, $y_i$, and the goal is to compute the average, $1/N \sum_{i=1}^{N} y_i$, at every node in the network.

A. Related Work

Gossip-based algorithms initially introduced by Tsitsiklis [4] and extensively studied by other researchers [1]–[3], [5], [6]. Randomized average consensus gossiping is an asynchronous time model where each node, at uniform random, wakes up, chooses a neighbor within the connectivity radius with some probability and performs a pairwise averaging. The sum is preserved at each iteration, so does the mean. It is shown that the algorithm converges to a consensus if the graph is strongly connected on the average. In this case, the transmitting node sends its packets to the chosen neighbor then wait for the neighbor’s packet. This scheme is vulnerable to packet collisions and yields a communication complexity (number of radio transmissions to achieve $\Theta(N^{-\alpha})$, for any $\alpha > 0$, distance to consensus) in the order of $\Theta(N^2)$ ($N$ denotes the number of nodes) over random geometric graphs [5].

Recently proposed geographic gossip algorithm combines gossip with geographic routing [7]. Similar to the gossip algorithms, a node randomly wakes up, chooses a node randomly in the whole network, rather than in its neighborhood and performs a pairwise averaging with this node, and, therefore, geographic gossiping increases the diversity of every pairwise averaging operations. The authors show that the communication complexity, which is in the order of $O(N^{3/2}/\log(N))$, is improved with respect to the standard gossiping algorithm. However, in addition to problems similar to the ones mentioned for the standard gossiping algorithms, in geographic gossip, each node needs to know its own location, learn and memorize locations of its neighbors. Moreover, at every update, nodes responsible for routing the packet, need to determine which of their neighboring nodes is closest to the target node, and route the packet in that direction which are nontrivial tasks requiring additional computations and resources growing with $N$.

B. Summary of Main Contributions

Geographic gossiping improves upon the convergence speed of the standard gossip by increasing the diversity of pairwise exchanges. However, it does not mitigate the major bottleneck associated with the fact that the messages between two peers need to be routed and exchanged to perform two updates. Setting up a two-way route successfully only aggravates the problem. Wireless media have the advantage of being broadcast and, at the cost of one transmission, one can reach several terminals. Our objective in this paper is to propose and analyze a broadcasting-based gossip algorithm that enables all nodes in range to perform an update by exploiting the wireless medium, and thereby avoiding the need of complex routing and pairwise exchanging operations.

Thus, to overcome the drawbacks of the standard packet based gossip algorithms, we propose broadcast gossip algorithms suitable for wireless sensor networks. In the proposed algorithm, a node in the network, at uniformly random, wakes up (asynchronous time model), broadcasts its value, and this value is successfully received by the nodes in the predefined radius of the broadcasting node, i.e., connectivity radius. The nodes that have received the broadcasted value updates their own state value, the remaining nodes sustain their value, and the algorithm iterates. It is shown here that this type of
gossiping algorithms are capable of achieving consensus over the network with probability one. Moreover, we show that the random consensus value is, in expectation, equal to the desired value, i.e., the average of initial node measurements.

However, the sum of the node state values is not preserved at each iteration, hence, the broadcasting type gossiping algorithms converge to a value that is in the neighborhood of the desired value. We provide theoretical results on the mean square error performance of the broadcast gossip algorithms. Finally, we present simulations highlighting the features of the proposed algorithm. Specifically, we show that broadcast gossip algorithms, in addition to being more practical and suitable for wireless sensor networks, provide significant decrease in communication complexity over randomized and geographic gossip algorithms.

C. Paper Organization

The remainder of this paper is organized as follows. Section II introduces the graph and time models adopted in this paper, and summarizes the average consensus problem. The proposed broadcast gossip algorithms are detailed in Section III along with their theoretical analysis and numerical examples highlighting their features. Finally, we conclude with Section VI.

II. GRAPH AND TIME MODELS

In the following, we briefly discuss the graph and time models adopted in this paper. Then, we summarize the distributed average consensus problem.

A. Graph Model

 Following previous work, we model our wireless sensor network as a random geometric graph. In this model, denoted $G(N, R)$, the $N$ sensor locations are chosen uniformly and independently in the unit square, and each pair of nodes is connected if their Euclidean distance is smaller than some transmission radius $R$, i.e., connectivity radius [8]. It is well known that in order to have good connectivity and minimize interference, the transmission radius $R$ has to scale like $\Theta(\sqrt{\log(N)/N})$ [5], [8]. The $N$-node topology is represented by the $N \times N$ adjacency matrix $\Phi$, where for $i \neq j$, $\Phi_{ij} = 1$ if nodes $i$ and $j$ are in their neighborhood, and $\Phi_{ij} = 0$, otherwise. Moreover, we define $\mathcal{N}_i = \{j \in \{1, 2, \ldots, N\} : \Phi_{ij} \neq 0\}$. For our analysis, we assume that communication within this transmission radius always succeeds and that radius $R$ yields, on the average, a strongly connected graph. Note however that the proposed algorithm is robust to communication and node failures.

B. Time Model

We use the asynchronous time model, which is well-matched to the distributed nature of sensor networks [5], [7]. More precisely, it is assumed that each sensor node has a clock which ticks independently as a rate $\mu$ Poisson process. Consequently, the inter-tick times are exponentially distributed, and independent across nodes and across time.

This set-up is equivalent to a single clock ticking according to a rate $N\mu$ Poisson process at times $Z_k$. On average, there are approximately $N$ clock ticks per unit of absolute time but we will always be measuring time in number of ticks of this (virtual) global clock. Time is discretized, and the interval $[Z_k; Z_{k+1})$ corresponds to the $k$-th timeslot. We can adjust time units relative to the communication time so that only one broadcasting occurs in the network at each time slot with high probability.

C. Average Consensus

At time slot $k \geq 0$, each node $i = 1, 2, \ldots, N$ has an estimate $x_i(k)$ of the global average, and we use $x(k)$ to denote the $N$-vector of these estimates. The ultimate goal is to drive the estimate $x(k)$ to the average vector $\bar{x}(0)$ (1, denoting the vector of ones) or as close as possible, where $\bar{x}(0) = 1/N \sum_{i=1}^{N} x_i(0)$, using the minimal amount of communication. For the algorithms of interest to us, the quantity $x(k)$ for $k > 0$ is a random vector, since the algorithms are randomized in their behavior.

III. BROADCAST BASED GossipING

Suppose $t$-th clock to tick belongs to node $i \in \{1, 2, \ldots, N\}$. Node $i$ activates and the followings occur over the network:

- Node $i$ broadcasts its current state value, $x_i(t)$ over the network.
- The broadcasted value is successfully received by the nodes that are in the radius $R$.
- The neighboring nodes, the index set of which is given by $\mathcal{N}_i$ receive the broadcasted value $x_i(j)$, and update their state values as $x_k(t+1) = \gamma x_k(t) + (1-\gamma)x_i(j)$, $\forall k \in \mathcal{N}_i$. (1) with $\gamma \in (0,1)$ denoting the mixing parameter.
- The remaining nodes in the network, including $i$, update their state values as $x_k(t+1) = x_k(t), \forall k \notin \mathcal{N}_i$. (2)

The asynchronous broadcast gossip algorithms are described as follows: In the $t$-th time-slot, suppose node $i$’s clock tick. Then, node $i$ broadcasts its own state value to neighboring nodes. Once the broadcasted value is received, all the neighboring nodes set their values equal to the (weighted) average of their current value and the value broadcasted by the node $i$. Formally, let $x(t)$ denote the vector of values at the end of the time-slot $t$. Then,

$$x(t+1) = W(t)x(t)$$

where the random matrix $W(t)$, with probability 1/$N$ is (assuming that the $i$-th clock ticks)

$$W_{jk}^{(i)} = \begin{cases} 
1 & j \notin \mathcal{N}_i, k = j \\
\gamma & j \in \mathcal{N}_i, k = j \\
1-\gamma & j \in \mathcal{N}_i, k = i \\
0 & \text{elsewhere}
\end{cases}$$

(4)
where \( W^{(i)} \) denotes the weight matrix corresponding to the case where node \( i \)'s clock ticks.

The following lemma discusses two important properties of the weight matrices.

**Lemma 1.** The weight matrices \( \{ W^{(i)} : i = 1, 2, \ldots, N \} \) satisfy the followings

(i) \( 1 \) is a right eigenvector of all \( W^{(i)} \), i.e., \( W^{(i)} \mathbf{1} = \mathbf{1} \), \( \forall i \).

(ii) \( 1 \) is not a left eigenvector of any \( W^{(i)} \), i.e., \( 1^T W^{(i)} \neq 1^T \), \( \forall i \).

**Proof:** Let us consider the first item. It suffices to show that all rows of all \( W^{(i)} \) matrices sum to unity. We have

\[
\sum_{k=1}^{N} W^{(i)}_{jk} = \mathbf{1} \left\{ j \in N_i \right\} (\gamma + (1 - \gamma)) + \mathbf{1} \left\{ j \notin N_i \right\} 1 = 1
\]

where \( \mathbf{1} \{ \cdot \} \) is the indicator function. Thus, the proof of first item is complete.

Let us consider the second item of the Lemma in the following. Note that

\[
\sum_{j=1}^{N} W^{(i)}_{jk} = 1 + (1 - \gamma)|N_i|, \quad \text{for } k = i.
\]

Since we assume strongly connected graphs, we have \( |N_i| \geq 1 \) indicating that \( \sum_{j=1}^N W^{(i)}_{jk} |k = i > 1 \). This, in turn shows that, \( \forall i \), there exists at least one column, namely the \( k = i \) column, with sum different than one, which implies that \( 1^T W^{(i)} \neq 1^T \), \( \forall i \) and concludes the proof.

The above lemma reveals that \( c \mathbf{1} \) for some \( c \in \mathbb{R} \) is a fixed point of the broadcasting gossip algorithm, thus, \( W^{(i)} \mathbf{1} = c \mathbf{1} \), \( \forall i \). If the algorithm converges to a consensus, the broadcasting gossip do not leave the consensus state. However, it also shows that the sum (and therefore the average) of the vector of node values is not preserved at each step.

In fact, suppose \( A(t) = k \) with \( A(t) \) denoting the node that is broadcasting at the time slot of interest, then it is easy to check that the discrepancy between the sum at the next and current time-slots is nonzero whenever \( x_k(t) \neq |N_k|^{-1} \sum_{i \in N_k} x_i(t) \). Moreover, the sum difference between time-slots is bounded by \( w.l.o.g. \) suppose that \( |x_k(t) - \min_{i \in N_k} x_i(t)| \leq \gamma \). Then, the sum difference between consecutive time-slots is bounded by \( (1 - \gamma) |N_k| x_k(t) - \min_{i \in N_k} x_i(t) \).

It is of interest to note that the sum difference between consecutive time-slots, thus, is small if the node state values are close to each other.

Let us denote the mean of i.i.d. \( W(t) \) as \( \mathbb{E}\{W(t)\} = W \). The following lemma gives some properties of the average weight matrix that would prove useful for the remainder of the paper. Moreover, they are also independently of interest.

**Lemma 2.** The average weight matrix \( W \) is given by

\[
W = I - \frac{1 - \gamma}{N} \text{diag}\{\Phi\} + \frac{1 - \gamma}{N} \Phi
\]

and, for all \( \gamma \), satisfies the followings:

\[
W_1 = 1, \quad 1^T W = 1^T, \quad \rho(W - J) < 1 \tag{9}
\]

where \( \rho(\cdot) \) denotes the spectral radius of its argument and \( J = (N)^{-1} 11^T \).

**Proof:** First we note that

\[
W_{jk} = \frac{1}{N} \sum_{i=1}^{N} W^{(i)}_{jk}.
\]

Then, by (4), we have

(11)

Therefore (8) follows. Note that (8) is the representation of the weight matrix in terms of graph Laplacian \( L \), i.e., \( W = I - \eta L \) where \( L = \text{diag}\{\Phi\} - \Phi \) and \( \eta = 1 - \gamma/N \). Since \( 0 < \eta < 1/N - 1 \) for all \( \gamma \in (0, 1) \), \( W \) satisfies the conditions given in (8).

The Lemma shows that, unlike the individual weight matrices, \( 1 \) is both a left and right eigenvector of the average weight matrix. Moreover, the spectral radius of the weight matrix is less than unity, a property that will prove useful throughout the rest of the paper.

**IV. CONVERGENCE OF BROADCAST Gossip**

In this section, we will study the convergence of asynchronous broadcast gossip algorithms. We will not restrict ourselves here to any particular algorithm; but rather consider convergence of the iteration governed by a product of random matrices, each of which satisfies certain properties discussed in the previous section.

**A. Convergence in the Expectation**

We consider the convergence in expectation of the broadcast gossip algorithm in the following. The next result reveals that, although the sum is not preserved per iteration, it is preserved in expectation.

**Proposition 1.** The limiting random vector obtained through broadcast gossip iterations is, in expectation, equal to the average of initial node measurements, i.e.,

\[
\mathbb{E}\left\{ \lim_{t \to \infty} x(t) \right\} = \frac{1}{N} \mathbf{1}^T x(0). \tag{12}
\]

**Proof:** By Lebesgue dominated convergence theorem, we have

\[
\mathbb{E}\left\{ \lim_{t \to \infty} x(t) \right\} = \lim_{t \to \infty} \mathbb{E}\{x(t)\}. \tag{13}
\]

Moreover, since \( W(t) \)’s are i.i.d., we obtain

\[
\lim_{t \to \infty} \mathbb{E}\{x(t+1)\} = \lim_{t \to \infty} \mathbb{E}\{W(t)x(0)\}. \tag{14}
\]

We, hence, see that we need to have \( \lim_{t \to \infty} W(t) = (N)^{-1} 11^T \). The conditions required to have this property are given by \( W_1 = 1, \ 1^T W = 1^T, \ \rho(W - J) < 1 \) [5]. Since
Lemma 2 indicates that these conditions are satisfied, the proof is complete.

The proposition indicates that the limiting random vector that the broadcasting gossip algorithm achieves is, as desired, in expectation, equal to the initial node measurement average vector.

B. Convergence in Second Moment

Since the sum is not preserved throughout broadcast gossip iterations, instead of tracking the distance to the average like the gossip-based algorithms, we use a more useful measure of consensus for the sequence $x(t)$. We define $\beta(t)$ to be the vector of deviations of the components of $x(t)$ from their average. This can be expressed in component form as $\beta_i(t) = x_i(t) - \overline{x}(t)$, or as

$$\beta(t) = x(t) - Jx(t) = (I - J)x(t). \quad (15)$$

This is a measure of relative deviation of the node values from their average. Let $\| \cdot \|_2$ and $\mathbb{N}$ denote the $\ell_2$ norm of its argument and the set of extended integer numbers, respectively. Now, consider the following Lemma that gives necessary and sufficient conditions on $\beta(t)$ that would indicate that there is a consensus.

**Lemma 3.** There is a consensus at time-slot $T \in \mathbb{N}$ of broadcast gossip iterations if and only if $\mathbb{E}\{\|\beta(T)\|_2^2\} = 0$.

**Proof:** If there is a consensus at time-slot $T$, then, $x(T) = c1$ for some $c \in \mathbb{R}$. Then,

$$\beta(T) = (I - J)c1 = c1 - \frac{c}{N}11^T1 = 0 \quad (16)$$

since $1^T1 = N$. By the properties of the norm and statistical expectation, we have $\beta(T) = 0 \Rightarrow \|\beta(T)\|_2^2 = 0 \Rightarrow \mathbb{E}\{\|\beta(T)\|_2^2\} = 0$. Now, if $\mathbb{E}\{\|\beta(T)\|_2^2\} = 0$, then, since $\|\beta(T)\|_2^2 \geq 0$, we have $\|\beta(T)\|_2^2 = 0 \Rightarrow \beta_i(T) = 0$, $\forall i$. Thus, $x_i(T) = \overline{x}(T) = 0$, $\forall i \Rightarrow x_i(T) = \overline{x}(T)$, $\forall i$ indicating that $x(T) = c1$ where $c = \overline{x}(T)$.

Thus, if the expectation of the norm of the deviation vector converges to zero then the node values converge to a consensus. Let $\lambda_i(\cdot)$ denote the $i$th ranked eigenvalue of its argument. In the following, we present a sufficient condition guaranteeing the convergence of the expectation of the deviation vector norm to zero.

**Lemma 4.** The expectation of the norm of the deviation vector of the broadcast gossip, i.e., $\mathbb{E}\{\|\beta(T)\|_2^2\}$, converges to zero if

$$\lambda_1(\mathbb{E}\{W(t)^T(I - J)W(t)\}) < 1 \quad (17)$$

where $I$ denotes the identity matrix.

**Proof:** Utilizing the properties of $W(t)$ matrices, we find that the deviation vector obey the following recursion with probability one:

$$\beta(t + 1) = (W(t) - JW(t))\beta(t). \quad (18)$$

Of note is that this iteration is different that the one tracking the distance to initial node measurements average in gossip-based algorithms as expected since the algorithm and the aim of the proof is different.

Denoting $Y(t) = (W(t) - JW(t))$ to simplify the presentation yields $\beta(t + 1) = Y(t)\beta(t)$. Now, taking the expected norm of $\beta(t + 1)$ given $\beta(t)$ and using the fact that $\|u\|_2^2 = u^Tu$ for $u \in \mathbb{R}^N$, yields

$$\mathbb{E}\{\|\beta(t + 1)\|_2^2|\beta(t)\} = (\beta(t)^T\mathbb{E}\{Y(t)^TY(t)\}\beta(t) \leq \lambda_1(\mathbb{E}\{Y(t)^TY(t)\})\|\beta(t)\|_2^2 \quad (19)$$

where the last line follows from the Rayleigh-Ritz theorem and using the fact that all $Y(t)^TY(t)$ matrices are symmetric. Then, repeatedly conditioning and using the linear iteration obtained above, we have,

$$\mathbb{E}\{\|\beta(t)\|_2^2\} \leq \lambda_1(\mathbb{E}\{Y(t)^TY(t)\})\|\beta(0)\|_2^2. \quad (20)$$

Thus, now we can note that $\lim_{t \to \infty} \mathbb{E}\{\|\beta(t)\|_2^2\} = 0$ if $\lambda_1(\mathbb{E}\{Y(t)^TY(t)\}) < 1$. The sufficiency condition then reduces to, after some algebraic manipulations, the one stated in the Lemma.

Note that this condition, $\lambda_1(\mathbb{E}\{W(t)^T(I - J)W(t)\}) < 1$, due to the fact that it is derived for ensuring the convergence of the deviation vector not the distance to initial node measurements average vector and for a different algorithm, is different than the convergence condition obtained for the standard gossip algorithms where one only need to have $\lambda_2(\mathbb{E}\{W(t)^TW(t)\}) < 1$ to ensure the second-order convergence to the initial node measurements average [5], [7].

The following lemma which consider the expectation of the second moment of the weight matrices will prove useful for stating the main result of this subsection.

**Lemma 5.** The second moment of the weights matrices, denoted as $W' = \mathbb{E}\{W(t)^TW(t)\}$, is given by

$$W' = I - \frac{2\gamma(1 - \gamma)}{N}\text{diag}\{\Phi\} + \frac{2\gamma(1 - \gamma)}{N}\Phi \quad (22)$$

and, for all $\gamma$, $W'$ satisfies the properties given in (9).

**Proof:** From the per-node weight matrices, we obtain

$$\{W(i)^TW(i)\}_{jk} = \begin{cases} 
1 + |N_i|(1 - \gamma)^2 & k = j = i \\
\gamma(1 - \gamma)^2 & k \in N_i, j = i \\
\gamma^2 & j \in N_i, k = j \\
\gamma(1 - \gamma) & j \in N_i, k = j \\
1 & j \notin N_i, k = j \\
0 & \text{elsewhere}
\end{cases} \quad (23)$$

where $W(i)$ denotes the weight matrix corresponding to the case where node $i$’s clock ticks. Therefore, the average is

$$W'_{jk} = \begin{cases} 
1 - \frac{|N_i|}{N}(1 - (1 - \gamma)^2 - \gamma^2) & k = j \\
\frac{2\gamma(1 - \gamma)}{N} & k \in N_j, j \in N_k \\
\frac{2\gamma(1 - \gamma)}{N} & j \notin N_i, k = j \\
0 & \text{elsewhere}
\end{cases} \quad (24)$$

Then, (22) follows. As we noted before, (22) is the representation of $W'$ in terms of graph Laplacian. Since $0 <
2γ(1 − γ)/N < 1/N − 1 for all γ, W′ satisfies the properties given in (9).

An interesting observation is that W′ is equal to W for γ = 1/2. In the following, we show that the broadcasting gossip algorithm satisfies the sufficiency condition required to achieve consensus in the second moment.

**Proposition 2.** The broadcast gossip algorithms satisfy the fact that λ1(E{W(t)T(I − J)W(t)}) < 1.

Proof: First, note that the eigenvalue of interest is the maximum eigenvalue of expectation over positive semidefinite matrices since (W(i))T(I − J)W(i) = ((I − J)W(i))T((I − J)W(i)). This indicates that λ1(E{W(t)T(I − J)W(t)}) ≥ 0. Moreover, let W′ = E{W(t)T(I − J)W(t)} and observe the following

![Math equation]

where the above follows from Lemma 5 and the variational definition of eigenvalues (Note that W′ − W′ is a symmetric matrix). Recall that max1≤i≤N ||w′i||2 = 1 for u = u1 = 1/√N1 which is the eigenvector corresponding to the unit eigenvalue. Of note is that for all \{u : u ∈ RN, ||u||2 = 1, u ≠ u1\}, we have uT W′u < 1 which implies that λ1(E{W(t)T(I − J)W(t)}) < 1 since uT W′u ≥ 0 for all u ∈ RN (Note that the expectation is taken over positive semidefinite matrices). Thus, the proof reduces to show that for u = u1, we still have λ1(W′ − W′) < 1. For u = u1, equation (25) reduces to

![Math equation]

where the last inequality follows from the fact that uT W′u1 > 0 since all entries of the W′ matrix is nonnegative (note that the expectation is taken over nonnegative entry matrices). Thus, uT W′u − uT W′u < 1 for all \{u : u ∈ RN, ||u||2 = 1\}, indicating that max1≤i≤N ||w′i||2 = 1

![Math equation]

\(\lim_{t\to\infty} x(t) = \mathbf{c} 1\)

for some \(c \in \mathbb{R}\) where

![Math equation]

\(\mathbb{E}\{c\} = \frac{1}{N} \mathbf{1}^T x(0)\).

Proof: At this stage of development, we just need to put the pieces together. Observe the followings for any \(\epsilon > 0\):

![Math equation]

where (a) follows from Markov’s Inequality, (b) follows from Lebesgue dominated convergence theorem and (c) from Lemma 3. The above set of equations indicate that the consensus is almost surely achieved, i.e., \(\mathbb{P}\{\lim_{t\to\infty} x(t) = \mathbf{c} 1\} = 1\). Finally, the proof, after recalling Proposition 1, is completed.

The theorem indicates that the broadcasting gossip algorithms achieve consensus with probability one, and the consensus value is, in expectation, equal to the desired value, i.e., average of initial node measurements.

V. PERFORMANCE ANALYSIS OF BROADCAST Gossip ALGORITHMS

In this section, we first consider the mean-square error performance of the broadcast gossip algorithms and study the effect of the mixing parameter on the convergence and mean-square error. We derive an upper bound on the limiting mean-square error performance. Moreover, we prove an upper bound on the discrete time (or equivalently, number of clock ticks) required to get within \(\epsilon\) of the consensus constraint \(c 1\), \(c \in \mathbb{R}\). Finally, we examine the communication complexity of the broadcast gossip algorithms to achieve a certain distance to consensus.

A. Mean Square Error

Since, in general, the broadcast gossip algorithms do not converge to the initial node measurements average, i.e., \(c ≠ (N)^{-1} \mathbf{1}^T x(0)\), it is of interest to consider the distance of the consensus value to \(\mathbb{1}(0)\). In the remaining, we use

![Math equation]

denote the difference between the state vector at time step \(t\) and the average of initial node measurements.

**Lemma 6.** Let \(\mathbb{E}\{||a(t)||_2^2\}\) denote the mean square error at time step \(t\). Then,

(i) The mean square error iteration obeys a recursion given as:

![Math equation]

(ii) The following holds:

![Math equation]

for some \(c \in \mathbb{R}\).
Proof: It is easy to see that $\alpha(t+1) = W(t)\alpha(t)$ yielding the following recursion for the second moment:

$$\mathbb{E}\{\alpha(t+1)^T\alpha(t+1)|\alpha(t)\} = \alpha(t)^T\mathbb{E}\{W(t)^TW(t)\}\alpha(t)$$

$$= \alpha(t)^TW'\alpha(t) \quad (37)$$

$$= y(t)^T\Lambda y(t) \quad (38)$$

where we utilize the eigendecomposition of $W' = \Lambda VV^T$ and define $y(t) = V^T\alpha(t)$. Of note is that given $\alpha(t)$, $y(t)$ is also given. Then, we have the followings:

$$\mathbb{E}\{\alpha(t+1)^T\alpha(t+1)|\alpha(t)\}$$

$$= \sum_{i=1}^{N} \lambda_i(W')|y_i(t)|^2 \quad (40)$$

$$= |y_1(t)|^2 + \sum_{i=2}^{N} \lambda_i(W')|y_i(t)|^2 \quad (41)$$

$$= (1 - \lambda_2(W'))|y_1(t)|^2 + \lambda_2(W')|y_1(t)|^2 + \sum_{i=2}^{N} \lambda_i(W')|y_i(t)|^2$$

$$\leq (1 - \lambda_2(W'))|y_1(t)|^2 + \lambda_2(W')\sum_{i=1}^{N} |y_i(t)|^2 \quad (42)$$

$$= (1 - \lambda_2(W'))|y_1(t)|^2 + \lambda_2(W')|y_1(t)|^2 \quad (43)$$

$$= (1 - \lambda_2(W'))||J\alpha(t)||_2^2 + \lambda_2(W')||\alpha(t)||_2^2 \quad (45)$$

where the last line follows from the facts that $||y(t)||_2^2 = y(t)^Ty(t) = \alpha(t)^TVV^T\alpha(t) = \alpha(t)^T\alpha(t) = ||\alpha(t)||_2^2$ due to unitary decomposition and $|y_1(t)|^2 = |u_1^T\alpha(t)|^2 = (N)^{-1}\alpha(t)^T11^T\alpha(t) = \alpha(t)^TJ\alpha(t) = ||J\alpha(t)||_2^2$ due to the fact that $v_1 = (\sqrt{N})^{-1}1$. This concludes the proof of the first item.

Let us now consider the second item. Note that $J$ is a paracontracting matrix with respect to $\ell_2$ norm since its symmetric and all its eigenvalues are in $(-1, 1)$. Thus, we have

$$Jx \neq x \Leftrightarrow ||Jx||_2^2 < ||x||_2^2. \quad (46)$$

Thus, if we can show that $J\alpha(t) = \alpha(t)$ if and only if $x(t) = c1$ for some $c \in \mathbb{R}$, we are done. If $x(t) = c1$, then

$$J\alpha(t) = Jx(t) - Jx(0) = x(t) - Jx(0) = \alpha(t) \quad (47)$$

where we used the facts that $J^2 = J$ and $Jx(t)|_{x(t)=c1} = x(t)$. Now, if $\alpha(t) = J\alpha(t)$, then $\alpha_i(t) = (N)^{-1}\sum_{i=1}^{N} \alpha_i(t)$. Thus, $\alpha(t) = \pi(t)1$. Since

$$x(t) = \alpha(t) + Jx(0) = \pi(t)1 + \pi(0)1 = (\pi(t) + \pi(0))1 \quad (48)$$

we are done. Therefore, the proof of the second item is complete.

The above Lemma reveals that the mean square error (MSE) recursion is a function strictly decreasing with time, and, converges when nodes converges to a consensus.

In the following, through simulations, we investigate the per-node MSE and variance performance of the broadcast gossip algorithms with respect to the parameter $\gamma$. The nodes are uniformly distributed over a unit square and their initial values are initialized as uniformly distributed random values with unit variance. Moreover, the connectivity radius is chosen as $R = \sqrt{\log(N)/N}$. Figure 1 gives the per-node MSE $(\langle N \rangle^{-1}||x(t) - Jx(0)||_2^2)$ and per-node variance $(\langle N \rangle^{-1}||x(t) - Jx(t)||_2^2)$ for varying $\gamma \in \{0.1, 0.3, \ldots, 0.9\}$ values at the 5000th clock tick. The plotted data points are ensemble average of 100 trials. It is interesting to note that per-node MSE performance appears to decrease with increasing $\gamma$. Conversely, the per-node variance at a given (large enough) iteration number appears to increase with increasing $\gamma$, i.e., the broadcast gossip is further away from consensus. Thus, the plot suggests that there is a trade-off between the per-node MSE and time taken to achieve consensus, and, in turn, total communication cost.

B. Communication Cost to Achieve Consensus

In gossip-type algorithms, it is of crucial importance to achieve consensus with minimal communication cost [5], [7], [8]. Since per iteration, broadcast gossip performs number of neighbors many state value update, one would expect significant decrease in the communication costs to achieve a given distance from consensus, compared to previously reported gossiping algorithms. Theoretical results proving this intuition is under investigation. However, in the following, we provide realistic numerical examples supporting this observation.

In the following, as in [7], we compare the number of radio transmissions to achieve a certain distance from consensus of broadcast gossiping (with $\gamma = 1/2$, since this value appears to give the best empirical trade-off between MSE and convergence speed) with standard [5] (with uniform neighbor selection probability giving the optimal scaling behavior) and geographic [7] gossiping algorithms for various network sizes. Figure 2 depicts per-node variance versus the number of radio transmissions (each data point is an ensemble average of 100...
trials). The simulation configuration and setup is as in Section V-A. Of note is that each iteration requires one, two and the number of hops many radio transmissions, respectively, for broadcast, standard and geographic gossiping. Simulation results suggest that broadcast gossiping outperforms both protocols from the communication cost perspective. We also note that we exclude some complexities in geographic gossiping protocol, such as costs due to memory and routing operations.

VI. CONCLUDING REMARKS

In this paper, we proposed and studied broadcast gossip algorithms capable of exploiting the wireless media. We present conditions on the weights matrices that would guarantee convergence to consensus, and we show that the broadcast gossip algorithms achieve consensus with probability one. Moreover, the random consensus value is, in expectation, equal to the desired value, i.e., the average of initial node measurements. Noting that the network sum is not preserved at each broadcasting time-slot, we provide some theoretical and simulation results on the mean square error performance. Finally, we present numerical examples evaluating and comparing the communication cost of gossiping algorithms required to achieve a given distance to consensus.

Of note is that, once can think some of the results presented here as a generalization of the gossip algorithms to the case where the communication is restricted to one-way but multiple state value updates (around the neighborhood of broadcasting node) occur at each iteration.

REFERENCES
