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

Broadcast Gossip Algorithms: Design and Analysis for Consensus

Tuncer C. Aysal, Mehmet E. Yildiz, Anand D. Sarwate, and Anna Scaglione

***Abstract*— Motivated by applications to wireless sensor, peer-to-peer, and ad hoc networks, we have recently proposed a broadcasting-based gossiping protocol to compute the (possibly weighted) average of the initial measurements of the nodes at every node in the network. The class of *broadcast gossip algorithms* achieve consensus almost surely at a value that is in the neighborhood of the initial node measurements' average.**

In this paper, we further study the broadcast gossip algorithms: we derive and analyze the optimal mixing parameter of the algorithm when approached from worst-case convergence rate, present theoretical results on limiting mean square error performance of the algorithm, and find the convergence rate order of the proposed protocol.

I. Introduction

A fundamental problem in decentralized networked systems is that of having nodes reach a state of agreement [1], [2]. Distributed agreement is a fundamental problem in ad hoc network applications, including distributed agreement and synchronization problems [3], distributed coordination of mobile autonomous agents [2], and distributed data fusion in sensor networks [1], [4]. It is also a central topic for load balancing (with divisible tasks) in parallel computers [5]. This paper focuses on a prototypical example of agreement in asynchronous networked systems, namely, the randomized average consensus problem.

A. Average Consensus

At time slot $t \geq 0$, each node $i = 1, 2, \ldots, N$ has an estimate $x_i(t)$ of the global average, and we use $x(t)$ to denote the N-vector of these estimates. The ultimate goal is to drive the estimate $x(t)$ to the average vector $\overline{x}(0)\mathbf{1}$ (with $\mathbf{1}$ denoting the vector of ones) or as close as possible, where

$$\overline{x}(0) = \frac{1}{N}\sum_{i=1}^{N} x_i(0) \qquad (1)$$

using minimal amount of communication. The quantity $x(t)$ for $t > 0$ is a random vector, since the algorithms are randomized in their behavior.

B. Related Work

Gossip-based algorithms were initially introduced by Tsitsiklis [6] to achieve consensus over a set of agents, and have recently received renewed attention from other researchers [1], [2], [7]–[11]. The standard *pairwise* randomized gossip algorithm uses an asynchronous time model in which a node chosen uniformly at random contacts a randomly chosen neighbor within its *connectivity radius*, and exchanges values with that neighbor. The two nodes then update their own values with the pairwise average of their values. This operation preserves both the total sum, and hence also the mean, of the node values. This algorithm converges to a consensus if the graph is strongly connected on the average.

Because the transmitting node must send a packet to the chosen neighbor and then wait for the neighbor's packet, this scheme is vulnerable to packet collisions and yields a communication complexity (measured by number of radio transmissions to drive the estimation error to within $\Theta(N^{-\alpha})$, for any $\alpha > 0$) on the order of $\Theta(N^2 \log N)$ over random geometric graphs [7].

Pairwise gossip was recently extended in the geographic *pairwise* gossip algorithm to include geographic routing [12]. As in the standard pairwise gossip algorithm, a node randomly wakes up, but instead chooses a node randomly in the *whole network*, rather than in its *neighborhood* and performs a pairwise averaging with this node. Geographic pairwise gossiping increases the diversity of every pairwise averaging operation. The authors show that the communication complexity is in the order of $O(N^{3/2}\sqrt{\log(N)})$, which is an improvement with respect to the standard gossiping algorithm. More recently, a variety of the algorithm that "averages along the way" has been shown to converge in $O(N \log N)$ transmissions [13].

C. Primary Motivations

The algorithms discussed above incur extra overhead because each node needs to know its own location as well as learn and memorize the locations of its neighbors. The problem of packet loss is exacerbated by the requirement that messages must be sent on long routes, creating congestion and complex routing issues. Finally, the routing protocol requires storage and computation resources that may grow with N.

Wireless media have the advantage of being broadcast and, at the cost of only 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 simultaneously perform an update by exploiting the wireless medium, thereby avoiding the need for complex routing and pairwise exchanging operations.

Tuncer C. Aysal, Mehmet E. Yildiz and Anna Scaglione are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, 14853, {tca27,mey7,as337}@cornell.edu

Anand D. Sarwate is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA, 94720, asarwate@eecs.berkeley.edu

This work is in part supported by NSF grant CCF 0729074.

D. Summary of Main Contributions

We recently proposed *broadcast gossip algorithms* as an alternative to the one discussed above, especially, for wireless sensor networks [14]. We have shown that the broadcast gossip achieves consensus almost surely at a value within a neighborhood of the initial node measurements' average. The expectation of this value is equal to the average of the initial node measurements, as desired. We further described some characteristics of the convergence of the mean square error (MSE) through the iterations.

In this paper we provide a more complete analysis and optimization for the broadcast gossip algorithm. We derive the optimal mixing parameter (*i.e.* the coefficient used to average) based on the worst-case convergence rate. Interestingly, the optimal mixing parameter depends on the graph Laplacian, and converges to zero as graph connectivity increases. We also present the limiting MSE expression for the broadcast gossip protocol and analyze its behavior with respect to the mixing parameter. Finally, we present theoretical and numerical results on the convergence rate of the broadcast gossip protocol.

E. Paper Organization

We introduce the broadcast gossip algorithm along with its convergence characteristics in Section II. The performance analysis of the proposed algorithm where we focus on the MSE and communication complexity is detailed in Section III. Finally, we conclude with Section IV.

II. Broadcast Based Gossiping

Following previous work, we model our wireless sensor network as a random geometric graph [7], [12], [15] and we represent the N–node topology by the $N \times N$ adjacency matrix Φ and denote the *connectivity radius* by R and the neighborhood of the node i by $\mathcal{N}_i = \{j : \Phi_{ij} = 1\}$. Moreover, we use the asynchronous time model which is well–matched to the distributed nature of sensor networks [7], [12]. For a more thorough treatment of the graph and time models, see [7], [12], [14].

Suppose at time step t, node $i \in \{1, 2, \ldots, N\}$ clock ticks. Then, node i activates and the following events occur in the network:

1) Node i broadcasts its current state value, $x_i(t)$ over the wireless medium.
2) The broadcasted value is successfully received by all nodes that are within a radius R of node i. These are precisely the nodes in the neighbor set $\mathcal{N}_i$.
3) Each node k in $\mathcal{N}_i$ uses the broadcasted value $x_i(t)$ to update its own state value according to:
$$x_k(t+1) = \gamma x_k(t) + (1-\gamma)x_i(t), \forall k \in \mathcal{N}_i \ , \quad (2)$$
where $\gamma \in (0, 1)$ is the mixing parameter of the algorithm.
4) The remaining nodes in the network, including i, update their state values as
$$x_k(t+1) = x_k(t), \forall k \notin \mathcal{N}_i \ . \quad (3)$$

Formally, let $x(t)$ denote the vector of values at the end of the time-slot t. Then, the network-wide update is given by
$$x(t+1) = W(t)x(t) \quad (4)$$
where the random matrix $W(t)$, with probability $1/N$ is (assuming that the i–th clock ticks)
$$W^{(i)}_{jk} = \begin{cases} 1 & j \notin \mathcal{N}_i, k = j \\ \gamma & j \in \mathcal{N}_i, k = j \\ 1-\gamma & j \in \mathcal{N}_i, k = i \\ 0 & \text{elsewhere} \end{cases} \quad (5)$$
where $W^{(i)}$ denotes the weight matrix corresponding to the case where node i's clock ticks. In the following, mostly borrowing from our previous work [14], we present the convergence properties of the broadcast gossip algorithms.

A. Convergence Properties

Through the analysis of per-node weight matrices, *i.e.*, $W^{(i)}$'s, we have recently shown that for some $c \in \mathbb{R}$, the vector $c\mathbf{1}$ is a fixed point of the broadcasting gossip algorithm. That is, $W^{(i)}c\mathbf{1} = c\mathbf{1}$ for all i [14]. If the algorithm converges to a consensus, the preceding algorithm will not leave the consensus state. In the same work, we have also shown that $\mathbf{1}^T W^{(i)} \neq \mathbf{1}^T$ for all i, which means that the sum (and therefore the average) of the vector of node values is not preserved at each step. Thus, the broadcast gossip algorithm does not converge to the initial node measurements' average in strict sense.

1) Convergence in Expectation: Although the broadcast gossip algorithm does not converge to the initial node measurements' average, in the following, we show that it does converge to the desired value in expectation.

Proposition 1 ([14]) *The limiting random vector obtained through broadcast gossip iterations, in expectation, is*
$$\mathbb{E}\left\{\lim_{t\to\infty} x(t)\right\} = \frac{1}{N}\mathbf{1}\mathbf{1}^T x(0) \ . \quad (6)$$

2) Convergence in Second Moment: It is of interest to consider the second order convergence properties of the broadcast gossip algorithm. Since the sum is not preserved throughout broadcast gossip iterations, instead of tracking the distance to the average as in pairwise gossip, 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
$$\beta(t) = x(t) - Jx(t) = (I - J)x(t) \quad (7)$$
where $J = (1/N)\mathbf{1}\mathbf{1}^T$. This is a measure of relative deviation of the node values from their average. The following Lemma that gives necessary and sufficient conditions on $\beta(t)$ guarantee that consensus is reached.

Lemma 1 ([14]) *There is a consensus at time-slot $T \in \overline{\mathbb{N}}$ of the broadcast gossip algorithm if and only if* $\mathbb{E}\{||\beta(T)||_2^2\} = 0$.

Thus, if the expectation of the norm of the deviation vector converges to zero then the node values converge to a consensus. Let $\lambda_i(\cdot)$ and $W(t)$ denote the ith ranked eigenvalue of its argument and the random weights matrix at time step t, respectively. In the following, we present a sufficient condition guaranteeing the convergence of the expectation of the deviation vector norm to zero and show that this condition is indeed satisfied.

Proposition 2 ([14]) *The following statements hold.*

(i) $\lim_{t\to\infty} \mathbb{E}\{||\beta(t)||_2^2\} = 0$ *(*i.e., $\mathbb{E}\{||\beta(t)||_2^2\}$ *converges to zero), if*

$$\lambda_1(\mathbb{E}\{W(t)^T(I-J)W(t)\}) < 1 \ , \tag{8}$$

where I denotes the identity matrix.

(ii) *The broadcast gossip algorithm satisfies the above condition.*

It is important to emphasize that the Proposition 2 gives a sufficient condition for any consensus protocol that does not preserve network sum. Moreover, note that the condition $\lambda_1(\mathbb{E}\{W(t)^T(I-J)W(t)\}) < 1$ is different than the convergence condition obtained for the standard pairwise gossip algorithms where one only need to have $\lambda_2(\mathbb{E}\{W(t)^T W(t)\}) < 1$ to ensure the second–order convergence to the initial node measurements average [7], [12].

Broadcast gossip is different from pairwise gossip and the condition ensures the convergence of the deviation vector and not the distance to initial node measurements average vector. However, it is of interest to note the sufficiency condition derived for the broadcast gossip algorithms reduces to the one for average-preserving gossip algorithms when $\mathbf{1}^T W^{(i)} = \mathbf{1}^T, \forall i$.

3) Almost-Sure Convergence: Given, the previous results, we can compile the following theorem, which was the main result in [14].

Theorem 1 ([14]) *The broadcast gossip algorithm converges, almost surely, to a consensus:*

$$\Pr\left\{\lim_{t\to\infty} x(t) = c\mathbf{1}\right\} = 1 \ , \tag{9}$$

for some $c \in \mathbb{R}$ where

$$\mathbb{E}\{c\} = \frac{1}{N}\mathbf{1}^T x(0). \tag{10}$$

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 nodes measurements.

B. Optimal Mixing Parameter

In the following, we focus on the mixing parameter $\gamma \in (0,1)$ and find the optimal γ when approached from the worst-cast convergence rate. The worst-case convergence rate is given by Proposition 2 as $\lambda_1(\mathbb{E}\{W(t)^T(I-J)W(t)\})$. We first present the following Lemma which will prove useful in characterizing this eigenvalue of significant interest. The proofs of Lemmas, Propositions, Theorems are omitted due to space constraints, but they can be found in [16].

Lemma 2 *Let $L = \mathrm{diag}(\Phi\mathbf{1}) - \Phi$ denote the graph Laplacian. The following two formulas hold:*

(i) *Let $W' \triangleq \mathbb{E}\{W(t)^T W(t)\}$. Then*

$$W' = I - \frac{2\gamma(1-\gamma)}{N} L. \tag{11}$$

(ii) *Let $W'' \triangleq \mathbb{E}\{W(t)^T J W(t)\}$. Then*

$$W'' = \frac{(1-\gamma)^2}{N^2} L^2 + J. \tag{12}$$

Given the Lemma above, the eigenvalue of interest can now be written as:

$$\lambda_1(W' - W'') = 1 - \frac{2\gamma(1-\gamma)}{N}\lambda_{N-1}(L) - \frac{(1-\gamma)^2}{N^2}\lambda_{N-1}(L)^2. \tag{13}$$

In the following, we investigate the effect of the mixing parameter on the eigenvalue of interest in Proposition 2, thereby revealing its effect on the convergence characteristic of the broadcast gossip algorithms.

Corollary 1 *Let us introduce $\lambda_1(W' - W''; \gamma)$ to show the dependency of the eigenvalue of interest to the mixing parameter γ. Then the following statements hold.*

(i) *$\lambda_1(W' - W''; \gamma)$ is convex in γ.*

(ii) *The optimal mixing parameter is given by*

$$\gamma^* = \frac{N - \lambda_{N-1}(L)}{2N - \lambda_{N-1}(L)}. \tag{14}$$

The above Corollary, thus, indicates that the optimal mixing parameter, interestingly, depends on the graph for finite N. However, a stronger result holds for large N as discussed in the following.

Corollary 2 *For graphs such that $\lambda_{N-1}(L) = \Theta(f(N))$ for some function $f(\cdot)$, with $\lim_{N\to\infty} f(N)/N = 0$, the optimal mixing parameter is given by*

$$\lim_{N\to\infty} \gamma^* = \frac{1}{2}. \tag{15}$$

Hence, for large enough N and standard radius connectivity considerations for random geometric graphs (*e.g.*, $R = \Theta(\sqrt{\log N/N})$ and $\lambda_{N-1}(L) = \Theta(\log N)$), the eigenvalue $\lambda_1(W' - W'')$ increases as $|\gamma - 1/2|$ increases. Therefore the worst-case convergence rate, characterized by $\lambda_1(W' - W'')$, decreases. In words, γ values that are closer to $1/2$ yield a faster worst-case convergence rate compared to the γ values closer to its boundaries, *i.e.*, zero and one.

In the following, we investigate the effect of the graph Laplacian on the optimal mixing parameter.

Corollary 3 *Let us introduce $\gamma^*(L) \triangleq \gamma^*$ to denote the dependency of the optimal γ on the graph Laplacian. Then, $\gamma^*(L)$ is monotonically decreasing function of $\lambda_{N-1}(L)$.*

Thus, the above corollary indicates that, as the graph connectivity increases, *i.e.*, the eigenvalues of the Laplacian increases, the optimal mixing parameter tends to zero. This result corroborates with intuition. For instance consider a fully connected graph, then clearly $\gamma = 0$ would result in a *consensus* at the first iteration.

III. Performance Analysis of Broadcast Gossip

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

A. Mean Square Error

Because the broadcast gossip algorithm does not in general converge to the initial node measurements average $(N)^{-1}\mathbf{1}^T x(0)$, it is of interest to consider the distance of the consensus value to $\overline{x}(0)$. In the remaining, we use

$$\alpha(t) = x(t) - Jx(0) \ . \tag{16}$$

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

Lemma 3 ([14]) *Let $\mathbb{E}\{\|\alpha(t)\|_2^2\}$ denote the mean square error at time step t. The following two statements hold:*

(i) *The MSE obeys the following recursion:*

$$\begin{aligned}\mathbb{E}\{\|\alpha(t+1)\|_2^2\} \leq (1-\lambda_2(W'))\mathbb{E}\{\|J\alpha(t)\|_2^2\} \\ + \lambda_2(W')\mathbb{E}\{\|\alpha(t)\|_2^2\}.\end{aligned} \tag{17}$$

(ii) *For some $c \in \mathbb{R}$,*

$$x(t) \neq c\mathbf{1} \Leftrightarrow \mathbb{E}\{\|\alpha(t+1)\|_2^2|\alpha(t)\} < \|\alpha(t)\|_2^2 \ . \tag{18}$$

The above Lemma reveals that the mean square error (MSE) is a *strictly decreasing* function of time and strict inequality becomes equality when the nodes converge to consensus. In the following, we consider the limiting MSE behavior of the broadcast gossip algorithms.

Proposition 3 *Let $W = \mathbb{E}\{W(t)\}$. The limiting MSE of the broadcast gossip algorithms is upper bounded by*

$$\mathbb{E}\{\lim_{t\to\infty} \|\alpha(t)\|_2^2\} \leq \|\alpha(0)\|_2^2 \left(1 - \frac{1-\lambda_2(W')}{1-\lambda_{N-1}^2(W-J)}\right) \ . \tag{19}$$

As in the worst-case convergence-rate case, it is of interest to characterize the effect of the mixing parameter γ on the limiting MSE performance. This is considered in the following Corollary.

Corollary 4 *Let $U_\infty(\gamma)$ be the upper-bound on the limiting MSE of the broadcast gossip iterations, given in Proposition 3, as a function of the mixing parameter γ. Then, the following statements hold.*

(i) *The boundary cases,* i.e., *$\gamma \to 0$ and $\gamma \to 1$, are given by*

$$\lim_{\gamma\to 0} U_\infty(\gamma) = ||\alpha(0)||_2^2 \tag{20}$$

and

$$\lim_{\gamma\to 1} U_\infty(\gamma) = ||\alpha(0)||_2^2 \left(1 - \frac{\lambda_{N-2}(L)}{\lambda_1(L)}\right) \tag{21}$$

respectively.

(ii) *$U_\infty(\gamma)$ is a monotonically decreasing function of γ.*

(iii) *$U_\infty(\gamma)$, for $\gamma = \gamma^*$, is given by*

$$U_\infty(\gamma^*) = ||\alpha(0)||_2^2 \left(1 - C(L)\frac{\lambda_{N-2}(L)}{\lambda_1(L)}\right) \tag{22}$$

where

$$C(L) = \frac{2N - 2\lambda_{N-1}(L)}{4N - 2\lambda_{N-1}(L) - \lambda_1(L)}. \tag{23}$$

The Corollary indicates that the limiting MSE value of the broadcast gossip algorithm decreases with increasing γ. Thus, when approached from the minimum MSE perspective, the optimal γ value is $1-\epsilon$ for some small ϵ. This is due to the fact that as γ approaches zero, the broadcasting nodes create a local dominance possibly shifting away from the desired mean, whereas for γ values closer to unity, the nodes receiving the broadcasted value adjust their own state only slightly, and given the large number of iterations, the end result resembles more closely its expectation, which is the average of the initial states. In addition, under the standard assumption on the connectivity radius and large enough N, it is interesting to note that the factor $C(L)$ tends to 1/2.

B. Communication Cost to Achieve Consensus

Deviating, but in a similar fashion, from the standard sum preserving gossip–based averaging algorithms, we define the ϵ–converging time in the following.

Definition 1 *Given $\epsilon > 0$, the ϵ–converging time is the earliest time at which the vector $x(t)$ is ϵ close to the normalized initial deviation with probability greater than $1-\epsilon$:*

$$T(N,\epsilon) = \sup_{x(0)} \inf \left\{ t : \Pr\left\{ \frac{\|x(t) - Jx(t)\|_2}{\|x(0) - Jx(0)\|_2} \geq \epsilon \right\} \leq \epsilon \right\}. \tag{24}$$

We will need the following lemma characterizing the behavior of the eigenvalue of interest before we present the main result of this section.

Lemma 4 *For the broadcast gossip algorithm,*

$$1 - O\left(\frac{\log^4 N}{N^2}\right) \leq \lambda_1\left(W' - W''\right) < 1 - \Omega\left(\frac{\sqrt{\log N}}{N^{5/2}}\right). \tag{25}$$

Proof: We would like to calculate the eigenvalue $\lambda_1(W' - W'')$. The matrix $W' - W''$ is given by:

$$W' - W'' = I - J - \frac{2\gamma(1-\gamma)}{N} L - \frac{(1-\gamma)^2}{N^2} L^2 \ . \quad (26)$$

First note that the vector $\mathbf{1}$ is an eigenvector of $W'-W''$ with eigenvalue 0. The vector $\mathbf{1}$ corresponds to the only nonzero eigenvalue of the matrix J and the only zero eigenvalue for the Laplacian matrix L. Therefore the eigenvectors of $W' - W''$ are exactly the eigenvectors of L, and the k-th eigenvalue of $W' - W''$ for $k = 1, 2, \ldots N - 1$ is:

$$\begin{aligned}\lambda_k(W' - W'') = 1 &- \frac{2\gamma(1-\gamma)}{N}\lambda_{N-k}(L) \\ &- \frac{(1-\gamma)^2}{N^2}\lambda_{N-k}(L)^2. \quad (27)\end{aligned}$$

Thus to characterize $\lambda_1(W' - W'')$ we must characterize the second-smallest eigenvalue of the Laplacian matrix L (*the algebraic connectivity* of the graph).

An upper bound on $\lambda_{N-1}(L)$ will yield a lower bound on the largest eigenvalue of $W' - W''$. A result of Alon and Milman [17, Theorem 2.7] shows that:

$$\lambda_{N-1}(L) \leq \frac{8d_{\max}}{\operatorname{diam}(G)^2}\log_2^2 N \quad (28)$$

where $\operatorname{diam}(G)$ denotes the graph diameter. If the communication radius is chosen large enough, for the random geometric graph with standard connectivity assumptions, $d_{\max} = \Theta(\log N)$ (see [13]). The diameter can be found as the number of hops to get from one corner to the diagonally opposite corner, so it is $\Theta(\sqrt{N/\log N})$. Thus the whole bound is:

$$\lambda_{N-1}(L) = O\left(\frac{\log^4 N}{N}\right) \ . \quad (29)$$

This gives the bound

$$\lambda_1(W' - W'') = 1 - O\left(\frac{\log^4 N}{N^2}\right) \ . \quad (30)$$

To upper bound $\lambda_1(W - W'')$ we need a nontrivial lower bound on $\lambda_{N-1}(L)$. A result of Mohar states that [18]:

$$\lambda_{N-1}(L) \geq \frac{4}{N \cdot \operatorname{diam}(G)} \ . \quad (31)$$

Therefore

$$\lambda_{N-1}(L) = \Omega\left(\frac{\sqrt{\log N}}{N^{3/2}}\right) \quad (32)$$

and

$$\lambda_1(W' - W'') = 1 - \Omega\left(\frac{\sqrt{\log N}}{N^{5/2}}\right) \quad (33)$$

completing the proof. ∎

Unfortunately, the upper and lower bounds do not coincide – they differ (ignoring logarithmic terms) by a $\sqrt{N}$ factor. It may be possible to tighten the upper bound by exploiting the fact that for a communication radius slightly larger than the threshold the random geometric graph is regular in an order sense with degree $\Theta(\log N)$ [13]. However, we do not pursue this here.

Given the convergence rate definition, and the previous lemma, we obtain, utilizing steps similar to [7], the following rate of convergence to a consensus for broadcast gossip.

Proposition 4 *The consensus time for the asynchronous broadcast average consensus, for $t \geq T(N, \epsilon)$ is bounded as follows:*

$$\Pr\left\{\frac{\|x(t) - Jx(t)\|_2}{\|x(0) - Jx(0)\|_2} \geq \epsilon\right\} \leq \epsilon \quad (34)$$

where

$$\Omega\left(\frac{N^2\log\epsilon^{-1}}{\log^4 N}\right) = T(N, \epsilon) = O\left(\frac{N^{5/2}\log\epsilon^{-1}}{\sqrt{\log N}}\right). \quad (35)$$

Moreover, note that if we set $\epsilon = 1/N^\alpha$ in the above equation, then we obtain $T(N, 1/N^\alpha) = \Omega(N^2/\log^3 N)$. Since the number of transmissions per iteration is 1 in the broadcast gossip algorithm, this result corresponds to the communication complexity. Of note is that broadcast gossip algorithms improves upon randomized gossip algorithms ($\Theta(N^2 \log N)$), but appears to be worse than the geographic gossip which has communication complexity in the order of $O(N^{3/2}\sqrt{\log N})$. However, as we will see very shortly through numerical examples, broadcast gossip significantly outperforms both algorithms for practical network sizes, revealing that the asymptotic scaling results may not be relevant for finite and practical network sizes.

C. Numerical Examples

To simulate the random geometric graph, we consider nodes that are uniformly distributed over a unit square. Their initial values are initialized as uniformly distributed random values with unit variance and zero mean. The connectivity radius is chosen as $R = \sqrt{\log(N)/N}$.

1) Communication Cost: In the following, as in [12], we compare the number of *radio transmissions* to achieve a certain distance from consensus of broadcast gossiping. We choose $\gamma = 1/2$, since this is the optimal value for large enough N and it furthermore offers a good trade-off for MSE. We present plots comparing the communication cost of standard pairwise gossip algorithm [7], geographic pairwise gossip algorithm [12], and the broadcast gossip algorithm for varying network sizes.

Figure 1 depicts per-node variance versus the number of radio transmissions (each data point is an ensemble average of 25 trials). Broadcast gossip requires a single radio transmission per iteration, whereas standard gossip requires two and geographic gossip requires the number of hops between the nodes. Simulation results suggest that broadcast gossiping significantly outperforms both protocols from the communication cost perspective for the given network sizes. Furthermore, broadcast gossip avoids some complexities in the geographic gossiping protocol, such as costs due to memory and routing operations (which are not incorporated into simulations).

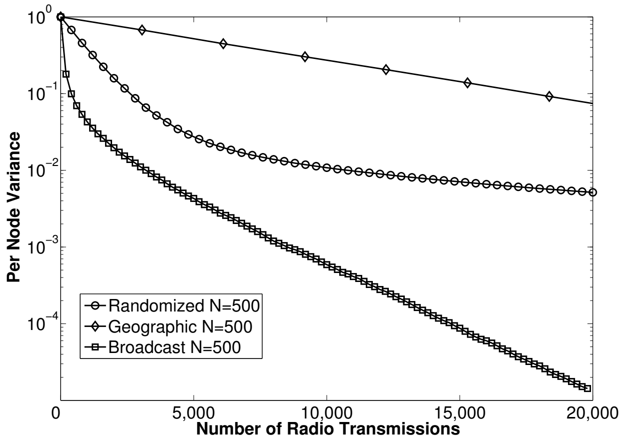

Fig. 1. Number of radio transmissions required to achieve a given distance (per node variance) from the consensus for $N = 500$.

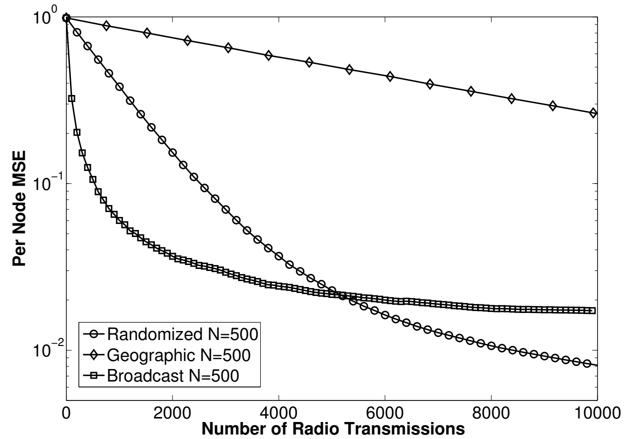

Fig. 2. Number of radio transmissions required to achieve a given MSE for $N = 500$.

2) Mean Square Error: Next, we consider the MSE performance of the broadcast gossip algorithm through iterations and compare the performance of our algorithm to those of randomized and geographic gossip algorithms. Recall that the MSE of randomized and geographic gossip algorithms achieve zero in the limit, whereas the MSE of the broadcast algorithm saturates to a non-zero value as the algorithm converges to a consensus.

Figures 2 depicts the MSE performance of the randomized, geographic and broadcast gossip algorithms through number of radio transmissions for $N = 500$, respectively. An interesting observation is that, for reasonable number of radio transmissions, the MSE performance of the broadcast gossip is better than those of the randomized and geographic. However, as the number of radio transmissions increases, the randomized and geographic gossip outperform the broadcasting one, as they tend to zero whereas the performance of the broadcast gossip saturates to a non-zero value.

These simulations results are in corroboration with the theoretical ones stating that the MSE strictly decreases as long as consensus is not achieved and the previous simulation results showing that the broadcast gossip achieves consensus significantly faster than randomized, and geographic pairwise gossip algorithms.

IV. Concluding Remarks

In this paper, we provided further analysis and optimization of the recently proposed broadcast gossip algorithm. By requiring only that nodes "lend their ears" to a broadcaster's transmission, these algorithms avoid some potential difficulties that could arise in the implementation of other gossip protocols. Specifically, we have derived the optimal mixing parameter when approached from the worst-case convergence rate perspective and we studied its effect of the limiting MSE performance. The results indicate that good MSE performance is achievable trading off the speed of convergence. Finally, we presented theoretical and numerical examples evaluating and comparing the communication cost of gossiping algorithms which demonstrated the fast convergence of our algorithm.

References

[1] C. C. Moallemi and B. V. Roy, "Consensus propagation," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4753–4766, Nov. 2006.

[2] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[3] W. Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, 2005.

[4] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proceedings of the IEEE/ACM Int. Symp. on Inf. Proc. in Sens. Netw.*, Los Angeles, CA, Apr. 2005.

[5] N. Lynch, *Distributed Algorithms*. Morgan Kaufmann Publishers, Inc., San Francisco, CA, 1996.

[6] J. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Dept. of Electrical Engineering and Computer Science, M.I.T., Boston, MA, 1984.

[7] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Info. Theory*, vol. 52, no. 6, pp. 2508–2530, June 2006.

[8] D. Kempe, A. Dobra, and J. Gehrke, "Computing aggregate information using gossip," in *Proc. Foundations of Computer Science*, Cambridge, MA, October 2003.

[9] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On distributed averaging algorithms and quantization effects," LIDS Report 2274, MIT, Tech. Rep., Nov. 2007.

[10] A. Nedić and A. Ozdaglar, "Convergence rate for consensus with delays," LIDS Report 2774, MIT, Tech. Rep., Oct. 2007.

[11] T. C. Aysal, M. J. Coates, and M. G. Rabbat, "Rates of convergence of distributed average consensus with probabilistic qantization," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep. 2007.

[12] A. G. Dimakis, A. D. Sarwate, and M. J. Wainwright, "Geographic gossip: Efficient averaging for sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 3, Mar. 2008.

[13] F. Bénézit, A. G. Dimakis, P. Thiran, and M. Vetterli, "Order-optimal consensus through randomized path averaging," February 2008, arXiv:0802.2587v1.

[14] T. C. Aysal, M. E. Yildiz, and A. Scaglione, "Broadcast gossip algorithms," in *Proceedings of the 2008 IEEE Information Theory Workshop*, Porto, Portugal, May 2008.

[15] A. Giridhar and P. R. Kumar, "Towards a theory of in-network computation in wireless sensor networks," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 98–107, Apr. 2006.

[16] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, "Broadcast gossip algorithms," *IEEE Trans. Signal Process.*, Mar. 2008, submitted for publication, www.crips.cornell.edu.

[17] N. Alon and V. Milman, "λ_1, isoperimetric inequalities for graphs, and superconcentrators," *Journal of Combinatorial Theory, Series B*, vol. 38, pp. 73–88, 1985.

[18] B. Mohar, *Graph Theory, Combinatorics, and Applications*. Wiley, 1991, vol. 2, ch. The Laplacian spectrum of graphs, pp. 871–898.