ON SPATIAL GOSSIP ALGORITHMS FOR AVERAGE CONSENSUS

Michael G. Rabbat

Dept. of Electrical and Computer Engineering
McGill University, Montréal, Québec
Email: michael.rabbat@mcgill.ca

ABSTRACT

This paper investigates the use of spatial gossip to compute the average consensus in networks such as grids or random geometric graphs, where connectivity is a function of proximity. Randomized gossip is a framework for distributed computation where, at each iteration, a random pair of nodes exchanges information, and then updates their local values by averaging. This simple protocol converges to an average consensus: every node obtains the average of the initial values across the network. In spatial gossip, if the distance between two nodes is d, then they communicate with probability proportional to $d^{-\beta}$ for some $\beta \geq 0$. The special case $\beta = 0$ corresponds to an algorithm known in the sensor network literature as geographic gossip. Dimakis et al. have shown that geographic gossip computes the average to accuracy n^{-1} in $O(n^{3/2}\sqrt{\log n})$ transmissions. In this paper we show that the same rates are achieved for $\beta = 2$ and $\beta = 3$. Each setting offers a different balance between the rate of convergence (in gossip rounds) and the average number of transmissions per gossip round. We illustrate, via simulation, that spatial gossip with $\beta = 2$ generally yields superior performance over geographic gossip by a constant factor.

***Index Terms*—** Sensor networks, aggregation, gossip algorithms, average consensus, distributed signal processing

1. INTRODUCTION

Efficiently computing the average consensus is a fundamental challenge at the heart of many distributed signal processing and sensor network applications. In a network of n nodes where each node has a scalar value y_i, average consensus is achieved when all nodes know the average, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. Our previous work illustrates how an efficient algorithm for average consensus can be used to solve a variety of conventional sensor network and distributed signal processing applications, including source localization, network health monitoring, and field estimation [1, 2].

Randomized gossip is a simple iterative algorithm for computing the average consensus. Consider a network of n nodes. Each node, $i = 1, \ldots, n$, initializes its gossip value to $x_i(0) = y_i$. In the kth gossip round, a node s is activated uniformly at random (e.g., via i.i.d. Poisson clocks ticking at each node). Then s draws another node, call it t, where the probability that s draws t is given by $P_{s,t}$. Nodes s and t exchange values and then perform the update,

$$x_s(k) \;=\; x_t(k) \;=\; \frac{1}{2}\big(x_s(k-1) + x_t(k-1)\big),$$

and all other nodes remain unchanged (i.e., $x_u(k) = x_u(k-1)$ for all $u \neq s, t$). Under mild conditions on the probabilities $\{P_{i,j}\}$ which ensure that information eventually diffuses over the entire network, the sequence of gossip values, $x_i(k)$, converges to the average $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i(0)$ of the initial values at every node, $i = 1, \ldots, n$ [3, 4].

This paper investigates spatial gossip, where the probability that nodes i and j exchange information decays as a power-law in terms of the distance, $d_{i,j}$, between them: $P_{i,j} \propto d_{i,j}^{-\beta}$ for some $\beta \geq 0$. Spatial gossip was introduced by Kempe et al. in [5], where it was studied in the context of resource location and not aggregation or average consensus. When $\beta = 0$, spatial gossip generalizes an algorithm called geographic gossip which computes the average at the best known rate of any randomized gossip algorithm in sensor networks: $O(n^{3/2}\sqrt{\log n})$ transmissions to attain accuracy n^{-1}. For the special case where the network is a two-dimensional grid, we show that the same rates (up to logarithmic factors) are achieved when $\beta = 2$ or $\beta = 3$. In particular, when $\beta \leq 2$, spatial gossip converges in $n \log n$ rounds, but each round uses an average of $O(\sqrt{n})$ transmissions so the overall communication complexity, measured in transmissions, is $O(n^{3/2}\log n)$ transmissions for accuracy n^{-1}. On the other hand, when $\beta = 3$, the expected number of transmissions per round decreases to $O(\log n)$, but the number of rounds required increases to $O(n^{3/2}\log n)$, resulting in a communication complexity of $O(n^{3/2}\log^2 n)$. Finally, we illustrate via simulation that spatial gossip with $\beta = 2$ offers a constant factor improvement in communication complexity over geographic gossip.

2. BACKGROUND

The popularity of randomized gossip in the sensor network and distributed signal processing literature stems from a number of factors: 1) nodes act autonomously, 2) there is no overhead required to establish or maintain complicated routes, 3) randomization makes these algorithms robust to unreliable networking conditions, and 4) there is no single point of failure an attacker can compromise to prevent the network from reaching consensus. For more, see [4, 6–8][1].

In nearest neighbor gossip, $P_{i,j} > 0$ only if nodes i and j are neighbors in the underlying communication network. This particular case seems attractive for wireless sensor networks because all information exchange is local, so nodes can act completely autonomously. However, Boyd et al. [4] show that nearest neighbor gossip converges slowly in networks such as grids or random geometric graphs, where connectivity is based on locality. In particular, $O(n^2)$ rounds of nearest neighbor gossip, and thus $O(n^2)$ transmissions, are needed to compute the average with accuracy n^{-1}.

Dimakis et al. [9] propose geographic gossip as a more efficient average consensus protocol. Rather than constraining information exchange to be purely local, geographic gossip makes the additional assumption that each node knows its geographic location relative to its neighbors and allows nodes to communicate with any other node in the network. In fact, in geographic gossip, $P_{i,j} = \frac{1}{n}$, so each node communicates equally often with every other node in the network. Assuming nodes know their geographic coordinates and the coordinates of their neighbors, greedy geographic routing provides a means of communicating over long distances without any overhead required in producing and maintaining routes. Dimakis et al. [9] prove that the number of geographic gossip rounds required to achieve averaging accuracy n^{-1} is $O(n \log n)$, a sharp improvement over the $O(n^2)$ required for nearest neighbor gossip. In a network of n nodes arranged in a square grid, the average number of transmissions per gossip round is proportional to $\sqrt{n}$ (similarly, $\sqrt{n/\log n}$ for a random geometric graph topology). Thus, the communication complexity of geographic gossip on a grid scales like $O(n^{3/2} \log n)$ transmissions.

Observe that geographic gossip, defined by $P_{i,j} = \frac{1}{n}$, is equivalent to spatial gossip with $\beta = 0$; i.e., the probability two nodes communicate is a constant, independent of node distance. The work presented in this paper was motivated by the question of whether one could achieve a better balance between the number of gossip rounds and number of transmissions per round by tempering the frequency of long-range communications.

3. ANALYSIS OF SPATIAL GOSSIP

In order to determine the communication complexity of spatial gossip, we need to characterize two quantities: 1) the number of gossip rounds required to achieve a desired level of accuracy, and 2) the number of transmissions per gossip round. To simplify our discussion, the analysis presented here is for a network of n nodes arranged in a square grid. Similar results (up to poly-logarithmic factors) can be derived for random geometric graphs. We adopt the Manhattan distance to measure proximity of nodes on the grid: for nodes i and j at coordinates (i_1, i_2) and (j_1, j_2), we have $d_{i,j} = |i_1 - j_i| + |i_2 - j_2|$.

Let $\mathbf{x}(k) = [x_1(k), \ldots, x_n(k)]^T$ denote the gossip values for all nodes after k gossip rounds, stacked into a vector. The number of gossip rounds required to achieve a desired level of accuracy is characterized by the ϵ-averaging time: for $\epsilon > 0$ and for a randomized gossip algorithm defined by transition matrix P, the ϵ-averaging time is given by

$$T_{\text{ave}}(\epsilon, P) \quad = \quad \sup_{\mathbf{x}(0)} \inf \left\{ k : \Pr\left(\frac{\|\mathbf{x}(k) - \bar{x}\mathbf{1}\|}{\|\mathbf{x}(0)\|} \geq \epsilon \right) \leq \epsilon \right\},$$

where $\|\cdot\|$ denotes Euclidean distance and $\mathbf{1}$ is the all-ones vector. We will refer to $\|\mathbf{x}(k) - \bar{x}\mathbf{1}\|/\|\mathbf{x}(0)\|$ as the relative error after k rounds.

Our first result provides general bounds for any gossip algorithm where every pair of nodes communicates with nonzero probability.

Proposition 1. *Suppose $P_{i,j} > 0$ for all $i, j = 1, \ldots, n$, and let $P_{\min} = \min\{P_{i,j} : i, j = 1, \ldots, n\}$. Then*

$$T_{ave}(\epsilon, P) \quad = \quad O\left(n \log n + \frac{\log n + \log \epsilon^{-1}}{P_{\min}} \right).$$

Proof. The proof of this result is based on characterizing the rate of convergence of a Markov chain whose transition matrix is related to the gossip matrix P. Since the transition matrices we deal with are symmetric, they all have uniform stationary distribution; i.e., $\lim_{t\to\infty} P^t \nu = (1/n)\mathbf{1}$, for any vector ν satisfying $\nu_i \geq 0$, $\sum_{i=1}^n \nu_i = 1$. The rate of convergence of a Markov chain with symmetric transition matrix P is captured by a quantity called the ϵ-mixing time, defined as

$$T_{\text{mix}}(\epsilon, P) = \sup_i \inf \left\{ t : \frac{1}{2} \sum_{j=1}^n \left| P_{i,j}^{t'} - \frac{1}{n} \right| \leq \epsilon,\ \forall t' \geq t \right\},$$

which measures the amount of time needed to guarantee each element of $P^t \nu - (1/n)\mathbf{1}$ has magnitude no larger than ϵ, for any ν. Note the resemblance to the definition of the ϵ-averaging time of a gossip algorithm. Boyd et al. explicitly relate between averaging and mixing times in Theorem 7 of [4], which gives[2]

$$T_{\text{ave}}(\epsilon, P) \quad = \quad \Theta\big(n \log n + n T_{\text{mix}}(\epsilon, \widetilde{P})\big),$$

[1] Due to the vast amount of work in this area, a comprehensive list of references is beyond the capacity of this space-limited abstract. The author apologizes in advance to any researchers whose work has been omitted from the bibliography.

[2] Theorem 7 in [4] is stated in terms of what they refer to as "absolute time" and must be scaled by a factor of n to be related back to the number of gossip rounds. In one unit of "absolute time", each node is activated once, on average.

where $\widetilde{P} = \frac{1}{2}(I - P)$. Thus, to bound the averaging time for spatial gossip, we need to characterize the mixing time of a Markov chain with transition matrix $\widetilde{P}$.

It is well-known that the mixing time of a Markov chain can be bounded in terms of its eigenvalues. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues of $\widetilde{P}$ arranged in decreasing order. Because $\widetilde{P}$ is a stochastic matrix, it has largest eigenvalue $\lambda_1 = 1$. Moreover, defining $\widetilde{P} = \frac{1}{2}(I - P)$ implies all remaining eigenvalues of $\widetilde{P}$ are non-negative since P also has eigenvalues in $(-1, 1]$. It follows that the mixing time of a Markov chain with transition matrix $\widetilde{P}$ is bounded in terms of its second largest eigenvalue, λ_2 (see, e.g., [10, 11]):

$$T_{\text{mix}}(\epsilon, \widetilde{P}) \quad \leq \quad \frac{\log n + \log \epsilon^{-1}}{1 - \lambda_2}. \tag{1}$$

Next, we direct our attention to bounding λ_2. We accomplish this task using the technique of canonical paths due to Diaconis and Stroock [10], and Sinclair [11]. A set of canonical paths in a Markov chain is a collection of paths, $\Gamma = \{\gamma_{i,j}\}$, containing one path $\gamma_{i,j}$ for each pair of nodes, and such that path $\gamma_{i,j}$ originates at node i, finishes at node j, and only makes use of transitions (x, y) if $P_{x,y} > 0$. For a given set of canonical paths, Theorem 5 of Sinclair [11] states that

$$\lambda_2 \quad \leq \quad 1 - \frac{1}{\bar{\rho}},$$

where

$$\bar{\rho} \quad = \quad \max_{(i,j)} \frac{1}{n P_{i,j}} \sum_{\gamma \in \Gamma : (i,j) \in \gamma} |\gamma|.$$

The sum in the expression above is over all paths involving a transition from i to j, and the term $|\gamma|$ denotes the number of transitions in the path γ. The quantity $\bar{\rho}$ is sometimes referred to as the *path congestion* and can be thought of as measuring the amount of loading in a network whose edges have capacity $P_{i,j}$ and each path carries one unit of traffic.

Since $\widetilde{P}_{i,j} > 0$ for all i, j in our setup, we may set $\gamma_{i,j} = (i, j)$, a direct transition from i to j. Then, every path in our canonical set only involves one transition (i.e., $|\gamma_{i,j}| = 1$), and there is only one path per transition. With this setup, the expression for $\bar{\rho}$ simplifies to

$$\begin{aligned} \bar{\rho} \quad &= \quad \max_{(i,j)} \frac{1}{n \widetilde{P}_{i,j}} \\ &= \quad (n P_{\min}/2)^{-1}. \end{aligned}$$

It follows that $1 - \lambda_2 \geq n P_{\min}/2$. Using this bound in (1) and plugging the result back into (1) yields the desired result. $\square$

3.1. Spatial Gossip with $\beta = 2$

Now, let us determine the number of gossip rounds required for spatial gossip with $\beta = 2$. In order to apply Proposition 1, we need to determine $P_{\min}$. When $\beta = 2$, we have $P_{i,j} \propto d_{i,j}^{-2}$. Note that, with Manhattan distance as our metric, each node has no more than $4d$ neighbors at distance d, and for a network of n nodes arranged in a square grid, the distance between any two nodes is bounded by $2\sqrt{n}$. It follows that

$$\begin{aligned} \sum_{j \neq i} d_{i,j}^{-2} \quad &\leq \quad \sum_{d=1}^{2\sqrt{n}} (4d)\ d^{-2} \\ &= \quad 4 \sum_{d=1}^{2\sqrt{n}} d^{-1} \\ &\leq \quad 4(1 + \log(2\sqrt{n})). \end{aligned}$$

Let $Z = 4 + 4\log(2\sqrt{n})$. Set $P_{i,j} = Z^{-1} d_{i,j}^{-2}$ if $i \neq j$, and set $P_{i,i} = 1 - \sum_{j \neq i} P_{i,j}$ so that $\{P_{i,j} : j = 1, \ldots, n\}$ form a distribution. Since $P_{i,j}$ decays monotonically as a function of $d_{i,j}$ and $d_{i,j} \leq 2\sqrt{n}$, it follows that $P_{\min} \succeq (n \log n)^{-1}$. Applying Proposition 1 leads to the following upper bound for the ϵ-averaging time:

$$T_{\text{ave}}(\epsilon, P) = O(n \log^2 n + n \log n \log \epsilon^{-1}).$$

For $\epsilon = n^{-1}$, this reduces to $O(n \log^2 n)$ rounds of gossip.

Next, to bound the expected number of hops per round, we calculate

$$\begin{aligned} \sum_{d=1}^{2\sqrt{n}} d\ \Pr(\text{transmit } d \text{ hops}) \quad &\leq \quad \sum_{d=1}^{2\sqrt{n}} d(4d) Z^{-1} d^{-2} \\ &= \quad 8 Z^{-1} \sqrt{n}. \end{aligned}$$

Thus, the expected number of hops is bounded above by a term proportional to $\sqrt{n}/\log n$. Putting this together with our bound on the ϵ-averaging time gives the following result.

Proposition 2. *For $\epsilon > 0$, the relative error of spatial gossip with $\beta = 2$ is less than ϵ with probability at least $1 - \epsilon$ after $O\big(n^{3/2}(\log n + \log \epsilon^{-1})\big)$ transmissions.*

Thus, spatial gossip with $\beta = 2$ has the same communication complexity (up to logarithmic factors) as geographic gossip. In fact, this is true of the entire range $\beta \leq 2$, where gossip mixes sufficiently fast enough that $T_{\text{ave}}(\epsilon, P)$ roughly grows like $n \log n$ gossip rounds, and the number of transmissions per round is bounded by $\sqrt{n}$.

3.2. Spatial gossip with $\beta = 3$

Again, we begin by characterizing $P_{\min}$ in order to determine a bound on the number of gossip rounds via Proposition 1. The probability that two nodes communicate is now $P_{i,j} = Z^{-1} d_{i,j}^{-3}$. As before, we first need to upper bound the normalization constant Z. Mimicking the calculation above, we have

$$\sum_{j\neq i} d_{i,j}^{-3} \leq \sum_{d=1}^{2\sqrt{n}} (4d)d^{-3} = 4\sum_{d=1}^{2\sqrt{n}} d^{-2}.$$

The series $\sum_{d=1}^{\infty} d^{-2}$ converges[3], so we may simply take Z to be a constant independent of n. Then, $P_{i,j} \propto d_{i,j}^{-3}$ and $P_{\min} \succeq (2\sqrt{n})^{-3}$, so according to Proposition 1, the ϵ-averaging time is bounded by

$$T_{\text{ave}}(\epsilon, P) = O\big(n^{3/2}(\log n + \log \epsilon^{-1})\big).$$

As expected, when $P_{i,j}$ decays too quickly as a function of distance, information does not diffuse as rapidly through the network.

On the other hand, when $P_{i,j}$ decays quickly as a function of distance, the expected number of transmissions per gossip round decreases:

$$\sum_{d=1}^{2\sqrt{n}} d \ \Pr(\text{transmit } d \text{ hops}) \leq \sum_{d=1}^{2\sqrt{n}} d(4d)d^{-3} = O(\log n).$$

This leads to a comparable communication complexity result for $\beta = 3$.

Proposition 3. *For $\epsilon > 0$, the relative error of spatial gossip with $\beta = 3$ is less then ϵ with probability at least $1-\epsilon$ after $O\big(n^{3/2}\log n(\log n + \log \epsilon^{-1})\big)$ transmissions.*

For $\beta > 3$, spatial gossip has poor communication complexity for reasons similar to nearest neighbor gossip. The probability that distant nodes communicate decays too rapidly. Consequently, information diffuses slowly through the network and many gossip rounds must be executed.

4. SIMULATION RESULTS

Figure 1 plots the relative error as a function of the number of transmissions. We simulate a network of 100 nodes arranged in a 10-by-10 grid. The simulations are initialized so that $x_i(0) = 0$ everywhere except one node, where $x_i(0) = 1$. This is the worst-case initialization for gossip, since the one unit of energy needs to diffuse over the entire network. The three curves shown are for nearest neighbor gossip, geographic gossip, and spatial gossip ($\beta = 2$). Each curve shown represents the average over 1000 trials.

Figure 2 shows the average number of transmissions needed to attain a relative accuracy of 10^{-3}, for networks of varying size. Each point in these curves represents the average over 100 trials. It is evident from both figures that geographic gossip and spatial gossip converge at the same rate, superior to that of nearest neighbor gossip, and that spatial gossip (with $\beta = 2$) improves upon geographic gossip by a constant factor.

[3]In fact, $\sum_{d=1}^{\infty} d^{-2} = \zeta(2)$, where $\zeta(x)$ denotes the Riemann zeta function. $\zeta(2) \approx 1.6449$.

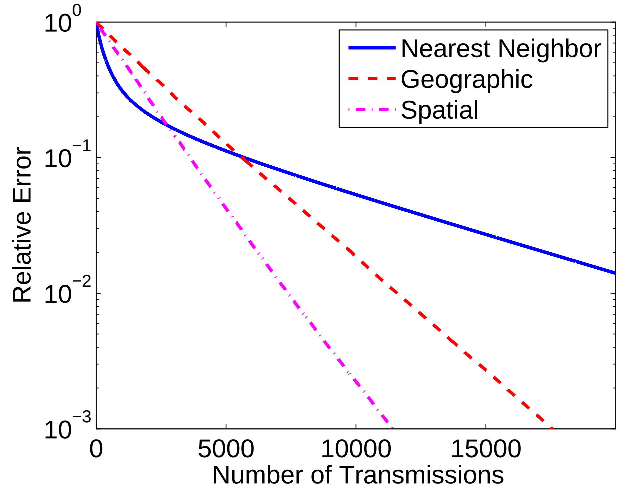

Fig. 1. Relative error as a function of the number of transmissions for a grid network of 100 nodes. Spatial gossip (with $\beta = 2$) converges to the average at a faster rate than geographic gossip. Each curve represents the average over 1000 trials.

5. DISCUSSION

In general, randomized gossip algorithms are slow to converge on networks such as random geometric graphs or grids, where connectivity is based on proximity (nodes communicate only if they are close enough). The reason for slow convergence on such graphs is that it takes many iterations for information to propagate across the network. In particular, if information is only exchanged between neighboring nodes then, in the worst case, $O(n^2)$ transmissions must be made to compute the average to within a factor of $1/n$ in a network of n nodes.

Geographic gossip, introduced in [9], is a variant of randomized gossip where nodes use knowledge of their location to facilitate information exchange over longer distances, and each node exchanges information equally often with every other node. Allowing for long-range information exchange significantly reduces the number of gossip iterations required to average (from $O(n^2)$ to $O(n \log n)$). However, long-range exchanges require more communication ($O(\sqrt{n})$, on average), so the resulting communication complexity of geographic gossip is $O(n^{3/2} \log n)$ on a grid topology.

In this paper we study spatial gossip, a generalization of geographic gossip where information is exchanged between

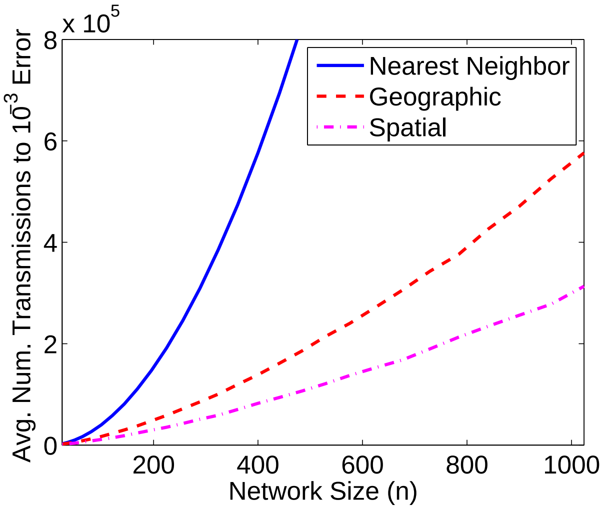

Fig. 2. The number of transmissions to reach a relative error of 10^{-3}, as a function of network size. Each point is the average over 100 trials. From this figure it is evident that the nearest neighbor curve is proportional to n^2, while geographic and spatial gossip are both proportional to $n^{3/2}$ and spatial gossip with $\beta = 2$ has a smaller constant.

any two nodes i and j at a rate, $(\text{dist}(i,j))^{-\beta}$, inversely proportional to the geographic distance between them. When $\beta = 0$, spatial gossip is equivalent to geographic gossip. The goal of this work was to study the tradeoff between the number of gossip rounds and the expected number of transmissions per gossip round as a function of β. We found that spatial gossip with $\beta = 2$ achieves the same rate of convergence as geographic gossip (up to a logarithmic factor). Intuitively, when $0 \leq \beta \leq 2$, long-range information exchange occurs frequently enough that relatively few gossip rounds need to be executed (roughly speaking, $O(n \log n)$), but the average number of transmissions per gossip round is high (roughly $O(\sqrt{n})$), so the overall communication complexity balances out at $O(n^{3/2})$ multiplied by a factor which is polylogarithmic in n.

We also identified spatial gossip with $\beta = 3$ as providing the same performance, but with a different tradeoff: many shorter-range transmissions as opposed to few long-range transmissions. That is, when $\beta = 3$, long-range information exchange is much less frequent, so more gossip rounds are required (up from $n \log n$ to $n^{3/2} \log n$), but each round only costs $\log n$ transmissions, on average. For $\beta > 3$, the communication complexity of spatial gossip grows at a rate larger than $n^{3/2} \log n$, since the number of gossip rounds tends to n^2 as β tends to ∞, returning to nearest neighbor gossip in the limit.

The range of spatial gossip algorithms for $\beta \in (2, 3)$ remains of interest. In this range the number of gossip rounds transitions up from $O(n)$ to $O(n^{3/2})$, ignoring logarithmic factors. At the same time, the average number of transmissions per round drops from $O(\sqrt{n})$ to $O(\log n)$. It remains unclear as to whether there is an optimal setting for β in this range which yields superior communication complexity. Empirical results indicate that the performance in this range is no better than at $\beta = 2$ or 3.

We note that, although the analysis in this paper has focused on a grid topology, similar results can be obtained for random geometric graphs (up to poly-logarithmic factors) using the binning strategy employed in [9].

6. REFERENCES

[1] M. Rabbat, R. Nowak, and J. Bucklew, "Robust decentralized source localization via averaging," in *Proc. IEEE ICASSP*, Philadelphia, PA, March 2005.

[2] M. Rabbat, J. Haupt, A. Singh, and R. Nowak, "Decentralized compression and predistribution via randomized gossiping," in *Proc. Information Processing in Sensor Networks (IPSN)*, Nashville, TN, April 2006.

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

[4] 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.

[5] D. Kempe, J. Kleinberg, and A. Demers, "Spatial gossip and resource location protocols," *J. of the Association for Computing Machinery*, vol. 51, no. 6, pp. 943–967, Nov. 2004.

[6] V. Saligrama, M. Alanyali, and O. Savas, "Distributed detection in sensor networks with packet loss and finite capacity links," *IEEE Trans. Signal Processing*, vol. 54, no. 11, pp. 4118–4132, November 2006.

[7] E. Yildiz and A. Scaglione, "Differential nested lattice encoding for consensus problems," in *Proc. Information Processing in Sensor Networks (IPSN)*, Cambridge, MA, April 2007.

[8] S.A. Aldosari and J.M.F. Moura, "Topology of sensor networks in distributed detection," in *Proc. IEEE ICASSP*, Toulouse, France, May 2006.

[9] A. Dimakis, A. Sarwate, and M. Wainwright, "Geographic gossip: Efficient aggregation for sensor networks," in *Proc. Information Processing in Sensor Networks (IPSN)*, Nashville, TN, April 2006.

[10] P. Diaconis and D. Stroock, "Geometric bounds for eigenvalues of markov chains," *Annals of Applied Probability*, vol. 1, no. 1, pp. 36–61, 1991.

[11] A. Sinclair, "Improved bounds for mixing rates of markov chains and multicommodity flow," *Combinatorics, Probability, and Computing*, vol. 1, pp. 351–370, 1992.