The asymptotical error of broadcast gossip averaging algorithms

Fabio Fagnani ∗ Paolo Frasca ∗

∗ Dipartimento di Matematica, Politecnico di Torino, Torino, Italy

Abstract: In problems of estimation and control which involve a network, efficient distributed computation of averages is a key issue. This paper presents theoretical and simulation results about the accumulation of errors during the computation of averages by means of iterative “broadcast gossip” algorithms. Using martingale theory, we prove that the expectation of the accumulated error can be bounded from above by a quantity which only depends on the mixing parameter of the algorithm and on few properties of the network: its size, its maximum degree and its algebraic connectivity. Both analytical results and computer simulations show that in several network topologies of applicative interest the accumulated error goes to zero as the size of the network grows larger.

1. INTRODUCTION

Distributed computation of averages is an important building block to solve problems of estimation and control over networks. As a reliable time-independent communication topology may be unlikely in the applications, a growing interest has been devoted to randomized “gossip” averaging algorithms. In such algorithms, at each time step a random subset of the nodes communicates and performs computations. Unfortunately, some of these iterative algorithms do not deterministically ensure that the average is preserved through iterations, and because of the accumulation of errors, in general there is no guarantee that the typical algorithm realization will converge to a value which is close to the desired average.

Motivated by wireless network applications, we focus on one of these randomized algorithms, the Broadcast Gossip Algorithm. In this algorithm, a node is randomly selected at each time step to broadcast its current value to its neighbors, which in turn update their values by a local averaging rule. Since these updates are not symmetric, it is clear that the average is not preserved, but instead is changed at each time step by some amount. Then, we want to study how these errors accumulate, and how much the convergence value of the algorithm differs from the original average to be computed.

Contribution. In the present paper we study the asymptotical error, or bias, committed by a Broadcasting Gossip Averaging algorithm, and we show that large neighborhoods in the communication graph and a large mixing parameter induce a large asymptotical error. By interpreting the average of states as a martingale, we prove that the expectation of the bias can not be larger than a constant, depending on the initial condition, times $q \frac{d^2_{\text{max}}}{1-qN\lambda_1}$, where $q$ is the “mixing parameter” of the algorithm, $N$ is the network size, $d_{\text{max}}$ is the maximum degree of the nodes, and $\lambda_1$ is the network algebraic connectivity. For some families of graphs (e.g, expander graphs), this is enough to prove that the bias goes to zero as $N$ goes to infinity. Moreover, by means of simulations we show that on several example graph topologies the mean bias is an increasing function of the mixing parameter and is proportional to the ratio between degree and size of the network. This implies that whenever $d_{\text{max}} = o(N)$, the bias goes to zero as $N$ goes to infinity.

Related works. The paper Fagnani and Zampieri [2008a] provides a general theory of randomized linear averaging algorithms and presents a few example algorithms, some of which do not preserve the average of the states. The variance of the asymptotic value, in such cases, can in principle be expressed using the formula in Tahbaz-Salehi and Jadbabaie [2010], which however does not give results of immediate interpretation. Among these non-average-preserving algorithms, the Broadcast Gossip Algorithm has been attracting a wide interest for its natural application to wireless networks: main references include the paper Aysal et al. [2009b] and the recent survey Dimakis et al. [2010]. While it is simple to give conditions to ensure that the expectation of the convergence value is equal to the initial average, the problem of estimating the difference between expectation and realizations is harder, and has received partial answers in a few papers. In Aysal et al. [2009a] the authors study a related communication model, in which the broadcasted values may not be received with a probability which depends on the transmitter and receiver nodes, and claim that “aggressive updating combined with large neighborhoods [...] result in more variance of the convergence value” within the short time to convergence. This intuition extends to the Broadcast Gossip Algorithm which we are considering in this paper. In Aysal et al. [2009b] the variance of the limit value has been estimated for general graphs, with an upper bound which is pro-
portional to \((1 - \lambda_1/N - 1/\sqrt{N})\), where \(\lambda_i\) is the \(i\)-th smallest non-zero eigenvalue of the graph Laplacian. This bound, however, is not sufficient to prove that the bias goes to zero as \(N\) grows. To date, this fact has only been proved –in Fagnani and Frasca [2011]– for sequences of Abelian Cayley graphs with bounded degree, using tools from algebra and Markov chain theory. Analogous problems can be studied for other randomized algorithms which do not preserve the average. For instance, in Fagnani and Zampieri [2008b] two related algorithms are studied, in which node values are sent to one random neighbor only. In the first of these algorithms, at each time step one random node sends its value to one neighbor. Then the bias has an upper bound which is proportional to \(\frac{q}{N}\). In the other algorithm, at each time step every node sends its value to one neighbor. Then the bias has an upper bound which is proportional to \(\frac{q^2}{\lambda_1 N}\).

2. PROBLEM STATEMENT

Let there be a (undirected) graph \(G = (V, E)\) with \(E \subset \{(u, v) : u, v \in V, u \neq v\}\), together with \(N = |V|\) real numbers \(\{y_v\}_{v \in V} \subset [0, L]\). For every node \(v \in V\), we denote its neighborhood by \(N_v = \{u \in V : \{u, v\} \in E\}\). The following Broadcasting Gossip Algorithm (BGA) is run in order to estimate the average \(y_{av} = N^{-1}\sum_{v \in V} y_v\).

At time \(t = 0\), we set \(x_v(0) = y_v\) for all \(v \in V\). Then, at each time step \(t \in \mathbb{Z}_{\geq 0}\), one node \(v\) is sampled from a uniform distribution over \(V\). The sampled node \(v\) broadcasts its state \(x_v(t)\) to its neighbors \(u \in N_v\), which in turn update their states as

\[
x_u(t+1) = (1 - q)x_u(t) + qx_v(t).
\] (1)

The parameter \(q \in (0, 1)\) is said to be the mixing parameter of the algorithm. If instead \(u \notin N_v\), there is no update:

\[x_u(t+1) = x_u(t).\]

It is known from [Fagnani and Zampieri, 2008a, Corollary 3.2] that the BGA converges in the sense that there exists a random variable \(x^*\) such that almost surely \(\lim_{t \to \infty} x(t) = x^*1\). Let now \(x_{av}(t) = N^{-1}\sum_{v \in V} x_v(t)\). Although \(E[x^*] = x_{av}(0)\), in general \(x^*\) is not equal to \(x_{av}(0)\). Then, it is worth to ask how far the convergence value from the initial average is. To study this bias in the computation of the average, we define

\[\beta(t) = |x_{av}(t) - x_{av}(0)|^2.\]

The goal of this work is to study this quantity, and in particular the limit \(E[\beta(\infty)] := \lim_{t \to \infty} E[\beta(t)]\), with a special attention to its dependence on the size of the network. In particular, we shall see that the algorithm is asymptotically unbiased when \(\lim_{N \to \infty} E[\beta(\infty)] = 0\).

3. ANALYSIS

3.1 A simplistic bound

Using (1), it is immediate to compute that, given \(v\) to be the broadcasting node at time \(t\),

\[x_{av}(t+1) - x_{av}(t) = \frac{q}{N} \sum_{u \in N_v} (x_u(t) - x_u(t)).\] (2)

Then, we can obtain the following deterministic bound on the error introduced at each time step,

\[|x_{av}(t+1) - x_{av}(t)| \leq \frac{q}{N} d_v L \leq \frac{q d_{max}}{N} L,\] (3)

where \(d_v = |N_v|\) is the degree of node \(v\), and \(d_{max} = \max_{v \in V} d_v\). This simple bound is worth some informal remarks. Indeed, Equation (3) suggests that choosing a low value for the mixing parameter \(q\), and a graph with low degree \(d_{max}\) and large size \(N\), may ensure a small bias in the computation of the average. However, by choosing a large \(N\) or \(d_{max}\), one affects the speed of convergence of the algorithm, not without consequences. Assume one is interested in an accurate computation, and chooses low values for \(q\) and \(d_{max}\), compared to \(N\). This choice would likely imply a slow convergence, and in turn a slow convergence enforces to run the algorithms for a larger number of steps, in order to meet the same precision requirement. These extra steps would introduce extra errors, thus possibly wasting the desired advantage in the accuracy. We argue from this discussion that a finer analysis is needed in order to understand the accumulation of errors.

3.2 The average as a martingale

In this section, we shall derive a general bound on \(E[\beta(\infty)]\) in terms of the topology of the graph. The derivation is based on applying the theory of martingales to the stochastic processes \(x_t(t)\) and \(x_{av}(t)\). The reader can find the essentials of martingale theory in Jacod and Protter [2003] or in Shiryaev [1989].

Definition 3.1. Given an increasing sequence (filtration) of \(\sigma\)-algebras \(\{\sigma_n\}_{n \in \mathbb{Z}_{\geq 0}}\), a sequence of random variables \(\{M_n\}_{n \in \mathbb{Z}_{\geq 0}}\) is a \(\sigma_n\)-martingale if \(E[M_n|\sigma_n] = M_n\), for any \(m \geq n\).

Our first result states that \(x_{av}(t)\) is a martingale with respect to the filtration induced by \(x_t(t)\). Given a set of random variables \(X\), let \(\sigma(X)\) denote the sigma-algebra generated by the random variables in \(X\).

Proposition 3.1. Let us consider the BGA algorithm and the filtration \(\mathcal{F}_t = \sigma\{x(s) : s \leq t\}\). Then the sequence of random variables \(\{x_{av}(t)\}_{t \in \mathbb{Z}_{\geq 0}}\) is a square-integrable \(\mathcal{F}_t\)-martingale.

Proof: First, note that \(x_{av}(t)\) is \(\mathcal{F}_t\)-measurable. Moreover, Equation (2) implies that for all \(t \geq 0\),

\[E[x_{av}(t+1) - x_{av}(t)] = \frac{1}{N} \sum_{v \in V} \left(\frac{q}{N} \sum_{u \in N_v} (x_u(t) - x_u(t))\right)\]

\[= \frac{q}{N^2} \sum_{v \in V} d_v x_v(t) - \sum_{v \in V} \sum_{u \in N_v} x_u(t)\]

\[= \frac{q}{N^2} \sum_{v \in V} d_v x_v(t) - \sum_{u \in V} d_u x_u(t) = 0.\]

Then, the sequence of random variables \(\{x_{av}(t)\}_{t \in \mathbb{Z}_{\geq 0}}\) is an \(\mathcal{F}_t\)-martingale. Moreover, since \(x_{av}(t) \in [0, L]\) the martingale is bounded in \(L^p\) for every \(p \geq 1\), and in particular square-integrable.
Let us define the distance from the current average, or disagreement, as
\[ d(t) := \frac{1}{N} \sum_{v \in V} (x_v(t) - x^\text{av}(t))^2. \]

Using this definition, we can prove an inequality which is a refinement of (3).

**Lemma 3.2.** The increments of the martingale \( \{x_v(t)\}_{t \in \mathbb{Z}^+} \) have bounded variance, namely for all \( t \geq 0 \),
\[ \mathbb{E}[(x_v(t + 1) - x^\text{av}(t))^2 | F_t] \leq \frac{q^2 d_{\text{max}}^2}{N^2} d(t). \]

**Proof:** By (2), we have
\[
\mathbb{E}[(x_v(t + 1) - x^\text{av}(t))^2 | F_t] \\
= \frac{1}{N} \sum_{u \in V} \left( \frac{q}{N} \sum_{u \in N_v} (x_v(t) - x_u(t))^2 \right) \\
\leq \frac{1}{N} \sum_{u \in V} \frac{q^2}{N^2} d_u \sum_{u \in N_v} (x_v(t) - x_u(t))^2 \\
\leq \frac{1}{N} \sum_{u \in V} \frac{q^2 d_{\text{max}}^2}{N^2} d_v(x_v(t) - x^\text{av}(t))^2 \\
\quad + \sum_{u \in N_v} (x_v(t) - x^\text{av}(t))^2 \\
\leq \frac{q^2 d_{\text{max}}^2}{N^2} d(t).
\]

This completes the proof. \( \square \)

Lemma 3.2 relates the increments of the martingale \( x_v(t) \) to the current disagreement. The convergence to zero of the disagreement has been widely studied, and the rate of convergence is known to depend on the properties of the network. Indeed, combining Lemmas 3 and 4 in Aysal et al. [2009b], we obtain that
\[ \mathbb{E}[d(t)] \leq \rho^t d(0), \]
and that\(^1\)
\[ \rho \leq 1 - \frac{2q(1-q)}{N} \lambda_1. \] (4)

Using the above facts, we are going to prove the next result about the asymptotic behavior of the bias as \( t \to +\infty \).

**Proposition 3.3.** With the above definitions,
\[ \mathbb{E}[\beta(\infty)] \leq 2 d(0) \frac{q}{1 - q} \frac{d_{\text{max}}^2}{N \lambda_1}. \]

**Proof:** Lemma 3.2 implies that
\[ \mathbb{E}[(x^\text{av}(s + 1) - x^\text{av}(s))^2] = \mathbb{E} \left[ \mathbb{E}[(x^\text{av}(s + 1) - x^\text{av}(s))^2 | F_s] \right] \\
\leq \frac{q^2 d_{\text{max}}^2}{N^2} \mathbb{E}[d(s)] \\
\leq 4 d(0) \frac{q^2 d_{\text{max}}^2}{N^2} \rho^s. \]

On the other hand, using the orthogonality of the increments of square-integrable martingales we observe that
\[ \mathbb{E}[(x^\text{av}(t) - x^\text{av}(0))^2] = \mathbb{E} \left[ \sum_{s=0}^{t-1} (x^\text{av}(s + 1) - x^\text{av}(s))^2 \right] \\
= \sum_{s=0}^{t-1} \mathbb{E}[(x^\text{av}(s + 1) - x^\text{av}(s))^2] \\
+ 2 \sum_{s=1}^{t-1} \sum_{r < s} \mathbb{E}[(x^\text{av}(s + 1) - x^\text{av}(s)) \cdot (x^\text{av}(r + 1) - x^\text{av}(r))] \\
= \sum_{s=0}^{t-1} \mathbb{E}[(x^\text{av}(s + 1) - x^\text{av}(s))^2] \\
\leq 4 d(0) \frac{q^2 d_{\text{max}}^2}{N^2} \frac{t-1}{1 - \rho} \rho^s. \]

Then, the inequality (4) implies that
\[ \lim_{t \to +\infty} \mathbb{E}[(x^\text{av}(t) - x^\text{av}(0))^2] \leq 2 d(0) \frac{q}{1 - q} \frac{d_{\text{max}}^2}{N \lambda_1}. \]
and the thesis follows by the Dominated Convergence Theorem. \( \square \)

Note that the proof of Proposition 3.3 also implies that
\[ \sup_{t \in \mathbb{N}} \mathbb{E}[(x^\text{av}(t) - x^\text{av}(0))^2] \leq 2 d(0) \frac{q}{1 - q} \frac{d_{\text{max}}^2}{N \lambda_1}. \]

On the other hand, we know by Doob’s maximal inequality that, if \( M_t \) is a convergent square-integrable martingale, then \( \mathbb{E} \left[ \sup_{t \in \mathbb{N}} M_t^2 \right] \leq 4 \mathbb{E} \left[ \lim_{t \to \infty} M_t^2 \right] \). Hence, we can immediately obtain the following finite-time counterpart of Proposition 3.3.

**Theorem 3.4.** With the above definitions,
\[ \mathbb{E} \left[ \sup_{t \in \mathbb{N}} \beta(t) \right] \leq 8 d(0) \frac{q}{1 - q} \frac{d_{\text{max}}^2}{N \lambda_1}. \]

Let us now consider a sequence of graphs \( G_N \) of increasing size \( N \). In such a sequence, both \( d_{\text{max}} \) and \( \lambda_1 \) depend on \( N \). In this setting, Proposition 3.3 implies the following corollary.

**Corollary 3.5.** Let \( I \subseteq \mathbb{N} \) and \( \{G_N\}_{N \in I} \) be a sequence of graphs such that \( G_N = (\mathcal{V}_N, \mathcal{E}_N) \) and \(|\mathcal{V}_N| = N \). If
\[ \frac{d_{\text{max}}^2}{\lambda_1} = o(N) \quad \text{as} \quad N \to +\infty, \]
then the BGA algorithm is asymptotically unbiased.

Note that, since \( x^\text{av}(t) \) converges almost surely and \( \lim_{t \to +\infty} x^\text{av}(t) \in [0, L] \), it is trivial to find an upper bound on \( \mathbb{E}[\beta(\infty)] \) which does not depend on \( N \). The interest of the above corollary is in giving a sufficient condition for \( \mathbb{E}[\beta(\infty)] \) to be \( o(1) \) as \( N \to +\infty \).

**Remark 3.6.** (Distributed estimation). In the applications, one is often interested in computing an average because the average is the maximum likelihood estimator of the expectation of a random variable. In such context, the average enjoys the property that the mean square error, committed by approximating the expectation by the average of \( N \) samples, is equal to \( 1/N \). For this reason, we would like to ensure that the bias introduced by the
Broadcast Gossip Algorithm is not larger than (a constant times) $1/N$. Applying Markov’s inequality, we see from Proposition 3.3 that for any $c > 0$, 
\[
\Pr[\beta(\infty) > c] \leq 2d(0) \frac{q}{1-q} \frac{d_{\max}^2}{N\lambda_1} c.
\]
Then, taking $c = 1/N$, we get 
\[
\Pr[\beta(\infty) > \frac{1}{N}] \leq 2d(0) \frac{q}{1-q} \frac{d_{\max}^2}{\lambda_1}.
\]
Provided the right-hand-side of this inequality does not diverge, we can choose the mixing parameter $q$ so that with a positive given probability the bias is below $1/N$. This ensures that the BGA is accurate and suitable to compute averages for estimation purposes in large networks. \[\blacksquare\]

4. EXAMPLES

In this section, we discuss the asymptotical bias of the BGA on several example topologies which have been considered in the literature. In our discussion, given a graph $G_N$ of size $N$, we denote by $\lambda_1(N)$ and $d_{\max}(N)$ the spectral gap and maximum degree, respectively.

The results presented in Section 3 are sufficient to prove asymptotical unbiasedness on graphs with a “large” spectral gap, like expanders and hypercube graphs.

Example 4.1. (Expander graphs). A sequence of graphs is said to be an expander sequence if there exist $d \in \mathbb{N}$ and $c > 0$ such that for every $N$, $\lambda_1(N) \geq c$ and $d_{\max}(N) \leq d$. In this case, $d_{\max}^2/N$ is bounded, and this fact implies by Proposition 3.3 that 
\[
\mathbb{E}[\beta(\infty)] = O\left(\frac{1}{N}\right) \quad \text{as } N \to +\infty.
\]
In particular, the BGA is asymptotically unbiased and Remark 3.6 applies.

Example 4.2. (Hypercube graphs). The $n$-dimensional hypercube graph is the graph obtained drawing the edges of a $n$-dimensional hypercube. It has $N = 2^n$ nodes which can be identified with the binary words of length $n$, and two nodes are neighbors if the corresponding binary words differ in only one component. For these graphs it is known, for instance from [Frasca et al., 2009, Example 7], that $\lambda_1(N) = \Theta(1/\log N)$ and $d_{\max}(N) = O(\log N)$. Then,
\[
\mathbb{E}[\beta(\infty)] = O\left(\frac{\log^3 N}{N}\right) \quad \text{as } N \to +\infty,
\]
and the BGA is asymptotically unbiased.

Instead, our bounds are not tight on sequences of lattices with fixed degree.

Example 4.3. ($k$-dimensional square lattices). We consider square lattices/grid obtained by tiling a $k$-dimensional torus, with $N = n^k$ nodes. For such $k$-torus graphs we know [Carli et al., 2008, Theorem 6] that $\lambda_1(N) = \Theta(1/N^{2/k})$ and $d_{\max}(N) = 2k$. Then,
\[
\mathbb{E}[\beta(\infty)] = O\left(\frac{1}{N^{1-2/k}}\right) \quad \text{as } N \to +\infty,
\]
and we argue that $k$–lattices are asymptotically unbiased if $k \geq 3$. However, in Fagnani and Frasca [2011] is proved by different techniques that the BGA is asymptotically unbiased also if $k = 1, 2$.

An example of biased topology is the complete graph.

Example 4.4. (Complete graphs). For these graphs, $\lambda_1(N) = N$ and $d_{\max}(N) = N - 1$. Then, we can not conclude from Proposition 3.3 that the BGA is asymptotically unbiased on complete graphs. Actually, in Fagnani and Zampieri [2008a] it is shown that the BGA is not asymptotically unbiased on complete graphs, and in particular
\[
\mathbb{E}[\beta(\infty)] = \frac{\text{Var}(x(0))}{2} = O\left(\frac{N-1}{N}\right),
\]
where $\text{Var}(x(0))$ is the sample variance of the initial condition.

We conclude with an example of random geometric graphs which is of major interest in the analysis of wireless networks. Although our bound is not sufficient to prove unbiasedness, the latter property is conjectured from the simulations results in the next section.

Example 4.5. (Random geometric graphs). Consider random sequences of geometric graphs constructed as follows. We sample $N$ points from a uniform distribution over the unit square, and we let nodes be connected if the two corresponding points in the square are less than $r(N)$ far apart, with $r(N) = 1.1\sqrt{\frac{\log N}{N}}$. We know from Penrose [2003] that this choice ensures that with “high” probability the graph is connected, $\lambda_1(N) = \Theta(1/N)$, and $d_{\max}(N) = O(\log N)$. Then, we argue that with high probability $\mathbb{E}[\beta(\infty)] = O(\log^2 N)$ as $N \to +\infty$, and we can not conclude asymptotical unbiasedness.

5. SIMULATIONS

We have extensively simulated the evolution of BGA algorithm on sequences of graphs, and in particular on the random topologies presented in Section 4. In this section, we account for our results about the dependence of the bias $\beta(\infty)$ on the size $N$ and on the parameter $q$. Our simulation setup is as follows. Let $q, N$ and the graph topology be chosen. For every run of the algorithm we generate a vector of initial conditions $x(0)$, sampling\(^2\) from a uniform distribution over $[0,1]^N$. Then, we run the algorithm until the disagreement $d(t)$ is below a small threshold $\varepsilon$, which we set at $10^{-4}$. At this time $T^\varepsilon = \inf\{t \geq 0 : d(t) \leq \varepsilon\}$, the algorithm is stopped, and $\beta(T^\varepsilon)$ is evaluated. In order to simulate the expectation $\mathbb{E}[\beta(\infty)]$, we average the outcome of 1000 realizations of $\beta(T^\varepsilon)$. Our results about the dependence on $N$ are summarized in Figure 1, which plots the average bias against $N$ in a log-log diagram. As expected, complete graphs are not asymptotically unbiased, while those topologies whose degree is $o(N)$ are asymptotically unbiased. More precisely, on sequences of torus graphs the bias is $\Theta(N^{-1})$, whereas on hypercube and random geometric graphs the bias is $\Theta(\log(N)/N)$. Overall, our set of simulations suggests that
\[
\mathbb{E}[\beta(\infty)] = \Theta\left(\frac{d_{\max}(N)}{N}\right) \quad \text{as } N \to \infty.
\]

\(^2\) If the topology is random – random geometric as described above – it is also sampled at this stage. Disconnected realizations are discarded: however, disconnected realizations are very few in our random geometric setting and their number decreases as $N$ grows, so that they are less than 2% when $N > 50$. 

10030
In this paper, we have proved that on any (undirected) graph, $E[\beta(\infty)] = O\left(\frac{d_{\text{max}}^2}{N\lambda_1}\right)$ as $N$ diverges, and in particular the BGA is asymptotically unbiased on expander graphs. On the other hand, simulations suggest that, on sequences of (almost) regular graphs with degree $d(N)$, the bias is such that $E[\beta(\infty)] = \Theta\left(\frac{d(N)}{N}\right)$ . This gap between simulations and theoretical understanding leaves the opportunity for further research on the topic. Our future research will be devoted to prove asymptotic unbiasedness on a wider set of topologies, looking for a tighter bound on $E[\beta(\infty)]$, and more in general to study the trade-offs between speed and accuracy in gossip algorithms.

REFERENCES

T. C. Aysal, A. D. Sarwate, and A. G. Dimakis. Reaching consensus in wireless networks with probabilistic broadcast. In Allerton Conf. on Communications, Control and Computing, pages 732–739, Monticello, IL, September 2009a.

T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione. Broadcast gossip algorithms for consensus. IEEE Transactions on Signal Processing, 57(7):2748–2761, 2009b.

S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. IEEE Transactions on Information Theory, 52(6):2508–2530, 2006.

R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri. Communication constraints in the average consensus problem. Automatica, 44(3):671–684, 2008.

A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione. Gossip algorithms for distributed signal processing. Proceedings of the IEEE, 98(11):1847–1864, 2010.

F. Fagnani and P. Frasca. Broadcast gossip averaging: interference and unbiasedness in large Abelian Cayley networks. IEEE Journal of Selected Topics in Signal Processing, 2011. doi: 10.1109/JSTSP.2011.2123078. In press.

F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. IEEE Journal on Selected Areas in Communications, 26(4):634–649, 2008a.

F. Fagnani and S. Zampieri. Asymmetric randomized gossip algorithms for consensus. In IFAC World Congress, pages 9052–9056, 2008b.

F. Frasca, R. Carli, F. Fagnani, and S. Zampieri. Average consensus on networks with quantized communication. International Journal of Robust and Nonlinear Control, 19(16):1787–1816, 2009.

J. Jacod and P. Protter. Probability Essentials. Springer, 2003.

M. Penrose. Random Geometric Graphs. Oxford Studies in Probability. Oxford University Press, 2003.

A. N. Shiryaev. Probability. Graduate Texts in Mathematics. Springer, 2 edition, 1989.

A. Tahbaz-Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. IEEE Transactions on Automatic Control, 55(1):225–230, 2010.