GENERALIZED CONSENSUS COMPUTATION IN NETWORKED SYSTEMS WITH ERASURE LINKS

Michael G. Rabbat, Robert D. Nowak, and James A. Bucklew

University of Wisconsin - Madison
Department of Electrical and Computer Engineering
1415 Engineering Drive, Madison, Wisconsin 53706
Email: rabbat@cae.wisc.edu, {nowak, bucklew}@engr.wisc.edu

ABSTRACT

We study consensus problems in networked systems with unreliable links. Our contributions are two-fold. First, we derive a family of decentralized consensus algorithms for minimizing a sum of convex functions, $\sum_{i=1}^{N} f_i(x)$, where each function f_i only depends on information at one node in the network. Computing the consensus average is a special case in this setting. Then, we construct a modified algorithm which is resilient in situations where the channels between nodes act as binary erasure channels. The flexibility and efficacy of our approach is demonstrated through an application of robust estimation.

1. LOCAL CONSENSUS ALGORITHMS

The problem of consensus computation in networked systems such as wireless sensor networks has recently received a great deal of attention in the research community [1, 2, 3, 4, 5]. Much of the work to date has focused on the special case of computing the average. In this paper we construct decentralized consensus algorithms for computing more general quantities in networked systems. Robust estimation is considered as an example application. A theoretical analysis demonstrates that the proposed algorithms converge to the desired quantity even when wireless communication links are lossy.

In many applications of wireless sensor networks it is desirable to carry out a computation of the form

$$\text{minimize} \quad \sum_{i=1}^{N} f_i(x), \tag{1}$$

over $x \in \mathbb{R}^d$, where the functions $f_i(x)$ only depend on information at node i and there are N nodes in the network [6, 7]. Such computations arise in a variety of applications including parameter estimation, aggregation, and distributed optimization. Similar problems may also arise as subcomponents of more complex tasks such as decentralized field estimation [8].

In this paper we propose a family of consensus algorithms addressing this problem, such that each node in the network has the solution, x^*, when computation terminates. We also impose the restriction that information is only exchanged between nodes in direct radio contact (*i.e.,* our algorithms are *local*). We represent these relationships in a binary N-by-N adjacency matrix, G, of the corresponding communication graph. Thus, $G_{i,j} = 1$ if and only if node j receives transmissions from node i, and $G_{i,j} = 0$ otherwise. The graph need not be symmetric; we merely make the usual assumption that it is *strongly connected* (*i.e.,* for each pair of nodes, i, j, there are directed paths in the graph both from i to j and from j to i). Later, we will be interested in the case where the links are unreliable. In this setting $G_{i,j} = 1$ only if there is a non-vanishing probability that node j will receive transmissions from node i, and $G_{i,j} = 0$ otherwise. We will also frequently refer to the set of neighbors, $\mathcal{N}_i$, of node i which consists of all nodes j for which $G_{j,i} = 1$; *i.e.,* node j is in $\mathcal{N}_i$ if node i receives transmissions from j.

This work was sponsored by Applied Signal Technology, the National Science Foundation, grants CCR-0310889, CCR-0325571, and ANI-0099148, and the Office of Naval Research, grant N00014-00-1-0966.

2. A PRIMAL-DUAL APPROACH TO LOCAL CONSENSUS

In this section we derive a primal-dual algorithm for solving the problem (1). The basic idea begins with formulating an equivalent problem to (1) with constraints that impose local communication and a consensual outcome. We then decompose this problem into N subproblems which can be executed in parallel at each node and verify that once each node solves its own subproblem the optimal consensus has been reached. Duality theory and decomposition methods plays a major role in the field of optimization. See [9] for a comprehensive treatment of the subject of duality, and see [10] and [11] for more on the use of duality and decomposition methods in parallel and distributed optimization.

We begin by introducing variables x_i, at each node, $i = 1, \ldots, N$, and observe that the problem

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{N} f_i(x_i) \\ \text{subject to} \quad & x_i = x \\ & x \in \mathbb{R}^d, \end{aligned} \tag{2}$$

is equivalent to (1). For the remainder of the paper we focus on the case where the f_i are strictly convex, continuously differentiable functions which are bounded below[1]. Although the constraints of

[1]Differentiability is not critical here; we assume it to simplify exposition. Extensions to nondifferentiable functions are straightforward and only involve a modification to the step, (7). Additionally, the assumption of strict convexity can be relaxed to plain convexity if one takes an augmented Lagrangian approach rather than the plain primal-dual approach

this new problem reflect our desired goal of arriving at a consensus, they also introduce a coupling which would require every node to communicate with every other node in order to reach a solution. We overcome this issue by rephrasing the problem in terms of local constraints as follows.

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{N} f_i(x_i) \\ \text{subject to} & x_i - x_j \geq 0, \text{ for all } j \in \mathcal{N}_i, \ i = 1, \ldots, N. \end{array} \tag{3}$$

The solution of this problem clearly coincides with the solution of (1) provided that the communication graph is strongly connected. To see this, note that in order for the equality constraints to hold at all nodes in the network, we require $x_i \geq x_j$ and $x_j \geq x_i$.

Next, we decompose this problem into N subproblems which are local to each node. These subproblems have the form

$$\begin{array}{ll} \text{minimize} & f_i(x_i) \\ \text{subject to} & x_i - x_j \geq 0, \ j \in \mathcal{N}_i, \end{array} \tag{4}$$

where optimization is being performed on the variable x_i. Let $\boldsymbol{\lambda}_i = \{\lambda_{i,j} : j \in \mathcal{N}_i\}$ denote a collection of Lagrange multipliers for the ith subproblem, and let $\boldsymbol{x} = \{x_i : i = 1, \ldots, N\}$, $\boldsymbol{\lambda} = \{\boldsymbol{\lambda}_i : i = 1, \ldots, N\}$. The Lagrangian function for the primal subproblem, (4), is

$$\mathcal{L}_i(x_i, \boldsymbol{\lambda}_i) = f_i(x_i) + \sum_{j \in \mathcal{N}_i} \lambda_{i,j}(x_i - x_j). \tag{5}$$

Let $q_i(\boldsymbol{\lambda}_i) = \inf_{x_i} \mathcal{L}_i(x_i, \boldsymbol{\lambda}_i)$. Then the dual subproblem is

$$\text{maximize} \quad q_i(\boldsymbol{\lambda}_i), \tag{6}$$

over $\boldsymbol{\lambda}_i$. The primal-dual approach to solving this problem alternates between iterates

$$x_i^k = \arg\min_{x_i} \mathcal{L}_i(\boldsymbol{x}, \boldsymbol{\lambda}^k) \tag{7}$$

$$\lambda_{i,j}^{k+1} = \lambda_{i,j}^k + \mu(x_i^k - x_j^k), \tag{8}$$

where $\mu > 0$ is a small constant step size [11]. Intuitively, this algorithm calculates the value of the dual cost function $q_i(\boldsymbol{\lambda}_i)$ near the current iterate $\boldsymbol{\lambda}_i^k$. Then we try to maximize $q_i(\boldsymbol{\lambda}_i)$ to solve the dual problem. Based on our assumption that f_i is strictly convex, one can compute x_i^{k+1} in the first step by solving $\frac{\partial}{\partial x_i} f_i(x_i) = -\sum_{j \in \mathcal{N}_i} \lambda_{i,j}^k$ for x_i. Then we perform a gradient ascent step on the Lagrange multipliers $\lambda_{i,j}^k$ to increase the dual cost function value.

The underlying principle behind this "primal-dual" approach is that, by solving the dual problem we simultaneously solve the primal problem which we were originally interested in. More concretely, because the f_i are strictly convex we are guaranteed that there exists a solution $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ for which

$$\inf_x f_i(x) = f_i(x_i^*) = \mathcal{L}_i(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = q_i(\boldsymbol{\lambda}^*) = \sup_{\boldsymbol{\lambda}} q_i(\boldsymbol{\lambda}),$$

and $\nabla_x \mathcal{L}_i(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0$ (see, *e.g.,* Appendix C in [10]). Observe that the Lagrangian function for the network-wide problem (3) is $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \sum_{i=1}^N \mathcal{L}_i(\boldsymbol{x}, \boldsymbol{\lambda})$. The condition $\nabla_x \mathcal{L}_i(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0$ for each subproblem implies that $\nabla_x \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0$; *i.e.,* if $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ simultaneously solves the subproblems at each node then it also solves the network-wide problem.

adopted here. However this modification is unnecessary in the applications considered here and is beyond the scope of this paper.

3. PRACTICAL IMPLEMENTATION OVER ERASURE CHANNELS

In a practical implementation of this algorithm, each node stores and maintains the set of local Lagrange multipliers, $\boldsymbol{\lambda}_i$. After performing the update (7), nodes exchange the values x_i^k with their neighbors (*e.g.,* via local broadcast) so that they can be used to compute (8). In what follows, we assume that these iterations occur on a regular schedule which could possibly be realized via a slotted TDMA protocol. However, similar results hold for an asynchronous scheme where nodes decide to transmit at random time intervals.

Asymmetry and link outage has been observed in a number of experiments with real wireless sensor devices [12]. We model this communication uncertainty by modelling each directed link as a binary erasure channel, independent of other channels, with i.i.d. realizations at each iteration. Formally, let $W_{i,j}^k$ be i.i.d. Bernoulli random variables with the event $W_{i,j}^k = 1$ indicating that a transmission on the link from j to i at iteration k arrives successfully. We make the restriction that $\alpha_{i,j} = 0$ if and only if $G_{i,j} = 0$ so that strong connectivity of the graph is probabilistically maintained, and we assume that the erasure channels are stationary with parameters $\alpha_{i,j} = P(W_{i,j}^k = 1)$ known at the receiving node, i. Assuming knowledge of the $\alpha_{i,j}$ is not unreasonable, as these parameters are commonly estimated in existing wireless sensor network systems [13]. This simple model captures many of the salient features of wireless sensor networks in this setup. One could, more generally, take the sequence of channel realizations, $\{W^k\}$, to be a Markov chain and the same analysis techniques and similar correction methodology would apply.

Next, we modify the update equations (7)-(8) to account for link outages so that the resulting convergence properties hold regardless of the sequence of channel realizations. Consider the modified Lagrange multiplier update

$$\lambda_{i,j}^{k+1} = \lambda_{i,j}^k + \mu \frac{x_i^k - x_j^k}{\alpha_{i,j}} W_{i,j}^k.$$

Thus, the multiplier $\lambda_{i,j}$ at node i is only updated when a transmission is received from node j (*i.e.,* $W_{i,j}^k = 1$), and the amount by which it is updated is now scaled by $1/\alpha_{i,j}$. Intuitively, our method of compensating for erasures is founded on the principle that the less a channel permits communication, the greater that information received over that channel should be valued. Of course, in practice, channels with $\alpha_{i,j}$ approaching zero could cause the algorithm to become numerically instable. A reasonable solution is to have nodes locally only admit neighbors with reception rates $\alpha_{i,j}$ greater than some threshold. This scheme works so long as the communication graph induced by thresholding is still strongly connected. In the following sections we consider two specific applications of this algorithm and find that by applying modification just described, the rate of convergence of our algorithm is not effected by erasures.

4. AVERAGE CONSENSUS

Average consensus computation has been a problem of particular interest in the literature. In this section we formulate a consensus averaging algorithm in the framework just described. Using tools developed in [14], we show that the modified primal-dual algorithm converges to the consensus average as desired, and we briefly discuss the relation of this algorithm to existing algorithms in the literature.

Suppose each sensor takes a scalar measurement, u_i. Computing the average, $\bar{u} = \frac{1}{N}\sum_{i=1}^{N} u_i$ is equivalent to minimizing the sum of squares, $\frac{1}{2}\sum_{i=1}^{N}(x_i - u_i)^2$. The corresponding modified primal-dual steps are

$$x_i^k = u_i - \sum_{j\in\mathcal{N}_i} \lambda_{i,j}^k \tag{9}$$

$$\lambda_{i,j}^{k+1} = \lambda_{i,j}^k + \mu \frac{x_i^k - x_j^k}{\alpha_{i,j}} W_{i,j}^k. \tag{10}$$

Let $y_i^k = \sum_{j\in\mathcal{N}_i} \lambda_{i,j}^k$. Then combining (9) and (10) we have

$$\begin{aligned} y_i^{k+1} &= y_i^k + \mu \sum_{j\in\mathcal{N}_i} \frac{x_i^k - x_j^k}{\alpha_{i,j}} W_{i,j}^k \\ &= y_i^k + \mu \sum_{j\in\mathcal{N}_i} \frac{u_i - y_i^k - (u_j - y_j^k)}{\alpha_{i,j}} W_{i,j}^k. \end{aligned} \tag{11}$$

This update fits the general algorithmic form studied in [14], so we can employ tools developed therein to analyze the convergence properties of our algorithm. In particular, we would like to confirm that the algorithm indeed converges, identify the value it converges to, and quantify the rate of convergence.

Let W^k denote the matrix with entries $W_{i,j}^k$, and let $\boldsymbol{y}^k$ and $\boldsymbol{u}$ be vectors with components y_i^k and u_i respectively. We focus our analysis on the update equation

$$H_i(\boldsymbol{y}, W) = \sum_{j\in\mathcal{N}_i} \frac{u_i - y_i - u_j + y_j}{\alpha_{i,j}} W_{i,j}.$$

Conditioning on $\boldsymbol{y}$, we average with respect to the distribution on erasure events, W, to obtain

$$\begin{aligned} &\bar{H}_i(\boldsymbol{y}) \\ &= \sum_{j\in\mathcal{N}_i} [u_i - y_i - u_j + y_j] \\ &= |\mathcal{N}_i| \left[\left(u_i - \sum_{j\in\mathcal{N}_i} \frac{u_j}{|\mathcal{N}_i|} \right) - \left(y_i - \sum_{j\in\mathcal{N}_i} \frac{y_j}{|\mathcal{N}_i|} \right) \right], \end{aligned}$$

where $|\mathcal{N}_i|$ denotes the cardinality of the set $\mathcal{N}_i$. Let D be a diagonal matrix with entries $D_{i,i} = |\mathcal{N}_i|$ and let I denote the identity matrix. Stacking the components $\bar{H}_i(\boldsymbol{y})$ into a vector and rewriting the above equation in vector-matrix notation we have

$$\bar{H}(\boldsymbol{y}) = D[(I - P)\boldsymbol{u} - (I - P)\boldsymbol{y}],$$

where P is a N-by-N matrix with entries

$$P_{i,j} = \begin{cases} \frac{1}{|\mathcal{N}_i|} & \text{if } G_{i,j} = 1 \\ 0 & \text{otherwise,} \end{cases}$$

corresponding to the natural random walk on the graph G.

The theory developed in [14] uses $\bar{H}(\boldsymbol{y})$ to make statements about a time scaled continuous version of $\boldsymbol{y}^k$, defined by $Y_\mu(t) = \boldsymbol{y}^{\lfloor t/\mu \rfloor}$, which one can think of as sampling and holding $\boldsymbol{y}^k$ for μ seconds. Theorem 1 in [14] tells us that as $\mu \to 0$, under a number of unrestrictive technical conditions which are all satisfied in our setting, $Y_\mu(t)$, converges almost surely to the trajectory

$$Y(t) = \boldsymbol{u} - \exp\{-D(P - I)t\}\boldsymbol{u},$$

after initializing $Y(0) = 0$, which we readily achieve by initializing $\lambda_{i,j}^0 = 0$. This implies convergence of the similarly defined process $X_\mu(t) = \boldsymbol{x}^{\lfloor t/\mu \rfloor}$ to

$$X(t) = \exp\{-D(P - I)t\}\boldsymbol{u},$$

as $\mu \to 0$. Observe that $P = D^{-1}G$. Then $-D(P - I) = D - G$. For a graph, G, the matrix $L = D - G$ is referred to as the graph Laplacian matrix. It is well known that if G is strongly connected (as we are assuming) then L has a minimum eigenvalue 0 with corresponding eigenvector being the vector of all 1's (or uniform, after normalization), L has rank $N - 1$, and thus all other eigenvalues of L are strictly positive [15, 16]. This implies that $X(t) \to (\mathbf{1}^T\boldsymbol{u}/n)\mathbf{1} = \bar{u}\mathbf{1}$ as $t \to \infty$, as desired.

Although the above results characterize the asymptotic behavior of the algorithm as $\mu \to 0$, we emphasize that in practice we fix a step size $\mu > 0$ and stick with it. One can interpret the asymptotic convergence results as suggesting that for reasonably small μ, the iterates $\boldsymbol{x}^k$ approximately follow

$$\boldsymbol{x}^k = \exp\{-D(P - I)k\mu\}\boldsymbol{u}. \tag{12}$$

In order to assess how closely the fixed step size algorithm ($\mu > 0$) will follow this trajectory, we analyze a properly scaled version of the error,

$$\mathcal{E}_\mu(t) = \big(X_\mu(t) - X(t)\big)/\sqrt{\mu},$$

using tools from [14]. Observe that

$$\begin{aligned} &H_i(\boldsymbol{y}, W) - \bar{H}_i(\boldsymbol{y}) \\ &\quad = \sum_{j\in\mathcal{N}_i} [u_i - y_i - u_j + y_j] \left(\frac{W_{i,j}}{\alpha_{i,j}} - 1 \right) \end{aligned}$$

and let $V(\boldsymbol{y}, W) = \big(H(\boldsymbol{y}, W) - \bar{H}(\boldsymbol{y})\big)\big(H(\boldsymbol{y}, W) - \bar{H}(\boldsymbol{y})\big)^T$. Again, we condition on $\boldsymbol{y}$ and smooth over the variables W in $V(\boldsymbol{y}, W)$ to obtain the matrix $\bar{V}(\boldsymbol{y})$ which is diagonal (based on the independence of different channels) with entries

$$\bar{V}_{i,i}(\boldsymbol{y}) = \sum_{j\in\mathcal{N}_i} \frac{(u_i - y_i - u_j + y_j)^2(1 - \alpha_{i,j})}{\alpha_i, j}.$$

Theorem 2 of [14] relates the error process $\mathcal{E}_\mu(t)$ to the solution of a stochastic differential equation being driven by a Wiener process with variance equal to $\bar{V}_{i,i}(\boldsymbol{y})$. Since $x_i = u_i - y_i \to \bar{u}$ for all i as $t \to \infty$, $\bar{V}_{i,i}(\boldsymbol{y}) \to 0$ also. This implies that the error process $\mathcal{E}_\mu(t)$ converges to zero in exactly the same manner that $X(t)$ converges to $\bar{u}\mathbf{1}$. Convergence to the consensus average was expected when there are no channel losses, however this result verifies that the modified consensus algorithm indeed converges to the average at an exponential rate even when communication occurs over an erasure channel.

4.1. Relation To Other Consensus Algorithms

Recall the update equation, (11), in the special case where $\alpha_{i,j} = 1$ for all i, j, so that there are no erasures. In vector notation we have

$$\boldsymbol{y}^{k+1} = \boldsymbol{y}^k + \mu D(I - P)(\boldsymbol{u} - \boldsymbol{y}^k).$$

Subtracting both sides of this equation from $\boldsymbol{u}$ gives

$$\boldsymbol{u} - \boldsymbol{y}^{k+1} = \boldsymbol{u} - \boldsymbol{y}^k + \mu D(P - I)(\boldsymbol{u} - \boldsymbol{y}^k),$$

or, equivalently, with initial condition $\boldsymbol{x}^0 = \boldsymbol{u}$,

$$\begin{aligned} \boldsymbol{x}^{k+1} &= \boldsymbol{x}^k + \mu L \boldsymbol{x}^k \\ &= A\boldsymbol{x}^k \end{aligned}$$

which bears strong resemblance to the algorithms studied in [1, 2, 3, 4, 5]. In fact, by applying additional weighting variables it is possible to mimic any of these algorithms in our framework. We also note that others have analyzed the performance of average consensus algorithms for networks with lossy links using techniques involving products of random matrices [17, 18]. Our framework allows us to view these issues from a different perspective.

A major drawback of other consensus algorithms is that after the initialization, $\boldsymbol{x}^0 = \boldsymbol{u}$, the data vector is never used again in the process. If, for example, the values $\boldsymbol{u}$ are also evolving (*e.g.*, sensors take new measurements on a slower time scale) then the other algorithms must be restarted with the new initialization. Alternatively, one can envision tracking a slowly time-varying parameter $\bar{u}(t)$ over time by simply modifying the values u_i in (9) without resetting the Lagrange multipliers the proposed primal-dual framework. We plan to investigate this and other extensions in our future work.

5. ROBUST CONSENSUS

To demonstrate the flexibility of our approach we have also designed and implemented a decentralized consensus algorithm for robust estimation. Wireless sensor networks are commonly envisioned as being composed of inexpensive devices which may be prone to hardware failures or security breach, and data integrity is a very relevant issue. Robust statistics is a mature field in which many problems related to detecting outliers and building resiliency against model mismatch have been thoroughly studied [19].

In this section, we replace the quadratic loss function, $f_i(x_i) = \frac{1}{2}(x_i - u_i)$, which led to average consensus, with the so-called "fair" loss function,

$$\rho_i(x_i - u_i) = 2\gamma^2 \left[\frac{|x_i - u_i|}{\gamma} - \ln\left(1 + \frac{|x_i - u_i|}{\gamma}\right)\right],$$

where γ is a user-specified parameter typically set around 1.5 [20]. This function not only has desirable statistical properties, such as a low sensitivity to the choice of γ, but it is also strictly convex. A primal-dual consensus algorithm for minimizing $\sum_{i=1}^N \rho_i(x_i - u_i)$ uses,

$$x_i^k = \begin{cases} u_i - \dfrac{\gamma \sum_{j\in\mathcal{N}_i} \lambda_{i,j}^k}{2\gamma - \sum_{j\in\mathcal{N}_i} \lambda_{i,j}^k} & \text{if } \sum_{j\in\mathcal{N}_i} \lambda_{i,j}^k \geq 0 \\ u_i - \dfrac{c \sum_{j\in\mathcal{N}_i} \lambda_{i,j}^k}{2\gamma + \sum_{j\in\mathcal{N}_i} \lambda_{i,j}^k} & \text{if } \sum_{j\in\mathcal{N}_i} \lambda_{i,j}^k < 0, \end{cases}$$

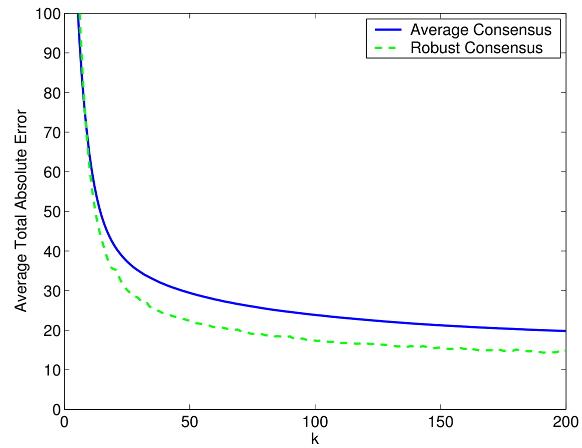

Fig. 1. Total absolute error over all nodes in the network averaged over 50 trials with erasure channels.

for its update rule, (7),

As an example of the utility of robust procedures in wireless sensor network applications, we have simulated a scenario in which 100 nodes, deployed according to a random geometric graph model, make measurements of a scalar parameter $s = 1$ in additive white Gaussian noise. Roughly two-thirds of the nodes function properly and make readings with variance $\sigma_1^2 = 1$. The other nodes are dysfunctional and make readings with variance $\sigma_2^2 = 20$. Two scenarios are simulated: lossless channels and erasure channels. In the case of erasure channels, all links have successful transmit probabilities $\alpha_{i,j} \geq 0.25$. For all results reported, the algorithms use a step size of $\mu = 0.01$, and a setting of $\gamma = 3.5$ was used in the "fair" loss function.

Figure 1 shows the total absolute error over all nodes in the network for the first 200 iterations, averaged over 50 simulations. Erasure channels were used in these simulations. The average consensus algorithm achieves a total error of roughly 21, and the robust consensus algorithm offers a 25% performance improvement, settling out at 15. Figure 2 shows similar curves for the case where channels are lossless, again, averaged over 50 simulations. As expected, lossy channels do not affect the performance of our modified primal-dual algorithm. Figure 3 compares trajectories taken at single node for lossless and erasure channels. The two trajectories follow roughly the same path, but some artifacts appear in the first few iterations due to transmissions received from a neighbor with a low $\alpha_{i,j}$. In this example, the receiving node has neighbors with $\alpha_{i,j}$'s ranging from 0.28 to 0.94. Figure 4 shows the trajectories taken by the robust consensus algorithm at each of the 100 nodes in the network, overlayed on one plot. Communication in this simulation was over erasure channels. The figure indicates that the algorithm does indeed achieve consensus as all trajectories are converging to the same point.

6. REFERENCES

[1] D.S. Scherber and H.C. Papadopoulos, "Locally constructed algorithms for distributed computations in ad-hoc networks," in *Proc. IEEE/ACM IPSN*, Berkeley, CA, April 2004.

[2] L. Xiao and S. Boyd, "Fast linear iterations for distributed

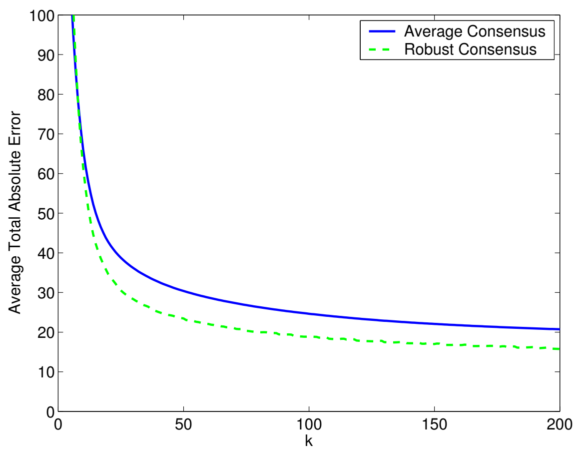

Fig. 2. Total absolute error over all nodes in the network averaged over 50 trials with lossless channels.

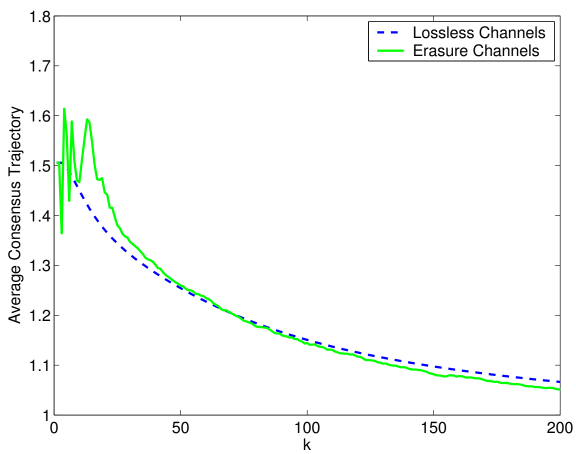

Fig. 3. Example trajectories at a single node for the same data realization, contrasting algorithm behavior over lossless and erasure channels.

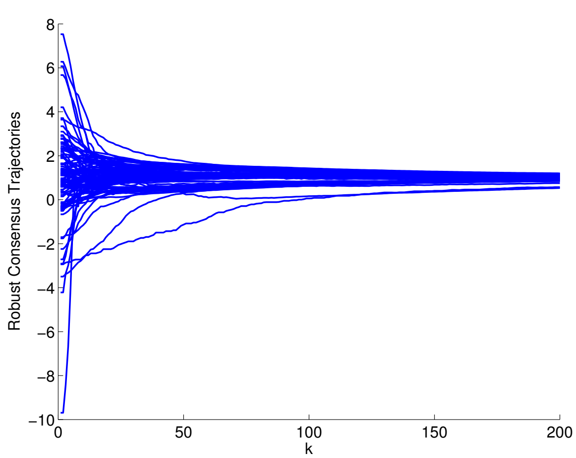

Fig. 4. Example trajectories taken by all nodes in a network running the robust consensus algorithm. Erasure links were used in this simulation.

averaging," *Systems and Control Letters*, vol. 53, no. 1, pp. 65–78, September 2004.

[3] J.Y. Chen, G. Pandurangan, and D. Xu, "Robust aggregates computation in wireless sensor networks: Distributed randomized algorithms and analysis," in *Proc. IEEE/ACM IPSN*, Los Angeles, CA, April 2005.

[4] R. Olfati-Saber and R.M. Murray, "Consensus problems in networks of agents with switching topology and time-delay," *IEEE Trans. Automatic Control*, vol. 49, no. 9, pp. 101–115, Sept. 2004.

[5] M. Alanyali, V. Saligrama, O. Savas, and S. Aeron, "Distributed bayesian hypothesis testing in sensor networks," in *Proc. American Control Conference*, 2004.

[6] M.G. Rabbat and R.D. Nowak, "Distributed optimization in sensor networks," in *Proc. ACM/IEEE IPSN*, Berkeley, CA, April 2004.

[7] M.G. Rabbat and R.D. Nowak, "Quantized incremental algorithms for distributed optimization," *IEEE J. on Selected Areas in Comm.*, vol. 23, no. 4, pp. 798–808, April 2005.

[8] R. Willett, R. Nowak, and U. Mitra, "Estimating inhomogeneous fields using wireless sensor networks," *IEEE J. on Selected Areas in Comm.*, vol. 22, no. 6, pp. 999–1006, Aug. 2004.

[9] D.P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 2nd edition, 2003.

[10] D.P. Bertsekas and J.N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Athena Scientific, 1997.

[11] S. Zenios and Y. Censor, *Parallel Optimization: Theory, Algorithms, and Applications*, Oxford University Press, 1997.

[12] A. Cerpa, J.L. Wong, L. Kuang, M. Potkonjak, and D. Estrin, "Statistical model of lossy links in wireless sensor networks," in *Proc. ACM/IEEE IPSN*, Los Angeles, CA, April 2005.

[13] P. Levis, S. Madden, D. Gay, J. Polastre, R. Szewczyk, A. Woo, E. Brewer, and D. Culler, "The emergence of networking abstractions and techniques in tinyos," in *Proc. USENIX NSDI*, San Francisco, CA, March 2004.

[14] J.A. Bucklew, T.G. Kurtz, and W.A. Sethares, "Weak convergence and local stability properties of fixed step size recursive algorithms," *IEEE Trans. Info. Theory*, vol. 30, no. 3, pp. 966–978, May 1993.

[15] N.L. Biggs, *Algebraic Graph Theory*, Cambridge University Press, 2nd edition, 1993.

[16] C. Godsil and G. Royale, *Algebraic Graph Theory*, Springer, 2001.

[17] L. Xiao, S. Boyd, and S. Lall, "A scheme for asynchronous sensor fusion based on average consensus," in *Proc. IEEE/ACM IPSN*, Los Angeles, CA, April 2005.

[18] V. Saligrama, M. Alanyali, and O. Savas, "Asynchronous distributed detection in sensor networks," (preprint), April 2005.

[19] P.J. Huber, *Robust Stastics*, Wiley, 1981.

[20] W.J.J. Rey, *Introduction to Robust and Quasi-Robust Statistical Methods*, Springer-Verlag, 1983.