Sensor Networks With Random Links: Topology Design for Distributed Consensus

Soummya Kar, *Student Member, IEEE*, and José M. F. Moura, *Fellow, IEEE*

***Abstract*—In a sensor network, in practice, the communication among sensors is subject to: 1) errors that can cause failures of links among sensors at random times; 2) costs; and 3) constraints, such as power, data rate, or communication, since sensors and networks operate under scarce resources. The paper studies the problem of designing the topology, i.e., assigning the probabilities of reliable communication among sensors (or of link failures) to maximize the rate of convergence of average consensus, when the link communication costs are taken into account, and there is an overall communication budget constraint. We model the network as a Bernoulli random topology and establish necessary and sufficient conditions for mean square sense (mss) and almost sure (a.s.) convergence of average consensus when network links fail. In particular, a necessary and sufficient condition is for the algebraic connectivity of the mean graph topology to be strictly positive. With these results, we show that the topology design with random link failures, link communication costs, and a communication cost constraint is a constrained convex optimization problem that can be efficiently solved for large networks by semidefinite programming techniques. Simulations demonstrate that the optimal design improves significantly the convergence speed of the consensus algorithm and can achieve the performance of a non-random network at a fraction of the communication cost.**

***Index Terms*—Consensus, convergence, distributed decision, graph, Laplacian, sensor networks, spectral graph theory, topology.**

I. Introduction

We DESIGN the optimal *topology* of a sensor network, i.e., of its communication configuration, under the following conditions: the optimality criterion is the convergence rate of the average consensus algorithm; the communication channels *fail at random times*; the communication among sensors incurs a *communication cost*; and the network operates under an *overall communication cost constraint*.

The consensus algorithm is an iterative distributed algorithm; we refer the reader to the existing considerable recent literature. We make a few brief comments. Agreement and consensus have been important problems in distributed computing, [1], [2]. The problem of dynamic load balancing for distributed multiprocessors leads to an algorithm that is essentially consensus. Reference [3] gives spectral conditions on the weight matrix of the network graph for its convergence. In the multi-agent and control literature, [4] develops a model for emergent behavior, schooling, and flocking described in [5], [6]. It presents conditions for alignment, i.e., for all agents to agree to a value that lies in the convex hull of the initial conditions. Consensus appears explicitly in [7], [8] that solve the average consensus problem specifically for distributed computation of functions over networks. This work identified the algebraic connectivity of the underlying graph as controlling the convergence rate of the continuous-time average-consensus algorithm. For additional comments and a survey of consensus for multi-agent coordination see [9], [10], and the references there in. Conditions for convergence of iterative distributed algorithms in a very generic framework have actually appeared in early work in [11], [12].

Realistic networks operate under stress: noise and errors cause links to fail at random times; communication among sensors entails a cost; and scarcity of resources constrain the operation of sensors and networks. To model the link failures, we assume a Bernoulli network: 1) at each iteration, a network link is active with some probability; 2) network links may have different link probabilities; 3) links fail or are alive independently of each other; and 4) the link probabilities remain constant across the consensus iterations. Several authors considered the consensus algorithm on Bernoulli like networks. Reference [13] shows that consensus with unreliable links can be cast as a special case of the more general problem of distributed minimization of the sum of convex functions over an erasure network, while [14] optimizes the gossip probabilities for a *given* network topology under the gossip protocol—only two sensors, randomly selected with gossip probability, can communicate at each iteration. The recent paper [15] proves almost sure convergence of the consensus algorithm in random networks using ergodicity of general doubly stochastic matrices. Our almost sure (a.s.) convergence result, which follows from [15], is more restrictive but applies to doubly stochastic matrices with a specific structure, and by exploiting this structure, leads to a simpler convergence condition. Other work on evolving topologies includes [8] that considers continuous time consensus in networks with switching topologies and communication delays, and [16] that studies distributed consensus on a complete graph with identical link probabilities. References [17] and [18] consider the impact of a fixed (nonrandom) topology on the convergence performance of the consensus algorithm. The networks are deterministic but drawn from several classes of graphs, including small-world graphs.

Manuscript received April 6, 2007; revised February 3, 2008. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Aleksandar Dogandzic. This work was supported by the Defense Advanced Research Projects Agency (DARPA) DSO Advanced Computing and Mathematics Program Integrated Sensing and Processing (ISP) Initiative under ARO Grant DAAD19-02-1-0180, by the National Science Foundation (NSF) under Grant ECS-0225449 and Grant CNS-0428404, by the Office of Naval Research under MURI N000140710747, and by an IBM Faculty Award.

The authors are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: soummyak@andrew.cmu.edu; moura@ece.cmu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2008.920143

All the above references do not consider the *design* of the topology of the network, the main concern here. The present paper studies the design of the topology of the sensor network that leads to improved convergence rate of the consensus algorithm when the communication among sensors incurs a cost, and when there is an overall budget constraint that taxes the communication in the network. This contrasts with the work on gossip algorithms in [14], which does not consider communication costs nor a network overall communication cost budget. Reference [19] presents our own preliminary work on this constrained topology design problem.

We consider several versions of the topology design problem.

1) Fixed topology with equal costs: The communication cost is the same for all links. The network communication cost constraint constrains the number of network links. A version of this simplistic topology design problem has a "closed form" solution—the class of non-bipartite Ramanujan graphs; for the precise statement and details on the solution, see [20]–[23].
2) Fixed topology with different costs: In this case, sensors communicate with different costs and there is an overall communication cost constraint. We still place a cap on the maximum number of links and the network should be sparse but connected. This is a difficult combinatorial optimization problem and there is no closed form solution in general.
3) Random topology with different costs: Since the network is random, designing the topology is the problem of distributing the available links among sensors and determining the link probabilities, or the fraction of time, that each link is used. The communication costs are link dependent, there is an infrastructure communication cost constraint, and a maximum number of links. Because the network is random, it makes sense to constrain the (network) *average*, or expected, communication cost per iteration. This version of the problem relaxes the combinatorial fixed topology problem described in item 2) and usually leads to solutions that are better than fixed topologies, especially under medium to low communication cost constraints. This is because a fixed topology forces the use always of the same network, while the random topology can occasionally, with a small nonzero probability, make use of very good, but costly links, still meeting the cost constraint, while accelerating the rate of convergence to consensus.

The paper establishes several spectral graph conditions for problem 3) that guarantee mean square sense and almost sure convergence of the consensus algorithm. Because problem 3) is not convex, we approximate it in two convex steps. We can then solve 3) and design the topology by semidefinite programming techniques. Simulations show that the resulting topology has very good convergence properties.

We outline the paper as follows. Section II summarizes spectral graph theory concepts and formalizes the Bernoulli network model. Sections III, IV, and V derive necessary and sufficient conditions for mean square sense (mss) and a.s. convergence of the state in the consensus algorithm in terms of the spectral properties of the graph topology and present bounds on the convergence rate. Section VI formulates the topology design for the randomized distributed consensus with a communication cost constraint (RCCC) problem, and then a convex constrained approximation to RCCC, to which we apply semidefinite programming (SDP) techniques. Section VII solves numerically this SDP optimization. We show that these designs can improve significantly the convergence rate, for example, by a factor of 3, when compared to geometric networks (networks where sensors communicate with every other sensor within a fixed radius) and that they can achieve practically the performance of a nonrandom network at a fraction, e.g., 50%, of the communication cost per iteration. Section VIII concludes the paper.

II. Random Topology and Distributed Average Consensus

This section discusses the Bernoulli random topology that we adopt to model a sensor network with link failures and recalls basic concepts from graphs and from distributed average consensus.

Random Topology: Bernoulli Model: In a nonrandom topology, the communication channels stay available whenever the sensors need to communicate. However, in typical sensor networks the bandwidth is constrained and the channels are imperfect. The nonrandom topology no longer is a realistic model; an alternative model that makes sense in many applications is the erasure link model: at each iteration, a link between sensors n and l is online or offline with probabilities P_{nl} and $1 - P_{nl}$, respectively. We refer to P_{nl} as the link probability. This models adequately, for example, networks using the ARQ protocol, where, if no acknowledgement packet is received within the protocol time window, the packet is assumed to be dropped or lost, corresponding to an erasure or link failure. Other similar situations include delayed transmissions (where a large delay is equivalent to a loss), or when, occasionally, the channel introduces large signal distortions.

We remark that the erasure link model is an approximation since it implicitly assumes that sensors exchange quantized data. On the other hand, the consensus algorithm involves communication of analog-amplitude data, in which case the effects of additive noise are very relevant. A more appropriate analysis, albeit more complex, considers quantized data, additive noise in the links, and link erasures. This analysis would divert the focus of this paper and is pursued elsewhere, see [24], [25].

We represent the sensor network by a graph $G = (V, \mathcal{E})$ where the sensors are vertices in the vertex set V and the network links or communication channels correspond to the edges in the edge set $\mathcal{E}$. We assume that the graph is simple and connected,[1] and take $|V| = N$ and $|\mathcal{E}| = M$. The set $\mathcal{E}$ collects all the channels that can be established directly among pairs of sensors, i.e., it is the set of *realizable* edges. These channels may fail at random times, but if $(n, l) \notin \mathcal{E}$ then sensors n and l cannot communicate directly—of course, they still communicate by rerouting their messages through one of the paths connecting them in G, since G is connected. We call G the *supergraph*.

[1]A graph is called simple if it is devoid of loops (self-edges) and multiple edges. It is connected if every vertex can be reached from any other vertex, which in network terms may require a routing protocol.

To model the random link failures, we take the state (online or offline) of each link $(n,l) \in \mathcal{E}$ to be a Bernoulli process with link probability, i.e., of the edge being online, to be P_{nl}; see also [14], [15], [19]. We assume that for distinct pairs of realizable edges (n,l), (m,k) the corresponding Bernoulli processes are statistically independent. Under this model, at each time i, the sensor network topology is a random graph $G(i) = (V, E(i))$, with $E(i) \subseteq \mathcal{E}$ a random subset. For example, the neighborhood, $\Omega_n(i)$, of vertex n in $G(i)$ that defines the network topology is

$$\forall 1 \leq n \leq N : \Omega_n(i) = \{l \in V : (n,l) \in E(i)\}. \quad (1)$$

The cardinality $d_n(i) = |\Omega_n(i)|$ is the node degree and is the number of edges online at time i and connected to vertex n. To the supergraph G and each random graph $G(i)$, we associate standard $N \times N$ matrices: adjacency matrices $\mathcal{A}$ and $A(i)$ ($A_{nl}(i) = 1$ if $(n,l) \in E(i)$), diagonal degree matrices D and $D(i) = \operatorname{diag}[d_1(i) \ldots d_N(i)]$, graph Laplacians $\mathcal{L}$ and $L(i) = D(i) - A(i)$, and weight matrices $\mathcal{W}$ and $W(i)$. Of particular interest is when the weights are all equal to α

$$W(i) = I - \alpha L(i) \quad (2)$$

For the Bernoulli random topology, the matrices in the sets $\{A(i)\}$, $\{L(i)\}$, and $\{W(i)\}$ are statistically independent, identically distributed (i.i.d.) random matrices. Note that $\mathbf{0} \preceq A(i) \preceq \mathcal{A}$, where $\mathbf{0}$ is the $N \times N$ zero matrix and $C \preceq D$ stands for $\forall 1 \leq i,j \leq N : C_{i,j} \leq D_{i,j}$. The probability of an instantiation $E(i)$, $A(i)$, $L(i)$, and $W(i)$ is $\Pi_{(n,l)\in\mathcal{E}} P_{nl}$. Define the link probability matrix

$$P = P^T = [P_{nl}], \quad 0 \leq P_{nl} \leq 1, \quad P_{nn} = 0.$$

The diagonal elements are zero because the graph is simple (no loops). The structure of P reflects the structure of the adjacency matrix $\mathcal{A}$ of the superset $\mathcal{E}$, i.e., $P_{nl} \neq 0$ if and only if $\mathcal{A}_{nl} = 1$. The matrix P is not stochastic since their row or column sums are not normalized to 1. Abusing notation, we will refer to P as the probability distribution of the $E(i)$, and, likewise, of $A(i)$, $L(i)$, and $W(i)$.

The matrix means are $\overline{A} = P$, $\overline{W} = I - \alpha\overline{L}$ where

$$\overline{L}_{nl} = \begin{cases} \sum_{m=1}^{N} P_{nm}, & \text{if } n = l \\ -P_{nl}, & \text{otherwise.} \end{cases} \quad (3)$$

Average Consensus: Average consensus, see [8] computes by a distributed algorithm the average of $x_n(0)$, $n = 1, \ldots, N$ where $x_n(0)$ is available at sensor n at time 0. At time i, each node exchanges its state $x_n(i)$, $i = 0, 1, \ldots$ synchronously with its neighbors specified by the graph edge neighborhood set, $\Omega_n(i)$. In vector form, the N states $x_n(i)$ are collected in the state vector $\mathbf{x}(i) \in \mathbb{R}^{N\times 1}$. Define the vector $\mathbf{1}$ of ones and the matrix $J = \mathbf{1}\mathbf{1}^T$. The vector of averages $\mathbf{x}_{\text{avg}}$ is written

$$\mathbf{x}_{\text{avg}} = \frac{\mathbf{1}^T \mathbf{x}(0)}{N}\mathbf{1} = \frac{1}{N} J \mathbf{x}(0). \quad (4)$$

For the sequence of random topologies, $\{G(i)\}_{i=0}^{\infty}$, state update by average consensus proceeds according to the iterative algorithm

$$\forall i \geq 0 : x_n(i+1) = W_{nn}(i) x_n(i) + \sum_{l \in \Omega_n(i)} W_{nl}(i) x_l(i) \quad (5)$$

$$\mathbf{x}(i+1) = W(i)\mathbf{x}(i) \quad (6)$$

where $W(i) = [W_{nl}(i)]$ is the matrix of weights. The sparsity of $W(i)$ is determined by the underlying network connectivity at time i, i.e., for $n \neq l$, the weight $W_{nl}(i) = 0$ if $(n,l) \notin E(i)$. Iterating (6),

$$\mathbf{x}(i) = \left(\prod_{j=0}^{i-1} W(j)\right) \mathbf{x}(0). \quad (7)$$

The state $\mathbf{x}(i)$ is random. Section IV analyzes the impact of the topology on the convergence of (7).

III. Preliminary Results

The section considers properties of the Laplacian and weight matrices, needed in Sections IV through VI when studying the random topology and random topology with communication cost constraint problems; for additional graph theory concepts see [26]–[28]. We will often drop the iteration index i; the context should make it clear when the quantities are iteration dependent.

The Laplacian $L(i)$ is a symmetric positive, semidefinite matrix with eigenvalues

$$0 = \lambda_1(L(i)) \leq \lambda_2(L(i)) \leq \cdots \leq \lambda_N(L(i)). \quad (8)$$

The normalized eigenvector $\mathbf{u}_1(L(i))$ corresponding to $\lambda_1(L(i))$ is

$$\mathbf{u}_1(L(i)) = \frac{1}{\sqrt{N}}\mathbf{1}. \quad (9)$$

The multiplicity of $\lambda_1(L(i))$ is the number of connected components of the graph. If $G(i)$ is connected, $\lambda_2(L(i)) > 0$, often referred to as the algebraic connectivity (or Fiedler value), see [29].

Lemma 1: Let the mean Laplacian be the weighted Laplacian for a graph $\overline{G}$

$$\lambda_2(\overline{L}) > 0 \Longleftrightarrow \overline{L} \text{ is irreducible} \Longleftrightarrow \overline{G} \text{ is connected.} \quad (10)$$

Proof: The proof is simple and follows from [30]. ■

The convergence results in Section IV involve the mean $\mathrm{E}[\lambda_2(L)]$, which is manifestly difficult to compute and manipulate. A much easier quantity to compute is $\lambda_2(\overline{L})$. By Jensen's inequality and recalling that $\lambda_2(L)$ is a concave function of L (see [31]), we have the following.

Lemma 2: $\mathrm{E}[\lambda_2(L)] \leq \lambda_2(\overline{L})$.

Let $\rho(X)$ be the spectral norm of the matrix X. For symmetric matrices the spectral radius is equal to the matrix 2-norm.

Lemma 3: For a given P, $\mathrm{E}[\rho(W-(1/N)J)]$ is convex on α.

For a given α

$$\mathrm{E}\left[\rho\left(W-\frac{1}{N}J\right)\right] \geq \rho\left(\overline{W}-\frac{1}{N}J\right). \tag{11}$$

Proof: To prove the first statement, note that, for a given L, $\rho(W-(1/N)J)$ is a convex function of α. In fact, let $\alpha_1, \alpha_2 \in \mathbb{R}$ and $0 \leq t \leq 1$. We then have

$$\begin{aligned}\rho&\left(I-(t\alpha_1+(1-t)\alpha_2)\,L-\frac{1}{N}J\right)\\ &=\left\|t\left(I-\alpha_1 L-\frac{1}{N}J\right)+(1-t)\left(I-\alpha_2 L-\frac{1}{N}J\right)\right\|_2\\ &\leq\left\|t\left(I-\alpha_1 L-\frac{1}{N}J\right)\right\|_2+\left\|(1-t)\left(I-\alpha_2 L-\frac{1}{N}J\right)\right\|_2\\ &=t\rho\left(I-\alpha_1 L-\frac{1}{N}J\right)+(1-t)\rho\left(I-\alpha_2 L-\frac{1}{N}J\right).\end{aligned} \tag{12}$$

Convexity and the first statement follow by taking expectation on both sides of (12).

The second statement is similarly proved and follows from Jensen's inequality. ∎

Lemma 3 and (12) provide an easily computable lower bound on $\mathrm{E}[\rho(W-(1/N)J)]$.

IV. Convergence of Average Consensus: Random Topology

We study convergence of the state in average consensus for random topologies

$$\forall \mathbf{x}(0) \in \mathbb{R}^{N\times 1}: \lim_{i\to\infty} \mathbf{x}(i) = \mathbf{x}_{\text{avg}} \tag{13}$$

in some appropriate probabilistic sense. Let $\|\cdot\|$ be the $\mathcal{L}_2$-norm. We start by giving next two equivalent necessary and sufficient conditions for convergence of the mean vector

$$\lim_{i\to\infty} \|\mathrm{E}\mathbf{x}(i) - \mathbf{x}_{\text{avg}}\| = 0. \tag{14}$$

Theorem 4: The mean converges iff either one of the following holds

1) $\rho(\overline{W}-(1/N)J) < 1$.
2) $\lambda_2(\overline{L}) > 0$ and $0 < \alpha < 2/\lambda_N(\overline{L})$.

If the mean converges, the choice of α that minimizes $\rho(\overline{W}-(1/N)J)$ is

$$\alpha^\star = \frac{2}{\lambda_2(\overline{L})+\lambda_N(\overline{L})}. \tag{15}$$

Proof: We sketch the first part of the proof. The matrix J satisfies

$$\left(\frac{1}{N}J\right)^n = \frac{1}{N}J, \quad n=1,2,\ldots.$$

Using the i.i.d. properties of $W(i)$, we obtain from (7)

$$\mathrm{E}[\mathbf{x}(i)] - \mathbf{x}_{\text{avg}} = \left(\overline{W}-\frac{1}{N}J\right)^i \mathbf{x}(0). \tag{16}$$

Convergence is now equivalent to condition 1. For the equivalence of 1. and 2. in Theorem 4 note that

$$\rho\left(\overline{W}-\frac{1}{N}J\right) = \max\left(1-\alpha\lambda_2(\overline{L}), 1-\alpha\lambda_N(\overline{L})\right). \tag{17}$$

The theorem also follows from a straightforward generalization to non-binary Laplacian matrices of the convergence results in deterministic consensus in [32]. ∎

We now consider convergence in the mean-square-sense (mss) in Section IV-A and almost sure convergence (convergence with probability 1) in Section IV-B.

A. Mean Square Convergence

This section studies mean-square convergence

$$\forall \mathbf{x}(0) \in \mathbb{R}^{N\times 1}: \lim_{i\to\infty} \mathrm{E}\|\mathbf{x}(i)-\mathbf{x}_{\text{avg}}\| = 0 \tag{18}$$

which implies convergence of the mean, but not the reverse.

Theorem 5: If $\mathrm{E}[\rho(W-(1/N)J)] < 1$, the state vector sequence $\{\mathbf{x}(i)\}_{i=0}^{\infty}$ converges in mss.

Proof: From a result in [19], stated here without proof, for any $\mathbf{x}(0) \in \mathbb{R}^{N\times 1}$

$$\|\mathbf{x}(i+1)-\mathbf{x}_{\text{avg}}\| \leq \left(\prod_{j=0}^{i} \rho\left(W(j)-\frac{1}{N}J\right)\right)\|\mathbf{x}(0)-\mathbf{x}_{\text{avg}}\|. \tag{19}$$

Taking expectation on both sides of (19) and using the i.i.d. properties of the $W(j)$'s

$$\mathrm{E}\|\mathbf{x}_i-\mathbf{x}_{\text{avg}}\| \leq \left(\mathrm{E}\left[\rho\left(W-\frac{1}{N}J\right)\right]\right)^{i-1}\|\mathbf{x}_0-\mathbf{x}_{\text{avg}}\|. \tag{20}$$

The theorem follows. We dropped the index i in $W(i)$ in (20). ∎

Theorem 5 shows that the smaller $\mathrm{E}[\rho(W-(1/N)J)]$ is, the faster the mss convergence is. The value of $\mathrm{E}[\rho(W-(1/N)J)]$ depends both on the probability distribution of the Laplacian $L(i)$ and on the constant weight α. However, the probability distribution of $L(i)$ must satisfy certain conditions to guarantee that there are values of α that lead to mss convergence. Otherwise, no choice of α will result in mss convergence. The next theorem considers this issue. Before stating the theorem, let $d_{\max}$ be the maximum degree of the graph with edge set $E=\mathcal{E}$ and define

$$\alpha_{\text{mss}} = \frac{1}{2d_{\max}}. \tag{21}$$

Theorem 6: There is an α such that the consensus algorithm converges in mss iff $\lambda_2(\overline{L}) > 0$. In other words, if $\lambda_2(\overline{L}) > 0$, we can find an α, in particular, $\alpha = \alpha_{\text{mss}}$ defined in (21), that leads to mss convergence. If $\lambda_2(\overline{L}) = 0$, no choice of α will result in mss convergence.

Proof: The proof of sufficiency is constructive. We show that, if $\lambda_2(\overline{L}) > 0$, we can find an α for which $\mathrm{E}[\rho(W - (J/N))] < 1$. Convergence then follows from Theorem 5.

Let $\lambda_2(\overline{L}) > 0$. By Lemma 1, $\overline{L}$ is irreducible. From the irreducibility of $\overline{L}$, with non-zero probability, we have graph realizations, G, for which $L(G)$ is irreducible and so $\lambda_2(L(G)) > 0$. In particular, with non-zero probability, we can have a realization for which the edge set $E = \mathcal{E}$; by assumption, this network is irreducible and hence connected (because the corresponding Laplacian matrix has the same sparsity pattern of $\overline{L}$ with non-zero entries of $\overline{L}$ replaced by ones.) Hence, with non-zero probability, $\lambda_2(L) > 0$, which makes $\mathrm{E}[\lambda_2(L)] > 0$. Thus, we have

$$\lambda_2(\overline{L}) > 0 \Longrightarrow \mathrm{E}\left[\lambda_2(L)\right] > 0. \tag{22}$$

Also, it follows from spectral graph theory (see [30]), that, for any realizable graph G,

$$\lambda_N\left(L(G)\right) \leq 2d_{\max}. \tag{23}$$

We now claim mss convergence for $\alpha = \alpha_{\mathrm{mss}}$. We note that

$$\begin{aligned}\rho\left(W(i) - \frac{1}{N}J\right) &= \max\left(\lambda_2\left(W(i)\right), -\lambda_N\left(W(i)\right)\right)\\ &= \max\left(1 - \alpha_{\mathrm{mss}}\lambda_2\left(L(i)\right),\right.\\ &\qquad \left.\alpha_{\mathrm{mss}}\lambda_N\left(L(i)\right) - 1\right)\\ &= 1 - \alpha_{\mathrm{mss}}\lambda_2\left(L(i)\right)\end{aligned} \tag{24}$$

where the last step follows because, from (23) and (21),

$$1 - \alpha_{\mathrm{mss}}\lambda_2\left(L(i)\right) \geq 0 \geq \alpha_{\mathrm{mss}}\lambda_N\left(L(i)\right) - 1. \tag{25}$$

Taking expectation on both sides of (24), and since $0 < \mathrm{E}[\lambda_2(L)] \leq 2d_{\max}$, we get

$$\begin{aligned}\mathrm{E}\left[\rho\left(W - \frac{1}{N}J\right)\right] &= 1 - \alpha_{\mathrm{mss}}\mathrm{E}\left[\lambda_2(L)\right]\\ &< 1\end{aligned} \tag{26}$$

and mss convergence then follows from Theorem 4. This proves sufficiency.

Necessity follows from the fact that, if $\lambda_2(\overline{L}) = 0$, Theorem 4 precludes convergence of the mean vector. Since, by Corollary 9 (see below), convergence of the mean is necessary for mss convergence, we conclude that, if $\lambda_2(\overline{L}) = 0$, no choice of α will result in mss convergence. ■

Theorem 6 gives necessary and sufficient conditions on the probability distribution of the Laplacian L, i.e., on $\overline{L}$ or P, for mean square convergence. This is significant as it relates mss convergence to the network topology. Because this condition is in terms of the algebraic connectivity of the mean Laplacian associated with the link probability distribution P, it is straightforward to check.

B. Almost Sure Convergence

We show that $\lambda_2(\overline{L}) > 0$ is also a necessary and sufficient condition for a.s. convergence of the sequence $\{\mathbf{x}(i)\}_{i=0}^{\infty}$. Before proceeding, we state the definition of a.s. convergence of (scalar) random variables.

Definition 7 (A.S. Convergence of Random Variables): Let $\{\boldsymbol{\xi}_i\}_{i=0}^{\infty}$ be a sequence of random variables defined on some common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $\{\boldsymbol{\xi}_i\}_{i=0}^{\infty}$ converges a.s. to another random variable $\boldsymbol{\xi}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ ($\boldsymbol{\xi}_i \rightarrow \boldsymbol{\xi}$ a.s.) iff

$$\mathbb{P}\left(\omega \in \Omega : \boldsymbol{\xi}_i(\omega) \underset{i\rightarrow\infty}{\longrightarrow} \boldsymbol{\xi}(\omega)\right) = 1. \tag{27}$$

For random vectors a.s. convergence means a.s. convergence of each component. We recall that mss convergence of a sequence of random variables $\{\mathbf{x}(i)\}_{i=0}^{\infty}$ implies convergence in probability through Chebyshev's inequality. Also, convergence in probability implies a.s. convergence of a subsequence (see [33].)

Theorem 8: The sequence $\{\mathbf{x}(i)\}_{i=0}^{\infty}$ converges a.s. iff $\lambda_2(\overline{L}) > 0$. In particular, if $\lambda_2(\overline{L}) > 0$, then for

$$\alpha = \alpha_{\mathrm{mss}} = \frac{1}{2d_{\max}}$$

$\mathbf{x}(i) \rightarrow \mathbf{x}_{\mathrm{avg}}$ a.s. On the contrary, if $\lambda_2(\overline{L}) = 0$, then no choice of α leads to a.s. convergence.

Proof: We consider sufficiency first. Since the $W(i)$'s in this paper satisfy the conditions $\mathbf{1}^T W(i) = \mathbf{1}^T$ and $W(i)\mathbf{1} = \mathbf{1}$, the assumptions of Theorem 6 in [15] are met and hence Theorem 8 falls under the purview of Theorem 6 in [15]. However, we provide a brief constructive proof here, which shows that $\alpha = \alpha_{\mathrm{mss}}$ leads to a.s. convergence if $\lambda_2(\overline{L}) > 0$. From Theorem 6 above, the sequence

$$\boldsymbol{\xi}_i = \left\|\mathbf{x}(i) - \mathbf{x}_{\mathrm{avg}}\right\|^{1/2} \overset{mss}{\longrightarrow} 0. \tag{28}$$

Thus, $\boldsymbol{\xi}_i \rightarrow 0$ in probability and there exists a subsequence $\{\boldsymbol{\xi}_{i_k}\}_{k=0}^{\infty}$ that converges to 0 a.s. Further, from (23) and (24), and using $\alpha = \alpha_{\mathrm{mss}}$,

$$\rho\left(W(i) - \frac{1}{N}J\right) \leq 1, \ \forall i \geq 0. \tag{29}$$

In a similar vein to (19) and using (29)

$$\begin{aligned}\boldsymbol{\xi}_i^2 &\leq \rho\left(W(i-1) - \frac{1}{N}J\right)\boldsymbol{\xi}_{i-1}^2\\ &\leq \boldsymbol{\xi}_{i-1}^2.\end{aligned} \tag{30}$$

Thus, $\{\boldsymbol{\xi}_i\}_{i=0}^{\infty}$ is a non-increasing sequence of random variables, a subsequence of which converges a.s. to 0. Clearly $\boldsymbol{\xi}_i \rightarrow 0$ a.s. and sufficiency follows.

Necessity uses the argument given in Theorem 6. ■

We comment on Theorems 6 and 8. In this paper, we consider only equal link weights, i.e., all the links are assigned the same weight α, see (2). However, whatever the weights are, in particular, different weights for different links, a necessary condition for mss convergence (and a.s. convergence) is $\lambda_2(\overline{L}) > 0$. This is because, if $\lambda_2(\overline{L}) = 0$, the network separates into two components with zero probability of communication between each other. Hence, no weight assignment can lead to mss convergence. Thus, the necessary condition established in Theorems 6 and 8 for mss convergence and a.s. convergence, respectively, in the constant link weight case also holds for the more general weight assignments.

V. MSS Convergence Rate

We study how fast the consensus algorithm can converge. We focus on mss convergence. A first result follows from Theorem 4 and Lemma 3.

Corollary 9: mss convergence cannot be faster than convergence of the mean vector.

To optimize the mss convergence rate, we note that, given a particular distribution P of the Laplacian L, hence $\overline{L}$, the actual choice of α plays a significant role in determining the convergence rate. To optimize the rate for a given P, or $\overline{L}$, we perform the minimization

$$\min_{\alpha} \mathrm{E}\left[\rho\left(W - \frac{1}{N}J\right)\right]. \tag{31}$$

We present the results in terms of the mss convergence rate formally defined next.

Definition 10 (mss Convergence Rate): Let $\lambda_2(\overline{L}) > 0$. If for a given α, $\mathrm{E}[\rho(W-(1/N)J)] < 1$, we call

$$S_{\mathrm{g}}(\alpha, \overline{L}) = \ln\left(\frac{1}{\mathrm{E}\left[\rho\left(W - \frac{1}{N}J\right)\right]}\right) \tag{32}$$

the mss convergence gain per iteration or the mss convergence rate of the consensus algorithm.

Because the ln is monotonic, the best achievable mss convergence rate for a given P or $\overline{L}$ is

$$\begin{aligned} S_{\mathrm{g}}^*(\overline{L}) &= \max_{\alpha} S_{\mathrm{g}}(\alpha, \overline{L}) \\ &= -\ln\left(\min_{\alpha} \mathrm{E}\left[\rho\left(W - \frac{1}{N}J\right)\right]\right). \end{aligned} \tag{33}$$

There is, in general, no closed form solution for the minimization in (31). It depends on the probability distribution P of the Laplacian L. By Lemma 3, $\mathrm{E}[\rho(W-(1/N)J)]$ is convex in α for a given P or $\overline{L}$; so, its minimum is attainable by numerical procedures. Denote the minimizing α by α^*. The following Lemma provides a range where the optimal α lies.

Lemma 11: Let $\lambda_2(\overline{L}) > 0$. Then

$$0 < \alpha^* < \frac{2}{\lambda_N(\overline{L})}. \tag{34}$$

Proof: We note that, a necessary condition for mss convergence is convergence of the mean vector. For a proof, let us assume, that, we have mss convergence, i.e.,

$$\lim_{i\to\infty} \mathrm{E}\left\|\mathbf{x}(i) - \mathbf{x}_{\mathrm{avg}}\right\| = 0. \tag{35}$$

Also, we have from Jensen's inequality

$$\left\|\mathrm{E}\mathbf{x}(i) - \mathbf{x}_{\mathrm{avg}}\right\| \le \mathrm{E}\left\|\mathbf{x}(i) - \mathbf{x}_{\mathrm{avg}}\right\|, \forall i \tag{36}$$

since, the $\mathcal{L}_2$-norm is a convex function on $\mathbb{R}^{N\times 1}$. It then follows from ((35), (36)), that

$$\lim_{i\to\infty} \left\|\mathrm{E}\mathbf{x}(i) - \mathbf{x}_{\mathrm{avg}}\right\| = 0 \tag{37}$$

thus, implying convergence of the mean vectors.

Hence, from Theorem 4, α^* leading to fastest mss convergence must belong to the range in (34). ∎

We can bound the optimal mss convergence rate $S_{\mathrm{g}}^*(\overline{L})$.

Lemma 12: If $\lambda_2(\overline{L}) > 0$, then

$$S_{\mathrm{g}}^*(\overline{L}) \ge \ln\left(\frac{1}{1 - \alpha_{\mathrm{mss}}\mathrm{E}\left[\lambda_2(L)\right]}\right). \tag{38}$$

Proof: By Theorem 6, if $\lambda_2(\overline{L}) > 0$, then $\alpha = \alpha_{\mathrm{mss}}$ leads to mss convergence and

$$\begin{aligned} S_{\mathrm{g}}^*(\overline{L}) &= \ln\left(\frac{1}{\min_{\alpha} \mathrm{E}\left[\rho\left(W - \frac{1}{N}J\right)\right]}\right) && (39) \\ &\ge \ln\left(\frac{1}{\mathrm{E}\left[\rho\left(I - \alpha_{\mathrm{mss}}L - \frac{1}{N}J\right)\right]}\right) && (40) \\ &= \ln\left(\frac{1}{1 - \alpha_{\mathrm{mss}}\mathrm{E}\left[\lambda_2(L)\right]}\right). && (41) \end{aligned}$$

∎

VI. Consensus With Communication Constraints: Topology Optimization

The previous sections analyzed the impact of the link probability distribution P on the convergence rate of the distributed consensus algorithm. This section studies the design of the sensor network topology that leads to the fastest rate of convergence but when there is a cost C_{nl} for the sensors n and l to communicate and there is an overall infrastructure communication cost constraint U. We collect the costs into a symmetric cost matrix C. We assume equal link weights throughout. The constraint U captures the common fact that networks usually have an overall *average* power budget to operate with.

A. Random Topology With Communication Cost Constraints (RCCC)

We restate the RCCC problem presented in Section I. The network has N sensors and follows the Bernoulli random model introduced in Section II. The communication costs between any pair of sensors at each iteration is specified by the $N \times N$ matrix $C = C^T$—entry $C_{nl} \ge 0, n \ne l$, is the cost incurred by a single communication between nodes n and l. If $C_{nl} = +\infty$, sensors n and l do not communicate directly, only through other sensors by a routing protocol. The total cost incurred at stage i is

$$\begin{aligned} u(i) &= -\frac{1}{2}\sum_{n\ne l} L_{nl}(i) C_{nl} && (42) \\ &= -\frac{1}{2}\mathrm{Tr}\left(CL(i)\right). && (43) \end{aligned}$$

Equation (43) follows because C is symmetric with zero diagonal entries. Equation (42) implicitly assumes that, if $C_{nl} = +\infty, \forall i : L_{nl}(i) = 0$; in other words, the edge (n, l) is not in the superset $\mathcal{E}$.

Let P be the link probability matrix. The diagonal entries of P are zero, although each node accesses its data with zero cost. The P matrix induces a probability distribution on the Laplacian $L(i)$. Since $L(i)$ is random, the cost u_i at step i is random and is between 0 (all links fail) and the total cost when all links with non-zero P_{nl} are active. From (43), the expected cost incurred at step i is

$$\forall i : \mathrm{E}\left[u_i\right] = -\frac{1}{2}\mathrm{Tr}(C\overline{L}). \tag{44}$$

Let $\mathcal{D}(U)$ be the set of feasible $\overline{L}$ (and hence P) given the expected network cost constraint per step U

$$\mathcal{D}(U) = \left\{ \overline{L} : -\frac{1}{2}\mathrm{Tr}(C\overline{L}) \leq U \right\}. \tag{45}$$

The optimal RCCC topology is obtained by solving the following optimization problem:

RCCC: Problem formulation.

$$\begin{aligned} &\max_{\alpha,\overline{L}} S_{\mathrm{g}}(\alpha,\overline{L}) \\ &\text{subject to} \quad \overline{L} = \overline{L}^T \in \mathbb{R}^{N\times N} - 1 \leq \overline{L}_{nl} \\ &\qquad\qquad \leq 0,\ n,l \in \{1,\ldots,N\},\ n \neq l \\ &\qquad \overline{L}\mathbf{1} = \mathbf{0} \\ &-\frac{1}{2}\mathrm{Tr}(C\overline{L}) \leq U \\ &\qquad \overline{L}_{nl} = 0 \quad \text{if } C_{nl} = +\infty. \end{aligned} \tag{46}$$

The second inequality constraint comes from the fact that $\overline{L}_{nl} = -P_{nl}$, $n \neq l$. The other conditions follow from the properties of the Laplacian and the cost constraint. We note that the RCCC optimal solution is a function of U; we denote it by

$$S_{\mathrm{g}}^{**}(U) = \max_{\alpha,\overline{L}\in\mathcal{D}(U)} S_{\mathrm{g}}\left(\alpha,\overline{L}\right). \tag{47}$$

B. Alternate Randomized Consensus Under Communication Cost Constraints (ARCCC)

The RCCC problem in (46) is difficult because it is nonconvex. Its objective function is the expected value of the spectral norm of a matrix. This expectation is with respect to the probability distribution of the Laplacian, which is the product of the entries of the link probability matrix P. Absorbing α in $\overline{L}$ does not make the RCCC objective function $S_{\mathrm{g}}(\alpha,\overline{L})$ convex on this distribution P of L, as can be easily seen by working a low order example. A second similar concern is that when α is absorbed into $\overline{L}$, the interpretation of the entries of $\overline{L}$ as probabilities is lost, and it is not possible to incorporate the cost constraint, which explicitly involves the probabilities.

We consider a suboptimal version, the alternate randomized consensus under communication cost constraint (ARCCC) problem. ARCCC decouples the optimization in $\overline{L}$ from the optimization in α and, when optimizing over $\overline{L}$, replaces the convergence rate $S_{\mathrm{g}}(\alpha,\overline{L})$ by the algebraic connectivity $\lambda_2(\overline{L})$. We show: (i) ARCCC is convex and can be solved by fast numerical optimization procedures; (ii) ARCCC is a good approximation to (46); and (iii) ARCCC leads to topologies with good convergence rates. Point (i) is next; points (ii) and (iii) are in Section VIII where we analyze the performance of ARCCC. But first, we present ARCCC formally.

ARCCC: Problem formulation.

$$\begin{aligned} \textbf{Step I}: \quad & \overline{L}^* = \arg\max_{\overline{L}} \lambda_2(\overline{L}) \\ \text{subject to} \quad & \overline{L} = \overline{L}^T \in \mathbb{R}^{N\times N} -1 \leq \overline{L}_{nl} \\ & \leq 0,\ n,l \in \{1,\ldots,N\},\ n \neq l \\ & \overline{L}\mathbf{1} = \mathbf{0} - \frac{1}{2}\mathrm{Tr}(C\overline{L}) \\ & \leq U \\ & \overline{L}_{nl} = 0 \quad \text{if} \quad C_{nl} = +\infty \end{aligned} \tag{48}$$

$$\textbf{Step II}: \qquad \max_{\alpha} S_g(\alpha,\overline{L}^*). \tag{49}$$

We show that ARCCC is a convex optimization problem.

Lemma 13: Each step in ARCCC is a convex optimization problem.

Proof: We consider (48). The objective $\lambda_2(\overline{L})$ is concave on $\overline{L}$. The set of $\overline{L}$ satisfying the constraints is convex. Concavity of (49) is equivalent to the minimization in Lemma 3, which is convex. ■

The optimization problem in Lemma 13 is a semidefinite programming (SDP) problem that can be solved numerically in efficient ways, see [34], [35] for SDP solving methods (see also [31], [36] on the problem of constrained optimization of graph Laplacian eigenvalues.)

In both the RCCC and ARCCC problems, the constraint is on the *average* network communication cost. This is appropriate in many practical situations but, as a consequence, the instantaneous communication cost per iteration in the ARCCC design will exceed from time to time the average communication constraint U. Applications where constraining the instantaneous costs is required will lead, in general, to integer constraints and to a much harder optimization problem.

VII. Topology Optimization: Performance Results

In this section, Section VII-A discusses in what sense the ARCCC problem presented in Section VI-B and (48) and (49) approximates well the RCCC problem described in Section VI-A and (46). Section VII-B establishes bounds on the optimal value of $\lambda_2(\overline{L})$ as a function of the communication budget constraint U. Finally, Section VII-C illustrates by simulation the fast rate of convergence of distributed consensus on the ARCCC derived topology.

A. ARCCC as a Good Approximation to RCCC

As noted, in RCCC, the joint optimization over α and $\overline{L}$ is not convex. This leads to an alternative formulation, the ARCCC topology optimization problem. We argue here why this is plausible and then present numerical results that confirm that ARCCC's are good performing topologies.

We arrive at ARCCC by successively approximating RCCC by the following steps.

1) We would like to decouple the joint nonconvex optimization over $\overline{L}$ and α into two convex optimization steps. This will in general result in a loss of optimality, as discussed below.
2) When optimizing over $\overline{L}$, start by replacing the RCCC convergence rate functional $S_g(\alpha,\overline{L})$ by the average algebraic connectivity $\mathrm{E}[\lambda_2(L)]$. To justify this, we note that (41) bounds $S_{\mathrm{g}}^*(\overline{L})$ from below, which suggests that larger values of $\mathrm{E}[\lambda_2(L)]$ lead to higher values of $S_{\mathrm{g}}^*(\overline{L})$. This suggests that, intuitively, $\mathrm{E}[\lambda_2(L)]$ and $S_{\mathrm{g}}^*(\overline{L})$ are monotonically related and the orderings they induce on the set of distributions $\mathcal{D}(U)$ are equivalent. This is strictly not true, but the numerical experiments below do confirm it as a general trend. Hence, it is to be expected that a topology that maximizes the quantity $\mathrm{E}[\lambda_2(L)]$ over the set $\mathcal{D}(U)$ will exhibit fast mss convergence rate, while satisfying the communication constraint U.

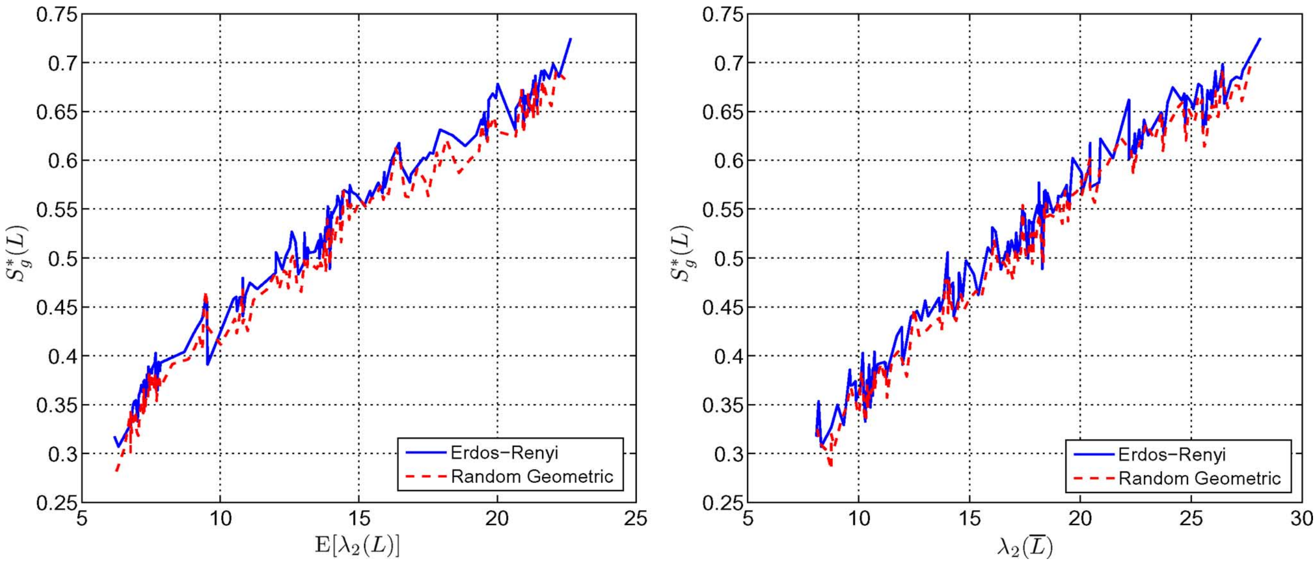


Fig. 1. Convergence rate $S_{\mathrm{g}}^*(\overline{L})$. Left: with varying $\mathrm{E}[\lambda_2(L)]$. Right: with varying $\lambda_2(\overline{L})$. The number of vertices is $N = 500$.

3) Replacing $S_g(\alpha, \overline{L})$ by $\mathrm{E}[\lambda_2(L)]$ does not lead, still, to a numerically simple topology optimization problem, since computing $\mathrm{E}[\lambda_2(L)]$ requires costly Monte Carlo simulations (see [19].) So, in ARCCC, the optimization of $\mathrm{E}[\lambda_2(L)]$ is replaced by maximization of $\lambda_2(\overline{L})$, which simply involves computing the second eigenvalue of $P = \overline{L}$, no Monte Carlo simulations being involved. This approximation is justified on the basis of Lemma 2, which upper-bounds $\mathrm{E}[\lambda_2(L)]$ by $\lambda_2(\overline{L})$ suggesting that, for $\mathrm{E}[\lambda_2(L)]$ to be large, $\lambda_2(\overline{L})$ should be large.

Putting together these steps, the RCCC problem in (46) is successively approximated by

$$\begin{aligned} S_{\mathrm{g}}^{**}(U) &= \max_{\alpha, \overline{L} \in \mathcal{D}(U)} S_{\mathrm{g}}(\alpha, \overline{L}) \\ &\approx \max_{\alpha} S_{\mathrm{g}}(\alpha, \overline{L}^*(U)) \\ &= \widehat{S_{\mathrm{g}}^{**}(U)} \end{aligned} \tag{50}$$

where $\overline{L}^*(U)$ is given by

$$\overline{L}^*(U) = \arg \max_{\overline{L} \in \mathcal{D}(U)} \lambda_2(\overline{L}). \tag{51}$$

In general, $\widehat{S_{\mathrm{g}}^{**}(U)} \leq S_{\mathrm{g}}^{**}(U)$. If $S_g^*(L)$ was a non-decreasing function of $\lambda_2(\overline{L})$, we would have $\widehat{S_{\mathrm{g}}^{**}(U)} = S_{\mathrm{g}}^{**}(U)$. We also note here, that, once we obtain $\overline{L}^*(U)$ from ARCCC, the optimal α is computed by minimizing $\mathrm{E}\rho[W - (1/N)J]$ with respect to α, as given in (31). This is a scalar convex minimization problem and can be done very efficiently using standard line search techniques, because the range of the optimal α is bounded by Lemma 11.

We verify by two sets of numerical studies how and in what sense $S_{\mathrm{g}}^*(\overline{L})$ increases with $\mathrm{E}[\lambda_2(L)]$ and $\lambda_2(\overline{L})$. For the first set of simulations, we choose a network with $N = 500$ sensors and let the average degree d_{avg} of the network vary in steps of 5 from 10 to 40. For each of these 7 values of d_{avg}, we construct 200 Erdös–Rényi random graphs by choosing at random $M = d_{\mathrm{avg}}N/2$ edges of the $N(N-1)/2$ possible pairings of vertices in the network. Each of these graphs fixes the structure of the superset $\mathcal{E}$. For each of these 200 superset (random) graphs, we generate randomly a link probability matrix P (hence a probability distribution of L, and so an $\overline{L}$) by choosing for each edge a number between 0 and 1 drawn from a uniform random distribution. With each such P matrix, we generate 400 random instantiations of the Laplacian $L(i)$, $i = 1, \ldots, 400$ and average the corresponding values of $\rho(W(i) - (1/N)J)$ and $\lambda_2(L(i))$ to estimate the convergence rate $S_{\mathrm{g}}^*(\overline{L})$ and the mean algebraic connectivity $\mathrm{E}[\lambda_2(L)]$. For each P, we obtain the corresponding $\lambda_2(\overline{L})$ by (3). In total, the simulation generated 560,000 Erdös–Rényi graphs of 500 vertices with the number of edges ranging from 10,000 to 40,000. Fig. 1 plots the numerically computed convergence rate $S_{\mathrm{g}}^*(\overline{L})$ with respect to $\mathrm{E}[\lambda_2(L)]$ (left plot, blue solid line) and with respect to $\lambda_2(\overline{L})$ (right plot, blue solid line.) These two plots are remarkably similar and both show that, except for local oscillations, the trend of the convergence rate $S_{\mathrm{g}}^*(\overline{L})$ is to increase with increasing $\mathrm{E}[\lambda_2(L)]$ and $\lambda_2(\overline{L})$. Of course, $\lambda_2(\overline{L})$ is much easier to evaluate than $\mathrm{E}[\lambda_2(L)]$.

We now present a similar set of simulations for the class of Random Geometric Graphs (see [37]), which are used as a reasonable model for wireless sensor networks. In particular, we consider a 35×35 square grid on the plane and deploy 500 sensors with a uniform distribution on the square. The random network deployment is repeated 10 times. For each deployment, the set $\mathcal{E}$ is determined by the radius of connectivity R defined as the maximum distance to which a sensor can possibly communicate with another sensor. For each deployment, we increase R in steps of 3.5 units starting from $R = 0$ to $R_{\max} = 35\sqrt{(2)}$, totalling 14 steps. For each value of R, we generate randomly 100 P matrices. Each entry P_{nl} of P, the link probability for edge $(n, l) \in \mathcal{E}$, is chosen randomly with a uniform distribution $0 \leq P_{nl} \leq 1$. To obtain by numerical average estimates of $S_{\mathrm{g}}^*(\overline{L})$ and $\mathrm{E}[\lambda_2(L)]$, we generate 120 random graphs $L(i)$ from each P. and $\lambda_2(\overline{L})$. Fig. 1 plots $S_{\rm g}^*(\overline{L})$ using red lines versus $\mathrm{E}[\lambda_2(L)]$ and $\lambda_2(\overline{L})$ averaged over the 10 sensor deployments.

Except for the local oscillatory behavior in Fig. 1, these plots confirm that, as a trend, given a class $\mathcal{D}(U)$ of probability distributions P of L, we can set an ordering in $\mathcal{D}(U)$ by evaluating the corresponding $\lambda_2(\overline{L})$'s. This ordering is in the sense that, given two very different values of $\lambda_2(\overline{L})$, the largest will most likely lead to a better convergence rate (see also [19], where part of these results were presented.) This study shows that optimal topologies with respect to ARCCC should be good topologies with respect to RCCC since $S_{\rm g}^*(\overline{L})$ is, as a trend, monotonic with respect to $\lambda_2(\overline{L})$.

B. ARCCC: Performance Analysis

To gain insight into ARCCC, we study the maximum value of its functional as a function of the communication cost constraint U

$$\phi(U)=\max_{\overline{L}\in\mathcal{D}(U)}\lambda_2(\overline{L})=\lambda_2(\overline{L}^*(U)). \tag{52}$$

Lemma 14: Given a cost matrix C, $\phi(U)$ is a concave function of U.

Proof: Let $0\leq U_1\leq U_2$ and $0\leq t\leq 1$. Consider the matrices $\overline{L}_1^*$ and $\overline{L}_2^*$, such that

$$\lambda_2\left(\overline{L}_1^*\right)=\phi(U_1)\quad\text{and}\quad\lambda_2\left(\overline{L}_2^*\right)=\phi(U_2).$$

It follows that $\overline{L}_1^*\in\mathcal{D}(U_1)$ and $\overline{L}_2^*\in\mathcal{D}(U_2)$. Let $\overline{L}=t\overline{L}_1^*+(1-t)\overline{L}_2^*$. Then,

$$\begin{aligned}-\frac{1}{2}\mathrm{Tr}\{C\overline{L}\}&=t\left(-\frac{1}{2}\mathrm{Tr}\left\{C\overline{L}_1^*\right\}\right)+(1-t)\left(-\frac{1}{2}\mathrm{Tr}\left\{C\overline{L}_2^*\right\}\right)\\&\leq tU_1+(1-t)U_2.\end{aligned} \tag{53}$$

Hence, $\overline{L}\in\mathcal{D}(tU_1+(1-t)U_2)$. From this, we conclude that

$$\phi\left(tU_1+(1-t)U_2\right)\geq\lambda_2(\overline{L}). \tag{54}$$

Now, since $\lambda_2(\overline{L})$ is a concave function of $\overline{L}$, we get

$$\begin{aligned}\lambda_2(\overline{L})&=\lambda_2\left(t\overline{L}_1^*+(1-t)\overline{L}_2^*\right)\\&\geq t\lambda_2\left(\overline{L}_1^*\right)+(1-t)\lambda_2\left(\overline{L}_2^*\right)\\&=t\phi(U_1)+(1-t)\phi(U_2).\end{aligned} \tag{55}$$

Finally, using (54) and (55), we get

$$\phi\left(tU_1+(1-t)U_2\right)\geq t\phi(U_1)+(1-t)\phi(U_2) \tag{56}$$

that establishes the concavity as a function of U of the ARCCC optimal cost $\phi(U)$. ■

We use the concavity of $\phi(U)$ to derive an upper bound on $\phi(U)$. Denote by $\mathcal{M}$ the edge set of the complete graph—the set of all possible $N(N-1)/2$ edges. Recall that the set of realizable edges

$$\mathcal{E}=\{(n,l)\in\mathcal{M}:C_{nl}<\infty\}\subset\mathcal{M} \tag{57}$$

with $L_{\mathcal{E}}$ the associated Laplacian. Also, let the total cost $C_{\rm tot}$, i.e., the communication cost per iteration when all the *realizable* links (links in $\mathcal{E}$) are used, be

$$C_{\rm tot}=\sum_{(n,l)\in\mathcal{E}}C_{nl}. \tag{58}$$

Lemma 15: Let C be a cost matrix and $U\geq C_{\rm tot}$. Then $\phi(U)=\lambda_2(L_{\mathcal{E}})$. If, in particular, $\mathcal{E}=\mathcal{M}$, then $\phi(U)=N$. Further,

$$S_{\rm g}^{**}(U)=\max_{\alpha\in\mathbb{R}}S_{\rm g}(\alpha,L_{\mathcal{E}}),\forall U\geq C_{\rm tot}. \tag{59}$$

Proof: The best case is when all the network links $(n,l)\in\mathcal{E}$ have link probability $P_{nl}=1$ (the links in the complement set $\mathcal{E}^c$ must have zero link probability to satisfy the cost constraint.) Then, $\overline{L}=L_{\mathcal{E}}$. Now, if $U\geq C_{\rm tot}$, then $L_{\mathcal{E}}\in\mathcal{D}(U)$ and hence the first part follows. The case $\mathcal{E}=\mathcal{M}$ follows from the fact that, for a complete graph, $\lambda_2(L_{\mathcal{M}})=N$ (see [26], [27].) The second part follows from an exactly similar argument. ■

Using the concavity of $\phi(U)$ (Lemma 14), we now derive a performance bound when $U\leq C_{\rm tot}$.

Lemma 16: Let C be a cost matrix. Then

$$\phi(U)\geq\left(\frac{U}{C_{\rm tot}}\right)\lambda_2(L_{\mathcal{E}}),\quad 0\leq U\leq C_{\rm tot}. \tag{60}$$

If, in particular, $\mathcal{E}=\mathcal{M}$, then

$$\phi(U)\geq\left(\frac{U}{C_{\rm tot}}\right)N,\quad 0\leq U\leq C_{\rm tot}. \tag{61}$$

Proof: From Lemma 15, $\phi(C_{\rm tot})=\lambda_2(L_{\mathcal{E}})$. Then, using the concavity of $\phi(U)$ (see Lemma 14) and the fact that $\phi(0)=0$, we have, for $0\leq U\leq C_{\rm tot}$,

$$\begin{aligned}\phi(U)&=\phi\left(\left(\frac{U}{C_{\rm tot}}\right)C_{\rm tot}\right)\\&\geq\left(\frac{U}{C_{\rm tot}}\right)\phi(C_{\rm tot})\\&=\left(\frac{U}{C_{\rm tot}}\right)\lambda_2(L_{\mathcal{E}}).\end{aligned} \tag{62}$$

This proves the Lemma. The case $\mathcal{E}=\mathcal{M}$ follows easily. ■

Lemma 15 states that $S_{\rm g}^{**}(U)$, $\widehat{S_{\rm g}^{**}(U)}$, and $\phi(U)$ as functions of U reach their maximum value at $C_{\rm tot}$. Equation (62) also shows that $\phi(U)$ lies above the straight line obtained by linearly interpolating between $(U=0,\phi(U)=0)$ and $(U=C_{\rm tot},\phi(U)=\lambda_2(L_{\mathcal{E}})$. Now, since we argue that $S_{\rm g}^*(\overline{L})$ is, in general, a non-decreasing function of $\lambda_2(\overline{L})$, we expect the $\widehat{S_{\rm g}^{**}}(U)$ versus the U curve to lie above the straight line between $(U=0,\widehat{S_{\rm g}^{**}}(U)=0)$ and $(U=C_{\rm tot},\widehat{S_{\rm g}^{**}}(U)=\max_{\alpha\in\mathbb{R}}S_{\rm g}(\alpha,L_{\mathcal{E}})$. Lemma 16 is interesting in this sense, since it states that the ARCCC optimal topology may achieve better performance (in terms of the convergence gain $\widehat{S_{\rm g}^{**}}(U)$) than the fraction of communication cost it uses would lead us to expect. The numerical study in the next section helps to quantify these qualitative assessments.

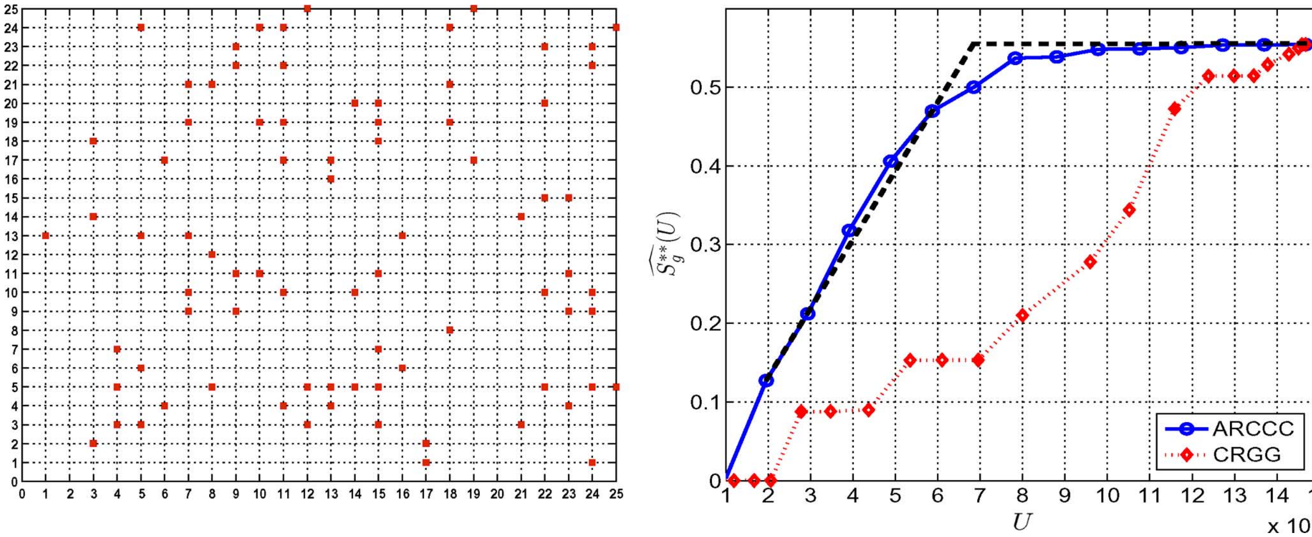


Fig. 2. Left: Sensor placement of $N = 80$ sensors in a 25×25 square grid ($\eta = 1$). Right: Convergence gain versus communication cost U: ARCCC optimal topology—top (blue) line; CRGG topology—bottom (red) line.

C. Numerical Studies: ARCCC

This section determines the probability distribution P and the weight α by solving the semidefinite programming optimization in ARCCC given by (48) and (49). The matrix P assigns to each realizable link its probability of error, or the fraction of time it is expected to be active. For fixed α, the maximization in (48) leads to P and $\overline{L}^{\star}$. For this $\overline{L}^{\star}$, (49) is a scalar convex minimization problem and Lemma 11 gives the range where the optimal α lies. Because of the boundedness of the search interval, as noted before, practically any line search algorithm will find the optimal α easily. In the paper, we discretize the interval with a very fine resolution, evaluate the objective at these points and take the minimum. A fine resolution gives very good performance (comparable to line search methods like Armijo rule) because the spectral functions of these type of matrices are well-behaved.

We compare the ARCCC optimal topology to a random geometric graph topology of fixed radius connectivity. We refer to them as constrained random geometric graphs (CRGG) because sensors only transmit to other sensors within a radius but for which there is a finite communication cost, i.e., an edge in $\mathcal{E}$. The sensor network, shown on the left of Fig. 2, deploys with uniform probability $N = 80$ sensors on a 25×25 square grid. The set $\mathcal{E}$ of realizable links is constructed by choosing $|\mathcal{E}|| = 9N$ edges randomly from the set $\mathcal{M}$ of all possible edges. We assume a geometric propagation model: the communication cost is proportional to the square of the Euclidean distance d_{nl} between sensors n and l

$$C_{nl} = \begin{cases} \eta d_{nl}^2, & \text{if } (n,l) \in \mathcal{E} \\ \infty, & \text{otherwise} \end{cases} \tag{63}$$

where η is an appropriately chosen constant. With this CRGG network, a sensor n communicates with all other sensors l with a finite communication cost $C_{nl} < \infty$ that lie within a radius R.

Fig. 2 on the right plots, as a function of the cost constraint U, the per step convergence gain $\widehat{S_{\mathrm{g}}^{**}}(U)$ for the ARCCC optimal topology (top blue solid line) and the per step convergence gain $S_{\mathrm{g}}^{\mathrm{CRGG}}(U)$ of the CRGG topology (bottom red dotted line). The ARCCC optimal topology converges much faster than the CRGG topology, with the improvement being more significant at medium to lower values of U.

The ARCCC topology has a markedly nonlinear behavior, with two asymptotes (the dotted black lines): for small U, the sharp increasing asymptote and for large U the horizontal asymptote (when all the realizable edges in $\mathcal{E}$ are used.) The two meet at the knee of the curve ($U = 6.9 \times 10^4, \widehat{S_{\mathrm{g}}^{**}}(U) = .555$.) For $U = 6.9 \times 10^4$, the ARCCC convergence rate is $\widehat{S_{\mathrm{g}}^{**}}(U) = .505$, while CRGG's is $S_{\mathrm{g}}^{\mathrm{CRGG}}(U) = .152$, showing that the ARCCC's topology is 3.3 times faster than the CRGG's. For this example, we compute $C_{\mathrm{tot}} = 14.7 \times 10^4$, which shows that the ARCCC's optimal topology achieves the asymptotic performance while using less than 50% of the communication cost.

VIII. Conclusion

The paper presents the design of the topology of a random sensor network to maximize the convergence rate of the consensus algorithm. We consider that the communication channels among sensors may fail at random times, that communication among sensors incurs a cost, and that there is an overall communication cost constraint in the network. The solution to this topology optimization specifies for each realizable link its probability of error, or the fraction of time the link is expected to be active. We first establish necessary and sufficient conditions for mss convergence and a.s. convergence of the consensus algorithm in terms of the expected value of the algebraic connectivity of the random graph defining the network topology and in terms of the algebraic connectivity of the average topology. We approximate the original optimization problem by an alternative two step algorithm, the approximate random communication cost constraint (ARCCC) problem. ARCCC is convex and we solve it by semidefinite programming techniques.

We discuss briefly some tradeoffs presented by the ARCCC design. Simulations show it can improve by about 300% the convergence speed of average consensus over more common designs, e.g., the CRGG design, i.e., geometric topologies where sensors communicate with sensors within a fixed distance. Our study also shows that the optimal random topology can achieve the convergence speed of a nonrandom network at a fraction of the cost. But, by maximizing the algebraic connectivity of the network to achieve these higher convergence rates, the ARCCC solution is not always local; it involves, even if with low probability, long distance communications, since the transmission radius is increased. This causes possibly added or stronger interference with further away sensors. This requires more complicated communication protocols (e.g., some form of multiplexing like time division multiple access protocols) than, for example, with a fixed-radius connectivity topology, in order to coordinate transmissions among sensors, avoid collisions, or combat the stronger interference. These effects can still be accounted for by the approach in this paper, but the price paid by the ARCCC design is to increase the costs of communication per iteration, i.e., the entries in the cost matrix C. This means that with the ARCCC design there is a price to pay. To be more realistic, in Fig. 2, the ARCCC design with communication cost constraint $U_{\rm ARCCC}$ should have been compared with a CRGG design with cost constraint $U_{\rm CRGG} < U_{\rm ARCCC}$, and, this may be a subject of future work.

REFERENCES

[1] M. J. Fischer, N. A. Lynch, and M. S. Paterson, "Impossibility of distributed consensus with one faulty process," *J. Assoc. Comput. Machinery*, vol. 32, no. 2, pp. 374–382, Apr. 1985.

[2] N. A. Lynch, *Distributed Algorithms*. San Francisco, CA: Morgan Kaufmann, 1997.

[3] G. V. Cybenko, "Dynamic load balancing for distributed memory multiprocessors," *J. Parallel Distrib. Comput.*, vol. 7, pp. 279–301, 1989.

[4] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Automat. Contr.*, vol. AC-48, no. 6, pp. 988–1001, Jun. 2003.

[5] C. Reynolds, "Flocks, birds, and schools: A distributed behavioral model," *Computer Graphics*, vol. 21, pp. 25–34, 1987.

[6] T. Vicsek, A. Czirok, E. B. Jacob, I. Cohen, and O. Schochet, "Novel type of phase transitions in a system of self-driven particles," *Phys. Rev. Lett.*, vol. 75, pp. 1226–1229, 1995.

[7] R. O. Saber and R. M. Murray, "Consensus protocols for networks of dynamic agents," in *Proc. 2003 American Control Conf.*, Denver, CO, Jun. 2003, vol. 2, pp. 951–956.

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

[9] W. Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proc. American Control Conf.*, Portland, OR, Jun. 2005, pp. 1859–1864.

[10] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[11] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, 1984.

[12] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Automat. Contr.*, vol. AC-31, no. 9, pp. 803–812, Sep. 1986.

[13] M. G. Rabbat, R. D. Nowak, and J. A. Bucklew, "Generalized consensus algorithms in networked systems with erasure links," in *Proc. 6th IEEE Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, New York, 2005, pp. 1088–1092.

[14] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 5, pp. 2508–2530, Jun. 2006.

[15] A. T. Salehi and A. Jadbabaie, "On consensus in random networks," presented at the Allerton Conf. Commun., Contr., and Comput., Allerton House, IL, Sep. 2006.

[16] Y. Hatano and M. Mesbahi, "Agreement over random networks," in *Proc. 43rd IEEE Conf. Decision and Control*, Dec. 2004, vol. 2, pp. 2010–2015.

[17] R. Olfati-Saber, "Ultrafast consensus in small-world networks," in *Proc. American Control Conf.*, Portland, OR, Jun. 2005, vol. 4, pp. 2371–2378.

[18] S. A. Aldosari and J. M. F. Moura, "Distributed detection in sensor networks: Connectivity graph and small-world networks," in *Proc. 39th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 2005, pp. 230–234.

[19] S. Kar and J. M. F. Moura, "Distributed average consensus in sensor networks with random link failures," presented at the 32nd IEEE Int. Conf. Acoust., Speech, Signal Process., Honolulu, HI, Apr. 2007.

[20] S. Kar and J. M. F. Moura, "Ramanujan topologies for decision making in sensor networks," presented at the 44th Allerton Conf. Commun., Contr., Comput., Monticello, IL, Sep. 2006.

[21] S. Kar and J. M. F. Moura, "Topology for global average consensus," presented at the 40th Asilomar Conf. Signals, Syst., Comput., Pacific Grove, CA, Oct. 2006.

[22] S. Kar, S. A. Aldosari, and J. M. F. Moura, "Topology for distributed inference on graphs," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2609–2613, Jun. 2008.

[23] R. Olfati-Saber, "Algebraic connecitivity ratio of Ramanujan graphs," presented at the American Control Conf., New York, Jul. 2007.

[24] S. Kar and J. M. F. Moura, "Distributed average consensus in sensor networks with quantized inter-sensor communication," presented at the IEEE Int. Conf. Signal Process. (ICASSP'08), Las Vegas, NV, Apr. 2008.

[25] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks with communication channel noise and random link failures," presented at the 41st Asilomar Conf. Signals, Syst., Comput., Pacific Grove, CA, Nov. 2007.

[26] F. R. K. Chung, *Spectral Graph Theory*. Providence, RI: Amer. Math. Soc., 1997.

[27] B. Mohar, "The Laplacian spectrum of graphs," in *Graph Theory, Combinatorics, and Applications*, Y. Alavi, G. Chartrand, O. R. Oellermann, and A. J. Schwenk, Eds. New York: Wiley, 1991, vol. 2, pp. 871–898.

[28] B. Bollobás, *Modern Graph Theory*. New York: Springer Verlag, 1998.

[29] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak. Math. J.*, vol. 23, no. 98, pp. 298–305, 1973.

[30] S. Camiz and S. Stefani, *Matrices and Graphs: Theory and Applications*. , Singapore: World Scientific, 1996.

[31] S. Boyd, "Convex optimization of graph Laplacian eigenvalues," in *Proc. Int. Congr. Mathematicians*, 2006, vol. 3, pp. 1311–1319.

[32] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Contr. Lett.*, vol. 53, pp. 65–78, Sep. 2004.

[33] O. Kallenberg, *Foundations of Modern Probability*, ser. Springer Series in Statistics, 2nd ed. New York: Springer, 2002.

[34] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge Univ. Press, 2004.

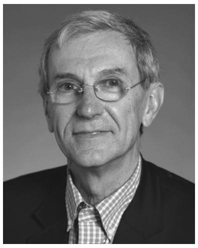
[35] H. Wolkowicz, R. Saigal, and L. Vandengerghe, *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*. Boston, MA: Kluwer, 2000.

[36] Y. Kim and M. Mesbahi, "On maximizing the second smallest eigenvalue of a state-dependent graph Laplacian," *IEEE Trans. Automat. Contr.*, vol. 51, no. 1, pp. 116–120, Jan. 2006.

[37] M. Penrose, *Random Geometric Graphs*. Oxford, U.K.: Oxford Univ. Press, 2003.

Soummya Kar (S'07) received the B.Tech. degree from the Indian Institute of Technology, Kharagpur, India, in May, 2005. He joined the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, in Fall 2005, where he is currently a third-year Ph.D. student.

His research interests include distributed computations in sensor networks, stochastic approximations, stochastic networks, and stochastic filtering in the context of sensor networks.



José M. F. Moura (S'71–M'75–SM'90–F'94) received the engenheiro electrotécnico degree from Instituto Superior Técnico (IST), Lisbon, Portugal, and the M.Sc., E.E., and D.Sc. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA.

He is a Professor of electrical and computer engineering and, by courtesy, of biomedical engineering, at Carnegie Mellon University (CMU), Pittsburgh, PA. He was on the faculty at IST, has held visiting faculty appointments at MIT, and was a research scholar at the University of Southern California, Los Angeles. He is a founding co-director of the Center for Sensed Critical Infrastructures Research (CenSCIR) and of the Information and Communications Technologies Institute, a joint venture between CMU and Portugal that manages a large education and research program between CMU and Portugal (www.icti.cmu.edu). His research interests include statistical and algebraic signal processing, image, bioimaging, and video processing, and digital communications. He has published over 300 technical journal and conference papers, is the co-editor of two books, holds six U.S. patents on image and video processing and digital communications, and has given numerous invited seminars at U.S. and European Universities and industrial and government laboratories.

Dr. Moura is the President (2008–2009) of the IEEE Signal Processing Society (SPS). He was Editor-in-Chief for the IEEE TRANSACTIONS IN SIGNAL PROCESSING, interim Editor-in-Chief for the *IEEE Signal Processing Letters*, and was on the Editorial Board of several journals, including the PROCEEDINGS OF THE IEEE, the *IEEE Signal Processing Magazine*, and the *ACM Transactions on Sensor Networks*. He was Vice-President for Publications for the IEEE Sensors Council. He chaired the IEEE TAB Transactions Committee, served on the IEEE TAB Periodicals Review Committee, and is currently Vice-Chair of the IEEE Publication Services and Products Board. He was on the steering and technical committees of several conferences and on the IEEE Press Board. He is a Fellow of the IEEE, a Fellow of the American Association for the Advancement of Science (AAAS), and a corresponding member of the Academy of Sciences of Portugal (Section of Sciences). He was awarded the 2003 IEEE Signal Processing Society Meritorious Service Award and in 2000 the IEEE Millennium Medal. In 2007 he received the CMU's College of Engineering Outstanding Research Award. He is affiliated with several IEEE societies, Sigma Xi, AMS, IMS, and SIAM.