

Topology for Distributed Inference on Graphs

Soumya Kar, Saeed Aldosari, and José M. F. Moura*

Abstract

Let N local decision makers in a sensor network communicate with their neighbors to reach a decision *consensus*. Communication is local, among neighboring sensors only, through noiseless or noisy links. We study the design of the network topology that optimizes the rate of convergence of the iterative decision consensus algorithm. We reformulate the topology design problem as a spectral graph design problem, namely, maximizing the eigenratio γ of two eigenvalues of the graph Laplacian L , a matrix that is naturally associated with the interconnectivity pattern of the network. This reformulation avoids costly Monte Carlo simulations and leads to the class of non-bipartite Ramanujan graphs for which we find a lower bound on γ . For Ramanujan topologies and noiseless links, the local probability of error converges much faster to the overall global probability of error than for structured graphs, random graphs, or graphs exhibiting small-world characteristics. With noisy links, we determine the optimal number of iterations before calling a decision. Finally, we introduce a new class of random graphs that are easy to construct, can be designed with arbitrary number of sensors, and whose spectral and convergence properties make them practically equivalent to Ramanujan topologies.

Key words: Sensor networks, consensus algorithm, distributed detection, topology optimization, Ramanujan, Cayley, small-world, random graphs, algebraic connectivity, Laplacian, spectral graph theory.

EDICS: SEN-DIST, SEN-FUSE

The 1st and 3rd authors are with the Dep. ECE, Carnegie Mellon University, Pittsburgh, PA, USA 15213 (e-mail: {soumyak,moura}@ece.cmu.edu, ph: (412)268-6341, fax: (412)268-3890.) The 2nd author is with EE Dept., King Saud University, P. O. Box 800, Riyadh, 11412, Saudi Arabia, (dosari@ksu.edu.sa, ph: +966-553367274, fax: + 966-1-4676757.)

Work supported by the DARPA DSO Advanced Computing and Mathematics Program Integrated Sensing and Processing (ISP) Initiative under ARO grant # DAAD19-02-1-0180 and by NSF under grants # ECS-0225449 and # CNS-0428404.

I. INTRODUCTION

The paper studies the problem of designing the topology of a graph network. As a motivational application we consider the problem of describing the connectivity graph of a sensor network, i.e., specifying with which sensors should each sensor in the network communicate. We will show that the topology of the network has a major impact on the convergence of distributed inference algorithms, namely, that these algorithms converge much faster for certain connectivity patterns than for others, thus requiring much less intersensor communication and power expenditure.

The literature on topology design for distributed detection is scarce. Usually, the underlying communication graph is specified ab initio as a structured graph, e.g., parallel networks where sensors communicate with a fusion center, [1], [2], [3], [4], or serial networks where communication proceeds sequentially from a sensor to the next; for these and other similar architectures, see Varshney [5] or [6], [7]. These networks may not be practical; e.g., a parallel network depends on the integrity of the fusion center.

We published preliminary results on topology design for distributed inference problems in [8], [9]. We restricted the class of topologies to structured graphs, random graphs obtained with the Erdős-Rényi construction, [10], [11], [12], see also [13], [14], or random constructions that exhibit small-world characteristics, see Watts-Strogatz [15], see also Kleinberg [16], [17]. We considered tradeoffs among these networks, their number of links M , and the number of bits b quantizing the state of the network at each sensor, under a global rate constraint, i.e., $Mb = K$, K fixed. We adopted as criterion the convergence of the average probability of error P_e , which required extensive simulation studies to find the desired network topology. Reference [18] designs Watts-Strogatz topologies in distributed consensus estimation problems, adopting as criterion the algebraic connectivity $\lambda_2(L)$ of the graph.

This paper designs good topologies for sensor networks, in particular, with respect to the rate of convergence of *iterative* consensus and *distributed* detection algorithms. We consider the two cases of noiseless and noisy network links. We assume that the total number M of communication links between sensors is fixed and that the graph weights are uniform across all network links. This paper shows that, for both the iterative average-consensus and the distributed detection problems, the topology design problem is equivalent to the problem of maximizing with respect to the network topology a certain graph spectral parameter γ . This parameter is the ratio of the algebraic connectivity of the graph over the largest eigenvalue $\lambda_N(L)$ of the graph Laplacian L . The algebraic connectivity of a graph, terminology introduced by [19], is the second smallest eigenvalue $\lambda_2(L)$ of its discrete Laplacian; see section II, the Appendix, and reference [20] for the definition of relevant spectral graph concepts. With this reinterpretation, we show that the class of Ramanujan graphs essentially provides the optimal network topologies, exhibiting

remarkable convergence properties, orders of magnitude faster than other structured or random small-world like networks. When the links are noisy, our analysis determines what is the optimal number of iterations to declare a decision. Finally, we present a new class of random regular graphs whose performance is very close to the performance of Ramanujan graphs. These graphs can be designed with arbitrary number of nodes, overcoming the limitation that the available constructions of Ramanujan graphs are restricted to networks whose number of sensors are limited to a sparse subset of the integers.

We now summarize the paper. Section II and the Appendix recall basic concepts and results from algebra and spectral graph theory. Section III presents the optimal *equal* weights consensus algorithm and establishes its convergence rate in terms of a spectral parameter. Section IV defines formally the topology design problem, shows that Ramanujan graphs provide essentially the optimal topologies, and presents explicit algebraic constructions available in the literature for the Ramanujan graphs. Section V considers distributed inference and shows that the average-consensus algorithm with noiseless links achieves asymptotically the optimal detection performance—that of a parallel architecture with a fusion center. This section shows that with noisy communication links there is an optimal maximum number of iterations to declare a decision. Section VI demonstrates with several experiments the superiority of the Ramanujan designs over other different alternative topologies, including structured networks, Erdős-Rényi random graphs, and small-world type topologies. Section VII presents the new class of random regular Ramanujan like graphs that are easy to design with arbitrary number of sensors and that exhibit convergence properties close to Ramanujan topologies. Finally, section VIII concludes the paper.

II. ALGEBRAIC PRELIMINARIES

Graph Laplacian The topology of the sensor network is given by an undirected graph $G = (V, E)$, with nodes $v_n \in V$, $n \in \mathcal{I} = \{1, \dots, N\}$, and edges the unordered pairs $e = (v_n, v_l) \in E$, or, simply, $e = (n, l)$, where v_n and v_l are called the edge endpoints. The edge $e = (n, l) \in E$ whenever sensor v_n can communicate with v_l , in which case the vertices v_n and v_l are adjacent and we write $v_n \sim v_l$.

We assume that the cardinality of E is $|E| = M$ and, when needed, label the edges by m , $m = 1, \dots, M$. The terms sensor, node, and vertex are assumed to be equivalent in this paper. A loop is an edge whose endpoints are the same vertex. Multiple edges are edges with the same pair of endpoints. A graph is *simple* if it has no loops or multiple edges. A graph with loops or multiple edges is called a multigraph. A *path* is a sequence v_{n_0}, \dots, v_{n_m} such that $e_l = (v_{n_{l-1}}, v_{n_l}) \in E$, $l = 1, \dots, m$, and the v_{n_l} , $l = 0, \dots, m - 1$, are all distinct. A graph is *connected* if there is a path from every sensor v_n to every other sensor v_l , $n, l = 1, \dots, N$. In this paper we assume the graphs to be *simple* and *connected*,

unless otherwise stated.

We can assign to a graph an $N \times N$ adjacency matrix A (where, we recall, $N = |V|$), defined by

$$a_{n,l} = \begin{cases} 1 & \text{if } (n,l) \in E \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The set of neighbors of node n is $\Omega_n = \{l : (n,l) \in E\}$ and its degree, $\deg(n)$, is the number of its neighbors, i.e., the cardinality $|\Omega_n|$. A graph is k -regular if all vertices have the same degree k .

The degree matrix, D is the $N \times N$ diagonal matrix $D = \text{diag}[d_{1,1} \cdots d_{N,N}]$ defined by

$$d_{n,n} = \deg(n) \quad (2)$$

The Laplacian L of the graph, [20], is the $N \times N$ matrix defined by

$$L = D - A \quad (3)$$

Spectral properties of graphs. We consider spectral properties of *connected regular* graphs. Since the adjacency matrix A is symmetric, all its eigenvalues are real. Arrange the eigenvalues of the adjacency matrix A as,

$$k = \lambda_1(A) > \lambda_2(A) \geq \dots \geq \lambda_N(A) \geq -k \quad (4)$$

It can be shown that the multiplicity of the largest eigenvalue $\lambda_1(A) = k$ equals the number of connected components in the graph. Then, for a connected graph, the multiplicity of the eigenvalue $\lambda_1(A) = k$ is 1, which explains the strict inequality on the left in (4). Also, $-k$ is an eigenvalue of A iff the graph is bipartite (please refer to the Appendix for the definition of bipartite graphs.) Hence, for non-bipartite graphs, $\lambda_N(A) > -k$. In this paper, we focus on connected, non-bipartite graphs.

The Laplacian is a symmetric, positive semi-definite matrix, and, consequently, all its eigenvalues are non-negative. It follows from (4) that the eigenvalues of the Laplacian satisfy

$$0 = \lambda_1(L) < \lambda_2(L) \leq \dots \leq \lambda_N(L) \quad (5)$$

The multiplicity of the zero eigenvalue of L equals the number of connected components in the graph, which explains the strict inequality on the left hand side of (5) for the case of connected graphs. For k -regular graphs, the eigenvalues of A and L are directly related by

$$\forall n \in \mathcal{I} : \lambda_n(L) = k - \lambda_n(A) \quad (6)$$

We write the eigendecomposition of the Laplacian L as

$$L = \mathbf{U}\Lambda\mathbf{U}^T \quad (7)$$

$$= [\mathbf{u}_1 \cdots \mathbf{u}_N] \text{diag} [\lambda_1(L) \cdots \lambda_N(L)] [\mathbf{u}_1 \cdots \mathbf{u}_N]^T \quad (8)$$

where the \mathbf{u}_n , $n = 1, \dots, N$, are orthonormal and $\text{diag}[\cdots]$ is a diagonal matrix. We note that, from the structure of L , each diagonal entry of D is the corresponding row sum of A , so the eigenvector \mathbf{u}_1 corresponding to the zero eigenvalue $\lambda_1(L)$ is the (normalized) vector of one's

$$\mathbf{u}_1 = \frac{1}{\sqrt{N}}\mathbf{1} = \frac{1}{\sqrt{N}}[1 \cdots 1]^T \quad (9)$$

III. AVERAGE CONSENSUS ALGORITHM

In this Section, we present the consensus algorithm in Subsection III-A, discuss the case of equal weights in Subsection III-B, and establish the convergence rate of the algorithm in Subsection III-C.

A. Consensus Algorithm Description

We review briefly the consensus algorithm that computes in a distributed fashion the average of N quantities r_n , $n = 1, \dots, N$. Assume a sensor network with interconnectivity graph $G = (V, E)$ defined by a neighborhood system $\Omega = \{\Omega_n : n \in \mathcal{I}\}$, and where Ω_n is the set of neighbors of sensor n . Initially, sensors take measurements r_1, \dots, r_N . It is desired to compute their mean in a distributed fashion,

$$\bar{r} = \frac{1}{N} \sum_{n=1}^N r_n \quad (10)$$

i.e., by only local communication among neighbors. Define the *state* at iteration $i = 0$ at sensor n by

$$x_n(i = 0) = r_n, \quad n = 1, \dots, N$$

Iterative consensus is carried out according to the following linear operation, [21],

$$x_n(i) = W_{nn}x_n(i-1) + \sum_{l \in \Omega_n} W_{nl}x_l(i-1) \quad (11)$$

where W_{nl} is a weight associated with edge (n, l) , if this edge exists. The weight $W_{nl} = 0$, $n \neq l$, when there is no link associated with it, i.e., if $(n, l) \notin E$. The value $x_n(i)$ stored at iteration i by sensor n is the *state* of v_n at i . The consensus (11) can be expressed in matrix form as

$$\mathbf{x}_i = W\mathbf{x}_{i-1} \quad (12)$$

where \mathbf{x}_i is the $N \times 1$ vector of all current states and $W = [W_{nl}]$ is the matrix of all the weights. The updating (12) can be written in terms of the initial states as

$$\mathbf{x}_i = W^i \mathbf{x}_0 \quad (13)$$

$$\mathbf{x}_0 = [x_1(0) \cdots x_N(0)]^T = [r_1 \cdots r_N]^T \quad (14)$$

Let the N -dimensional vector $\mathbf{1} = [1 \cdots 1]^T$. Convergence to consensus occurs if

$$\forall n : \lim_{i \rightarrow \infty} x_n(i) = \bar{r}$$

$$\lim_{i \rightarrow \infty} \mathbf{x}_i = \bar{\mathbf{x}} = \bar{r} \mathbf{1} \quad (15)$$

$$\lim_{i \rightarrow \infty} W^i = \frac{\mathbf{1}\mathbf{1}^T}{N} \quad (16)$$

B. Link Weights

The convergence speed of the iterative consensus depends on the choice of the link weights, W_{nl} . In this paper, we consider only the case of *equal* weights, i.e., we assign an equal weight α to all network links. I and L be the N -dimensional identity matrix and the graph Laplacian. The weight matrix becomes

$$W = I - \alpha L \quad (17)$$

For a particular network topology, the value of α that maximizes the convergence speed is, [21],

$$\alpha^* = \frac{2}{\lambda_2(L) + \lambda_N(L)} \quad (18)$$

For proofs of these statements and other weight design techniques, the reader is referred to [21] and [18].

We now consider the eigendecomposition of the weight matrix W . From (17), with the optimal weight (18), using the eigendecomposition (8) of L , we have that

$$W = [\mathbf{u}_1 \cdots \mathbf{u}_N] \text{diag}[\gamma_1 \cdots \gamma_N] [\mathbf{u}_1 \cdots \mathbf{u}_N]^T \quad (19)$$

$$= \sum_{n=1}^N \gamma_n \mathbf{u}_n \mathbf{u}_n^T, \quad (20)$$

where: $\mathbf{u}_n, n = 1, \cdots, N$, are the orthonormal eigenvectors of L , and à fortiori of W ; and $\text{diag}[\gamma_1 \cdots \gamma_N]$ is the diagonal matrix of the eigenvalues γ_n of W . These eigenvalues are

$$\gamma_n = 1 - \alpha^* \lambda_n(L) \quad (21)$$

From the spectral properties of the Laplacian of a connected graph, and the choice of α^* , the eigenvalues

of W satisfy

$$1 = \gamma_1 > \gamma_2 \geq \dots \geq \gamma_N \quad (22)$$

$$\forall n > 1 : \quad |\gamma_n| \leq \gamma_2 < 1 \quad (23)$$

C. Consensus Algorithm: Convergence Rate

We now study the convergence rate of the consensus algorithm.

Result 1 For any connected graph G , the convergence rate of the consensus algorithm (12) or (13) is

$$\|\mathbf{x}_i - \bar{\mathbf{x}}\| \leq \|\mathbf{x}_0 - \bar{\mathbf{x}}\| \gamma_2^i \quad (24)$$

where $\bar{\mathbf{x}}$ and \mathbf{x}_0 are given in (15) and (13) and

$$\gamma_2 = \frac{1 - \gamma}{1 + \gamma} \quad (25)$$

$$\gamma = \frac{\lambda_2(L)}{\lambda_N(L)} \quad (26)$$

Proof: Represent the vector \mathbf{x}_0 in (14) in terms of the eigenvectors \mathbf{u}_n of L

$$\mathbf{x}_0 = \sum_{n=1}^N d_n \mathbf{u}_n \quad (27)$$

where $d_n = \mathbf{x}_0^T \mathbf{u}_n$. From the value of \mathbf{u}_1 in (9) it follows that

$$d_1 = \sqrt{N} \bar{r} \quad (28)$$

Replacing (20) and (27) in (13) and using (28) and the orthonormality of the eigenvectors of L (and W), we obtain

$$\begin{aligned} \mathbf{x}_i &= W^i \mathbf{x}_0 \\ &= \sum_{l=1}^N \gamma_l^i \mathbf{u}_l \mathbf{u}_l^T \sum_{n=1}^N d_n \mathbf{u}_n \end{aligned} \quad (29)$$

$$\begin{aligned} &= \sum_{n,l=1}^N d_n \gamma_l^i \mathbf{u}_l \mathbf{u}_l^T \mathbf{u}_n \\ &= \sum_{l=1}^N d_l \gamma_l^i \mathbf{u}_l \\ &= d_1 \gamma_1^i \mathbf{u}_1 + \sum_{l=2}^N d_l \gamma_l^i \mathbf{u}_l \\ &= \bar{\mathbf{x}} + \sum_{l=2}^N d_l \gamma_l^i \mathbf{u}_l \end{aligned} \quad (30)$$

where $\bar{\mathbf{x}}$ is given in eqn (15). From these it follows that

$$\begin{aligned} \|\mathbf{x}_i - \bar{\mathbf{x}}\| &= \left\| \sum_{l=2}^N d_l \gamma_l^i \mathbf{u}_l \right\| \\ &\leq |\gamma_2^i| \left\| \sum_{l=2}^N d_l \mathbf{u}_l \right\| \end{aligned} \quad (31)$$

$$= \|\mathbf{x}_0 - \bar{\mathbf{x}}\| \gamma_2^i \quad (32)$$

To get (31), we used the bounds given by (23). To obtain (32) we used the fact that from (30), for $i = 0$, it follows that

$$\left\| \sum_{l=2}^N d_l \gamma_l^i \mathbf{u}_l \right\| = \|\mathbf{x}_0 - \bar{\mathbf{x}}\| \quad (33)$$

From (33) it follows that, to obtain the optimal convergence rate, γ_2 should be as small as possible. From the expression for γ_n in (21), and using the optimal choice for α in (18), we get successively

$$\gamma_2 = 1 - \alpha \lambda_2(L) \quad (34)$$

$$\begin{aligned} &= 1 - \frac{2\lambda_2(L)}{\lambda_2(L) + \lambda_N(L)} \\ &= \frac{\lambda_N(L) - \lambda_2(L)}{\lambda_N(L) + \lambda_2(L)} \\ &= \frac{1 - \lambda_2(L)/\lambda_N(L)}{1 + \lambda_2(L)/\lambda_N(L)} \end{aligned} \quad (35)$$

Thus, the minimum value of γ_2 is attained when the ratio

$$\lambda_2(L)/\lambda_N(L) \tag{36}$$

is maximum, i.e.,

$$\text{max convergence rate} \sim \min \gamma_2 \sim \max \gamma = \max \frac{\lambda_2(L)}{\lambda_N(L)} \tag{37}$$

■

IV. TOPOLOGY DESIGN: RAMANUJAN GRAPHS

In this section, we consider the problem of designing the topology of a sensor network that maximizes the rate of convergence of the average consensus algorithm. Using the results of Section III, in Subsection IV-A, we reformulate the average consensus topology design as a spectral graph topology design problem by restating it in terms of the design of the topology of the network that maximizes an eigenratio of two eigenvalues of the graph Laplacian, namely, the graph parameter γ given by (26). We then consider in Subsection IV-B the class of Ramanujan graphs and show in what sense they are good topologies. Finally, Subsection IV-C describes algebraic constructions of Ramanujan graphs available in the literature, see [22].

A. Topology Optimization

We formulate the design of the topology of the sensor network for the average consensus algorithm as the optimization of the spectral eigenratio parameter γ , see (26). From our discussion in Section III, it follows that the topology that optimizes the convergence rate of the consensus algorithm can be restated as the following graph optimization problem:

$$\max_{G \in \mathcal{G}} \gamma = \max_{G \in \mathcal{G}} \frac{\lambda_2(L)}{\lambda_N(L)} \tag{38}$$

where \mathcal{G} denotes the set of all possible simple connected graphs with N vertices and M edges.

We remark that (38) will be significant because we will be able to use spectral properties of graphs to propose a class of graphs—the Ramanujan graphs—for which we can present a lower bound on the spectral parameter γ . This avoids the lengthy and costly Monte Carlo simulations used to evaluate the performance of other topologies as done, for example, in our previous work, see [8], [9] or in [18].

B. Ramanujan Graphs

In this section, we consider k -regular graphs. Before introducing the class of Ramanujan graphs, we discuss several bounds on eigenvalues of graphs. We first state a well-known result from algebraic graph theory.

Theorem 2 (Alon and Boppana [23], [22]) Let $G = G_{N,k}$ be a k -regular graph on N vertices. Denote by $\lambda_G(A)$, the absolute value of the largest eigenvalue (in absolute value) of the adjacency matrix A of the graph G , which is distinct from $\pm k$; in other words, $\lambda_G^2(A)$ is the next to largest eigenvalue of A^2 . Then

$$\liminf_{N \rightarrow \infty} \lambda_G(A) \geq 2\sqrt{k-1} \quad (39)$$

A second result, [23], also shows that, for an infinite family of k -regular graphs G_m , $m \in \{1, 2, \dots\}$, for which the number of nodes diverges as m becomes large, the algebraic connectivity $\lambda_2(L)$ of the graphs is asymptotically bounded by

$$\liminf_{N \rightarrow \infty} \lambda_2(L) \leq k - 2\sqrt{k-1} \quad (40)$$

Note that (40) is a direct upperbound on the limiting behavior of $\lambda_2(L)$ itself, while from (39) we may derive an upperbound on the limiting behavior of $\lambda_2(A)$ or of $\lambda_N(A)$, depending if $\lambda_2(A) \leq |\lambda_N(A)|$ or $\lambda_2(A) \geq |\lambda_N(A)|$ in the limit. We consider each of these two cases separately.

$$1) \liminf_{N \rightarrow \infty} \lambda_2(A) \leq \liminf_{N \rightarrow \infty} |\lambda_N(A)| : \liminf_{N \rightarrow \infty} |\lambda_N(A)| \geq 2\sqrt{k-1}.$$

Since $\lambda_N(A) \leq 0$, it follows that for k -regular connected simple graphs

$$\liminf_{N \rightarrow \infty} \lambda_N(A) \leq -2\sqrt{k-1}$$

From this, we have

$$\liminf_{N \rightarrow \infty} \lambda_N(L) \geq k + 2\sqrt{k-1} \quad (41)$$

Combining (41) with (40), we get using standard results from limits of series of real numbers

$$\liminf_{N \rightarrow \infty} \gamma(N) = \liminf_{N \rightarrow \infty} \frac{\lambda_2(L)}{\lambda_N(L)} \leq \frac{k - 2\sqrt{k-1}}{k + 2\sqrt{k-1}} \quad (42)$$

Eqn (42) is an asymptotic upper bound on the spectral eigenratio parameter $\gamma = \lambda_2(L)/\lambda_N(L)$ for the family of non-bipartite graphs for which $\liminf \lambda_2(A) \leq \liminf |\lambda_N(A)|$.

$$2) \liminf_{N \rightarrow \infty} \lambda_2(A) \geq \liminf_{N \rightarrow \infty} |\lambda_N(A)| : \liminf_{N \rightarrow \infty} |\lambda_N(A)| \leq 2\sqrt{k-1}.$$

Now Theorem (2) is inconclusive with respect to $\liminf_{N \rightarrow \infty} \lambda_N(A)$. From the fact that $-k \leq$

$\lambda_N(A) \leq 0$, we can promptly deduce that $k \leq \lambda_N(L) \leq 2k$. Combining this with (40), we get

$$\liminf_{N \rightarrow \infty} \frac{\lambda_2(L)}{\lambda_N(L)} \leq \frac{k - 2\sqrt{k-1}}{k} \quad (43)$$

which gives an asymptotic upper bound for the eigenratio parameter $\gamma = \lambda_2(L)/\lambda_N(L)$ for the family of non-bipartite graphs satisfying $\liminf \lambda_2(A) \geq \liminf |\lambda_N(A)|$.

We now consider the class of Ramanujan graphs.

Definition 3 (Ramanujan Graphs) A graph $G = G_{N,k}$ will be called Ramanujan if

$$\lambda_G(A) \leq 2\sqrt{k-1} \quad (44)$$

Graphs with small $\lambda_G(A)$ (often called graphs with large spectral gap in the literature) are called expander graphs, and the Ramanujan graphs are one of the best explicit expanders known. Note that Theorem 2 and (39) show that, for general graphs, $\lambda_G(A)$ is in the limit *lower* bounded by $2\sqrt{k-1}$, while for Ramanujan graphs $\lambda_G(A)$ is, for every finite N , *upper* bounded by $2\sqrt{k-1}$.

From (44), it follows that, for non-bipartite Ramanujan graphs,

$$\lambda_2(A) \leq 2\sqrt{k-1} \quad (45)$$

$$\lambda_N(A) \geq -2\sqrt{k-1} \quad (46)$$

Equations (45) and (46) together with eqn (6) give, for non-bipartite Ramanujan graphs,

$$\lambda_2(L) \geq k - 2\sqrt{k-1}$$

$$\lambda_N(L) \leq k + 2\sqrt{k-1}$$

and, hence, for non-bipartite Ramanujan graphs

$$\gamma = \frac{\lambda_2(L)}{\lambda_N(L)} \geq \frac{k - 2\sqrt{k-1}}{k + 2\sqrt{k-1}} \quad (47)$$

This is a key result and shows that for non-bipartite Ramanujan graphs the eigenratio parameter γ is lower bounded by (47). It will explain in what sense we take Ramanujan graphs to be “optimal” with respect to the topology design problem stated in Subsection IV-A as we discussed next. To do this, we compare the lower bound (47) on γ for Ramanujan graphs with the asymptotic upper bounds (42) and (43) on γ for generic graphs. We consider the two cases separately again.

1) Generic graphs for which $\liminf_{N \rightarrow \infty} \lambda_2(A) \leq \liminf_{N \rightarrow \infty} |\lambda_N(A)|$. Here, the lower bound

on (47) and the upper bound on (42) are the same. Since for any value of N , (47) shows that γ is above the bound, we conclude that, in the limit of large N , the eigenratio parameter γ for non-bipartite Ramanujan graphs approaches the bound from above. This contrasts with non-bipartite non-Ramanujan graphs for which in the limit of large N the eigenratio parameter γ stays below the bound.

- 2) Generic graphs for which $\liminf_{N \rightarrow \infty} \lambda_2(A) \geq \liminf_{N \rightarrow \infty} |\lambda_N(A)|$. Now the bound (43) does not help in asserting that Ramanujan graphs have faster convergence than these generic graphs.

This is because

$$\frac{k - 2\sqrt{k-1}}{k + 2\sqrt{k-1}} < \frac{k - 2\sqrt{k-1}}{k}$$

i.e., the lower bound (47) for Ramanujan graphs is smaller than the upper bound (43) for generic graphs. We should note that the ratio of two quantities is usually much more sensitive to variations in the numerator than to variations of the denominator. Because Ramanujan graphs optimize the algebraic connectivity of the graph, i.e., $\lambda_2(L)$, we still expect γ to be much larger for Ramanujan graphs than for these graphs. We show in Section VI this to be true for broad classes of graphs, including, structured graphs, small-world graphs, and Erdős-Renýi random graphs.

C. Ramanujan graphs: Explicit Algebraic Construction

We now provide explicit constructions of Ramanujan graphs available in the literature. We refer the reader to the Appendix for the definitions of the various terms used in this section. The explicit constructions presented next are based on the construction of Cayley graphs. The following paragraph gives a brief overview of the Cayley graph construction.

Cayley Graphs. The Cayley graph construction gives a simple procedure for constructing k -regular graphs using group theory. Let X be a finite group with $|X| = N$, and S a k -element subset of X . For the graphs used in this paper, we assume that S is a symmetric subset of X , in the sense that $s \in S$ implies $s^{-1} \in S$. We now construct a graph $G = G(X, S)$ by having the vertex set to be the elements of X , with (u, v) as an edge if and only if $vu^{-1} \in S$. It can be easily verified that, for a symmetric subset S , the graph constructed above is k -regular on $|X|$ vertices. The subset S is often called the set of generators of the Cayley graph G , over the group X . Explicit constructions of Ramanujan graphs for a fixed k and varying N , [24], have been described for the cases $k - 1$ is a prime, [22], [25], or a prime power, [26]. The Ramanujan graphs used in this paper are obtained using the Lubotzky-Phillips-Sarnak (LPS) construction, [22]. We describe two constructions of non-bipartite Ramanujan graphs in this section, [22], and refer to them as LPS-I and LPS-II, respectively.

LPS-I Construction. We consider two unequal primes p and q , congruent to 1 modulo 4, and further let the Legendre symbol $\left(\frac{p}{q}\right) = 1$. The LPS-I graphs are Cayley graphs over the $\text{PSL}(2, \mathbb{Z}/q\mathbb{Z})$ group (Projective Special Linear group over the field of integers modulo q .) (Precise definitions and explanations of these terms are provided in the Appendix.) Hence, in this case, the group X is the $\text{PSL}(2, \mathbb{Z}/q\mathbb{Z})$ group. It can be shown that the number of elements in X is given by

$$|X| = \frac{q(q^2 - 1)}{2},$$

see [22]. To get the symmetric subset S of generators, we consider the equation,

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = p,$$

where a_0, a_1, a_2, a_3 are integers. Let

$$\beta = (a_0, a_1, a_2, a_3),$$

be a solution of the above equation. From a formula by Jacobi, [27], there are a total of $8(p+1)$ solutions of this equation, and, out of them, $p+1$ solutions are such that $a_0 > 0$ and odd, and a_j even for $j = 1, 2, 3$. Also, let i be an integer satisfying

$$i^2 \equiv -1 \pmod{q}.$$

For each of these $p+1$ solutions, β , we define the matrix $\tilde{\beta}$ in $\text{PSL}(2, \mathbb{Z}/q\mathbb{Z})$ as,

$$\tilde{\beta} = \begin{pmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{pmatrix} \quad (48)$$

The Appendix shows that these $p+1$ matrices belong to the $\text{PSL}(2, \mathbb{Z}/q\mathbb{Z})$ group. These $p+1$ matrices constitute the subset S , and S acts on the $\text{PSL}(2, \mathbb{Z}/q\mathbb{Z})$ group to produce the $p+1$ -regular Ramanujan graphs on $\frac{1}{2}q(q^2 - 1)$ vertices. The Ramanujan graphs thus obtained are non-bipartite, see [22]. As an example of a LPS-I graph, we may choose $p = 17$ and $q = 13$. We note that p and q are congruent to 1 modulo 4, and the Legendre symbol $\left(\frac{17}{13}\right) = 1$. The LPS-I graph with these values of p and q will be a regular graph with degree $k = p + 1 = 18$ and has $\frac{q(q^2-1)}{2} = 1092$ vertices.

The only problem with the LPS-I graphs is that the number of vertices grows as $O(q^3)$, which limits the use of such graphs. In the next section the explicit construction of a second-class of Ramanujan graphs is presented that avoids this difficulty.

LPS-II Construction. The LPS-II graphs are obtained in a slightly different way. Here also, we start

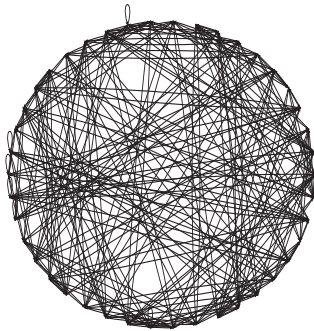


Fig. 1. LPS-II graph with number of vertices $N = 42$ and degree $k = 6$.

with two unequal primes p and q congruent to 1 mod 4, such that the Legendre symbol $\left(\frac{p}{q}\right) = 1$. We define the set $P^1(F_q) = \{0, 1, \dots, q-1, \infty\}$, called Projective line over F_q , and which is basically the set of integers modulo q , with an additional “infinite” element inserted in it. It follows that $|P^1(F_q)| = q+1$. The LPS-II graphs are produced by the action of the set S of the $p+1$ generators defined above (LPS-I) on $P^1(F_q)$, in a linear fractional way. More information about linear fractional transformations is provided in the Appendix. The Ramanujan graphs obtained in this way, are non-bipartite $p+1$ -regular graphs on $q+1$ vertices [22]. The LPS-II graphs thus obtained, may few loops [28], which does not pose any problem because their removal does not affect the Laplacian matrix and hence its spectrum in any way (this is because the Laplacian $L = D - A$, and a loop at vertex n adds the same term to both D_{nn} and A_{nn} , which gets canceled while taking the difference.) The LPS-II offers a larger family of Ramanujan graphs than LPS-I, because in the former, the number of vertices grows only linearly with q . As an example of a LPS-II Ramanujan graph, we take $p = 5$ and $q = 41$. (It can be verified that $p, q \equiv 1 \pmod{4}$ and the Legendre symbol, $\left(\frac{p}{q}\right) = 1$.) Thus, we have a non-bipartite Ramanujan graph, which is 6-regular and has 42 vertices. Fig. 1 shows the graph, thus obtained.

V. DISTRIBUTED INFERENCE

In this Section, we apply the average-consensus algorithm to inference in sensor networks, in particular, to detection. This continues our work in [8], [9] where we compared small-world topologies to Erdős-Rényi random graphs and structured graphs. Subsection V-A formalizes the problem and Subsection V-B presents the noise analysis.

A. Distributed Detection

We study in this Section the simple binary hypothesis test where the state of the environment takes one of two possible alternatives, H_0 (target absent) or H_1 (target present). The true state H is monitored by a network G of N sensors. These collect measurements $\mathbf{y} = (y_1, \dots, y_N)$ that are independent and identically distributed (i.i.d.) conditioned on the true state H ; their known conditional probability density is $f_i(y) = f(y|H_i)$, $i = 0, 1$. We first consider a parallel architecture where the sensors communicate to a single fusion center their local decisions.

Each sensor v_n , $n = 1, \dots, N$, starts by computing the (local) log-likelihood ratio (LLR)

$$r_n = \ln \frac{\Pr(y_n|H_1)}{\Pr(y_n|H_0)} \quad (49)$$

of its measurement y_n . The local decisions are then transmitted to a fusion center. The central decision is

$$\ell = \frac{1}{N} \sum_{n=1}^N r_n \underset{\hat{H}=0}{\overset{\hat{H}=1}{\geq}} v \quad (50)$$

where v denotes an appropriate threshold derived for example from a Bayes' criteria that minimizes the average probability of error P_e .

To be specific, we consider the simple binary hypothesis problem

$$H_m : y_n = \mu_m + \xi_n, \xi_n \sim \mathcal{N}(0, \sigma^2), m = 0, 1 \quad (51)$$

where, without loss of generality, we let $\mu_1 = -\mu_0 = \mu$.

Parallel architecture: fusion center. Under this model, the local likelihoods r_n are also Gaussian, i.e.,

$$H_m : r_n \sim \mathcal{N}\left(\frac{2\mu\mu_m}{\sigma^2}, \frac{4\mu^2}{\sigma^2}\right) \quad (52)$$

From (50), the test statistic for the parallel architecture fusion center is also Gauss

$$H_m : \ell \sim \mathcal{N}\left(\frac{2\mu\mu_m}{\sigma^2}, \frac{4\mu^2}{N\sigma^2}\right) \quad (53)$$

The error performance of the minimum probability of error P_e Bayes' detector (threshold $v = 0$ in (50)) is given by

$$P_e = \operatorname{erfc}^*\left(\frac{d}{2}\right) = \int_{d/2}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad (54)$$

where the equivalent signal to noise ratio d^2 that characterizes the performance is given by, [29],

$$d = \frac{2\mu\sqrt{N}}{\sigma} \quad (55)$$

Distributed detection. We now consider a distributed solution where the sensor nodes reach a global common decision \hat{H} about the true state H based on the measurements collected by all sensors but through *local exchange* only of information over the network G . By *local exchange*, we mean that the sensor nodes do not have the ability to *route* their data to parts of the network other than their immediate neighbors. Such algorithms are of course of practical significance when using power and complexity constrained sensor nodes since such sensor networks may not be able to handle the high costs associated with routing or flooding techniques. We apply the average-consensus algorithm described in Section III-A. This distributed average-consensus detector achieves asymptotically (in the number of iterations) the same optimal error performance P_e of the parallel architecture given by (54), see [8], [9].

Actually, we consider a more general problem than the average-consensus algorithm in (13), namely, we assume that the communications among sensors is through noisy channels. Let the network state, i.e., the likelihood vector, at iteration i be $\mathbf{x}_i \in \mathbb{R}^N$. We modify (13), by taking into account the communication channel noise in each iteration. The distributed detection average-consensus algorithm is modeled by

$$\mathbf{x}_{i+1} = W\mathbf{x}_i + \mathbf{n}_i \quad (56)$$

The weight matrix is as given by (17) using the weight in (18)

$$W = I - \frac{2}{\lambda_2(L) + \lambda_N(L)}L \quad (57)$$

The initial condition \mathbf{x}_0 that collects the local LLRs r_n given in (14), herein repeated,

$$\mathbf{x}_0 = [r_1 \cdots r_N]^T$$

has statistics

$$H_m : \mathbf{x}_0 \sim \mathcal{N}\left(\frac{2\mu\mu_m}{\sigma^2}\mathbf{1}, \Sigma_0 = \frac{4\mu^2}{\sigma^2}I\right), \quad m = 0, 1 \quad (58)$$

The communications noise at iteration i is zero mean Gauss white noise with covariance R given by

$$\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, R) \quad (59)$$

$$R = \text{diag}[\phi_1^2, \dots, \phi_N^2] \quad (60)$$

The *communication channel* noise \mathbf{n}_i is assumed to be independent of the *measurement* noise $\xi_n, \forall i, n$.

The final decision at each sensor is

$$x_n(i) \underset{\hat{H}(n)=0}{\overset{\hat{H}(n)=1}{\gtrless}} v$$

where $\hat{H}(n)$ denotes the decision of sensor v_n .

B. Noise Analysis

In this Subsection we carry out the statistical analysis of the distributed average-consensus detector.

Theorem 4 The local state $x_n(i)$ has mean

$$H_m : E[x_n(i)] = \frac{2\mu\mu_m}{\sigma^2} \quad (61)$$

where $E[\cdot]$ stands for the expectation operator and μ_m is either $\mu_1 = \mu$ or $\mu_0 = -\mu$.

Proof: From the distributed detection (56)

$$x_n(i) = \sum_{j=1}^N (W^i)_{n,j} r_j \quad (62)$$

Hence,

$$E[x_n(i)] = \frac{2\mu\mu_m}{\sigma^2} \sum_{j=1}^N (W^i)_{n,j} \quad (63)$$

It follows:

$$\begin{aligned} \sum_{j=1}^N (W^i)_{n,j} &= (W^i \mathbf{1})_{n,1} \\ &= \mathbf{1}_{n,1} \\ &= 1 \end{aligned} \quad (64)$$

(since $\mathbf{1}$ is an eigenvector of W with eigenvalue 1, it is also an eigenvector of W^i with eigenvalue 1.)

Replacing this result in (63) leads to the Theorem and (61). \blacksquare

We now consider the variance $\text{var}_n(i)$ of the state $x_n(i)$ of the sensor n at iteration i . The following Theorem provides an upper bound.

Theorem 5 The variance $\text{var}_n(i)$ of the state $x_n(i)$ of the sensor n at iteration i is bounded by

$$\text{var}_n(i) \leq \frac{4\mu^2}{\sigma^2} \left[\frac{1}{N} + \gamma_2^{2i} \left(1 - \frac{1}{N} \right) \right] + \phi_{\max}^2 \left[\frac{i}{N} + \frac{1 - \gamma_2^{2i}}{1 - \gamma_2^2} \left(1 - \frac{1}{N} \right) \right] \quad (65)$$

where γ_2 is given in (35).

Proof: Let the covariance of the network state at iteration i be

$$\Sigma_i = \text{covar}\{\mathbf{x}_i\}$$

From eqn. (56) and using standard stochastic processes analysis

$$\Sigma_i = W^i \Sigma_0 W^i + \sum_{k=0}^{i-1} W^k R W^k \quad (66)$$

Thus the variance at the n -th sensor is given by,

$$\text{var}_n(i) = (W^i \Sigma_0 W^i)_{n,n} + \sum_{k=0}^{i-1} (W^k R W^k)_{n,n} \quad (67)$$

Let $\mathbf{w}_j^{(k)}$ be the columns of W^k , $j \in [1, \dots, N]$. Then,

$$W^k R W^k = \sum_{j=1}^N \phi_j^2 \mathbf{w}_j^{(k)} \mathbf{w}_j^{(k)T} \quad (68)$$

It follows

$$(W^k R W^k)_{n,n} = \sum_{j=1}^N \phi_j^2 (w_{j,n}^{(k)})^2 \quad (69)$$

where $w_{j,n}^{(k)}$ represents the n -th component of the vector $\mathbf{w}_j^{(k)}$. Denote by

$$\phi_{\max} = \max(\phi_1, \dots, \phi_N)$$

From eqn. (69), we get

$$\begin{aligned} (W^k R W^k)_{n,n} &\leq \phi_{\max}^2 \sum_{j=1}^N (w_{j,n}^{(k)})^2 \\ &= \phi_{\max}^2 (W^{2k})_{n,n} \end{aligned} \quad (70)$$

We now use the eigendecomposition of W in (20). This leads to

$$W^{2k} = \sum_{m=1}^N \gamma_m^{2k} \mathbf{u}_m \mathbf{u}_m^T \quad (71)$$

from which

$$\begin{aligned}
(W^{2k})_{n,n} &= \sum_{m=1}^N \gamma_m^{2k} (\mathbf{u}_{m,n})^2 \\
&= \frac{1}{N} + \sum_{m=2}^N \gamma_m^{2k} (\mathbf{u}_{m,n})^2 \\
&\leq \frac{1}{N} + \gamma_2^{2k} \sum_{m=2}^N (\mathbf{u}_{m,n})^2 \\
&= \frac{1}{N} + \gamma_2^{2k} \left(1 - \frac{1}{N}\right)
\end{aligned} \tag{72}$$

Hence, from eqn. (70),

$$\begin{aligned}
(W^k R W^k)_{n,n} &\leq \phi_{\max}^2 (W^{2k})_{n,n} \\
&\leq \phi_{\max}^2 \left[\frac{1}{N} + \gamma_2^{2k} \left(1 - \frac{1}{N}\right) \right]
\end{aligned} \tag{73}$$

Through a similar set of manipulations,

$$\begin{aligned}
(W^i \Sigma_0 W^i)_{n,n} &= \frac{4\mu^2}{\sigma^2} (W^{2i})_{n,n} \\
&\leq \frac{4\mu^2}{\sigma^2} \left[\frac{1}{N} + \gamma_2^{2i} \left(1 - \frac{1}{N}\right) \right]
\end{aligned} \tag{74}$$

Finally from eqn. (67) we obtain,

$$\begin{aligned}
\text{var}_n(i) &\leq \frac{4\mu^2}{\sigma^2} \left[\frac{1}{N} + \gamma_2^{2i} \left(1 - \frac{1}{N}\right) \right] + \sum_{k=0}^{i-1} \phi_{\max}^2 \left[\frac{1}{N} + \gamma_2^{2k} \left(1 - \frac{1}{N}\right) \right] \\
&= \frac{4\mu^2}{\sigma^2} \left[\frac{1}{N} + \gamma_2^{2i} \left(1 - \frac{1}{N}\right) \right] + \phi_{\max}^2 \sum_{k=0}^{i-1} \left[\frac{1}{N} + \gamma_2^{2k} \left(1 - \frac{1}{N}\right) \right] \\
&= \frac{4\mu^2}{\sigma^2} \left[\frac{1}{N} + \gamma_2^{2i} \left(1 - \frac{1}{N}\right) \right] + \phi_{\max}^2 \left[\frac{i}{N} + \frac{1 - \gamma_2^{2i}}{1 - \gamma_2^2} \left(1 - \frac{1}{N}\right) \right]
\end{aligned} \tag{75}$$

which gives an upper bound on the variance of the n -th sensor at iteration i and proves Theorem 5. ■

If the channels are noiseless, we immediately obtain a Corollary to Theorem 5 that bounds the variance of the state of sensor n at iteration i .

Corollary 6 With noiseless communication channels, the variance of the state of sensor n at iteration i is bounded by

$$\text{var}_n(i) \leq \frac{4\mu^2}{\sigma^2} \left[\frac{1}{N} + \gamma_2^{2i} \left(1 - \frac{1}{N}\right) \right] \tag{76}$$

We now interpret Theorems 4 and 5, and Corollary 6. Theorem 4 shows that the mean of the local state is the same as the mean of the global statistic ℓ of the fusion center in the parallel architecture. Then to compare the local probability of error $P_e(i, n)$ at sensor n and iteration i in the distributed detector with the probability of error P_e of the fusion center in the parallel architecture we need to compare the variances of the sufficient statistics in each detector. With noiseless communication channels, we see that the upper bound in (76) in Corollary 6 converges to We now interpret Theorems 4 and 5, and Corollary 6. Theorem 4 shows that the mean of the local state is the same as the mean of the global statistic ℓ of the fusion center in the parallel architecture. Then to compare the local probability of error $P_e(i, n)$ at sensor n and iteration i in the distributed detector with the probability of error P_e of the fusion center in the parallel architecture we need to compare the variances of the sufficient statistics in each detector. With noiseless communication channels, we see that the upper bound in (76) in Corollary 6 converges to

$$\frac{4\mu^2}{\sigma^2} \left[\frac{1}{N} + \gamma_2^{2i} \left(1 - \frac{1}{N} \right) \right] \rightarrow \frac{4\mu^2}{N\sigma^2}$$

which is the variance of the parallel architecture test statistic (50). This shows that

$$\lim_{i \rightarrow \infty} P_e(i, n) = P_e \quad (77)$$

The rate of convergence is again controlled by

$$\gamma_2^{2i} = \left(1 - \frac{2\lambda_2(L)}{\lambda_2(L) + \lambda_N(L)} \right)^{2i}$$

and maximizing this rate is equivalent to minimizing γ_2 , which in turn, see (35), is equivalent to maximizing the eigenratio parameter $\gamma = \lambda_2(L)/\lambda_N(L)$ like for the average-consensus algorithm.

For noisy channels, it is interesting to note that there is a linear trend $\phi_{\max}^2 i/N$ that makes $\text{var}_n(i)$ to become arbitrarily large as the number of iterations i grows to ∞ . We no longer have the convergence of the probability of error $P_e(i, n)$ as in (77). The average minimum probability of error is still given by (54), with now the equivalent SNR parameter d^2 bounded below by Theorem 5.

C. Optimal number of iterations

With noisy communication channels, the performance of the distributed detector no longer achieves the performance of the fusion center in a parallel architecture. This is no surprise, since each iteration corrupts the inter communicated state of the sensor. However, there is an interesting tradeoff between sensing signal to noise ratio (S-SNR) and the communication noise. Intuitively, the local sensors perceive

better the global state of the environment as they obtain information through their neighbors from more remote sensors. However, this new information is counter balanced by the additional noise introduced by the communication links. This leads to an interesting tradeoff that we now exploit and leads to an optimal number of iterations to carry out the consensus through noisy channels before a decision is declared by each sensor.

The upper bound in eqn. (75) is a function of the number of iterations i . We rewrite it, replacing the integer valued iteration number i by a continuous variable z , as

$$f(z) = \left(\frac{4\mu^2}{N\sigma^2} + \frac{\phi_{\max}^2 (1 - \frac{1}{N})}{1 - \gamma_2^2} \right) + \left(1 - \frac{1}{N} \right) \left(\frac{4\mu^2}{\sigma^2} - \frac{\phi_{\max}^2}{1 - \gamma_2^2} \right) \gamma_2^{2z} + \frac{\phi_{\max}^2}{N} z \quad (78)$$

We consider only the case when

$$\frac{4\mu^2}{\sigma^2} > \frac{\phi_{\max}^2}{1 - \gamma_2^2} \quad (79)$$

This is reasonable. For example, if $\frac{4\mu^2}{\sigma^2} > \phi_{\max}^2$, which is the case when the communication noise is smaller than the equivalent sensing noise power and iterating among sensors can be reasonably expected to improve upon decisions based solely on the local measurement. Secondly, if γ_2 , which is bounded above by 1, is small, then the right-hand-side of (79) is more likely to be satisfied. This means that topologies like the Ramanujan graphs where γ_2 is minimized (which, from (35) means that the eigenratio parameter γ is maximized) will satisfy better this assumption.

We now state the result on the number of iterations.

Theorem 7 If (79) holds, $f(z)$ has a global minimum at

$$z^* = \frac{1}{2\ln \gamma_2} \ln \left(\frac{\phi_{\max}^2}{\left(2\ln \frac{1}{\gamma_2} \right) (N - 1) \left(\frac{4\mu^2}{\sigma^2} - \frac{\phi_{\max}^2}{1 - \gamma_2^2} \right)} \right) \quad (80)$$

Proof: When (79) holds, $f(z)$ is convex. Hence there exists a global minimum, say attained at z^* .

We find z^* by rooting the first derivative, successively obtaining

$$\frac{df}{dz}(z^*) = (2\ln \gamma_2) \left(1 - \frac{1}{N} \right) \left(\frac{4\mu^2}{\sigma^2} - \frac{\phi_{\max}^2}{1 - \gamma_2^2} \right) \gamma_2^{2z^*} + \frac{\phi_{\max}^2}{N} = 0 \quad (81)$$

$$\gamma_2^{2z^*} = - \frac{\phi_{\max}^2}{(N - 1) (2\ln \gamma_2) \left(\frac{4\mu^2}{\sigma^2} - \frac{\phi_{\max}^2}{1 - \gamma_2^2} \right)} \quad (82)$$

$$z^* = \frac{1}{2\ln \gamma_2} \ln \left(\frac{\phi_{\max}^2}{\left(2\ln \frac{1}{\gamma_2} \right) (N - 1) \left(\frac{4\mu^2}{\sigma^2} - \frac{\phi_{\max}^2}{1 - \gamma_2^2} \right)} \right) \quad (83)$$

■

From Theorem 7, we conclude that, if $z^* > 0$, then the variance upper bound will decrease till $i^* = \lfloor z^* \rfloor$. The iterative distributed detection algorithm should be continued till i^* if

$$\min(f(\lfloor z^* \rfloor), f(\lceil z^* \rceil)) < \text{var}_n(0) = \frac{4\mu^2}{\sigma^2} \quad (84)$$

Numerical Examples. We illustrate Theorem 7 with two numerical examples. We consider a network of $N = 1,000$ sensors, $\mu^2/\sigma^2 = 1$ (0 db), and $\gamma_2 = .7$. The initial likelihood variance before fusion is $\text{var}_n(0) = 4$. We first consider $\phi_{\max} = .1$. Then, $z^* = 17.6$ and $\text{var}_n(17) \leq f(\lfloor z^* \rfloor) = .0238 = f(\lceil z^* \rceil)$. The variance reduction achieved with iterative distributed detection over the single measurement decision is $\frac{\text{var}_n(0)}{\text{var}_n(i^*)} \geq 168 = 22$ dB, a considerable improvement. We now consider a second case where the communication noise is $\phi_{\max} = .3162$. It follows that $z^* = 14.3$, and the improvement by iterating till $i^* = 14$ with the distributed detection is $\frac{\text{var}_n(0)}{\text{var}_n(14)} \geq 20 = 13$ dB.

VI. EXPERIMENTAL RESULTS

This section shows how Ramanujan graph topologies outperform other topologies. We first describe the graph topologies to be contrasted with the Ramanujan LPS-II constructions described in Section IV. We start by defining the average degree k_{avg} of a graph G as

$$k_{\text{avg}} = \frac{2|E|}{|V|}$$

where $|E| = M$ denotes the number of edges and $|V| = N$ is the number of vertices of the graph. In this section, we use the symbols and terms k and k_{avg} interchangeably. For, k -regular graphs, it follows that $k_{\text{avg}} = k$. This means, that, when we work with general graphs, k refers to the average degree, while with regular graphs, it refers to both the average degree and the degree of each vertex.

A. Structured graphs, Watts-Strogatz Graphs, and Erdős-Rényi Graphs

We compare Ramanujan graphs, which are regular graphs, with regular and non regular graphs. The symbol k will stand in this Section for the degree of the graph for regular graphs and for the average degree for non regular graphs. We describe briefly the three classes of graphs used to benchmark the Ramanujan graphs. Structured graphs usually have high clustering but large average path length. Erdős-Rényi graphs are random graphs, they have small average path length but low clustering. Small-world graphs generated with a rewiring probability above a phase transition threshold have both high clustering and small average path length.

Structured graphs: Regular ring lattice (RRL). This is a highly structured network. The nodes are numbered sequentially (for simplicity, display them uniformly placed on a ring.) Starting from node # 1,

connect each node to $k/2$ nodes to the left and $k/2$ nodes to the right. The resulting graph is regular with degree k .

Small world networks: Watts-Strogatz (WS-I). We explain briefly the Watts-Strogatz construction of a small world network, [15]. It starts from a highly structured regular network where the nodes are placed uniformly around a circle, with each node connected to its k nearest neighbors. Then, random rewiring is conducted on all graph links. With probability p_w , a link is rewired to a different node chosen uniformly at random. Notice that the p_w parameter controls the “randomness” of the graph in the sense that $p_w = 0$ corresponds to the original highly structured network while $p_w = 1$ results in a random network. Self and parallel links are prevented in the rewiring procedure and the number of links is kept constant, regardless of the value of p_w . In [8], distributed detection was studied with two slightly different versions of the Watts-Strogatz model. In both versions, the rewiring procedure is such that the nodes are considered one by one in a fixed direction along the circle (clockwise or counter clockwise.) For each node, the $k/2$ edges connecting it to the following nodes (in the same direction) are rewired with probability p_w . In the first version of the Watts-Strogatz model, called Watts-Strogatz-I (WS-I) in the sequel, the edges are kept connected to the current node while their other ends are rewired with probability p_w . In the second version, called Watts-Strogatz-II (WS-II), the particular vertex to be disconnected is chosen randomly between the two ends of the rewired edges. It was shown in [8] that the WS-I graphs yield better convergence rates among the different models of small world graphs considered in that paper (WS-I, WS-II, and the Kleinberg model, [16], [17].) Hence, we restrict attention here to WS-I graphs.

Erdős-Renýi random graphs (ER). In these graphs, we randomly choose $\frac{Nk}{2}$ edges out of a total of $\frac{N(N-1)}{2}$ possible edges. These are not regular graphs, their degree distribution follows a binomial distribution, which in the limit of large N approaches the Poisson law.

B. Comparison Studies

We present numerical studies that will show the superiority of the Ramanujan graphs (RG) over the other three classes of graphs: Regular ring lattice (RRL), Watts-Strogatz-I (WS-I), and Erdős-Renýi (ER) graphs. We carry out three types of comparisons: (1) Convergence speed S_c ; (2) The γ parameters for the RG and each of the other three classes of graphs; (3) The algebraic connectivity $\lambda_2(L)$ for the RG and each of the other three classes of graphs. In Section V, we considered a distributed detection problem based on the average-consensus algorithm. Here we present results for the noiseless link case. We define the convergence time T_c of the distributed detector, as the number of iterations required to reach within 10% of the global probability of error, averaged over all sensor nodes. Rather than using T_c , the results

are presented in terms of the convergence speed, $S_c = 1/T_c$. To simplify the comparisons, we subscript the γ parameter by the corresponding acronym, e.g., γ_{RG} to represent the eigenratio of the Ramanujan graph. We also define the following comparison parameters

$$\psi(\text{RRL}) = \frac{S_{c, \text{RG}}}{S_{c, \text{RRL}}}, \quad \nu(\text{RRL}) = \frac{\gamma_{\text{RG}}}{\gamma_{\text{RRL}}}, \quad \text{and} \quad \eta(\text{RRL}) = \frac{\lambda_{2, \text{RG}}(L)}{\lambda_{2, \text{RRL}}(L)} \quad (85)$$

Ramanujan graphs and regular ring lattices. Fig. 2 compares RG with RRL graphs. The panel on the right plots $\psi(\text{RRL})$, the center panel displays $\nu(\text{RRL})$, and the right panel shows $\eta(\text{RRL})$ when

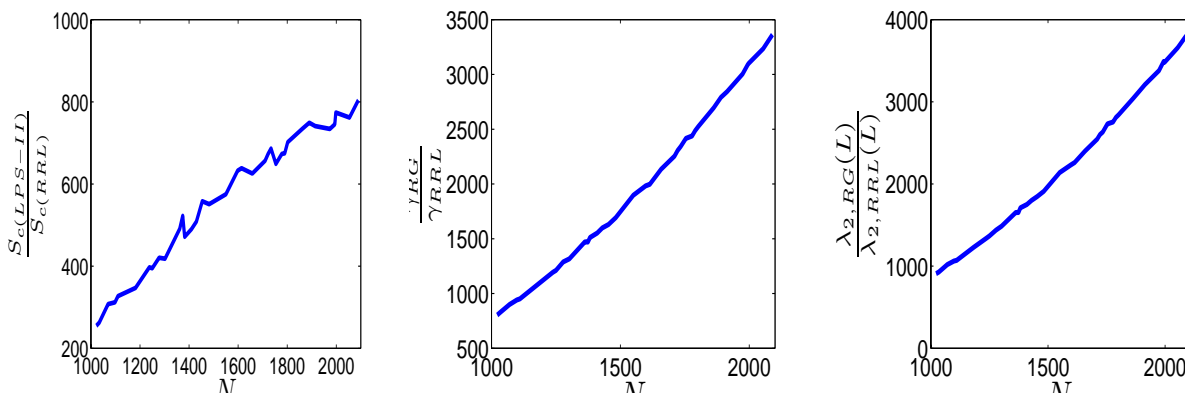


Fig. 2. Spectral properties of LPS-II and RRL graphs, $k = 18$, varying N : Left: Ratio of convergence speed $\psi(\text{RRL})$; Center: Ratio $\nu(\text{RRL})$ of $\frac{\lambda_2(L)}{\lambda_N(L)}$; Right: Ratio $\eta(\text{RRL})$ of $\lambda_2(L)$.

the degree $k = 18$ and the number of nodes N varies. We conclude that the RGs converge 3 orders of magnitude faster than the RRLs, the γ parameters can be up to 3,500 times faster, and the algebraic connectivity for the RGs can be up to 4,000 times larger than for the RRLs.

Ramanujan graphs and Watts-Strogatz graphs. Fig. 3 contrasts the RG with the WS-I graphs. Because the WS-I graphs are randomly generated, we fix the number of nodes $N = 6038$ and the degree $k = 18$ and vary on the horizontal axis the rewiring probability $0 \leq p_w \leq 1$. The Figure shows on the left panel the convergence speed S_c . The top horizontal line is S_c for the RG—it is flat because the graph is the same regardless of p_w . The three lines below correspond to the WS-I topologies. For each value of p_w , we generate 150 WS-I graphs. Of the WS-I three lines, the top line corresponds, at each p_w , to the topologies (among the 150 generated) with maximum convergence rate, the medium line to the average convergence rate (averaged over the 150 random topologies generated), and the bottom line to the topologies (among the 150 generated) with worst convergence rate. Similarly, the center and right panels on Fig. 3 compare the eigenratio parameters γ (center panel) and the algebraic connectivity λ_2

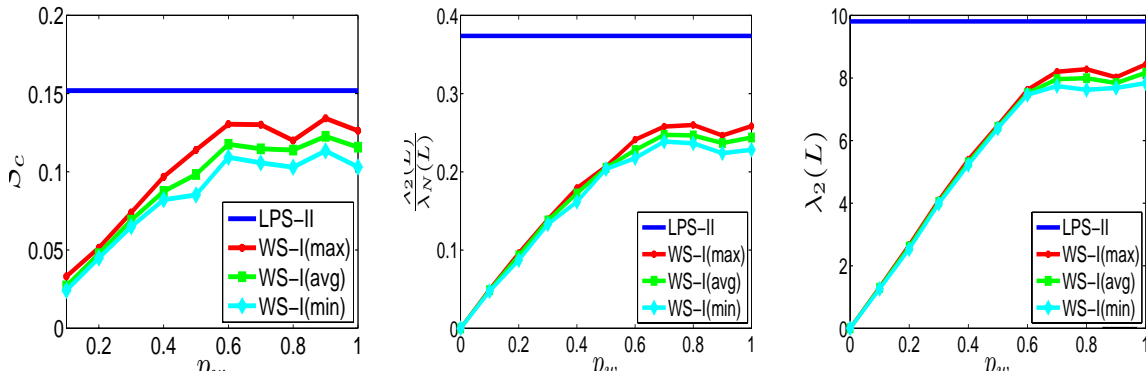


Fig. 3. Spectral properties of LPS-II and WS-I graphs, $N = 6038$, $k = 18$, varying p_w . Left: S_c ; Center: eigenratio $\gamma = \frac{\lambda_2(L)}{\lambda_N(L)}$; Right: algebraic connectivity λ_2 .

(right panel). For example, the RG improves by 50 % the γ eigenratio over the best WS-I topology (in this case for $p_w = .8$.)

Ramanujan graphs and Erdős-Renýi graphs. We conclude this section by comparing the LPS-II graphs with the Erdős-Renýi graphs in Figs. 4 and 4. Fig. 4 shows the results for topologies with different number of sensors N (plotted in the horizontal axis.) For each value of N , we generated 200 random Erdős-Renýi graphs. In the panels of both Figures, the top line illustrates the results for the RG, while the three lines below show the results for the Erdős-Renýi graphs—among these three, the top line is the topology with best convergence rate among the 200 ER topologies, the middle plot is the averaged convergence rate, averaged over the 200 topologies, and the bottom line corresponds to the worst topologies. Again, for example, the γ parameter of the RG is about twice as large than the γ parameter for the ER.

VII. RANDOM REGULAR RAMANUJAN-LIKE GRAPHS

Section IV-C explains the construction of the Ramanujan graphs. These graphs can be constructed only for certain values of N , which may limit their application in certain practical scenarios. We describe here briefly biased random graphs that can be constructed with arbitrary number of nodes N and average degree, and whose performance closely matches that of Ramanujan graphs. Reference [30] argues that, in general, heterogeneity in the degree distribution reduces the eigenratio $\gamma = \frac{\lambda_2(L)}{\lambda_N(L)}$. Hence, we try to construct graphs that are regular in terms of the degree. There exist constructions of random regular graphs, but these are difficult to implement especially for very large number of vertices, see, e.g., [31], [32], [33], [34], which are good references on the construction and application of random regular graphs.

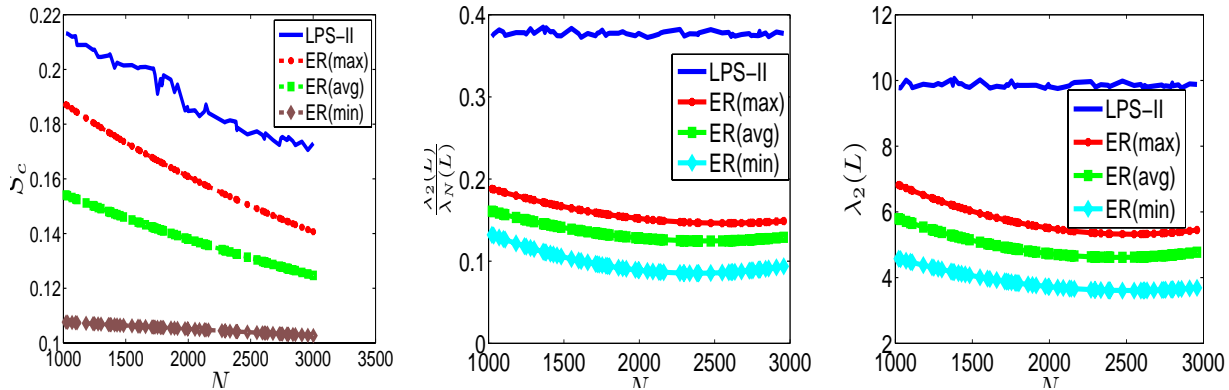


Fig. 4. Spectral properties of LPS-II and ER graphs, $k = 18$, varying N : Left: Convergence speed S_c ; Center: eigenratio $\gamma = \frac{\lambda_2(L)}{\lambda_N(L)}$; Right: algebraic connectivity λ_2 .

Ours is a procedure that is simple to implement and constructs random regular graphs, which we refer to as Random Regular Ramanujan-Like (R3L) graphs. Suppose, we want to construct a random regular graph with N vertices and degree k . Our construction starts from a regular graph of degree k , which we call the seed. The seed can be any regular graph of degree k , for example, the regular ring lattice with degree k (see Section VI.) We start by randomly choosing (uniformly) a vertex (call it v_1 .) In the next step, we randomly choose a neighbor of v_1 (call it v_2), and we also randomly choose a vertex not adjacent to v_1 (call it v_3 .) We now choose a neighbor of v_3 (call it v_4). The next step consists of removing the edges between v_1 and v_2 , and between v_3 and v_4 . Finally we add edges between v_1 and v_3 and between v_2 and v_4 . (Care is taken so that no conflict arises in the process of removing and forming the edges.) It is quite clear that after this sequence of steps, the degree of each vertex remains the same and hence the resulting graph remains k -regular. We repeat this sequence of steps a sufficiently large number of times, which makes the resulting graph to become random. Thus, starting with any k -regular graph, we get a random regular graph with degree k .

We now present numerical studies of the R3L graphs, which show that these graphs have convergence properties that are very close to those of LPS-II graphs. Specifically, we focus on the eigenratio parameter $\gamma = \frac{\lambda_2(L)}{\lambda_N(L)}$. Fig. 5 plots the eigenratio $\gamma = \frac{\lambda_2(L)}{\lambda_N(L)}$ for the RG and the R3L graphs for varying number of nodes N and degree $k = 18$. We generate 100 R3L graphs for each value of N . The top three lines correspond to the RG, the best R3L topologies, and the average value of γ over the 100 R3L graphs. We observe that the maximum values of $\gamma = \frac{\lambda_2(L)}{\lambda_N(L)}$ are sometimes higher than those obtained with the LPS-II graphs. Note also that, on average, the R3L graphs are quite close to the LPS-II graphs in terms

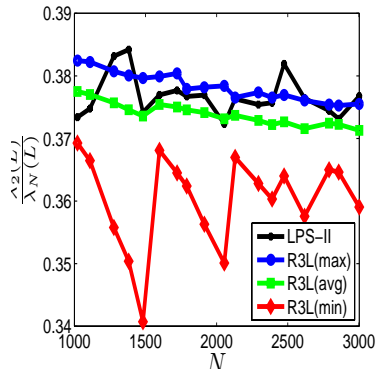


Fig. 5. LPS-II and R3L graphs, $k = 18$, varying N : Eigenratio $\gamma = \frac{\lambda_2(L)}{\lambda_N(L)}$.

of the $\gamma = \frac{\lambda_2(L)}{\lambda_N(L)}$ ratio, even for large values of N . This study shows that the R3L graphs are a good alternative to the LPS-II graphs with the added advantage that they can be generated for arbitrary number of nodes N and degree k .

VIII. CONCLUSION

The paper studies the impact of network topology on the convergence speed of distributed inference and average-consensus. We derive that the convergence speed is governed by a graph spectral parameter, the eigenratio $\gamma = \lambda_2(L)/\lambda_N(L)$ of the second largest and the largest eigenvalues of the graph Laplacian. We show that the class of non-bipartite Ramanujan graphs is essentially optimal. Numerical simulations verify the Ramanujan LPS-II graphs outperform the highly structured graphs, the Erdős-Renýi random graphs, and graphs exhibiting the small-world property. We considered average-consensus and distributed detection with noiseless and noisy links. We derived for the distributed inference problem an analytical upper bound on the likelihood variance. For noiseless links, this bound shows that the local likelihood variances (and hence the local probability of errors) converge to the global likelihood variance (global probability of error) at a rate determined by γ . With noisy links, we demonstrate that there is a maximum, optimal number of iterations before declaring a decision. Finally, we introduced a novel biased construction of random regular graphs (R3L graphs) and showed by numerical results that their convergence performance tracks very closely that of the Ramanujan LPS-II graphs. R3L graphs address a main limitation of Ramanujan graphs that can be constructed only for very restricted number of nodes. In contrast, R3L graphs are simple to construct and can have an arbitrary number of nodes N and degree k .

APPENDIX

Definition 8 (Group) : A group X is a non-empty collection of elements, with a binary operation “.” defined on them, such that the following properties are satisfied:

- 1) If $a, b \in X$, then $a.b \in X$ (closure property.)
- 2) If $a, b, c \in X$, then $a.(b.c) = (a.b).c$ (associative property.)
- 3) There exists an element $e \in X$, such that for any element $a \in X$, $a.e = e.a = a$ (identity element.)
- 4) $\forall a \in X$, there exists $a^{-1} \in X$, the inverse of a , such that $a.a^{-1} = a^{-1}.a = e$ (inverse.)

The group X is called abelian if the “.” operation is commutative, that is, for any $a, b \in X$, $a.b = b.a$.

Definition 9 (Field) : A field F is a non-empty collection of elements, with the following properties:

There exists a binary operation “+” on the elements of F such that,

- 1) If $a, b \in F$, then $a + b \in F$.
- 2) If $a, b \in F$, then $a + b = b + a$.
- 3) If $a, b, c \in F$, then $a + (b + c) = (a + b) + c$.
- 4) There exists an element 0 (zero) $\in F$, such that for any element $a \in F$, $a + 0 = a$.
- 5) If $a \in F$, then there exists an element $(-a) \in F$, such that $a + (-a) = (-a) + a = 0$.

There exists another binary operation “.” on the elements of F such that,

- 1) If $a, b \in F$, then $a.b \in F$.
- 2) If $a, b \in F$, then $a.b = b.a$.
- 3) If $a, b, c \in F$, then $a.(b.c) = (a.b).c$.
- 4) There exists a non-zero element 1 (one) $\in F$, such that for any element $a \in F$, $a.1 = a$.
- 5) For every non-zero element $a \in F$, there exists an element $a^{-1} \in F$, such that $a.a^{-1} = 1$.
- 6) If $a, b, c \in F$, then $a.(b + c) = a.b + a.c$.

Congruence. For integers a, b, c , the statement a is congruent to b modulo c , or $a \equiv b \pmod{c}$ implies that $(a - b)$ is divisible by c .

Quadratic Residue. For integers a, b , the statement a is a quadratic residue modulo b implies that there exists an integer c such that $c^2 \equiv a \pmod{b}$.

Definition 10 (Legendre Symbol) : For an integer a and a prime p , the Legendre symbol $\left(\frac{a}{p}\right)$ is

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \text{ divides } a \\ 1 & \text{if } a \text{ is a quadratic residue modulo } p \\ -1 & \text{if } a \text{ is a quadratic non-residue modulo } p \end{cases} \quad (86)$$

$PSL(2, Z/qZ)$. For a prime q , the set $Z/qZ = \{0, 1, \dots, q-1\}$ is the field of integers modulo q . To define the group $PSL(2, Z/qZ)$ (Projective Special Linear Group), first consider the set of 2×2 matrices over the field Z/qZ , whose determinants are non-zero quadratic residues modulo q . Next, define an equivalence relation on this set, such that two matrices are in the same equivalence class, if one is a non-zero scalar multiple of the other. The $PSL(2, Z/qZ)$ group is then the set of all these equivalence classes. Think of each element of $PSL(2, Z/qZ)$ as a 2×2 matrix over the field Z/qZ , whose determinant is a non-zero quadratic residue modulo q , and whose second row can be represented as either $(0,1)$ or $(1,a)$, where a being any element of Z/qZ , [35]. The $p+1$ generators discussed in the paper, belong to the $PSL(2, Z/qZ)$ group, because their determinants are $p \pmod{q}$ and by assumption, p is a quadratic residue modulo q or $\left(\frac{p}{q}\right) = 1$ for the non-bipartite Ramanujan graphs we use in this paper.

Linear Fractional Transformation. Let $P^1(F_q) = \{0, 1, \dots, q-1, \infty\}$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a 2×2 matrix. Then a linear fractional transformation on $P^1(F_q)$ is defined by the mapping,

$$x \mapsto \frac{ax + b}{cx + d} \pmod{q} \quad (87)$$

for every element $x \in P^1(F_q)$, with the usual assumptions that $\frac{z}{0} = \infty$ for $z \neq 0$, and $\frac{a\infty+b}{c\infty+d} = \frac{a}{c}$.

Definition 11 (Bipartite graph) : A bipartite graph is a graph in which the vertex set can be partitioned into two disjoint subsets, such that no two vertices in the same subset are adjacent.

REFERENCES

- [1] R. R. Tenney and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 98–101, July 1981.
- [2] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *MCSS*, vol. 1, no. 2, pp. 167–182, 1988.
- [3] —, "Decentralized detection," In *"Advances in Statistical Signal Processing: Vol 2 - Signal Detection,"* H. V. Poor, and John B. Thomas, eds., JAI Press, Greenwich, CT., pp. 297–344, Nov. 1993.
- [4] P. K. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: distributed detection of a known signal in dependent gaussian noise," *IEEE Trans. Signal Processing*, vol. 48, p. 32663279, Dec. 2000.
- [5] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer-Verlag, 1996.
- [6] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors: Part II—Advanced topics," *Proc. IEEE*, vol. 85, pp. 64–79, Jan. 1997.
- [7] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Processing*, vol. 51, pp. 407–416, Feb. 2003.
- [8] S. A. Aldosari and J. M. F. Moura, "Distributed detection in sensor networks: connectivity graph and small-world networks," in *Asilomar Conference on Signals, Systems, and Computers*, 2005.

- [9] —, “Topology of sensor networks in distributed detection,” in *ICASSP’06, IEEE International Conference on Signal Processing*, May 2006.
- [10] P. Erdős and A. Rényi, “On random graphs,” *Publ. Math. Debrecen*, vol. 6, pp. 290–291, 1959.
- [11] —, “On the evolution of random graphs,” *Publ. Math. Inst. Hung. Acad. Sciences (Magyar Tud. Akad. Mat. Kutató Int. Közl.)*, vol. 5, pp. 17–61, 1960.
- [12] —, “On the evolution of random graphs,” *Bull. Inst. Internat. Statist.*, vol. 38, pp. 343–347, 1961.
- [13] E. N. Gilbert, “Random graphs,” *Annals of Mathematical Statistics*, vol. 30, pp. 1141–1144, 1959.
- [14] B. Bollobás, *Modern Graph Theory*. New York, NY: Springer Verlag, 1998.
- [15] D. J. Watts and S. H. Strogatz, “Collective dynamics of small-world networks,” *Nature*, vol. 393, pp. 440–442, 1998.
- [16] J. M. Kleinberg, “Navigation in a small world,” *Nature*, vol. 406, p. 845, Aug. 2000.
- [17] —, “The small-world phenomenon: An algorithmic perspective,” in *Proc. of the thirty-second annual ACM symposium on Theory of computing*, vol. 2, Portland, Oregon, 2000, pp. 63–170.
- [18] 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, pp. 1520–1533, 2004.
- [19] M. Fiedler, “Algebraic connectivity of graphs,” *Czechoslovak. Mathematical Journal*, vol. 23, no. 98, pp. 298–305, 1973.
- [20] F. R. K. Chung, *Spectral Graph Theory*. American Mathematical Society, 1997.
- [21] L. Xiao and S. Boyd, “Fast linear iteration for distributed averaging,” *Syst. Contr. Lett.*, vol. 53, pp. 65–78, Sept. 2004.
- [22] A. Lubotzky, R. Phillips, and P. Sarnak, “Ramanujan graphs,” *Combinatorica*, vol. 8, no. 3, pp. 261–277, 1988.
- [23] N. Alon, “Eigenvalues and expanders,” *Combinatorica*, vol. 6, pp. 83–96, 1986.
- [24] M. R. Murty, “Ramanujan graphs,” *J. Ramanujan Math. Soc.*, vol. 18, no. 1, pp. 1–20, 2003.
- [25] G. Margulis, “Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators,” *J. Probl. Inf. Transm.*, vol. 24, no. 1, pp. 39–46, 1988.
- [26] M. Morgenstern, “Existence and explicit construction of $q + 1$ regular Ramanujan graphs for every prime power q ,” *J. Comb. Theory, ser. B*, vol. 62, pp. 44–62, 1994.
- [27] G. Andrews, S. B. Ekhad, and D. Zeilberger, “A short proof of Jacobi’s formula for the number of representations of an integer as a sum of four squares,” *Amer. Math. Monthly*, vol. 100, pp. 273–276, 1993.
- [28] A. Berger and J. Lafferty, “Probabilistic decoding of low-density Cayley codes.”
- [29] H. L. V. Trees, *Detection, Estimation, and Modulation Theory: Part I*. New York, NY: John Wiley & Sons, 1968.
- [30] A. . E. Motter, C. Zhou, and J. Kurths, “Network synchronization, diffusion, and the paradox of heterogeneity,” *Physical Review E*, vol. 71, 2005.
- [31] E. A. Bender and E. R. Canfield, “The asymptotic number of non-negative integer matrices with given row and column sums,” *Journal of Combinatorial Theory, Series A*, vol. 24, pp. 296–307, 1978.
- [32] B. Bollobás, “A probabilistic proof of an asymptotic formula for the number of labelled regular graphs,” *European Journal of Combinatorics*, vol. 1, pp. 311–316, 1980.
- [33] N. C. Wormald, “Generating random regular graphs,” *Journal of Algorithms*, vol. 5, pp. 247–280, 1984.
- [34] —, “Models of random regular graphs,” *London Mathematical Society Lecture Note Series*, vol. 267, pp. 239–298, 1999.
- [35] Y. Kohayakawa, V. Rodl, and L. Thoma, “An optimal algorithm for checking regularity,” *SIAM J. on Computing*, vol. 32, no. 5, pp. 1210–1235, 2003.