On Optimality of Sparse Long-Range Links in Circulant Consensus Networks

Makan Fardad

***Abstract*—We consider spatially invariant consensus networks in which the directed graph describing the interconnection topology, the link weights, and the temporal dynamics, are all characterized by circulant matrices. We seek the best new links, subject to budget constraints, whose addition to the network maximally improves its rate of convergence to consensus. Motivated by small-world networks, we apply the optimal link creation problem to circulant networks with local communication links. We observe that the optimal new links are sparse and long-range, and have an increasingly pronounced effect on the convergence rate of the network as its size grows. To further investigate the properties of optimal links analytically, we restrict attention to the creation of links with small weights, referred to as weak links. We employ perturbation methods to reformulate the problem of optimal weak link creation, and uncover conditions on the network architecture which guarantee sparse and long-range solutions to this optimization problem.**

***Index Terms*—Circulant matrices, consensus, convex optimization, long-range links, perturbation methods, small-world networks, social ties, sparse interconnection topology, weak communication links.**

I. Introduction

Small-world networks [1] have been the topic of active research in the past decade and are observed in biological, social, and technological networks [2]–[4]. Intuitively, these networks are characterized by their regular short-range, and sparse long-range links; nodes are only locally clustered and yet almost any node can be reached from any other via a small number of hops. In an influential paper [5], numerical evidence was used to suggest that the addition of sparse long-range links to regular networks leads to a dramatic improvement in their rate of convergence to consensus. It was thus conjectured that small-world networks could be used to achieve ‘ultrafast consensus’ in a spectrum of applications. Given the hypothesis that the complex structures of real-world networks are optimal or nearly optimal for the function they serve [6, Chap. 14], we are driven to investigate whether networks with sparse long-range links [7]–[11] are optimal or nearly optimal in the way they balance link creation and collective performance.

In this technical note, we consider the problem of optimal link creation for the purpose of consensus facilitation in circulant networks [12], [13]. We frame our network design problems in the context of the DeGroot model [14] as a motivational application and concrete instance of collaborative behavior. We focus on regular networks described by circulant matrices, noting that a common method of generating small-world networks is by augmenting regular lattices [1], [15]. From a physical point of view, a circulant network can be interpreted as a homogenous one where every node represents an aggregation of many agents (e.g., a small community in the case of a social network). Results obtained for circulant networks can provide important insights and guidelines for the design of efficient communication architectures for more general classes of networks [16].

Solutions of the optimal link creation problem indicate that, when the budget for link creation is small, the optimal links have the property of being both *sparse* and *long-range* for large classes of networks with local communication links. Motivated by this observation, we use perturbation methods to analytically investigate the topology of optimal *weak* links, i.e., communication links of small strength. From a physical standpoint, weak links can be interpreted as infrequent communication between nodes [17]. Perturbation methods allow us to uncover mild conditions on the architecture of the original network which guarantee the sparsity and long-range property of optimal weak links.

This work supplements the growing literature on topology design and consensus in dynamical networks. In [18], after assigning a measure of distance between all pairs of nodes, the communication efficiency of a network is defined as the average value of the inverses of node distances, and the cost of a network as the sum of the lengths of its links. These quantities are then used in numerical experiments to show that small-world networks both are efficient at spreading information and are cheap to build. In [5] it is shown that the speed of convergence in an unweighted and undirected consensus network can be improved by orders of magnitude through the random rewiring of links in a ring lattice. In [19] the authors demonstrate that the addition of an Erdős-Rényi random graph to a ring lattice results in a significant increase in the speed of convergence to consensus, regardless of the probability of the random links. In [20] the topology of certain classes of expander graphs is proven to be asymptotically optimal for achieving fast consensus in the limit of large graph size. In [21] the structure of circulant Laplacian matrices is exploited to examine the problem of probabilistic switching in consensus networks. In [22] the authors consider the problem of how to best add a fixed number of new links to a given network in order to maximally increase its number of spanning trees. It is proven in particular that for the problem of adding a single shortcut to a ring network, the number of spanning trees is an increasing function of the length of the added link. In [23] the design of fastest mixing Markov chains is examined, which is equivalent to the design of fastest converging consensus networks with undirected links. After demonstrating convexity of the design problem, various types of symmetries are exploited to solve the optimization problem analytically or to reduce its number of optimization variables. In [24] the design of optimal communication topologies for consensus is considered under the assumption of binary links. The resulting combinatorial optimization problems are solved using mixed-integer

Manuscript received June 8, 2016; accepted September 30, 2016. Date of publication October 20, 2016; date of current version July 26, 2017. This work was supported by the National Science Foundation under awards EAGER ECCS-1545270 and CNS-1329885. Recommended by Associate Editor M. Margaliot.

The author is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 (e-mail: makan@syr.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/TAC.2016.2619673

programming methods. In [25] an analytical framework is developed for finding the best links with which to augment consensus networks whose graphs are spanning trees. It is shown in particular that the best single link to add to a spanning tree is the one that minimizes the graph's diameter.

Compared to existing literature, the distinguishing features of the present work are its approach to link design within a budget-constrained optimization framework and the exploitation of spatial invariance and perturbation methods to analytically demonstrate the optimality of sparse long-range links.

II. Problem Formulation

We consider a network of n nodes, also referred to as agents, whose values evolve according to the DeGroot model [14]

$$x(t+1) = Tx(t), \quad \text{with } T\mathbb{1} = \mathbb{1},\ T \geq 0. \tag{1}$$

Here, $x(t)$ is a column n-vector composed of node values at time t, T is a (not necessarily symmetric) $n \times n$ matrix [26] whose nonnegative entries sum to one in every row, $\mathbb{1}$ is the column n-vector of all ones, and the inequality is elementwise. These dynamics imply that, at every time instant, nodes update their values by taking a weighted average of the values of nodes they receive information from. In particular, if the network is initialized at some $x(0)$ whose entries all belong to the interval [0, 1] then the entries of $x(t)$ remain within this interval for all $t \geq 0$. If all diagonal entries of T are positive and the directed graph described by T is strongly connected then all eigenvalues of T except for one at $\lambda = 1$ belong to the open unit disk [27] and the network reaches consensus asymptotically [14], i.e., $x(t) \to \chi\mathbb{1}$ as $t \to \infty$ for some $\chi \in [0, 1]$.

Reference [16] considers a design problem where the network as a whole is allotted a budget ϱ with which to create new links for the purpose of improving its efficiency of reaching consensus. Let σ_i denote the sum of weights corresponding to the new information-receiving communication links created by the ith agent, $\sum_{i=1}^{n} \sigma_i \leq \varrho$. For the ith agent to create new links of total weight σ_i, it has to 'make room' in its current weighted averaging scheme by scaling down with a factor of $1 - \sigma_i$ the weight of its existing links. Let the elementwise nonnegative matrix U describe the weight and distribution of the new links, so that the (i, j)th entry of U signifies a new link made by agent i to receive information from agent j. Then the entries in the ith row of U sum to σ_i, or equivalently $U\mathbb{1} = \sigma$, where σ is a vector whose ith entry is σ_i. Node values in this augmented network evolve according to (1) with T replaced by $ST + U$ and $S := I - \operatorname{diag}(\sigma)$. We can thus formulate an optimal link creation problem subject to a constraint on the total weight of new links

$$\begin{array}{ll} \text{maximize} & f(ST + U) \\ \text{subject to} & S = I - \operatorname{diag}(U\mathbb{1}) \\ & \mathbb{1}^* U\mathbb{1} \leq \varrho,\ U \geq 0,\ U\mathbb{1} \leq \mathbb{1}, \end{array}$$

where the optimization variables are the matrices U, S, all inequalities are elementwise, and $*$ represents the complex conjugate transpose. The objective function f characterizes the collective performance of the network, which in the present work will be the efficiency with which the network achieves consensus.

In contrast to [12], [16], and motivated in part by the investigation of small-world networks [1]–[4], in this work we restrict attention to spatially invariant networks described by circulant matrices. A circulant matrix is one in which every row is a right cyclic shift of the row above it [28]. We associate with every circulant matrix a graph whose nodes are arranged around a circle and labeled consecutively. Let $X_{i,j}$ denote the (i, j)th entry of matrix X. If X is circulant then $X_{i,j} \neq 0$ indicates the weight of the link that ends at node i and starts at the $|i - j|$th neighbor of node i on one side and equivalently the $(n - |i - j|)$th neighbor of node i on the other side. The following is a standing assumption in the rest of this technical note.

Assumption 1: The matrix T is real circulant with $T\mathbb{1} = \mathbb{1}$, $T \geq 0$, $T_{i,i} > 0$, and its corresponding directed graph is strongly connected.

For a network described by the circulant matrix T, the preceding optimization problem can be rewritten as

$$\begin{array}{ll} \text{maximize} & f(ST + U) \\ \text{subject to} & S = (1 - d)I,\ d \leq \delta \\ & U\mathbb{1} = d\,\mathbb{1},\ U \geq 0,\ UR = RU, \end{array} \tag{2}$$

where the optimization variables are the matrices U, S and the scalar d, with $\delta := \varrho/n \leq 1$. The function f measures the speed of the network in reaching average consensus. The parameter δ represents the total amount of resources available to each node with which to make new links. The circulant matrix R is zero everywhere except on its first lower subdiagonal and its $(1, n)$th entry, where it takes the value one; the constraint $UR = RU$ is a necessary and sufficient condition for the matrix U to be circulant. To see this, note that multiplication of U from the right by R results in a (cyclic) shift of the columns of U leftward, and multiplication of U from the left by R results in a (cyclic) shift of the rows of U downward; thus $UR = RU$ is equivalent to $U_{i,j+1} = U_{i-1,j}$, as needed to guarantee a circulant U.

We note that the possibilities in the solution of (2) range from agents making many new links with small weights (corresponding to a *dense* matrix U with many small entries) to agents making a few links with large weights (corresponding to a *sparse* matrix U with few significant entries). In the rest of this technical note, we will demonstrate that when f encapsulates the rate of convergence to consensus then the solution of (2) corresponds to the generation of sparse long-range links for all but very large values of δ.

Before closing this section, we make concrete what we mean by a sparse matrix, a banded matrix, and a long-range link. We say that a matrix is sparse when most of its entries are zero, and refer to it as *s-sparse* when at most s of its entries are nonzero. We say that a circulant matrix is *k-banded* when it has at most k nonzero upper and k nonzero lower (circulant) subdiagonals. Finally, we say that a link is *long-range* if the distance between the nodes it connects is of the order of the network's diameter.

III. Optimal Link Creation for Consensus Facilitation

We begin this section by reviewing a performance objective that captures the speed with which the network converges to average consensus. We use this objective together with (2) to formulate the optimal link creation problem, and rewrite it as a semidefinite program (SDP). We employ numerical experiments to show the sparse and long-range property of optimal links, as well as to demonstrate the graceful scaling with system size of the augmented network's temporal and graph-theoretic characteristics. An analytic treatment of the problem is deferred until Section IV.

Consider a design problem in which we wish to augment the communication graph of a network in order to achieve the fastest possible convergence to average consensus. Let $x_{\text{ave}}(t) := \frac{1}{n}\sum_i x_i(t) = \frac{1}{n}\mathbb{1}^* x(t)$, and consider the distance from average consensus at time t,

$$\psi(t) := \sum_{i=1}^{n} (x_i(t) - x_{\text{ave}}(t))^2 = x(t)^* Q\, x(t), \tag{3}$$

where $Q := I - \frac{1}{n}\mathbb{1}\mathbb{1}^*$. From the properties of T and Q, and the fact that all circulant matrices commute, it follows that

$$\psi(t) = x(0)^* Q \left(T^*T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*\right)^t Q\, x(0),$$

and thus

$$\psi(t) \le \|x(0) - x_{\text{ave}}(0)\mathbb{1}\|_2^2 \lambda_0^t, \qquad \lambda_0 := \lambda_{\max}(T^*T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*). \tag{4}$$

It is easy to show that $0 \le \lambda_0 < 1$, and that the inequality in (4) is satisfied with equality if $Qx(0)$ is aligned with the eigenvector corresponding to λ_0. Hence ψ decreases along the trajectories of the evolution (1). If T is symmetric then λ_0 is the square of the second-largest eigenvalue modulus (often referred to as SLEM in the literature) [23] of T. We conclude that to optimize the speed of convergence to consensus in the augmented network, we need to minimize the largest eigenvalue of the matrix $(ST + U)^*(ST + U) - \frac{1}{n}\mathbb{1}\mathbb{1}^*$. This determines the objective f in problem (2).

We hereafter replace the inequality constraint $d \le \delta$ in (2) with the equality constraint

$$d = \delta,$$

and claim that the resulting optimization problem will render the same augmented network. To elaborate, let $(U^{\text{opt}}, d^{\text{opt}})$ with $d^{\text{opt}} < \delta$ be a solution of (2). Then it is easy to see that there always exists a circulant matrix $U \ge 0$, satisfying $U\mathbb{1} = \delta\mathbb{1}$, such that the dynamics of the augmented system resulting from (U, δ) is identical to that resulting from $(U^{\text{opt}}, d^{\text{opt}})$; indeed, the circulant matrix $U = (\delta - d^{\text{opt}})T + U^{\text{opt}} \ge 0$ satisfies $U\mathbb{1} = \delta\mathbb{1}$ and uniquely solves the equation $(1 - \delta)T + U = (1 - d^{\text{opt}})T + U^{\text{opt}}$. In particular, both scenarios render the same argument for the objective function f in (2).

For simplicity of notation, we define the dynamics of the augmented system as

$$T_u := (1 - \delta)T + U,$$

which allows the elimination of the first two constraints in (2). We now formulate the main optimization problem of this technical note, namely the problem of optimal circulant link creation

$$\begin{array}{ll} \text{minimize} & \lambda_{\max}\left(T_u^* T_u - \frac{1}{n}\mathbb{1}\mathbb{1}^*\right) \\ \text{subject to} & U\mathbb{1} = \delta\mathbb{1},\ U \ge 0,\ UR = RU, \end{array} \tag{5}$$

recalling that $\delta \le 1$. Note that when resources are abundant and $\delta = 1$, the globally optimal solution of (5) is given by $U = \frac{1}{n}\mathbb{1}\mathbb{1}^*$, which renders the complete graph with homogeneous links $T_u = \frac{1}{n}\mathbb{1}\mathbb{1}^*$, [23].

Problem (5) can be rewritten as

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & \gamma I \succeq T_u^* T_u - \frac{1}{n}\mathbb{1}\mathbb{1}^* \\ & U\mathbb{1} = \delta\mathbb{1},\ U \ge 0,\ UR = RU, \end{array}$$

or equivalently, using the Schur complement, as the semidefinite program

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & \begin{bmatrix} \gamma I + \frac{1}{n}\mathbb{1}\mathbb{1}^* & T_u^* \\ T_u & I \end{bmatrix} \succeq 0 \\ & U\mathbb{1} = \delta\mathbb{1},\ U \ge 0,\ UR = RU. \end{array} \tag{6}$$

Henceforth in this work, problems (5)–(6) will be the main focus of our investigations. These problems are closely related to the fastest mixing Markov chain problem formulated in [23]. In contrast to [23], however, we incorporate a budget δ for the creation of new links at every node and eliminate any additional architectural constraints on the new links other than their circulant structure. We will see that this framework allows for the uncovering of pivotal communication topologies in the network, as characterized by the nonzero entries of the optimal matrix U. In the rest of this section using numerical examples, and in Section IV using analytical methods, we will investigate the sparse long-range and scaling properties of the solutions of (6).

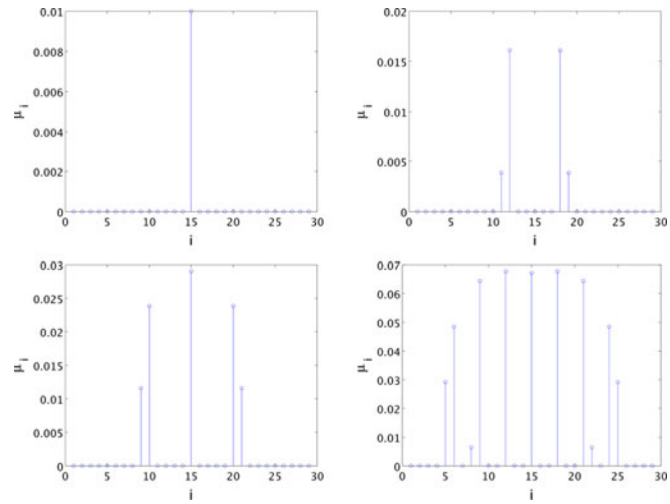

Fig. 1. Different plots correspond to different δ values, $\delta = 1/100, 1/25, 1/10, 1/2$. The horizontal axis in each plot represents the indices $i = 1, \ldots, n$ of the entries of a vector in $\mathbb{R}^n$. The circular points give the values of the entries of the first column μ of the optimal U resulting from problem (6). The value for $i = 1$ is repeated at $i = n + 1$ to emphasize the circulant structure. In the case of $\delta = 1$, we have $T_u = U = \frac{1}{n}\mathbb{1}\mathbb{1}^*$ and all entries of μ are equal to $1/n$ (not shown). The optimal γ values corresponding to the different plots are $\gamma = 0.9362, 0.8426, 0.6952, 0.1371$.

A. Illustrative Example

We consider a network with $n = 28$ nodes and nearest neighbor interactions described by a circulant matrix T whose first column τ is given by

$$\tau = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \cdots & 0 & \frac{1}{4} \end{bmatrix}^*.$$

We emphasize that T need not be symmetric and our choice of a symmetric T here is merely for the sake of clarity of exposition, as it leads to visually simpler plots. The entries of the first column μ of the optimal circulant U resulting from problem (6) appear in Fig. 1. For all computations we use CVX, a package for specifying and solving convex programs [29], [30].

It can be seen in particular that when δ is small and thus resources for the creation of new links are scarce, the optimal solution corresponds to each node making a link of strength δ to the node farthest away from itself, hence maximally reducing the graph's diameter. We emphasize that this solution is not obvious; for example, it is reasonable to expect that the optimal link-creation policy for every node is to distribute its recourses democratically and make links to all other nodes in the network. In Section IV we will revisit the small-δ problem, which we will refer to as the creation of 'weak' links, and will bring to bear perturbation methods to further explore the topology of optimal links rendered by (5)–(6).

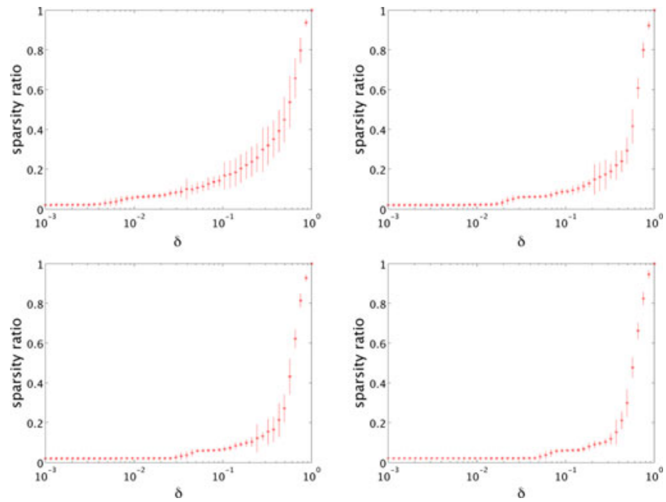

Fig. 2. Different plots correspond to different k values, $k = 1, 2, 3, 4$. The horizontal axes represent the value of δ. Every point on a plot gives the average value of the sparsity ratios for all networks in an ensemble, and the vertical bar through the point depicts one standard deviation on either side.

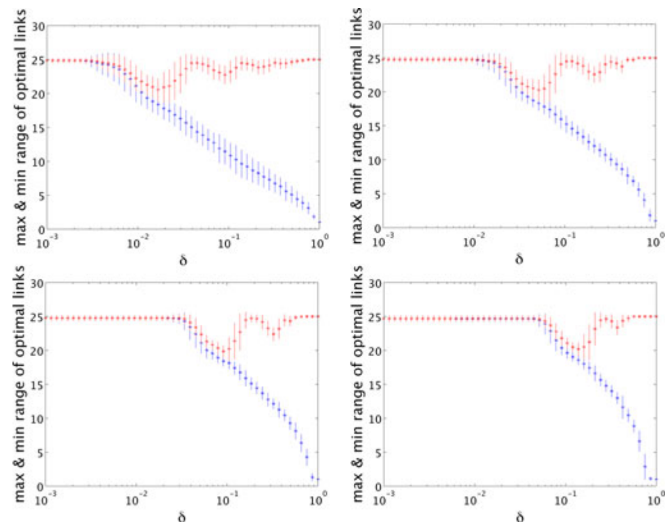

Fig. 3. Different plots correspond to different k values, $k = 1, 2, 3, 4$. The horizontal axes represent the value of δ. Every red (resp. blue) point on a plot gives the average value of the maximum (resp. minimum) range of optimal links for all networks in an ensemble, and the vertical bar through the point depicts one standard deviation on either side.

B. Ensembles of Random Networks

In this section we show that the sparse and long-range properties of optimal new links, demonstrated in Fig. 1 for a specific example, hold for general classes of circulant networks with local links. We consider ensembles of 50 randomly and independently generated circulant networks of size $n = 50$, where all nodes of the networks in a given ensemble have the same number of nearest neighbor interactions. We generate four ensembles $k = 1, 2, 3, 4$ of k-banded matrices, recalling that k signifies the maximum number of nearest neighbors with which every node interacts on each side. Thus the first column of the matrix T for any network that belongs to the kth ensemble has the form

$$\tau = \Big[\, \star \; \underbrace{\star \,\cdots\, \star}_{k} \; 0 \,\cdots\, 0 \; \underbrace{\star \,\cdots\, \star}_{k} \,\Big]^{*},$$

where the stars represent nonnegative random numbers that sum to one, such that $\mathbb{1}^{*}\tau = 1$.

Fig. 2 demonstrates the degree of sparsity of newly generated links as a function of δ, where each plot corresponds to an ensemble with fixed k. For a given network we define the *sparsity ratio* of an optimal U resulting from problem (6) to be the ratio of the number of nonzero entries in U to the total number of possible entries n^2. Every point on a plot gives the average value of the sparsity ratios for all networks in an ensemble, and the vertical bar through the point depicts one standard deviation on either side. For small values of δ the optimal U for all networks in an ensemble are sparse, and they remain sparse for higher values of δ as the number k of nearest neighbor interactions in the original network increases.

Fig. 3 demonstrates the length of newly generated links as a function of δ, where as before each plot corresponds to an ensemble with fixed k. For a given network we define the *maximum* (resp. *minimum*) *range of optimal links* as the length of the longest (resp. shortest) link among all optimal links resulting from problem (6). Every red (resp. blue) point on a plot gives the average value of the maximum (resp. minimum) range of optimal links for all networks in an ensemble, and the vertical bar through the point depicts one standard deviation on either side. For small values of δ the optimal U for all networks in an ensemble correspond to the generation of long-range links. The red plots show that long-range links persist through the entire range of δ values, while the blue plots illustrate the decrease in length of the shortest links as δ increases.

C. Scaling-in-n of Convergence Rate & Average Path Length

Issues of scaling are prominent in network analysis and design. We consider the scaling-in-n of the solutions of (6) with respect to two characteristic quantities of dynamical networks: convergence rate and average path length.

We define the (system-theoretic) notion of *rate of convergence-to-consensus* of the circulant network associated with T as

$$\xi(T) := -\log\left(\lambda_{\max}\left(T^{*}T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^{*}\right)\right). \tag{7}$$

We have that $\xi(T) \to 0$ as $\lambda_{\max}(T^{*}T - \frac{1}{n}\mathbb{1}\mathbb{1}^{*}) \to 1$ and $\xi(T) \to \infty$ as $\lambda_{\max}(T^{*}T - \frac{1}{n}\mathbb{1}\mathbb{1}^{*}) \to 0$.[1] To see how the convergence properties of the solution T_u of (6) compare to that of T, we normalize the rate of convergence of T_u with that of T and consider $\xi(T_u)/\xi(T)$. This normalization will be particularly revealing when we consider the scaling of $\xi(T_u)$ with network size.

We also employ the (graph-theoretic) notion of the *average path length* of the graph associated with T, defined as the average of the shortest paths between all pairs of nodes in the graph. For a given T we define its associated binary matrix A as

$$A_{i,j} := \begin{cases} 1 & \text{if } T_{i,j} \neq 0, \\ 0 & \text{if } T_{i,j} = 0. \end{cases}$$

The graph corresponding to A is the unweighted version of the directed graph corresponding to T. Let $\Delta(T)$ represent the matrix of node

[1] Part of the motivation for definition (7) comes from the continuous-time consensus problem $\dot{x} = -L\,x$ considered in [5], where changes to the second smallest eigenvalue ξ of the symmetric Laplacian matrix L of a network was investigated as a result of the addition of new links. Given the temporal convergence-to-consensus of $\mathrm{e}^{-\xi t}$ for such a system and comparing it to the convergence behavior $\lambda^{t} \equiv \mathrm{e}^{(\log\lambda)t}$ for a discrete-time consensus network [see (4)] one obtains the correspondence $\xi \sim -\log\lambda$. Reference [5] further normalizes ξ by considering ξ/ξ_0, where ξ_0 characterizes the second smallest eigenvalue of the Laplacian of the original (non-augmented) network.

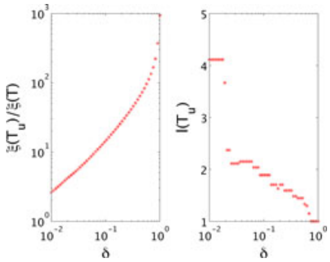

Fig. 4. The plot on the left shows $\xi(T_u)/\xi(T)$ for U resulting from problem (6), and the one on the right shows $\ell(T_u)$. The horizontal axes represent the value of δ. In the case of $\delta = 0$, we have $\ell(T_u) = \ell(T) = 7.26$ corresponding to the average path length of the original network (not shown).

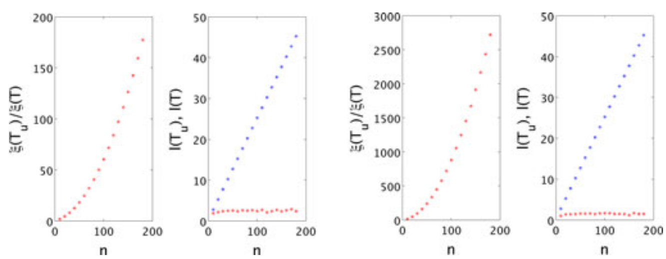

Fig. 5. Different pairs of plots correspond to different δ values, $\delta = 1/25, 1/2$. In each pair of plots, the one on the left shows $\xi(T_u)/\xi(T)$ for U resulting from problem (6), and the one on the right shows $\ell(T_u)$ [in red] versus $\ell(T)$ [in blue]. The horizontal axes represent the size n of the network.

distances

$$(\Delta(T))_{i,j} := \text{length of shortest path from node } j \text{ to node } i \text{ in directed graph defined by } A,$$

where every link has unit length and $(\Delta(T))_{i,i} = 0$. The average path length of T is now given by

$$\ell(T) := \frac{1}{n(n-1)} \sum_{i,j=1}^{n} (\Delta(T))_{i,j}.$$

Inspired by [5], we return to the network considered in Section III-A and investigate changes in $\xi(T_u)/\xi(T)$ and $\ell(T_u)$ resulting from variations in δ. Fig. 4 demonstrates these changes as a function of δ. Similar to the findings in [5], we observe that even small values of δ in (6) lead to dramatic improvements in the rate of convergence-to-consensus for the augmented system.

As another example, we consider networks with nearest neighbor interactions described by circulant matrices $T \in \mathbb{R}^{n\times n}$ whose first column τ is given by

$$\tau = \left[\, \tfrac{10}{20} \;\; \tfrac{3}{20} \;\; 0 \;\; \cdots \;\; 0 \;\; \tfrac{7}{20} \,\right]^*.$$

While maintaining the same nonzero entries for T (and therefore the same local interaction topology) and fixing the value of δ, we increase the network size n and observe the scaling properties of the solution of (6). In particular, we examine the scaling-in-n of the convergence rate and average path length corresponding to the augmented system T_u, and compare it with those of T. As before, all optimization problems are solved using CVX [29], [30].

Each pair of plots in Fig. 5 corresponds to a different value of δ and demonstrates how $\xi(T_u)/\xi(T)$ [on left] and $\ell(T_u)$, $\ell(T)$ [on right] scale with n. Plots remain qualitatively similar for the entire range of δ values. The plots in Fig. 5 reveal that the newly created links have an increasingly prominent effect on the convergence rate as n grows. Furthermore, it can be seen that while $\ell(T)$ scales linearly in n, the solution T_u of (6) is such that $\ell(T_u)$ remains essentially independent of n.

IV. Creation of Optimal Weak Links: Sparse & Long-Range Communications

In this section we consider the problem of finding optimal weak links, where we refer to a link as weak if it has small weight. Weak links can be thought of as a proxy for modeling infrequent communication between the nodes they connect. The cohesive power of weak ties and their importance in the diffusion of information in social networks was illustrated in the seminal work [17]. We use perturbation methods to find conditions on the network architecture which guarantee optimal weak links that are both sparse and long-range. For improved readability we first state our main results in Props. 1–4 and then proceed to their proofs.

A. Optimal Weak Links

We reconsider problem (5) under the assumption that δ is a small positive number ε, and hence the elementwise-positive matrix U has small entries

$$\begin{array}{ll} \text{minimize} & \lambda_{\max}\left(T_u^* T_u - \frac{1}{n}\mathbb{1}\mathbb{1}^*\right) \\ \text{subject to} & U\mathbb{1} = \varepsilon\mathbb{1},\; U \geq 0,\; UR = RU. \end{array} \tag{8}$$

The conceivable scenarios resulting from solutions of (8) range from every node making only one link of weight ε to every node making many links whose weights add up to ε. Prop. 3 and Prop. 4 in what follows will establish conditions on T such that the optimal U resulting from (8) corresponds to the generation of sparse and long-range links. We begin, however, by exploiting a perturbation framework in Prop. 1 to simplify (8).

Proposition 1: For $U =: \varepsilon V$, $V\mathbb{1} = \mathbb{1}$, and small enough values of ε we have

$$\lambda_{\max}\left(T_{\varepsilon v}^* T_{\varepsilon v} - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*\right) = \lambda_{\max}\left(T^*T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*\right) + \varepsilon g(V) + O(\varepsilon^2),$$

where

$$\begin{aligned} T_{\varepsilon v} &:= (1-\varepsilon)T + \varepsilon V, \\ g(V) &:= \lambda_{\max}\left(P^*(T^*V + V^*T)P\right) - 2\lambda_0, \end{aligned}$$

and $P \in \mathbb{C}^{n\times r}$ is a matrix whose columns constitute normalized and mutually orthogonal eigenvectors corresponding to the largest eigenvalue $\lambda_0 > 0$ of $T^*T - \frac{1}{n}\mathbb{1}\mathbb{1}^*$ with multiplicity r.

We remark that perturbed repeated eigenvalues can in general have non-analytic dependence on the perturbation parameter [31]. However, our restriction to circulant matrices in the present work allows for a simplification of the spectral perturbation problem and renders eigenvalues that are all analytic in ε.

Motivated by Prop. 1, we replace the objective in (8) with its first-order approximation to form the optimal weak link creation problem

$$\begin{array}{ll} \text{minimize} & \lambda_{\max}\left(P^*(T^*V + V^*T)P\right) \\ \text{subject to} & V\mathbb{1} = \mathbb{1},\; V \geq 0,\; VR = RV. \end{array} \tag{9}$$

To elaborate on the relationship between problems (9) and (8), let $\delta = \varepsilon$ be a given small number signifying the total weight of the new information-receiving links to be created at each node, and let V^{opt} solve (9). Then

$$U = \varepsilon V^{\text{opt}}$$

is an approximate solution to (8).

B. Optimality of Sparse & Long-Range Links

We begin by reviewing some definitions and notation pertaining to Fourier domain representations. We then employ frequency methods to uncover conditions under which optimal weak links will be sparse and long-range.

Let the unitary matrix F indicate the discrete Fourier transform,

$$F_{l+1,m+1} := \tfrac{1}{\sqrt{n}}\, \mathrm{e}^{-\mathrm{i}2\pi lm/n}, \quad l,m = 0,\ldots,n-1. \tag{10}$$

We will henceforth use hatted variables to represent the Fourier transform of matrices and vectors,

$$\hat{x} := Fx, \quad \hat{X} := FXF^*,$$

for any n-vector x and $n \times n$ matrix X. It is well-known that all circulant matrices are simultaneously diagonalized by F [26], so that if X is circulant then $\hat{X}$ is diagonal. Let τ and $\hat{\tau}$ respectively characterize the first column of the circulant matrix T and its Fourier symbol,

$$\tau = Te_1, \quad \hat{\tau} = F\tau, \tag{11}$$

where e_1 denotes the 1st standard basis vector of $\mathbb{R}^n$. Then

$$\hat{\tau}_i = \tfrac{1}{\sqrt{n}}\, \hat{T}_{i,i}, \quad i = 1,\ldots,n, \tag{12}$$

with $\hat{T} = FTF^*$.[2] Finally, let

$$\theta := \angle \hat{T}_{2,2} = \angle \hat{\tau}_2, \quad \rho := |\hat{T}_{2,2}| = \sqrt{n}\,|\hat{\tau}_2|. \tag{13}$$

and take e_i to represent the ith standard basis vector of $\mathbb{R}^n$.

Assumption 2: The Fourier symbol of the matrix T satisfies

$$|\hat{\tau}_2| > |\hat{\tau}_i|, \quad i = 3,\ldots,\lfloor \tfrac{n}{2} \rfloor + 1,$$

with $\hat{\tau}$ defined as in (12)–(11).

The assumption implies that every column/row of T as a vector has a larger component in the direction of the first spatial harmonic $\varphi_2 = F^* e_2$ (and $\varphi_n = F^* e_n$) than in the direction of any other harmonic, i.e., $|\varphi_2^* \tau| > |\varphi_i^* \tau|$ or equivalently $|\hat{\tau}_2| > |\hat{\tau}_i|$ for $i \neq 1, 2, n$. In many physical systems, for example in a social network, it is reasonable to expect that T satisfies Assumption 2. Roughly, this is because the influence of agents on one another deteriorates as they grow more distant, and therefore the nonnegative entries of T decay as one moves from the main diagonal to the $\lfloor \frac{n}{2} \rfloor$th upper and lower subdiagonals. For instance, it is often the case that T is a banded matrix with only a few nonzero subdiagonals around its main diagonal, corresponding to local communication only between every node and its closest neighbors.

Proposition 2: If Assumptions 1–2 hold, then (9) is equivalent to the optimization problem

$$\begin{array}{ll} \text{minimize} & \mathrm{trace}(K^* V) \\ \text{subject to} & V\mathbb{1} = \mathbb{1},\ V \geq 0,\ VR = RV, \end{array} \tag{14}$$

with

$$K := TPP^*, \quad P = F^* \big[\, e_2, e_n \,\big].$$

One can interpret K as a 'filtered' version of T, where only the first spatial harmonic of the columns/rows of T survives the filter PP^*, and the columns/rows of K inherit a pure cosine structure of period n and frequency $l = 1$, namely $K_{i,j} = \frac{2\rho}{n} \cos(\frac{2\pi}{n}(i-j) + \theta)$; see (18)–(19) for details.

[2] We have $\hat{\tau} = F\tau = FTe_1 = \frac{1}{\sqrt{n}} FTF^* \mathbb{1} = \frac{1}{\sqrt{n}} \hat{T}\mathbb{1}$ and therefore $\hat{\tau}_i = e_i^* \hat{\tau} = \frac{1}{\sqrt{n}} e_i^* \hat{T}\mathbb{1} = \frac{1}{\sqrt{n}}\, \hat{T}_{i,i}$.

The following two propositions constitute the main results of this section.

Proposition 3: If Assumptions 1–2 hold, then the solution of (14) is an n-sparse matrix.

Proposition 4: If Assumptions 1–2 hold and

$$\tfrac{n-1}{2} - \tfrac{n}{2\pi}\theta \leq q < \tfrac{n+1}{2} - \tfrac{n}{2\pi}\theta \tag{15}$$

for θ defined as in (13) and some integer q, then the solution of (14) corresponds to the generation of a link between every node and its qth neighbor. Furthermore, if T is a k-banded matrix with $k \ll n$, this solution corresponds to the generation of long-range links.

The results of Prop. 4 are particularly easy to verify for a symmetric and banded matrix T. In this case $\hat{\tau} \in \mathbb{R}^n$, $\theta = 0$, $K_{i,j} = \frac{2\rho}{n} \cos(\frac{2\pi}{n}(i-j))$, and therefore the smallest entries of K occur at its $|i-j| = \frac{n}{2}$ subdiagonal if n is even and its $|i-j| = \frac{n-1}{2}$ subdiagonal if n is odd, thus maximally reducing the network's diameter. This is reminiscent of the results in [25].

In summary, Assumption 2 and Prop. 3 ensure the sparsity of optimal weak links. However, while Assumption 2 restricts the columns/rows of T to have a strong first harmonic component it does not restrict spatial shifts of this component. This motivates Prop. 4, where the exact range of optimal weak links and conditions for their long-range property are derived.

C. Proofs of Propositions 1–4

Proof of Prop. 1: Replace U in (8) with εV, and let $T_{\varepsilon v} = (1-\varepsilon)T + \varepsilon V$. Clearly

$$\begin{aligned}
\lambda_{\max}\big(T_{\varepsilon v}^* T_{\varepsilon v} - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*\big) &= \lambda_{\max}(\hat{T}_{\varepsilon v}^* \hat{T}_{\varepsilon v} - e_1 e_1^*) \\
&= \lambda_{\max}\big(\hat{T}^*\hat{T} - e_1 e_1^* + \varepsilon(\hat{T}^*\hat{V} + \hat{V}^*\hat{T} - 2\,\hat{T}^*\hat{T}) \\
&\qquad + \varepsilon^2(\hat{T}^*\hat{T} - \hat{T}^*\hat{V} - \hat{V}^*\hat{T} + \hat{V}^*\hat{V})\big),
\end{aligned}$$

where all hatted matrices are diagonal. For small enough ε the maximum eigenvalue of $T_{\varepsilon v}^* T_{\varepsilon v} - \frac{1}{n}\mathbb{1}\mathbb{1}^*$ is thus found from a perturbation of the largest (and possibly repeated) diagonal entries $\lambda_0 > 0$ of the matrix $\hat{T}^*\hat{T} - e_1 e_1^*$. Therefore,

$$\begin{aligned}
&\lambda_{\max}(T_{\varepsilon v}^* T_{\varepsilon v} - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*) \\
&\quad = \lambda_0 + \varepsilon \lambda_{\max}\big(\hat{P}^*(\hat{T}^*\hat{V} + \hat{V}^*\hat{T} - 2\lambda_0 I)\hat{P}\big) + O(\varepsilon^2),
\end{aligned}$$

where $\hat{P} = \big[\, e_{i_1}, \ldots, e_{i_r} \,\big]$ is a matrix whose columns are composed of standard basis vectors of $\mathbb{R}^n$, and the indices $\{i_1, \ldots, i_r\}$ correspond to the $r \geq 1$ locations where λ_0 appears on the diagonal of $\hat{T}^*\hat{T} - e_1 e_1^*$. Returning to the spatial domain we obtain the expression for $\lambda_{\max}$ given in the statement of the proposition, with

$$P = F^* \hat{P}.$$

The proof of the proposition is now complete. ■

The following lemma will be instrumental in the proofs of Props. 2–4.

Lemma 5: If Assumptions 1–2 hold, then the largest eigenvalue $\lambda_0 > 0$ of $T^*T - \frac{1}{n}\mathbb{1}\mathbb{1}^*$ has multiplicity $r = 2$ and the matrix P of orthonormal eigenvectors of $T^*T - \frac{1}{n}\mathbb{1}\mathbb{1}^*$ corresponding to λ_0 is given by

$$P = F^* \big[\, e_2, e_n \,\big]. \tag{16}$$

Proof of Lemma 5: From (12) we have for $i = 2, \ldots, n$,

$$
\begin{aligned}
n|\hat{\tau}_i|^2 &= \|\hat{T} e_i\|^2 \\
&= \|T F^* e_i\|^2 \\
&= \|(T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*)F^* e_i\|^2,
\end{aligned}
$$

where in the last equality we have used the fact that $\mathbb{1}$ is orthogonal to the ith column of F^* for $i \geq 2$. Thus

$$
\begin{aligned}
n|\hat{\tau}_i|^2 &= e_i^* F M F^* e_i \qquad (17)\\
&= \hat{M}_{i,i},
\end{aligned}
$$

with $M := T^*T - \frac{1}{n}\mathbb{1}\mathbb{1}^*$.

Since M is a real symmetric circulant matrix, $\hat{M}$ is real diagonal with $\hat{M}_{i,i} = \hat{M}_{n-i+2,n-i+2}$ for $i = 2, \ldots, \lfloor \frac{n}{2} \rfloor + 1$. Then the largest eigenvalue λ_0 of M appears in pairs at the ith and $(n-i+2)$th diagonal entries of $\hat{M}$ for (possibly multiple) $i \in \{2, \ldots, \lfloor \frac{n}{2} \rfloor + 1\}$, and the eigenvectors corresponding to λ_0 occur in pairs at the ith and $(n-i+1)$th columns of F^*. Note also that $T\mathbb{1} = \mathbb{1}$ implies $\hat{M}_{1,1} = 0$.

From Assumption 2, equations (17), and the argument in the previous paragraph, it follows that

$$
\hat{M}_{2,2} > \hat{M}_{i,i}, \quad i = 3, \ldots, \lfloor \tfrac{n}{2} \rfloor + 1
$$

and therefore λ_0 has multiplicity two, and that the eigenvectors of $T^*T - \frac{1}{n}\mathbb{1}\mathbb{1}^*$ corresponding to λ_0 are F^*e_2 and F^*e_n. The proof of the lemma is now complete. ■

Proof of Prop. 2: Since $T^*V + V^*T$ is a real symmetric circulant matrix, the 2nd and nth diagonal entries of $\hat{T}^*\hat{V} + \hat{V}^*\hat{T}$ are real and equal. Let these diagonal entries have value $\zeta \in \mathbb{R}$. Then

$$
\hat{P}^*(\hat{T}^*\hat{V} + \hat{V}^*\hat{T})\hat{P} = \zeta I,
$$

where $\hat{P} = \left[\, e_2, e_n \,\right]$ from Lemma 5 and the identity matrix on the right has dimension 2. This implies $P^*(T^*V + V^*T)P = \zeta I$ and

$$
\begin{aligned}
\lambda_{\max}\big(P^*(T^*V + V^*T)P\big) &= \frac{1}{2}\,\mathrm{trace}\big(P^*(T^*V + V^*T)P\big) \\
&= \mathrm{trace}(K^*V),
\end{aligned}
$$

with $K = TPP^*$. The proof of the proposition is now complete. ■

Proof of Prop. 3: Since T is a real circulant matrix, $\hat{T}$ is diagonal with $\hat{T}_{i,i} = (\hat{T}_{n-i+2,n-i+2})^*$ for $i = 2, \ldots, \lfloor \frac{n}{2} \rfloor + 1$. Recalling from Lemma 5 that $\hat{P} = \left[\, e_2, e_n \,\right]$, we have from Prop. 2

$$
\begin{aligned}
\hat{K} &= \hat{T}\hat{P}\hat{P}^* \qquad (18)\\
&= \hat{T}_{2,2}\, e_2 e_2^* + \hat{T}_{n,n}\, e_n e_n^* \\
&= \rho\, \mathrm{e}^{\mathrm{i}\theta}\, e_2 e_2^* + \rho\, \mathrm{e}^{-\mathrm{i}\theta} e_n e_n^*.
\end{aligned}
$$

Multiplying $\hat{K}$ from left and right respectively by F^* and F, we obtain

$$
K = \rho\, \mathrm{e}^{\mathrm{i}\theta} F^* e_2 e_2^* F + \rho\, \mathrm{e}^{-\mathrm{i}\theta} F^* e_n e_n^* F,
$$

and for $i, j = 1, \ldots, n$,

$$
\begin{aligned}
K_{i,j} &= \tfrac{\rho}{n}\, \mathrm{e}^{\mathrm{i}\theta} \mathrm{e}^{\mathrm{i}2\pi(i-1)/n}\, \mathrm{e}^{-\mathrm{i}2\pi(j-1)/n} \qquad (19)\\
&\quad + \tfrac{\rho}{n}\, \mathrm{e}^{-\mathrm{i}\theta} \mathrm{e}^{\mathrm{i}2\pi(n-1)(i-1)/n}\, \mathrm{e}^{-\mathrm{i}2\pi(n-1)(j-1)/n} \\
&= \tfrac{2\rho}{n} \cos\!\big(\tfrac{2\pi}{n}(i-j) + \theta\big).
\end{aligned}
$$

Thus K is circulant, as its ijth entry depends only on $i - j$ (modulo n).

Furthermore, the cosine structure of the rows (and columns) of K implies that its entries achieve a distinct minimum value in every row (and column). Let ν denote the value of the smallest entries of K,

$$
\nu = \min\{K_{i,j},\ i, j = 1, \ldots, n\},
$$

and let i_ν denote the index of some circulant subdiagonal of K that achieves this minimum. Take V^{opt} to be the circulant matrix that takes the value one on its i_νth subdiagonal and zero elsewhere. The matrix V^{opt} achieves the smallest value of the objective in (14) among all elementwise-nonnegative V for which $V\mathbb{1} = \mathbb{1}$, and is n-sparse. Thus, in the set of nonzero circulant matrices of size n, V^{opt} is maximally sparse. The proof of the proposition is now complete. ■

Proof of Prop. 4: Recall from the proof of Prop. 3 that

$$
K_{i,j} = \tfrac{2\rho}{n}\ \cos\big(\tfrac{2\pi}{n}(i-j) + \theta\big).
$$

Thus the constant value on the $(i-j)$th circulant subdiagonal of K is given by sampling the cosine function at $\frac{2\pi}{n}(i-j) + \theta$. It is convenient to make a change of variables by setting $i - j =: q$ with $q \in \mathbb{Z}$. The minimum value of the entries of K is then achieved on its qth subdiagonal if

$$
\pi - \frac{\pi}{n} \leq \frac{2\pi}{n} q + \theta < \pi + \frac{\pi}{n}, \qquad (20)
$$

which is equivalent to inequality (15). It follows that V^{opt} defined as the circulant matrix with ones on its qth subdiagonal and zeros elsewhere solves (14).

Furthermore, if T is k-banded then

$$
\begin{aligned}
\hat{\tau}_2 &= \tfrac{1}{\sqrt{n}} \sum_{m=0}^{n-1} \mathrm{e}^{-\mathrm{i}2\pi m/n}\, \tau_{m+1} \\
&= \tfrac{1}{\sqrt{n}} \sum_{m=-k}^{k} \mathrm{e}^{-\mathrm{i}2\pi m/n}\, \tau_{m+1},
\end{aligned}
$$

where $\tau_i := \tau_{n+i}$ for $i = -k, \ldots, 0$. The last expression above is the weighted sum of $2k + 1$ unit-modulus complex numbers whose phases are equally spaced in $[-\frac{2\pi}{n}k, \frac{2\pi}{n}k]$. This, together with the assumptions $k \ll n$ and $\tau_1 = T_{1,1} > 0$, implies that

$$
|\theta| = |\angle \hat{\tau}_2| < \frac{2\pi}{n} k. \qquad (21)
$$

Inequalities (15) and (21) render

$$
\frac{n-1}{2} - k < q < \frac{n+1}{2} + k.
$$

Since $k \ll n$ we conclude that $|q|$ is $O(n)$, which guarantees the generation of a long-range link at every node. The proof of the proposition is now complete. ■

V. Conclusion

We consider the problem of budget-constrained optimal link creation for the enhancement of convergence to consensus in circulant networks. We demonstrate numerically that when the budget for link creation is small, optimal links have the property of being both sparse and long-range for large classes of networks with local communications. We use perturbation methods to analytically investigate the topology of optimal communication links of small strength, and uncover conditions on the network architecture which guarantee their sparsity and long-range property.

Acknowledgement

The author would like to thank anonymous reviewers for their valuable comments that helped improve this technical note.

References

[1] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, 1998.

[2] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.

[3] M. O. Jackson, *Social and Economic Networks*. Princeton University Press, 2010.

[4] M. E. J. Newman, *Networks: An Introduction*. Oxford University Press, 2010.

[5] R. Olfati-Saber, "Ultrafast consensus in small-world networks," in *Proc. American Control Conf.*, 2005, pp. 2371–2378.

[6] S. N. Dorogovtsev, *Lectures on Complex Networks*. Clarendon Press, 2010.

[7] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, "Classes of small-world networks," *Proc. National Academy Sci.*, vol. 97, pp. 11 149–11 152, 2000.

[8] K. A. Seaton and L. M. Hackett, "Stations, trains and small-world networks," *Physica A: Statistical Mech. Appl.*, vol. 339, pp. 635–644, 2004.

[9] O. Sporns, D. R. Chialvo, M. Kaiser, and C. C. Hilgetag, "Organization, development and function of complex brain networks," *Trends Cognitive Sci.*, vol. 8, pp. 418–425, 2004.

[10] M. D. Humphries, K. Gurney, and T. J. Prescott, "Is there a brainstem substrate for action selection?," *Philosophical Trans. Royal Society B*, vol. 362, pp. 1627–1639, 2007.

[11] S. Yu, D. Huang, W. Singer, and D. Nikolić, "A small world of neuronal synchrony," *Cerebral Cortex*, vol. 18, pp. 2891–2901, 2008.

[12] M. Fardad, X. Zhang, F. Lin, and M. R. Jovanović, "On the properties of optimal weak links in social networks," in *Proc. 53rd IEEE Conf. Decision Control*, 2014, pp. 2124–2129.

[13] M. Fardad, "On the optimality of sparse long-range links in circulant consensus networks," in *Proc. American Control Conf.*, Jul. 2015, pp. 2075–2080.

[14] M. H. DeGroot, "Reaching a consensus," *J. American Statistical Assoc.*, vol. 69, pp. 118–121, 1974.

[15] J. Kleinberg, "Navigation in a small world," *Nature*, vol. 406, p. 845, 2000.

[16] M. Fardad, F. Lin, and M. R. Jovanović, "On optimal link creation for facilitation of consensus in social networks," *Proc. American Control Conf.*, 2014, pp. 3790–3795.

[17] M. Granovetter, "The strength of weak ties," *American J. Sociology*, vol. 78, pp. 1360–1380, 1973.

[18] V. Latora and M. Marchiori, "Economic small-world behavior in weighted networks," *Eur. Physical J. B*, vol. 32, pp. 249–263, 2003.

[19] A. Tahbaz-Salehi and A. Jadbabaie, "Small world phenomenon, rapidly mixing markov chains, and average consensus algorithms," in *Proc. 46th IEEE Conf. Decision Control*, 2007, pp. 276–281.

[20] S. Kar, S. Aldosari, and J. M. F. Moura, "Topology for distributed inference on graphs," *IEEE Trans. Signal Processing*, vol. 56, pp. 2609–2613, 2008.

[21] P. Hovareshti, J. S. Baras, and V. Gupta, "Average consensus over small world networks: A probabilistic framework," *Proc. 47th IEEE Conf. Decision Control*, 2008, pp. 375–380.

[22] J. S. Baras and P. Hovareshti, "Efficient and robust communication topologies for distributed decision making in networked systems," *Proc. 48th IEEE Conf. Decision Control*, 2009, pp. 3751–3756.

[23] S. Boyd, P. Diaconis, P. Parrilo, and L. Xiao, "Fastest mixing markov chain on graphs with symmetries," *SIAM J. Optim.*, vol. 20, pp. 792–819, 2009.

[24] R. Dai and M. Mesbahi, "Optimal topology design for dynamic networks," in *Proc. 50th IEEE Conf. Decision Control*, 2011, pp. 1280–1285.

[25] D. Zelazo, S. Schuler, and F. Allgöwer, "Performance and design of cycles in consensus networks," *Syst. Control Lett.*, vol. 62, no. 1, pp. 85–96, 2013.

[26] P. J. Davis, *Circulant Matrices*. Wiley, 1970.

[27] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," *IEEE J. Selected Areas Commun.*, vol. 26, pp. 634–649, 2008.

[28] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations Trends Commun. Inform. Theory*, vol. 2, pp. 155–239, 2006.

[29] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," http://cvxr.com/cvx/, 2013.

[30] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," *Recent Advances in Learning and Control*, ser. Lecture Notes Control Inform. Sci., V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110.

[31] T. Kato, *Perturbation Theory for Linear Operators*. Springer-Verlag, 1995.