2015 American Control Conference
Palmer House Hilton
July 1-3, 2015. Chicago, IL, USA

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

Makan Fardad

***Abstract*— We consider spatially invariant consensus networks in which the link weights, the directed graph describing the interconnection topology, 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. We show that the optimal circulant link creation problem is convex and can be written as a semidefinite program. Motivated by small-world networks, we apply the link creation problem to circulant networks which possess only local communication links. We observe that the optimal new links are always 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, we restrict attention to the creation of links with small strengths, which we refer to as weak links. We employ perturbation methods to reformulate the optimal weak link creation problem, and uncover conditions on the network architecture which guarantee sparse and long-range solutions to this optimization problem.**

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

I. Introduction

Small-world networks [1] have been the topic of intense 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 (unweighted and undirected) 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 answer the following questions: Are small-world networks optimal or nearly optimal in the way they balance link creation and collective performance? Is nature solving an optimization problem when it arranges in the form of a small-world network the anatomical connections in the brain and the visual system [7]–[9]? Are humans optimizing an objective when we configure our transportation networks and power grids in a small-world topology [10], [11]?

In this paper, we consider the problem of link creation for the purpose of consensus facilitation in circulant networks [12], [13]. We frame our network design problems in the context of social networks and consensus formation via the DeGroot model [14] as a motivational application and concrete instance of collaborative behavior. We focus on regular networks described by circulant matrices, given that small-world networks have been generated within the context of regular lattices in the literature [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). 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].

We demonstrate that the optimal link creation problem is convex and can be formulated as a semidefinite program. Solutions of the optimization 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 conditions on the network architecture which guarantee the sparsity and long-range property of optimal weak links. The obtained results suggest that small-world networks are optimal in the way they balance parsimonious link creation and harmonious collective behavior.

Our contributions are most related to the work of Latora and Marchiori [18] on economic small-world networks, Olfati-Saber [5] on ultrafast consensus in small-world networks, Baras and Hovareshti [19] on efficient communication topologies in networked systems, Zelazo *et al.* [20] on the design of cycles in consensus networks, and Boyd *et al.* [21] on fastest mixing Markov chains and the exploitation of symmetry groups in network graphs. Except for [18], the mentioned papers are restricted to symmetric communication links corresponding to undirected graphs (and in the case of [5], [19], [20], also unweighted), whereas the communication links in this work are allowed to be both directed and weighted. Our work further differs from [5], [18] in

Financial support from the National Science Foundation under award CNS-1329885 is gratefully acknowledged.

M. Fardad is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244. E-mail: makan@syr.edu.

978-1-4799-8686-6/$31.00 ©2015 AACC

that new links result from the solution of an optimization problem rather begin generated randomly. As opposed to [18], [19], where distance-based global and local measures of efficiency and the graph-theoretic notion of number of spanning trees respectively are used to characterize performance, we measure a network's performance by the speed with which it achieves consensus. While [20] develops an analytical framework for finding the best links with which to augment graphs that are spanning trees, in this work our only assumption on the original graph is its circulant structure. Similar to [21], we exploit problem symmetry, namely the spatial invariance of the network, to obtain optimal values of link weights. However, in contrast to [21], our primary focus is on detecting pivotal communication topologies in the network. We perform this task by not assuming a predetermined communication topology and rather incorporating a budget constraint for new link creation. Moreover, a distinguishing feature of this work compared to existing literature is that through exploitation of problem structure and use of a perturbation framework, we demonstrate the sparsity and long-range property of optimal weak links.

II. Problem Formulation

We consider a network of n agents whose beliefs evolve according to the DeGroot model [14]

$$x(k+1) = Tx(k), \quad \text{with} \quad T\mathbb{1} = \mathbb{1},\ T \geq 0, \qquad (1)$$

where the inequality is elementwise. Here, $x(k)$ is a column n-vector composed of values that represent the beliefs of agents at time k; T is a (not necessarily symmetric) matrix [22] whose nonnegative entries sum to one in every row; and $\mathbb{1}$ is the column n-vector of all ones. These dynamics imply that, at every time instant, agents update their beliefs by taking a weighted average of those they interact with. In particular, if the network is initialized at some $x(0)$ whose entries all belong to the interval $[0, 1]$, then the entries of $x(k)$ remain within this interval for all $k \geq 0$. If all diagonal entries of T are positive and the graph described by T is strongly connected then all eigenvalues of T except for one at $\lambda = 1$ belong to the open unit disk [23], and the network (1) reaches consensus asymptotically [14].

Reference [16] considers a framework where the network as a whole is allotted a budget ϱ with which to create new links for the purpose of improving its collective consensus performance. Let σ_i denote the sum of weights corresponding to the new 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 $S := I - \text{diag}\{\sigma\}$. We can thus formulate an optimal link creation problem subject to a constraint on the total weight of new links

$$\begin{array}{ll} \text{maximize} & f(U) \\ \text{subject to} & S = I - \text{diag}\{U\mathbb{1}\} \\ & U\mathbb{1} \leq \mathbb{1},\ U \geq 0,\ \|U\|_{\ell_1} \leq \varrho, \end{array}$$

where the optimization variables are the matrices U, S, all inequalities are elementwise, and $\|X\|_{\ell_1} := \sum_{i,j} |X_{ij}|$ for any matrix X. The objective function f characterizes the collective performance of the network, which in the present work is taken to be the efficiency with which the network achieves consensus.

In contrast to [12], [16], and motivated by the investigation of the properties of small-world networks [1]–[4], in this work we restrict attention to spatially invariant networks described by circulant matrices. This in particular implies that the matrices T, U, S are circulant and the preceding optimization problem can be rewritten as

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

where the optimization variables are the matrices U, S and the scalar d, and $1 \geq \delta := \varrho/n$. The parameter δ represents the total amount of resources available to each node with which to make new links. The 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 U to be circulant.

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 paper, 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.

III. Optimal Link Creation for Consensus Facilitation

In this section we first describe 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 graphical characteristics.

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 $x_{\text{ave}}(k) := \frac{1}{n}\sum_i x_i(k) = \frac{1}{n}\mathbb{1}^* x(k)$, and consider the distance from average consensus at time k,

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

where $Q := I - \frac{1}{n}\mathbb{1}\mathbb{1}^*$ and $*$ denotes the complex conjugate transpose.

Lemma 1: We have

$$\psi(k) = x(0)^* Q \,(T^*T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*)^k Q\, x(0). \tag{4}$$

The proof is simple and follows from the fact that all circulant matrices commute, and that $Q^2 = Q$, $Q\mathbb{1} = 0$, $T\mathbb{1} = \mathbb{1}$, $\mathbb{1}^*T = \mathbb{1}^*$.

It follows trivially from (4) and the definition of Q that

$$\psi(k) \le \|x(0) - x_{\rm ave}(0)\mathbb{1}\|_2^2\, \lambda_0^k, \;\; \lambda_0 := \lambda_{\max}(T^*T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*). \tag{5}$$

Furthermore, it is easy to show that $0 \le \lambda_0 < 1$, and that the inequality in (5) is satisfied with equality if $x(0)$ is aligned with the eigenvector corresponding to λ_0. Note that λ_0 is the square of the second-largest eigenvalue modulus (SLEM) [21] of the matrix 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 thus formulate the optimization problem

$$\begin{array}{ll} \text{minimize} & \lambda_{\max}\big((ST+U)^*(ST+U) - \frac{1}{n}\mathbb{1}\mathbb{1}^*\big) \\ \text{subject to} & S = (1-d)I, \; d \le \delta \\ & U\mathbb{1} = d\,\mathbb{1}, \; U \ge 0, \; UR = RU. \end{array} \tag{6}$$

In what follows, we demonstrate the convexity of (6) and rewrite it as a semidefinite program. Problem (6) is closely related to the fastest mixing Markov chain problem formulated in [21]. In contrast to [21], 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. Indeed, 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.

We hereafter replace the inequality constraint $d \le \delta$ in (6) with the equality constraint

$$d = \delta,$$

and claim that the resulting optimization problem is equivalent to (6), in the sense that both problems render the same augmented network. To elaborate, let $(U^{\rm opt}, d^{\rm opt})$ with $d^{\rm opt} < \delta$ be a solution of (6). Then it is easy to see that there always exists a matrix U, 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 $(U^{\rm opt}, d^{\rm opt})$; indeed, the matrix $U = (\delta - d^{\rm opt})T + U^{\rm opt}$ satisfies $U\mathbb{1} = \delta\mathbb{1}$ and uniquely solves the equation $(1-\delta)T + U = (1 - d^{\rm opt})T + U^{\rm opt}$.

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 (6). We now formulate the main optimization problem of this paper, namely the problem of optimal circulant link creation

$$\begin{array}{ll} \text{minimize} & \lambda_{\max}\big(T_u^*T_u - \frac{1}{n}\mathbb{1}\mathbb{1}^*\big) \\ \text{subject to} & U\mathbb{1} = \delta\mathbb{1}, \; U \ge 0, \; UR = RU, \end{array} \tag{7}$$

recalling that $\delta \le 1$. Note that when resources are abundant and $\delta = 1$, the globally optimal solution of (7) is given by $U = \frac{1}{n}\mathbb{1}\mathbb{1}^*$, which renders the complete graph with homogeneous links $T_u = \frac{1}{n}\mathbb{1}\mathbb{1}^*$, [21].

Problem (7) 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 (SDP)

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & \begin{bmatrix} \gamma I + \frac{1}{n}\mathbb{1}\mathbb{1}^* & T_u^* \\ T_u & I \end{bmatrix} \succeq 0 \\ & U\mathbb{1} = \delta\mathbb{1}, \; U \ge 0, \; UR = RU. \end{array} \tag{8}$$

Hereafter in this work, problems (7)–(8) will be the main focus of our investigations. In the rest of this section, we will demonstrate the topological and scaling properties of the solutions of (8) using numerical examples.

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

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

Indeed, $\xi(T) \to 0$ 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 (8) 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. We use the notation $X_{i,j}$ to refer to the (i,j)th entry of a matrix X. 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 $\Delta(T)$ denote the

[1] Part of the motivation for the definition of ξ 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 a temporal convergence-to-consensus of $\mathrm{e}^{-\xi t}$ for such a system, and comparing it to the convergence behavior $\lambda_0^k \equiv \mathrm{e}^{(\log\lambda_0)k}$ for a discrete-time consensus network (5), one obtains the correspondence $\xi \sim -\log\lambda_0$. Ref. [5] further normalizes ξ by considering ξ/ξ_0, where ξ_0 denotes the second smallest eigenvalue of the Laplacian of the original network.

matrix of node 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

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

Illustrative Example

We consider a graph with $n = 28$ 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 (8) appears in Fig. 1. For all computations we use CVX, a package for specifying and solving convex programs [24], [25].

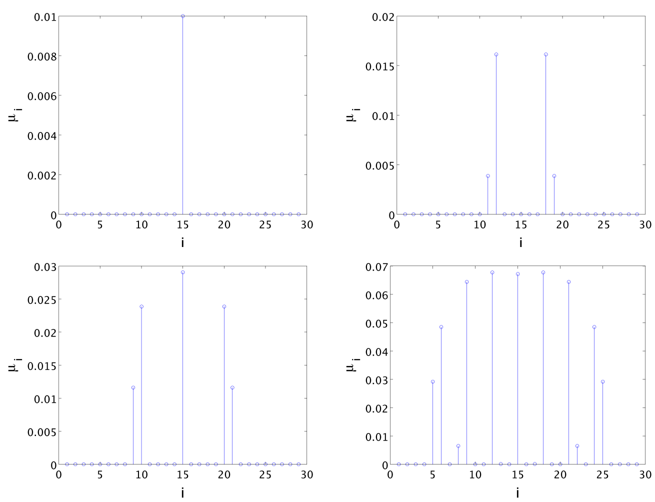

Fig. 1: The horizontal axis represents the indices $i = 1, \ldots, n$ of the entries of a vector in $\mathbb{R}^n$. The circular points give the values of the entries of the first column μ of the optimal U resulting from problem (8). The value for $i = 1$ is repeated at $i = n + 1$ to emphasize the circulant structure. The different plots correspond to different δ values, $\delta = 1/100, 1/25, 1/10, 1/2$. 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).

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 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 Sec. IV we will revisit the small-δ problem, which we refer to as the creation of 'weak' links, and will bring to bear perturbation methods to further explore the topology of optimal links.

Moreover, inspired by [5], we investigate changes in $\xi(T_u)/\xi(T)$ and $\ell(T_u)$ resulting from variations in δ. Fig. 2 demonstrates these changes for δ varying in $[0, 1]$, where T is taken as the same circulant matrix defined earlier. Fig. 2 is similar to findings in [5] and illustrates that even small values of δ in (8) lead to dramatic improvements in the rate of convergence-to-consensus for the augmented system.

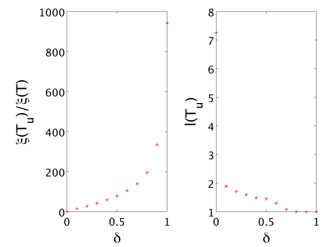

Fig. 2: The horizontal axes represent the value of δ. The plot on the left shows $\xi(T_u)/\xi(T)$ for U resulting from problem (8), and the one on the right shows $\ell(T_u)$. The values of δ vary in $1/10$ increments from $\delta = 0$ to $\delta = 1$.

Scaling-in-n of $\xi(T_u)$ and $\ell(T_u)$

Issues of scaling are prominent in network analysis and design. We consider the scaling-in-n of the solutions of (8) with respect to two characteristic quantities of dynamical networks: convergence rate and average path length.

We consider graphs with nearest neighbor interactions described by circulant matrices $T \in \mathbb{R}^{n \times n}$ whose first column τ is given by

$$\tau = \begin{bmatrix} \frac{10}{20} & \frac{3}{20} & 0 & \cdots & 0 & \frac{7}{20} \end{bmatrix}^*.$$

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 (8). 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 [24], [25].

Each subfigure in Fig. 3 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. 3 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 (8) is such that $\ell(T_u)$ remains approximately 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 modeling infrequent communication between the nodes they connect. The cohesive power of weak ties and their importance in the

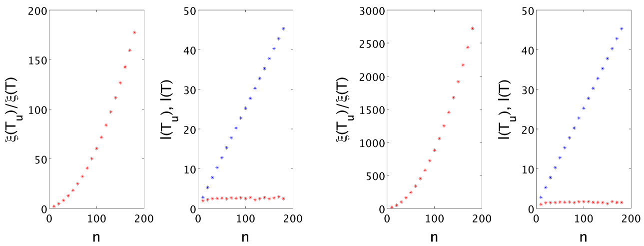

Fig. 3: The horizontal axes represent the size n of the network. In each pair of plots, the one on the left shows $\xi(T_u)/\xi(T)$ for U resulting from problem (8), and the one on the right shows $\ell(T_u)$ [in red] versus $\ell(T)$ [in blue]. The different plots correspond to different δ values, $\delta = 1/25, 1/2$.

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. All proofs are omitted due to space limitations and will be reported elsewhere.

We reconsider problem (7) under the assumption that δ is a small positive number ε, and hence the elementwise-positive matrix U has small entries

$$\begin{aligned} &\text{minimize} \quad \lambda_{\max}(T_u^* T_u - \tfrac{1}{n}\mathbb{1}\mathbb{1}^*) \\ &\text{subject to} \quad U\mathbb{1} = \varepsilon\mathbb{1},\ U \geq 0,\ UR = RU. \end{aligned} \tag{9}$$

The conceivable scenarios resulting from solutions of (9) range from every node making only one link of weight ε to every node making many links whose weights add up to ε. Prop. 5 and Prop. 6 in what follows will establish conditions on T such that the optimal U resulting from (9) corresponds to the generation of sparse and long-range links. We begin, however, by exploiting a perturbation framework in Props. 2–4 to simplify (9).

We replace the objective in (9) with its first-order Taylor approximation around $U = 0$,

$$\lambda_{\max}(T_{0+\varepsilon v}^* T_{0+\varepsilon v} - \frac{1}{n}\mathbb{1}\mathbb{1}^*) \approx \lambda_{\max}(T^*T - \frac{1}{n}\mathbb{1}\mathbb{1}^*) + \varepsilon g(V),$$

with matrix variable V and scalar-valued linear function g, to form the problem

$$\begin{aligned} &\text{minimize} \quad g(V) \\ &\text{subject to} \quad V\mathbb{1} = \mathbb{1},\ V \geq 0,\ VR = RV. \end{aligned} \tag{10}$$

To elaborate on the relationship between problems (10) and (9), let $\delta = \varepsilon$ be a given small number signifying the total weight of the new links to be created by each node, and let V solve (10). Then $U := \varepsilon V$ is an approximate solution to (9) for small enough values of ε.

Proposition 2: We have

$$g(V) = \lambda_{\max}\big(P^*(T^*V + V^*T)P\big) - 2\lambda_0,$$

where $P = \big[p_1, \ldots, p_r\big] \in \mathbb{C}^{n\times r}$ 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 thus formulate the optimal weak link creation problem

$$\begin{aligned} &\text{minimize} \quad \lambda_{\max}\big(P^*(T^*V + V^*T)P\big) \\ &\text{subject to} \quad V\mathbb{1} = \mathbb{1},\ V \geq 0,\ VR = RV. \end{aligned} \tag{11}$$

We state some definitions and notation before stating our next result. Let the unitary matrix F denote the discrete Fourier transform,

$$F_{l+1,m+1} := \tfrac{1}{\sqrt{n}}\, \mathrm{e}^{-\mathrm{i}2\pi lm/n}, \quad l, m = 0, \ldots, n-1. \tag{12}$$

We will hereafter use hatted variables to denote the Fourier transform of matrices and vectors,

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

for any n-vector x and n-by-n matrix X. It is well-known that all circulant matrices are simultaneously diagonalized by F [22], so that if X is circulant then $\hat{X}$ is diagonal. Let τ and $\hat{\tau}$ respectively denote the first column of the circulant matrix T and its Fourier symbol,

$$\tau := Te_1, \quad \hat{\tau} := F\tau, \tag{13}$$

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

$$\hat{\tau}_i = \tfrac{1}{\sqrt{n}}\hat{T}_{i,i}, \quad i = 1, \ldots, n, \tag{14}$$

with $\hat{T} := FTF^*$.[2]

Proposition 3: Let the real circulant matrix T be such that

$$|\hat{\tau}_2| > |\hat{\tau}_i|, \quad i = 3, \ldots, \lceil \tfrac{n+1}{2} \rceil, \tag{15}$$

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 $P = F^*\big[e_2, e_n\big]$, where e_i denotes the ith standard basis vector in $\mathbb{R}^n$.

Motivated by Prop. 3, hereafter in the paper we make the following assumption.

Assumption 1: The real circulant matrix T with $T\mathbb{1} = \mathbb{1}$ and $T \geq 0$ satisfies $|\hat{\tau}_2| > |\hat{\tau}_i|$ for $i = 3, \ldots, \lceil \frac{n+1}{2} \rceil$, with $\hat{\tau}_i$ defined as in (14)–(13).

When T is taken to represent social interactions for example, it is reasonable to expect that it satisfies Assumption 1. 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. This implies that any (appropriately shifted) row of T 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., $\|T\varphi_2\|_2 > \|T\varphi_i\|_2$ or equivalently $|\hat{\tau}_2| > |\hat{\tau}_i|$ for $i \neq 1, 2, n$.

[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 $\hat{\tau}_i = e_i^*\hat{\tau} = \frac{1}{\sqrt{n}}e_i^*\hat{T}\mathbb{1} = \frac{1}{\sqrt{n}}\hat{T}_{i,i}$.

Proposition 4: If Assumption 1 holds, then (11) is equivalent to the optimization problem

$$\begin{array}{ll} \text{minimize} & \text{trace}(K^* V) \\ \text{subject to} & \frac{1}{n}\|V\|_{\ell_1} = 1, \ V \geq 0, \ VR = RV, \end{array} \tag{16}$$

with $K := TPP^*$, and P defined as in Prop. 3.

It is not difficult to show that (16) is a linear program. The following proposition is a consequence of Prop. 4 and is one of the main results of this work.

Proposition 5: If Assumption 1 holds, then the solution of (11) corresponds to the generation of sparse links.

One can interpret K as a 'filtered' version of T, where only the first spatial harmonic of the rows/columns of T survives the filter PP^*, and the rows/columns of K inherit a pure cosine structure of period n and frequency $l = 1$.

Once again, if T is taken to represent social interactions where 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, the solution of (11) corresponds to the generation of long-range communication links. In particular, if T is additionally assumed to be symmetric, then $\hat{\tau} \in \mathbb{R}^n$, $K_{i,j} = \frac{2\sigma}{n}\cos(\frac{2\pi}{n}(i-j))$, and therefore the smallest entries of K occur at its $|i-j| \approx \lfloor \frac{n}{2} \rfloor$ subdiagonal. This motivates the following assumption.

Assumption 2: Given the real circulant matrix T with $T\mathbb{1} = \mathbb{1}$ and $T \geq 0$, there exists some positive integer $\overline{\eta} \ll n$ such that $|\theta| < \frac{2\pi}{n}\overline{\eta}$, with $\theta := \angle \hat{T}_{2,2} = \angle \hat{\tau}_2$.

We can now state another main result of this work.

Proposition 6: If Assumption 1 holds, and the real circulant matrix T is such that θ satisfies

$$-\frac{2\pi}{n}\eta - \frac{\pi}{n} < \theta < -\frac{2\pi}{n}\eta + \frac{\pi}{n} \tag{17}$$

for some $\eta \in \mathbb{Z}$, then the solution of (11) corresponds to the generation of a link between every node and its $(\lfloor \frac{n}{2} \rfloor + \eta)$th neighbor. Furthermore, if Assumption 2 holds, then for large n the solution of (11) corresponds to the generation of long-range links.

In summary, Assumption 1 and Prop. 5 ensure the sparsity of optimal weak links. However, while Assumption 1 restricts the rows (and columns) of T to have a strong first harmonic component, it does not restrict spatial shifts of this component. This motivates Assumption 2, which further limits the rows of T to being cosine-like, in the sense that T should attain its largest and smallest entries respectively close to its main diagonal and its $\lfloor \frac{n}{2} \rfloor$th subdiagonal. Together, Assumption 2 and Prop. 6 ensure the long-range property of optimal weak links.

V. Conclusions

We consider the problem of optimal link creation for the enhancement of convergence to consensus in circulant networks. We formulate this problem as a convex program and show the favorable scaling properties of its solutions. We use a perturbation framework to further derive conditions on the network architecture that guarantee optimal weak links to be sparse and long-range. Our findings give credibility to the conjecture that the small-world networks observed across different scientific disciplines are a result of balancing the use of resources, as expended in the creation of new communication or interconnection pathways, and achieving high collective performance.

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 *Proceedings of the 2005 American Control Conference*, 2005, pp. 2371–2378.

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

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

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

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

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

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

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

[13] M. Fardad, "On the optimality of sparse long-range links in circulant consensus networks," submitted to *2015 American Control Conference*, 2014.

[14] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, 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," in *Proceedings of the 2014 American Control Conference*, 2014, pp. 3790–3795.

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

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

[19] J. S. Baras and P. Hovareshti, "Efficient and robust communication topologies for distributed decision making in networked systems," in *Proceedings of the 48th IEEE Conference on Decision and Control*, 2009, pp. 3751–3756.

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

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

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

[23] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 634–649, 2008.

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

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