2015 IEEE 54th Annual Conference on Decision and Control (CDC)
December 15-18, 2015. Osaka, Japan

On Consensus-Based Community Detection

Makan Fardad

Abstract— **We consider networks in which every node updates its value in discrete time by taking a weighted average of the values of the nodes it interacts with. Using an objective function that quantifies the efficiency with which clusters of interacting nodes converge to consensus internally, we formulate an optimization problem that identifies distinct communities in the network. The optimal community detection problem is combinatorial in nature and intractable in general, and we use convex relaxations to reformulate the problem as a semidefinite program. We demonstrate the utility of our algorithm by applying it to some benchmark graphs from the network science literature.**

Index Terms— **Clustering, community detection, consensus, convex relaxation, dynamical systems, graph partitioning, optimization, semidefinite programming, social networks.**

I. INTRODUCTION

Community detection algorithms are of interest in a range of biological, social, and technological networks; examples include the identification of proteins groups that collectively serve a certain function within the cell [1]–[3], the detection of criminal organizations using cell phone metadata [4], and the partitioning of power networks for identifying slow-coherent areas [5], [6] and for islanding under contingencies [7]. Community detection and clustering problems are inherently combinatorial and require approximations and/or heuristics to make them tractable [8].

Among the many existing community detection algorithms [9], [10], those based on modularity maximization and spectral methods have dominated the literature [8], [11]–[16]. It has been argued that optimizing modularity suffers from a resolution limit and may over- or under-partition the network [9], [10], [17], [18], thus missing its natural community architectures. Furthermore, these methods are rooted only in the topological aspects of a network's graph, are insensitive to its temporal dynamics and time scales, and are sometimes incapable of detecting even weakly connected clusters [19]–[21].

Reference [17] employs random walks on graphs to provide a dynamical approach to community detection. Using Markov chains and temporal autocovariances, a partitioning of a graph is categorized as high quality for a given time horizon if a random walk starting within a cluster spends most of its time there before escaping. Although [17] uses temporal dynamics in its characterization of cluster quality, it still requires an exhaustive search over all possible partitions of the network. The authors thus rely on heuristic-based methods to obtain tractable optimization algorithms. Reference [22] considers an extension of the Krause opinion dynamics model to characterize and detect communities as asymptotically connected components of the network, while [23] employs the maximum entropy principle to formulate and solve dynamic clustering problems.

In this paper we employ convex optimization, and a consensus-based objective function, to identify communities in networks of dynamical agents. We consider networks in which links are weighted and directed, and nodes execute a consensus protocol based on DeGroot dynamics [24]. We first quantify the efficiency with which clusters of nodes converge to consensus internally. We then formulate an optimization problem that identifies communities which maximize this efficiency. Our formulation employs the Lyapunov equation to characterize cumulative temporal behavior while effectively eliminating the time variable. The resulting optimization problem is combinatorial and nonconvex, and we use convex relaxations to reformulate the problem as a semidefinite program. Finally, we apply our detection algorithm to some benchmark graphs from the network science literature.

Given that the proposed detection algorithm allows for weighted and directed graphs, accounts for the network's temporal dynamics, and finds communities optimally based on their ability to reach consensus, it is expected that it will perform better than other existing clustering methods in applications where consensus behavior (such as synchronization performance in power networks) is of primary concern.

II. PROBLEM FORMULATION

In this section we begin by describing a linear dynamical model for the evolution of node values in a network. We derive a performance objective that quantifies the efficiency with which communities of nodes converge to consensus internally. We then formulate an optimization problem that uses this consensus-based performance objective to find the best clustering of the network. Hereafter in this paper we use the words 'cluster' and 'community' interchangeably.

A. Temporal Dynamics

We consider a network in which node values evolve according to DeGroot dynamics [24]

$$x(t+1) = T\,x(t), \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the vector of node values at time t and $T \in \mathbb{R}^{n\times n}$ is a (not necessarily symmetric) matrix such that

$$T\,\mathbb{1}_n = \mathbb{1}_n, \quad T \geq 0. \tag{2}$$

Here, $\mathbb{1}_n$ is defined as the column n-vector of all ones and $\geq$ denotes elementwise matrix inequality. Henceforth we

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.

978-1-4799-7885-4/15/$31.00 ©2015 IEEE

will drop the subscript n and write $\mathbb{1}$ instead of $\mathbb{1}_n$, but will reinstate the subscript for emphasis or when the vector of all ones has dimension other than n.

If we interpret nodes as agents and node values as agents' beliefs, then the DeGroot model (1)-(2) implies that at every time instant agents update their beliefs by taking a weighted average of their own beliefs and the beliefs of those they interact with. If all diagonal entries of T are positive and the graph described by T is strongly connected, then all eigenvalues of T except for the one at $\lambda = 1$ belong to the open unit disk, and the network (1)-(2) will reach consensus asymptotically [24], [25]. Finally, if the system 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 $t \geq 0$.

B. Quadratic Measure for Efficiency of Reaching Consensus

Reference [26] considers a measure of consensus-reaching efficiency that accounts for the cumulative-over-time fluctuation of beliefs around consensus. When the matrix Q is chosen such that

$$Q \succeq 0, \quad \mathbb{1}^T Q\, \mathbb{1} = 0, \tag{3}$$

with $\succeq$ denoting matrix inequality with respect to the positive semidefinite cone, then the quadratic form $x(t)^T Q\, x(t)$ characterizes the *deviation of beliefs from consensus* at t. In particular, if $Q = I - \frac{1}{n}\mathbb{1}\mathbb{1}^T$ then $x(t)^T Q\, x(t)$ yields the deviation from average at time t, i.e., $x^T Q\, x = \sum_{i=1}^n (x_i - x_{\mathrm{ave}})^2$ with $x_{\mathrm{ave}} := \frac{1}{n}\sum_{i=1}^n x_i$. Defining J as the average of $\sum_{t=0}^\infty x(t)^T Q\, x(t)$ as $x(0)$ changes over a set of orthonormal vectors that span $\mathbb{R}^n$, it is shown in the appendix that

$$J = \mathrm{trace}(P), \tag{4}$$

where $P \succeq 0$ satisfies the algebraic Lyapunov equation

$$P = (T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^T)^T P\, (T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^T) + Q, \tag{5}$$

and $P\,\mathbb{1} = 0$. For a given Q, the smaller J is the more efficient the convergence of beliefs to the consensus value.

C. Efficiency of Clusters in Reaching Consensus Internally

We utilize (3)-(5) to quantify the efficiency with which communities of nodes reach consensus internally, where Q will now additionally serve to characterize the clusters.

Consider a network with m clusters and n_k nodes in cluster k. Without loss of generality, let nodes be numbered such that those with indices $1, \ldots, n_1$ belong to cluster 1, nodes with indices $n_1 + 1, \ldots, n_1 + n_2$ belong to cluster 2, and so on for all clusters $k = 1, \ldots, m$. Let Q be a block-diagonal matrix $Q = \mathrm{diag}\{Q_k\}_{k=1}^m$, $Q_k \in \mathbb{R}^{n_k \times n_k}$. This special structure of Q implies that $x^T Q\, x = \sum_{k=1}^m x_k^T Q_k\, x_k$, where x_k denotes the state vector corresponding to nodes in cluster k. In particular, if $Q_k = I_{n_k} - \frac{1}{n_k}\mathbb{1}_{n_k}\mathbb{1}_{n_k}^T$ then $x_k^T Q_k\, x_k$ quantifies the deviation of beliefs in cluster k from the cluster average, where I_{n_k} denotes the identity matrix of dimension n_k.

Let us now consider an exaggerated scenario in which the links between clusters are few and weak while the links within clusters are many and strong. In this case the value of $\sum_{t=0}^\infty x(t)^T Q\, x(t)$ is small for a network that has been initialized to some initial vector of beliefs $x(0)$. To elaborate, within any given cluster all nodes are strongly coupled and therefore will quickly converge to a (perhaps slowly varying) common value. Therefore, within cluster k the deviations of node values from the cluster average, as quantified by $x_k(t)^T Q_k\, x_k(t)$, remains small for all t. Summing over all clusters gives that $x(t)^T Q\, x(t)$ is small for all time.

Of course, for a general large network and desired number of clusters m, it is unlikely that a partitioning with weak inter-cluster coupling exists or can be easily identified by inspection. The problem thus becomes one of finding m clusters that minimize J in (4)-(5), with the matrix Q acting as an optimization variable whose structure describes the clusters.

We define the auxiliary variable $K \in \{0, 1\}^{n \times m}$ to be a binary matrix that characterizes both Q and the m clusters in the network as follows. Let

$$K_{\eta\kappa} = \begin{cases} 1 & \text{if node } \eta \text{ belongs to cluster } \kappa \\ 0 & \text{otherwise,} \end{cases}$$

where $K_{\eta\kappa}$ denotes the (η, κ)th entry of K, and set

$$X = KK^T, \quad D = \mathrm{diag}(X\mathbb{1}).$$

Now define $Q = I - D^{-1}X$, or equivalently

$$\begin{aligned} Q &:= D^{-1}(D - X) \\ &= D^{-1/2}(D - X)D^{-1/2}. \end{aligned}$$

The matrix Q, as formed above, has the following properties:

- Q satisfies (3), and
- with an appropriate relabeling of the nodes, Q is block-diagonal with kth block $Q_k = I_{n_k} - \frac{1}{n_k}\mathbb{1}_{n_k}\mathbb{1}_{n_k}^T$.

We use an example to elucidate the equations relating Q, D, X, and K. Consider a network of five nodes in which we have identified the optimal two clusters to be composed of nodes $\{1, 2, 3\}$ and $\{4, 5\}$, so that $n_1 = 3$ and $n_2 = 2$. Then

$$K = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix},\ X = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix},\ D = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \tag{6}$$

$$D - X = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix},\ Q = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & 0 & 0 \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

where $Q = \mathrm{diag}\{Q_1, Q_2\}$ with $Q_1 = I_3 - \frac{1}{3}\mathbb{1}_3\mathbb{1}_3^T$ and $Q_2 = I_2 - \frac{1}{2}\mathbb{1}_2\mathbb{1}_2^T$.

Finally, due to the special structure of K, the positive semidefinite matrix $X = KK^T$ always satisfies

$$X \circ I = I, \quad X \in \{0, 1\}^{n \times n}, \tag{7}$$

where $\circ$ denotes elementwise matrix multiplication.

D. Consensus-Based Community Detection

We now formulate the problem of finding the optimal m clusters that minimize a consensus-based objective in a network of n nodes as

$$\begin{array}{ll}\text{minimize} & \operatorname{trace}(P)\\ \text{subject to} & P-(T-\tfrac{1}{n}\mathbb{1}\mathbb{1}^T)^TP(T-\tfrac{1}{n}\mathbb{1}\mathbb{1}^T)=D^{-1}(D-X)\\ & P\succeq 0,\ D=\operatorname{diag}(X\mathbb{1}),\ X=KK^T\\ & K\mathbb{1}_m=\mathbb{1}_n,\ K\in\{0,1\}^{n\times m},\end{array} \tag{8}$$

where the optimization variables are the matrices P, D, X, and K.

Problem (8) is challenging due to the nonlinearity of the Lyapunov equality constraint in the optimization variable D, the nonlinearity of the equality constraint $X = KK^T$, and the existence of binary constraints. In order to restrict the enlargement of the feasible set when the nonconvex constraints are relaxed later in Sec. III, we endow problem (8) with the redundant constraints in (7) to obtain

$$\begin{array}{ll}\text{minimize} & \operatorname{trace}(P)\\ \text{subject to} & P-(T-\tfrac{1}{n}\mathbb{1}\mathbb{1}^T)^TP(T-\tfrac{1}{n}\mathbb{1}\mathbb{1}^T)=D^{-1}(D-X)\\ & P\succeq 0,\ D=\operatorname{diag}(X\mathbb{1}),\ X=KK^T\\ & X\circ I=I,\ X\in\{0,1\}^{n\times n}\\ & K\mathbb{1}_m=\mathbb{1}_n,\ K\in\{0,1\}^{n\times m}.\end{array} \tag{9}$$

Problem (9) is the main focus of this work. We note that if the number of clusters m was allowed to vary and was itself an optimization variable, then the solution of (9) would be given by $K = X = D = I$, every node would be a cluster on its own, the right-hand side of the Lyapunov equation in (9) would be $Q = D^{-1}(D - X) = 0$, and $P = 0$ would minimize the objective.

III. Relaxed Reformulation of Consensus-Based Community Detection

In this section we relax the nonconvex constraints in (9) in such a way that the resulting problem is a semidefinite program (SDP).

Returning to problem (9) we convexify the Lyapunov constraint by dropping the scaling factor D^{-1} from the right hand side, we replace the nonconvex equality constraint $X = KK^T$ with the convex inequality constraint $X \succeq KK^T$, and we relax both of the binary constraints $X \in \{0,1\}^{n\times n}$ and $K \in \{0,1\}^{n\times m}$ to their convex hulls. We thus arrive at the convex optimization problem

$$\begin{array}{ll}\text{minimize} & \operatorname{trace}(P)\\ \text{subject to} & P-(T-\tfrac{1}{n}\mathbb{1}\mathbb{1}^T)^TP\,(T-\tfrac{1}{n}\mathbb{1}\mathbb{1}^T)=D-X\\ & P\succeq 0,\ D=\operatorname{diag}(X\mathbb{1}),\ X\succeq KK^T\\ & X\circ I=I,\ X\in[0,1]^{n\times n}\\ & K\mathbb{1}_m=\mathbb{1}_n,\ K\in[0,1]^{n\times m}.\end{array} \tag{10}$$

Using the Schur complement, the inequality constraint $X \succeq KK^T$ can further be rewritten as

$$\begin{bmatrix} X & K\\ K^T & I\end{bmatrix}\succeq 0,$$

rendering problem (10) a semidefinite program that can be solved efficiently.

To elaborate on dropping D^{-1} from the Lyapunov equality constraint, we observe that $D - X$ still satisfies the two conditions (3) required of the matrix on the right hand side of the Lyapunov equation, namely $D - X \succeq 0$ and $\mathbb{1}^T(D - X)\mathbb{1} = 0$. Furthermore, if the clusters are roughly the same size then D will be close to a multiple of the identity; therefore, its elimination from the Lyapunov equation can be approximated by a rescaling of the matrix P and will not significantly affect the solution of the optimization problem.

Remark: Another method for forming a linear approximation of the Lyapunov equation in (9) is to replace its right hand side with $\widetilde{D}^{-1/2}(D-X)\widetilde{D}^{-1/2}$, where $\widetilde{D}^{-1/2}$ serves as a proxy for $D^{-1/2}$. Starting from $\widetilde{D}^{-1/2} = I$, we iteratively solve (10) with its Lyapunov equation replaced by

$$P-(T-\tfrac{1}{n}\mathbb{1}\mathbb{1}^T)^TP(T-\tfrac{1}{n}\mathbb{1}\mathbb{1}^T)=\widetilde{D}^{-1/2}(D-X)\widetilde{D}^{-1/2}. \tag{11}$$

The solution D from every iteration is then employed, by setting $\widetilde{D} := D$, to form $\widetilde{D}^{-1/2}$ for the next iteration. The reason for using $\widetilde{D}^{-1/2}(D - X)\widetilde{D}^{-1/2}$ rather than $\widetilde{D}^{-1}(D - X)$ in (11) is that, when the entries of X are not purely binary, the latter matrix may fail to be symmetric whereas the former is symmetric by construction. Finally, if the iterations converge and X is binary, then (11) and the Lyapunov equation in (9) are identical. We demonstrate an application of this method in Example 3 of Sec. IV.

IV. Examples

In this section we illustrate the utility of problem (10) through some examples. For all computations we use `CVX`, a package for specifying and solving convex programs [27], [28]. We use `Cytoscape` to plot all network graphs.

Example 1: 5-Node Network

We begin with a simple 5-node network described by

$$T=\begin{bmatrix}\frac{2}{3} & \frac{1}{3} & 0 & 0 & 0\\ \frac{2}{5} & \frac{3}{5}-\frac{1}{6} & \frac{1}{6} & 0 & 0\\ 0 & \frac{1}{5} & 1-\frac{1}{5}-\frac{1}{7} & \frac{1}{7} & 0\\ 0 & 0 & \frac{1}{8} & \frac{3}{5}-\frac{1}{8} & \frac{2}{5}\\ 0 & 0 & 0 & \frac{1}{3} & \frac{2}{3}\end{bmatrix}. \tag{12}$$

Here, nodes $1, 2$ have strong interaction between them, as do nodes $4, 5$, whereas the interaction between node 3 and either pair is relatively weaker. Furthermore, node 3 is more strongly coupled with nodes $1, 2$ than it is with nodes $4, 5$.

Solving (10) for $m = 2$ yields the exact X matrix given in (6), which identifies the optimal two clusters $\{1, 2, 3\}$ and $\{4, 5\}$. Fig. 1 (left) shows a color plot of X; we adopt this pictorial representation of matrices hereafter.

Solving (10) for $m = 3$ yields the X matrix shown in Fig. 1 (right), which identifies the optimal three clusters

$\{1,2\}$, $\{3\}$, and $\{4,5\}$, as expected. In this case we also observe that some entries of X slightly deviate from purely binary values.

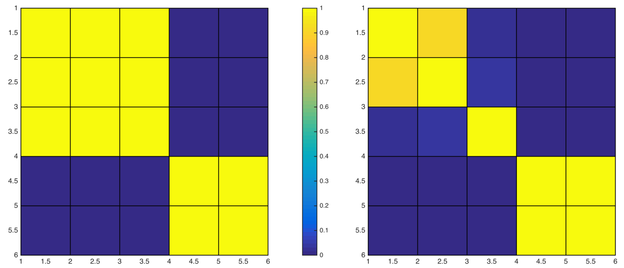

Fig. 1: Community detection for 5-node network described by (12). Left: X matrix resulting from (10) for $m = 2$. Right: X matrix resulting from (10) for $m = 3$.

Example 2: Zachary's Karate Club

We consider the well-known 34-node karate club network [29] in Fig. 2 (top). This network describes a real-life university-based karate club whose evolution was documented for three years, and in which nodes 1 and 34 respectively represent the club's instructor and its president. Due to a dispute over lesson prices, the group separated into two clubs centered around the instructor and the president. Thus, for this particular example there exists a de facto partitioning of the network with which we can compare the result of our clustering algorithm.

We use the weighted adjacency matrix A given in [29] to obtain T from[1]

$$T := S^{-1}(A + 5I), \quad S := \mathrm{diag}((A + 5I)\mathbb{1}).$$

The role of the diagonal matrix S is to normalize the rows of $A + 5I$ so that $T\mathbb{1} = \mathbb{1}$. The rationale for adding $5I$ to the zero-diagonal matrix A is as follows. Since almost all nonzero entries of A in [29] have values between 1 and 5, by adding $5I$ to A we insure that at every time instant each agent will update its belief by giving at least as much value to its own belief as it gives to anyone else's.

Solving (10) for $m = 2$ yields the X matrix shown in Fig. 2 (center), which identifies the optimal two clusters demonstrated in Fig. 2 (bottom). All entries of X are binary. Except for node 9, our clusters are in perfect agreement with the actual splitting of the network reported in [29]. In [29] Zachary used a minimum cut analysis of the graph to investigate whether he could have predicted the split; the clustering that results from his analysis is identical to that in Fig. 2 (bottom). Indeed, Zachary himself comments on the anomaly of agent 9, and that at the time of the dispute between agent 1 (the instructor) and agent 34 (the president), agent 9 was just a few weeks away from his test for black belt; aligning himself with the president would have been tantamount to him giving up his rank and starting out once again as a white belt after the splitting of the club [29]. Therefore, while agent 9 was a supporter of the president, he chose to side with the instructor, a diplomatic move not reflected in the adjacency matrix A.

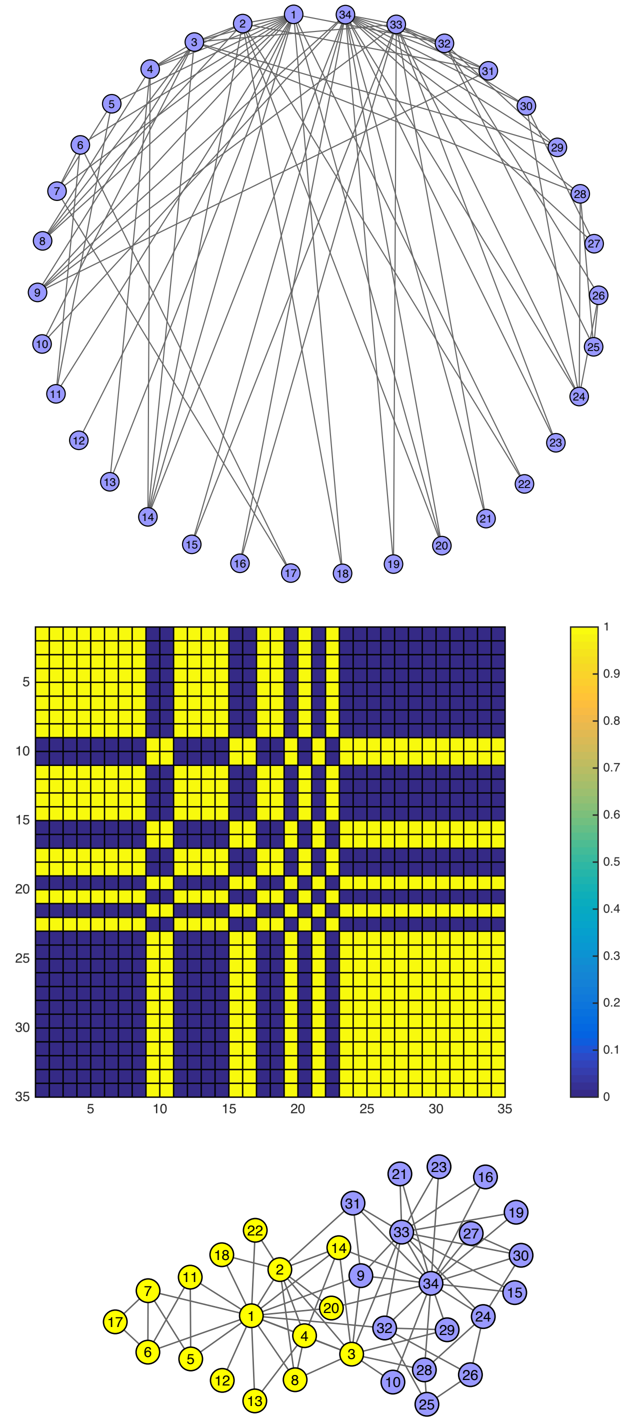

Fig. 2: Community detection for Zachary's karate club network [29]. Top: Graphical representation of network. Center: X matrix resulting from (10) for $m = 2$. Bottom: Graphical representation of two clusters.

Finally, since in many practical scenarios information about the values of link weights is not available, we repeat the community detection exercise assuming only knowledge

[1] Zachary's paper [29] contains inconsistencies in the network's adjacency matrix, in particular $A(23, 34) = 0$ and $A(34, 23) \neq 0$, rendering $A \neq A^T$. Since in much of the network science literature it is assumed that there exists a link between nodes 23 and 34, we make the same assumption here.

of the standard (unweighted) adjacency matrix. We let

$$T := S^{-1}(A + I), \quad S := \mathrm{diag}((A + I)\mathbb{1}).$$

By adding I to A we insure that at every time instant each agent updates its belief by taking an unweighted average of its own belief and the beliefs of those it interacts with.

Solving (10) for $m = 2$ yields the same clustering as in Fig. 2 (bottom). However, row/column 9 of X in this case incurs a small deviation from purely binary values.

Example 3: Dolphin Social Network

We consider the 62-node dolphin social network [30]–[32] in Fig. 3 (top). This network describes a social network of frequent associations between 62 dolphins in a community living off Doubtful Sound, New Zealand. We employ the procedure described in the preceding example to find the evolution matrix T from the adjacency matrix A, namely[2]

$$T := S^{-1}(A + I), \quad S := \mathrm{diag}((A + I)\mathbb{1}).$$

Solving (10) iteratively for $m = 2$ and with its Lyapunov equation replaced by (11), as described in the remark in Sec. III, yields the X matrix shown in Fig. 2 (center), which identifies the optimal two clusters demonstrated in Fig. 2 (bottom). The algorithm converges in five iterations and all entries of X are binary. We note that if we just solve (10) as we did in previous examples (i.e., without iterations and change of its Lyapunov equation) then some rows/columns of X incur deviations from binary values; furthermore, nodes $43, 48$ will belong to the complementary cluster, resulting in an objective value $\mathrm{trace}(P)$ that is larger than that corresponding to the clustering in Fig. 2 (bottom).

V. Conclusions & Future Directions

We use a consensus- and optimization-based framework to detect communities in networks of dynamical agents. We consider networks whose temporal dynamics are governed by the DeGroot model. We propose an optimization problem that seeks clusters of nodes capable of achieving efficient consensus internally. Given that our proposed detection algorithm allows for weighted and directed graphs, accounts for the network's temporal dynamics, and finds clusters optimally based on their ability to reach consensus, it is expected that it will perform better than existing methods in applications where consensus behavior is of primary concern.

There are multiple directions for future research. We would like to gain an understanding of the scenarios in which the matrix X resulting from (10) is binary. It is of interest to investigate the classes of networks for which other clustering algorithms yield communities similar to those given by the algorithm presented here, and also classes of networks for which the obtained communities are very different. Application of the developed clustering algorithm to the partitioning of power networks is an area of immediate appeal. Another direction for future work is to explore the effect of time scales on the detection of

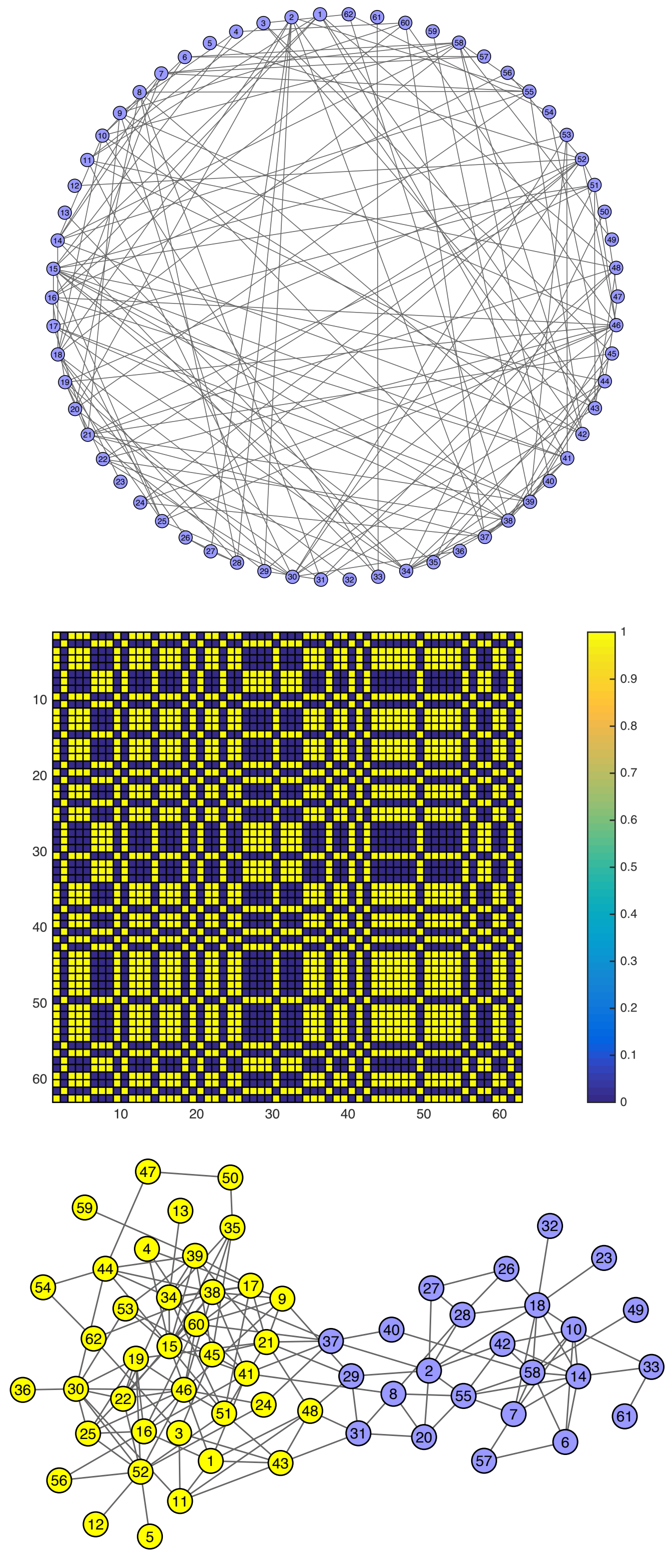

Fig. 3: Community detection for the dolphin social network [30]–[32]. Top: Graphical representation of network. Center: X matrix resulting from iteratively solving (10)-(11) for $m = 2$. Bottom: Graphical representation of two clusters.

[2]References [30]–[32] refer to the dolphins by their names. Here, we follow Mark Newman's labeling of the dolphins with numbers. The correspondence between dolphin names and their numbers can be found at `http://www-personal.umich.edu/~mejn/netdata/` with the caveat that node η in Newman's labeling corresponds to node $\eta + 1$ in this paper (Newman labels nodes starting from 0 whereas here we start from 1).

communities. To account for different time scales, one can replace the matrix $T - \frac{1}{n}\mathbb{1}\mathbb{1}^T$ with $\tau(T - \frac{1}{n}\mathbb{1}\mathbb{1}^T)$ in the Lyapunov equation; by manipulating the positive scalar τ, and thus rescaling the spectrum of $T - \frac{1}{n}\mathbb{1}\mathbb{1}^T$, one can vary the speed of convergence to consensus both within the clusters and across the network.

Accounting for temporal dynamics at the nodes and finding clusters based on their collaborative performance comes at a cost; namely, our SDP-based community detection algorithm is more computationally intensive than, and does not scale nearly as well as, some of the best methods reported in the literature, e.g. Blondel *et al.* A first step toward improving computational efficiency is the development of customized methods to replace off-the-shelf SDP solvers. In the long term, however, it is desired to find other measures of consensus-reaching efficiency that are more amenable to large-scale optimization.

APPENDIX

Derivation of (4)-(5): Let e_i denote the ith standard basis vector in $\mathbb{R}^n$, and let $x(0) = e_i$. This can be interpreted as a network in which all nodes have zero initial value except for the ith node, whose value is equal to one. From (1) it follows that the propagation of this initial value through the network is described by $x(t) = T^t e_i$, and

$$\begin{aligned} x(t)^T Q\, x(t) &= e_i^T T^{tT} Q\, T^t e_i \\ &= e_i^T (T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^T)^{tT} Q\, (T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^T)^t e_i \end{aligned}$$

measures quadratic deviation from average consensus at time t; here we have used $Q\,\mathbb{1} = 0$ and $T\,\mathbb{1} = \mathbb{1}$ to replace T^t with $(T - \frac{1}{n}\mathbb{1}\mathbb{1}^T)^t$. Summing over t, i, and defining

$$J := \textstyle\sum_{i=1}^n \sum_{t=0}^\infty e_i^T (T - \frac{1}{n}\mathbb{1}\mathbb{1}^T)^{tT} Q\, (T - \frac{1}{n}\mathbb{1}\mathbb{1}^T)^t e_i$$

as a cumulative measure of consensus-reaching efficiency, it is easy to show that $J = \operatorname{trace}(P)$, where P is a positive semidefinite matrix that satisfies the Lyapunov equation

$$P = (T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^T)^T P\, (T - \tfrac{1}{n}\mathbb{1}\mathbb{1}^T) + Q.$$

Note that the initial conditions $x(0)$ need not be restricted to varying only over $\{e_i\}_{i=1}^n$; indeed, any complete set of orthonormal vectors $\{\varphi_i\}_{i=1}^n$ would give the same result. Finally, the reason for subtracting $\frac{1}{n}\mathbb{1}\mathbb{1}^T$ from T is to eliminate the eigenvalue $\lambda = 1$ of T and render the spectrum of $T - \frac{1}{n}\mathbb{1}\mathbb{1}^T$ inside the open unit disk, thus guaranteeing that (5) has a well-defined solution $P \succeq 0$.

REFERENCES

[1] J. Chen and B. Yuan, "Detecting functional modules in the yeast protein-protein interaction network," *Bioinformatics*, vol. 22, pp. 2283–2290, 2006.

[2] V. Spirin and L. A. Mirny, "Protein complexes and functional modules in molecular networks," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 12 123–12 128, 2003.

[3] A. W. Rives and T. Galitski, "Modular organization of cellular networks," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 1128–1133, 2003.

[4] E. Ferrara, P. de Meo, S. Catanese, and G. Fiumara, "Detecting criminal organizations in mobile phone networks," 2014, arXiv:1404.1295.

[5] J. H. Chow, J. Cullum, and R. A. Willoughby, "A sparsity-based technique for identifying slow-coherent areas in large power systems," *IEEE Transactions on Power Apparatus and Systems*, vol. 103, pp. 463–473, 1984.

[6] J. H. Chow and P. V. Kokotovic, "Time scale modeling of space dynamic networks," *IEEE Transactions on Automatic Control*, vol. 30, pp. 714–722, 1985.

[7] G. B. Giannakis, V. Kekatos, N. Gatsis, S.-J. Kim, H. Zhu, and B. F. Wollenberg, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 30, pp. 107–128, 2013.

[8] M. E. J. Newman, "Spectral methods for community detection and graph partitioning," *Physical Review E*, vol. 88, p. 042822, 2013.

[9] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, pp. 75–174, 2010.

[10] M. A. Porter, J.-P. Onnela, and P. J. Mucha, "Communities in networks," *Notices of the AMS*, vol. 56, pp. 1082–1166, 2009.

[11] K. Chaudhuri, F. Chung, and A. Tsiatas, "Spectral clustering of graphs with general degrees in the extended planted partition model," in *Proceedings of the 25th Annual Conference on Learning Theory*, 2012, pp. 35.1–35.23.

[12] E. A. Leicht and M. E. J. Newman, "Community structure in directed networks," *Physical Review Letters*, vol. 100, p. 118703, 2008.

[13] V. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 10, p. P10008, 2008.

[14] M. E. J. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 8577–8582, 2006.

[15] R. Kannan, S. Vempala, and A. Vetta, "On clusterings: Good, bad and spectral," *Journal of the ACM*, vol. 51, pp. 497–515, 2004.

[16] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 888–905, 2000.

[17] J.-C. Delvenne, S. N. Yaliraki, and M. Barahona, "Stability of graph communities across time scales," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 12 755–12 760, 2010.

[18] S. Fortunato and M. Barthelemy, "Resolution limit in community detection," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 36–41, 2007.

[19] A. Lancichinetti and S. Fortunato, "Consensus clustering in complex networks," *Scientific Reports*, vol. 2, p. 336, 2012.

[20] ——, "Limits of modularity maximization in community detection," *Physical Review E*, vol. 84, p. 066122, 2011.

[21] B. H. Good, Y.-A. de Montjoye, and A. Clauset, "Performance of modularity maximization in practical contexts," *Physical Review E*, vol. 81, p. 046106, 2010.

[22] I.-C. Morărescu and A. Girard, "Opinion dynamics with decaying confidence: Application to community detection in graphs," *IEEE Transactions on Automatic Control*, vol. 56, pp. 1862–1873, 2011.

[23] P. Sharma, S. M. Salapaka, and C. L. Beck, "Entropy-based framework for dynamic coverage and clustering problems," *IEEE Transactions on Automatic Control*, vol. 57, pp. 135–150, 2012.

[24] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, pp. 118–121, 1974.

[25] 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.

[26] 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.

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

[28] ——, "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.

[29] W. W. Zachary, "An information flow model for conflict and fission in small groups," *Journal of Anthropological Research*, vol. 33, pp. 452–473, 1977.

[30] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, "The bottlenose dolphin community of doubtful sound features a large proportion of long-Lasting associations," *Behavioral Ecology and Sociobiology*, vol. 54, pp. 396–405, 2003.

[31] D. Lusseau, "The emergent properties of a dolphin social network," *Proceedings of the Royal Society of London B (Suppl.)*, vol. 270, pp. S186–S188, 2003.

[32] ——, "Evidence for social role in a dolphin social network," http://arxiv.org/abs/q-bio/0607048, 2006.