Identification of Overlapping Communities via Constrained Egonet Tensor Decomposition

Fatemeh Sheikholeslami and Georgios B. Giannakis

***Abstract*—Detection of overlapping communities in real-world networks is a generally challenging task. Upon recognizing that a network is in fact the union of its egonets, a novel network representation using multiway data structures is advocated in this contribution. The introduced sparse tensor-based representation exhibits richer structure compared to its matrix counterpart and, thus, enables a more robust approach to community detection. To leverage this structure, a constrained tensor approximation framework is introduced using PARAFAC decomposition. The arising constrained trilinear optimization is handled via alternating minimization, where intermediate subproblems are solved using the alternating direction method of multipliers to ensure convergence. The factors obtained provide soft community memberships, which can further be exploited for crisp, and possibly-overlapping community assignments. The framework is further broadened to include time-varying graphs, where the edgeset as well as the underlying communities evolve through time. The performance of the proposed approach is assessed via tests on benchmark synthetic graphs as well as real-world networks. As corroborated by numerical tests, the proposed tensor-based representation captures multihop nodal connections, that is, connectivity patterns within single-hop neighbors, whose exploitation yields a more robust community identification in the presence of mixing as well as overlapping communities.**

***Index Terms*—Community detection, overlapping communities, egonet subgraphs, tensor decomposition, constrained PARAFAC, sparse tensors.**

I. Introduction

GRAPH representation of complex real-world networks provides an invaluable tool for analysis and discovery of intrinsic attributes present in social, biological, and financial networks. One such attribute is the presence of small subgraphs, referred to as "communities" or "clusters," whose dense intra-connections and sparse inter-connections often represents a potential "association" among the participating entities (nodes). The task of community identification targets the discovery of such highly-interwoven nodes, and is of paramount interest in areas as diverse as unveiling functional modules in biological networks such as brain [1], trend analysis in social media [2], [3], and clustering of costumers in recommender systems [4].

Past works on community detection include those based on modularity-maximization [5], [6], generative and statistical models such as mixed membership stochastic block models (MMSB) [7]–[12], local-metric optimization[13], spectral clustering [14], and matrix factorization [15]–[19]; see e.g. [20], [21] for a comprehensive overview. With recent exploratoty studies over contemporary real-world networks, new challenges have been raised in community identification, addressing the presence of overlapping communities [22]–[24], multimodal interaction of nodes over multiview networks [25], [26], exploitation of nodal and edge-related side-information [27], as well as dynamic interactions [28], [29].

In handling such new challenges, reliance on the adjacency matrix representation of networks limits their capabilities in capturing higher-order interactions, which can potentially provide critical information via temporal, multi-modal, or even multi-hop connectivity among nodes. To this end, *tensors* as multi-way data structures provide a viable alternative, whose increased representational capabilities can potentially lead to a more informative community identification [25], [26], [28], [30], [31]. For instance, [32] and [33] construct higher-order tensors whose entries are non-zero if a tuple of nodes jointly belongs to a cycle or a clique, while [28] captures temporal dynamics of communities via tensors. Under certain conditions, tensor decompositions are unique [34]–[36], and can guarantee identifiability of the community structure.

The present work develops a novel tensor-based network representation by recognizing that a network is the union of its *egonets*. An egonet is defined per node as the subgraph induced by the node itself, its one-hope neighbors, and all their connections, whose structure has been exploited in anomaly detection [37], and user-specific community identification [38], [39]. By concatenating egonet adjacency matrices along the 3-rd dimension of a three-way tensor, the proposed network representation, named *egonet-tensor*, captures information per node beyond its one-hop connectivity patterns. In fact, in a number of practical networks only adjacency matrix of the network is given, rendering egonets a unique candidate for enhancing community identification performance when extra nodal features are kept private, e.g., Amazon costumer graphs. By construction, egonet-tensor exhibits richer structure compared to its matrix counterpart, which is further exploited by casting the community detection task in a constrained tensor decomposition framework. Building on preliminary results in [31], solvers with convergence guarantees are developed for the proposed constrained non-convex optimization, whose solution yields the community-revealing components, utilized for soft as well as crisp community assignments unveiling possibly *overlapping* communities.

Manuscript received July 13, 2017; revised February 19, 2018, May 11, 2018, and September 10, 2018; accepted September 11, 2018. Date of publication September 24, 2018; date of current version October 1, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Pierre Borgnat. This work was supported by the National Science Foundation under Grants 1500713, 1442686, and 1711471. *(Corresponding author: Georgios B. Giannakis.)*

The authors are with the Department of Electrical and Computer Engineering and Digital Technology Center, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: sheik081@umn.edu; georgios@umn.edu).

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

Digital Object Identifier 10.1109/TSP.2018.2871383

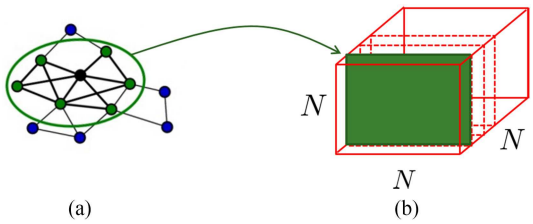


Fig. 1. Construction of *egonet-tensor* with frontal slabs as egonet adjacencies.

The rest of the paper is organized as follows. Section II introduces the novel tensor-based network representation, and Section III presents the constrained tensor decomposition, and its efficient solver for the task of community identification. Section IV provides analytical results corroborating the underlying intuition. Performance metrics for evaluating the quality of detected communities in networks with and without ground-truth communities is the subject of Section V, and Section VI introduces identification of time-varying communities using the proposed algorithm. Section VII provides numerical tests, while Section VIII concludes the paper.

Notation: Lower- (upper-) case boldface letters denote column vectors (matrices), and underlined upper-case boldface letters stand for tensor structures. Calligraphic symbols are reserved for sets, while ^{T} stands for transposition. Symbols $\circ$, $\otimes$ and $\odot$ are reserved for outer-product, Kronecker-product and Khatri-Rao-product, respectively, $\mathrm{Tr}\{\mathbf{X}\}$ denotes the trace of matrix $\mathbf{X}$, and x_{ij} denotes the (i,j)-th entry of matrix $\mathbf{X}$.

II. Egonet-Tensor Construction

Consider a graph $\mathcal{G}=(\mathcal{V},\mathcal{E},\mathbf{W})$, where $\mathcal{V}$, $\mathcal{E}$, and $\mathbf{W}\in\mathbb{R}^{N\times N}$ respectively denote the set of N nodes, i.e., $|\mathcal{V}|=N$, edges, and the adjacency matrix. In the case of binary networks, $w_{ij}:=1$ if $(i,j)\in\mathcal{E}$, and $w_{ij}:=0$ otherwise. Furthermore, the *egonet of node* n is defined as the subgraph induced by node n, its single-hop neighbors, and all their edges. Let $\mathcal{G}^{(n)}:=(\mathcal{V},\mathcal{E}^{(n)},\mathbf{W}^{(n)})\subset\mathcal{G}$ be the subgraph with $\mathcal{E}^{(n)}$ the egonet edgeset, and $\mathbf{W}^{(n)}\in\mathbb{R}^{N\times N}$ the corresponding adjacency matrix whose non-zero support captures the edges in $\mathcal{E}^{(n)}$; that is,

$$w_{ij}^{(n)}:=\begin{cases}w_{ij} & \text{if } (i,j)\in\mathcal{E}^{(n)}\\ 0 & \text{otherwise.}\end{cases}$$

where w_{ij} denotes the nonzero edge between nodes i and j, and $\mathcal{E}^{(n)}:=\{(u,v)|\forall(u,v)\in\mathcal{E},(n,v)\in\mathcal{E},(n,u)\in\mathcal{E}\}$. Fig. 1(a) illustrates such subgraphs, where the black node corresponds to the central node of the egonet, and the single-hop neighbors are colored green. Typically, the center node n is excluded from $\mathcal{G}^{(n)}$, but it is included here for convenience.

As Fig. 1(b) depicts, graph $\mathcal{G}$ can now be fully described by a three-way *egonet-tensor* $\underline{\mathbf{W}}\in\mathbb{R}^{N\times N\times N}$, where frontal slabs correspond to egonet adjacency matrices $\{\mathbf{W}^{(n)}\}_{n=1}^N$, stacked one after the other. In tensor parlance, that is tantamount to setting the n-th frontal slab as $\underline{\mathbf{W}}_{:,:,n}:=\mathbf{W}^{(n)}$, where : is a free index that spans its range.

The advantage of presenting a graph with its egonet-tensor lies in the fact that tensors as higher-order structures are capable of capturing "useful information" along their different dimensions, a.k.a., "modes." In a network with underlying community

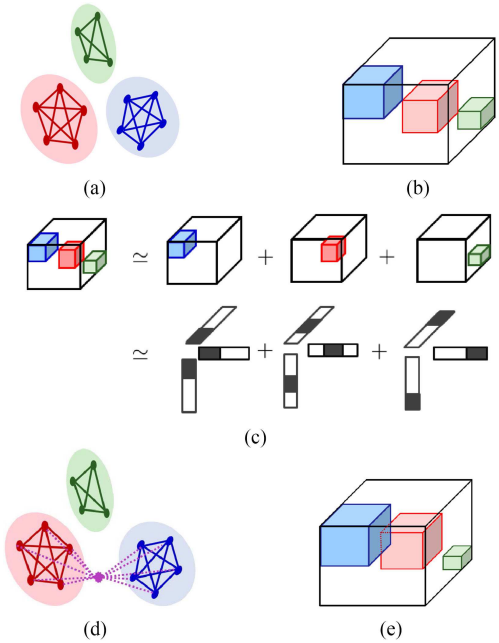


Fig. 2. (a) A toy network with three fully connected non-overlapping communities; (b) corresponding egonet-tensor; (c) its community-revealing factorization via CPD; (d) overlapping communities; and (e) corresponding egonet-tensor.

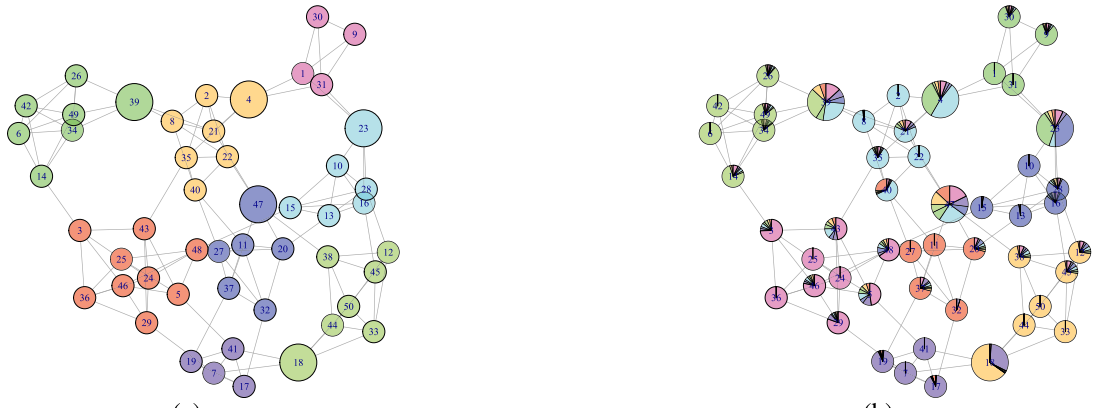
structure, the proposed egonet-tensor representation is indeed capable of preserving the inherent "similarities" among egonet adjacencies along the 3-rd mode, whose extraction is crucial in tasks such as community detection. Such representation is of particular interest for various settings where no nodal features are provided (that is, only the adjacency matrix is given), making egonets very appealing for an improved network representation as well as the development of robust schemes for the task of interest. The following toy example clarifies how such similarities induce a structure over the egonet tensor, and intuitively discusses how its exploitation can lead to an improved performance.

A. Toy Example

Let us consider a toy network with three fully-connected communities, illustrated in Fig. 2(a). Since each node is a member of a *fully-connected* community, the binary adjacency matrix of its egonet is identical to that of any other node in its resident community. Furthermore, after permutation, this egonet adjacency matrix consists of a single block of nonzero entries (with zero diagonal entries if the network is free of self-loops), and zeros elsewhere. This implies that the egonet-tensor can be permuted similarly into a block-diagonal tensor with three nonzero blocks, as illustrated in Fig. 2(b).

Adopting the well-known canonical polyadic decomposition (CPD) to decompose $\underline{\mathbf{W}}$ into its constituent rank-one tensors, the model naturally approximates the tensor with rank three, thus revealing the number of underlying communities. In fact, if the diagonal entries were all set to 1, i.e. considering self loops for all the neighboring nodes in an egonet, this approximation would be exact; see Fig. 2(c).

In practice, real-world networks often demonstrate overlapping community structure, where some nodes are associated



(a) (b)

Fig. 3. Visualization of an LFR network with $N = 50$, $\mu = 0.2$, and five "shared" nodes $\{4, 18, 23, 39, 47\}$, represented by a larger radius, with (a) hard community association via Infomap; and (b) soft association via EgoTen. Pie-charts depict association ratios.

with multiple communities rather than a single one. To address such cases, consider the augmented network in Fig. 2(d), where a new node (or a super node corresponding to more than one node) associated with two communities is added. Once the corresponding egonet tensor is constructed, the presence of overlapping communities manifests itself in overlapping diagonal blocks in the egonet tensor; see Fig. 2(e) in comparison with disjoint blocks in Fig. 2(b). As it will become evident in Section II, by exploiting a strutured CPD on the arising egonet-tensor, it can be shown that the frontal slab corresponding to the egonet adjacency of an overlapping node is approximated by multiple summands, each corresponding to one of its resident communities. The egonet-tensor representation naturally trades off flexibility for increased redundancy and memory costs. Nevertheless, resulting tensor is extremely sparse, and off-the-shelf tools for sparse tensor computations can be readily utilized; see e.g., [40]–[42].

Unfortunately, such idealistic assumptions where each community is fully-connected within and well-separated from other communities, are not fulfilled in real-world networks. Nevertheless, the inherent similarities among egonet adjacencies induce a potentially useful reinforced structure along the 3rd mode of the proposed egonet-tensor representation. For instance, nodes in a community often exhibit dense (rather than full) connections among themselves, and fewer connections with other communities. This property is consequently reflected in the egonet-tensor (as well as the traditional matrix adjacency) representation by dense diagonal blocks, whose clear separation fades aways as *out-of-community connections* increase. However, the presence of overlapping communities can further smear the block-structure as overlapping nodes are "well-connected" with multiple communities. The performance of traditional community detection methods often dramatically degrades in networks with such properties, whereas exploiting the *structured redundancy* offered via the proposed egonet-tensor representation and casting the problem in a community-revealing tensor decomposition framework increases robustness against the aforementioned phenomena; see Fig. 3.

III. Constrained Egonet Tensor CPD

Given $\underline{\mathbf{W}}$, this section leverages the canonical polyadic decomposition (CPD) [40] in order to factorize the egonet tensor into its constituent community-revealing factors. Assuming that the number of communities is upperbounded by K, a rank-K CPD model is sought by solving the following constrained least-squares (LS) problem

$$\{\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}\} = \arg\min_{\mathbf{A},\mathbf{B},\mathbf{C}} \left\| \underline{\mathbf{W}} - \sum_{k=1}^{K} \mathbf{a}_k \circ \mathbf{b}_k \circ \mathbf{c}_k \right\|_F^2 \quad \text{s.t.} \quad \mathbf{A} \geq \mathbf{0}, \mathbf{B} \geq \mathbf{0}, \mathbf{C} \geq \mathbf{0} \tag{1}$$

where $\mathbf{A} := [\mathbf{a}_1, \ldots, \mathbf{a}_K] \in \mathbb{R}^{N \times K}$, $\mathbf{B} := [\mathbf{b}_1, \ldots, \mathbf{b}_K] \in \mathbb{R}^{N \times K}$, and $\mathbf{C} := [\mathbf{c}_1, \ldots, \mathbf{c}_K] \in \mathbb{R}^{N \times K}$; while the term $(\mathbf{a}_k \circ \mathbf{b}_k \circ \mathbf{c}_k)$ is the outer product of the three vectors, which induces the k-th rank-one tensor component in the rank-K decomposition; see [40] for further details on CPD. The constraint $\mathbf{A} \geq \mathbf{0}$ denotes entry-wise nonnegativity constraints, i.e., $a_{nk} \geq 0$ for $n = 1, \ldots, N$ and $k = 1, \ldots, K$; and similarly for factors $\mathbf{B}$ and $\mathbf{C}$. These constraints enforce the nonnegativity of egonet adjacency matrices, thus inducing structure in the sought CPD and providing interpretation of the decomposition factors.

It is possible to re-write (1) as, see e.g., [40]

$$\{\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}\} = \arg\min_{\mathbf{A},\mathbf{B},\mathbf{C}} \sum_{n=1}^{N} \|\mathbf{W}^{(n)} - \mathbf{A}\text{diag}(\tilde{\mathbf{c}}_n)\mathbf{B}^\top\|_F^2 \quad \text{s.t.} \quad \mathbf{A} \geq \mathbf{0}, \mathbf{B} \geq \mathbf{0}, \mathbf{C} \geq \mathbf{0}$$

where $\text{diag}(\tilde{c}_n)$ is a diagonal matrix holding the n-th row of $\mathbf{C}$ on its diagonal. Focusing on the n-th frontal slab of the egonet-tensor, CPD provides the approximation

$$\mathbf{W}^{(n)} = \sum_{k=1}^{K} c_{nk} (\mathbf{a}_k \mathbf{b}_k^\top) \tag{2}$$

where c_{nk} denotes the (n, k)-th entry of factor $\mathbf{C}$. Such decomposition can be interpreted as a weighted sum over K "basis", $\{\mathbf{a}_k \mathbf{b}_k^\top\}_{k=1}^K$, where $(\mathbf{a}_k \mathbf{b}_k^\top)$ captures the "connectivity structure" within the k-th community. Consequently, c_{nk} can be viewed as *association level* of node n to community k for $k = 1, \ldots, K$. Furthermore, one can easily realize that since $(\mathbf{a}_k \mathbf{b}_k^\top)$ is viewed as connectivity structure of community k, the elements in $\mathbf{a}_k := [a_{1k}, \ldots, a_{nk}]^\top$ and $\mathbf{b}_k := [b_{1k}, \ldots, b_{nk}]^\top$ can be consequently viewed as the *contribution levels* of nodes $n = 1, \ldots, N$ to community k. This interpretation of the

factors prompts us to further leverage the structure and introduce additional constraints on the CPD factors.

A. Structured CPD

Real-world networks often involve nodes which are associated with more than one community, resulting in multiple nonzero entries in the association vector $[c_{n1}, c_{n2}, \ldots, c_{nK}]$ corresponding to a generic node n. To return a normalized association vector, we augment the optimization in (1) by the constraints $\sum_{k=1}^{K} c_{nk} = 1$ for $n = 1, 2, \ldots, N$, which together with $\mathbf{C} \geq \mathbf{0}$ acts as a simplex constraint on the rows of $\mathbf{C}$. Upon imposing simplex constraints (1) is further regularized as

$$\{\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}\} = \arg\min_{\mathbf{A},\mathbf{B},\mathbf{C}} \Bigg\{ \|\underline{\mathbf{W}} - \sum_{k=1}^{K} \mathbf{a}_k \circ \mathbf{b}_k \circ \mathbf{c}_k\|_F^2 + \lambda(\|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2) \Bigg\}$$
$$\text{s.t.} \quad \mathbf{A} \geq \mathbf{0}, \mathbf{B} \geq \mathbf{0}, \mathbf{C} \geq \mathbf{0}$$
$$\|\tilde{\mathbf{c}}_n\|_1 = 1 \qquad \forall n = 1, 2, \ldots, N \quad (3)$$

Different from [43] and [44], the regularization term $\|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2$ does not play the role of rank-regularization for subspace learning, instead it solves the scalar ambiguity in factors $\mathbf{A}$ and $\mathbf{B}$.

Note that the proposed algorithm requires an upperbound on the number of communities K to solve for the optimization in (3). In principle, one can readily regularize the objective function with a complexity term similar to that used by (Bayesian) Akaike's information-theoretic, see e.g. [45], or, minimum description length criteria [46]. At the expense of complexity, such augmented criteria can also return estimates of the number of communities.

The CPD problem formulated in (3) is a tri-linear constrained LS problem, whose minimization can be tackled by alternating optimization. In the ensuing subsection, the proposed solver is developed using by alternating optimization with ADMM intermediate steps; see e.g. [47] and [48].

B. Solving the Egonet Structured CPD

In the proposed alternating optimization scheme, each step consists of fixing two factors and minimizing the arising subproblem with respect to the third factor. In this subsection, we study the emerging sub-problems and propose efficient solvers for tackling those.

1) Factor $\mathbf{A}$ *Update:* Consider first the update of factor $\mathbf{A}$ at iteration k, obtained after fixing $\mathbf{B} = \mathbf{B}^{(k-1)}$ and $\mathbf{C} = \mathbf{C}^{(k-1)}$ and solving the corresponding minimization. The arising subproblem, after algebric manipulation can be re-written as

$$\mathbf{A}^{(k)} = \arg\min_{\mathbf{A} \geq \mathbf{0}} \|\mathbf{W}_1 - (\mathbf{B}^{(k-1)} \odot \mathbf{C}^{(k-1)})\mathbf{A}^\top\|_F^2 + \lambda\|\mathbf{A}\|_F^2 \quad (4)$$

where $\mathbf{W}_1 := [\text{vec}(\underline{\mathbf{W}}_{1,:,:}), \ldots, \text{vec}(\underline{\mathbf{W}}_{N,:,:})] \in \mathbf{R}^{N^2 \times N}$ is a matricized reshaping of the tensor $\underline{\mathbf{W}}$. Also $\mathbf{B}^{(k-1)} \odot \mathbf{C}^{(k-1)} := [\mathbf{b}_1^{(k-1)} \otimes \mathbf{c}_1^{(k-1)}, \ldots, \mathbf{b}_K^{(k-1)} \otimes \mathbf{c}_K^{(k-1)}]$ is the Khatri-Rao product of $\mathbf{B}^{(k-1)}$ and $\mathbf{C}^{(k-1)}$, where $\mathbf{b}_i^{(k-1)}$ ($\mathbf{c}_i^{(k-1)}$) denotes column i of $\mathbf{B}^{(k-1)}$ (resp. $\mathbf{C}^{(k-1)}$), and $\otimes$ denotes the Kronecker product operator; see also [40].

Following the steps in [48], auxiliary variable $\bar{\mathbf{A}}$ is introduced to account for the nonnegativity constraint, and the augmented Lagrangian of (4) is

$$\mathcal{L}_A^{(k)}(\mathbf{A}, \bar{\mathbf{A}}, Y) = \|\mathbf{W}_1 - \mathbf{H}_A^{(k)}\mathbf{A}^\top\|_F^2 + \lambda\text{Tr}\{\mathbf{A}\mathbf{A}^\top\} + r_+(\bar{\mathbf{A}}) + (\rho/2)\|\mathbf{Y} + \mathbf{A} - \bar{\mathbf{A}}\|_F^2 \quad (5)$$

where $\bar{\mathbf{A}}, \mathbf{Y} \in \mathbb{R}^{N \times K}$, $\mathbf{H}_A^{(k)} := \mathbf{C}^{(k-1)} \odot \mathbf{B}^{(k-1)}$, and $r_+(\bar{\mathbf{A}})$ is the regularizer corresponding to the nonnegativity constraint,

$$r_+(\bar{\mathbf{A}}) := \begin{cases} 0 & \text{if } \bar{\mathbf{A}} \geq \mathbf{0} \\ +\infty & \text{o.w.} \end{cases}$$

The ADMM solver then proceeds by iteratively updating blocks of variables $\mathbf{A}, \bar{\mathbf{A}}, \mathbf{Y}$ as

$$\begin{cases} \mathbf{A}^{(r)} = \arg\min_{\mathbf{A}} \mathcal{L}_A^{(k)}(\mathbf{A}, \bar{\mathbf{A}}^{(r-1)}, \mathbf{Y}^{(r-1)}) \\ \bar{\mathbf{A}}^{(r)} = \mathcal{P}_+(\mathbf{Y}^{(r-1)} + \mathbf{A}^{(r)}) \\ \mathbf{Y}^{(r)} = \mathbf{Y}^{(r-1)} - \rho(\mathbf{A}^{(r)} - \bar{\mathbf{A}}^{(r)}) \\ r = r + 1 \end{cases} \quad (6)$$

until a convergence criterion is met, namely whether the maximum number of iterations is exceeded, i.e., $r = I_{\text{max,ADMM}}$, or a prescribed ϵ-accuracy is met, i.e., $\|\mathbf{A}^{(r)} - \mathbf{A}^{(r-1)}\|_F / \|\mathbf{A}^{(r-1)}\|_F < \epsilon$. Operator $\mathcal{P}_+(.)$ in (6) denotes the element-wise projection of the input matrix onto the positive orthant, and its use enables the $\bar{\mathbf{A}}^{(r)}$ update to be carried at a very low cost. The Lagrange multiplier is set to $\rho = \|\mathbf{H}_A\|_F^2/K$- a value that is empirically shown to yield similar performance to that of the optimal value [48]. The final $\bar{\mathbf{A}}^{(r)}$ iterate in the ADMM solver will be used to update $\mathbf{A}^{(k)}$.

2) Factor $\mathbf{B}$ *update:* Update of factor $\mathbf{B}$ can be similarly carried out by solving the subproblem

$$\mathbf{B}^{(k)} = \arg\min_{\mathbf{B} \geq \mathbf{0}} \|\mathbf{W}_2 - \mathbf{H}_B^{(k)}\mathbf{B}^\top\|_F^2 + \lambda\|\mathbf{B}\|_F^2 \quad (7)$$

where $\mathbf{W}_2 := [\text{vec}(\underline{\mathbf{W}}_{:,1,:}), \ldots, \text{vec}(\underline{\mathbf{W}}_{:,N,:})]$, and $\mathbf{H}_B^{(k)} := \mathbf{C}^{(k-1)} \odot \mathbf{A}^{(k)}$, yielding a similar optimization problem as in (4). Algorithm 2 tabulates the explicit update rules for solving (4) and similarly (7) using a general framework.

3) Factor $\mathbf{C}$ *Update:* Update of factor $\mathbf{C}$ is obtained by fixing $\mathbf{A}$ and $\mathbf{B}$ at their most recent values, and solving the subproblem

$$\mathbf{C}^{(k)} = \arg\min_{\mathbf{C}} \|\mathbf{W}_3 - (\mathbf{A}^{(k)} \odot \mathbf{B}^{(k)})\mathbf{C}^\top\|_F^2$$
$$\text{s.t.,} \quad \mathbf{C} \geq \mathbf{0} \quad \|\tilde{\mathbf{c}}_n\|_1 = 1 \quad \forall n = 1, \ldots, N \quad (8)$$

where $\mathbf{W}_3 := [\text{vec}(\underline{\mathbf{W}}_{:,:,1}), \ldots, \text{vec}(\underline{\mathbf{W}}_{:,:,N})]$. Utilizing an ADMM approach, the augmented Lagrangian is formed as

$$\mathcal{L}_C^{(k)}(\mathbf{C}, \bar{\mathbf{C}}, Y) = \|\mathbf{W}_3 - \mathbf{H}_C^{(k)}\mathbf{C}^\top\|_F^2 + r_{\text{simp}}(\bar{\mathbf{C}}) + (\rho/2)\|\mathbf{Y} + \mathbf{C} - \bar{\mathbf{C}}\|_F^2 \quad (9)$$

where $\bar{\mathbf{C}}, \mathbf{Y} \in \mathbb{R}^{N \times K}$, $\mathbf{H}_C^{(k)} := (\mathbf{A}^{(k)} \odot \mathbf{B}^{(k)})$, and $r_{\text{simp}}(\bar{\mathbf{C}})$ is the regularizer corresponding to the simplex constraint on the rows of matrix $\bar{\mathbf{C}}$ as

$$r_{\text{simp}}(\bar{\mathbf{C}}) := \begin{cases} 0 & \text{if } \bar{\mathbf{C}} \geq \mathbf{0}, \sum_{k=1}^{K} \bar{c}_{n,k} = 1 \ \forall n \\ +\infty & \text{o.w.} \end{cases}$$

Algorithm 1: Constrained Tensor Decomposition via Alternating Least-Squares (ALS).

Input $\underline{\mathbf{W}}, K, I_{\max}, \lambda$
Initialize $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{N\times K}$ at random and set $k=0$
Form Matrix reshapes $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ of the tensor as
$\mathbf{W}_1 := [\text{vec}(\underline{\mathbf{W}}_{1,:,:}), \ldots, \text{vec}(\underline{\mathbf{W}}_{N,:,:})]$
$\mathbf{W}_2 := [\text{vec}(\underline{\mathbf{W}}_{:,1,:}), \ldots, \text{vec}(\underline{\mathbf{W}}_{:,N,:})]$
$\mathbf{W}_3 := [\text{vec}(\underline{\mathbf{W}}_{:,:,1}), \ldots, \text{vec}(\underline{\mathbf{W}}_{:,:,N})]$
while $k < I_{\max}$ **do** or not-convergenced
  $\mathbf{H}_A^{(k)} = \mathbf{C}^{(k-1)} \odot \mathbf{B}^{(k-1)}$
  $\mathbf{A}^{(k)} \leftarrow$ Algorithm 2 with input $\{\mathbf{H}_A^{(k)}, \mathbf{W}_1, \mathbf{A}^{(k-1)}\}$
  $\mathbf{H}_B^{(k)} = \mathbf{C}^{(k-1)} \odot \mathbf{A}^{(k)}$
  $\mathbf{B}^{(k)} \leftarrow$ Algorithm 2 with input $\{\mathbf{H}_B^{(k)}, \mathbf{W}_2, \mathbf{B}^{(k-1)}\}$
  $\mathbf{H}_C^{(k)} = \mathbf{B}^{(k)} \odot \mathbf{A}^{(k)}$
  $\mathbf{C}^{(k)} \leftarrow$ Algorithm 3 with input $\{\mathbf{H}_C^{(k)}, \mathbf{W}_3, \mathbf{C}^{(k-1)}\}$
  $k \leftarrow k+1$
end while
Retrun $\mathbf{A}^{(k)}, \mathbf{B}^{(k)}, \mathbf{C}^{(k)}$

Algorithm 2: ADMM Solver for 1st and 2nd Mode Subproblems.

Input $\mathbf{H}, \mathbf{W}, \mathbf{Z}_{\text{init}}$
Goal is to solve

$$\mathbf{Z}^* = \arg\min_{\mathbf{Z}\geq\mathbf{0}} \text{Tr}\Big\{\mathbf{Z}(\mathbf{H}^\top\mathbf{H} + \lambda\mathbf{I}_{K\times K})\mathbf{Z}^\top - 2\mathbf{W}^\top\mathbf{H}\mathbf{Z}^\top\Big\}$$

Set $\rho = \frac{\|\mathbf{Z}_{\text{init}}\|_F^2}{K}, \mathbf{Z}^{(0)} = \mathbf{Z}_{\text{init}}, \bar{\mathbf{Z}}^{(0)} = \mathbf{0}_{N\times K}, \mathbf{Y}^{(0)} = \mathbf{0}_{N\times K}, r = 0$
while $r < I_{\text{max,ADMM}}$ **do**

$$\begin{aligned}
\mathbf{Z}^{(r)} &= (\mathbf{H}^\top\mathbf{H} + (\lambda+\rho/2)\mathbf{I}_{K\times K})^{-1} \\
&\quad \times \Big(\mathbf{W}^\top\mathbf{H} + \frac{\rho}{2}(\mathbf{Y}^{(r-1)} - \bar{\mathbf{Z}}^{(r-1)})\Big) \\
\bar{\mathbf{Z}}^{(r)} &= \mathcal{P}_+\Big(\mathbf{Z}^{(r)} + \mathbf{Y}^{(r-1)}\Big) \\
\mathbf{Y}^{(r)} &= \mathbf{Y}^{(r-1)} - \rho(\mathbf{Z}^{(r)} - \bar{\mathbf{Z}}^{(r)}) \\
r &= r+1
\end{aligned}$$

end while
Retrun $\bar{\mathbf{Z}}^{(r)}$

ADMM solver then proceeds with iterative updates as

$$\begin{cases}
\mathbf{C}^{(r)} = \arg\min_{\mathbf{C}} \mathcal{L}_C^{(k)}(\mathbf{C}, \bar{\mathbf{C}}^{(r-1)}, \mathbf{Y}^{(r-1)}) \\
\bar{\mathbf{C}}^{(r)} = \mathcal{P}_{\text{simp}}(\mathbf{Y}^{(r-1)} + \mathbf{C}^{(r)}) \\
\mathbf{Y}^{(r)} = \mathbf{Y}^{(r-1)} - \rho(\mathbf{C}^{(r)} - \bar{\mathbf{C}}^{(r)}) \\
r = r+1
\end{cases} \tag{10}$$

where $\mathcal{P}_{\text{simp}}(.)$ denotes the projection of rows of the input matrix onto the simplex set. This projection has been widely studied and can be efficiently accommodated by the algorithm discussed in [49]. Explicit update steps of the ADMM solver for (8) are tabulated under Algorithm 3.

Proposition 1: If the sequence generated by Algorithm 4 is bounded, then the sequence $\{\mathbf{A}^{(k)}, \mathbf{B}^{(k)}, \mathbf{C}^{(k)}\}$ converges to a stationary point of (3).

Algorithm 3: ADMM Solver for 3rd Mode Subproblem.

Input $\mathbf{H}, \mathbf{W}, \mathbf{Z}_{\text{init}}$
Goal is to solve

$$\mathbf{Z}^* = \arg\min_{\mathbf{Z}\geq\mathbf{0}\|\tilde{\mathbf{z}}_n\|_1=1\ \forall n=1,\ldots,N} \text{Tr}\Big\{\mathbf{Z}\mathbf{H}^\top\mathbf{H}\mathbf{Z}^\top - 2\mathbf{W}^\top\mathbf{H}\mathbf{Z}^\top\Big\}$$

Set $\rho = \frac{\|\mathbf{Z}_{\text{init}}\|_F^2}{K}, \mathbf{Z}^{(0)} = \mathbf{Z}_{\text{init}}, \bar{\mathbf{Z}}^{(0)} = \mathbf{0}_{N\times K}, \mathbf{Y}^{(0)} = \mathbf{0}_{N\times K}, r = 0$
while $r < I_{\text{max,ADMM}}$ **do**

$$\begin{aligned}
\mathbf{Z}^{(r)} &= (\mathbf{H}^\top\mathbf{H} + \rho/2\,\mathbf{I}_{N\times N})^{-1} \\
&\quad \times \Big(\mathbf{W}^\top\mathbf{H} + \frac{\rho}{2}(\mathbf{Y}^{(r-1)} - \bar{\mathbf{Z}}^{(r-1)})\Big) \\
\bar{\mathbf{Z}}^{(r)} &= \mathcal{P}_{\text{simp}}\Big(\mathbf{Z}^{(r)} + \mathbf{Y}^{(r-1)}\Big) \\
\mathbf{Y}^{(r)} &= \mathbf{Y}^{(r-1)} - \rho(\mathbf{Z}^{(r)} - \bar{\mathbf{Z}}^{(r)}) \\
r &= r+1
\end{aligned}$$

end while
Retrun $\bar{\mathbf{Z}}^{(r)}$

Proof: The convergence follows from [48, Theorem 1]; also c.f. [50], [51]. ■

Proposition 1 provides convergence guarantees on the sequence $\{\mathbf{A}^{(k)}, \mathbf{B}^{(k)}, \mathbf{C}^{(k)}\}$, where the necessary assumption is on boundedness of the generated iterates. This assumption is easily guaranteed here due to Frobenius regularization on factors $\mathbf{A}$ and $\mathbf{B}$, as well as the simplex constraint on factor $\mathbf{C}$; thus, the results readily carry over.

IV. Low-Rank Egonet Tensor Analysis

In this section, the intuitive low-rank property of the egonet adjacency matrix and tensor are justified analytically. To this end, and similar to the analysis of spectral clustering methods in community detection, consider modeling a network of N nodes and two communities by a stochastic blockmodel (SBM) [52], [53]. That is, consider an undirected network with two communities of equal size $N/2$, assuming N is even for simplicity, where the intra-community edge probability is p and the inter-community edge probability is q, denoted as SBM$(N; 2; p; q)$, with $p > q$. Define also the *expected egonet-adjacency of community* $\mathcal{C}_k$ as

$$\overline{\mathbf{W}}_k := \mathbb{E}\left[\frac{1}{|\mathcal{C}_k|}\sum_{v_i\in\mathcal{C}_k}\mathbf{W}^{(i)}\right], k = 1, 2$$

$\mathbf{W}^{(i)}$, as in Section II, denotes the egonet-adjacency matrix of node v_i, and the expectation is taken with respect to the randomness of the graph.

Proposition 2: In an SBM$(N; 2; p; q)$ network with communities of equal size and $q < p$, the expected egonet-adjacency matrix $\overline{\mathbf{W}}_1$ of community $\mathcal{C}_1$ can be approximated by a rank-2 matrix $\widehat{\overline{\mathbf{W}}}_1$, where the approximation error is bounded as

$$\|\overline{\mathbf{W}}_1 - \widehat{\overline{\mathbf{W}}}_1\|_2 \leq p^3 + \frac{4}{N}p + \sqrt{2}q.$$

Furthermore, for networks with large N and $p > \sqrt{2}q$, selecting the $N/2$ largest-amplitude entries of the eigenvector

corresponding to the largest eigenvalue unravels the members of community $\mathcal{C}_1$.

Proof: See Appendix A for detailed proof. ■

Building on this and upon proper permutation, similar result can be provided for the expected egonet-adjacency matrix $\overline{\mathbf{W}}_2$ of community $\mathcal{C}_2$. That is, $\overline{\mathbf{W}}_2$ in an SBM$(N;2;p;q)$ network with communities of equal size can be approximated by a rank-2 matrix $\widehat{\overline{\mathbf{W}}}_2$, and the approximation error is bounded as

$$\|\overline{\mathbf{W}}_2 - \widehat{\overline{\mathbf{W}}}_2\|_2 \leq p^3 + \frac{4}{N}p + \sqrt{2}q.$$

As the exact number of nodes in a community is unknown in practice, hard community detection is performed by thresholding the association coefficients. This is common in non-negative matrix and tensor factorization approaches to community detection [9], [14], [18], [24], [27], and has been empirically demonstrated to yield high-quality results. For this reason, it has been adopted here too.

Definition: In an SBM$(N;2;p;q)$ network, the *expected egonet-adjacency tensor* $\underline{\overline{\mathbf{W}}}$ of size $N \times N \times 2$ is defined as $\underline{\overline{\mathbf{W}}}_{(:,:,1)} := \overline{\mathbf{W}}_1$ and $\underline{\overline{\mathbf{W}}}_{(:,:,2)} := \overline{\mathbf{W}}_2$.

Proposition 3: Tensor $\underline{\overline{\mathbf{W}}}$ in an SBM$(N;2;p;q)$ network with equally-sized communities can be approximated by a rank-3 tensor, where the approximation error is upper-bounded as

$$\min_{\underline{\widehat{\mathbf{W}}},\text{rank}(\underline{\widehat{\mathbf{W}}})=3} \frac{1}{N}\|\underline{\overline{\mathbf{W}}} - \underline{\widehat{\mathbf{W}}}\|_F \leq \sqrt{\frac{(p^3+4p/N)^2+p^2q^4}{N}} + \frac{q(1-pq)}{N} \tag{11}$$

which diminishes with the network size as $\mathcal{O}(1/\sqrt{N})$. Furthermore, such approximation can be achieved by

$$\underline{\widehat{\mathbf{W}}}_{N\times N\times 2} = \beta \mathbf{v}_0 \circ \mathbf{v}_0 \circ \begin{bmatrix}1\\1\end{bmatrix} + (\alpha-\beta)\mathbf{v}_1 \circ \mathbf{v}_1 \circ \begin{bmatrix}1\\0\end{bmatrix} + (\alpha-\beta)\mathbf{v}_2 \circ \mathbf{v}_2 \circ \begin{bmatrix}0\\1\end{bmatrix} \tag{12}$$

where $\mathbf{v}_0 = \mathbf{1}_{N\times 1}$, $\mathbf{v}_1 = [\mathbf{1}_{1\times N/2}, \mathbf{0}_{1\times N/2}]^\top$, and $\mathbf{v}_2 = [\mathbf{0}_{1\times N/2}, \mathbf{1}_{1\times N/2}]^\top$ and $\circ$ denotes the outer product.

Proof: See Appendix B for proof. ■

Remark: The first term in decomposition (12) captures the inter-community edges, while the other two terms correspond to intra-community connections, and $\mathbf{v}_1$ and $\mathbf{v}_2$ reflect community structures. The following Corollary studies the asymptotic case where the probability of inter-community edges q approaches 0.

Corollary 2: The expected adjacency tensor of an SBM $(N;2;p;q)$ network approaches a rank-2 tensor as $q \to 0$; that is

$$\lim_{q\to 0}\frac{1}{N}\left\|\underline{\widehat{\mathbf{W}}} - p^3\mathbf{v}_1\circ\mathbf{v}_1\circ\begin{bmatrix}1\\0\end{bmatrix} + p^3\mathbf{v}_2\circ\mathbf{v}_2\circ\begin{bmatrix}0\\1\end{bmatrix}\right\|_F = 0$$

and the approximation error for sufficiency large N can be approximated by

$$\min_{\underline{\widehat{\mathbf{W}}},\text{rank}(\underline{\widehat{\mathbf{W}}})=2}\frac{1}{N}\|\underline{\overline{\mathbf{W}}} - \underline{\widehat{\mathbf{W}}}\|_F \simeq \sqrt{\frac{p^3}{N}}.$$

Proof: The proof readily follows from the fact that the approximation term corresponsing to $\mathbf{v}_0$ in (12) has Frobenius-norm $\sqrt{2}\beta N$ and its normalized (by N) Frobenius-norm goes to zero as $q \to 0$. Furthermore, the right-hand-side of the inequality in Prop. 3 for sufficiently large N and $q \to 0$ is dominated by the first term, and thus it can be approximated by $\sqrt{p^3/N}$. ■

So far, in Propositions 2 and 3 we have justified the intuition behind the rank-revealing decomposition of the expected egonet-adjacency tensor $\underline{\overline{\mathbf{W}}}$. In practice however, we never have access to $\overline{\mathbf{W}}_1$ and $\overline{\mathbf{W}}_2$. Instead, the egonet adjacency tensor in Section II provides $N/2$ samples of the egonet adjacency for each of the two communities. Thus, one is motivated to approximate the expected egonet-tensor $\underline{\overline{\mathbf{W}}}$ with its sample mean.

Specifically, were the community membership of the vertices provided, one could approximate $\overline{\mathbf{W}}_1$ and $\overline{\mathbf{W}}_2$ as

$$\overline{\mathbf{W}}_1 \simeq \widetilde{\mathbf{W}}_1 := \frac{2}{N}\sum_{v_i\in\mathcal{C}_1}\mathbf{W}^{(i)},\ \overline{\mathbf{W}}_2 \simeq \widetilde{\mathbf{W}}_2 := \frac{2}{N}\sum_{v_i\in\mathcal{C}_2}\mathbf{W}^{(i)}$$

and define the $N \times N \times 2$ *sample mean egonet-adjacency tensor* $\underline{\widetilde{\mathbf{W}}}$ with two slabs $\underline{\widetilde{\mathbf{W}}}_{(:,:,1)} = \widetilde{\mathbf{W}}_1$ and $\underline{\widetilde{\mathbf{W}}}_{(:,:,2)} = \widetilde{\mathbf{W}}_2$.

Without community memberships available, we only have access to the egonet adjacency $\underline{\mathbf{W}}$. The following proposition proves that the CPD of $\underline{\mathbf{W}}$ (without the oracle knowledge) provides an upperbound on the CPD of $\underline{\widetilde{\mathbf{W}}}$.

Proposition 4: The rank-r CPD objective of $\underline{\widetilde{\mathbf{W}}}_{N\times N\times 2}$ is upper bounded by the rank-r CPD objective of $\underline{\mathbf{W}}_{N\times N\times N}$; that is,

$$\arg\min_{\{\mathbf{A},\mathbf{B},\check{\mathbf{C}}\}\geq\mathbf{0}}\frac{1}{2}\sum_{k=1}^{2}\|\widetilde{\mathbf{W}}_k - \mathbf{A}\text{diag}(\check{\mathbf{c}}_k)\mathbf{B}^\top\|_F^2 \leq \arg\min_{\{\mathbf{A},\mathbf{B},\mathbf{C}\}\geq\mathbf{0}}\frac{1}{N}\sum_{n=1}^{N}\|\mathbf{W}^{(n)} - \mathbf{A}\text{diag}(\tilde{\mathbf{c}}_n)\mathbf{B}^\top\|_F^2$$

where $\mathbf{A}, \mathbf{B}, \mathbf{C} := [\tilde{\mathbf{c}}_1, \ldots, \tilde{\mathbf{c}}_N]^\top$ are of size $N \times r$, and $\check{\mathbf{C}} := [\check{\mathbf{c}}_1, \check{\mathbf{c}}_2]^\top$ is of size $2 \times r$.

Proof: For detailed proof, see Appendix C. ■

According to Proposition 4, rank-r CPD of $\underline{\mathbf{W}}$ provides an upper-bound on the rank-r CPD of the oracle-given $\underline{\widetilde{\mathbf{W}}}$. The analysis can be generalized to networks with K communities, and further regularization and simplex constraints can help better utilize the intrinsic properties of the community detection task as well as resolve scalar ambiguities among CPD factors. The merits of our novel tensor decomposition with respect to the conventional adjacency matrix factorization will be corroborated by test results on synthetic as well as real-world networks. To this end, common metrics for quality evaluation of the detected communities are discussed next.

V. Community Assignment and Evaluation

Once the proposed solver returns the solution of (3), the rows of factor $\hat{\mathbf{C}}$ provide a "soft" or "fuzzy" community membership for the nodes in the network. In the special case of networks with non-overlapping communities, using the n-th row of matrix $\hat{\mathbf{C}}$, node n will be assigned to the community k^* where $k^* = \arg\max_{k=1,\ldots,K}\hat{c}_{nk}$.

In order to provide a "crisp" community association in networks *with* overlapping communities, where a node can be associated with more than one community, the entries $\hat{c}_{nk}$ are compared with a threshold τ and node n is associated with community k if $\hat{c}_{nk} > \tau$. Thus, crisp community membership

matrix $\mathbf{\Gamma} \in \mathbb{R}^{N\times K}$ is obtained as

$$[\mathbf{\Gamma}]_{nk} := \begin{cases} 1 & \text{if } \hat{c}_{nk} > \tau \\ 0 & \text{o.w.} \end{cases} \qquad \forall n,k \tag{13}$$

There are a number of metrics available for evaluating the quality of a detected *cover*, that is, a set of communities, in networks with underlying community structure. Depending on whether the ground-truth communities are available or not, two categories of metrics are considered.

A. Networks With Ground-Truth Communities

Normalized mutual information and F1-score are the most commonly-used metrics for performance evaluation over networks with ground truth communities. Let cover $\hat{\mathcal{S}} := \{\hat{\mathcal{C}}_1, \ldots, \hat{\mathcal{C}}_{|\hat{\mathcal{S}}|}\}$ denote the set of detected communities, where $\mathcal{C}_i$ is the set of nodes associated with community i for $i = 1, 2, \ldots, |\hat{\mathcal{S}}|$, and let the ground truth communities be denoted by $\mathcal{S}^* := \{\mathcal{C}_1^*, \ldots, \mathcal{C}_{|\mathcal{S}^*|}^*\}$.

Normalized mutual information (NMI) [20]: NMI is an information-theoretic metric defined as (cf. [20])

$$\text{NMI}(\mathcal{S}^*, \hat{\mathcal{S}}) := \frac{2\text{I}(\mathcal{S}^*, \hat{\mathcal{S}})}{\text{H}(\mathcal{S}^*) + \text{H}(\hat{\mathcal{S}})}$$

where $\text{H}(\hat{\mathcal{S}})$ denotes the entropy of set $\hat{\mathcal{S}}$ defined as

$$\text{H}(\hat{\mathcal{S}}) := -\sum_{i=1}^{|\hat{\mathcal{S}}|} p(\hat{\mathcal{C}}_i) \log p(\hat{\mathcal{C}}_i) = -\sum_{i=1}^{|\hat{\mathcal{S}}|} \frac{|\hat{\mathcal{C}}_i|}{N} \log \frac{|\hat{\mathcal{C}}_i|}{N}$$

and similarly for $\text{H}(\mathcal{S}^*)$. Furthermore, $\text{I}(\mathcal{S}^*, \hat{\mathcal{S}})$ denotes the mutual information between the detected and ground-truth communities, and is defined as

$$\text{I}(\mathcal{S}^*, \hat{\mathcal{S}}) := \sum_{i=1}^{|\mathcal{S}^*|} \sum_{j=1}^{|\hat{\mathcal{S}}|} p(\mathcal{C}_i^* \cap \hat{\mathcal{C}}_j^*) \log \frac{p(\mathcal{C}_i^* \cap \hat{\mathcal{C}}_j)}{p(\mathcal{C}_i^*)p(\hat{\mathcal{C}}_j)} \tag{14}$$

$$= \sum_{i=1}^{|\mathcal{S}^*|} \sum_{j=1}^{|\hat{\mathcal{S}}|} \frac{|\mathcal{C}_i^* \cap \hat{\mathcal{C}}_j|}{N} \log \frac{N|\mathcal{C}_i^* \cap \hat{\mathcal{C}}_j|}{|\mathcal{C}_i^*||\hat{\mathcal{C}}_j|} \tag{15}$$

Intuitively, mutual information $\text{I}(\mathcal{S}^*, \hat{\mathcal{S}})$ reflects a measure of similarity between the two community sets, while entropy $\text{H}(\hat{\mathcal{S}})$ ($\text{H}(\mathcal{S}^*)$) denotes the level of uncertainty in community affiliation of a random node in cover $\hat{\mathcal{S}}$ (resp. $\mathcal{S}^*$). Thus, high values of NMI, namely its maximum at 1, reflect *predictability* of $\hat{\mathcal{S}}$ based on $\mathcal{S}^*$ which readily translates into correct community identification in the detected cover $\hat{\mathcal{S}}$, whereas low values of NMI, namely its minimum at 0, reflects poor discovery of the true underlying communities. This measure has been generalized for overlapping communities in [54], and will be utilized for performance assessment in such networks.

Average F1-score [9]: F1-score is a measure of binary classification accuracy, specifically, the harmonic mean of *precision* and *recall*, taking its highest value at 1 and lowest value at 0. To obtain the average F1-score for $\hat{\mathcal{S}}$, one needs to find which detected community $\hat{\mathcal{C}}_i \in \hat{\mathcal{S}}$ corresponds to a given true community $\mathcal{C}_j^* \in \mathcal{S}^*$, i.e., maximizes the corresponding F1-score. The average F1-score is then given by

$$\bar{F1} := \frac{1}{|\mathcal{S}^*|} \sum_{i=1}^{|\mathcal{S}^*|} F1(\mathcal{C}_i^*, \hat{\mathcal{C}}_{I(i)})$$

where

$$I(i) = \arg\max_j F1(\mathcal{C}_i^*, \hat{\mathcal{C}}_j)$$

in which $F1(\mathcal{C}_i^*, \hat{\mathcal{C}}_j) := \frac{2\,|\mathcal{C}_i^* \cap \hat{\mathcal{C}}_j|}{|\mathcal{C}_i^*| + |\hat{\mathcal{C}}_j|}$.

B. Networks without ground-truth communities

A general metric for evaluating the "quality" of detected communities, regardless of whether the ground-truth memberships are available or not, is to measure *conductance* [55].

Conductance: Conductance of a detected community $\hat{\mathcal{C}}_k$ in graph $\mathcal{G}$ is defined as

$$\phi(\hat{\mathcal{C}}_k) := \frac{\sum_{i\in\hat{\mathcal{C}}_k, j\notin\hat{\mathcal{C}}_k} \mathbf{W}_{ij}}{\min\{\text{vol}(\hat{\mathcal{C}}_k), \text{vol}(\mathcal{V}\setminus\hat{\mathcal{C}}_k)\}}$$

where

$$\text{vol}(\hat{\mathcal{C}}_k) := \sum_{i\in\hat{\mathcal{C}}_k, \forall j} \mathbf{W}_{ij}$$

and $(\mathcal{V}\setminus\hat{\mathcal{C}}_k)$ is the complement of $\hat{\mathcal{C}}_k$. According to this measure, *high-quality* communities yield small conductance scores as they exhibit dense connections among the nodes within the community and sparse connections with the rest.

Furthermore, the weighted-average $\bar{\phi}(\hat{\mathcal{S}})$ is defined as the average conductance of the detected communities weighted by their (normalized) community size, that is,

$$\bar{\phi}(\hat{\mathcal{S}}) := \sum_{k=1}^{|\hat{\mathcal{S}}|} \frac{|\hat{\mathcal{C}}_k|}{N} \phi(\hat{\mathcal{C}}_k). \tag{16}$$

VI. Community Detection on Time-Varying Graphs

In this section, we extend the introduced overlapping community identification approach over networks for which the connectivity evolves over time [56]. For instance, consider the emergence of a new sports club giving rise to a new community of individuals whose newly-formed interactions in the club will be reflected in their connections over the social media, or, the network of brain regions where the activation/deactivation of different regions during a certain task can be captured by a time-varying graph. The goal is to utilize the proposed EgoTen approach for identification of dynamic communities, as well as the corresponding time-varying association of nodes.

To this end, consider graph $\mathcal{G}_t := (\mathcal{V}, \mathcal{E}_t, \mathbf{W}_t)$, where the subscript t denotes time index $t = 1, \ldots, T$. The set of nodes $\mathcal{V}$ is assumed fixed across time, while the edgeset $\mathcal{E}_t$ as well as the corresponding adjacency matrix $\mathbf{W}_t$ are allowed to vary. The introduced EgoTen community identification algorithm can be readily applied to time-varying graphs as follows.

For any slot t, the procedure of egonet tensor construction can be carried out, giving rise to a 3-dimensional egonet-tensor denoted by $\underline{\mathbf{W}}_t$. Subsequently, the overall 4-dimensional egonet-tensor is constructed by stacking $\underline{\mathbf{W}}_t$ $\forall t$ along the fourth

dimension of $\underline{\mathbf{W}} \in \mathbb{R}^{N\times N\times N\times T}$; that is, according to the tensor parlance we have $\underline{\mathbf{W}}_{:,:,:,t} = \underline{\mathbf{W}}_t$ for $t = 1,\ldots,T$. Having formed the overall egonet-tensor $\underline{\mathbf{W}}$, dynamic community detection is now cast as

$$\{\widehat{\mathbf{A}},\widehat{\mathbf{B}},\widehat{\mathbf{C}},\widehat{\mathbf{D}}\} = \underset{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}}{\arg\min}\left\{\|\underline{\mathbf{W}} - \sum_{k=1}^{K}\mathbf{a}_k\circ\mathbf{b}_k\circ\mathbf{c}_k\circ\mathbf{d}_k\|_F^2 + \lambda(\|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2)\right\}$$

$$\text{s.t.} \quad \mathbf{A}\geq\mathbf{0},\mathbf{B}\geq\mathbf{0},\mathbf{C}\geq\mathbf{0},\mathbf{D}\geq\mathbf{0}$$
$$\|\tilde{\mathbf{c}}_n\|_1 = 1 \qquad \forall n = 1,2,\ldots,N$$
$$\|\tilde{\mathbf{d}}_n\|_1 = 1 \qquad \forall n = 1,2,\ldots,N \tag{17}$$

where $\mathbf{A},\mathbf{B},\mathbf{C}\in\mathbb{R}^{N\times K}$, and $\mathbf{D} := [\tilde{\mathbf{d}}_1^\top,\ldots,\tilde{\mathbf{d}}_T^\top]^\top \in \mathbb{R}^{T\times K}$. The LS cost in (17) is the generalization of that for the 3-dimensional tensor decomposition to a higher dimension, while nonnegativity and simplex constrains are similarly carried over for $\mathbf{D}$. To clarify the simplex constraints on the rows of $\mathbf{D}$, consider the decomposition

$$\underline{\mathbf{W}}_t = \sum_{k=1}^{K} d_{tk}\Big(\mathbf{a}_k\circ\mathbf{b}_k\circ\mathbf{c}_k\Big). \tag{18}$$

For a fixed slot t, the rank-one tensors of $(\mathbf{a}_k\circ\mathbf{b}_k\circ\mathbf{c}_k)$ can be perceived as building blocks of the 3-dimensional egonet-tensors, and the "degree of presence" of community k at time t is captured by d_{kt}, while the constraint $\|\tilde{\mathbf{d}}_t\|_1 = 1\ \forall t$ resolves the scalar ambiguity by normalization. Indeed, stationary graphs with $T = 1$ can be subsumed by this model, for which the additional constraints on $\mathbf{D}$ reduce to the trivial solution for the fourth factor as $\mathbf{D} = \mathbf{1}_{1\times K}$. Focusing on the n-th slab of $\underline{\mathbf{W}}_t$, the egonet adjacency of node n at time t is decomposed as

$$\mathbf{W}_t^{(n)} = \sum_{k=1}^{K} d_{tk}c_{nk}\Big(\mathbf{a}_k\circ\mathbf{b}_k\Big) \tag{19}$$

where the product $d_{tk}c_{nk}$ provides the association of node n to community k at time t.

Similar to (3), the decomposition in (17) is solved by alternating least-squares for updating the factors. The overall solver is provided in Algorithm 4, within which we have utilized Algorithm 2 and Algorithm 3 for handling the emerging subproblems.

Note the proposed decomposition is in fact unraveling the evolution of detected communities across time via columns of factor D. That is, rather than explicit discovery of temporal community association of node n, the k-th column of factor D is utilized to modulate the association of node n to community k across time, with $d_{tk} * c_{nk}$ for time t. Thus, association of different nodes will be modulated in the same way at time t, while the final product is influenced by both d_{tk} and c_{nk}. Furthermore, one could jointly utilize factors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ for discovering node-community association, where similar constraints such as simplex should be imposed on (at least one of) the other two factors. For faster convergence however, we have not imposed such additional structure, and have solely utilized factor $\mathbf{C}$ for this purpose. Thus, joint utilization of the three factors remains an open direction.

Algorithm 4: Constrained ALS for Time-Varying Graphs.

Input $\underline{\mathbf{W}}, K, I_{\max}, \lambda$
Initialize $\mathbf{A},\mathbf{B},\mathbf{C}\in\mathbb{R}^{N\times K}$ and $\mathbf{D}\in\mathbb{R}^{T\times K}$ and set $k = 0$
Form matrix reshapes $\mathbf{W}_1,\mathbf{W}_2,\mathbf{W}_3,\mathbf{W}_4$ of the tensor as
$\mathbf{W}_1 := [\text{vec}(\underline{\mathbf{W}}_{1,:,:,:}),\ldots,\text{vec}(\underline{\mathbf{W}}_{N,:,:,:})]$
$\mathbf{W}_2 := [\text{vec}(\underline{\mathbf{W}}_{:,1,:,:}),\ldots,\text{vec}(\underline{\mathbf{W}}_{:,N,:,:})]$
$\mathbf{W}_3 := [\text{vec}(\underline{\mathbf{W}}_{:,:,1,:}),\ldots,\text{vec}(\underline{\mathbf{W}}_{:,:,N,:})]$
$\mathbf{W}_4 := [\text{vec}(\underline{\mathbf{W}}_{:,:,:,1}),\ldots,\text{vec}(\underline{\mathbf{W}}_{:,:,:,T})]$
while $k < I_{\max}$ **do** or not-convergenced
  $\mathbf{H}_A^{(k)} = \mathbf{D}^{(k-1)}\odot\mathbf{C}^{(k-1)}\odot\mathbf{B}^{(k-1)}$
  $\mathbf{A}^{(k)} \leftarrow$ Algorithm 2 with input $\{\mathbf{H}_A^{(k)},\mathbf{W}_1,\mathbf{A}^{(k-1)}\}$
  $\mathbf{H}_B^{(k)} = \mathbf{D}^{(k-1)}\odot\mathbf{C}^{(k-1)}\odot\mathbf{A}^{(k)}$
  $\mathbf{B}^{(k)} \leftarrow$ Algorithm 2 with input $\{\mathbf{H}_B^{(k)},\mathbf{W}_2,\mathbf{B}^{(k-1)}\}$
  $\mathbf{H}_C^{(k)} = \mathbf{D}^{(k-1)}\odot\mathbf{B}^{(k)}\odot\mathbf{A}^{(k)}$
  $\mathbf{C}^{(k)} \leftarrow$ Algorithm 3 with input $\{\mathbf{H}_C^{(k)},\mathbf{W}_3,\mathbf{C}^{(k-1)}\}$
  $\mathbf{H}_D^{(k)} = \mathbf{C}^{(k)}\odot\mathbf{B}^{(k)}\odot\mathbf{A}^{(k)}$
  $\mathbf{D}^{(k)} \leftarrow$ Algorithm 3 with input $\{\mathbf{H}_D^{(k)},\mathbf{W}_4,\mathbf{D}^{(k-1)}\}$
  $k \leftarrow k+1$
end while
Retrun $\mathbf{A}^{(k)},\mathbf{B}^{(k)},\mathbf{C}^{(k)},\mathbf{D}^{(k)}$

VII. Simulated Tests

In this section, the performance of the proposed EgoTen community detection algorithm is assessed via benchmark synthetic networks, as well as real-world datasets. Experiments over Lancicchinetti-Fortunatoand-Radicci (LFR) synthetic benchmarks enables us to simulate networks with different levels of community mixing, as well as number of overlapping nodes. This proves helpful in highlighting the enhanced capacity of community detection achieved by exploiting the "higher-order" properties of vertices captured in the proposed egonet-tensor.

A. Benchmark Networks

LFR benchmark networks provide synthetic graphs with ground-truth communities, in which certain properties of real-world networks, namely power-law distribution for nodal degrees as well as community sizes are preserved. LFR networks are configured to have a total number of N nodes (vertices), nodal average degree $\bar{d}$, exponent of degree distribution γ_1, and exponent of community-size distribution γ_2. Furthermore, *mixing parameter* μ controls the community cohesion, where larger μ induces more connections among nodes in different communities, thus generating less cohesive communities. Moreover, parameters $\{o_n, o_m\}$ set the number of overlapping nodes, that is, nodes belonging to more than one community, and the number of communities with which these nodes are associated, respectively. We have compared the performance of the proposed EgoTen with (similarly-constrained) nonnegative matrix factorization (NMF) schemes over the adjacency matrix, as well as other state-of-the-art community detection schemes.

1) Matrix Versus Tensor Factorization: The experiments here focus on comparison between EgoTen and its matrix counterpart. In particular, we will demonstrate how the higher-order connectivity patterns as well as structured redundancy offered through EgoTen can increase the robustness of community detection in the case of overlapping, and highly-mixed communities.

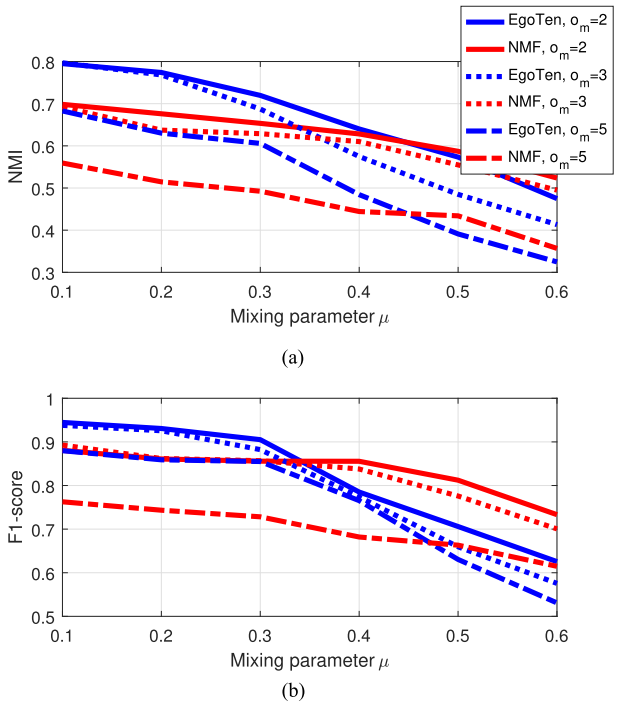


Fig. 4. Performance of constrained NMF and EgoTen in terms of (a) NMI; and (b) average F1-score, versus μ for LFR networks of $N = 1{,}000$, with $o_n = 200$.

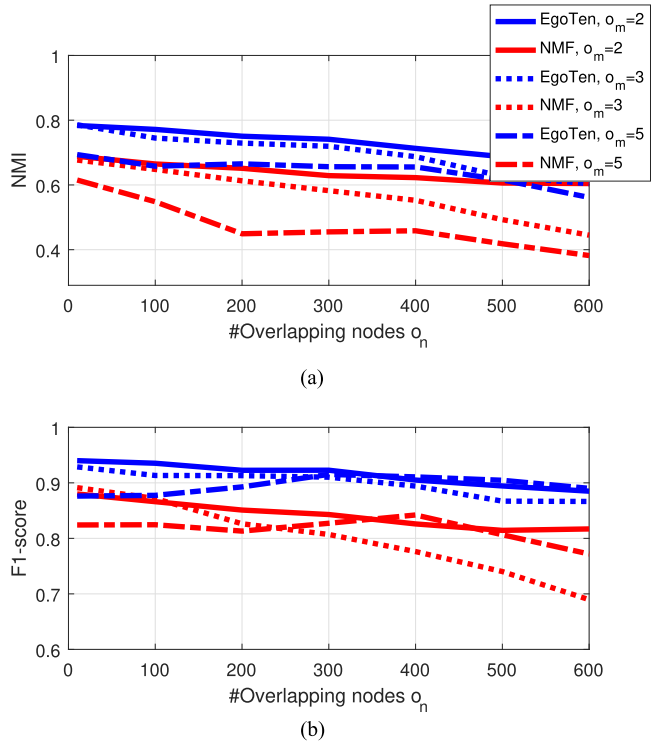


Fig. 5. Performance of constrained NMF and EgoTen in terms of (a) NMI; and (b) average F1-score, versus o_n for LFR networks with $N = 1{,}000$, and $\mu = 0.2$.

To this end, the corresponding NMF approach aims at factorizing matrix $\mathbf{W}$ by solving

$$\min_{\mathbf{U},\mathbf{V}} \|\mathbf{W} - \mathbf{U}\mathbf{V}^\top\|_F^2$$
$$\text{s.t.} \qquad \|\mathbf{u}_n\|_1 = 1 \,\forall n = 1,\ldots,N,\ \mathbf{U} \geq \mathbf{0}\,,\mathbf{V} \geq \mathbf{0} \qquad (20)$$

where $\mathbf{U},\mathbf{V} \in \mathbb{R}^{N\times K}$, and $\mathbf{U}^\top := [\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_N]$. Thus, similar to the nonnegative tensor decomposition in (3), the n-th row of matrix $\mathbf{U}$ contains community association coefficients of node n, and is subject to a simplex constraint. This minimization is solved via the AO-ADMM toolbox in [48]. Hard community assignments resulting from the factor $\mathbf{U}$ can be achieved similar to the procedure over the $\hat{\mathbf{C}}$ factor in EgoTen, as discussed in Section IV.

In our experiments, we have generated LFR networks with $N = 1{,}000, \gamma_1 = 2, \gamma_2 = 1$, average degree $\bar{d} = 100$, and $o_m = \{2,3,5\}$. To demonstrate the robustness of EgoTen, Fig. 4 shows the performance of EgoTen in comparison with constrained NMF over the adjacency matrix versus different values of μ, in terms of the NMI, and the F1-score. The experiments have been carried out for $o_n = 200$ (20%) number of overlapping nodes, upperbound on the number of communities is set as $K = 3|\mathcal{C}|$, and τ is chosen so that the highest NMI is achieved for both methods. Furthermore, Fig. 5 depicts the performance of NMF and EgoTen across different levels of overlap, controlled through the number of overlapping nodes o_n, while fixing $\mu = 0.2$. In addition, since a common practical concern is the lack of knowledge over the number of communities, Fig. 6 studies the robustness of NMF and EgoTen to this factor, by setting $K = \kappa|\mathcal{C}|$, and varying $\kappa \in [1.2, 5]$. As the plots in Figs. 4–6 corroborate, capturing the higher-order connectivity patterns

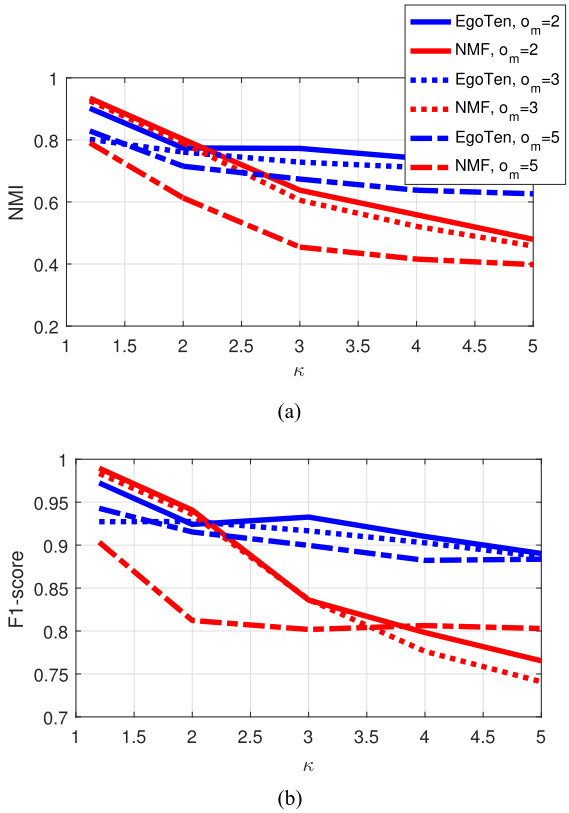
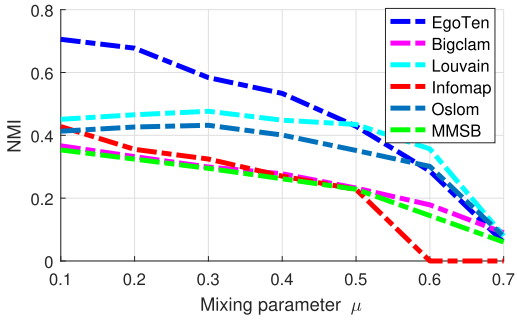
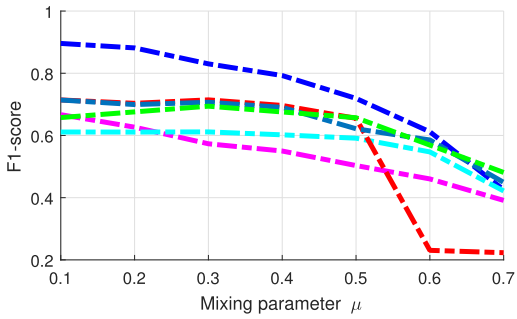


Fig. 6. Performance of constrained NMF and EgoTen in terms of (a) NMI; and, (b) average F1-score, versus upperbound on community number $K = \kappa|\mathcal{C}|$ parametrized by κ for LFR networks with $N = 1{,}000, o_n = 300$, and $\mu = 0.2$.

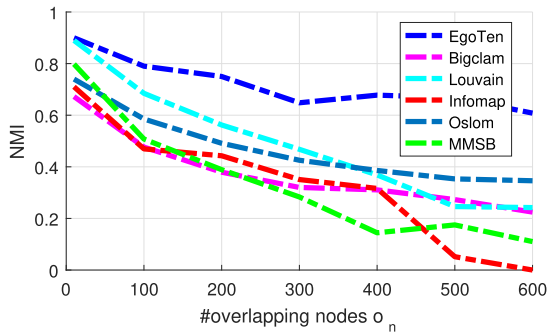


(a)

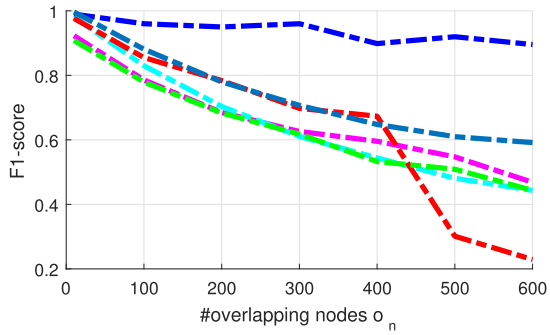


(b)

Fig. 7. Performance of EgoTen and state-of-the-art algorithms in terms of (a) NMI; and (b) average F-1 score, for LFR networks of $N = 1,000$, $o_n = 300$, and $o_m = 5$ versus μ.



(a)



(b)

Fig. 8. Performance of EgoTen and state-of-the-art algorithms in terms of (a) NMI; and (b) average F1-score, for LFR networks of $N = 1,000$, $\mu = 0.2$, and $o_m = 5$ versus o_n.

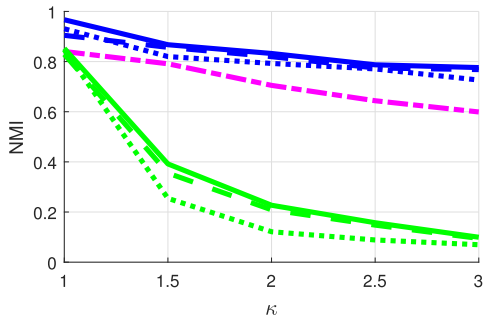
of the vertices and the structured decomposition of reinforced egonet-tensor improve robustness of nonnegative factorization methods against community coherence, presence of overlapping nodes, as well as rough estimates of the number of communities.

2) EgoTen vs. State-of-the-Art Methods: In this subsection, we compare the performance of EgoTen with state-of-the-art competitors, namely Oslom [57], Bigclam [9], Infomap [58], and Louvain [5], and MMSB [10]. Similar to the previous subsection, we have generated LFR networks with $N = 1,000$, $\gamma_1 = 2$, $\gamma_2 = 1$ and $o_m = \{2, 3, 5\}$. Number of overlapping nodes o_n as well as mixing parameter μ are varied in the range $[10, 600]$ and $[0.1, 0.7]$, respectively. Unless otherwise stated, threshold parameter for assigning hard community memberships of EgoTen is chosen as[1] $\tau = 1/K$.

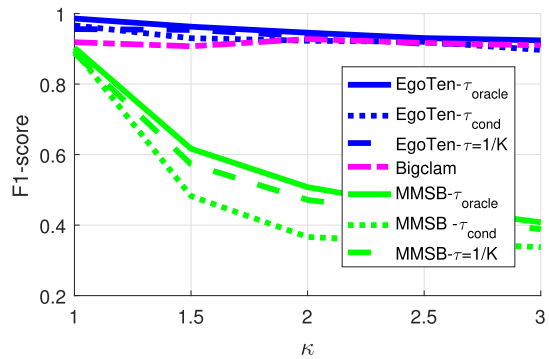
Performance is reported in terms of NMI, and F1-score. As Figs. 7 and 8 corroborate, EgoTen offers the highest robustness for a wide range of μ, as well as o_n, thanks to the reinforced structure of the egonet-tensor.

Furthermore, to investigate the influence of parameters K (and τ) on the performance of Bigclam, MMSB, and Egoten, Fig. 9 depicts NMI and F1-score for three choices of thresholding while κ in $K = \kappa|\mathcal{C}|$ is varied in $[1, 3]$. The three choices of thresholding are as follows: (a) $\tau = 1/K$, (b) τ_{cond} is selected by exhaustive search over $[0.01, 0.5]$ with a discretization of 0.01, and the value corresponding to the minimum conductance

[1]Since we have imposed a simplex constraint over the K association indices for any given node, $\tau = 1/K$ could be interpreted as having an association index higher than an equal association with all detected communities. Other selection schemes for τ, for instance setting to the value providing the community cover with the smallest average conductance, are also viable.



(a)



(b)

Fig. 9. Performance of EgoTen, Bigclam, and MMSB vs κ in $K = \kappa|\mathcal{C}|$, in terms of (a) NMI; and (b) average F1-score, in LFR network of $N = 1{,}000$, $\mu = 0.2$, $o_m = 5$, and $o_n = 200$.

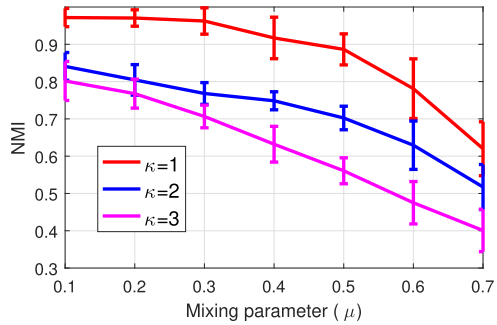


Fig. 10. Variance of NMI via EgoTen versus μ and κ averaged in LFR networks of $N = 1,000$, $o_m = 2$, and $o_n = 200$.

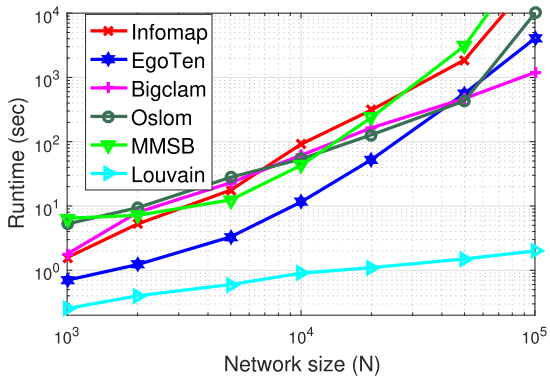


Fig. 11. Scalability of different algorithms as N grows.

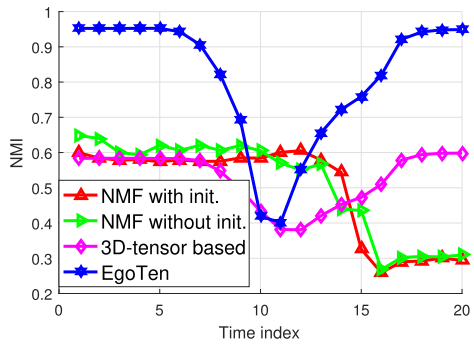


Fig. 12. NMI for the detected communities across time for synthetic time-varying graphs.

is finally selected, and (c) τ_{oracle} corresponds to the selection of τ giving the highest NMI. The last choice is assumed to be given by an oracle who knows the ground truth, and is provided to benchmark performance. As the plot suggests, while MMSB-based approaches provide interesting modeling of the problem and have attracted a lot of attention in the past years, their performance is highly influenced by how good of an estimate one has on the number of communities K. Thus, MMSB-modeling provides high accuracy when $K \simeq |\mathcal{C}|$, while its performance degrades dramatically as the mismatch increases. In contrast, EgoTen, followed by Bigclam, provide higher level of robustness, which we attribute to the enhanced structure of the egonet tensor inherent to its construction. To illustrate the effect of initialization, variance of the algorithm in terms of NMI is provided by the error-bar plot in Fig. 10, versus different values of μ and κ.

Regarding the scalability of the proposed EgoTen algorithm, Fig. 11 depicts run-time versus network size N while average

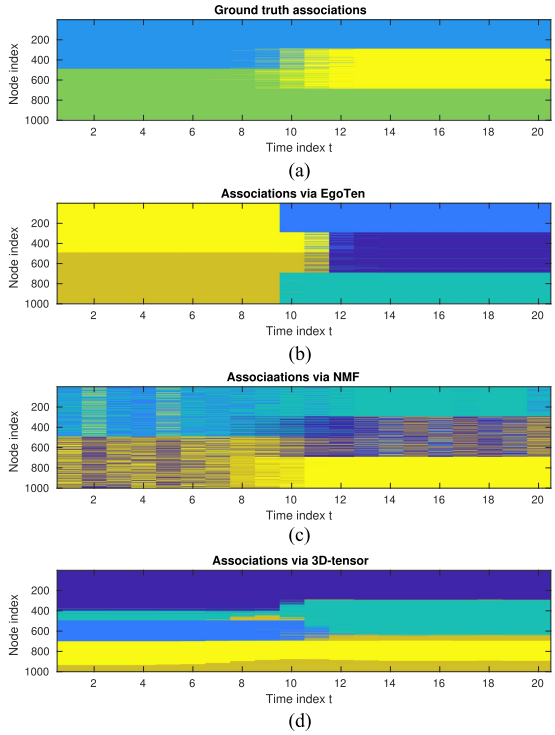


Fig. 13. Time-varying community association of nodes with transition times drawn from $\mathcal{N}(10, 1)$. Communities are color-coded; thus, for a given t on the x-axis, nodes in a community have the same color in (a) ground truth; and results via (b) EgoTen; (c) NMF; and (d) 3D-tensor decomposition.

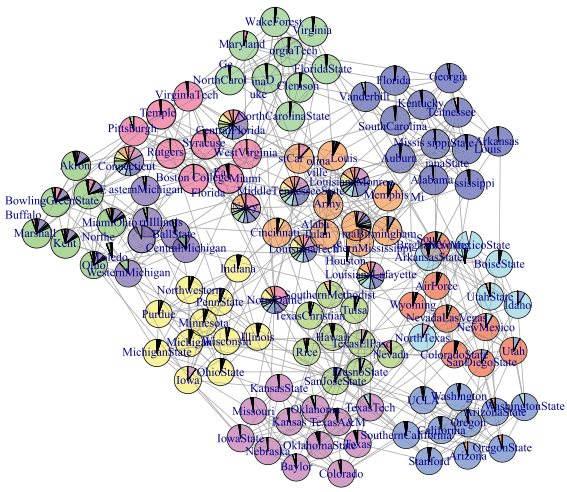


Fig. 14. Visualization of the American College Football Network with $N = 115$ and $K = 12$. Different colors correspond to different detected communities, and the pie-charts reveal soft community association of the nodes.

nodal degree is $\bar{d} = 20$. The experiment was run on an Intel (R) Core (TM) i7 4.00 GHz CPU with 8 cores and 32 GB of RAM. EgoTen, Bigclam, Louvain, and MMSB take effective use of parallelization in their implementation. As the plot corroborates, the introduced redundancy in EgoTen can be fairly alleviated via exploitation of sparsity as well as parallelization, thus endowing it with scalability, while the same cannot be claimed for all other competitors.

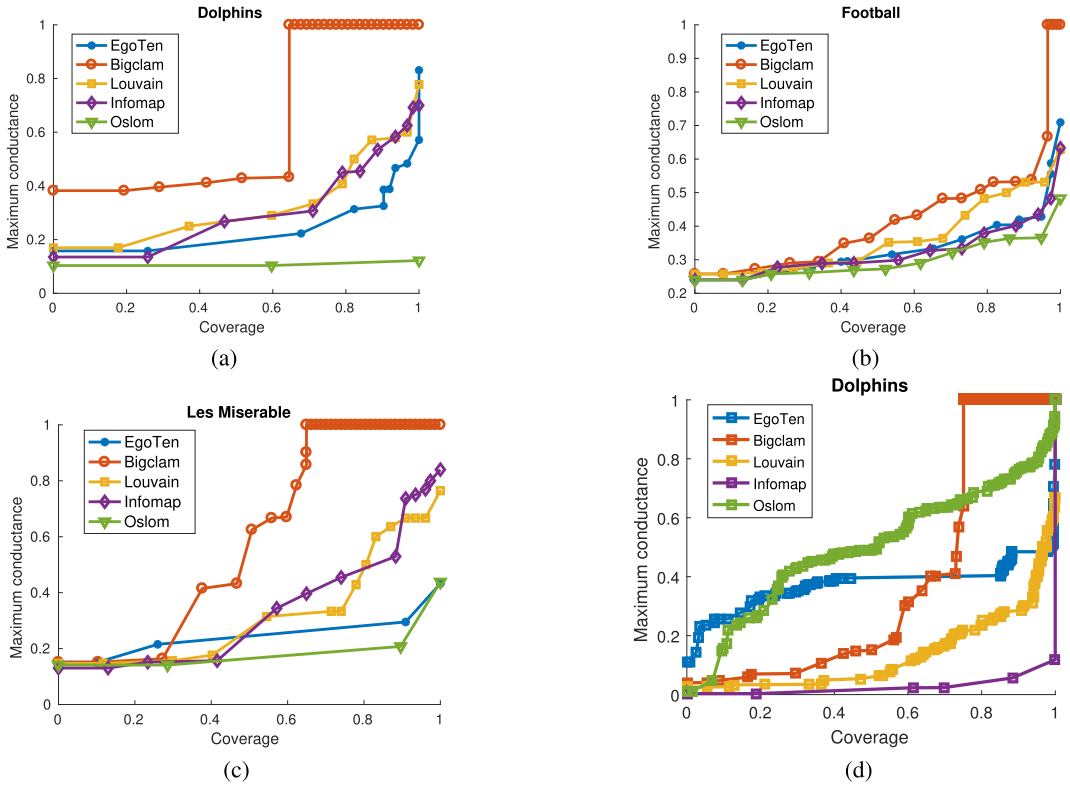


Fig. 15. Maximum-conductance versus coverage for real-world networks with underlying community structure. Lower curves correspond to better performance.

B. Time-Varying Graphs

In this subsection, the performance of the proposed EgoTen in Algorithm 4 for community identification over time-varying networks is assessed. To this end, we have generated a synthetic network with $N = 1,000$ nodes for a span of $T = 20$ slots. Initially at $t = 1$, the networks is generated with two distinct communities, each containing of 500 nodes. For $t > 1$, the community association of 600 randomly selected nodes remains unchanged, whereas the other 400 nodes migrate from their original community to a third newly-formed community. Transition slot τ_n for each of these nodes is identically drawn from a normal distribution $\mathcal{N}(10, 2)$. For any time slot t, the network edges are drawn according to a block stochastic model, where nodes within the same community are connected with probability 0.3, and out-of-community edges are drawn with probability 0.1. EgoTen's performance is compared with that of constrained NMF, for which $\mathbf{U}$ and $\mathbf{V}$ per t is used as initialization for $t + 1$ to provide NMF with consistency across time. We also provide the results of NMF when no initialization is utilized. Furthermore, performance is compared with the result of community detection via rank-K tensor decomposition of a 3D-tensor; the t-th slab of the tensor is the adjacency matrix of the network at time t. In addition to the nonegativity constraints, first and second factors are regularized with Frobenius norm to resolve scalar ambiguity, and the rows of the third term is subject to simplex constraints. This comparison is to highlight the advantage of the proposed enhanced network representation via the tensor of egonets.

Performance is measured in terms of NMI, and it is averaged over 100 realizations of the network in Fig. 12. Furthermore, Fig. 13 illustrates the identified communities for different nodes across time for a realization. That is, for any t on the x-axis, nodes associated with the same community are shown with the same color. The plot depicts ground truth as well as 3D-tensor, EgoTen and NMF results, where we have perturbed ordering of nodes for a better visualization. Clearly, EgoTen successfully identifies the two initial communities, as well as three communities after the migration of a subset of nodes, presenting solid blocks similar to those in the ground truth, while communities detected via constrained NMF and 3D-tensor decomposition are of lower quality.

TABLE I
REAL-WORLD NETWORKS

	Size N	Edges
Dolphins	62	159
Les miserable	77	254
Football	115	613
Facebook	4039	88,234

C. Real World Networks

In this section EgoTen is utilized for performing community detection on a number of real-world networks, tabulated in Table I. To study the quality of detected communities for real-world networks -whose ground truth community association is often unavailable- we examine the conductance of detected communities. In particular, given a cover $\mathcal{C} = \{\mathcal{C}_1, ..., \mathcal{C}_K\}$, let us compute the conductance $\phi(\mathcal{C}_i)$ for $i = 1, \ldots, K$. Corresponding to a value $\nu \in [0\ 1]$, let us define the set of communities whose conductance is less than ν, i.e., $\mathcal{S}_\nu := \{\mathcal{C}_j | \phi(\mathcal{C}_j) < \nu\}$. Then, coverage($\nu$) is defined as

$$\text{coverage}(\nu) := \frac{\left|\cup_{\mathcal{C}_i \in \mathcal{S}_\nu} \mathcal{C}_i\right|}{N} \tag{21}$$

TABLE II
QUALITY OF DETECTED COVER ON REAL-WORLD NETWORKS IN TERMS OF AUC AND AVERAGE CONDUCTANCE

	EgoTen			Bigclam			Louvain		Infomap			Oslom		
	$\lvert\mathcal{C}\rvert$	AUC	$\bar{\phi}(\mathcal{C})$	$\lvert\mathcal{C}\rvert$	AUC	$\bar{\phi}(\mathcal{C})$	$\lvert\mathcal{C}\rvert$	AUC (or $\bar{\phi}(\mathcal{C})$)	$\lvert\mathcal{C}\rvert$	AUC	$\bar{\phi}(\mathcal{C})$	$\lvert\mathcal{C}\rvert$	AUC	$\bar{\phi}(\mathcal{C})$
Dolphins	10	0.2984	0.4584	5	0.6176	0.6176	11	0.3902	10	0.3201	0.4012	2	0.1034	0.1034
Les miserable	5	0.2803	0.2803	10	0.6042	0.4666	15	0.3343	12	0.3534	0.4052	3	0.2127	0.2182
Football	15	0.4085	0.3480	15	0.4989	0.4101	15	0.3752	12	0.3974	0.3468	11	0.3430	0.3037
Facebook	100	0.3768	0.4931	100	0.3798	0.4973	100	0.1329	5	0.0370	0.0360	110	0.4998	0.4955

Consequently, conductance-coverage curve is plotted by varying the value ν from 0 to 1 on the y-axis and reporting the corresponding coverage value on the x-axis. As low values of conductance correspond to more cohesive communities, smaller area under curve (AUC) implies better performance. Fig. 15 plots the coverage-conductance curve and Table II tabulates AUC as well as average conductance defined in (16). Since Louvain does not allow for overlapping nodes, the two metrics coincide, hence the corresponding one column in Tabel II. As the results corroborate, the quality of detected communities via EgoTen is closely competing or outperforming the ones provided by other methods, while it remains robust to "resolution limit" [59] observed in Oslom in Fig. 15(a) and Infomap in Fig. 15(d) whose performance are limited to detecting only large communities.

VIII. CONCLUSION AND REMARKS

By viewing networks as the union of nodal egonets, a novel tensor-based representation for capturing high-order nodal connectivities has been introduced. The induced redundancy in the constructed egonet-tensor bestows the novel representation with rich structure, and is utilized for community detection by casting the problem as a constrained tensor decomposition task. Utilization of tensor sparsity as well as parallel computation endow the algorithm with scalability, while the structured redundancy enhances the performance against overlapping and highly-mixed communities. The proposed framework is broadened to accommodate time-varying graphs, where a four-dimensional tensor enables simultaneous community identification over the entire horizon yielding an improved performance.

As a natural extension, one can generalize the tensor-based representation to account for adjacency matrices capturing the connectivity of $2, 3, \ldots, d_{\max}$-hop neighbors. This approach indeed highlights the tradeoff between flexibility and redundancy, as memory and computational intensity of the corresponding CPD as well as proper tuning of parameters will influence the quality of the detected communities. One could analyze this tradeoff to further characterize how the quality of detected communities evolves as the coverage of *extended-egonets* increases; however, this goes beyond the scope of this work, and is left for future investigation.

APPENDIX A

Before proving Proposition 2, we will establish the following Lemma.

Lemma: For the $N \times N$ matrix

$$\mathbf{M} = \begin{bmatrix} \alpha \mathbf{1}_{N/2\times N/2}, \beta \mathbf{1}_{N/2\times N/2} \\ \beta \mathbf{1}_{N/2\times N/2}, \beta \mathbf{1}_{N/2\times N/2} \end{bmatrix} \tag{22}$$

with $0 < \beta < \alpha$, it holds that $\text{rank}(\mathbf{M}) = 2$. Furthermore, for the two eigenvectors $\lambda_1(\mathbf{M}) > \lambda_2(\mathbf{M}) > 0$, we have:

i) $\lambda_1(\mathbf{M}) > N\alpha/2$;
ii) $\lambda_2(\mathbf{M}) < N\beta/2$;
iii) the eigen-gap is at least $\lambda_1(\mathbf{M}) - \lambda_2(\mathbf{M}) > N(\alpha-\beta)/2$.

Proof of Lemma: The claim on the rank is due to the decomposition

$$\mathbf{M} = (\alpha-\beta)[\mathbf{1}_{1\times N/2}, \mathbf{0}_{1\times N/2}]^\top [\mathbf{1}_{1\times N/2}, \mathbf{0}_{1\times N/2}] + \beta \mathbf{1}_{1\times N}^\top \mathbf{1}_{1\times N}.$$

Furthermore, due to the symmetry of $\mathbf{M}$, the two eigenvectors take the form

$$\mathbf{v} = [\gamma \mathbf{1}_{1\times N/2}, \eta \mathbf{1}_{1\times N/2}]^\top.$$

Solving $\mathbf{M}\mathbf{v} = \lambda \mathbf{v}$ for λ yields

$$\lambda^2 - \frac{N}{2}(\beta+\alpha)\lambda + \frac{N^2}{4}\beta(\alpha-\beta) = 0$$

from which the two eigenvalues are

$$\lambda_{1,2}(\mathbf{M}) = \frac{N}{4}(\alpha+\beta) \pm \frac{N}{4}\sqrt{(\alpha+\beta)^2 - 4\beta(\alpha-\beta)} > 0.$$

Thus,

$$\lambda_1(\mathbf{M}) = \frac{N}{4}(\alpha+\beta) + \frac{N}{4}\sqrt{(\alpha-\beta)^2 + 4\beta^2} \geq \frac{N}{2}\alpha,$$

$$\lambda_2(\mathbf{M}) = \frac{N}{4}(\alpha+\beta) - \frac{N}{4}\sqrt{(\alpha-\beta)^2 + 4\beta^2} \leq \frac{N}{2}\beta$$

where we have used the inequality $\sqrt{(\alpha-\beta)^2+4\beta^2} \geq \alpha - \beta$ for $\alpha > \beta$. Clearly, their difference satisfies

$$\lambda_1(\mathbf{M}) - \lambda_2(\mathbf{M}) > \frac{N}{2}(\alpha-\beta).$$

Furthermore, the eigenvector $\mathbf{v}_1$ corresponding to the top eigenvalue λ_1 is

$$\mathbf{v}_1 = [\gamma_1 \mathbf{1}_{N/2\times 1}^\top, \eta_1 \mathbf{1}_{N/2\times 1}^\top]^\top$$

where using $\mathbf{M}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, $\lambda_1 \geq N\alpha/2$, and given $\alpha > 2\beta$

$$\frac{\eta_1}{\gamma_1} = \frac{N\beta}{2\lambda_1 - \beta N} < \frac{\beta}{\alpha-\beta} < 1.$$

Thus selecting the $N/2$ largest-amplitude entries of $\mathbf{v}_1$ unravels the member nodes in community $\mathcal{C}_1$. ■

Proof of Proposition 2: To start, consider the structure of the egonet adjacency of nodes in community $\mathcal{C}_1$. Without loss of generality, suppose that the vertices are indexed such that $\mathcal{C}_1 = \{v_1, v_2, \ldots, v_{N/2}\}$, $\mathcal{C}_2 = \{v_{N/2+1}, \ldots, v_N\}$, and partition the adjacency $\overline{\mathbf{W}}_1$ into four submatrices each of size $N/2 \times N/2$, as

$$\overline{\mathbf{W}}_1 = \begin{bmatrix} \mathbf{\Omega}_{11}, \mathbf{\Omega}_{12} \\ \mathbf{\Omega}_{21}, \mathbf{\Omega}_{22} \end{bmatrix}. \tag{23}$$

Next, focus on the block $\mathbf{\Omega}_{11}$ that captures the edges among the within-community neighbors of 'ego node' $v_n \in \mathcal{C}_1$, i.e., central nodes in an egonet in expectation. For $1 \leq i, j \leq N/2$ and $i \neq j$ we have

$$\begin{aligned}\overline{\mathbf{W}}_1(i,j) &= \mathbb{E}\left[\frac{1}{N/2}\sum_{n=1}^{N/2} 1_{(w_{i,j}^{(n)}=1)}\right] = \frac{1}{N/2}\sum_{n=1}^{N/2}\Pr\{w_{i,j}^{(n)}=1\} \\ &= \frac{1}{N/2}\Bigg(\sum_{n=1,n\neq i,j}^{N/2}\Pr\{w_{i,n}=1, w_{j,n}=1, w_{i,j}=1\} \\ &\qquad + \Pr\{w_{i,j}=1\} + \Pr\{w_{i,j}=1\}\Bigg) \\ &= p^3 + \frac{4}{N}p - \frac{4}{N}p^3\end{aligned}$$

where we have split the summation over three parts (i) $n = 1, \ldots, N/2, n \neq i, j$, (ii) $n = i$, and (iii) $n = j$; and have also used the fact that since nodes $i, j, n \leq N/2$ are in community $\mathcal{C}_1$, their pair-wise edge probability is equal to p, and independent. However, the diagonal entries of $\mathbf{\Omega}_{11}$ are zero, due to lack of self loops in the graph. Thus, with $\alpha := p^3 + 4p/N - 4p^3/N$, the $\mathbf{\Omega}_{11}$ submatrix can be briefly rewritten as $\mathbf{\Omega}_{11} = \alpha \mathbf{1}_{N/2\times 1}\mathbf{1}_{N/2\times 1}^\top - \alpha \mathbf{I}_{N/2}$.

Similarly, the submatrix $\mathbf{\Omega}_{22}$ capturing the $\overline{\mathbf{W}}_1(i,j)$ entries for $i, j > N/2$ and $i \neq j$ yields

$$\begin{aligned}\overline{\mathbf{W}}_1(i,j) &= \mathbb{E}\left[\frac{1}{N/2}\sum_{n=1}^{N/2} 1_{(w_{i,j}^{(n)}=1)}\right] = \frac{1}{N/2}\sum_{n=1}^{N/2}\Pr\{w_{i,j}^{(n)}=1\} \\ &= \frac{2}{N}\sum_{n=1}^{N/2}\Pr\{w_{i,n}=1, w_{j,n}=1, w_{i,j}=1\} = pq^2\end{aligned}$$

where we have used $i, j \in \mathcal{C}_2$ and $n \in \mathcal{C}_1$. Furthermore, the diagonal entries of $\mathbf{\Omega}_{22}$ are zero due to the absence of self loops. Thus, with $\beta := pq^2$, $\mathbf{\Omega}_{22}$ can be rewritten as $\mathbf{\Omega}_{22} = \beta \mathbf{1}_{N/2\times 1}\mathbf{1}_{N/2\times 1}^\top - \beta \mathbf{I}_{N/2}$.

Finally, regarding entries of $\mathbf{\Omega}_{12}$, the (i,j) entry of $\overline{\mathbf{W}}_1$ for $1 \leq i \leq N/2$ and $N/2 < j$ is by definition

$$\begin{aligned}\overline{\mathbf{W}}_1(i,j) &= \mathbb{E}\left[\frac{1}{N/2}\sum_{n=1}^{N/2} 1_{(w_{i,j}^{(n)}=1)}\right] = \frac{1}{N/2}\sum_{n=1}^{N/2}\Pr\{w_{i,j}^{(n)}=1\} \\ &= \frac{2}{N}\Bigg(\sum_{n=1,n\neq i}^{N/2}\Pr\{w_{i,n}=1, w_{j,n}=1, w_{i,j}=1\} \\ &\qquad + \Pr\{w_{i,j}=1\}\Bigg) = pq^2 + \frac{2}{N}q - \frac{2}{N}pq^2\end{aligned}$$

where we have used that $i, n \in \mathcal{C}_1$ and $j \in \mathcal{C}_2$, as well as the edge independence. Similar argument holds for the submatrix $\mathbf{\Omega}_{21}$ due to symmetry, which can be rewritten as

$$\mathbf{\Omega}_{21} = \mathbf{\Omega}_{12} := (\beta + \Delta\beta)\mathbf{1}_{N/2\times 1}\mathbf{1}_{N/2\times 1}^\top$$

with $\Delta\beta := \frac{2}{N}q - \frac{2}{N}pq^2$.

Putting together the unraveled structure of matrix $\overline{\mathbf{W}}_1$, we obtain $\overline{\mathbf{W}}_1 = \mathbf{M} + \Delta\mathbf{M} + \mathbf{\Gamma}$, where

$$\Delta\mathbf{M} := \left[\begin{array}{c|c} \mathbf{0} & \Delta\beta\mathbf{1}_{N/2\times 1}\mathbf{1}_{N/2\times 1}^\top \\ \hline \Delta\beta\mathbf{1}_{N/2\times 1}\mathbf{1}_{N/2\times 1}^\top & \mathbf{0}\end{array}\right]$$

$$\mathbf{\Gamma} := \begin{bmatrix} -\alpha\mathbf{I}_{N/2} & \mathbf{0} \\ \mathbf{0} & -\beta\mathbf{I}_{N/2}\end{bmatrix}.$$

It can be easily shown that $\|\mathbf{\Gamma}\|_2 = \alpha$ due to $\beta < \alpha$, and $\|\Delta\mathbf{M}\|_2 \leq \|\Delta\mathbf{M}\|_F \leq \sqrt{2}(q - pq^2)$. Thus, using the Cauchy-Schwarz inequality, while substituting α and omitting the negative terms, we arrive at

$$\|\overline{\mathbf{W}}_1 - \mathbf{M}\|_2 \leq \|\Delta\mathbf{M}\|_2 + \|\mathbf{\Gamma}\|_2 \leq p^3 + \frac{4}{N}p + \sqrt{2}q.$$

Furthermore, following the proof of Lemma 1, in networks with large N and given $p > \sqrt{2}q$, the $N/2$ largest-amplitude entries of the eigenvector corresponding to the largest eigenvalue reveal the members of community $\mathcal{C}_1$. ■

Intuitively, the ℓ_2-norm of matrices $\Delta\mathbf{M}$ and $\mathbf{D}$ are to be compared with the largest eigenvalue of $\lambda_1(\mathbf{M}) > \alpha N/2$. These two terms can be interpreted as small perturbations on the main $\mathbf{M}$ matrix. Hence, for sufficiently large network, the eigenvalues and vector pairs are slightly perturbed due to these terms, but the overall structure of community $\mathcal{C}_1$ can be inferred from the eigenvector corresponding to the largest eigenvalue.

APPENDIX B

Proof of Proposition 3: Under arguments similar to those as in the proof of Prop. 2, it can be shown that the expected egonet-adjacency for community $\mathcal{C}_2$ is

$$\overline{\mathbf{W}}_2 = \begin{bmatrix}\mathbf{\Omega}'_{11}, \mathbf{\Omega}'_{1,2} \\ \mathbf{\Omega}'_{21}, \mathbf{\Omega}'_{22}\end{bmatrix}$$

where $\mathbf{\Omega}'_{11} = \beta\mathbf{1}_{N/2\times 1}\mathbf{1}_{N/2\times 1}^\top - \beta\mathbf{I}_{N/2}$, $\mathbf{\Omega}'_{12} = \mathbf{\Omega}'_{21} := (\beta + \Delta\beta)\mathbf{1}_{N/2\times 1}\mathbf{1}_{N/2\times 1}^\top$, $\mathbf{\Omega}'_{22} = \alpha\mathbf{1}_{N/2\times 1}\mathbf{1}_{N/2\times 1}^\top - \alpha\mathbf{I}_{N/2}$, and α and β are defined in Appendix A. Thus using the Lemma with appropriate permutation matrix, $\overline{\mathbf{W}}_2$ can be effectively approximated by a rank-2 matrix

$$\mathbf{M}' = \beta\mathbf{v}_0 \circ \mathbf{v}_0 + (\alpha - \beta)\mathbf{v}_2 \circ \mathbf{v}_2$$

while matrix $\overline{\mathbf{W}}_1$ can be approximated by a rank-2 matrix

$$\mathbf{M} = \beta\mathbf{v}_0 \circ \mathbf{v}_0 + (\alpha - \beta)\mathbf{v}_1 \circ \mathbf{v}_1$$

Building on the common term in $\mathbf{M}$ and $\mathbf{M}'$, and utilizing a rank-3 CPD triplet of factors

$$\mathbf{A} = \mathbf{B} = [\sqrt{\beta}\mathbf{v}_0, \sqrt{\alpha - \beta}\mathbf{v}_1, \sqrt{\alpha - \beta}\mathbf{v}_2], \mathbf{C} = \begin{bmatrix}1, 1, 0 \\ 1, 0, 1\end{bmatrix},$$

the Forbenius cost of rank-3 approximation of tensor $\underline{\widehat{\overline{\mathbf{W}}}}$ can be bounded as

$$\begin{aligned}
&\min_{\underline{\widehat{\overline{\mathbf{W}}}},\text{rank}(\underline{\widehat{\overline{\mathbf{W}}}})=3} \frac{1}{N}\|\underline{\overline{\mathbf{W}}}-\underline{\widehat{\overline{\mathbf{W}}}}\|_F \\
&\quad\le \frac{1}{N}\Big(\|\overline{\mathbf{W}}_1-\mathbf{A}\text{diag}(1,1,0)\mathbf{B}^\top\|_F^2 \\
&\qquad\quad +\|\overline{\mathbf{W}}_2-\mathbf{A}\text{diag}(1,0,1)\mathbf{B}^\top\|_F^2\Big)^{1/2} \\
&\quad\le \frac{1}{N}\sqrt{\|\overline{\mathbf{W}}_1-\mathbf{M}\|_F^2+\|\overline{\mathbf{W}}_2-\mathbf{M}'\|_F^2} \\
&\quad\le \frac{\sqrt{2}}{N}\|\overline{\mathbf{W}}_1-\mathbf{M}\|_F \\
&\quad\le \frac{\sqrt{2}}{N}\Big(\|\mathbf{\Gamma}\|_F+\|\Delta\mathbf{M}\|_F\Big) \\
&\quad\le \frac{\sqrt{2}}{N}\left(\sqrt{\frac{N}{2}(\alpha^2+\beta^2)}+\frac{\sqrt{2}N}{2}\Delta\beta\right) \\
&\quad\le \frac{\sqrt{(p^3+4p/N)^2+p^2q^4}}{\sqrt{N}}+\frac{q(1-pq)}{N}. \quad \blacksquare
\end{aligned}$$

APPENDIX C

Proof of Proposition 4: One can readily rewrite the CPD objective corresponding to $\underline{\mathbf{W}}$ as

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N}\|\mathbf{W}^{(n)}-\mathbf{A}\text{diag}(\tilde{\mathbf{c}}_n)\mathbf{B}^\top\|_F^2 \\
&=\frac{1}{2}\left(\frac{1}{(N/2)}\sum_{v_n\in\mathcal{C}_1}\|\mathbf{W}^{(n)}-\mathbf{A}\text{diag}(\tilde{\mathbf{c}}_n)\mathbf{B}^\top\|_F^2\right. \\
&\quad\left.+\frac{1}{(N/2)}\sum_{v_n\in\mathcal{C}_2}\|\mathbf{W}^{(n)}-\mathbf{A}\text{diag}(\tilde{\mathbf{c}}_n)\mathbf{B}^\top\|_F^2\right) \\
&\ge\frac{1}{2}\left(\left\|\frac{1}{(N/2)}\sum_{v_i\in\mathcal{C}_1}\mathbf{W}^{(i)}-\mathbf{A}\frac{1}{(N/2)}\sum_{v_n\in\mathcal{C}_1}(\text{diag}(\tilde{\mathbf{c}}_n))\mathbf{B}^\top\right\|_F^2\right. \\
&\quad\left.+\left\|\frac{1}{(N/2)}\sum_{v_i\in\mathcal{C}_2}\mathbf{W}^{(i)}-\mathbf{A}\frac{1}{(N/2)}\sum_{v_n\in\mathcal{C}_2}(\text{diag}(\tilde{\mathbf{c}}_n))\mathbf{B}^\top\right\|_F^2\right) \\
&=\frac{1}{2}\left(\|\widetilde{\mathbf{W}}_1-\mathbf{A}\text{diag}(\check{\mathbf{c}}_1)\mathbf{B}^\top\|_F^2+\|\widetilde{\mathbf{W}}_2-\mathbf{A}\text{diag}(\check{\mathbf{c}}_2)\mathbf{B}^\top\|_F^2\right) \\
&\ge\underset{\{\mathbf{A},\mathbf{B},\check{\mathbf{c}}_1,\check{\mathbf{c}}_2\}\ge\mathbf{0}}{\arg\min}\frac{1}{2}\sum_{k=1}^{2}\|\widetilde{\mathbf{W}}_k-\mathbf{A}\text{diag}(\check{\mathbf{c}}_k)\mathbf{B}^\top\|_F^2 \qquad (24)
\end{aligned}$$

where the first inequality follows from Jenson's inequality over the (convex) Frobenius norm square, and in the last line, $\check{\mathbf{c}}_1=\frac{1}{(N/2)}\sum_{v_n\in\mathcal{C}_1}\text{diag}(\tilde{\mathbf{c}}_n)$ and $\check{\mathbf{c}}_2=\frac{1}{(N/2)}\sum_{v_n\in\mathcal{C}_2}\text{diag}(\tilde{\mathbf{c}}_n)$. Minimizing the left hand side with respect to $\{\mathbf{A},\mathbf{B},\mathbf{C}\}\ge\mathbf{0}$ completes the proof. ■

REFERENCES

[1] J. D. Power *et al.*, "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, Nov. 2011.

[2] Y. R. Lin, Y. Chi, S. Zhu, H. Sundaram, and B. L. Tseng, "Analyzing communities and their evolutions in dynamic social networks," *ACM Trans. Knowledge Discovery Data*, vol. 3, no. 2, pp. 8:1–8:31, Apr. 2009.

[3] S. Papadopoulos, Y. Kompatsiaris, A. Vakali, and P. Spyridonos, "Community detection in social media," *Data Mining Knowledge Discovery*, vol. 24, no. 3, pp. 515–554, 2012.

[4] P. K. Reddy, M. Kitsuregawa, P. Sreekanth, and S. S. Rao, "A graph-based approach to extract a neighborhood customer community for collaborative filtering," in *Proc. Int. Workshop Databases Netw. Inf. Syst.*, 2002, pp. 188–200.

[5] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Statist. Mech., Theory Exp.*, vol. 2008, no. 10, 2008, Art. no. P10 008.

[6] J. Duch and A. Arenas, "Community detection in complex networks using extremal optimization," *Phys. Rev. E*, vol. 72, no. 2, 2005, Art. no. 027 104.

[7] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing, "Mixed membership stochastic blockmodels," in *Proc. Adv. Neural Inf. Process. Syst.*, Vancouver, Canada, Dec. 2009, pp. 33–40.

[8] A. Anandkumar, R. Ge, D. Hsu, and S. M. Kakade, "A tensor approach to learning mixed membership community models," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 2239–2312, Jan. 2014.

[9] J. Yang and J. Leskovec, "Overlapping community detection at scale: A nonnegative matrix factorization approach," in *Proc. ACM Int. Conf. Web Search Data Mining*, Rome, Italy, Feb. 2013, pp. 587–596.

[10] X. Mao, P. Sarkar, and D. Chakrabarti, "On mixed memberships and symmetric nonnegative matrix factorizations," in *Proc. Int. Conf. Mach. Learn.*, Sydney, Australia, 2017, pp. 2324–2333.

[11] P. K. Gopalan, S. Gerrish, M. Freedman, D. M. Blei, and D. M. Mimno, "Scalable inference of overlapping communities," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 2249–2257.

[12] I. El-Helw, R. Hofman, W. Li, S. Ahn, M. Welling, and H. Bal, "Scalable overlapping community detection," in *Proc. IEEE Int. Parallel Distrib. Process. Symp. Workshops.*, Chicago, IL, USA, 2016, pp. 1463–1472.

[13] I. Derényi, G. Palla, and T. Vicsek, "Clique percolation in random networks," *Phys. Rev. Lett.*, vol. 94, no. 16, Apr. 2005, Art. no. 160202.

[14] U. Von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, Aug. 2007.

[15] F. Wang, T. Li, X. Wang, S. Zhu, and C. Ding, "Community discovery using nonnegative matrix factorization," *ACM Trans. Data Mining Knowledge Discovery*, vol. 22, no. 3, pp. 493–521, Jul. 2011.

[16] X. Cao, X. Wang, D. Jin, Y. Cao, and D. He, "Identifying overlapping communities as well as hubs and outliers via nonnegative matrix factorization," *Sci. Rep.*, vol. 3, 2013, Art. no. 2993.

[17] I. Psorakis, S. Roberts, M. Ebden, and B. Sheldon, "Overlapping community detection using Bayesian non-negative matrix factorization," *Phys. Rev. E*, vol. 83, no. 6, 2011, Art. no. 066114.

[18] Y. Zhang and D.-Y. Yeung, "Overlapping community detection via bounded nonnegative matrix TRI-factorization," in *Proc. ACM Int. Conf. Knowledge Discovery Data Mining*, 2012, pp. 606–614.

[19] Z.-Y. Zhang, Y. Wang, and Y.-Y. Ahn, "Overlapping community detection in complex networks using symmetric binary matrix factorization," *Phys. Rev. E*, vol. 87, no. 6, 2013, Art. no. 062803.

[20] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, no. 3, pp. 75–174, Feb. 2010.

[21] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Phys. Rep.*, vol. 659, pp. 1–44, 2016.

[22] J. J. Whang, D. F. Gleich, and I. S. Dhillon, "Overlapping community detection using seed set expansion," in *Proc. ACM Int. Conf. Inf. Knowledge Manage.*, 2013, pp. 2099–2108.

[23] K. He, Y. Sun, D. Bindel, J. Hopcroft, and Y. Li, "Detecting overlapping communities from local spectral subspaces," in *Proc. IEEE Int. Conf. Data Mining*, 2015, pp. 769–774.

[24] J. J. Whang, I. S. Dhillon, and D. F. Gleich, "Non-exhaustive, overlapping k-means," in *Proc. SIAM Int. Conf. Data Mining*, 2015, pp. 936–944.

[25] E. E. Papalexakis, L. Akoglu, and D. Ience, "Do more views of a graph help? Community detection and clustering in multi-graphs," in *Proc. IEEE Int. Conf. Inf. Fusion*, 2013, pp. 899–905.

[26] E. E. Papalexakis, N. D. Sidiropoulos, and R. Bro, "From k-means to higher-way co-clustering: Multilinear decomposition with sparse latent factors," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 493–506, Jan. 2013.

[27] J. Yang, J. McAuley, and J. Leskovec, "Community detection in networks with node attributes," in *Proc. IEEE Int. Conf. Data Mining*, 2013, pp. 1151–1156.
[28] M. Araujo *et al.*, "Com2: Fast automatic discovery of temporal (comet) communities," in *Proc. Pac. Asia Conf. Knowledge Discovery Data Mining*, 2014, pp. 271–283.
[29] B. Baingana and G. B. Giannakis, "Joint community and anomaly tracking in dynamic networks," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2013–2025, Apr. 2016.
[30] F. Huang, U. Niranjan, M. U. Hakeem, and A. Anandkumar, "Online tensor methods for learning latent variable models," *J. Mach. Learn. Res.*, vol. 16, pp. 2797–2835, 2015.
[31] F. Sheikholeslami, B. Baingana, G. B. Giannakis, and N. D. Sidiropoulos, "Egonet tensor decomposition for community identification," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Washingto, DC, USA, Dec. 2016, pp. 341–345.
[32] A. R. Benson, D. F. Gleich, and J. Leskovec, "Tensor spectral clustering for partitioning higher-order network structures," in *Proc. SIAM Int. Conf. Data Mining*, Vancouver, Canada, Feb. 2015, pp. 118–126.
[33] T. G. Kolda, B. W. Bader, and J. P. Kenny, "Higher-order web link analysis using multilinear algebra," in *Proc. IEEE Int. Conf. Data Mining*, 2005, 8 pp.
[34] J. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.
[35] N. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *J. Chemometrics*, vol. 14, no. 3, pp. 229–239, Jun. 2000.
[36] L. Chiantini and G. Ottaviani, "On generic identifiability of 3-tensors of small rank," *SIAM J. Matrix Anal. Appl.*, vol. 33, no. 3, pp. 1018–1037, 2012.
[37] L. Akoglu, M. McGlohon, and C. Faloutsos, "Oddball: Spotting anomalies in weighted graphs," in *Proc. Adv. Knowledge Discovery Data Mining*, pp. 410–421, 2010.
[38] J. Leskovec and J. J. Mcauley, "Learning to discover social circles in ego networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 539–547.
[39] C. Lan, Y. Yang, X. Li, B. Luo, and J. Huan, "Learning social circles in ego-networks based on multi-view network structure," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 8, pp. 1681–1694, Aug. 2017.
[40] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
[41] E. E. Papalexakis, C. Faloutsos, and N. D. Sidiropoulos, "Parcube: Sparse parallelizable tensor decompositions," in *Proc. Joint Eur. Conf. Mach. Learn. Knowledge Discovery Databases*, Bristol, U.K., 2012, pp. 521–536.
[42] S. Smith and G. Karypis, "SPLATT: The Surprisingly ParalleL spArse Tensor Toolkit." [Online]. Available: http://cs.umn.edu/ splatt/
[43] M. Mardani, G. Mateos, and G. B. Giannakis, "Subspace learning and imputation for streaming big data matrices and tensors," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2663–2677, May 2015.
[44] J. A. Bazerque, G. Mateos, and G. B. Giannakis, "Rank regularization and Bayesian inference for tensor completion and extrapolation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5689–5703, Nov. 2013.
[45] H. Bozdogan, "Model selection and akaike's information criterion (AIC): The general theory and its analytical extensions," *Psychometrika*, vol. 52, no. 3, pp. 345–370, 1987.
[46] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
[47] G. B. Giannakis, Q. Ling, G. Mateos, I. D. Schizas, and H. Zhu, "Decentralized learning for wireless communications and networking," *Splitting Methods in Communication, Imaging, Science, and Engineering*. Berlin, Germany: Springer, 2016, pp. 461–497.
[48] K. Huang, N. D. Sidiropoulos, and A. P. Liavas, "A flexible and efficient algorithmic framework for constrained matrix and tensor factorization," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5052–5065, Oct. 2016.
[49] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient projections onto the l 1-ball for learning in high dimensions," in *Proc. Int. Conf. Mach. Learn.*, 2008, pp. 272–279.
[50] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[51] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.
[52] B. Karrer and M. E. Newman, "Stochastic blockmodels and community structure in networks," *Phys. Rev. E*, vol. 83, no. 1, Jan. 2011, Art. no. 016107.
[53] E. Abbe, "Community detection and stochastic block models: Recent developments," *J. Mach. Learn. Res.*, vol. 18, no. 177, pp. 1–86, 2018.
[54] A. Lancichinetti, S. Fortunato, and J. Kertész, "Detecting the overlapping and hierarchical community structure in complex networks," *New J. Phys.*, vol. 11, no. 3, 2009, Art. no. 033015.
[55] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney, "Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters," *Internet Math.*, vol. 6, no. 1, pp. 29–123, 2009.
[56] L. Gauvin, A. Panisson, and C. Cattuto, "Detecting the community structure and activity patterns of temporal networks: A non-negative tensor factorization approach," *PloS One*, vol. 9, no. 1, 2014, Art. no. e86028.
[57] A. Lancichinetti, F. Radicchi, J. J. Ramasco, and S. Fortunato, "Finding statistically significant communities in networks," *PloS One*, vol. 6, no. 4, 2011, Art. no. e18961.
[58] L. Bohlin, D. Edler, A. Lancichinetti, and M. Rosvall, "Community detection and visualization of networks with the map equation framework," *Measuring Scholarly Impact*. Berlin, Germany: Springer, 2014, pp. 3–34.
[59] S. Fortunato and M. Barthelemy, "Resolution limit in community detection," *Proc. Nat. Academy Sci.*, vol. 104, no. 1, pp. 36–41, 2007.

Fatemeh Sheikholeslami (S'13) received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Tehran, Tehran, Iran, and the Sharif University of Technology, Tehran, Iran, in 2010 and 2012, respectively. Since September 2013, she has been working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA. Her research interests include machine learning and network science.

Georgios B. Giannakis (F'97) received the Diploma degree in electrical engineering from the National Technical University of Athens, Athens, Greece, 1981 and the M.Sc. degree in electrical engineering in 1983, the M.Sc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986 from University of Southern California (USC), CA, USA.

From 1982 to 1986, he was with the University of Southern California, Los Angeles, CA, USA, He was with the University of Virginia from 1987 to 1998, and since 1999, he has been a Professor with the University of Minnesota, Minneapolis, MN, USA, where he holds an Endowed Chair in Wireless Telecommunications, a University of Minnesota McKnight Presidential Chair in Electrical and Computer Engineering, and is the Director of the Digital Technology Center. His general interests include communications, networking, and statistical learning—on which he has authored or coauthored more than 430 journal papers, 720 conference papers, 25 book chapters, two edited books, and two research monographs (h-index 133). Current research focuses on data science, Internet of things, and network science with applications to social, brain, and power networks with renewables. He is the (co-) inventor of 32 patents issued, and the (co-) recipient of nine best journal paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He was the recipient of Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, the G. W. Taylor Award for Distinguished Research from the University of Minnesota, and the IEEE Fourier Technical Field Award (inaugural recipient in 2015). He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SPS.