Community Detection in Dynamic Networks: Exact Recovery under Two Link Evolution Models

Javad Zahedi Moghaddam, Aria Nosratinia, *Fellow, IEEE,*

Abstract—This paper studies edge persistence and memory in time-varying (dynamic) networks, and their effect on community detection. In particular, we focus on two models representing network memory, using which we study the asymptotic behavior of community detection in dynamic networks of large size, manifested through their phase transition threshold. In the first part, we adopt a Markovian stochastic block model (SBM) in which the edge probabilities in each network snapshot depend not only on the respective node attributes but also on the previous network snapshot. Under this model, semi-definite relaxation achieves the optimal phase transition bound for exact recovery, as observed in other community detection problems. The adverse effect of edge persistence on perfect recovery is analyzed and highlighted. In the second part, we study networks where underlying communities change slowly compared with network measurements. We model this scenario via a time series of SBM wherein the node attributes are fixed within a window of a certain size and vary independently across windows. The phase transitions are calculated via semidefinite programming which, once again, is asymptotically optimal. Numerical simulations conducted on finite-size networks interpret the asymptotic results.

Index Terms—Community detection, stochastic block model, dynamic network, maximum likelihood detection, semidefinite programming, edge persistence, clustering.

1 Introduction

Community detection is the discovery of underlying communities among individuals in a network by analyzing their connections. It is a core problem in network analysis, and it has many applications in real-world networks including social networks [1], [2], [3], identification of protein complexes in biology [4], [5], [6], and transportation networks [7], among others.

Most studies in community detection utilize a network model involving a single, static instance of a graph. However, many real-world networks are temporal networks, where communities and the connections among individuals undergo variations over time. For example, the transfer season in the national football league can cause substantial changes in the communities and connections within the friendship graph of football players. The connections between players in different teams can experience significant change or remain relatively stable. Similar behaviors are commonly observed in many other networks that cannot adequately be captured within a static framework.

In dynamic community detection, research has focused primarily on developing computationally effective algorithms to recover communities under different evolving settings in temporal networks [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18]. Compared with static networks, there has been little work on the fundamental limits (also known as phase transitions) of community detection in dynamic networks [19], [20], [21], [22]. Furthermore, there are no studies that provide rigorous bounds on the impact of edge evolution on the fundamental limits of dynamic community detection under the exact recovery criterion, where the community detection algorithm aims to recover labels with a probability that converges to one.

• *Javad Zahedi Moghaddam and Aria Nosratinia are with the Department of Electrical and Computer Engineering, Erick Johnson School of Engineering and Computer Science, University of Texas at Dallas, Richardson, TX 75080, United States of America*
E-mail: {moghaddam.jz; aria}@utdallas.edu.

1.1 Contribution

This paper focuses on exact recovery in temporal graphs. We study the exact recovery phase transition under two link evolution models. The first involves edge persistence, and the second involves node label dynamics that are much slower than the frequency of graph observations.

Under the edge persistence model, some node pairs maintain the status of their interactions (edges) across time, beyond what is predicted by the underlying communities. For example, high-school classmates often maintain their friendship over an extended period, despite subsequent differences in political perspectives, religion, social class, or fandom of sports teams. To model graphs with edge persistence, we consider a stochastic block model (SBM) with Markovian edges such that current observations depend on prior snapshot observations as well as current hidden labels. We calculate the phase transition and propose a semi-definite programming (SDP) relaxation algorithm that achieves this phase transition.

The second model applies to scenarios in which variations in underlying communities are slow compared with the frequency of graph observations (snapshots). For example, in social networks, node attributes such as political affiliation or education level remain stationary on the scale of months or years, while the interactions of nodes (edges) can change within that time scale due to finding new friends,

new activities, or new preferences. Inspired by the utility of block models in a variety of settings, we employ a model in which time is divided into blocks or windows; the labels for each node remain constant within each window, and vary independently from one window to the next. We extend the idea of SBM to account for the sequence of edge observations between the consecutive snapshots occurring in each time window. We then determine the phase transition and detect communities using SDP relaxation. We show that SDP produces the same phase transition as the optimal estimator for the two models studied in this paper.

1.2 Related Works

Finding the exact recovery limits is a fundamental problem in statistical learning and information theory. To compute these thresholds, a stochastic framework is required that models observations based on the underlying affiliations in the network. The stochastic block model a widely known and utilized framework that offer such a probabilistic configuration. An SBM(n, p, q) is a generative model that assigns probabilities (p, q) to observations (intra-community and inter-community edges) based on node labels [23], where n is the number of nodes in the graph. Studies on SBM have shown exact recovery is possible when these probabilities are in dense regimes, i.e. $p = a\frac{\log(n)}{n}$, $q = \frac{b\log(n)}{n}$. In this regime, exact recovery is possible for static networks if and only if $\sqrt{a} - \sqrt{b} > \sqrt{k}$ where k is the number of communities [24]. In addition, computationally effective algorithms have been developed that can achieve the exact recovery threshold [25], [26], [27], [28], [29], [30], [31], [32], [33].

Different dynamic models such as preferential attachment graph [34], growing model [35], and graph convolutional networks [36] have been proposed in the literature to study dynamic community detection. However, an SBM equipped with some evolutionary parameters is the most prevailing one to study the statistical behavior and information-theoretic limits of dynamic community detection in networks. In [37], [38], [39], [40], [41], dynamic behavior is captured by modeling the evolution of connectivity parameters of an SBM as the system states, and an extended Kalman filter is employed for state prediction and updating to recover communities. These studies did not explicitly mention phase transition. Static models can also be adapted for evolving graphs by introducing a Markov Switching Model for each of the node labels across time [19], [22], [37], [42], [43], [44]. Ghasemian et al. [19] utilized temporal label dependencies for the improvement of detectability, but no formal analysis or bounds were offered. In [22], the exact recovery fundamental limits under label dependency for a CBM was established, showing that semi-definite programming can achieve phase transition. Other studies have modeled the appearance and disappearance of edges in temporal snapshots using either random rates [45], or a Markov dependency [20], [21]. In [20] and [21], both edge and node persistence were incorporated in the SBM via Markov dependencies between consecutive snapshots in evolving graphs, wherein it was conjectured that temporal link persistence diminishes the detectability, while temporal node persistence enhances it similarly to [19]. Previous work on edge persistence graphs has not formally addressed phase transition, in particular in the exact recovery regimes.

2 Edge Evolving Models

At each time instance $t = 1, 2, \ldots$, a graph with n nodes is assumed, whose node labels are characterized by the $n \times 1$ vector $\boldsymbol{g}(t)$ and the edges by the corresponding $n \times n$ matrix $\mathbf{A}(t)$. (See Fig. 1). We consider two variations of this model. In the first model, the nodes and edges of the graph vary on the same timescale, while in the second model, the node labels vary more slowly than the graph edges.

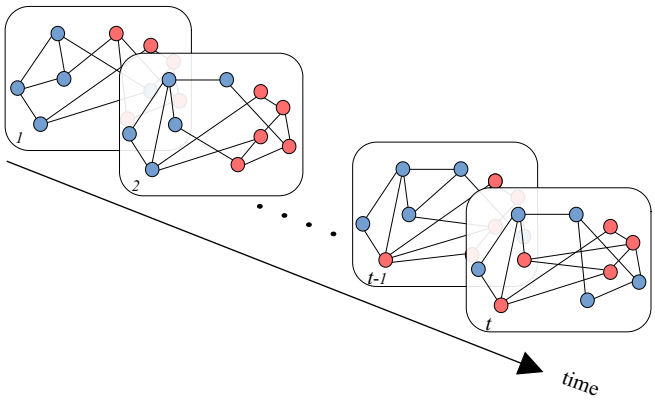


Fig. 1. Label and edge changes among nodes in temporal snapshots of an evolving network with binary communities

2.1 Markov SBM Graphs

We explore a situation where nodes are expected to maintain some of their connections across the sequence of graph observations, even as the underlying groups change. This implies the existence of persistent edges within sequential graph snapshots. The model is represented by probabilities of current values of the network edges, conditioned on the earlier instances of the network edges, and the network node labels in the current instance of the network. Under a balanced binary community detection setting, let $\mathbf{Z}(t) \triangleq \boldsymbol{g}(t)[\boldsymbol{g}(t)]^\top$ and:

$$\begin{aligned}\mathbb{P}\big(A_{ij}(k) \mid A_{ij}(k-1), Z_{ij}(k)\big) = (1-\eta)\delta_{A_{ij}(k),A_{ij}(k-1)} + \\ \eta p^{A_{ij}(k)\frac{1+Z_{ij}(k)}{2}}(1-p)^{(1-A_{ij}(k))\frac{1+Z_{ij}(k)}{2}} q^{A_{ij}(k)\frac{1-Z_{ij}(k)}{2}} \\ \times (1-q)^{(1-A_{ij}(k))\frac{1-Z_{ij}(k)}{2}} \qquad (1)\end{aligned}$$

where $Z_{ij}(k) = g_i(k)g_j(k)$, $p = \frac{a\ln(n)}{n}$, $q = \frac{b\ln(n)}{n}$, $a \geq b > 0$, and $\ln(n) = \log_e(n)$. In (1), $\eta \in [0, 1]$ is the edge persistence parameter Shown in Fig. 2.

The model reduces to a common SBM model at each snapshot for $\eta = 1$, while $\eta = 0$ means a fixed network at all snapshots. We now utilize the common asymptotic model for network probabilities that reveals the phase transition phenomenon, i.e., the network has edge probabilities

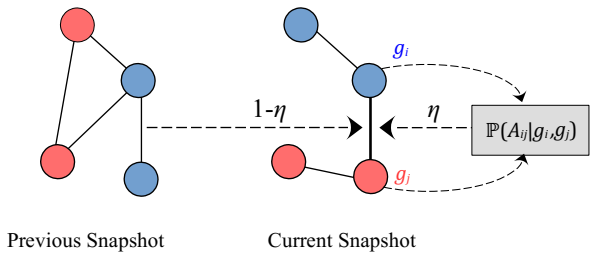


Fig. 2. Edge and node evolution in Markov SBM graphs with edge persistence parameter η

proportional to $\frac{\ln(n)}{n}$ and operates in the so-called dense regime.[1] Let:

$$\mathbb{P}(e_{ij}(t-1)=1 \mid Z_{ij}(t)=1)=\tilde{p}=\frac{\tilde{a}\ln(n)}{n} \tag{2}$$

$$\mathbb{P}(e_{ij}(t-1)=1 \mid Z_{ij}(t)=-1)=\tilde{q}=\frac{\tilde{b}\ln(n)}{n} \tag{3}$$

where $\tilde{a}$ and $\tilde{b}$ are non-negative constants. Considering binary node labels:

$$\begin{aligned}\mathbb{P}(e_{ij}(t)=1)&=\frac{1}{2}\big((1-\eta+\eta p)\tilde{p}+\eta p(1-\tilde{p})\big)+\\&\quad\frac{1}{2}\big((1-\eta+\eta q)\tilde{q}+\eta q(1-\tilde{q})\big)\\&=\frac{1}{2}\big((1-\eta)(\tilde{p}+\tilde{q})+\eta(p+q)\big)\end{aligned}$$

Since $\mathbb{P}(e_{ij}(t)=1)=\mathbb{P}(e_{ij}(t-1)=1$, we have:

$$\tilde{p}+\tilde{q}=p+q \tag{4}$$

Without loss of generality, we assume $\tilde{a}=a-\epsilon\geq\tilde{b}=b+\epsilon$ and $\epsilon\in[0,\frac{a-b}{2}]$.

2.2 Multi-observation Models

We now consider the case where the graph observations are frequent, or the node labels vary slowly relative to the sampling of the graph observations. To model such graphs, we use time-window block models that allow us to capture the persistent features of the network while providing a flexible framework to adapt to evolving interactions. Such scenarios are modeled in this paper as follows: we model the slowness of the variation of the node labels according to a piecewise constant model, i.e., the label for each node $g_i(t)$ is constant for a block of $\tau+1$ time units, and is independent from block to block[2] as shown in Fig. 3. Therefore, the values of $\mathbf{g}(t)$ for $t=k\tau$ carry all the information contained in the random process. We now characterize the model via the distribution $\mathbb{P}(A(t),A(t-1),\ldots,A(t-\tau)|Z(t))$. For ease of exposition, let s be a variable representing the state of the edge between nodes i,j over the block of length τ, namely,

$$[A_{ij}(t),A_{ij}(t-1),\ldots,A_{ij}(t-\tau)]$$

For a non-weighted graph model, graph edges take binary values, therefore without loss of generality, one may assume s to be an integer taking values in $[0,2^{\tau+1}-1]$. Then, graph observations depend on the node labels according to the following SBM:

$$\mathbb{P}(s\mid Z_{ij}(t))=\prod_{\ell}\big[p_\ell^{\frac{(1+Z_{ij}(t))}{2}}q_\ell^{\frac{(1-Z_{ij}(t))}{2}}\big]^{\mathbb{1}_{\{s=\ell\}}} \tag{5}$$

where $\mathbb{1}_{\{s=\ell\}}$ is the indicator function of set $\{s=\ell\}$, and p_ℓ is the probability that the edges in the block $[A_{ij}(t),A_{ij}(t-1),\ldots,A_{ij}(t-\tau)]$ taking the outcome configuration ℓ, conditioned on nodes i,j belonging to the same community. This means $Z_{ij}=1$. Obviously, $p_\ell\geq 0$ and $\sum_\ell p_\ell=1$. q_ℓ is the probability of edges between nodes i,j taking configuration ℓ, subject to nodes i,j belonging to different communities, in our case this means $Z_{ij}=-1$. Once again, $\sum_\ell q_\ell=1$. We assume graphs are in the so-called dense regime, i.e., $p_\ell=\frac{a_\ell\ln n}{n}$ and $q_\ell=\frac{b_\ell\ln n}{n}$ for all $\ell\neq 0$. The model in (5) does not capture any dynamics within the block, therefore we call this simply a *multi-observation model*.

It is obvious that the parameter space grows exponentially with the size of the block τ. To control the parameter space, we also focus on the special case where the dependence of the node labels with the edges s can be described through the temporal frequency of the appearance of the edge.

$$s'=\sum_{i=0}^{\tau}A_{ij}(t-i)$$

This will allow the parameter space to collapse so that it grows only linearly with the block size, and the representative state variable s' for the probabilities can take values between 0 and τ, such that. In either case, since $\mathbf{g}(t)$ is constant over blocks of length τ, we utilize τ graph observations for estimating the value of $\mathbf{g}(t)$.

3 Community Detection via SDP

To find the optimal estimator under each of the model settings in Section 2, we need first to formulate the maximum a-posteriori problem.

3.1 Markov SBM Graphs

Under binary, balanced communities, Eq. (1) shows that the maximum a-posteriori estimator and maximum likelihood estimator are equivalent. The maximum likelihood estimator can be formulated as follows:

$$\hat{\mathbf{Z}}(t)=\underset{\mathbf{Z}(t)}{\arg\max}\quad\mathbb{P}\big(\mathbf{A}(t),\mathbf{A}(t-1)\mid\mathbf{Z}(t)\big) \tag{6a}$$

$$\text{s.t.}\quad Z_{ii}(t)=1\quad i\in[n] \tag{6b}$$

$$\langle\mathbf{Z}(t),\mathbf{J}\rangle_F=0. \tag{6c}$$

where $\langle\cdot,\cdot\rangle_F$ indicates the Frobenius inner product of two matrices, $[n]\triangleq\{1,2,\cdots,n\}$, and $\mathbf{J}$ denotes a matrix where all entries are one. Optimization problem (6) is nonconvex

1. The closest characterization in the literature compared to this model is arguably [20], with two main differences. First, our model focuses on edge memory and isolates its effect. Second, our SBM graph parameters do not vary with time.

2. This structure has been used for modeling and analysis of other time-varying stochastic phenomena, for example, the block fading model for wireless channels [46]

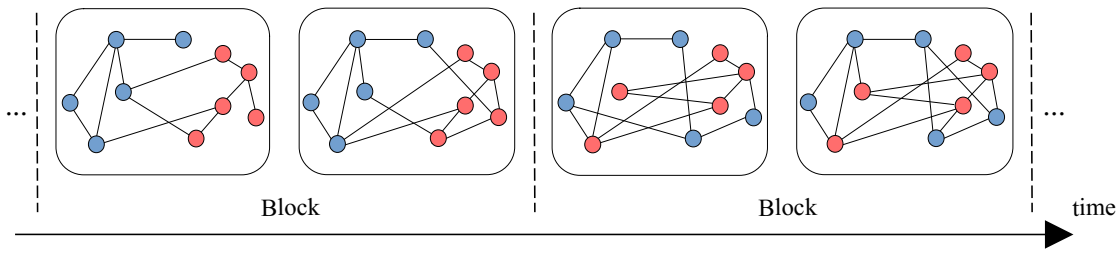


Fig. 3. Edge and node evolution of evolving binary network under Multi-observation configuration with two snapshots in each block

due to rank-one constraint on $\mathbf{Z}(t)$, therefore it is not directly tractable. Relaxing this constraint to $\mathbf{Z}(t) \succeq 0$, we have:

$$\check{\mathbf{Z}}(t) = \underset{\mathbf{Z}(t)}{\arg\max} \quad \big\langle \mathbf{B}(t,t-1), \mathbf{Z}(t) \big\rangle_F \tag{7a}$$

$$\text{s.t.} \quad \mathbf{Z}(t) \succeq 0 \tag{7b}$$

$$Z_{ii}(t) = 1, \quad i \in [n] \tag{7c}$$

$$\langle \mathbf{Z}(t), \mathbf{J} \rangle_F = 0. \tag{7d}$$

where:

$$\mathbf{B}(t,t-1) = \ln\big(\frac{a}{b}\big)\big(\mathbf{A}(t) - \mathbf{A}(t) \circ \mathbf{A}(t-1)\big) + \ln\big(\frac{a-\epsilon}{b+\epsilon}\big)\mathbf{A}(t-1) \tag{8}$$

where $\circ$ is the Hadamard product.

Theorem 1 (Necessary Condition)**.** *For a sequence of estimators $\hat{\mathbf{Z}}(t)$ defined via* (6)*, if*

$$\eta\big(\sqrt{a} - \sqrt{b}\big)^2 + \big(\sqrt{a-\epsilon} - \sqrt{b+\epsilon}\big)^2 < 2$$

where $\eta \in (0,1]$, $\epsilon \in [0, \frac{a-b}{2}]$, then,

$$\lim_{n\to\infty} \mathbb{P}\big(\hat{\mathbf{Z}}(t) = \mathbf{Z}^*(t)\big) = 0$$

for every t, where $\mathbf{Z}^*(t)$ is the ground truth.

Proof. See Appendix A. □

Theorem 2 (Sufficient Condition)**.** *The semidefinite programming estimator of* (7) *is asymptotically optimal for sufficiently large n, i.e., $\lim_{n\to\infty} \mathbb{P}\big(\check{\mathbf{Z}}(t) = \mathbf{Z}^*(t)\big) = 1$, if*

$$\eta\big(\sqrt{a} - \sqrt{b}\big)^2 + \big(\sqrt{a-\epsilon} - \sqrt{b+\epsilon}\big)^2 \geq 2$$

where $\eta \in (0,1]$ and $\epsilon \in [0, \frac{a-b}{2}]$.

Proof. See Appendix B. □

Remark 1. *The assortativity, $\alpha = \frac{a}{b}$, of the SBM parameters is a key parameter in determining the exact recovery detection threshold. The higher the assortativity, the better the threshold. Fig. 4 shows under Markov SBM graphs recoverable regions are larger when $\alpha = 9$ (left panel) compared with $\alpha = 4.5$ (right panel).*

Remark 2. *Higher edge persistence, characterized by smaller η, shrinks the perfect recovery region. With sufficiently high edge persistence, exact community detection will become infeasible, as illustrated in Fig. 4.*

Remark 3. *A weak bond between current labels and past observations may lead to unrecoverability of labels. See Fig. 4, where smaller values of ϵ are more favorable for perfect recovery.*

3.2 Multi-observation Model

Recall that in this model, graph observations over a window of length τ are used. For convenience, we represent τ binary graphs $\{\mathbf{A}(t), \ldots, \mathbf{A}(t-\tau)\}$ (having the same nodes) equivalently with a compact multi-valued graph $\boldsymbol{A}$ whose every edge has a binary vector value $[A_{ij}(t), \ldots, A_{ij}(t-\tau)]$.[3] A binary vector edge weight is equivalent to a positive integer $s \in [0, 2^\tau - 1]$. We then consider $2^\tau - 1$ binary graphs $\boldsymbol{M}(s), s = 1, \ldots, 2^\tau - 1$ defined as:

$$\boldsymbol{M}_{ij}(s) = \begin{cases} 1 & \boldsymbol{A}_{ij} = s \\ 0 & \text{otherwise} \end{cases}$$

The maximum a-posteriori estimator and maximum likelihood estimator are equivalent due to the uniform distribution of binary balanced-size communities. Therefore, the optimal estimator of (5) is the solution of the following optimization:

$$\hat{\mathbf{Z}}(t) = \underset{\mathbf{Z}(t)}{\arg\max} \quad \Big\langle \sum_{s\neq 0} \ln\big(\frac{a_s}{b_s}\big)\boldsymbol{M}(s), \mathbf{Z}(t) \Big\rangle_F \tag{9a}$$

$$\text{s.t.} \quad Z_{ii}(t) = 1, \quad i \in [n] \tag{9b}$$

$$\langle \mathbf{J}, \mathbf{Z}(t) \rangle_F = 0. \tag{9c}$$

where $\mathbf{Z}(t) = \boldsymbol{g}(t)[\boldsymbol{g}(t)]^\top$. Since $\mathbf{Z}(t)$ is a rank-one matrix, (9) is nonconvex optimization. Therefore, we relax this constraint, and the estimator will be as follows:

$$\check{\mathbf{Z}}(t) = \underset{\mathbf{Z}(t)}{\arg\max} \quad \Big\langle \sum_{s\neq 0} \ln\big(\frac{a_s}{b_s}\big)\boldsymbol{M}(s), \mathbf{Z}(t) \Big\rangle_F \tag{10a}$$

$$\text{s.t.} \quad \mathbf{Z}(t) \succeq 0 \tag{10b}$$

$$Z_{ii}(t) = 1, \quad i \in [n] \tag{10c}$$

$$\langle \mathbf{J}, \mathbf{Z}(t) \rangle_F = 0. \tag{10d}$$

Theorem 3. *For any sequence of estimators $\hat{\mathbf{Z}}(t)$ in* (9)*, if*

$$\sum_{s\neq 0} \big(\sqrt{a_s} - \sqrt{b_s}\big)^2 < 2,$$

then

$$\lim_{n\to\infty} \mathbb{P}\big(\hat{\mathbf{Z}}(t) = \mathbf{Z}^*(t)\big) = 0$$

for every t, where $\mathbf{Z}^*(t)$ is the ground truth.

Proof. See Appendix C. □

3. In effect, τ binary graphs have been stacked into one multi-valued graph.

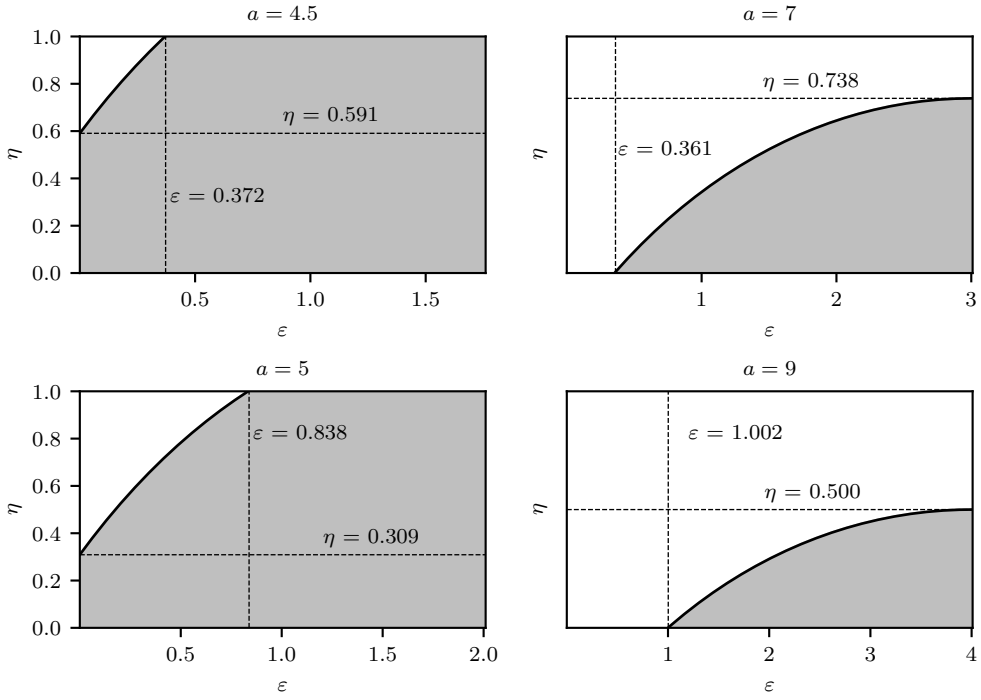


Fig. 4. Edge persistence impact on exact recovery threshold. The gray regions show where exact recovery is impossible

Theorem 4. *The semidefinite programming estimator of* (10) *is asymptotically optimal for sufficiently large* n*, i.e.,* $\lim_{n\to\infty} \mathbb{P}\big(\tilde{\mathbf{Z}}(t) = \mathbf{Z}^*(t)\big) = 1$*, when*

$$\sum_{s\neq 0} \left(\sqrt{a_s} - \sqrt{b_s}\right)^2 \geq 2$$

where $\mathbf{Z}^*(t)$ *is the ground truth*

Proof. See Appendix D. □

4 Numerical Results

This section conducts numerical simulations that provide insight into the range of situations where the earlier asymptotic findings in the paper are valid. These simulations were done utilizing Python and an academic-licensed edition of MOSEK for the execution of the SDP tasks.

4.1 Markov SBM Graphs

Graph community detection was performed on a graph with $n = 1000$ nodes, featuring two equally sized communities. The SDP algorithm was executed for a variety of parameter settings. For each setting 25 graphs were simulated and community value detected for all nodes. Using the 25,000 node samples thus simulated, we estimate the probability of error (by averaging) and also its variance. Using a Gaussian approximation, we calculated the 95% confidence intervals. Fig. 5 shows pictorially results for these simulations under a variety of graph parameters. In Fig. 5, the red circles represent graph parameters for which the zero node error probability is outside the 95% confidence interval for the calculated average error. These are parameter values under which, with high confidence, the graph is not fully informative with respect to the underlying communities. For certain parameter sets, the entire simulation did not result in any node errors, these values are represented with green. Other outcomes are shown in white.

TABLE 1
Probability of error and lower bound of 95% confidence under Markov SBM graphs

α	η	ϵ	$\mathbb{P}(Error)$	(95% CI)/2
4.5	0.1	0.5	1.88×10^{-2}	3.86×10^{-4}
4.5	0.1	1	2.79×10^{-1}	3.46×10^{-4}
4.5	0.1	1.5	4.53×10^{-1}	3.46×10^{-4}
5	0.1	1.5	3.19×10^{-1}	3.46×10^{-4}
7	0.1	1.5	9.44×10^{-3}	1.67×10^{-3}
7	0.2	1.5	2.52×10^{-3}	7.84×10^{-4}
7	0.3	1.5	1.12×10^{-3}	4.74×10^{-4}
9	0.3	2.25	2.4×10^{-4}	1.67×10^{-4}
9	0.4	2.25	4×10^{-5}	7.68×10^{-5}

It can be observed in Fig. 5 that the green circles (zero node error) align with the theoretical results in asymptotic regimes that are established in Theorems 1 and 2 and depicted in Fig. 4. Fig. 5 obviously indicates that the higher the assortativity, the bigger the green region and the smaller the red region. Results associated with some parameter configurations are also reported in Table 1. These simulation outcomes concur with the theoretical conclusions, indicating that larger values of η and smaller values of ϵ are associated with smaller misclassification errors for all scenarios.

4.2 Multi-observation Model

We considered $\tau = 1$ and five parameter sets, as outlined in Table 2. In two snapshots of a binary graph, the edge between nodes i, j can have four distinct outcomes (states). Without loss of generality, we assume a start time $t = 0$ for

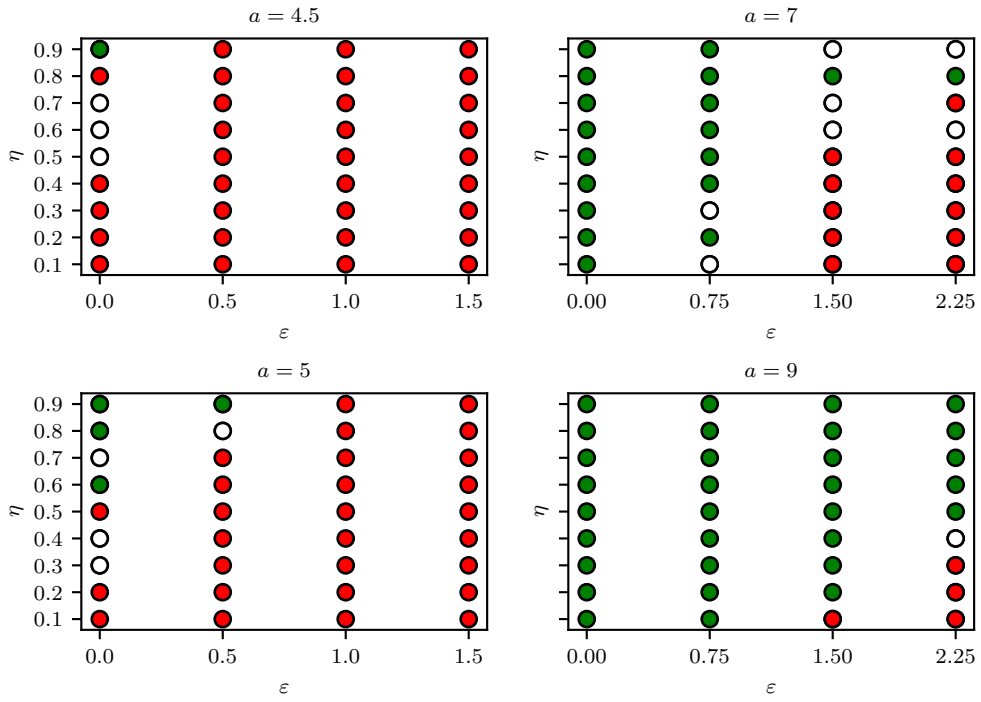


Fig. 5. Parameter configuration of the Markov SBM where exact recovery is achieved (green), not achieved with 95% probability (red), and inconclusive (white)

TABLE 2
Probability of error and lower bound of 95% confidence interval under multi-observation model

a_1	a_2	a_3	$\mathbb{P}(Error)$	(95% CI)/2
1.5	1.25	1.25	4.6808×10^{-1}	8.7784×10^{-3}
2	1.5	1.5	2.6168×10^{-1}	3.2267×10^{-2}
3	2	2	8.92×10^{-3}	4.2709×10^{-3}
4	2.5	2.5	2×10^{-4}	1.9203×10^{-4}
4.5	3	3	8×10^{-5}	1.0634×10^{-4}

our simulation, therefore our observations consist of graphs $A(0), A(1)$. For the edge between each pair of nodes i, j our states our defined as:

$$s = \begin{cases} 0 & A_{ij}(1) = A_{ij}(2) = 0 \\ 1 & A_{ij}(1) = A_{ij}(2) = 1 \\ 2 & A_{ij}(1) = 0,\ A_{ij}(2) = 1 \\ 3 & A_{ij}(1) = 1,\ A_{ij}(2) = 0 \end{cases}$$

Consider a_s denotes parameters associated with inter-community edges, while b_s represents parameters related to intra-community edges. Without loss of generality, we assume $a_s \geq b_s$ for each $s \neq 0$, $b_1 = b_2 = b_3 = 1$, and $a_1 \geq a_2 = a_3 \geq 0$ in all experiments. Note that, $a_0 = 1 - \sum_{s \neq 0} a_s$ and $b_0 = 1 - \sum_{s \neq 0} b_s$. Similar to Markov SBM graphs, 25 graphs were simulated and community value was detected for all nodes under each setting. Simulation results reveal that the probability of error decreases with an increasing value of $\alpha_s = \frac{a_s}{b_s}$ for $s = 1, 2, 3$. This reduction in error is particularly significant when the system's parameters satisfy the conditions outlined in Theorems 3 and 4, as demonstrated by the last set of parameters in Table 2.

5 Conclusion

This paper explores the influence of edge evolution on dynamic community detection under two dynamical graph models. We considered an SBM with Markovian edge dependency for temporal snapshots, in which our approach employed semi-definite relaxation of the maximum likelihood problem. We showed that the achievability and the converse for this problem are asymptotically tight. A second set of problems are also considered, where graph edges evolve faster than labels. Inspired by the well-known block fading model from wireless communication, our modeling of this problem approximates the evolution of the graph with node labels that remain stationary (constant) within a time interval, and change independently across these intervals. Throughout, the graph edges obey an SBM model at each time instance. Our analysis of this problem is enabled by the construction of an equivalent weighted adjacency matrix in each block. We introduce a semidefinite programming algorithm and show that its operation with different graph parameters is consistent with the phase transition for the optimal estimator. In both problems, we demonstrate that greater relative assortativity leads to a bigger recovery region. Numerical results verify the asymptotic theoretical predictions in the first part of the paper and illustrate the impact of different parameter settings in the second part of the paper.

Appendix A
Proof of Theorem 1

Lemma 1. *Recall* $\mathbf{B}(t, t-1)$ *from* (8) *and let* $\boldsymbol{g}^*(t)$ *be the ground truth,* $\mathcal{A} = \{i \in [n] : g_i^*(t) = 1\}$, *and* $\mathcal{B} =$

$\{j \in [n] : g_j^*(t) = -1\}$. *Define:*

$$\mathcal{F} \triangleq \{\hat{\mathbf{Z}}(t) \neq \mathbf{Z}^*(t)\}$$
$$\mathcal{F}_{\mathcal{A}} \triangleq \Big\{\exists i \in \mathcal{A} : -[\boldsymbol{g}^*(t)]^\top \mathbf{B}_i(t,t-1) \geq \ln\big(\frac{a}{b}\big)\Big\}$$
$$\mathcal{F}_{\mathcal{B}} \triangleq \Big\{\exists j \in \mathcal{B} : [\boldsymbol{g}^*(t)]^\top \mathbf{B}_j(t,t-1) \geq \ln\big(\frac{a}{b}\big)\Big\}$$

where $\mathbf{B}_i(t,t-1)$ is the i^{th} column of $\mathbf{B}(t,t-1)$. Then, $\mathcal{F}_{\mathcal{A}} \cap \mathcal{F}_{\mathcal{B}} \to \mathcal{F}$, where $\mathcal{F}$ is the event that maximum likelihood estimator fails.

Proof. Define $\bar{\boldsymbol{g}}(t)$ as a label vector that disagrees with the ground truth only in nodes, i, j, i.e.,

$$\bar{g}_\ell(t) = \begin{cases} -g_\ell^*(t) & \text{if } \ell = i, j \\ g_\ell^*(t) & \text{otherwise} \end{cases}$$

We explore conditions in which $\bar{\mathbf{g}}(t)$ would have a better likelihood metric than $\boldsymbol{g}^*(t)$.

$$\big[\bar{\boldsymbol{g}}(t) - \boldsymbol{g}^*(t)\big]^\top \mathbf{B}(t,t-1)\big[\bar{\boldsymbol{g}}(t) - \boldsymbol{g}^*(t)\big] = -8B_{ij}(t,t-1) + 4\boldsymbol{g}^{*\top}(t)\big(\mathbf{B}_j(t,t-1) - \mathbf{B}_i(t,t-1)\big) \tag{11}$$

Since $\max B_{ij}(t,t-1) = \ln\big(\frac{a}{b}\big)$, the following is then the sufficient condition for the failure of maximum likelihood:

$$\boldsymbol{g}^{*\top}(t)\big(\mathbf{B}_j(t,t-1) - \mathbf{B}_i(t,t-1)\big) \geq 2\ln\big(\frac{a}{b}\big)$$

Equivalently, when simultaneously:

$$\begin{cases} -[\boldsymbol{g}^*(t)]^\top \mathbf{B}_i(t,t-1) \geq \ln\big(\frac{a}{b}\big), \\ [\boldsymbol{g}^*(t)]^\top \mathbf{B}_j(t,t-1) \geq \ln\big(\frac{a}{b}\big) \end{cases}$$

then $\mathcal{F}$ occurs. This completes the proof. □

Definition 1. *Let $\mathcal{H} \cup \bar{\mathcal{H}} = \{i : g_i^*(t) = 1\}$ provided that $\mathcal{H} \cap \bar{\mathcal{H}} = \phi$. Also, define $f(n) \triangleq \ln\big(\frac{a}{b}\big) + \frac{\ln(n)}{\ln(\ln(n))}$. For each node $i \in \{i : g_i^*(t) = 1\}$ define the following events:*

$$\Delta_i^{\mathcal{H}} \triangleq \Big\{\sum_{k\in\mathcal{H}} B_{ik}(t,t-1) \leq \frac{\ln(n)}{\ln(\ln(n))}\Big\}$$
$$F_i^{\mathcal{H}} \triangleq \Big\{\sum_{k\notin\mathcal{H}\cup\bar{\mathcal{H}}} B_{ik}(t,t-1) - \sum_{k\in\bar{\mathcal{H}}} B_{ik}(t,t-1) \leq f(n)\Big\}$$
$$F_i \triangleq \Big\{\sum_{k\notin\mathcal{H}\cup\bar{\mathcal{H}}} B_{ik}(t,t-1) - \sum_{k\in\mathcal{H}\cup\bar{\mathcal{H}}} B_{ik}(t,t-1) \leq f(n)\Big\}$$

where superscript $\mathcal{H}$ indicates $i \in \mathcal{H} \subset \{i : g_i^*(t) = 1\}$. In addition, define $\Delta^{\mathcal{H}} = \cap_{i\in\mathcal{H}}\Delta_i^{\mathcal{H}}$ and $F^{\mathcal{H}} = \cup_{i\in\mathcal{H}}F_i^{\mathcal{H}}$.

Lemma 2. *If $\mathbb{P}(F^{\mathcal{H}}) = 1-\delta$ and $\mathbb{P}(\Delta^{\mathcal{H}}) = 1-\delta$ for $\delta < 0.25$, then there exists a positive δ' such that $\mathbb{P}(\mathcal{F}) \geq \delta'$.*

Proof. We know $\Delta^{\mathcal{H}} \cap F^{\mathcal{H}} = \mathcal{F}_{\mathcal{A}}$. Therefore,

$$\mathbb{P}(\mathcal{F}_{\mathcal{A}}) \geq \mathbb{P}(F^{\mathcal{H}}) + \mathbb{P}(\Delta^{\mathcal{H}}) - 1 \geq 1 - 2\delta$$

Similarly, we have $\mathbb{P}(\mathcal{F}_{\mathcal{B}}) \geq 1-2\delta$. Due to Lemma 1, maximum likelihood fails if $\mathcal{F}_{\mathcal{A}}$ and $\mathcal{F}_{\mathcal{B}}$ occur simultaneously, i.e. $\mathcal{F}_{\mathcal{A}} \cap \mathcal{F}_{\mathcal{B}} \to \mathcal{F}$. Hence, the probability of error will be bounded away from zero for $\delta < 0.25$ since:

$$\mathbb{P}(\mathcal{F}) \geq \mathbb{P}(\mathcal{F}_{\mathcal{A}}) + \mathbb{P}(\mathcal{F}_{\mathcal{B}}) - 1 \geq 1 - 4\delta$$

□

The following Lemma states under what conditions Δ happens with high probability asymptotically.

Lemma 3. *Let $|\mathcal{H}| = \frac{n}{\ln^2(n)}$ then $\lim_{n\to\infty}\mathbb{P}(\Delta) = 1$.*

Proof. If $\bar{\Delta}_i^{\mathcal{H}}$ be the complement of $\Delta_i^{\mathcal{H}}$, then:

$$\begin{aligned}\mathbb{P}(\bar{\Delta}_i^{\mathcal{H}}) &= \mathbb{P}\Big(\sum_{\substack{k\in\mathcal{H}\\k\neq i}} B_{ik}(t,t-1) \geq \frac{\ln(n)}{\ln(\ln(n))}\Big) \\ &\leq \mathbb{P}\Big(\sum_{k\in\mathcal{H}} B_{ik}(t,t-1) \geq \frac{\ln(n)}{\ln(\ln(n))}\Big) \\ &\leq e^{-\frac{\ln(n)}{\ln(\ln(n))}(1-o(1))}\end{aligned}$$

where the last inequality holds due to applying the Chernoff bound and using this fact that $\{\forall x \in \mathbb{R} : \quad e^x > 1+x\}$. The union bound results in:

$$\begin{aligned}\mathbb{P}(\Delta^{\mathcal{H}}) &= 1 - \mathbb{P}(\bar{\Delta}^{\mathcal{H}}) \\ &\geq 1 - \frac{n}{\ln^2(n)}e^{-\ln(n)+\ln(\ln(n))} \\ &= 1 - e^{\ln(\ln(n))-\ln(\ln^2(n))}\end{aligned} \tag{12}$$

□

The following Lemma provides conditions under which $\mathbb{P}(F_i^{\mathcal{H}}) \geq 1-\delta$ asymptotically.

Lemma 4. *For any $\delta \in (0,1)$, if $\mathbb{P}(F_i^{\mathcal{H}}) \geq \frac{\ln^2(n)}{n}\ln(\frac{1}{\delta})$, then*

$$\lim_{n\to\infty}\mathbb{P}(F^{\mathcal{H}}) \geq 1-\delta$$

Proof. Since $F_i^{\mathcal{H}}$ are i.i.d.:

$$\begin{aligned}\mathbb{P}(F^{\mathcal{H}}) &= \mathbb{P}(\cup_{i\in\mathcal{H}}F_i^{\mathcal{H}}) = 1 - \mathbb{P}(\cap_{i\in\mathcal{H}}(F_i^{\mathcal{H}})^c) \\ &= 1 - \big[\big(1-\mathbb{P}(F_i^{\mathcal{H}})\big)^{\frac{1}{\mathbb{P}(F_i^{\mathcal{H}})}}\big]^{\left(\frac{n\mathbb{P}(F_i^{\mathcal{H}})}{\ln^2(n)}\right)} \\ &\geq 1 - \big[\big(1-\mathbb{P}(F_i^{\mathcal{H}})\big)^{\frac{1}{\mathbb{P}(F_i^{\mathcal{H}})}}\big]^{-\ln\delta}\end{aligned} \tag{13}$$

The result follows because for all $\mathbb{P}(F_i^{\mathcal{H}}) < 1$, we have $\delta > \big(1-\mathbb{P}(F_i^{\mathcal{H}})\big)^{\frac{-\ln\delta}{\mathbb{P}(F_i^{\mathcal{H}})}}$. □

Based on Defintion 1, $\mathbb{P}(F_i^{\mathcal{H}}) \geq \mathbb{P}(F_i)$. Let:

$$\begin{aligned}\Gamma(u) =& \frac{1}{2}\Big(\tilde{a}\big(\frac{\tilde{a}}{\tilde{b}}\big)^{-u} + \tilde{b}\big(\frac{\tilde{a}}{\tilde{b}}\big)^{u} + \eta a\big(\frac{a}{b}\big)^{-u} + \eta b\big(\frac{a}{b}\big)^{u}\Big) - \\ & \frac{1}{2}(\tilde{a}+\tilde{b}+\eta(a+b)) \\ =& \frac{1}{2}\Big(\tilde{a}\big(\frac{\tilde{a}}{\tilde{b}}\big)^{-u} + \tilde{b}\big(\frac{\tilde{a}}{\tilde{b}}\big)^{u} + \eta a\big(\frac{a}{b}\big)^{-u} + \eta b\big(\frac{a}{b}\big)^{u}\Big) - \\ & \frac{1+\eta}{2}(a+b)\end{aligned} \tag{14}$$

where the last term derived from $\tilde{a}+\tilde{b} = a+b$. Then:

$$\mathbb{P}(F_i) \leq \inf_{u>0} e^{-\ln(n)\left(\frac{uf(n)}{\ln(n)} - \frac{1}{2}\Gamma(u)\right)} \tag{15}$$

where the inequality holds due to applying the Chernoff bound on i.i.d. random variables, and because $e^x > 1+x$. Finding infimum of (15) is equivalent to minimizing $\bar{\Gamma}(u,n) = \frac{1}{2}\Gamma(u) - \frac{uf(n)}{\ln(n)}$, where $\Gamma(u)$ is convex with respect to u. If $\frac{uf(n)}{\ln(n)} = o(1)$, then for sufficiently large n, $u^* = \frac{1}{2}$, and this leads to:

$$\mathbb{P}(F_i) \leq e^{-\frac{\ln(n)}{2}\mu + o(1)} \tag{16}$$

where:

$$\mu = \eta\big(\sqrt{a}-\sqrt{b}\big)^2 + a + b - 2\sqrt{\tilde{a}\tilde{b}}$$
$$= \eta\big(\sqrt{a}-\sqrt{b}\big)^2 + \big(\sqrt{a-\epsilon}-\sqrt{b+\epsilon}\big)^2$$

Finally, the proof of Theorem 1 follows by combining (16) and Lemmas 1, 2, 3, and 4.

APPENDIX B
PROOF OF THEOREM 2

Lemma 5. *Consider the scalar σ^*, the diagonal matrix $\mathbf{D}^*$ and :*

$$\mathbf{R}^* = \mathbf{D}^* + \sigma^*\mathbf{J} - \mathbf{B}(t,t-1) \tag{17}$$

If $\mathbf{R}^* \succeq 0$, $\lambda_2(\mathbf{R}^*) \geq 0$ and $\mathbf{R}^*\boldsymbol{g}^*(t) = \mathbf{0}$, then $(\sigma^*, \mathbf{D}^*, \mathbf{R}^*)$ is the dual optimal solution of (7), whose unique primal solution is $\tilde{\mathbf{Z}}(t) = \boldsymbol{g}^*(t)[\boldsymbol{g}^*(t)]^\top$.

Proof. The Lagrangian of Eq. (7) is :

$$L(\mathbf{Z}(t),\mathbf{R},\mathbf{D},\sigma) = \langle \mathbf{B}(t,t-1),\mathbf{Z}(t)\rangle_F + \langle \mathbf{R},\mathbf{Z}(t)\rangle_F - \langle \mathbf{D},\mathbf{Z}(t)-\mathbf{I}\rangle_F - \sigma\langle \mathbf{J},\mathbf{Z}(t)\rangle_F,$$

where $\mathbf{R} \succeq 0$, $\mathbf{D} = \mathrm{diag}(d_i)$ and σ are Lagrange multipliers. For $\mathbf{Z}(t)$ that satisfies the constraints in (7):

$$\begin{aligned}
\langle \mathbf{B}(t,t-1),\mathbf{Z}(t)\rangle_F &\leq L(\mathbf{Z}(t),\mathbf{R}^*,\mathbf{D}^*,\sigma^*) && \text{(18a)}\\
&= \langle \mathbf{D}^*,\mathbf{I}\rangle_F && \text{(18b)}\\
&= \langle \mathbf{D}^*,\mathbf{Z}^*(t)\rangle_F && \text{(18c)}\\
&= \langle \mathbf{B}(t,t-1),\mathbf{Z}^*(t)\rangle_F + \langle \mathbf{R}^*-\sigma^*\mathbf{J},\mathbf{Z}^*(t)\rangle_F && \text{(18d)}\\
&= \langle \mathbf{B}(t,t-1),\mathbf{Z}^*(t)\rangle_F && \text{(18e)}
\end{aligned}$$

where (18a) holds since $\langle \mathbf{R}^*,\mathbf{Z}(t)\rangle \geq 0$, (18b) because of (17) and optimization constraint $\langle \mathbf{J},\mathbf{Z}(t)\rangle_F = 0$ in (7), (18c) due to $Z^*_{ii}(t) = 1$ for all $i \in [n]$, (18d) owing to replacing $\mathbf{D}^*$ from (17), and (18e) because $\langle \mathbf{J},\mathbf{Z}^*(t)\rangle_F = 0$, $\langle \mathbf{R}^*,\mathbf{Z}^*(t)\rangle_F = [\boldsymbol{g}^*(t)]^\top\mathbf{R}^*\boldsymbol{g}^*(t)$, and $\mathbf{R}^*\boldsymbol{g}^*(t) = \mathbf{0}$ based on Lemma 5.

Now, we need to show the uniqueness of the optimal solution to (7). Assume the contrary, i.e., that $\tilde{\mathbf{Z}}(t) = [\tilde{\boldsymbol{g}}(t)]^\top\tilde{\boldsymbol{g}}(t)$ is another optimal solution to (7), thus $\tilde{\mathbf{Z}}(t) \succeq 0$ and:

$$\langle \mathbf{B}(t,t-1),\mathbf{Z}^*(t)\rangle_F = \langle \mathbf{B}(t,t-1),\tilde{\mathbf{Z}}(t)\rangle_F,$$
$$\langle \mathbf{J},\mathbf{Z}^*(t)\rangle_F = \langle \mathbf{J},\tilde{\mathbf{Z}}(t)\rangle_F,$$
$$Z^*_{ii}(t) = \tilde{Z}_{ii}(t) = 1 \quad \forall i \in [n].$$

This results in:

$$\begin{aligned}
\langle \mathbf{R}^*,\tilde{\mathbf{Z}}(t)\rangle_F &= \langle \mathbf{D}^* + \sigma^*\mathbf{J} - \mathbf{B}(t,t-1),\tilde{\mathbf{Z}}(t)\rangle_F && \text{(19a)}\\
&= \langle \mathbf{D}^* + \sigma^*\mathbf{J} - \mathbf{B}(t,t-1),\mathbf{Z}^*(t)\rangle_F && \text{(19b)}\\
&= \langle \mathbf{R}^*,\mathbf{Z}^*(t)\rangle_F && \text{(19c)}\\
&= 0 && \text{(19d)}
\end{aligned}$$

where (19d) holds because $\mathbf{R}^*\boldsymbol{g}^*(t) = \mathbf{0}$. Therefore, we have $[\tilde{\boldsymbol{g}}(t)]^\top\mathbf{R}^*\tilde{\boldsymbol{g}}(t) = 0$. Since $\tilde{\boldsymbol{g}}(t)$ is a nonzero vector whose elements are -1 and 1, and because of balanced sized communities, there are two possible scenarios for $\mathbf{R}^*\tilde{\boldsymbol{g}}(t)$ as follows:

- $\mathbf{R}^*\tilde{\boldsymbol{g}}(t) = \mathbf{1}$. This contradicts with optimality of $\tilde{\mathbf{Z}}(t)$.
- $\mathbf{R}^*\tilde{\boldsymbol{g}}(t) = \mathbf{0} \implies \mathbf{R}^*\big(\boldsymbol{g}^*(t) - \tilde{\boldsymbol{g}}(t)\big) = \mathbf{0}$. Because $\mathbf{R}^* \succeq 0$, and its second smallest eigenvalue $\lambda_2(\mathbf{R}^*)$ is positive, $\tilde{\boldsymbol{g}}(t)$ is a multiple of $\boldsymbol{g}^*(t)$. This means $\tilde{\mathbf{Z}}(t)$ is a multiple of $\mathbf{Z}^*(t)$, accordingly $\tilde{\mathbf{Z}}(t) = \mathbf{Z}^*(t)$ due to $Z^*_{ii}(t) = \tilde{Z}_{ii}(t) = 1$ for all $i \in [n]$.

□

We now proceed to show the conditions of Lemma 5 can be satisfied. Let:

$$D^*_{ii} = \sum_{j=1}^{n} B_{ij}(t,t-1)g^*_j(t)\,g^*_i(t) \tag{20}$$

It follows that $\mathbf{D}^*\boldsymbol{g}^*(t) = 0$ and $\mathbf{R}^*\boldsymbol{g}^*(t) = \mathbf{0}$. It remains to show that $\mathbf{R}^* \succeq 0$ and $\lambda_2(\mathbf{R}^*) > 0$ with high probability. This means:

$$\mathbb{P}\Big(\inf_{\mathbf{v}\perp\boldsymbol{g}^*(t),\|\mathbf{v}\|=1}\mathbf{v}^\top\mathbf{R}^*\mathbf{v} > 0\Big) \geq 1 - o(1), \tag{21}$$

where $\mathbf{v} \in \mathbb{R}^n$. Define $T \triangleq \ln\left(\frac{a}{b}\right)$, $T_\epsilon \triangleq \ln\left(\frac{a-\epsilon}{b+\epsilon}\right)$. Under the conditions $\mathbf{v} \perp \boldsymbol{g}^*(t)$ and $\|\mathbf{v}\| = 1$ we have:

$$\begin{aligned}
\mathbf{v}^\top\mathbf{R}^*\mathbf{v} &= \langle \mathbf{v}\mathbf{v}^\top,\mathbf{D}^* + \sigma^*\mathbf{J} - \mathbf{B}(t,t-1)\rangle_F\\
&= \langle \mathbf{v}\mathbf{v}^\top,\mathbf{D}^* + \sigma^*\mathbf{J} - \mathbb{E}[\mathbf{B}(t,t-1)]\rangle_F -\\
&\quad \langle \mathbf{v}\mathbf{v}^\top,\mathbf{B}(t,t-1) - \mathbb{E}[\mathbf{B}(t,t-1)]\rangle_F\\
&\geq \min_{i\in[n]} d^*_i - T\|\mathbf{B}(t,t-1) - \mathbb{E}[\mathbf{B}(t,t-1)]\| -\\
&\quad \langle \mathbf{v}\mathbf{v}^\top,\mathbb{E}[\mathbf{B}(t,t-1)] - \sigma^*\mathbf{J}\rangle_F
\end{aligned} \tag{22}$$

where the last inequality holds due to:

$$\langle \mathbf{v}\mathbf{v}^\top,\mathbf{D}^*\rangle_F \geq \min_{i\in[n]} D^*_{ii},$$
$$\langle \mathbf{v}\mathbf{v}^\top,\mathbf{W}(t,t-1)\rangle_F \leq \|\mathbf{W}(t,t-1)\| \tag{23}$$

where $\mathbf{W}(t,t-1) = \mathbf{B}(t,t-1) - \mathbb{E}[\mathbf{B}(t,t-1)]$. We also have:

$$\begin{aligned}
\frac{n}{\ln(n)}\langle \mathbf{v}\mathbf{v}^\top,\mathbb{E}[\mathbf{B}(t,t-1)]\rangle_F &= \frac{(a+b)(\eta T + T_\epsilon)}{2}\langle \mathbf{v}\mathbf{v}^\top,\mathbf{J}\rangle_F\\
&\quad + \frac{(\eta T + T_\epsilon)(a-b) - 2\epsilon T_\epsilon}{2}\langle \mathbf{v}\mathbf{v}^\top,\mathbf{Z}(t)\rangle_F\\
&\quad - \big(a(\eta T + T_\epsilon) - \epsilon T_\epsilon\big)\mathbf{I}
\end{aligned} \tag{24}$$

Lemma 6. *[27, Theorem 9] For any $c > 0$, there exist $c' > 0$, $c'' > 0$ and $c''' > 0$ such that for any $n \geq 1$:*

$$\mathbb{P}\big(\|\mathbf{A}(t) - \mathbb{E}[\mathbf{A}(t)]\|_2 \leq c'\sqrt{np}\big) \geq 1 - n^{-c},$$
$$\mathbb{P}\big(\|\mathbf{A}(t-1) - \mathbb{E}[\mathbf{A}(t-1)]\|_2 \leq c''\sqrt{np}\big) \geq 1 - n^{-c},$$
$$\mathbb{P}\big(\|\bar{\mathbf{A}}(t,t-1) - \mathbb{E}[\bar{\mathbf{A}}(t,t-1)]\|_2 \leq c'''\sqrt{np}\big) \geq 1 - n^{-c}$$

where $\bar{\mathbf{A}}(t,t-1) = \mathbf{A}(t) \circ \mathbf{A}(t-1)$.

Lemma 7. *For sufficiently large n,*

$$\mathbb{P}\big(\langle \mathbf{v}\mathbf{v}^\top,\mathbf{Z}(t)\rangle_F \leq \sqrt{\ln(n)}\big) \leq 1 - o(1)$$

Proof. This follows from applying the Chernoff bound. □

For convenience define:

$$\sigma^* \triangleq \frac{(a+b)(T+\eta T_\epsilon)}{2}$$
$$\bar{T}_{\eta,\epsilon} \triangleq a(\eta T + T_\epsilon) + 2\epsilon T_\epsilon$$

Then, applying Lemmas 6 and 7 on (22) results in:

$$\mathbf{v}^\top \mathbf{R}^* \mathbf{v} \geq \min_{i\in[n]} D^*_{ii} - (T(c'+c'')+T_\epsilon c''')\sqrt{\ln(n)} + \bar{T}_{\eta,\epsilon} \quad (25)$$

with a probability converging to one. Now, we need to show the right hand of (25) is positive. This is accomplished by bounding $\min_{i\in[n]} D^*_{ii}$.

$$H = \sum_{j=1}^{n} B_{ij}(t,t-1) g^*_i(t)$$

Then (14) and $1-x \leq e^x$ for all $x \geq 0$ gives $\mathbb{E}(e^{-uH}) \leq e^{\Gamma(u)}$. Since $\Gamma(u)$ is convex w.r.t u, then for sufficiently large n define $\delta \triangleq \frac{\ln(n)}{\ln(\ln(n))}$. Therefore:

$$\begin{aligned}\mathbb{P}(D^*_{ii} \leq \delta) &= \mathbb{P}\big(H \leq \delta\big) \\ &\leq e^{-\frac{\ln(n)}{2}\mu + o(1)} \end{aligned} \quad (26)$$

where

$$\mu = \eta\big(\sqrt{a}-\sqrt{b}\big)^2 + a + b - 2\sqrt{(a-\epsilon)(b+\epsilon)}$$

If $\mu \geq 2$, then $\mathbb{P}\big(H \geq \delta\big) = 1 - o(1)$. This completes the proof of Theorem 2.

APPENDIX C
PROOF OF THEOREM 3

Lemma 8. *Consider* (9) *and let* $\boldsymbol{g}^*(t)$ *be the ground truth,* $\mathcal{A} = \{i \in [n] : \mathrm{g}^*_i(t) = 1\}$, *and* $\mathcal{B} = \{j \in [n] : \mathrm{g}^*_j(t) = -1\}$. *Define:*

$$\mathcal{F} \triangleq \{\hat{\mathbf{Z}}(t) \neq \mathbf{Z}^*(t)\}$$
$$\mathcal{F}_{\mathcal{A}} \triangleq \Big\{\exists i \in \mathcal{A} : -[\boldsymbol{g}^*(t)]^\top \sum_{s\neq 0} \ln\big(\frac{a_s}{b_s}\big)\boldsymbol{M}_i(s) \geq \tilde{T}\Big\}$$
$$\mathcal{F}_{\mathcal{B}} \triangleq \Big\{\exists j \in \mathcal{B} : [\boldsymbol{g}^*(t)]^\top \sum_{s\neq 0} \ln\big(\frac{a_s}{b_s}\big)\boldsymbol{M}_j(s) \geq \tilde{T}\Big\}$$

where $\tilde{T} = \max_s \big\{\ln\big(\frac{a_s}{b_s}\big)\big\}$ *and* $\boldsymbol{M}_i(s)$ *is the* i^{th} *column of* $\boldsymbol{M}(s)$. *Then,* $\mathcal{F}_{\mathcal{A}} \cap \mathcal{F}_{\mathcal{B}} \to \mathcal{F}$.

Proof. Similar to Lemma 1 □

Definition 2. *Let* $\mathcal{Q} \cup \bar{\mathcal{Q}} = \{i : g^*_i(t) = 1\}$ *provided that* $\mathcal{Q} \cap \bar{\mathcal{Q}} = \phi$. *In addition, let* $f(n) = \tilde{T} + \frac{\ln(n)}{\ln(\ln(n))}$, $T_s = \ln\big(\frac{a_s}{b_s}\big)$, *and define the following events for each node* $j \in \mathcal{Q}$ *:*

$$\Delta^{\mathcal{Q}}_j \triangleq \Big\{\sum_{k\in\mathcal{Q}}\sum_{s\neq 0} T_s M_{ik}(s) \leq \frac{\ln(n)}{\ln(\ln(n))}\Big\}$$
$$F^{\mathcal{Q}}_j \triangleq \Big\{\sum_{s\neq 0} T_s\Big[\sum_{k\notin\mathcal{Q}\cup\bar{\mathcal{Q}}} M_{ik}(s) - \sum_{k\in\bar{\mathcal{Q}}} M_{ik}(s)\Big] \geq f(n)\Big\}$$

In addition, define $\Delta^{\mathcal{Q}} \triangleq \cap_{j\in\mathcal{Q}}\Delta^{\mathcal{Q}}_j$ *and* $F^{\mathcal{Q}} \triangleq \cup_{j\in\mathcal{Q}} F^{\mathcal{Q}}_j$.

Lemma 9. *If* $\mathbb{P}(F^{\mathcal{Q}}) = 1-\delta$ *and* $\mathbb{P}(\Delta^{\mathcal{Q}}) = 1-\delta$ *for* $\delta < 0.25$, *then there exists a positive* δ' *such that* $\mathbb{P}(\mathcal{F}) \geq \delta'$.

Proof. Similar to Lemma 2. □

Lemma 10. *Let* $|\mathcal{Q}| = \frac{n}{\ln^2(n)}$ *then* $\lim_{n\to\infty}\mathbb{P}(\Delta^{\mathcal{Q}}) = 1$.

Proof. Similar to Lemma 3. □

Lemma 11. *For any* $\delta \in (0,1)$, *if* $\mathbb{P}(F^{\mathcal{Q}}_j) \geq \frac{\ln^2(n)}{n}\ln(\frac{1}{\delta})$, *then*

$$\lim_{n\to\infty}\mathbb{P}(F^{\mathcal{Q}}) \geq 1-\delta$$

Proof. Similar to Lemma 4, omitted for brevity. □

Define:

$$\bar{M}_i(s) \triangleq \sum_{k\notin\mathcal{Q}\cup\bar{\mathcal{Q}}} M_{ik}(s) - \sum_{k\in\mathcal{Q}\cup\bar{\mathcal{Q}}} M_{ik}(s)$$

To calculate $\mathbb{P}(F^{\mathcal{Q}}_j)$, we have:

$$\mathbb{P}(F^{\mathcal{Q}}_j) \geq \mathbb{P}\Big(\sum_{s\neq 0} T_s \bar{M}_i(s) \geq f(n)\Big) \quad (27)$$

Define:

$$\Gamma(u) \triangleq \frac{\ln(n)}{2}\sum_{s\neq 0}\Big[a_s\big(\frac{a_s}{b_s}\big)^{-u} + b_s\big(\frac{a_s}{b_s}\big)^{u} - a_s - b_s\Big] \quad (28)$$

Then:

$$\mathbb{P}\Big(\sum_{s\neq 0}\ln\big(\frac{a_s}{b_s}\big)\bar{M}_i(s) \geq f(n)\Big) \leq \inf_{u>0} e^{-\ln(n)\left(\frac{uf(n)}{\ln(n)} - \frac{1}{2}\Gamma(u)\right)} \quad (29)$$

due to applying the Chernoff bound on *i.i.d.* random variables, and $e^x > 1+x$. Finding infimum of (29) is equivalent to minimizing $\bar{\Gamma}(u,n) = \frac{1}{2}\Gamma(u) - \frac{uf(n)}{\ln(n)}$ where $\Gamma(u)$ is convex with respect to u. If $\frac{uf(n)}{\ln(n)} = o(1)$, then for sufficiently large n, $u^* = \frac{1}{2}$, and this leads to:

$$\mathbb{P}\Big(\sum_{s\neq 0}\ln\big(\frac{a_s}{b_s}\big)\bar{M}_i(s) \geq f(n)\Big) \leq e^{-\frac{\ln(n)}{2}\nu + o(1)} \quad (30)$$

where:

$$\nu = \sum_{s\neq 0}\big(\sqrt{a_s}-\sqrt{b_s}\big)^2$$

This completes the proof of Theorem 3.

APPENDIX D
PROOF OF THEOREM 4

Lemma 12. *Assume the scalar* σ^*, *the diagonal matrix* $\mathbf{D}^*$ *and:*

$$\mathbf{R}^* = \mathbf{D}^* + \sigma^*\mathbf{I} - \sum_{s\neq 0}\ln\big(\frac{a_s}{b_s}\big)\boldsymbol{M}(s)$$

such that $\mathbf{R}^* \succeq 0$, $\lambda_2(\mathbf{R}^*) \geq 0$ *and* $\mathbf{R}^*\boldsymbol{g}^*(t) = \mathbf{0}$. *Then* $(\sigma^*, \mathbf{D}^*, \mathbf{R}^*)$ *is the dual optimal solution of* (10), *whose unique primal solution is* $\check{\mathbf{Z}}(t) = \boldsymbol{g}^*(t)[\boldsymbol{g}^*(t)]^\top$.

Proof. The Lagrangian is :

$$\begin{aligned}L(\mathbf{Z}(t),\mathbf{R},\mathbf{D},\sigma) =& \langle\sum_{s\neq 0}\ln\big(\frac{a_s}{b_s}\big)\boldsymbol{M}(s), \mathbf{Z}(t)\rangle_F + \langle\mathbf{R},\mathbf{Z}(t)\rangle_F \\ &- \langle\mathbf{D},\mathbf{Z}(t)-\mathbf{I}\rangle_F - \sigma\langle\mathbf{J},\mathbf{Z}(t)\rangle_F,\end{aligned}$$

where $\mathbf{R} \succeq 0$, $\mathbf{D} = \mathrm{diag}(d_i)$ and σ are Lagrange multipliers. For $\mathbf{Z}(t)$ that satisfies optimization in (10):

$$\begin{aligned}\Big\langle\sum_{s\neq 0}\ln\big(\frac{a_s}{b_s}\big)\boldsymbol{M}(s), \mathbf{Z}(t)\Big\rangle_F &\leq L(\mathbf{Z}(t),\mathbf{R}^*,\mathbf{D}^*,\sigma^*) \\ &= \Big\langle\sum_{s\neq 0}\ln\big(\frac{a_s}{b_s}\big)\boldsymbol{M}(s), \mathbf{Z}^*(t)\Big\rangle_F\end{aligned}$$

where the last equality holds due to definitions and constraints in Lemma 12. Similar to the proof of Lemma 5,

we can show that if $\tilde{\mathbf{Z}}(t)$ is an optimal solution, then $\tilde{\mathbf{Z}}(t) = \mathbf{Z}^*(t)$, thus the solution is unique. □

We now have to prove that $\mathbf{R}^*$ in Lemma 12 satisfies the other constraints with high probability, $1 - o(1)$. To do so, let:

$$D^*_{ii} = \sum_{j=1}^{n} \Big[\sum_{s\neq 0} \ln\big(\frac{a_s}{b_s}\big) M_{ij}(s) \Big] g^*_j(t)\, g^*_i(t) \tag{31}$$

Therefore $\mathbf{D}^* \boldsymbol{g}^*(t) = 0$ and matirx $\mathbf{R}^*$ satisfies the condition $\mathbf{R}^* \boldsymbol{g}^*(t) = \mathbf{0}$. To complete the proof, we need to show $\mathbf{R}^* \succeq 0$ and $\lambda_2(\mathbf{R}^*) > 0$ with high probability. Let $T_s = \ln\big(\frac{a_s}{b_s}\big)$. This results in:

$$\begin{aligned} \mathbf{v}^\top \mathbf{R}^* \mathbf{v} &= \mathbf{v}^\top \Big(\mathbf{D}^* + \sigma^* - \sum_{s\neq 0} \ln\big(\frac{a_s}{b_s}\big) \boldsymbol{M}(s) \Big) \mathbf{v} \\ &\geq \min_{i\in[n]} D^*_{ii} - \sum_{s\neq 0} \Big(\|\boldsymbol{M}(s) - \mathbb{E}[\boldsymbol{M}(s)]\| - T_s a_s \Big) \end{aligned} \tag{32}$$

where (32) holds due to $\sigma^* = \frac{\sum_{s\neq 0} T_s(a_s+b_s)}{2}$ and (23). Then, applying Lemma 12 on (32) gives:

$$\mathbb{P}(D^*_{ii} \leq \delta) \leq e^{-\frac{\ln(n)}{2}\nu + o(1)} \tag{33}$$

where $\nu = \sum_{s\neq 0} \big(\sqrt{a_s} - \sqrt{b_s}\big)^2$. If $\nu \geq 2$, then $\mathbb{P}\big(D^*_{ii} \geq \delta\big) = 1 - o(1)$, and this completes the proof.

REFERENCES

[1] T. Ji, C. Luo, Y. Guo, Q. Wang, L. Yu, and P. Li, "Community detection in online social networks: A differentially private and parsimonious approach," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 1, pp. 151–163, 2020.

[2] D. Jin, X. Wang, M. Liu, J. Wei, W. Lu, and F. Fogelman-Soulié, "Identification of generalized semantic communities in large social networks," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2966–2979, 2020.

[3] M. Li, S. Lu, L. Zhang, Y. Zhang, and B. Zhang, "A community detection method for social network based on community embedding," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 2, pp. 308–318, 2021.

[4] W. P. Grant, S. E. Ahnert, and E. Estrada, "Modular decomposition of protein structure using community detection," *Journal of Complex Networks*, vol. 7, no. 1, pp. 101–113, 2018.

[5] M. Li, X. Meng, R. Zheng, F.-X. Wu, Y. Li, Y. Pan, and J. Wang, "Identification of protein complexes by using a spatial and temporal active protein interaction network," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 3, pp. 817–827, 2020.

[6] N. Zaki, H. Singh, and E. A. Mohamed, "Identifying protein complexes in protein-protein interaction data using graph convolutional network," *IEEE Access*, vol. 9, pp. 123 717–123 726, 2021.

[7] L. Huang, Y. Yang, H. Gao, X. Zhao, and Z. Du, "Comparing community detection algorithms in transport networks via points of interest," *IEEE Access*, vol. 6, pp. 29 729–29 738, 2018.

[8] G. Rossetti and R. Cazabet, "Community discovery in dynamic networks: a survey," *ACM computing surveys (CSUR)*, vol. 51, no. 2, pp. 1–37, 2018.

[9] P. Wagenseller, F. Wang, and W. Wu, "Size matters: A comparative analysis of community detection algorithms," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 4, pp. 951–960, 2018.

[10] P. Chunaev, "Community detection in node-attributed social networks: a survey," *Computer Science Review*, vol. 37, p. 100286, 2020.

[11] K. Taha, "Static and dynamic community detection methods that optimize a specific objective function: A survey and experimental evaluation," *IEEE Access*, vol. 8, pp. 98 330–98 358, 2020.

[12] Z. Wang, C. Wang, X. Li, C. Gao, X. Li, and J. Zhu, "Evolutionary markov dynamics for network community detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 3, pp. 1206–1220, 2020.

[13] N. Alotaibi and D. Rhouma, "A review on community structures detection in time evolving social networks," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, Part B, pp. 5646–5662, 2022.

[14] G. Xue, M. Zhong, J. Li, J. Chen, C. Zhai, and R. Kong, "Dynamic network embedding survey," *Neurocomputing*, vol. 472, pp. 212–223, 2022.

[15] O. F. Robledo, M. Klepper, E. v. Boven, and H. Wang, "Community detection for temporal weighted bipartite networks," in *International Conference on Complex Networks and Their Applications*. Springer, 2022, pp. 245–257.

[16] T. Jia, C. Cai, X. Li, X. Luo, Y. Zhang, and X. Yu, "Dynamical community detection and spatiotemporal analysis in multilayer spatial interaction networks using trajectory data," *International Journal of Geographical Information Science*, vol. 36, no. 9, pp. 1719–1740, 2022.

[17] H. Safdari, M. Contisciani, and C. De Bacco, "Reciprocity, community detection, and link prediction in dynamic networks," *Journal of Physics: Complexity*, vol. 3, no. 1, p. 015010, 2022.

[18] J. Singh, D. Pandey, and A. K. Singh, "Event detection from real-time twitter streaming data using community detection algorithm," *Multimedia Tools and Applications*, vol. 83, no. 8, pp. 23 437–23 464, 2024.

[19] A. Ghasemian, P. Zhang, A. Clauset, C. Moore, and L. Peel, "Detectability thresholds and optimal algorithms for community structure in dynamic networks," *Phys. Rev. X*, vol. 6, p. 031005, Jul 2016.

[20] P. Barucca, F. Lillo, P. Mazzarisi, and D. Tantari, "Disentangling group and link persistence in dynamic stochastic block models," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2018, no. 12, p. 123407, dec 2018.

[21] L. Dall'Amico, R. Couillet, and N. Tremblay, "Community detection in sparse time-evolving graphs with a dynamical bethe-hessian," *Advances in Neural Information Processing Systems*, vol. 33, pp. 7486–7497, 2020.

[22] J. Z. Moghaddam, M. Esmaeili, and A. Nosratinia, "Exact recovery threshold in dynamic binary censored block model," in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 1088–1093.

[23] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," *Social networks*, vol. 5, no. 2, pp. 109–137, 1983.

[24] E. Abbe, "Community detection and stochastic block models: Recent developments," *Journal of Machine Learning Research*, vol. 18, no. 177, pp. 1–86, 2018.

[25] L. Massoulié, "Community detection thresholds and the weak ramanujan property," in *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, 2014, pp. 694–703.

[26] E. Abbe, A. S. Bandeira, A. Bracher, and A. Singer, "Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery," *IEEE Transactions on Network Science and Engineering*, vol. 1, no. 1, pp. 10–22, 2014.

[27] B. Hajek, Y. Wu, and J. Xu, "Achieving exact cluster recovery threshold via semidefinite programming," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2788–2797, 2016.

[28] ——, "Achieving exact cluster recovery threshold via semidefinite programming: Extensions," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5918–5937, 2016.

[29] C. Gao, Z. Ma, A. Y. Zhang, and H. H. Zhou, "Achieving optimal misclassification proportion in stochastic block models," *Journal of Machine Learning Research*, vol. 18, no. 60, pp. 1–45, 2017.

[30] H. Saad and A. Nosratinia, "Community detection with side information: Exact recovery under the stochastic block model," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 5, pp. 944–958, 2018.

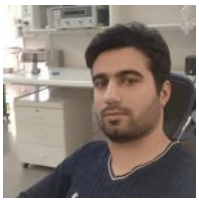
[31] P. Wang, Z. Zhou, and A. M.-C. So, "A nearly-linear time algorithm for exact community recovery in stochastic block model," in *International Conference on Machine Learning*. PMLR, 2020, pp. 10 126–10 135.

[32] Q. Duchemin, "Inference in the stochastic block model with a markovian assignment of the communities." *CoRR*, 2020.

[33] M. Esmaeili, H. M. Saad, and A. Nosratinia, "Semidefinite programming for community detection with side information," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1957–1973, 2021.

[34] B. E. Hajek and S. Sankagiri, "Community recovery in a preferential attachment graph," *IEEE Transactions on Information Theory*, vol. 65, pp. 6853–6874, 2018.

[35] K.-k. Shang, B. Yang, J. M. Moore, Q. Ji, and M. Small, "Growing networks with communities: A distributive link model," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 4, 2020.
[36] C. Gao, J. Zhu, F. Zhang, Z. Wang, and X. Li, "A novel representation learning for dynamic graphs based on graph convolutional networks," *IEEE Transactions on Cybernetics*, 2022.
[37] E. P. Xing, W. Fu, and L. Song, "A state-space mixed membership blockmodel for dynamic network tomography," *The Annals of Applied Statistics*, vol. 4, no. 2, pp. 535–566, 06 2010.
[38] K. S. Xu and A. O. Hero, "Dynamic stochastic blockmodels: Statistical models for time-evolving networks," in *Proceedings of the 6th International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction*. Springer-Verlag, 2013, pp. 201–210.
[39] ——, "Dynamic stochastic blockmodels for time-evolving social networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 552–562, 2014.
[40] K. S. Xu, "Stochastic block transition models for dynamic networks," *ArXiv*, vol. abs/1411.5404, 2014.
[41] M. Arastuie, *Generative Models of Link Formation and Community Detection in Continuous-Time Dynamic Networks*. The University of Toledo, 2020.
[42] T. Yang, Y. Chi, S. Zhu, Y. Gong, and R. Jin, "Detecting communities and their evolutions in dynamic social networks—a Bayesian approach," *Machine Learning*, vol. 82, no. 2, pp. 157–189, Feb 2011.
[43] M. Kim and J. Leskovec, "Nonparametric multi-group membership model for dynamic networks," in *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 1385–1393.
[44] A. Karaaslanlı and S. Aviyente, "Community detection in dynamic networks: Equivalence between stochastic blockmodels and evolutionary spectral clustering," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 7, pp. 130–143, 2021.
[45] X. Zhang, C. Moore, and M. E. J. Newman, "Random graph models for dynamic networks," *The European Physical Journal B*, vol. 90, pp. 1–14, 2017.
[46] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1273–1289, 2001.



Javad Zahedi Moghaddam received the B.Sc. degree in electrical engineering-control from Isfahan University of Technology, Isfahan, Iran, in 2014, and the M.Sc. degree in electrical engineering-control from Tarbiat Modares University, Tehran, Iran, in 2017. He is currently working toward a Ph.D. degree in electrical engineering with the University of Texas at Dallas, Richardson, TX, USA. His research interests include machine learning, control, information theory, and wireless communication.

Aria Nosratinia (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 1996. He is currently the Erik Jonsson Distinguished Professor and Associate Head of the Electrical and computer Engineering Department, The University of Texas at Dallas, Richardson, TX, USA. He has held visiting appointments with Princeton University, Rice University, and UCLA. His interests are in the broad areas of information theory and signal processing. He was Associate Editor of the IEEE Transactions on Information Theory, IEEE Transactions on Communications, IEEE Signal Processing Letters, IEEE Transactions on Image Processing, and IEEE Wireless Communications. He was named a Highly Cited Researcher by Clarivate Analytics (formerly Thompson Reuters).