

Detectability Thresholds and Optimal Algorithms for Community Structure in Dynamic Networks

Amir Ghasemian,¹ Pan Zhang,² Aaron Clauset,^{1,3,4} Cristopher Moore,⁴ and Leto Peel^{5,6}

¹*Department of Computer Science, University of Colorado, Boulder, Colorado 80309, USA*

²*CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*

³*BioFrontiers Institute, University of Colorado, Boulder, Colorado 80305, USA*

⁴*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

⁵*ICTEAM, Université catholique de Louvain, Avenue George Lemaître 4, B-1348 Louvain-la-Neuve, Belgium*

⁶*naXys, University of Namur, Rempart de la Vierge 8, 5000 Namur, Belgium*

(Received 2 November 2015; revised manuscript received 17 May 2016; published 13 July 2016)

The detection of communities within a dynamic network is a common means for obtaining a coarse-grained view of a complex system and for investigating its underlying processes. While a number of methods have been proposed in the machine learning and physics literature, we lack a theoretical analysis of their strengths and weaknesses, or of the ultimate limits on when communities can be detected. Here, we study the fundamental limits of detecting community structure in dynamic networks. Specifically, we analyze the limits of detectability for a dynamic stochastic block model where nodes change their community memberships over time, but where edges are generated independently at each time step. Using the cavity method, we derive a precise detectability threshold as a function of the rate of change and the strength of the communities. Below this sharp threshold, we claim that no efficient algorithm can identify the communities better than chance. We then give two algorithms that are optimal in the sense that they succeed all the way down to this threshold. The first uses belief propagation, which gives asymptotically optimal accuracy, and the second is a fast spectral clustering algorithm, based on linearizing the belief propagation equations. These results extend our understanding of the limits of community detection in an important direction, and introduce new mathematical tools for similar extensions to networks with other types of auxiliary information.

DOI: [10.1103/PhysRevX.6.031005](https://doi.org/10.1103/PhysRevX.6.031005)

Subject Areas: Complex Systems,
Interdisciplinary Physics,
Statistical Physics

I. INTRODUCTION

Many complex systems can be represented as networks, that is, as a set of elements characterized by pairwise interactions. Examples of networks are plentiful and include friendships or communication in a social network, regulatory interactions among genes, transportation between cities, and hyperlinks between documents or websites. Furthermore, many, perhaps even most, of these networks are dynamic in nature, and their evolving structure is often represented as a sequence of graphs [1–8].

A common step in analyzing the structure of such networks is the detection of communities, in which we seek to divide a network into groups of nodes that play similar structural roles. A good division should provide a structural

coarse graining of the network, revealing the large-scale structure of the system. In the simple case of static networks, we now have rigorous methods for accomplishing this task, using Bayesian techniques and probabilistic generative models [9–15]. Importantly, we also have a precise mathematical understanding of when they can or cannot succeed [14,16]. However, real-world networks are rarely simple, and are often accompanied by auxiliary data, such as weights on edges [17] or metadata on nodes [18–22]. Extending the rigorous results for static networks to these richer graph structures remains an important direction of study. Here, we focus on the question of dynamic networks, in which each node's connections may change over time.

Community detection in dynamic networks inherits many of the challenges of static networks, including learning the number of communities, their sizes and node membership, and the pattern of connections among communities, e.g., assortative or disassortative (or, in physical terms, ferromagnetic or antiferromagnetic). However, it also poses new challenges, as both the network topology and the community memberships may evolve over time.

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Community detection in dynamic networks has a long history, and a number of techniques have been previously developed. For instance, there are variants of multilayer or temporal modularity maximization [5,23,24], non-negative matrix or tensor factorization [3,6,8,25,26], minimum description length [27,28], and probabilistic models [4,7,29–33]. References [34,35] provide more comprehensive reviews of this work. Approaches for detecting communities in multiplex networks are also relevant [36–45], as dynamic networks are a special case of multiplex networks, in which the layers are organized in a linear sequence. However, despite these varied efforts, up to now we have lacked a theoretical understanding of the optimality of these techniques, when or how they tend to fail, or whether there are fundamental limits to detecting community structure in dynamic networks.

Here, we answer these questions by deriving a precise threshold on the detectability of communities in dynamic networks, whose location depends only on the rate of change of the community structure and on its strength. Below this sharp threshold, we claim no efficient algorithm can recover the true communities better than chance. Furthermore, we give two algorithms that are optimal in the sense that they succeed all the way down to this threshold. These results generalize the theoretical insights of Refs. [14,15] for community detection in static networks to the dynamic setting, in which detectability depends on both spatial and temporal coupling between nodes. The mathematical tools we use to obtain these results are also general, and could be used to obtain similar extensions to networks with other types of auxiliary information.

Our approach exploits the powerful tools of probabilistic generative models and Bayesian inference, which we use to study the limits of the community detection problem using the cavity method of statistical physics. We begin with the well-known stochastic block model [46,47], a generative model for static networks with community structure. We note that there are several dynamic variants of this model [29–31] and its mixed-membership version [7], and the variant that we analyze here is a special case of some of these models. Specifically, we mathematically study a model in which nodes change their community membership over time according to a Markov process, and edges are generated independently at each time step. As a result, the network of connections between nodes at different times is locally treelike.

In many real-world systems, edge occurrence can correlate across time [1]. In this case, however, our model can still be applied if the edges have a time scale that is short relative to the time windows over which the interactions are aggregated, which returns us to a setting in which the dynamic network will be locally treelike. For instance, consider a network of phone calls or emails where the autocorrelation time governing conversations (or sequences of successive calls) is on the order of days, but where each

network snapshot aggregates these calls over a month. In this case, belief propagation (BP) algorithms, like the ones we develop here, are often asymptotically optimal, and we may use the cavity method to compute the detectability threshold exactly. While our results are not mathematically rigorous, we believe that they can be made so using the techniques of Refs. [16,48–50].

Finally, we give two principled and efficient algorithms for detecting communities in real dynamic networks. The first algorithm uses BP to pass messages between neighbors both within a network at a particular time and between consecutive networks in order to integrate information over the network's time series in an optimal way. We then linearize BP to obtain a second spectral algorithm, based on a dynamical version of the nonbacktracking matrix [50,51]. Through numerical experiments, we confirm our theoretical calculations by showing that these algorithms accurately recover the true community structure in dynamic networks all the way down to the generalized detectability threshold.

II. DYNAMIC STOCHASTIC BLOCK MODEL

The stochastic block model is a classic model of community structure in static networks [46,47]. To obtain a theoretical understanding of detectability in dynamic networks, we use a variant of the stochastic block model in which the community labels of nodes change over time, but where edges are independent conditioned on these labels. This particular model is also a special case of several models previously introduced for community detection in dynamic networks [4,7,29–31]. A crucial feature of the variant we study is that it captures the dynamic behavior of changing community labels but is analytically tractable.

Our model generates a dynamic sequence of graphs $G(t) = (V, E(t))$, with $1 \leq t \leq T$. There are $|V| = n$ nodes divided into k groups. Each graph has its own group assignment, represented by an n -dimensional vector of labels $\{g_i(t) \in \{1, \dots, k\} | i \in V\}$. To generate this sequence, we start by drawing $g_i(1)$ from a prior distribution, where each node has initial probability q_r of being in community $1 \leq r \leq k$. In successive steps $t > 1$, each node updates its label according to a transition matrix τ , moving to group r from group s with probability τ_{rs} . Finally, the edges $E(t)$ are generated independently for each t according to a $k \times k$ matrix p , connecting each pair of nodes i, j at time t with probability $p_{g_i(t), g_j(t)}$. The likelihood of the graph sequence is then

$$\begin{aligned} P(\{G(t)\}, \{g(t)\} | p, q, \tau) \\ = P(g(1)) \prod_{t=2}^T P(g(t) | g(t-1)) \\ \times \prod_{t=1}^T \left[\prod_{(i,j) \in E(t)} p_{g_i(t), g_j(t)} \prod_{(i,j) \notin E(t)} (1 - p_{g_i(t), g_j(t)}) \right], \end{aligned}$$

and

$$P(g(1)) = \prod_i q_{g_i(1)},$$

$$P(g(t)|g(t-1)) = \prod_i \tau_{g_i(t), g_i(t-1)}.$$

In our analysis, we focus on a uniform initial prior $q_r = 1/k$ and the popular special case, where $p_{rs} = c_{\text{in}}/n$ if $r = s$ and c_{out}/n if $r \neq s$ for constants $c_{\text{in}}, c_{\text{out}}$. The average degree of each graph is then $c = [c_{\text{in}} + (k-1)c_{\text{out}}]/k$. For simplicity, we also assume the transition matrix τ has a special form, where the node keeps its label with probability η and chooses a uniformly random label with probability $1 - \eta$. In that case,

$$\tau = \eta \mathbb{1} + (1 - \eta) \frac{J}{k}, \quad (1)$$

where $\mathbb{1}$ is the identity matrix and J is the all-1's matrix.

III. GENERALIZED DETECTABILITY THRESHOLD

In this context, the community detection task consists of recovering the labels $\{g_i(t)\}$ given the parameters p, q, η and the sequence of graphs $\{G(t)\}$. We now consider under what conditions we can perform this task better than chance. For static networks, previous work has shown that there exists a phase transition below which no algorithm can succeed [14,15]; for the case $k = 2$, this is now known rigorously [16]. This threshold occurs at a critical value of $c_{\text{in}} - c_{\text{out}}$, which depends on the average degree, namely, $|c_{\text{in}} - c_{\text{out}}| = k\sqrt{c}$.

In a dynamic network where community memberships change slowly, we can learn more about a network and its large-scale structure by integrating its edges over time. Summing $G(t)$ to form a single graph yields a denser network, in which case we would expect to be able to detect its community structure whenever $c_{\text{in}} - c_{\text{out}} \neq 0$. On the other hand, if node labels at successive steps are uncorrelated, we can do no better than to treat each graph in $G(t)$ separately as a static graph. We thus expect the community detection threshold in dynamic networks to interpolate between its static value at $\eta = 0$ and zero at $\eta = 1$.

To facilitate our analysis, we define a *spatiotemporal* graph with Tn vertices $i(t)$, one for each node at each time step. In addition to the “spatial” edges $(i(t), j(t)) \in E(t)$ for each t , we add “temporal” edges $(i(t), i(t \pm 1))$ connecting each node with its time-adjacent copies. Since the spatial edges $E(t)$ are independent and sparse, short loops in this spatiotemporal graph are rare, implying that it is locally treelike.

Now consider the neighborhood of a particular node $i(t)$. Moving outward in space and time, there is a tree with $i(t)$

as its root: each node in this tree has “children” consisting of its spatial and temporal neighbors. Using the cavity method, we can think of inference as a reconstruction problem on this tree, where each child’s label is transmitted, with some noise, to its parent. As stated above, we assume that node labels are copied along temporal edges with probability η and replaced with uniformly random labels with probability $1 - \eta$. Similarly, since each edge in $E(t)$ exists with probability c_{in}/n if the labels are the same, and with probability c_{out}/n otherwise, Bayes’s rule implies that labels are copied along a spatial edge with probability

$$\lambda = \frac{c_{\text{in}} - c_{\text{out}}}{kc}$$

and replaced with a random label with probability $1 - \lambda$. Thus, we can think of the labels on spatial and temporal edges as following a Markov process with stochastic transition matrices σ and τ , respectively, where

$$\sigma = \frac{np}{kc} = \lambda \mathbb{1} + (1 - \lambda) \frac{J}{k}, \quad (2)$$

and τ is given by Eq. (1).

We now consider the question of whether information from distant leaves on this tree is transmitted to the root. The tree is generated by a two-type branching process: following a temporal edge leads to a node with one temporal child, while following a spatial edge leads to a node with two temporal children, and in both cases the number of spatial children is Poisson distributed with mean c . The transition matrix describing the expected number of children of each type is then $\begin{pmatrix} c & c \\ 2 & 1 \end{pmatrix}$. On the other hand, besides the trivial eigenvalue 1 corresponding to the uniform distribution, the eigenvalues of the transition matrices σ and τ are λ and η , respectively. The results of Ref. [52] then imply that the detectability transition occurs when the largest eigenvalue of the matrix $\begin{pmatrix} c\lambda^2 & c\lambda^2 \\ 2\eta^2 & \eta^2 \end{pmatrix}$ crosses unity. This yields

$$c\lambda^2 = \frac{1 - \eta^2}{1 + \eta^2} \quad \text{or} \quad |c_{\text{in}} - c_{\text{out}}| = k\sqrt{c \frac{1 - \eta^2}{1 + \eta^2}}, \quad (3)$$

which ranges from the static threshold $k\sqrt{c}$ when $\eta = 0$ to zero when $\eta = 1$, as expected.

The expression of Eq. (3) holds in the limit $T \rightarrow \infty$. We can compute the corresponding finite-time threshold for a fixed T by diagonalizing a $(3T - 2)$ -dimensional matrix, where we have a branching process with states corresponding to moving along spatial, forward-temporal, or backward-temporal edges at each time step. In particular, the threshold then ranges from the static value for $\eta = 0$ to $|c_{\text{in}} - c_{\text{out}}| = k\sqrt{c/T}$ for $\eta = 1$.

In spin glass theory, this type of threshold is called the Almeida-Thouless line [53]; in probability and information theory, it is known as the Kesten-Stigum bound or the robust reconstruction threshold [52]. In the static case, it has been shown rigorously for $k = 2$ that below this point community detection is information theoretically impossible [16]. For $k > 4$ groups (and $k = 4$ in the disassortative or antiferromagnetic case $c_{\text{in}} - c_{\text{out}} < 0$), it was first conjectured [14,15], and later proved [54,55], that there is an additional region below the threshold where community detection is information theoretically possible, but exponentially hard, so that no efficient algorithm can do better than chance. We make the same claims for the dynamic case.

IV. BAYESIAN INFERENCE AND BELIEF PROPAGATION

Given an observed graph sequence $G(t)$, we wish to infer the posterior distribution of group assignments $\{g(t)\}$. For fixed p , q , and η , Bayes's rule gives us

$$P(\{g(t)\}|\{E(t)\}) = \frac{P(\{E(t)\}, \{g(t)\})}{\sum_{\{g'(t)\}} P(\{E(t)\}, \{g'(t)\})}. \quad (4)$$

We are especially interested in the one-point marginals of this distribution, i.e., the probability distribution of $g_i(t)$ for each node i at each step t . We denote this as

$$\begin{aligned} \mu_r^i(t) &= P(g_i(t) = r|\{E(t)\}) \\ &= \sum_{\{g(t)\}} P(\{g(t)\}|\{E(t)\}) \delta_{g_i(t), r}. \end{aligned}$$

As always, computing the denominator in Eq. (4) is difficult, as it is a sum over k^{Tn} terms. In physical terms, this quantity is a partition function and thus we do not expect to be able to compute $P(\{g(t)\}|\{E(t)\}, p, \eta)$ exactly. However, we can approximate it with variational methods. Since the spatiotemporal graph is locally treelike, we can make a Bethe approximation, corresponding to the cavity method in physics or belief propagation in machine learning. This allows us to approximate the marginals and the free energy of the model in an efficient and asymptotically optimal way.

In belief propagation, vertices send their neighbors “messages” consisting of estimates of their marginal distributions. Each vertex updates its message to each of its neighbors, based on the message it receives from its other neighbors. It does this using Bayes's rule, assuming that its neighbors are independent of each other (conditioned on that vertex's state). We then update the messages until they reach a fixed point.

In our dynamic setting, we have two kinds of messages, passing along the spatial and temporal edges of the spatiotemporal graphs (Fig. 1). We denote the spatial

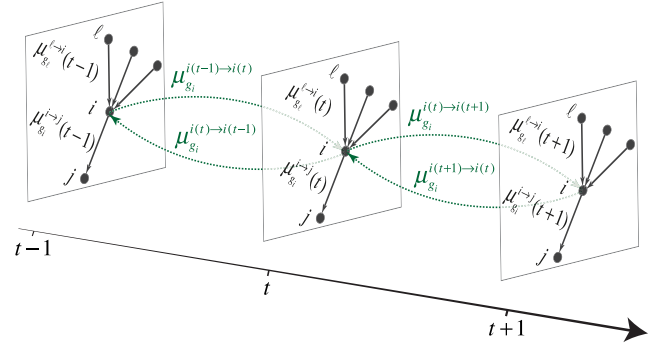


FIG. 1. A schematic representation of belief propagation messages [see Eqs. (5) and (6)] being passed along spatial and temporal edges in the spatiotemporal graph.

messages $\mu_r^{i \rightarrow j}(t)$: this is i 's estimate, sent to j , of the probability that i belongs to group r at time t . Similarly, the temporal message $\mu_r^{i(t) \rightarrow i(t \pm 1)}$ is i 's estimate of this probability sent to its past and future selves.

For general values of the connection probabilities p_{rs} and the transition matrix τ , the update equation for the spatial messages is

$$\begin{aligned} \mu_r^{i \rightarrow j}(t) &= \frac{1}{Z^{i \rightarrow j}(t)} \left(\sum_s \tau_{rs} \mu_s^{i(t-1) \rightarrow i(t)} \right) \left(\sum_s \tau_{sr} \mu_s^{i(t+1) \rightarrow i(t)} \right) \\ &\times \prod_{\substack{\ell: (i, \ell) \in E(t) \\ \ell \neq j}} \sum_s c_{rs} \mu_s^{\ell \rightarrow i}(t) \prod_{\substack{\ell: (i, \ell) \notin E(t) \\ \ell \neq j}} \sum_s (1 - p_{rs}) \mu_s^{\ell \rightarrow i}(t), \end{aligned} \quad (5)$$

and the update equation for the temporal messages is

$$\begin{aligned} \mu_r^{i(t) \rightarrow i(t+1)} &= \frac{1}{Z^{i(t) \rightarrow i(t+1)}} \left(\sum_s \tau_{rs} \mu_s^{i(t-1) \rightarrow i(t)} \right) \\ &\times \prod_{\ell: (i, \ell) \in E(t)} \sum_s c_{rs} \mu_s^{\ell \rightarrow i}(t) \\ &\times \prod_{\ell: (i, \ell) \notin E(t)} \sum_s (1 - p_{rs}) \mu_s^{\ell \rightarrow i}(t), \end{aligned} \quad (6)$$

where $Z^{i \rightarrow j}(t)$ and $Z^{i(t) \rightarrow i(t \pm 1)}$ are normalization factors, and similarly for $\mu_r^{i(t) \rightarrow i(t-1)}$ with τ transposed. Finally, once the messages reach a fixed point, we compute the marginals at each vertex by taking all its incoming messages into account:

$$\begin{aligned} \mu_r^i(t) &= \frac{1}{Z^i(t)} \left(\sum_s \tau_{rs} \mu_s^{i(t-1) \rightarrow i(t)} \right) \left(\sum_s \tau_{sr} \mu_s^{i(t+1) \rightarrow i(t)} \right) \\ &\times \prod_{\ell: (i, \ell) \in E(t)} \sum_s c_{rs} \mu_s^{\ell \rightarrow i}(t) \\ &\times \prod_{\ell: (i, \ell) \notin E(t)} \sum_s (1 - p_{rs}) \mu_s^{\ell \rightarrow i}(t). \end{aligned} \quad (7)$$

Of course, when $t = 1$ or $t = T$, we remove the term corresponding to the temporal edge coming from outside the domain of t .

In the update equations given here, we have $O(Tn^2)$ messages, with spatial messages between both neighboring and non-neighboring pairs of nodes. As in Refs. [13–15], in the sparse case $p_{rs} = c_{rs}/n$, we can approximate the effect of non-neighboring pairs with an external field, so that we only need to keep track of $O(Tn)$ messages between nodes and their spatiotemporal neighbors. This amounts to writing

$$\prod_{\ell} \sum_s (1 - p_{rs}) \mu_s^{\ell \rightarrow i}(t) = e^{-h_r(t)},$$

where

$$h_r(t) = \frac{1}{n} \sum_{\ell} \sum_s c_{rs} \mu_s^{\ell \rightarrow i}(t). \quad (8)$$

Furthermore, in our special case where τ and σ are given by Eqs. (1) and (2), we have

$$\begin{aligned} \sum_s \tau_{rs} \mu_s^{i(t-1) \rightarrow i(t)} &= \eta \mu_r^{i(t-1) \rightarrow i(t)} + \frac{1 - \eta}{k}, \\ \sum_s \tau_{sr} \mu_s^{i(t+1) \rightarrow i(t)} &= \eta \mu_r^{i(t+1) \rightarrow i(t)} + \frac{1 - \eta}{k}, \\ \sum_s c_{rs} \mu_s^{\ell \rightarrow i}(t) &= \lambda \mu_r^{\ell \rightarrow i}(t) + \frac{1 - \lambda}{k}. \end{aligned}$$

As in Refs. [14,15], the BP equations have a trivial fixed point where all the messages are uniform: $\mu^{i \rightarrow j}(t) = \mu^{i(t) \rightarrow i(t \pm 1)} = 1/k$ for all i, j , and t . The Kesten-Stigum transition computed above is precisely where this fixed point becomes unstable. Below the transition, BP converges to the trivial fixed point, all marginals are uniform, and the algorithm performs no better than chance.

However, above this transition, the trivial fixed point is unstable, and BP converges to a nontrivial fixed point. If the network is generated by our model and we know the correct parameters, we expect this nontrivial fixed point to give an asymptotically correct estimate of the marginals Eq. (7) up to a permutation of the groups. In physical terms, we are on the Nishimori line, so there is no static replica symmetry breaking and no spin glass phase [56]. Then, if we assign each node its most likely label at each time, setting $\widehat{g_i}(t) = \arg\max_r \mu_r^i(t)$, this assignment maximizes the fraction of correct labels. Thus, our BP algorithm succeeds all the way down to the detectability threshold given by Eq. (3), and is asymptotically optimal in terms of its accuracy. We expect that this can be made rigorous, at least in the case $k = 2$, using the techniques of Refs. [16,48,49].

For $k \leq 3$, the detectability transition is second order, with the optimal accuracy going to zero continuously at the transition. In analogy with the static block model [14,15], we believe that for $k > 4$ (or $k \geq 4$ for the disassortative case) the detectability transition becomes first order. Then there is an additional regime where there are at least two competing fixed points, the trivial one and an accurate one, both of which are locally stable. However, the basin of attraction of the accurate fixed point is exponentially small, so that BP with random initial messages will almost always converge to the trivial fixed point. In this regime, community detection is information theoretically possible, but it would require exponential time to search the space of possible fixed points. Physically, there is a free-energy barrier between the trivial fixed point, which corresponds to a paramagnetic phase, and the accurate fixed point, which corresponds to a ferromagnetic one.

V. SPECTRAL CLUSTERING

In dense networks, a common approach to detecting communities is spectral clustering, which is accomplished by examining the eigenvectors of either the adjacency or Laplacian matrix. In sparse networks, approach fails strictly above the detectability threshold [51,57]. However, it is known that this difficulty can be circumvented, in static networks, by using a nonbacktracking matrix or Hashimoto operator [50,51], which prevents localization of the eigenvectors to the high-degree vertices. Here, we extend these techniques to derive a spectral algorithm for sparse dynamic networks.

The idea is simply to linearize the BP update equations around the trivial fixed point described above, expanding all the messages to first order around $1/k$. In the static case, this linearization yields the nonbacktracking matrix. Its second through k th eigenvectors are correlated with the true community structure all the way down to the detectability transition [50,51], so that we can label nodes using a clustering technique in $\mathbb{R}^{k-1}$ such as the k -means algorithm.

For the dynamic block model, linearizing the BP update equations around the trivial fixed point gives a $2km \times 2km$ matrix, where m is the total number of edges in the spatiotemporal graph. Analogous to Ref. [51], this is the tensor product of a $k \times k$ matrix with a $2m \times 2m$ matrix, which we can simplify further by writing it in terms of the total incoming and outgoing messages at each vertex. This gives a $4nT \times 4nT$ matrix:

$$B = \begin{pmatrix} \lambda \mathbb{A}^{\text{spatial}} & -\lambda \mathbb{1} & \lambda \mathbb{A}^{\text{spatial}} & 0 \\ \lambda(\mathbb{D}^{\text{spatial}} - \mathbb{1}) & 0 & \lambda \mathbb{D}^{\text{spatial}} & 0 \\ \eta \mathbb{A}^{\text{temp}} & 0 & \eta \mathbb{A}^{\text{temp}} & -\eta \mathbb{1} \\ \eta \mathbb{D}^{\text{temp}} & 0 & \eta(\mathbb{D}^{\text{temp}} - \mathbb{1}) & 0 \end{pmatrix}. \quad (9)$$

Here, $\mathbb{1}$ denotes the nT -dimensional identity matrix, $\mathbb{A}^{\text{temp}}$ is the adjacency matrix of temporal edges, $\mathbb{D}^{\text{temp}}$ is the diagonal matrix of temporal degrees, $\mathbb{A}^{\text{spatial}}$ is the adjacency matrix of spatial edges, and $\mathbb{D}^{\text{spatial}}$ is the diagonal matrix of spatial degrees. That is,

$$\begin{aligned}\mathbb{A}_{(u,t),(v,t')}^{\text{temp}} &= \delta_{uv}(\delta_{t,t'+1} + \delta_{t,t'-1}), \\ \mathbb{D}_{(u,t),(u,t)}^{\text{temp}} &= \begin{cases} 2 & \text{if } 1 < t < T \\ 1 & \text{if } t = 1 \text{ or } t = T, \end{cases} \\ \mathbb{A}_{(u,t),(v,t')}^{\text{spatial}} &= \bigoplus_t A(t) = \begin{cases} 1 & \text{if } t = t' \text{ and } (u,v) \in E(t) \\ 0 & \text{otherwise,} \end{cases} \\ \mathbb{D}_{(u,t),(u,t)}^{\text{spatial}} &= \bigoplus_t D(t) = \sum_v A_{u,v}(t), \end{aligned}$$

where the symbol $\bigoplus$ denotes the matrix direct sum (diagonal concatenation). The terms $\lambda \mathbb{A}^{\text{spatial}}$ and $\eta \mathbb{A}^{\text{temp}}$ in Eq. (9) correspond to attenuation of the messages along the spatial and temporal edges by the second eigenvalues of σ and τ , respectively. The terms $-\mathbb{1}$ correspond to the nonbacktracking nature of belief propagation, and are known in physics as Onsager reaction terms.

This analysis gives us a spectral algorithm for dynamic networks. We form the spatiotemporal graph, construct the matrix B , compute the $k-1$ eigenvectors with the largest eigenvalue (in absolute value), and finally perform k -means clustering on the resulting n vectors in $\mathbb{R}^{k-1}$. In the case $k=2$, we can simply label nodes according to the sign of the second eigenvector to separate nodes into two communities.

As in the static case [51], the instability of the trivial fixed point corresponds exactly to where the community-correlated eigenvectors emerge from the bulk of B 's spectrum in the complex plane. Thus, we claim that, while it is somewhat less accurate than belief propagation, this spectral algorithm is optimal in the sense that it works all the way down to the dynamical detectability transition. Finally, we also expect that this result can be made rigorous, as Ref. [50] did for the static case.

VI. NUMERICAL EXPERIMENTS

To verify our claims of the detectability transition in dynamic networks and the accuracy of our algorithms, we conduct the following numerical experiment. Using the dynamic block model, we generate a number of dynamic networks with various rates of change and various strengths of community structure. We then use the BP and spectral algorithms to infer the group assignments, assuming that the true parameters are available to the algorithm, and measure their accuracy against the known underlying structure.

Following past work, we parametrize the strength of the community structure by $\epsilon = c_{\text{out}}/c_{\text{in}}$. For $\epsilon = 0$, nodes connect only to others in the same group, while for $\epsilon = 1$,

the network at each time is an Erdős-Rényi random graph with no community structure. In terms of ϵ , the detectability transition in Eq. (3) occurs at

$$\lambda = \frac{1 - \epsilon}{1 + (k-1)\epsilon} = \frac{1}{\sqrt{c}} \sqrt{\frac{1 - \eta^2}{1 + \eta^2}}. \quad (10)$$

In our experiments, we explore networks in the (ϵ, η) plane, while keeping the average degree c fixed.

We measure the accuracy of the inferred labels by the *overlap* between the true assignment g^* and the inferred one $\hat{g}$. This is the fraction of nodes labeled correctly, averaged over all nodes and all times, normalized so that it is 1 if $\hat{g} = g^*$ and 0 if $\hat{g}$ is uniformly random. (To break the permutation symmetry, we maximize over all $k!$ permutations of the groups.) In Fig. 2, we show the overlap obtained by BP for dynamic networks as a function of ϵ for several choices of η , with $n = 512$, $T = 40$, $k = 2$, and $c = 16$. For each η , the critical value of ϵ for $T = 40$ is shown as a vertical line in the lower panel. As predicted, the critical ϵ increases with η ; we find numerically that the finite-size effects scale as $n^{-1/2}$, typical of phase transitions in infinite-dimensional systems.

Figure 3 shows the overlap throughout the (ϵ, η) plane, using both BP and our spectral algorithm, again for $k = 2$ and $c = 16$. The dashed curve shows the detectability transition for $T = 40$; note that close to $\eta = 1$ it diverges from the transition for $T = \infty$, which is shown as the magenta curve. While BP achieves a higher overlap throughout the (ϵ, η) plane, both algorithms achieve a large overlap when ϵ is small or η is large, i.e., when the community structure is strong or when the group memberships change slowly. As we approach the critical curve, both algorithms undergo a second-order transition, with their accuracy going to zero continuously. (As stated

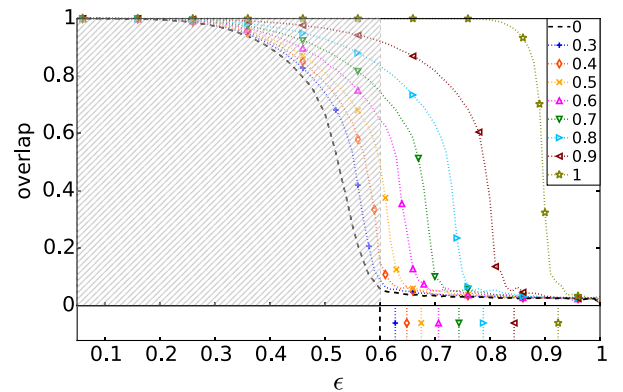


FIG. 2. Overlap as a function of ϵ for different values of η (given in the legend). For each η , the critical value of ϵ for $T = 40$ is shown as a vertical line in the lower panel, and the hatched area shows the region of detectability for static networks [14,15]. Each data point is the average of 100 instances, with $n = 512$, $T = 40$, $k = 2$ groups, and average degree $c = 16$.

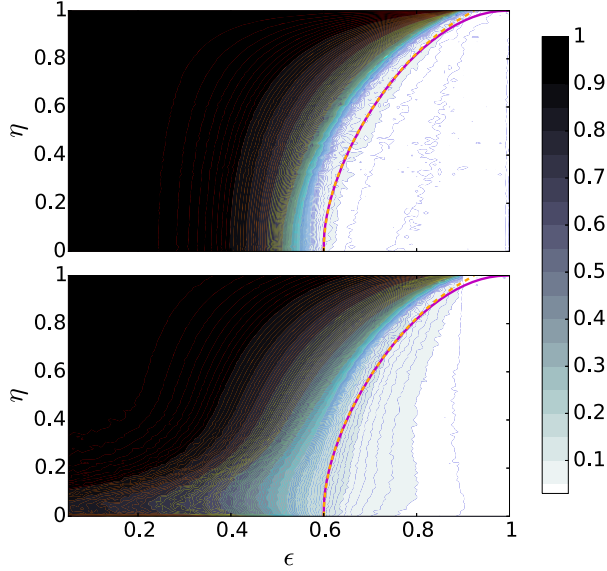


FIG. 3. The overlap for (top) belief propagation and (bottom) our spectral algorithm. The detectability transition in Eq. (10) for $T = \infty$ is shown as a solid line. The dashed curve shows the detectability transition for $T = 40$; the magenta curve shows the transition for $T = \infty$. Each point shows the average over 100 dynamic networks generated by our model with $n = 512$, $T = 40$, $k = 2$ groups, and average degree $c = 16$. The overlap here is calculated by averaging the maximum overlap at each time slot over all permutations. This maximization step implies that the expected overlap in the undetectable region is $O(n^{-1/2})$, and this produces a small deviation away from overlap = 0 in our numerical experiments.

above, for some larger values of k we expect this transition to become first order, with the accuracy jumping to zero discontinuously.) Moreover, Fig. 4 shows that the convergence time of BP, i.e., the number of times we need to

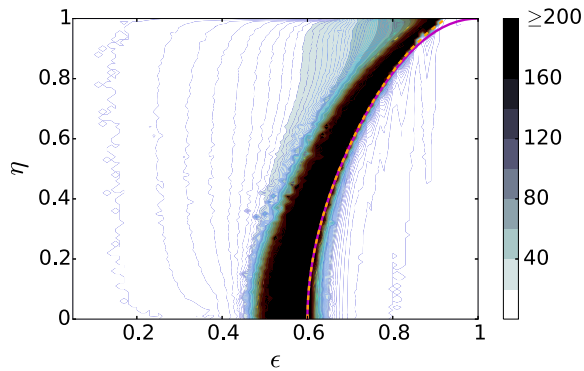


FIG. 4. The convergence time of belief propagation diverges as we approach the transition. This heat map shows the number of iterations it takes BP to converge to a fixed point, with the same parameters as in Fig. 3. As before, the dashed curve shows the detectability transition for $T = 40$, and the magenta curve shows the transition for $T = \infty$.

iterate the update equations to reach a fixed point, diverges near the critical curve.

VII. CONCLUSION

Although we now have a rigorous theoretical understanding of the strengths, weaknesses, and limits of community detection in static networks, comparable theoretical insights for networks with more general structures have been lacking. Here, we derive a mathematically precise understanding of the limits of detectability for communities in dynamic networks, under a model in which group memberships are correlated over time but where the edges at each time are generated independently. Using the cavity method, we generalize the static-case detectability limit to depend both on the strength of the communities and on the rate at which community membership changes over time. Because this model is an analytically tractable special case of several previously published models, we expect qualitatively similar behavior to occur within more elaborate models of dynamic networks, including those where the edges at successive times are correlated.

Our two efficient algorithms for detecting communities in dynamic networks, one based on belief propagation and one on spectral clustering, are optimal in the sense that they succeed all the way down to the detectability threshold. Furthermore, the belief propagation algorithm is asymptotically optimal in terms of its accuracy, implying that no algorithm can perform better at detecting communities in dynamic networks in which edges are generated independently at each time step.

We believe that all of our results can be made rigorous, at least for two groups, using the methods of Refs. [16, 48–50]. We also note that the mathematical tools we introduce to obtain our results for dynamic networks are quite general, and could be used to obtain similar results for other, more general types of networks. Examples of such future directions include the case where the matrix p of connection probabilities changes over time (a situation similar to change-point detection in networks [58]), or where edges are persistent across time [1], or where networks have edge weights [13] or additional metadata on the nodes [19–22]. The latter represents a particularly interesting case, as recent numerical results by Newman and Clauset [22] suggest that metadata on the nodes may also serve to shift the location of the detectability threshold in static networks.

ACKNOWLEDGMENTS

The authors thank Elchanan Mossel and Andrey Lokhov for helpful conversations, and acknowledge the BioFrontiers Computing Core at the University of Colorado Boulder for providing High Performance Computing resources (NIH 1S10OD012300) supported by BioFrontiers IT. This work was supported in part by

Grant No. IIS-1452718 (A. G., A. C.) from the National Science Foundation, Grant No. FA9550-12-1-0432 from the U.S. Air Force Office of Scientific Research (AFOSR) and the Defense Advanced Research Projects Agency (DARPA) (L. P.), Contract No. W911NF-12-R-0012 from the Army Research Office (ARO) (C. M.), and the John Templeton Foundation (P. Z., C. M.). A. G. and P. Z. are joint principal authors, with the remaining authors appearing alphabetically.

-
- [1] A. Clauset and N. Eagle, *Persistence and Periodicity in a Dynamic Proximity Network*, in *Proceedings of the DIMACS Workshop on Computational Methods for Dynamic Interaction Networks* (Rutgers University, Piscataway, NJ, 2007).
 - [2] T. Berger-Wolf, C. Tantipathananandh, and D. Kempe, *Dynamic community identification*, in *Link Mining: Models, Algorithms, and Applications* (Springer, New York, 2010), pp. 307–336.
 - [3] 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* **9**, e86028 (2014).
 - [4] M. Kim and J. Leskovec, *Nonparametric Multi-Group Membership Model for Dynamic Networks*, in *Advances in Neural Information Processing Systems 26*, edited by C. Burges (Curran Associates, Inc., Lake Tahoe, Nevada, 2013), pp. 1385–1393.
 - [5] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, *Community Structure in Time-Dependent, Multiscale, and Multiplex Networks*, *Science* **328**, 876 (2010).
 - [6] R. Rossi, B. Gallagher, J. Neville, and K. Henderson, Lawrence Livermore National Laboratory Technical Report No. LLNL-TR-514271, 2011.
 - [7] E. P. Xing, W. Fu, and L. Song, *A State-Space Mixed Membership Blockmodel for Dynamic Network Tomography*, *Ann. Appl. Stat.* **4**, 535 (2010).
 - [8] L. Zhu, D. Guo, J. Yin, G. V. Steeg, and A. Galstyan, *Scalable Link Prediction in Dynamic Networks via Non-Negative Matrix Factorization*, [arXiv:1411.3675](https://arxiv.org/abs/1411.3675).
 - [9] A. Goldenberg, A. X. Zheng, S. E. Fienberg, and E. M. Airolidi, *A Survey of Statistical Network Models*, *Found. Trends Mach. Learn.* **2**, 129 (2010).
 - [10] M. B. Hastings, *Community Detection as an Inference Problem*, *Phys. Rev. E* **74**, 035102 (2006).
 - [11] J. M. Hofman and C. H. Wiggins, *Bayesian Approach to Network Modularity*, *Phys. Rev. Lett.* **100**, 258701 (2008).
 - [12] B. Karrer and M. E. J. Newman, *Stochastic Blockmodels and Community Structure in Networks*, *Phys. Rev. E* **83**, 016107 (2011).
 - [13] C. Aicher, A. Z. Jacobs, and A. Clauset, *Learning Latent Block Structure in Weighted Networks*, *J. Complex Netw.* **3**, 221 (2015).
 - [14] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, *Asymptotic Analysis of the Stochastic Block Model for Modular Networks and Its Algorithmic Applications*, *Phys. Rev. E* **84**, 066106 (2011).
 - [15] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, *Inference and Phase Transitions in the Detection of Modules in Sparse Networks*, *Phys. Rev. Lett.* **107**, 065701 (2011).
 - [16] E. Mossel, J. Neeman, and A. Sly, *Reconstruction and Estimation in the Planted Partition Model*, *Probab. Theory Relat. Fields* **162**, 431 (2015).
 - [17] A. C. Thomas and J. K. Blitzer, *Valued Ties Tell Fewer Lies: Why not to Dichotomize Network Edges with Thresholds*, [arXiv:1101.0788](https://arxiv.org/abs/1101.0788).
 - [18] M. E. J. Newman, *Mixing Patterns in Networks*, *Phys. Rev. E* **67**, 026126 (2003).
 - [19] L. Peel, *Topological Feature Based Classification*, in *Proceedings of the 14th International Conference on Information Fusion (FUSION)* (IEEE, New York, 2011), pp. 1–8.
 - [20] L. Peel, *Supervised Blockmodelling*, in *Proceedings of the ECML/PKDD Workshop on Collective Learning and Inference on Structured Data*, <http://www.cse.iitm.ac.in/CoLISD/CoLISD.html>.
 - [21] P. Zhang, C. Moore, and L. Zdeborová, *Phase Transitions in Semisupervised Clustering of Sparse Networks*, *Phys. Rev. E* **90**, 052802 (2014).
 - [22] M. E. J. Newman and A. Clauset, *Structure and Inference in Annotated Networks*, *Nat. Commun.* **7**, 11863 (2016).
 - [23] D. S. Bassett, M. A. Porter, N. F. Wymbs, S. T. Grafton, J. M. Carlson, and P. J. Mucha, *Robust Detection of Dynamic Community Structure in Networks*, *Chaos* **23**, 013142 (2013).
 - [24] M. Bazzi, M. A. Porter, S. Williams, M. McDonald, D. J. Fenn, and S. D. Howison, *Community Detection in Temporal Multilayer Networks, with an Application to Correlation Networks*, *Multiscale Model. Simul.* **14**, 1 (2016).
 - [25] E. Acar, D. M. Dunlavy, and T. G. Kolda, *Link Prediction on Evolving Data Using Matrix and Tensor Factorizations*, in *Proceedings of the ICDM'09 Workshop on Large Scale Data Mining Theory and Applications (LDMTA'09)* (IEEE, New York, 2009), pp. 262–269.
 - [26] D. M. Dunlavy, T. G. Kolda, and E. Acar, *Temporal Link Prediction Using Matrix and Tensor Factorizations*, *ACM Trans. Knowl. Discovery Data* **5**, 10 (2011).
 - [27] J. Sun, C. Faloutsos, S. Papadimitriou, and P. S. Yu, *GraphScope: Parameter-Free Mining of Large Time-Evolving Graphs*, in *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'07)* (ACM, San Jose, CA, 2007), pp. 687–696.
 - [28] M. Rosvall and C. T. Bergstrom, *Mapping Change in Large Networks*, *PLoS One* **5**, e8694 (2010).
 - [29] T. Yang, Y. Chi, S. Zhu, Y. Gong, and R. Jin, *A Bayesian Approach Toward Finding Communities and Their Evolutions in Dynamic Social Networks*, in *Proceedings of the SIAM International Conference on Data Mining (SDM'09)* (SIAM, Sparks, Nevada, 2009), pp. 990–1001.
 - [30] K. S. Xu and A. O. Hero, *Dynamic Stochastic Blockmodels for Time-Evolving Social Networks*, *IEEE J. Sel. Top. Signal Process.* **8**, 552 (2014).
 - [31] Q. Han, K. S. Xu, and E. M. Airolidi, *Consistent Estimation of Dynamic and Multi-Layer Networks*, in *Proceedings of the 32nd International Conference on Machine Learning (JMLR Workshop and Conference Proceedings (JMLR: W&CP), Lille, 2015)*.

- [32] T. P. Peixoto, *Inferring the Mesoscale Structure of Layered, Edge-Valued, and Time-Varying Networks*, *Phys. Rev. E* **92**, 042807 (2015).
- [33] T. Vallès-Català, F. A. Massucci, R. Guimerà, and M. Sales-Pardo, *Multilayer Stochastic Block Models Reveal the Multilayer Structure of Complex Networks*, *Phys. Rev. X* **6**, 011036 (2016).
- [34] C. Aggarwal and K. Subbian, *Evolutionary Network Analysis: A Survey*, *Comput. Surv.* **47**, 10 (2014).
- [35] T. Hartmann, A. Kappes, and D. Wagner, *Clustering Evolving Networks*, *arXiv:1401.3516*.
- [36] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas, *Diffusion Dynamics on Multiplex Networks*, *Phys. Rev. Lett.* **110**, 028701 (2013).
- [37] V. Nicosia, G. Bianconi, V. Latora, and M. Barthélemy, *Growing Multiplex Networks*, *Phys. Rev. Lett.* **111**, 058701 (2013).
- [38] A. Solé-Ribalta, S. Gómez, and A. Arenas, *Congestion Induced by the Structure of Multiplex Networks*, *Phys. Rev. Lett.* **116**, 108701 (2016).
- [39] Y. Hu, S. Havlin, and H. A. Makse, *Conditions for Viral Influence Spreading through Multiplex Correlated Social Networks*, *Phys. Rev. X* **4**, 021031 (2014).
- [40] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivela, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, *Mathematical Formulation of Multilayer Networks*, *Phys. Rev. X* **3**, 041022 (2013).
- [41] M. De Domenico, A. Lancichinetti, A. Arenas, and M. Rosvall, *Identifying Modular Flows on Multilayer Networks Reveals Highly Overlapping Organization in Interconnected Systems*, *Phys. Rev. X* **5**, 011027 (2015).
- [42] E. Valdano, L. Ferreri, C. Poletto, and V. Colizza, *Analytical Computation of the Epidemic Threshold on Temporal Networks*, *Phys. Rev. X* **5**, 021005 (2015).
- [43] L. K. Gallos, D. Rybski, F. Liljeros, S. Havlin, and H. A. Makse, *How People Interact in Evolving Online Affiliation Networks*, *Phys. Rev. X* **2**, 031014 (2012).
- [44] C. Granell, S. Gómez, and A. Arenas, *Dynamical Interplay between Awareness and Epidemic Spreading in Multiplex Networks*, *Phys. Rev. Lett.* **111**, 128701 (2013).
- [45] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Avalanche Collapse of Interdependent Networks*, *Phys. Rev. Lett.* **109**, 248701 (2012).
- [46] P. W. Holland, K. B. Laskey, and S. Leinhardt, *Stochastic Blockmodels: First Steps*, *Soc. Networks* **5**, 109 (1983).
- [47] K. Nowicki and T. A. B. Snijders, *Estimation and Prediction for Stochastic Blockstructures*, *J. Am. Stat. Assoc.* **96**, 1077 (2001).
- [48] E. Mossel, J. Neeman, and A. Sly, *Belief Propagation, Robust Reconstruction and Optimal Recovery of Block Models*, in *Proceedings of the 27th Conference on Learning Theory* (2014), pp. 356–370.
- [49] L. Massoulié, *Community Detection Thresholds and the Weak Ramanujan Property*, in *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC'14)* (ACM, New York, NY, 2014), pp. 694–703.
- [50] C. Bordenave, M. Lelarge, and L. Massoulié, *Non-Backtracking Spectrum of Random Graphs: Community Detection and Non-Regular Ramanujan Graphs*, in *Proceedings of the 56th Annual Symposium on the Foundations of Computer Science (IEEE, New York, 2015)*, pp. 1347–1357.
- [51] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang, *Spectral Redemption in Clustering Sparse Networks*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20935 (2013).
- [52] S. Janson and E. Mossel, *Robust Reconstruction on Trees Is Determined by the Second Eigenvalue*, *Ann. Probab.* **32**, 2630 (2004).
- [53] J. R. L. de Almeida and D. J. Thouless, *Stability of the Sherrington-Kirkpatrick Solution of a Spin Glass Model*, *J. Phys. A* **11**, 983 (1978).
- [54] E. Abbe and C. Sandon, *Detection in the Stochastic Block Model with Multiple Clusters: Proof of the Achievability Conjectures, Acyclic BP, and the Information-Computation Gap*, *arXiv:1512.09080*.
- [55] J. Banks, C. Moore, J. Neeman, and P. Netrapalli, *Information-Theoretic Thresholds for Community Detection in Sparse Networks*, in *Proceedings of the 29th Conference on Learning Theory* (to be published).
- [56] Y. Iba, *The Nishimori Line and Bayesian Statistics*, *J. Phys. A* **32**, 3875 (1999).
- [57] P. Zhang, F. Krzakala, J. Reichardt, and L. Zdeborová, *Comparative Study for Inference of Hidden Classes in Stochastic Block Models*, *J. Stat. Mech.* (2012) P12021.
- [58] L. Peel and A. Clauset, *Detecting Change Points in the Large-Scale Structure of Evolving Networks*, in *Proceedings of the 29th International Conference on Artificial Intelligence (AAAI)* (AAAI, Palo Alto, CA, 2015), pp. 2914–2920.