Phase Transitions in Spectral Community Detection of Large Noisy Networks

Pin-Yu Chen and Alfred O. Hero III, *Fellow*, *IEEE*

Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, USA

Email : {pinyu,hero}@umich.edu

Abstract—In this paper, we study the sensitivity of the spectral clustering based community detection algorithm subject to a Erdos-Renyi type random noise model. We prove phase transitions in community detectability as a function of the external edge connection probability and the noisy edge presence probability under a general network model where two arbitrarily connected communities are interconnected by random external edges. Specifically, the community detection performance transitions from almost perfect detectability to low detectability as the inter-community edge connection probability exceeds some critical value. We derive upper and lower bounds on the critical value and show that the bounds are identical when the two communities have the same size. The phase transition results are validated using network simulations. Using the derived expressions for the phase transition threshold we propose a method for estimating this threshold from observed data.

Index Terms—community detectability, noisy graph

I. INTRODUCTION

Community detection is a graph signal processing problem [1]–[9] where the goal is to cluster the nodes on a graph into different communities by inspecting the connectivity structure of the graph. Consider an undirected regular graph consisting of two node-disjoint communities interconnected by some external edges. Let n denote the total number of nodes in the network. The network topology can be characterized by its symmetric adjacency matrix $\mathbf{A}$, where $\mathbf{A}$ is an $n \times n$ matrix, with $\mathbf{A}_{ij} = 1$ if an edge exists between nodes i and j, and $\mathbf{A}_{ij} = 0$ otherwise.

Since community detection can be viewed as a graph partitioning problem that can be solved by identifying the graph cut that correctly separates the communities, spectral clustering [10], [11] approaches to community detection are natural [12]–[15]. Spectral clustering specifies a graph cut by inspecting the eigenstructure of the graph. Let $\mathbf{1}_n(\mathbf{0}_n)$ be the n-dimensional all-one (all-zero) vector. Define $\mathbf{L} = \mathbf{D} - \mathbf{A}$ as the graph Laplacian matrix of the graph, where $\mathbf{D} = \text{diag}(\mathbf{A}\mathbf{1}_n)$ is the diagonal degree matrix. Let $\lambda_i(\mathbf{L})$ denote the i-th smallest eigenvalue of $\mathbf{L}$. It is well-known that $\lambda_1(\mathbf{L}) = 0$ since $\mathbf{L}\mathbf{1}_n = \mathbf{0}_n$ and $\mathbf{L}$ is a positive semidefinite (PSD) matrix [16], [17]. The second smallest eigenvalue, $\lambda_2(\mathbf{L})$, is known as the algebraic connectivity. The eigenvector associated with $\lambda_2(\mathbf{L})$ is called the Fiedler vector [18]. A mathematical representation of the algebraic connectivity is

$$\lambda_2(\mathbf{L}) = \min_{\|\mathbf{x}\|_2 = 1, \mathbf{1}_n^T \mathbf{x} = 0} \mathbf{x}^T \mathbf{L} \mathbf{x}. \tag{1}$$

The principle underlying spectral clustering for community detection [12]–[15] is summarized as follows:

1) Compute the graph Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$.
2) Compute the Fiedler vector $\mathbf{y}$.
3) Perform K-means clustering [19] on the entries of $\mathbf{y}$ to cluster the nodes into two groups. To detect more than two

communities, we can use successive spectral clustering on the discovered communities [1], [20].

Most literature on community detectability [21]–[28] focuses on the noiseless setting where the edges are not subject to random insertions or deletions. However, in practice the network data can be corrupted by incorrect measurements or background noises (e.g., bio-informatics data) that can produce such random insertions and deletions. Consequently, analyzing the sensitivity of community detection algorithms to noise is an important task. In this paper, we prove the existence of abrupt phase transitions in community detectability for spectral community detection under a Erdos-Renyi type random noise model. Our network model includes the widely used stochastic block model [29] as a special case. We show that at some critical value of random external edge connection probability the community detection performance transitions from almost perfect detectability to low detectability in the large network limit (large n). We provide asymptotic upper and lower bounds on this critical value. The bounds become equal to each other when these two community sizes are identical. This framework can be generalized to community detection on more than two communities by aggregating multiple communities into two larger communities.

We use simulated networks to validate the asymptotic expressions for the phase transitions. Using our theory, we propose an empirical estimator of the critical phase transition threshold that can be applied to data. These empirical estimates are used to test whether the detector is operating in a reliable detection regime, i.e., below the phase transition threshold.

II. NETWORK MODEL AND RELATED WORKS

Consider two arbitrarily connected communities with internal adjacency matrices $\mathbf{A}_{S_1}$ and $\mathbf{A}_{S_2}$ and network sizes n_1 and n_2, respectively. The external connections between these two communities are characterized by an $n_1 \times n_2$ adjacency matrix $\mathbf{C}_S$, where each entry in $\mathbf{C}_S$ is a Bernoulli(p) random variable. Let $n = n_1 + n_2$. The overall $n \times n$ adjacency matrix of the community structure can be represented as

$$\mathbf{A}_S = \begin{bmatrix} \mathbf{A}_{S_1} & \mathbf{C}_S \\ \mathbf{C}_S^T & \mathbf{A}_{S_2} \end{bmatrix}. \tag{2}$$

The widely used stochastic block model [29] is a special case of (2) when the two community structures are generated by connected Erdos-Renyi random graphs parameterized by the within-community connection probability p_i ($i = 1, 2$). Our network model is more general since we only assume random connection probability p on the external edges and we allow the within-community adjacency matrices $\mathbf{A}_{Si}$ to be arbitrary. In this paper we consider the noisy setting in which the adjacency matrix $\mathbf{A}_S$ is corrupted by a random adjacency matrix $\mathbf{A}_N$ such that the observed adjacency matrix is $\mathbf{A} = \mathbf{A}_S + \mathbf{A}_N$. The adjacency matrix $\mathbf{A}_N$ is generated by a Erdos-Renyi random graph with edge connection probability q. Note that this model only allows random insertions and not deletions of edges.

Community detectability has been studied under the stochastic block model with restricted assumptions such as $n_1 = n_2$, $p_1 = p_2$ and fixed average degree as the network size n increases [23]–[26],

[30]. The planted clique detection problem in [31] is a special case of the stochastic block model when $p_1 = 1$ and $p_2 = p$. A less restricted stochastic block model is studied in [28] where a universal phase transition in community detectability is established for which the critical value does not depend on the community sizes. A similar model to our network model is studied in [32] for interconnected networks. However, in [32] the subnetworks are of equal size and the external edges are known (i.e., non-random). Phase transitions in spectral community detection under noiseless network setting is studied in [27].

III. PHASE TRANSITION ANALYSIS

Let $\mathbf{1}_{n_i}$ be the n_i-dimensional all-one vector and let $\mathbf{D}_{S_1} = \mathrm{diag}\left(\mathbf{C}_S \mathbf{1}_{n_2}\right)$ and $\mathbf{D}_{S_2} = \mathrm{diag}\left(\mathbf{C}_S^T \mathbf{1}_{n_1}\right)$. The graph Laplacian matrix of the noiseless graph can be represented as

$$\mathbf{L}_S = \begin{bmatrix} \mathbf{L}_{S_1} + \mathbf{D}_{S_1} & -\mathbf{C}_S \\ -\mathbf{C}_S^T & \mathbf{L}_{S_2} + \mathbf{D}_{S_2} \end{bmatrix}, \qquad (3)$$

where $\mathbf{L}_{Si}$ is the graph Laplacian matrix of i-th community. Similarly, the graph Laplacian matrix of the noise matrix can be represented as

$$\mathbf{L}_N = \begin{bmatrix} \mathbf{L}_{N_1} + \mathbf{D}_{N_1} & -\mathbf{C}_N \\ -\mathbf{C}_N^T & \mathbf{L}_{N_2} + \mathbf{D}_{N_2} \end{bmatrix}, \qquad (4)$$

where $\mathbf{L}_{Ni}$ is the graph Laplacian matrix of the noise matrix in i-th community, $\mathbf{C}_N$ is the adjacency matrix of noisy edges between two communities, $\mathbf{D}_{N_1} = \mathrm{diag}\left(\mathbf{C}_N \mathbf{1}_{n_2}\right)$ and $\mathbf{D}_{N_2} = \mathrm{diag}\left(\mathbf{C}_N^T \mathbf{1}_{n_1}\right)$. Therefore the overall graph Laplacian matrix is $\mathbf{L} = \mathbf{L}_S + \mathbf{L}_N$.

Let $\mathbf{x} = [\mathbf{x}_1 \ \mathbf{x}_2]^T$, where $\mathbf{x}_1 \in \mathbb{R}^{n_1}$ and $\mathbf{x}_2 \in \mathbb{R}^{n_2}$. By (1) we have $\lambda_2(\mathbf{L}) = \min_{\mathbf{x}} \mathbf{x}^T \mathbf{L} \mathbf{x}$ subject to the constraints $\mathbf{x}_1^T \mathbf{x}_1 + \mathbf{x}_2^T \mathbf{x}_2 = 1$ and $\mathbf{x}_1^T \mathbf{1}_{n_1} + \mathbf{x}_2^T \mathbf{1}_{n_2} = 0$. Using Lagrange multipliers μ, ν and (3), the Fiedler vector $\mathbf{y} = [\mathbf{y}_1 \ \mathbf{y}_2]^T$ of $\mathbf{L}$, with $\mathbf{y}_1 \in \mathbb{R}^{n_1}$ and $\mathbf{y}_1 \in \mathbb{R}^{n_2}$, satisfies $\mathbf{y} = \arg\min_{\mathbf{x}} \Gamma(\mathbf{x})$, where

$$\begin{aligned} \Gamma(\mathbf{x}) = & \ \mathbf{x}_1^T (\mathbf{L}_{S_1} + \mathbf{D}_{S_1} + \mathbf{L}_{N_1} + \mathbf{D}_{N_1}) \mathbf{x}_1 - 2\mathbf{x}_1^T (\mathbf{C}_S + \mathbf{C}_N) \mathbf{x}_2 \\ & + \mathbf{x}_2^T (\mathbf{L}_{S_2} + \mathbf{D}_{S_2} + \mathbf{L}_{N_2} + \mathbf{D}_{N_2}) \mathbf{x}_2 \\ & - \mu(\mathbf{x}_1^T \mathbf{x}_1 + \mathbf{x}_2^T \mathbf{x}_2 - 1) - \nu(\mathbf{x}_1^T \mathbf{1}_{n_1} + \mathbf{x}_2^T \mathbf{1}_{n_2}). \end{aligned} \qquad (5)$$

Differentiating (5) with respect to $\mathbf{x}_1$ and $\mathbf{x}_2$ respectively, and substituting $\mathbf{y}$ to the equations, we obtain

$$\begin{aligned} & 2(\mathbf{L}_{S_1} + \mathbf{D}_{S_1} + \mathbf{L}_{N_1} + \mathbf{D}_{N_1}) \mathbf{y}_1 - 2(\mathbf{C}_S + \mathbf{C}_N) \mathbf{y}_2 - 2\mu \mathbf{y}_1 - \nu \mathbf{1}_{n_1} \\ & = \mathbf{0}_{n_1}, \qquad (6) \end{aligned}$$

$$\begin{aligned} & 2(\mathbf{L}_{S_2} + \mathbf{D}_{S_2} + \mathbf{L}_{N_2} + \mathbf{D}_{N_2}) \mathbf{y}_2 - 2(\mathbf{C}_S + \mathbf{C}_N)^T \mathbf{y}_1 - 2\mu \mathbf{y}_2 - \nu \mathbf{1}_{n_2} \\ & = \mathbf{0}_{n_2}. \qquad (7) \end{aligned}$$

Left multiplying (6) by $\mathbf{1}_{n_1}^T$ and left multiplying (7) by $\mathbf{1}_{n_2}^T$, we have

$$\begin{aligned} & 2\mathbf{1}_{n_1}^T (\mathbf{D}_{S_1} + \mathbf{D}_{N_1}) \mathbf{y}_1 - 2\mathbf{1}_{n_1}^T (\mathbf{C}_S + \mathbf{C}_N) \mathbf{y}_2 - 2\mu \mathbf{1}_{n_1}^T \mathbf{y}_1 - \nu n_1 \\ & = 0, \qquad (8) \end{aligned}$$

$$\begin{aligned} & 2\mathbf{1}_{n_2}^T (\mathbf{D}_{S_2} + \mathbf{D}_{N_2}) \mathbf{y}_2 - 2\mathbf{1}_{n_2}^T (\mathbf{C}_S + \mathbf{C}_N)^T \mathbf{y}_1 - 2\mu \mathbf{1}_{n_2}^T \mathbf{y}_2 - \nu n_2 \\ & = 0. \qquad (9) \end{aligned}$$

Since by definition $\mathbf{1}_{n_1}^T \mathbf{D}_{S_1} = \mathbf{1}_{n_2}^T \mathbf{C}_S^T$, $\mathbf{1}_{n_1}^T \mathbf{C}_S = \mathbf{1}_{n_2}^T \mathbf{D}_{S_2}$, $\mathbf{1}_{n_1}^T \mathbf{D}_{N_1} = \mathbf{1}_{n_2}^T \mathbf{C}_N^T$ and $\mathbf{1}_{n_1}^T \mathbf{C}_N = \mathbf{1}_{n_2}^T \mathbf{D}_{N_2}$, adding (8) and (9) we obtain $\nu = -\frac{2\mu}{n}(\mathbf{y}_1^T \mathbf{1}_{n_1} + \mathbf{y}_2^T \mathbf{1}_{n_2}) = 0$ by the fact that the Fiedler vector $\mathbf{y}$ has the property $\mathbf{y}^T \mathbf{1} = 0$. Applying $\nu = 0$ and left multiplying (6) by $\mathbf{y}_1^T$ and left multiplying (7) by $\mathbf{y}_2^T$, we have

$$\begin{aligned} & \mathbf{y}_1^T (\mathbf{L}_{S_1} + \mathbf{D}_{S_1} + \mathbf{L}_{N_1} + \mathbf{D}_{N_1}) \mathbf{y}_1 - \mathbf{y}_1^T (\mathbf{C}_S + \mathbf{C}_N) \mathbf{y}_2 - \mu \mathbf{y}_1^T \mathbf{y}_1 \\ & = 0, \qquad (10) \end{aligned}$$

$$\begin{aligned} & \mathbf{y}_2^T (\mathbf{L}_{S_2} + \mathbf{D}_{S_2} + \mathbf{L}_{N_2} + \mathbf{D}_{N_2}) \mathbf{y}_2 - \mathbf{y}_2^T (\mathbf{C}_S + \mathbf{C}_N)^T \mathbf{y}_1 - \mu \mathbf{y}_2^T \mathbf{y}_2 \\ & = 0. \qquad (11) \end{aligned}$$

Adding (10) and (11) and by (1) and (3) we obtain $\mu = \lambda_2(\mathbf{L})$.

Let $\bar{\mathbf{C}}_S = p\mathbf{1}_{n_1}\mathbf{1}_{n_2}^T$, a matrix whose elements are the means of entries in $\mathbf{C}_S$. Let $\sigma_i(\mathbf{M})$ denote the i-th largest singular value of a rectangular matrix $\mathbf{M}$[1] and write $\mathbf{C}_S = \bar{\mathbf{C}}_S + \boldsymbol{\Delta}_S$, where $\boldsymbol{\Delta}_S = \mathbf{C}_S - \bar{\mathbf{C}}_S$. By Latala's theorem [33], $\mathbb{E}\left[\sigma_1\left(\frac{\boldsymbol{\Delta}_S}{\sqrt{n_1 n_2}}\right)\right] \to 0$. This is proved in Appendix VII-A of [27]. Furthermore, by Talagrand's concentration inequality [34], almost surely,

$$\sigma_1\left(\frac{\mathbf{C}_S}{\sqrt{n_1 n_2}}\right) \to p; \ \sigma_i\left(\frac{\mathbf{C}_S}{\sqrt{n_1 n_2}}\right) \to 0 \ \forall \ i \geq 2 \qquad (12)$$

when $n_1, n_2 \to \infty$ and $\frac{n_1}{n_2} \to c > 0$. This is proved in Appendix VII-B of [27]. Note that the convergence rate is maximal when $n_1 = n_2$ because $n_1 + n_2 \geq 2\sqrt{n_1 n_2}$ and the equality holds if $n_1 = n_2$. Similarly, let $\bar{\mathbf{C}}_N = q\mathbf{1}_{n_1}\mathbf{1}_{n_2}^T$, a matrix whose elements are the means of entries in $\mathbf{A}_N$. We have $\sigma_1\left(\frac{\mathbf{C}_N}{\sqrt{n_1 n_2}}\right) \to q$ and $\sigma_i\left(\frac{\mathbf{C}_N}{\sqrt{n_1 n_2}}\right) \to 0 \ \forall \ i \geq 2$ when $n_1, n_2 \to \infty$ and $\frac{n_1}{n_2} \to c > 0$.

As proved in [35], the singular vectors of $\mathbf{C}_S$ ($\mathbf{C}_N$) and $\bar{\mathbf{C}}_S$ ($\bar{\mathbf{C}}_N$) are close to each other in the sense that the squared inner product of their left/right singular vectors converges to 1 almost surely when $\sqrt{n_1 n_2} p \to \infty$ ($\sqrt{n_1 n_2} q \to \infty$). Consequently, we have, almost surely,

$$\frac{(\mathbf{D}_{S_1} + \mathbf{D}_{N_1})\mathbf{1}_{n_1}}{n_2} = \frac{(\mathbf{C}_S + \mathbf{C}_N)\mathbf{1}_{n_2}}{n_2} \to (p+q)\mathbf{1}_{n_1}; \qquad (13)$$

$$\frac{(\mathbf{D}_{S_2} + \mathbf{D}_{N_2})\mathbf{1}_{n_2}}{n_1} = \frac{(\mathbf{C}_S + \mathbf{C}_N)^T\mathbf{1}_{n_1}}{n_1} \to (p+q)\mathbf{1}_{n_2}. \qquad (14)$$

Applying (12), (13) and (14) to (8) and (9) and recalling that $\nu = 0$ and $\frac{n_1}{n_2} = c > 0$, we have, almost surely,

$$\frac{1}{\sqrt{c}}(p+q)\mathbf{1}_{n_1}^T \mathbf{y}_1 - \sqrt{c}(p+q)\mathbf{1}_{n_2}^T \mathbf{y}_2 - \frac{\mu \mathbf{1}_{n_1}^T \mathbf{y}_1}{\sqrt{n_1 n_2}} \to 0; \qquad (15)$$

$$\sqrt{c}(p+q)\mathbf{1}_{n_2}^T \mathbf{y}_2 - \frac{1}{\sqrt{c}}(p+q)\mathbf{1}_{n_1}^T \mathbf{y}_1 - \frac{\mu \mathbf{1}_{n_2}^T \mathbf{y}_2}{\sqrt{n_1 n_2}} \to 0. \qquad (16)$$

By the fact that $\mathbf{1}_{n_1}^T \mathbf{y}_1 + \mathbf{1}_{n_2}^T \mathbf{y}_2 = 0$, we have, almost surely,

$$\left(\sqrt{c} + \frac{1}{\sqrt{c}}\right)\left(p + q - \frac{\mu}{n}\right)\mathbf{1}_{n_1}^T \mathbf{y}_1 \to 0; \qquad (17)$$

$$\left(\sqrt{c} + \frac{1}{\sqrt{c}}\right)\left(p + q - \frac{\mu}{n}\right)\mathbf{1}_{n_2}^T \mathbf{y}_2 \to 0. \qquad (18)$$

Consequently, as $\mu = \lambda_2(\mathbf{L})$, at least one of the two cases have to be satisfied:

$$\text{Case 1: } \frac{\lambda_2(\mathbf{L})}{n} \xrightarrow{\text{a.s.}} p + q =: t, \qquad (19)$$

$$\text{Case 2: } \mathbf{1}_{n_1}^T \mathbf{y}_1 \to 0 \ \text{ and } \ \mathbf{1}_{n_2}^T \mathbf{y}_2 \to 0 \ \text{ almost surely.} \qquad (20)$$

We will show that the algebraic connectivity $\lambda_2(\mathbf{L})/n$ and the Fiedler vector $\mathbf{y}$ undergo a phase transition between Case 1 and Case 2 as a function of $t = p+q$. That is, a transition from Case 1 to Case 2 occurs when p exceeds a certain threshold p^*. In Case 1, observe that asymptotically $\frac{\lambda_2(\mathbf{L})}{n}$ grows linearly with t while the asymptotic Fiedler vector remains the same (unique up to its sign). Furthermore, from (10), (11), (12), (19), $\mu = \lambda_2(\mathbf{L})$ and $\mathbf{1}_{n_1}^T \mathbf{y}_1 + \mathbf{1}_{n_2}^T \mathbf{y}_2 = 0$, the Fiedler vector $\mathbf{y}$ in Case 1 has the following property. Almost surely,

$$\frac{\mathbf{y}_1^T(\mathbf{L}_{S_1} + \mathbf{L}_{N_1})\mathbf{y}_1}{\sqrt{n_1 n_2}} + \frac{p+q}{\sqrt{n_1 n_2}}(\mathbf{1}_{n_1}^T \mathbf{y}_1)^2 - \sqrt{c}(p+q)\mathbf{y}_1^T \mathbf{y}_1 \to 0, \qquad (21)$$

$$\frac{\mathbf{y}_2^T(\mathbf{L}_{S_2} + \mathbf{L}_{N_2})\mathbf{y}_2}{\sqrt{n_1 n_2}} + \frac{p+q}{\sqrt{n_1 n_2}}(\mathbf{1}_{n_1}^T \mathbf{y}_1)^2 - \frac{1}{\sqrt{c}}(p+q)\mathbf{y}_2^T \mathbf{y}_2 \to 0. \qquad (22)$$

---

[1]Note that for convenience, we use $\lambda_i(\mathbf{M}_1)$ to denote the i-th smallest eigenvalue of a square matrix $\mathbf{M}_1$ and use $\sigma_i(\mathbf{M}_2)$ to denote the i-th largest singular value of a rectangular matrix $\mathbf{M}_2$.

Adding (21) and (22), we have

$$\frac{1}{\sqrt{n_1 n_2}}\left[\mathbf{y}_1^T(\mathbf{L}_{S_1}+\mathbf{L}_{N_1})\mathbf{y}_1+\mathbf{y}_2^T(\mathbf{L}_{S_2}+\mathbf{L}_{N_2})\mathbf{y}_2\right]+$$
$$\left[\frac{2(\mathbf{1}_{n_1}^T\mathbf{y}_1)^2}{\sqrt{n_1 n_2}}-\left(\sqrt{c}\mathbf{y}_1^T\mathbf{y}_1+\frac{1}{\sqrt{c}}\mathbf{y}_2^T\mathbf{y}_2\right)\right](p+q)\xrightarrow{\text{a.s.}}0. \quad (23)$$

As the two bracketed terms in (23) converge to finite constants for all $t=p+q$ in Case 1; almost surely,

$$\frac{1}{\sqrt{n_1 n_2}}\left[\mathbf{y}_1^T(\mathbf{L}_{S_1}+\mathbf{L}_{N_1})\mathbf{y}_1+\mathbf{y}_2^T(\mathbf{L}_{S_2}+\mathbf{L}_{N_2})\mathbf{y}_2\right]\to 0; \quad (24)$$

$$\frac{2(\mathbf{1}_{n_1}^T\mathbf{y}_1)^2}{\sqrt{n_1 n_2}}-\left(\sqrt{c}\mathbf{y}_1^T\mathbf{y}_1+\frac{1}{\sqrt{c}}\mathbf{y}_2^T\mathbf{y}_2\right)\to 0. \quad (25)$$

By the PSD property of the graph Laplacian matrix, $\mathbf{y}_1^T(\mathbf{L}_{S_1}+\mathbf{L}_{N_1})\mathbf{y}_1+\mathbf{y}_2^T(\mathbf{L}_{S_2}+\mathbf{L}_{N_2})\mathbf{y}_2>0$ if and only if $\mathbf{y}_1$ and $\mathbf{y}_2$ are not constant vectors. Therefore (24) implies $\mathbf{y}_1$ and $\mathbf{y}_2$ converge to constant vectors. By the constraints $\mathbf{y}_1^T\mathbf{y}_1+\mathbf{y}_2^T\mathbf{y}_2=1$ and $\mathbf{1}_{n_1}^T\mathbf{y}_1+\mathbf{1}_{n_2}^T\mathbf{y}_2=0$, we have, almost surely,

$$\sqrt{\frac{n n_1}{n_2}}\mathbf{y}_1\to\pm\mathbf{1}_{n_1}\quad\text{and}\quad\sqrt{\frac{n n_2}{n_1}}\mathbf{y}_2\to\mp\mathbf{1}_{n_2}. \quad (26)$$

Consequently, in Case 1 $\mathbf{y}_1$ and $\mathbf{y}_2$ tend to be constant vectors with opposite signs. More importantly, (26) suggests a phase transition in spectral community detectability. In Case 1, spectral clustering can almost correctly identify these two communities since $\mathbf{y}_1$ and $\mathbf{y}_2$ are constant vectors with opposite signs. On the other hand, in Case 2, $\mathbf{1}_{n_1}^T\mathbf{y}_1\to 0$ and $\mathbf{1}_{n_2}^T\mathbf{y}_2\to 0$ almost surely. The entries of $\mathbf{y}_1$ and $\mathbf{y}_2$ tend to have opposite signs in their entries. Therefore in Case 2 spectral clustering results in very poor community detection.

IV. UPPER AND LOWER BOUNDS ON THE CRITICAL VALUE

Next we derive an upper bound on the critical value p^* of the phase transition. From (1) and (3) we know that

$$\lambda_2(\mathbf{L})=\mathbf{y}_1^T(\mathbf{L}_{S_1}+\mathbf{D}_{S_1}+\mathbf{L}_{N_1}+\mathbf{D}_{N_1})\mathbf{y}_1-2\mathbf{y}_1^T(\mathbf{C}_S+\mathbf{C}_N)\mathbf{y}_2$$
$$+\mathbf{y}_2^T(\mathbf{L}_{S_2}+\mathbf{D}_{S_2}+\mathbf{L}_{N_2}+\mathbf{D}_{N_2})\mathbf{y}_2 \quad (27)$$

subject to $\mathbf{1}_{n_1}^T\mathbf{y}_1+\mathbf{1}_{n_2}^T\mathbf{y}_2=0$ and $\mathbf{y}_1^T\mathbf{y}_1+\mathbf{y}_2^T\mathbf{y}_2=1$. In Case 2, since $\mathbf{1}_{n_1}^T\mathbf{y}_1\to 0$ and $\mathbf{1}_{n_2}^T\mathbf{y}_2\to 0$ almost surely, recalling the definition $\mathbf{\Delta}_S=\mathbf{C}_S-\bar{\mathbf{C}}_S$ and let $\mathbf{\Delta}_N=\mathbf{C}_N-\bar{\mathbf{C}}_N$,

$$\frac{\mathbf{y}_1^T(\mathbf{C}_S+\mathbf{C}_N)\mathbf{y}_2}{\sqrt{n_1 n_2}}=\frac{\mathbf{y}_1^T(\bar{\mathbf{C}}_S+\bar{\mathbf{C}}_N)\mathbf{y}_2+\mathbf{y}_1^T\mathbf{\Delta}_S\mathbf{y}_2+\mathbf{y}_1^T\mathbf{\Delta}_N\mathbf{y}_2}{\sqrt{n_1 n_2}}$$
$$\le\frac{\mathbf{y}_1^T(\bar{\mathbf{C}}_S+\bar{\mathbf{C}}_N)\mathbf{y}_2+\|\mathbf{y}_1\|_2\|\mathbf{y}_2\|_2\cdot[\sigma_1(\mathbf{\Delta}_S)+\sigma_1(\mathbf{\Delta}_N)]}{\sqrt{n_1 n_2}}$$
$$\xrightarrow{\text{a.s.}}0 \quad (28)$$

by the fact that $\sigma_1\left(\frac{\mathbf{\Delta}_S}{\sqrt{n_1 n_2}}\right)\xrightarrow{\text{a.s.}}0$ and $\sigma_1\left(\frac{\mathbf{\Delta}_N}{\sqrt{n_1 n_2}}\right)\xrightarrow{\text{a.s.}}0$ in Appendix VII-B of [27] and $\bar{\mathbf{C}}_S=p\mathbf{1}_{n_1}\mathbf{1}_{n_2}^T$ and $\bar{\mathbf{C}}_N=q\mathbf{1}_{n_1}\mathbf{1}_{n_2}^T$. Furthermore, since $\mathbf{D}_{S_1}=\text{diag}(\mathbf{C}_S\mathbf{1}_{n_2})$, $\mathbf{D}_{S_2}=\text{diag}(\mathbf{C}_S^T\mathbf{1}_{n_1})$, $\mathbf{D}_{N_1}=\text{diag}(\mathbf{C}_N\mathbf{1}_{n_2})$ and $\mathbf{D}_{N_2}=\text{diag}(\mathbf{C}_N^T\mathbf{1}_{n_1})$, (12) gives, almost surely,

$$\frac{1}{n_2}\mathbf{y}_1^T(\mathbf{D}_{S_1}+\mathbf{D}_{N_1})\mathbf{y}_1\to(p+q)\mathbf{y}_1^T\mathbf{y}_1; \quad (29)$$

$$\frac{1}{n_1}\mathbf{y}_1^T(\mathbf{D}_{S_2}+\mathbf{D}_{N_2})\mathbf{y}_1\to(p+q)\mathbf{y}_2^T\mathbf{y}_2. \quad (30)$$

Therefore in Case 2 we have

$$\frac{\lambda_2(\mathbf{L})}{n}\xrightarrow{\text{a.s.}}\min_{\mathbf{x}\in\mathcal{S}}\left\{\frac{\mathbf{x}_1^T\mathbf{L}_1\mathbf{x}_1+\mathbf{x}_2^T\mathbf{L}_2\mathbf{x}_2+n_2 t\mathbf{x}_1^T\mathbf{x}_1+n_1 t\mathbf{x}_2^T\mathbf{x}_2}{n}\right\}, \quad (31)$$

where $\mathbf{L}_i=\mathbf{L}_{Si}+\mathbf{L}_{Ni}$, $t=p+q$, and

$$\mathcal{S}=\left\{\mathbf{x}=[\mathbf{x}_1\ \mathbf{x}_2]^T:\mathbf{1}_{n_1}^T\mathbf{x}_1=\mathbf{1}_{n_2}^T\mathbf{x}_2=0,\ \mathbf{x}_1^T\mathbf{x}_1+\mathbf{x}_2^T\mathbf{x}_2=1\right\}. \quad (32)$$

Define two sets

$$\mathcal{S}_1=\left\{\mathbf{x}:\mathbf{1}_{n_1}^T\mathbf{x}_1=\mathbf{1}_{n_2}^T\mathbf{x}_2=0,\ \mathbf{x}_1^T\mathbf{x}_1=1,\ \mathbf{x}_2^T\mathbf{x}_2=0\right\}; \quad (33)$$

$$\mathcal{S}_2=\left\{\mathbf{x}:\mathbf{1}_{n_1}^T\mathbf{x}_1=\mathbf{1}_{n_2}^T\mathbf{x}_2=0,\ \mathbf{x}_1^T\mathbf{x}_1=0,\ \mathbf{x}_2^T\mathbf{x}_2=1\right\}, \quad (34)$$

and define

$$\mu_i(\mathbf{L})=\min_{\mathbf{x}\in\mathcal{S}_i}\left\{\frac{\mathbf{x}_1^T\mathbf{L}_1\mathbf{x}_1+\mathbf{x}_2^T\mathbf{L}_2\mathbf{x}_2+n_2 t\mathbf{x}_1^T\mathbf{x}_1+n_1 t\mathbf{x}_2^T\mathbf{x}_2}{n}\right\}. \quad (35)$$

Since $\mathcal{S}_1,\mathcal{S}_2\subseteq\mathcal{S}$, we have, almost surely,

$$\frac{\lambda_2(\mathbf{L})}{n}\le\min\{\mu_1(\mathbf{L}),\mu_2(\mathbf{L})\}$$
$$=\min\left\{\frac{\lambda_2(\mathbf{L}_1)+n_2 t}{n},\frac{\lambda_2(\mathbf{L}_2)+n_1 t}{n}\right\}$$
$$=\frac{t}{2}+\frac{\lambda_2(\mathbf{L}_1)+\lambda_2(\mathbf{L}_2)-|\lambda_2(\mathbf{L}_1)-\lambda_2(\mathbf{L}_2)+(n_2-n_1)t|}{2n}$$
$$\le\frac{t}{2}+\frac{|n_1-n_2|t}{2n}+\frac{\lambda_2(\mathbf{L}_1)+\lambda_2(\mathbf{L}_2)-|\lambda_2(\mathbf{L}_1)-\lambda_2(\mathbf{L}_2)|}{2n}, \quad (36)$$

where we use the facts that $\min\{a,b\}=\frac{a+b-|a-b|}{2}$ and $|a-b|\ge|a|-|b|$. Note that the last equality in (36) holds if $n_1=n_2$. Let $t^*=p^*+q$ be the critical value for phase transition from Case 1 to Case 2. There is a phase transition on the asymptotic value of $\frac{\lambda_2(\mathbf{L})}{n}$ since the slope of $\frac{\lambda_2(\mathbf{L})}{n}$ converges to 1 almost surely when $t\le t^*$, whereas from (36) $\frac{\lambda_2(\mathbf{L})}{n}-t\le\frac{(|n_1-n_2|-n)t}{2n}+\frac{\lambda_2(\mathbf{L}_1)+\lambda_2(\mathbf{L}_2)-|\lambda_2(\mathbf{L}_1)-\lambda_2(\mathbf{L}_2)|}{2n}$ when $t\ge t^*$. From (19), we obtain an asymptotic upper bound p_{UB} on the critical value p^* by substituting $t^*=p^*+q$ to (36).

$$p_{\text{UB}}=\frac{\lambda_2(\mathbf{L}_1)+\lambda_2(\mathbf{L}_2)-|\lambda_2(\mathbf{L}_1)-\lambda_2(\mathbf{L}_2)|}{n-|n_1-n_2|}-q. \quad (37)$$

To derive a lower bound on p^*, we have that in Case 2,

$$\frac{\lambda_2(\mathbf{L})}{n}\xrightarrow{\text{a.s.}}\min_{\mathbf{x}\in\mathcal{S}}\left\{\frac{\mathbf{x}_1^T\mathbf{L}_1\mathbf{x}_1+\mathbf{x}_2^T\mathbf{L}_2\mathbf{x}_2+n_2 p\mathbf{x}_1^T\mathbf{x}_1+n_1 p\mathbf{x}_2^T\mathbf{x}_2}{n}\right\}$$
$$\ge\min_{\mathbf{x}\in\mathcal{S}}\left\{\frac{\mathbf{x}_1^T\mathbf{L}_1\mathbf{x}_1+\mathbf{x}_2^T\mathbf{L}_2\mathbf{x}_2}{n}\right\}+\min_{\mathbf{x}\in\mathcal{S}}\left\{\frac{n_2 t\mathbf{x}_1^T\mathbf{x}_1+n_1 t\mathbf{x}_2^T\mathbf{x}_2}{n}\right\}$$
$$=\min\left\{\frac{\lambda_2(\mathbf{L}_1)}{n},\frac{\lambda_2(\mathbf{L}_2)}{n}\right\}+\min\left\{\frac{n_1 t}{n},\frac{n_2 t}{n}\right\}. \quad (38)$$
$$=\frac{t}{2}-\frac{|n_1-n_2|t}{2n}+\frac{\lambda_2(\mathbf{L}_1)+\lambda_2(\mathbf{L}_2)-|\lambda_2(\mathbf{L}_1)-\lambda_2(\mathbf{L}_2)|}{2n}. \quad (39)$$

Substituting $t^*=p^*+q$ to (39), we obtain an asymptotic lower bound p_{LB} on the critical value p^*.

$$p_{\text{LB}}=\frac{\lambda_2(\mathbf{L}_1)+\lambda_2(\mathbf{L}_2)-|\lambda_2(\mathbf{L}_1)-\lambda_2(\mathbf{L}_2)|}{n+|n_1-n_2|}-q. \quad (40)$$

Note that when $n_1=n_2$, the equality in (38) holds. This means when $n_1=n_2$, $\frac{\lambda_2(\mathbf{L})}{n}\xrightarrow{\text{a.s.}}\frac{t}{2}+\frac{\lambda_2(\mathbf{L}_1)+\lambda_2(\mathbf{L}_2)-|\lambda_2(\mathbf{L}_1)-\lambda_2(\mathbf{L}_2)|}{2n}=:\frac{t}{2}+c^*$ in Case 2, and the critical value

$$p^*\xrightarrow{\text{a.s.}}\frac{\lambda_2(\mathbf{L}_1)+\lambda_2(\mathbf{L}_2)-|\lambda_2(\mathbf{L}_1)-\lambda_2(\mathbf{L}_2)|}{n}-q. \quad (41)$$

Here we derive the bounds on the critical value p^* for the stochastic block model, where the internal adjacency matrix $\mathbf{A}_i$ in (2) is generated by a Erdos-Renyi random graph with edge connection prob-
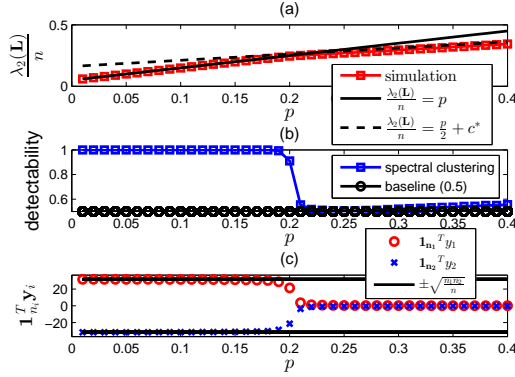
Fig. 1. Two communities generated by the stochastic block model [29]. The results are averaged over 100 trials. $n_1 = n_2 = 2000$, $p_1 = p_2 = 0.25$, and $q = 0.05$. The theoretical critical value from (41) is $p^* = 0.2229$.

noise level (q)		0	0.002	0.01	0.05	0.1
detectability	mean	0.8571	0.8548	0.8004	0.6325	0.5038
	std	0	0.006	0.1227	0.1597	0.0823
$\widehat{p}_{\text{LB}}$	mean	0.0127	0.0116	0.0076	0.00016	0
	std	0	0.0021	0.0039	0.001	0
$\widehat{p}$	meaen	0.0073	0.0095	0.0173	0.0513	0.0835
	std	0	0.001	0.0025	0.011	0.0209
$\widehat{p}_{\text{UB}}$	mean	0.013	0.0124	0.0633	0.1422	0.1494
	std	0	0.0021	0.1493	0.3199	0.3213
fraction of $\widehat{p} \leq \widehat{p}_{\text{LB}}$		1	0.98	0.01	0	0
fraction of $\widehat{p}_{\text{UB}} < \widehat{p} < \widehat{p}_{\text{UB}}$		0	0.02	0.75	0.2	0.2
fraction of $\widehat{p} \geq \widehat{p}_{\text{UB}}$		0	0	0.24	0.8	0.8

TABLE I
Sensitivity of spectral community detection to noisy edge insertions for Amazon American political books co-purchasement data [36]. The network contains 105 nodes and 441 edges. The oracle detectability is 0.8762. The noisy edges are randomly generated for 100 trials.

ability p_i. It is proved in Appendix VII-C of [27] that $\lambda_2 \left(\frac{\mathbf{L}_i}{n_i} \right) \xrightarrow{\text{a.s.}} p_i + q$. Therefore $p_{\text{UB}} = \frac{cp_1 + p_2 - |cp_1 - p_2 + (c-1)q| - |1-c|q}{1+c-|1-c|}$ and $p_{\text{LB}} = \frac{cp_1 + p_2 - |cp_1 - p_2 + (c-1)q| - |1-c|q}{1+c+|1-c|}$. When $n_1 = n_2$ (i.e., $c = 1$), the critical value $p^* \xrightarrow{\text{a.s.}} \frac{p_1 + p_2 - |p_1 - p_2|}{2}$. This suggests that in the largest network limit when $n \to \infty$ and $c = 1$ the performance of spectral community detection is independent of the noise parameter q.

V. PERFORMANCE EVALUATION

A. Simulated Networks

We use the stochastic block model [29] to generate network graphs for community detection. The detectability is defined as the fraction of nodes that are correctly identified and the baseline detectability is 0.5 for random guesses. In Fig. 1, when $p_1 = p_2 = 0.25$, $n_1 = n_2 = 2000$ and $q = 0.05$, the theoretical critical value from (41) is $p^* = 0.2229$. Note that p^* will converge to 0.25 as we increase n as predicted in Sec. IV.

Fig. 1 (a) verifies the phase transition in $\frac{\lambda_2(\mathbf{L})}{n}$ empirically confirming that $\frac{\lambda_2(\mathbf{L})}{n}$ approaches $p + q$ when $p \leq p^*$ and $\frac{\lambda_2(\mathbf{L})}{n}$ approaches $\frac{p+q}{2} + c^*$ when $p > p^*$, where $c^* = \frac{\lambda_2(\mathbf{L}_1) + \lambda_2(\mathbf{L}_2) - |\lambda_2(\mathbf{L}_1) - \lambda_2(\mathbf{L}_2)|}{2n}$. Fig. 1 (b) shows that the community detectability transitions from almost perfect detectability when $p < p^*$ to low detectability when $p > p^*$. Moreover, as derived in (26), the Fiedler vector components $\mathbf{y}_1$ and $\mathbf{y}_2$ are constant vectors with opposite signs for $p < p^*$, and $\mathbf{1}_{n_1}^T \mathbf{y}_1 \to 0$ and $\mathbf{1}_{n_2}^T \mathbf{y}_2 \to 0$ for $p > p^*$, as shown in Fig. 1 (c).

B. Empirical Estimators of Phase Transition Bounds on Real-world Dataset

Here we show that the critical phase transition threshold p^* can be empirically estimated to empirically test the reliability of spectral community detection. Let $\widehat{\mathbf{L}}_i$ be the graph Laplacian matrix of the estimated community i obtained by applying spectral clustering to the observed adjacency matrix $\mathbf{A}$ and let $\widehat{n}_i$ denote the estimated network size of community i. Using (37) and (40), the empirical estimators of these parameters are defined as

$$\widehat{p} = \text{number of identified external edges}/\widehat{n}_1\widehat{n}_2, \quad (42)$$

$$\widehat{p}_{\text{LB}} = \frac{\lambda_2(\widehat{\mathbf{L}}_1) + \lambda_2(\widehat{\mathbf{L}}_2) - \left| \lambda_2(\widehat{\mathbf{L}}_1) - \lambda_2(\widehat{\mathbf{L}}_2) \right|}{n + |\widehat{n}_1 - \widehat{n}_2|}, \quad (43)$$

$$\widehat{p}_{\text{UB}} = \frac{\lambda_2(\widehat{\mathbf{L}}_1) + \lambda_2(\widehat{\mathbf{L}}_2) - \left| \lambda_2(\widehat{\mathbf{L}}_1) - \lambda_2(\widehat{\mathbf{L}}_2) \right|}{n - |\widehat{n}_1 - \widehat{n}_2|}. \quad (44)$$

Based on these empirical estimates, the performance of community detection can be classified into three categories. If $\widehat{p} \leq \widehat{p}_{\text{LB}}$, the network is in the reliable detection region. If $\widehat{p}_{\text{LB}} < \widehat{p} < \widehat{p}_{\text{UB}}$, the network is in the intermediate detection region. If $\widehat{p} \geq \widehat{p}_{\text{UB}}$, the network is in the unreliable detection region.

The co-purchasement data between 105 American political books sold on Amazon [36] are used to estimate the parameters p_{LB}, p_{UB} and p. For the corresponding network graph nodes represent political books and edges represent co-purchasements. An edge exists between two books if they are frequently purchased by the same buyer. Three labels, *liberal*, *conservative* and *neutral*, were determined by Newman [36]. We perform community detection by separating the books into two groups since there are only 13 books with neutral labels (i.e., the oracle detectability is 0.8762). To investigate the sensitivity of spectral community detection to noisy edge insertions, for each edge not present in the original graph, an edge is added with probability q. The community detection results are summarized in Table I. Observe that for small q ($q=0$ or 0.002) the network is mostly in the reliable detection region ($\widehat{p} < \widehat{p}_{\text{LB}}$), which indicates that spectral community detection achieves high detectability. When $q = 0.01$, the network is mostly in the intermediate detection region ($\widehat{p}_{\text{LB}} < \widehat{p} < \widehat{p}_{\text{UB}}$), indicating that the community detectability has large variation. When q is large ($q=0.05$ or 0.1), the network is mostly in the unreliable detection region resulting in low detectability. The large standard deviation of $\widehat{p}_{\text{UB}}$ for large q is due to the fact that spectral community detection may mistakenly detect two communities with extremely imbalanced community sizes such that the denominator of the estimator $\widehat{p}_{\text{UB}}$ is small.

VI. CONCLUSION

We establish asymptotic phase transition bounds on the critical value p^* under a general network setting corrupted by a Erdos-Renyi type noise model. The communities are proven to be almost perfectly detectable below the phase transition threshold and to be undetectable above the phase transition threshold. The phase transition bounds are used to establish empirical estimators to evaluate the reliability of spectral community detection, where the detector is said to be operating in the reliable, intermediate, or unreliable detection regime based on the empirical estimates. Simulated networks generated by the stochastic block model validate the phase transition theory for community detectability. An empirical estimator of the phase transition is proposed that can be used to explore sensitivity of the spectral community detection algorithm on real data.

References

[1] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3-5, pp. 75–174, 2010.

[2] B. Miller, N. Bliss, and P. J. Wolfe, "Subgraph detection using eigenvector L1 norms," in *Advances in Neural Information Processing Systems (NIPS)*, 2010, pp. 1633–1641.

[3] A. Sandryhaila and J. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.

[4] A. Bertrand and M. Moonen, "Seeing the bigger picture: How nodes can learn their place within a complex ad hoc network topology," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 71–82, May 2013.

[5] D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. –98, May 2013.

[6] B. Miller, N. Bliss, and P. Wolfe, "Toward signal processing theory for graphs and non-Euclidean data," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2010, pp. 5414–5417.

[7] P.-Y. Chen and A. O. Hero, "Deep community detection," `arXiv:1407.6071`, 2014.

[8] S. Chen, A. Sandryhaila, G. Lederman, Z. Wang, J. Moura, P. Rizzo, J. Bielak, J. Garrett, and J. Kovacevic, "Signal inpainting on graphs via total variation minimization," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 8267–8271.

[9] P.-Y. Chen and A. O. Hero, "Local Fiedler vector centrality for detection of deep and overlapping communities in networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 1120–1124.

[10] U. Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, Dec. 2007.

[11] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, 2000.

[12] S. White and P. Smyth, "A spectral clustering approach to finding communities in graph," in *SDM*, 2005, pp. 274–285.

[13] Y. van Gennip, H. Hu, B. Hunter, and M. A. Porter, "Geosocial graph-based community detection," in *IEEE International Conference on Data Mining Workshops*, 2012, pp. 754–758.

[14] S. Tsironis, M. Sozio, T. Paristech, and M. Vazirgiannis, "Accurate spectral clustering for community detection in MapReduce," in *Advances in Neural Information Processing Systems (NIPS) Workshops*, 2013.

[15] L. Huang, R. Li, H. Chen, X. Gu, K. Wen, and Y. Li, "Detecting network communities using regularized spectral clustering algorithm," *Artificial Intelligence Review*, vol. 41, no. 4, pp. 579–594, 2014.

[16] R. Merris, "Laplacian matrices of graphs: a survey," *Linear Algebra and its Applications*, vol. 197-198, pp. 143–176, 1994.

[17] F. R. K. Chung, *Spectral Graph Theory*. American Mathematical Society, 1997.

[18] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Mathematical Journal*, vol. 23, no. 98, pp. 298–305, 1973.

[19] J. A. Hartigan and M. A. Wong, "A k-means clustering algorithm," *JSTOR: Applied Statistics*, vol. 28, no. 1, pp. 100–108, 1979.

[20] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Phys. Rev. E*, vol. 74, p. 036104, Sep 2006.

[21] P. J. Bickel and A. Chen, "A nonparametric view of network models and newmangirvan and other modularities," *Proceedings of the National Academy of Sciences*, vol. 106, no. 50, pp. 21 068–21 073, 2009.

[22] Y. Zhao, E. Levina, and J. Zhu, "Consistency of community detection in networks under degree-corrected stochastic block models," *The Annals of Statistics*, vol. 40, no. 4, pp. 2266–2292, 08 2012.

[23] R. R. Nadakuditi and M. E. J. Newman, "Graph spectra and the detectability of community structure in networks," *Phys. Rev. Lett.*, vol. 108, p. 188701, May 2012.

[24] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborov, and P. Zhang, "Spectral redemption in clustering sparse networks," *Proceedings of the National Academy of Sciences*, vol. 110, no. 52, pp. 20 935–20 940, 2013.

[25] F. Radicchi, "Detectability of communities in heterogeneous networks," *Phys. Rev. E*, vol. 88, p. 010801, Jul 2013.

[26] ——, "A paradox in community detection," *EPL (Europhysics Letters)*, vol. 106, no. 3, p. 38001, 2014.

[27] P.-Y. Chen and A. O. Hero, "Phase transitions in spectral community detection," `arXiv:1409.3207`, 2014.

[28] ——, "Universal phase transition in community detectability under a stochastic block model," `arXiv:1409.2186`, 2014.

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

[30] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, "Inference and phase transitions in the detection of modules in sparse networks," *Phys. Rev. Lett.*, vol. 107, p. 065701, Aug 2011.

[31] R. R. Nadakuditi, "On hard limits of eigen-analysis based planted clique detection," in *IEEE Statistical Signal Processing Workshop (SSP)*, Aug 2012, pp. 129–132.

[32] F. Radicchi and A. Arenas, "Abrupt transition in the structural formation of interconnected networks," *Nature Physics*, vol. 9, no. 11, pp. 717–720, Nov. 2013.

[33] R. Latala, "Some estimates of norms of random matrices." *Proc. Am. Math. Soc.*, vol. 133, no. 5, pp. 1273–1282, 2005.

[34] M. Talagrand, "Concentration of measure and isoperimetric inequalities in product spaces," *Publications Mathmatiques de l'Institut des Hautes tudes Scientifiques*, vol. 81, no. 1, pp. 73–205, 1995.

[35] F. Benaych-Georges and R. R. Nadakuditi, "The singular values and vectors of low rank perturbations of large rectangular random matrices," *Journal of Multivariate Analysis*, vol. 111, no. 0, pp. 120–135, 2012.

[36] M. E. J. Newman, "Modularity and community structure in networks," *Proc. National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.