Non-backtracking spectrum of random graphs: community detection and non-regular Ramanujan graphs

Charles Bordenave
Institut de Mathématiques de Toulouse
CNRS and University of Toulouse III
Toulouse, France
Email: bordenave@math.univ-toulouse.fr

Marc Lelarge
INRIA-ENS
Paris, France
Email: marc.lelarge@inria.fr

Laurent Massoulié
Microsoft Research-Inria Joint Centre
Saclay, France
Email: laurent.massoulie@inria.fr

Abstract

A non-backtracking walk on a graph is a directed path such that no edge is the inverse of its preceding edge. The non-backtracking matrix of a graph is indexed by its directed edges and can be used to count non-backtracking walks of a given length. It has been used recently in the context of community detection and has appeared previously in connection with the Ihara zeta function and in some generalizations of Ramanujan graphs. In this work, we study the largest eigenvalues of the non-backtracking matrix of the Erdős-Rényi random graph and of the Stochastic Block Model in the regime where the number of edges is proportional to the number of vertices. Our results confirm the "spectral redemption conjecture" that community detection can be made on the basis of the leading eigenvectors above the feasibility threshold.

Keywords

Stochastic Block Model; Ramanujan graphs; non-backtracking matrix.

I. Introduction

Given a finite (simple, non-oriented) graph $G = (V, E)$, several matrices of interest can be associated to G, besides its adjacency matrix $A = (\mathbf{1}_{\{u,v\} \in E})_{u,v \in V}$. In this work we are interested in the so-called *non-backtracking* matrix of G, denoted by B. It is indexed by the set $\vec{E} = \{(u, v) : \{u, v\} \in E\}$ of *oriented* edges in E and defined by

$$B_{ef} = \mathbf{1}(e_2 = f_1)\mathbf{1}(e_1 \neq f_2) = \mathbf{1}(e_2 = f_1)\mathbf{1}(e \neq f^{-1}),$$

where for any $e = (u, v) \in \vec{E}$, we set $e_1 = u$, $e_2 = v$, $e^{-1} = (v, u)$. This matrix was introduced by Hashimoto [1]. A non-backtracking walk is a directed path of directed edges of G such that no edge is the inverse of its preceding edge. It is easily seen that for any $k \geq 1$, B^k_{ef} counts the number of non-backtracking walks of $k+1$ edges on G starting with e and ending with f.

Our focus is the spectrum of B, referred to in the sequel as the non-backtracking spectrum of G, when G is a sparse random graph. Specifically we shall characterize the asymptotic behavior of the leading eigenvalues and associated eigenvectors in the non-backtracking spectrum of sparse random graphs in the limit $n \to \infty$ where $n = |V|$.

The random graphs we consider are drawn according to the Stochastic Block Model, a generalization of Erdős-Rényi graphs due to Holland et al. [2]. In this model nodes $v \in V$ are partitioned into r groups. We focus on the sparse case: an edge between two nodes u, v is drawn with probability $W(\sigma(u), \sigma(v))/n$, where $\sigma(u) \in [r]$ denotes the group node u belongs to. Thus when the $r \times r$ matrix W is fixed the expected node degrees remain of order 1 as $n \to \infty$. We focus moreover on instances where the fraction of nodes in group i converges to a limit $\pi(i)$ as $n \to \infty$.

An informal statement of our results for eigenvalues is as follows. Let G be drawn according to a Stochastic Block Model with fixed number r of node groups such that all nodes have same fixed expected degree $\alpha > 1$. Let $\mu_1, \dots, \mu_r$ denote the leading eigenvalues of the expected adjacency matrix $\bar{A} := \mathbb{E}(A)$, ordered so that $\mu_1 = \alpha \geq |\mu_2| \geq \dots \geq |\mu_r|$. Let $r_0 \leq r$ be such that $|\mu_{r_0+1}| \leq \sqrt{\alpha} < |\mu_{r_0}|$. Then the r_0 leading eigenvalues of B are asymptotic to $\mu_1, \dots, \mu_{r_0}$, the remaining eigenvalues λ satisfying $|\lambda| \leq (1 + o(1))\sqrt{\alpha}$.

Community detection

Our primary motivation stems from the problem of *community detection*, namely: how to estimate a clustering of graph nodes into groups close to the underlying blocks, based on the observation of such a random graph G? Decelle et al. [3] conjectured a phase transition phenomenon on detectability, namely: the underlying block structure could be detected if and only if $|\mu_2| > \sqrt{\alpha}$.

In the case of two communities with roughly equal sizes ($\pi(1) = \pi(2) = 1/2$) and symmetric matrix W, the negative part (impossibility of detection when $|\mu_2| \leq \sqrt{\alpha}$) was proven by Mossel et al [4]. As for the positive part (feasibility of detection

0272-5428/15 $31.00 © 2015 IEEE
DOI 10.1109/FOCS.2015.86

when $|\mu_2| > \sqrt{\alpha}$), it was conjectured in [3] that the so-called belief propagation algorithm would succeed. Krzakala et al. [5] then made their so-called "spectral redemption conjecture" according to which detection based on the second eigenvector of the non-backtracking matrix B would succeed for the case of two communities.

Recently a variant of the spectral redemption conjecture was proven by Massoulié [6]: the spectrum of a matrix counting *self-avoiding paths* in G allows us to detect communities through thresholding of the second eigenvector. More recently and independently of [6], an alternative proof of the positive part of the conjecture in [3] was given by Mossel et al. [7], based on an elaborate construction involving countings of non-backtracking paths in G.

The two approaches of [6] and [7] to proving the positive part of the conjecture in [3], while both relying ultimately on properties of specific path counts, differ however in the following respects. The method in [6] relies on a clear spectral separation property but its implementation is computationally expensive, as the counts of self-avoiding walks it relies upon take super-linear (though polynomial) time. The method in [7] is computationally efficient as it runs in quasi-linear time, but the proof does not establish a spectral separation property. The other two methods conjectured to achieve successful reconstruction, namely belief propagation and analysis of non-backtracking spectrum, are computationally efficient and they are motivated by a clear intuition as described in the spectral redemption conjecture. In particular the spectral algorithm in [5] has $O(n \log n)$ running time with a small constant in the O and an efficient implementation.

Our present work proves the spectral redemption conjecture. More generally by characterizing all the leading eigen-elements it determines the limits of community detection based on the non-backtracking spectrum in the presence of an arbitrary number of communities.

Weak Ramanujan property

Our result also has an interpretation from the standpoint of Ramanujan graphs, introduced by Lubotzky et al. [8] (see Murty [9] for a recent survey). These are by definition k-regular graphs such that the second largest modulus of its eigenvalues is at most $2\sqrt{k-1}$. By a result of Alon and Boppana (see [10]) for fixed k, k-regular graphs on n nodes must have their second largest eigenvalue at least $2\sqrt{k-1} - o(1)$ as $n \to \infty$. Hence Ramanujan graphs are regular graphs with maximal spectral gap between the first and second eigenvalue moduli. A celebrated result of Friedman [11] states that random k-regular graphs achieve this lower bound with high probability as their number of nodes n goes to infinity.

Lubotzky [12] has proposed an extension of the definition of Ramanujan graphs to the non-regular case. Specifically, G is Ramanujan according to this definition if and only if

$$\max\{|\lambda| : \lambda \in \text{spectrum}(A),\ |\lambda| \neq \rho\} \leq \sigma,$$

where A is the adjacency matrix of G, ρ its spectral radius, and σ the spectral radius of the *adjacency operator on the universal covering tree of* G.

Using the analogy between the Ihara zeta function and the Riemann zeta function, Stark and Terras (see [13]) have defined a graph to satisfy the *graph Riemann hypothesis* if its non-backtracking matrix B has no eigenvalues λ such that $|\lambda| \in (\sqrt{\rho_B}, \rho_B)$, where ρ_B is the Perron-Frobenius eigenvalue of B. Interestingly, a regular graph G is Ramanujan if and only if it satisfies the graph Riemann hypothesis (see [9] and [13]). Thus the graph Riemann hypothesis can also be seen as a generalization of the notion of Ramanujan graphs to the non-regular case, phrased in terms of non-backtracking spectra rather than on spectra of universal covers as in the definition of Lubotzky [12].

Our results imply that for fixed $\alpha > 1$, Erdős-Rényi graphs $\mathcal{G}(n, \alpha/n)$ have an associated non-backtracking matrix B such that $\rho_B \sim \alpha$ and all its other eigenvalues λ verify $|\lambda| \leq \sqrt{\alpha} + o(1)$ with high probability as $n \to \infty$. In this sense, Erdős-Rényi graphs asymptotically satisfy the graph Riemann hypothesis, which itself is a plausible extension of the notion of Ramanujan graphs to the non-regular case. This may be seen as an analogue of Friedman's Theorem [11] in the context of Erdős-Rényi graphs. Similarly, for the Stochastic Block Model, our main result is analogous to recent results on the eigenvalues of random n-lifts of base graphs, see [14], [15]. Interestingly, in [15], the methods developed in the present paper are adapted to lead to a new and simpler proof of Friedman's Theorem and its extensions to random n-lifts. The random graphs studied here will require a more delicate analysis.

II. Main results

We now state our results on the non-backtracking spectra of Erdős-Rényi graphs first, and Stochastic Block Models next.

A. Erdős-Rényi graphs

Let the vector χ on $\mathbb{R}^{\vec{E}}$ be defined as $\chi(e) = 1,\ e \in \vec{E}$. The Euclidean norm of a vector $x \in \mathbb{R}^d$ will be denoted by $\|x\|$. We have the following theorem.

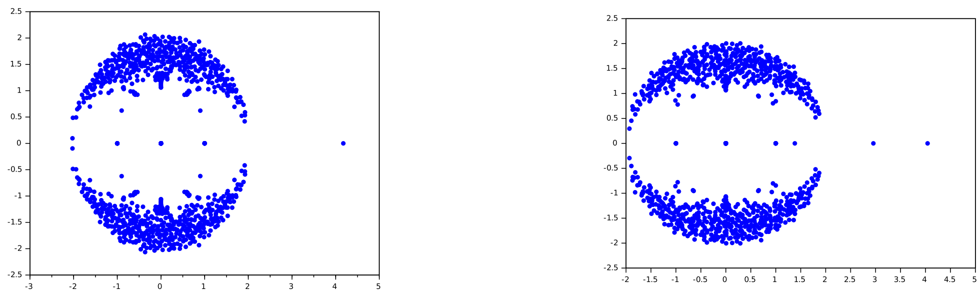

Figure 1. Left : eigenvalues of B for a realization of an Erdős-Rényi graph with parameters $(n, \alpha/n)$ with $n = 500$, $\alpha = 4$. Right : eigenvalues of B for Example 1 (symmetric case) with $n = 500$, $a = 7$, $b = 1$.

Theorem 1. *Let G be an Erdős-Rényi graph with parameters $(n, \alpha/n)$ for some fixed parameter $\alpha > 1$. Then, with probability tending to 1 as $n \to \infty$, the eigenvalues $\lambda_i(B)$ of its non-backtracking matrix B satisfy*

$$\lambda_1(B) = \alpha + o(1) \ \text{ and } \ |\lambda_2(B)| \leq \sqrt{\alpha} + o(1).$$

Moreover the normalized Perron-Frobenius eigenvector associated to $\lambda_1(B)$ is asymptotically aligned with

$$\frac{B^\ell B^{*\ell}\chi}{\|B^\ell B^{*\ell}\chi\|},$$

where $\ell \sim \kappa \log_\alpha n$ for any $0 < \kappa < 1/6$.

Theorem 1 is illustrated by Figure 1. We conjecture that the lower bound $|\lambda_2(B)| \geq \sqrt{\alpha} - o(1)$ holds, it is reasonable in view of Figure 1.

B. Stochastic Block Model

For integer $k \geq 1$, we set $[k] = \{1, \dots, k\}$. We consider a random graph $G = (V, E)$ on the vertex set $V = [n]$ defined as follows. Each vertex $v \in [n]$ is given a *type* $\sigma_n(v)$ from the set $[r]$ where the number of types r is assumed fixed and the map $\sigma_n : [n] \to [r]$ is such that, for all $i \in [r]$,

$$\pi_n(i) := \frac{1}{n} \sum_{v=1}^{n} \mathbf{1}(\sigma_n(v) = i) = \pi(i) + o(1), \tag{1}$$

for some probability vector $\pi = (\pi(1), \cdots, \pi(r))$. For ease of notation, we often write σ in place of σ_n.

Given a symmetric $r \times r$ matrix with non-negative entries W we assume that there is an edge between vertices u and v independently with probability

$$\frac{W(\sigma(u), \sigma(v))}{n} \wedge 1.$$

We set $\Pi = \mathrm{diag}(\pi(1), \dots, \pi(r))$ and introduce the mean progeny matrix $M = \Pi W$. Note that the eigenvalues of M are the same as the ones of the symmetric matrix $S = \Pi^{1/2} W \Pi^{1/2}$ and in particular are real-valued. They are also the same as the non-zero eigenvalues of the expected adjacency matrix $\bar{A} := \mathbb{E}(A)$ conditioned on the vertex types. We denote them by μ_k and order them by their absolute value, $|\mu_r| \leq \cdots \leq |\mu_2| \leq \mu_1$. We shall make the following assumptions:

$$\mu_1 > 1 \text{ and } M \text{ is positively regular,} \tag{2}$$

i.e. for some integer $k \geq 1$, M^k has positive coefficients. In particular, $\mu_1 > \max_{k \geq 2} |\mu_k|$ is the Perron-Frobenius eigenvalue. It implies notably that for all $i \in [r]$, $\pi(i) > 0$. We define r_0 by

$$\mu_{r_0}^2 > \mu_1 \text{ and } \quad \mu_{r_0+1}^2 \leq \mu_1,$$

(with $\mu_{r+1} = 0$). Since $M = \Pi^{1/2} S \Pi^{-1/2}$, the matrix M is diagonalizable. Let $\{u_i\}_{i \in [r]}$ be an orthonormal basis of eigenvectors of S such that $S u_i = \mu_i u_i$. Then $\phi_i := \Pi^{-1/2} u_i$ and $\psi_i = \Pi^{1/2} u_i$ are the left and right-eigenvectors associated to eigenvalue μ_i, $\phi_i^* M = \mu_i \phi_i^*$, $M \psi_i = \mu_i \psi_i$. We get

$$\langle \phi_i, \psi_j \rangle = \delta_{ij}, \quad \text{and,} \quad \langle \phi_i, \phi_j \rangle_\pi = \delta_{ij}, \tag{3}$$

where $\langle x, y\rangle_\pi = \sum_k \pi(k) x_k y_k$ denotes the usual inner product in $\ell^2(\pi)$. The following spectral decompositions will also be useful

$$M = \sum_{k=1}^{r} \mu_k \psi_k \phi_k^* \quad \text{and} \quad W = \sum_{k=1}^{r} \mu_k \phi_k \phi_k^*, \tag{4}$$

where the second identity comes from $\psi_k = \Pi\phi_k$ and $W = \Pi^{-1} M$.

We will make the further assumption that each vertex type has the same asymptotic average degree $\alpha > 1$, i.e.,

$$\alpha = \sum_{i=1}^{r} \pi(i) W_{ij} = \sum_{i=1}^{r} M_{ij} \qquad \text{for all } j \in [r]. \tag{5}$$

This entails that M^*/α is a stochastic matrix and we then have

$$\mu_1 = \alpha > 1, \quad \phi_1 = \mathbf{1} \quad \text{and} \quad \psi_1 = \pi^*. \tag{6}$$

We will also assume that a quantitative version of (1) holds, namely that for some $\gamma \in (0, 1]$,

$$\|\pi - \pi_n\|_\infty = \max_{i \in [r]} |\pi(i) - \pi_n(i)| = O(n^{-\gamma}). \tag{7}$$

The random graph G is usually called the stochastic block model (SBM for short) or inhomogeneous random graph, see Bollobás, Janson and Riordan [16] and Holland, Laskey and Leinhardt [2]. A popular case is when the map σ is itself random and $\sigma(v)$ are i.i.d. with distribution $(\pi(1), \cdots, \pi(r))$. In this case, with probability one, condition (7) is met for any $\gamma < 1/2$.

Example 1. *If $r = 2$, then we have $\pi(1) = 1 - \pi(2)$. Under condition (5), we have $W_{22} = (\pi(1)W_{11} + (1 - 2\pi(1))W_{12})/(1 - \pi(1))$ so that $\mu_1 = \alpha = \pi(1)W_{11} + (1 - \pi(1))W_{12}$ and $\mu_2 = \pi(1)(W_{11} - W_{12})$.*

In the symmetric case, $\pi(i) = 1/2$ and $W_{ii} = a \neq b = W_{ij}$ for all $i \neq j$, we have $\mu_1 = \alpha = (a+b)/2$ and $\mu_2 = (a-b)/2$.

For $k \in [r]$, we introduce the vector on $\mathbb{R}^{\vec{E}}$,

$$\chi_k(e) = \phi_k(\sigma(e_2)) \quad \text{for all } e \in \vec{E}. \tag{8}$$

In particular, $\chi_1 = \chi$. If $x \in \mathbb{R}^{\vec{E}}$, we define $\check{x}$ by

$$\check{x}(e) = x(e^{-1}). \tag{9}$$

Our main result is the following generalization of Theorem 1.

Theorem 2. *Let G be a SBM as above such that hypotheses (2,5,7) hold. Then with probability tending to 1 as $n \to \infty$,*

$$\lambda_k(B) = \mu_k + o(1) \;\text{ for } k \in [r_0], \;\; \text{and for } k > r_0, \;\; |\lambda_k(B)| \leq \sqrt{\alpha} + o(1).$$

Moreover, if μ_k is a simple eigenvalue of M for some $k \in [r_0]$, then a normalized eigenvector, say ξ_k, of $\lambda_k(B)$ is asymptotically aligned with

$$\frac{B^\ell B^{*\ell} \check{\chi}_k}{\|B^\ell B^{*\ell} \check{\chi}_k\|}, \tag{10}$$

where $\ell \sim \kappa \log_\alpha n$ for any $0 < \kappa < \gamma/6$. Finally, the vectors ξ_k of these simple eigenvalues are asymptotically orthogonal.

It follows from this result that a non-trivial estimation of the node types $\sigma(v)$ is feasible on the basis of the eigenvectors $\{\xi_k\}_{2 \leq k \leq r_0}$ provided $r_0 > 1$. More precisely, for vertex type estimators $\hat{\sigma}(v) : [n] \to [r]$ based on the observed random graph G, following Decelle et al. [3], define the overlap $\mathrm{ov}(\hat{\sigma}, \sigma)$ as the minimum over permutations $p : [r] \to [r]$ of the quantity

$$\frac{1}{n} \sum_{v=1}^{n} \mathbf{1}_{\hat{\sigma}(v) = p \circ \sigma(v)} - \max_{k \in [r]} \pi(k).$$

We shall say that $\hat{\sigma}$ has asymptotic overlap δ if $\mathrm{ov}(\hat{\sigma}, \sigma)$ converges in probability to δ as n grows. It has asymptotic positive overlap if for some $\delta > 0$, $\mathrm{ov}(\hat{\sigma}, \sigma) > \delta$ with probability tending to 1 as n grows. Note that an asymptotic overlap of zero is always achievable by assigning to each vertex the type k^* that maximizes $\pi(k)$. In the case where all communities have asymptotically the same size, i.e. $\pi(i) \equiv 1/r$, zero overlap is also achieved by assigning types at random.

As conjectured in [3] and proven in [7], in the setup of Example 1 with 2 communities of equal size, the best possible overlap is $o(1)$ with high probability when $r_0 = 1$, i.e. when $\mu_2 \leq \sqrt{\mu_1}$. Conversely, adapting the argument in [6], when $r_0 > 1$, we have the following

Theorem 3. *Let G be an SBM as above such that hypotheses (2,5,7) hold. Assume further that $\pi(i) \equiv 1/r$, that $r_0 > 1$ and that for some $k \in \{2, \ldots, r_0\}$, μ_k is a simple eigenvalue of M. Let $\xi_k \in \mathbb{R}^{\vec{E}}$ be a normalized eigenvector of B associated with $\lambda_k(B)$.*

Then, there exists a deterministic threshold $\tau \in \mathbb{R}$, a partition (I^+, I^-) of $[r]$ and a random signing $\omega \in \{-1,1\}^V$ dependent of ξ_k such that the following estimation procedure yields asymptotically positive overlap: assign to each vertex v a label $\hat{\sigma}(v)$ picked uniformly at random from I^+ if $\omega(v)\sum_{e:e_2=v} \xi_k(e) > \tau/\sqrt{n}$ and from I^- otherwise.

The reason for the existence of the signing $\omega \in \{-1,1\}^V$ in the above statement is that we do not know a priori whether the vector ξ_k or $-\xi_k$ is asymptotically close to (10). In the simplest case, we will be able to estimate this sign and the vector ω will be equal to $-\mathbf{1}$ or $\mathbf{1}$ and $I^+ = \{i \in [r] : \phi_k(i) > 0\}$, $I^- = [r]\backslash I^+$.

III. Erdős-Rényi graph: proof strategy for Theorem 1

We now give a sketch of the proof for Theorem 1 and will explain how to adapt the argument for the stochastic block model in Section IV. We refer to [17] for all details and proofs.

We start with an important remark: despite B not being a normal matrix, it contains some symmetry. More precisely, we observe that $(B^*)_{ef} = B_{fe} = B_{e^{-1}f^{-1}}$. Recall that for all $x \in \mathbb{R}^{\vec{E}}$ we define:

$$\check{x}_e = x_{e^{-1}},\ e \in \vec{E}. \tag{11}$$

It is then easy to check that for $x, y \in \mathbb{R}^{\vec{E}}$ and integer $k \geq 0$,

$$\langle y, B^k x\rangle = \langle B^k \check{y}, \check{x}\rangle. \tag{12}$$

In other words, if P denotes the involution on $\mathbb{R}^{\vec{E}}$, $Px = \check{x}$, we have for any integer $k \geq 0$,

$$B^k P = P B^{*k}.$$

Hence $B^k P$ is a symmetric matrix. If $(\sigma_{j,k})$, $1 \leq j \leq m$, are the eigenvalues of $B^k P$ and $(x_{j,k})$, $1 \leq j \leq m$, is an orthonormal basis of eigenvectors, we deduce that

$$B^k = \sum_{j=1}^{m} \sigma_{j,k} x_{j,k} \check{x}_{j,k}^*. \tag{13}$$

We order the eigenvalues,

$$\sigma_{1,k} \geq |\sigma_{2,k}| \geq \ldots \geq |\sigma_{m,k}|.$$

From Perron-Frobenius theorem, $x_{1,k}$ can be chosen to have non-negative entries. Since P is an orthogonal matrix, $(\check{x}_{j,k})$, $1 \leq j \leq m$, is also an orthonormal basis of $\mathbb{R}^{\vec{E}}$. In particular, (13) gives the singular value decomposition of B^k. Indeed, if $s_{j,k} = |\sigma_{j,k}|$ and $y_{j,k} = \operatorname{sign}(\sigma_{j,k})\check{x}_{j,k}$, we get

$$B^k = \sum_{j=1}^{m} s_{j,k} x_{j,k} y_{j,k}^*. \tag{14}$$

This is precisely the singular value decomposition of B^k.

For example, for $k = 1$, it is a simple exercise to compute $(\sigma_{j,1})_{1\leq j\leq m}$. We find that the eigenvalues of BP are $(\deg(v) - 1)$, $1 \leq v \leq n$, and -1 with multiplicity $m - n$. In particular, the singular values of B contain only information on the degree sequence of the underlying graph G as noted in [5].

For large k however, we may expect that the decomposition (13) carries more structural information on the graph. This will be the underlying principle in the proof of our main results. For the moment, we simply note the following. Assume that B is irreducible. From the Perron-Frobenius theorem, if ξ is the Perron eigenvector of B, $\|\xi\| = 1$, then for any n fixed,

$$\lambda_1(B) = \lim_{k\to\infty} \sigma_{1,k}^{1/k} \quad \text{and} \quad \lim_{k\to\infty} \|x_{1,k} - \xi\| = 0. \tag{15}$$

A quantitative version of the above limits is given in Section 4 of [17] as an extension of the Bauer-Fike theorem. Another consequence of (13) is that, for $i \neq j$, $x_{i,k}$ and $\check{x}_{j,k}$ should be nearly orthogonal if these vectors converge as $k \to \infty$. Indeed, a heuristic computation gives

$$\langle x_{i,k}, \check{x}_{j,k} \rangle = \frac{\langle B^k \check{x}_{i,k}, B^{*k} x_{j,k} \rangle}{\sigma_{i,k}\sigma_{j,k}} = \frac{\langle B^{2k} \check{x}_{i,k}, x_{j,k} \rangle}{\sigma_{i,k}\sigma_{j,k}} \simeq \frac{\langle B^{2k} \check{x}_{i,2k}, x_{j,2k} \rangle}{\sigma_{i,k}\sigma_{j,k}} = \frac{\sigma_{i,2k}\langle x_{i,2k}, x_{j,2k} \rangle}{\sigma_{i,k}\sigma_{j,k}} = 0.$$

We will exploit this general phenomenon in the proof of our main results as explained below.

In what follows, we consider a sequence $\ell = \ell(n) \sim \kappa \log_\alpha n$ for some $\kappa \in (0, 1/6)$ as in Theorem 1. Let

$$\varphi = \frac{B^\ell \chi}{\|B^\ell \chi\|}, \quad \theta = \|B^\ell \check{\varphi}\|, \text{ and } \zeta = \frac{B^\ell \check{\varphi}}{\theta} = \frac{B^\ell B^{*\ell} \chi}{\|B^\ell B^{*\ell} \chi\|},$$

(if $\theta = 0$, we set $\zeta = 0$). The proof relies on the following two propositions.

Proposition 4. *For some $c_1, c_0 > 0$, w.h.p. $\langle \zeta, \check{\varphi} \rangle \geq c_0$ and $c_0 \alpha^\ell \leq \theta \leq c_1 \alpha^\ell$.*

Proposition 5. *For some $c > 0$, w.h.p. $\sup_{x:\langle x, \check{\varphi}\rangle = 0, \|x\|=1} \|B^\ell x\| \leq (\log n)^c \alpha^{\ell/2}$.*

Let us check that the last two propositions 4 and 5 imply Theorem 1 . Let $R = B^\ell - \theta \zeta \check{\varphi}^*$ and $y \in \mathbb{R}^{\vec{E}}$ with $\|y\| = 1$. We write $y = s\check{\varphi} + x$ with $x \in \check{\varphi}^\perp$ and $s \in \mathbb{R}$. We find

$$\|Ry\| = \|B^\ell x + s(B^\ell \check{\varphi} - \theta \zeta)\| \leq \sup_{x:\langle x, \check{\varphi}\rangle = 0, \|x\|=1} \|B^\ell x\|.$$

Hence, Proposition 5 implies that w.h.p.

$$\|R\| \leq (\log n)^c \alpha^{\ell/2}. \tag{16}$$

We may now apply the Bauer-Fike theorem. If $\lambda_i = \lambda_i(B)$, we find that w.h.p.

$$|\lambda_1 - \alpha| = O(1/\ell), \qquad |\lambda_2| \leq \Big(C(\log n)^c \alpha^{\ell/2}\Big)^{1/\ell} = \sqrt{\alpha} + O\bigg(\frac{\log\log n}{\log n}\bigg),$$

and the normalized Perron eigenvector ξ of B satisfies w.h.p. $\|\xi - \zeta\| = O((\log n)^c \alpha^{-\ell/2})$. This concludes the proof of Theorem 1.

Proposition 4 will follow from a *local analysis*. Namely the statistics of node neighborhoods up to distance ℓ in the original random graph will be related by coupling to a Galton-Watson branching process; relevant properties of the corresponding Galton-Watson process will be established; finally we shall deduce weak laws of large numbers for the ℓ-neighborhoods of the random graph from the estimations performed on the branching process combined with some asymptotic decorrelation property between distinct node neighborhoods. This is done in Section 9 of [17] which contains a proof of Proposition 4.

The proof of Proposition 5 relies crucially on a matrix expansion given in Proposition 6, which extends the argument introduced in [6] for matrices counting self-avoiding walks to the present setup where non-backtracking walks instead are considered. We now introduce some notation to state it.

A. Matrix expansion for B^ℓ

For convenience we extend matrix B and vector χ to $\mathbb{R}^{\vec{E}(V)}$ where $\vec{E}(V) = \{(u, v) : u \neq v \in V\}$ is the set of directed edges of the *complete graph*. We set for all $e, f \in \mathbb{R}^{\vec{E}(V)}$, $\chi(e) = 1$ and

$$B_{ef} = A_e A_f \mathbf{1}(e_2 = f_1)\mathbf{1}(e_1 \neq f_2),$$

where A is the graph's adjacency matrix. For integer $k \geq 1$, $e, f \in \vec{E}(V)$, we define Γ^k_{ef} as the set of non-backtracking walks $\gamma = (\gamma_0, \ldots, \gamma_k)$ of length k starting from $(\gamma_0, \gamma_1) = e$ and ending at $(\gamma_{k-1}, \gamma_k) = f$ in the complete graph on the vertex set V. We have that

$$(B^k)_{ef} = \sum_{\gamma \in \Gamma^{k+1}_{ef}} \prod_{s=0}^{k} A_{\gamma_s \gamma_{s+1}}.$$

We associate to each walk $\gamma = (\gamma_0, \ldots, \gamma_k)$, a graph $G(\gamma) = (V(\gamma), E(\gamma))$ with vertex set $V(\gamma) = \{\gamma_i, 0 \leq i \leq k\}$ and edge set $E(\gamma)$ the set of distinct visited edges $\{\gamma_i, \gamma_{i+1}\}, 0 \leq i \leq k-1$. Following [7], we say that a graph H is *tangle-free* (or ℓ-tangle free to make the dependence in ℓ explicit) if every neighborhood of radius ℓ in H contains at most one cycle.

Otherwise, H is said to be tangled. We say that γ is tangle-free or tangled if $G(\gamma)$ is. Obviously, if G is tangle-free and $1 \le k \le \ell$ then $B^k = B^{(k)}$, where

$$B_{ef}^{(k)} = \sum_{\gamma \in F_{ef}^{k+1}} \prod_{s=0}^{k} A_{\gamma_s \gamma_{s+1}},$$

and F_{ef}^{k+1} is the subset of tangle-free paths in Γ_{ef}^{k+1}. For $u \neq v$, we set $\underline{A}_{uv} = A_{uv} - \frac{\alpha}{n}$.

We define similarly the matrix $\Delta^{(k)}$ on $\mathbb{R}^{\vec{E}(V)}$

$$\Delta_{ef}^{(k)} = \sum_{\gamma \in F_{ef}^{k+1}} \prod_{s=0}^{k} \underline{A}_{\gamma_s \gamma_{s+1}}.$$

The matrix $\Delta^{(k)}$ can be thought of as an attempt to center the non-backtracking matrix B^k when the underlying graph is tangle-free. We use the convention that a product over an empty set is equal to 1. We also set

$$\Delta_{ef}^{(0)} = \mathbf{1}(e = f)\underline{A}_e \quad \text{and} \quad B_{ef}^{(0)} = \mathbf{1}(e = f)A_e. \tag{17}$$

Notably, $B^{(0)}$ is the projection on $\vec{E}$. We have the following telescopic sum decomposition.

$$B_{ef}^{(\ell)} = \Delta_{ef}^{(\ell)} + \sum_{t=0}^{\ell} \sum_{\gamma \in F_{ef}^{\ell+1}} \prod_{s=0}^{t-1} \underline{A}_{\gamma_s \gamma_{s+1}} \left(\frac{\alpha}{n}\right) \prod_{s=t+1}^{\ell} A_{\gamma_s \gamma_{s+1}}. \tag{18}$$

Indeed,

$$\prod_{s=0}^{\ell} x_s = \prod_{s=0}^{\ell} y_s + \sum_{t=0}^{\ell} \prod_{s=0}^{t-1} y_s (x_t - y_t) \prod_{s=t+1}^{\ell} x_s.$$

We denote by K the non-backtracking matrix of the complete graph on V. For $0 \le t \le \ell$, we define $R_t^{(\ell)}$ via

$$(R_t^{(\ell)})_{ef} = \sum_{\gamma \in F_{t,ef}^{\ell+1}} \prod_{s=0}^{t-1} \underline{A}_{\gamma_s \gamma_{s+1}} \prod_{s=t+1}^{\ell} A_{\gamma_s \gamma_{s+1}},$$

where for $1 \le t \le \ell - 1$, $F_{t,ef}^{\ell+1} \subset \Gamma_{ef}^{\ell+1}$ is the set of non-backtracking tangled paths $\gamma = (\gamma_0, \ldots, \gamma_{\ell+1}) = (\gamma', \gamma'') \in \Gamma_{ef}^{\ell+1}$ with $\gamma' = (\gamma_0, \ldots, \gamma_t) \in F_{eg}^t$, $\gamma'' = (\gamma_{t+1}, \ldots, \gamma_{\ell+1}) \in F_{g'f}^{\ell-t}$ for some $g, g' \in \vec{E}(V)$. For $t = 0$, $F_{0,ef}^{\ell+1}$ is the set of non-backtracking tangled paths $\gamma = (\gamma', \gamma'')$ with $\gamma' = e_1$, $\gamma'' = (\gamma_1, \ldots, \gamma_{\ell+1}) \in F_{g'f}^{\ell}$ for some $g' \in \vec{E}(V)$ (necessarily $g_1' = e_2$). Similarly, for $t = \ell$, $F_{\ell,ef}^{\ell+1}$ is the set of non-backtracking tangled paths $\gamma = (\gamma_0, \ldots, \gamma_{\ell+1}) = (\gamma', \gamma'')$ with $\gamma'' = f_2$, $\gamma' = (\gamma_0, \ldots, \gamma_\ell) \in F_{eg}^{\ell}$ for some $g \in \vec{E}(V)$ (necessarily $g_2 = f_1$).

We define $L = K^2 - \chi\chi^*$, (L is nearly the orthogonal projection of K^2 on $\chi^\perp$). We further denote for $1 \le t \le \ell - 1$ $S_t^{(\ell)} = \Delta^{(t-1)} L B^{(\ell-t-1)}$. A simple argument then gives:

Proposition 6. *With the above notations matrix $B^{(\ell)}$ admits the following expansion*

$$B^{(\ell)} = \Delta^{(\ell)} + \frac{\alpha}{n} K B^{(\ell-1)} + \frac{\alpha}{n} \sum_{t=1}^{\ell-1} \Delta^{(t-1)} K^2 B^{(\ell-t-1)} + \frac{\alpha}{n} \Delta^{(\ell-1)} K - \frac{\alpha}{n} \sum_{t=0}^{\ell} R_t^{(\ell)}. \tag{19}$$

If G is tangle-free, for any normed vector $x \in \mathbb{C}^{\vec{E}(V)}$, one has

$$\begin{aligned} \|B^\ell x\| \le\; & \|\Delta^{(\ell)}\| + \frac{\alpha}{n} \|K B^{(\ell-1)}\| + \frac{\alpha}{n} \sum_{t=1}^{\ell-1} \|\Delta^{(t-1)} \chi\| \left| \langle \chi, B^{(\ell-t-1)} x \rangle \right| \\ & + \frac{\alpha}{n} \sum_{t=1}^{\ell-1} \|S_t^{(\ell)}\| + \alpha \|\Delta^{(\ell-1)}\| + \frac{\alpha}{n} \sum_{t=0}^{\ell} \|R_t^{(\ell)}\|. \end{aligned} \tag{20}$$

B. Proof of Proposition 5

The following proposition is established in Section 6 of [17] using path counting combinatorial arguments.

Proposition 7. *Let $\ell \sim \kappa \log_\alpha n$ with $\kappa \in (0, 1/6)$. With high probability, the following norm bounds hold for all k, $0 \le k \le \ell$:*

$$\|\Delta^{(k)}\| \le (\log n)^{10} \alpha^{k/2}, \tag{21}$$

$$\|\Delta^{(k)}\chi\| \le (\log n)^{5} \alpha^{k/2}\sqrt{n}, \tag{22}$$

$$\|R_k^{(\ell)}\| \le (\log n)^{25} \alpha^{\ell - k/2}, \tag{23}$$

$$\|B^{(k)}\| \le (\log n)^{10} \alpha^{k} \quad \text{and} \quad \|KB^{(k)}\| \le \sqrt{n}(\log n)^{10} \alpha^{k}, \tag{24}$$

and the following bound holds for all k, $1 \le k \le \ell - 1$:

$$\|S_k^{(\ell)}\| \le \sqrt{n}(\log n)^{20} \alpha^{\ell - k/2}. \tag{25}$$

Together with Propositions 6 and 7, we shall also need the next two results, established by local analysis in Section 9 of [17].

Lemma 8. *For $\ell \sim \kappa \log_\alpha n$ with $\kappa < 1/2$, w.h.p. the random graph G is ℓ-tangle-free.*

For the Erdős-Rényi graph, we have:

Proposition 9. *For $\ell \sim \kappa \log_\alpha n$ with $\kappa < 1/2$, w.h.p. , for any $0 \le t \le \ell - 1$, it holds that*

$$\sup_{\|x\|=1, \langle B^\ell \chi, x\rangle = 0} \left|\langle B^t \chi, x\rangle\right| \le (\log n)^5 n^{1/2} \alpha^{t/2}.$$

We now have all the ingredients necessary to prove Proposition 5. In view of Lemma 8, we may use the bound (20) of Proposition 6 and take the supremum over of all x, $\|x\| = 1$, $\langle \check{\varphi}, x\rangle = \langle \chi, B^\ell x\rangle = 0$. By the norm bounds (21)-(23)-(25) of Proposition 7, w.h.p.

$$\alpha\|\Delta^{(\ell-1)}\| + \|\Delta^{(\ell)}\| + \frac{\alpha}{n}\sum_{t=0}^{\ell}\|R_t^{(\ell)}\| + \frac{\alpha}{n}\sum_{t=1}^{\ell-1}\|S_t^{(\ell)}\| \le C(\log n)^c \alpha^{\ell/2}(1 + \alpha^{\ell/2}/\sqrt{n}) = O((\log n)^c \alpha^{\ell/2}).$$

Also, from (12), since $\check{\chi} = \chi$,

$$\sup_{\|x\|=1, \langle \chi, B^\ell x\rangle = 0} \left|\langle \chi, B^{(t)} x\rangle\right| = \sup_{\|x\|=1, \langle \chi, B^\ell \check{x}\rangle = 0} \left|\langle \chi, B^{(t)} \check{x}\rangle\right| = \sup_{\|x\|=1, \langle B^\ell \chi, x\rangle = 0} \left|\langle B^{(t)} \chi, x\rangle\right|.$$

Hence, from Proposition 9 and norm bound (22), w.h.p.

$$\|\Delta^{(t-1)}\chi\| |\langle \chi, B^{(\ell-t-1)} x\rangle| \quad \le \quad C(\log n)^c n \alpha^{\ell/2}$$

Hence, w.h.p.

$$\frac{\alpha}{n}\sum_{t=1}^{\ell-1}\|\Delta^{(t-1)}\| |\langle \chi, B^{(\ell-t-1)} x\rangle| = O\Big((\log n)^{c+1} \alpha^{\ell/2}\Big).$$

It remains to use norm bound (24) to deal with the term $\|KB^{(\ell-1)}\|/n$ in (20) to conclude the proof of Proposition 5.

IV. Stochastic Block Model: proof strategy for Theorem 2

In this section, we give a high level overview of the proof structure in the case of the stochastic block model. We refer to [17] for all details and proofs.

As in the Erdős-Rényi case, we use *perturbation theory of linear operator*: we apply an extension of the Bauer-Fike theorem to B^k for a sufficient large k. The main intuition is that starting with a vector correlated with a particular eigenvector and *iterating powers* of the matrix will give a vector quite close to the actual eigenvector. Further, if the eigenvectors are orthogonal this can be done simultaneously for all vectors. In our case, the candidate eigenvectors are the extensions χ_k of the eigenvectors ϕ_k of the progeny matrix M. Note that:

$$\langle \chi_k, \chi_\ell\rangle \quad = \quad \sum_{v\in V} \deg(v)\phi_k(\sigma(v))\phi_\ell(\sigma(v)) \simeq n\alpha \sum_j \pi(j)\phi_k(j)\phi_\ell(j) = 0,$$

where the last equality follows from (3).

More precisely, let $\ell = \ell(n) \sim \kappa \log_\alpha n$ for some $\kappa \in (0, \gamma/6)$ as in Theorem 2. Recalling the definition (8) of vector χ_k, we further introduce for all $k \in [r]$:

$$\varphi_k = \frac{B^\ell \chi_k}{\|B^\ell \chi_k\|}, \quad \theta_k = \|B^\ell \check{\varphi}_k\|, \text{ and } \zeta_k = \frac{B^\ell \check{\varphi}_k}{\theta_k} = \frac{B^\ell B^{*\ell} \check{\chi}_k}{\|B^\ell B^{*\ell} \check{\chi}_k\|}. \tag{26}$$

(in the above, if $\theta_k = 0$, we set $\zeta_k = 0$). We also define

$$H = \mathrm{span}(\check{\varphi}_k, k \in [r]).$$

The generalization of propositions 4 and 5 are now:

Proposition 10. *For some $b, c > 0$, w.h.p.*

(i) $b|\mu_k^\ell| \le \theta_k \le c|\mu_k^\ell|$ *if* $k \in [r_0]$,
(ii) $\mathrm{sign}(\mu_k^\ell)\langle \zeta_k, \check{\varphi}_k \rangle \ge b$ *if* $k \in [r_0]$,
(iii) $\theta_k \le (\log n)^c \alpha^{\ell/2}$ *if* $k \in [r] \backslash [r_0]$,
(iv) $|\langle \varphi_j, \varphi_k \rangle| \le (\log n)^c \alpha^{3\ell/2} n^{-\gamma/2}$ *if* $k \neq j \in [r]$,
(v) $|\langle \zeta_j, \check{\varphi}_k \rangle| \le (\log n)^c \alpha^{2\ell} n^{-\gamma/2}$ *if* $k \neq j \in [r_0]$.
(vi) $|\langle \zeta_j, \zeta_k \rangle| \le (\log n)^c \alpha^{5\ell/2} n^{-\gamma/2}$ *if* $k \neq j \in [r_0]$.

Proposition 10 will follow from the local analysis done in Section 9 of [17]. The next Proposition will be established in Section 10 of [17] using a matrix expansion together with norm bounds derived by combinatorial arguments parallel to the proof of Proposition 5 for the Erdős-Rényi graph.

Proposition 11. *For some $c > 0$, w.h.p.*

$$\sup_{x \in H^\perp, \|x\|=1} \|B^\ell x\| \le (\log n)^c \alpha^{\ell/2}.$$

We now check that the two preceding propositions imply Theorem 2. We consider $(\bar{\varphi}_1, \cdots, \bar{\varphi}_{r'})$ obtained by the Gram-Schmidt orthonormalization of $(\check{\varphi}_1, \cdots, \check{\varphi}_r)$. By Proposition 10(iv), w.h.p. $r' = r$ and for all $k \in [r]$,

$$\|\check{\varphi}_k - \bar{\varphi}_k\| = O((\log n)^c \alpha^{3\ell/2} n^{-\gamma/2}). \tag{27}$$

Similarly, for $k \in [r_0]$, we denote by $\tilde{\zeta}_k$ the orthogonal projection of ζ_k on the orthogonal of the vector space spanned by $\bar{\varphi}_j$, $j \in [r_0]$, $j \neq k$ and $\tilde{\zeta}_j$, $j < k$. We set $\bar{\zeta}_k = \tilde{\zeta}_k / \|\tilde{\zeta}_k\|$. From Proposition 10(v)-(vi), we find w.h.p. for $k \in [r_0]$,

$$\|\zeta_k - \bar{\zeta}_k\| = O((\log n)^c \alpha^{5\ell/2} n^{-\gamma/2}). \tag{28}$$

We then set

$$D_0 = \sum_{k=1}^{r_0} \theta_k \bar{\zeta}_k \bar{\varphi}_k^*.$$

Since $\|\check{\varphi}_k - \bar{\varphi}_k\| = o(1)$, from Proposition 10(i)-(iii), we find by induction on $k \in [r]$, w.h.p. for all $k \in [r]$,

$$\|B^\ell \bar{\varphi}_k\| = O(\alpha^\ell).$$

Consequently, from Proposition 11, we have w.h.p.

$$\|B^\ell\| = O(\alpha^\ell).$$

In particular, since $D_0 \bar{\varphi}_k = \theta_k \bar{\zeta}_k = B^\ell \check{\varphi}_k + \theta_k(\bar{\zeta}_k - \zeta_k)$, we get for $k \in [r_0]$,

$$\begin{aligned} \|B^\ell \bar{\varphi}_k - D_0 \bar{\varphi}_k\| &\le \|B^\ell\| \|\bar{\varphi}_k - \check{\varphi}_k\| + \|B^\ell \check{\varphi}_k - D_0 \bar{\varphi}_k\| + \theta_k \|\bar{\zeta}_k - \zeta_k\| \\ &= O((\log n)^c \alpha^{7\ell/2} n^{-\gamma/2}). \end{aligned}$$

We have $\alpha^{7\ell/2} n^{-\gamma/2} = n^{7\kappa/2 + o(1) - \gamma/2}$. Since $0 < \kappa < \gamma/6$, $7\kappa/2 - \gamma/2 < \kappa/2$, we thus obtain, if P_0 is the orthogonal projection of $H_0 = \mathrm{span}(\bar{\varphi}_k, k \in [r_0])$,

$$\|B^\ell P_0 - D_0\| = O(\alpha^{\ell/2}). \tag{29}$$

We also set $D_1 = B^\ell P_1$ where P_1 is the orthogonal projection of $H_1 = \mathrm{span}(\bar{\varphi}_k, k \in [r] \backslash [r_0])$ and $C = B^\ell - D_0 - D_1$. Arguing similarly, from Proposition 10(iii), w.h.p. , for $k \in [r] \backslash [r_0]$,

$$\|D_1 \bar{\varphi}_k\| = \|B^\ell \bar{\varphi}_k\| \le \|B^\ell\| \|\bar{\varphi}_k - \check{\varphi}_k\| + \|B^\ell \check{\varphi}_k\| = O((\log n)^c \alpha^{\ell/2}).$$

Hence

$$\|D_1\| = O((\log n)^c \alpha^{\ell/2}). \tag{30}$$

Also, let $y \in \mathbb{R}^{\vec{E}}$ with $\|y\| = 1$. We write $y = x + h_0 + h_1$ with $x \in H^\perp$, $h_1 \in H_1$, $h_0 \in H_0 = \mathrm{span}(\varphi_k, k \in [r_0])$. We find

$$\|Cy\| = \|B^\ell x + (B^\ell - D_0)h_0\| \le \sup_{x \in H^\perp, \|x\|=1} \|B^\ell x\| + \|B^\ell P_0 - D_0\|.$$

Hence, Proposition 11 and (29)-(30) imply that w.h.p.

$$\|C\| = O((\log n)^c \alpha^{\ell/2}).$$

We decompose $B^\ell = D_0 + R$ with $R = C + D_1$, from what precedes w.h.p.

$$\|R\| = O((\log n)^c \alpha^{\ell/2}).$$

We are now in position to apply the Bauer-Fike theorem (see Proposition 8 in [17]). From (28), the statement of Proposition 10(ii) also holds with ζ_k replaced by $\bar{\zeta}_k$. It readily implies Theorem 2.

The *community detection* algorithm is essentially the same as in [6]. To make the argument more transparent, let's assume that we have obtained estimates of the eigendecomposition of the expected adjacency matrix $\mathbb{E}(A)$. Using the membership matrix $Z \in \{0,1\}^{n \times r}$, where $Z_{u,\ell} = \mathbf{1}(\sigma(u) = \ell)$, we can write:

$$\mathbb{E}(A) = \frac{1}{n} Z W Z^* = \frac{1}{n} \sum_{k=1}^{r} \mu_k (Z\phi_k)(Z\phi_k)^*.$$

Note that $\frac{1}{n}\langle Z\phi_k, Z\phi_j\rangle = \phi_k \Pi \phi_j^* = \delta_{kj}$ by (3). The vector $Z\phi_k$ is constant within each class and further $\langle 1, Z\phi_k\rangle = 0$ for $k \neq 1$ by (6). Hence, $Z\phi_k$ will have both positive and negative coordinates. Let I^+ be the indices of $Z\phi_k$ with positive and respectively I^- the indices with negative coordinates, then partitioning the subset of vertices with coordinates in I^+ and I^- respectively allows us to identify a sub-partition of the true assignment. Thus, if one recover a vector that is 'aligned' with one of the eigenvector $Z\phi_k$ with $k \neq 1$, then it is possible to recover an assignment that is correlated with the true assignment. The same reasoning is applied to the eigenvector of the non-backtracking matrix after averaging for each vertex over all incident edges.

Acknowledgment

M.L. acknowledges the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-11-JS02-005-01 (GAP project).

References

[1] K.-i. Hashimoto, "Zeta functions of finite graphs and representations of p-adic groups," in *Automorphic forms and geometry of arithmetic varieties*, ser. Adv. Stud. Pure Math. Academic Press, Boston, MA, 1989, vol. 15, pp. 211–280.

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

[3] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, "Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications," *Physics Review E*, vol. 84:066106, 2011.

[4] E. Mossel, J. Neeman, and A. Sly, "Reconstruction and estimation in the planted partition model," Feb. 2012, available at: http://arxiv.org/abs/1202.1499.

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

[6] L. Massoulié, "Community detection thresholds and the weak Ramanujan property," *ACM Symposium on the Theory of Computing (STOC)*, 2014, see also http://arxiv.org/abs/arXiv:1311.3085v1.

[7] E. Mossel, J. Neeman, and A. Sly, "A proof of the block model threshold conjecture," 2013, arXiv:1311.4115v2.

[8] A. Lubotzky, R. Phillips, and P. Sarnak, "Ramanujan graphs," *Combinatorica*, no. 9, pp. 261–277, 1988.

[9] R. Murty, "Ramanujan graphs," *J. Ramanujan Math. Soc.*, vol. 18, no. 1, pp. 1–20, 2003.

[10] A. Nilli, “On the second eigenvalue of a graph,” *Discrete Math.*, vol. 91, no. 2, pp. 207–210, 1991. [Online]. Available: http://dx.doi.org/10.1016/0012-365X(91)90112-F

[11] J. Friedman, “A proof of Alon’s second eigenvalue conjecture and related problems,” *Mem. Amer. Math. Soc.*, vol. 195, no. 910, pp. viii+100, 2008. [Online]. Available: http://dx.doi.org/10.1090/memo/0910

[12] A. Lubotzky, “Cayley graphs: eigenvalues, expanders and random walks,” *Surveys in Combinatorics, London Math. Soc. Lecture Notes*, vol. 218, pp. 155–189, 1995.

[13] M. D. Horton, H. M. Stark, and A. A. Terras, “What are zeta functions of graphs and what are they good for?” *Contemporary Mathematics, Quantum Graphs and Their Applications; Edited by Gregory Berkolaiko, Robert Carlson, Stephen A. Fulling, and Peter Kuchment*, vol. 415, pp. 173–190, 2006.

[14] J. Friedman and D.-E. Kohler, “The relativized second eigenvalue conjecture of alon,” 2014, arXiv:1403.3462.

[15] C. Bordenave, “A new proof of friedman’s second eigenvalue theorem and its extension to random lifts,” 2015, arXiv:1502.04482.

[16] B. Bollobás, S. Janson, and O. Riordan, “The phase transition in inhomogeneous random graphs,” *Random Structures Algorithms*, vol. 31, no. 1, pp. 3–122, 2007. [Online]. Available: http://dx.doi.org/10.1002/rsa.20168

[17] C. Bordenave, M. Lelarge, and L. Massoulié, “Non-backtracking spectrum of random graphs: community detection and non-regular Ramanujan graphs,” *arXiv preprint arXiv:1501.06087*, 2015.