Non-Backtracking Spectrum of Degree-Corrected Stochastic Block Models

Lennart Gulikers[1], Marc Lelarge[2], and Laurent Massoulié[3]

1 **Microsoft Research - INRIA Joint Centre, Palaiseau and École Normale Supérieure, Paris, France**
`lennart.gulikers@inria.fr`

2 **INRIA Paris and École Normale Supérieure, Paris, France**
`marc.lelarge@ens.fr`

3 **Microsoft Research - INRIA Joint Centre, Palaiseau, France**
`laurent.massoulie@inria.fr`

Abstract

Motivated by community detection, we characterise the spectrum of the non-backtracking matrix B in the Degree-Corrected Stochastic Block Model.

Specifically, we consider a random graph on n vertices partitioned into two asymptotically equal-sized clusters. The vertices have i.i.d. weights $\{\phi_u\}_{u=1}^n$ with second moment $\Phi^{(2)}$. The intra-cluster connection probability for vertices u and v is $\frac{\phi_u \phi_v}{n} a$ and the inter-cluster connection probability is $\frac{\phi_u \phi_v}{n} b$.

We show that with high probability, the following holds: The leading eigenvalue of the non-backtracking matrix B is asymptotic to $\rho = \frac{a+b}{2}\Phi^{(2)}$. The second eigenvalue is asymptotic to $\mu_2 = \frac{a-b}{2}\Phi^{(2)}$ when $\mu_2^2 > \rho$, but asymptotically bounded by $\sqrt{\rho}$ when $\mu_2^2 \leq \rho$. All the remaining eigenvalues are asymptotically bounded by $\sqrt{\rho}$. As a result, a clustering positively-correlated with the true communities can be obtained based on the second eigenvector of B in the regime where $\mu_2^2 > \rho$.

In a previous work we obtained that detection is impossible when $\mu_2^2 < \rho$, meaning that there occurs a phase-transition in the sparse regime of the Degree-Corrected Stochastic Block Model.

As a corollary, we obtain that Degree-Corrected Erdős-Rényi graphs asymptotically satisfy the graph Riemann hypothesis, a quasi-Ramanujan property.

A by-product of our proof is a weak law of large numbers for local-functionals on Degree-Corrected Stochastic Block Models, which could be of independent interest.

1998 ACM Subject Classification E.1 Graphs and Networks

Keywords and phrases Degree-Corrected Stochastic Block Model, Non-backtracking Matrix, Ramanujan Graphs

Digital Object Identifier 10.4230/LIPIcs.ITCS.2017.44

1 Introduction

The non-backtracking matrix B of a graph $G = (V, E)$ is indexed by the set of its oriented edges $\vec{E} = \{(u, v) : \{u, v\} \in E\}$. For $e = (e_1, e_2), f = (f_1, f_2) \in \vec{E}$, B is defined as

$$B_{ef} = 1_{e_2 = f_1} 1_{e_1 \neq f_2}.$$

This matrix was introduced by Hashimoto [10] in 1989.

We study the spectrum of B when G is a random graph generated according to the Degree-Corrected Stochastic Block Model (DC-SBM) [11]. We characterise its leading eigenvalues

8th Innovations in Theoretical Computer Science Conference (ITCS 2017).
Editor: Christos H. Papadimitrou; Article No. 44; pp. 44:1–44:27
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

and corresponding eigenvectors when the number of vertices in G tends to infinity. Our motivation stems from community detection problems: experiments in [14] show that the spectral method based on the non-backtracking matrix seems to work well on real datasets. We test the robustness of this method and show in particular that, above a certain threshold, the second eigenvector of B is correlated with the underlying communities.

The DC-SBM [11] is an extension of the *ordinary* Stochastic Block Model (SBM) [8]. The latter model has as a drawback that vertices in the same community are stochastically indistinguishable and it therefore fails to accurately describe networks with high heterogeneity. Compare this to fitting a straight line on intrinsically curved data, which is doomed to miss important information. The DC-SBM is a more realistic model: it allows for very general degree-sequences.

The special case of the DC-SBM under consideration here is defined as follows: It is a random graph on n vertices partitioned into two asymptotically equal-sized clusters. The vertices have bounded i.i.d. weights $\{\phi_u\}_{u=1}^n$ with second moment $\Phi^{(2)}$. The intra-cluster connection probability for vertices u and v is $\frac{\phi_u\phi_v}{n}a$ and the inter-cluster connection probability is $\frac{\phi_u\phi_v}{n}b$, for two constants $a, b > 0$.

Note that those graphs are thus sparse, which is a challenging regime for community detection. Indeed, in the *ordinary* SBM (obtained by putting $\phi_1 = \ldots = \phi_n = 1$), an instance of the graph might not contain enough information to distinguish between the two clusters if the difference between a and b is small. More precisely, reconstruction is impossible when $(a-b)^2 \leq 2(a+b)$ [18]. Interestingly, positively-correlated reconstruction can be obtained by thresholding the second-eigenvector of B [2] immediately above the threshold (i.e., $(a-b)^2 > 2(a+b)$). The SBM thus has a phase-transition in its sparse regime.

Does the DC-SBM exhibit a similar behaviour? We showed in an earlier work [6] that detection is impossible when $(a-b)^2\Phi^{(2)} \leq 2(a+b)$. In our current work we analyse the regime where $(a-b)^2\Phi^{(2)} > 2(a+b)$. We answer the following questions: is detection possible in this regime and if so, can we use again the non-backtracking matrix or do we need to modify it? *A priori this is unclear, because an algorithm solely based on B cannot use any information on the weights as input.* Our main result shows that the spectral method based on the non-backtracking matrix (thus the same method as in [2]) successfully detects communities in the regime $(a-b)^2\Phi^{(2)} > 2(a+b)$. Surprisingly, no modification of the matrix, nor information about the weights is needed (compare this to the adjacency matrix, which needs to be adapted to the degree-corrected setting [7]), which shows the robustness of the method. Moreover as in the standard SBM, the algorithm is optimal in the sense that it works all the way down to the detectability-threshold.

Informally, we have the following results: With high probability, the leading eigenvalue of the non-backtracking matrix B is asymptotic to $\rho = \frac{a+b}{2}\Phi^{(2)}$. The second eigenvalue is asymptotic to $\mu_2 = \frac{a-b}{2}\Phi^{(2)}$ when $\mu_2^2 > \rho$, but asymptotically bounded by $\sqrt{\rho}$ when $\mu_2^2 \leq \rho$. All the remaining eigenvalues are asymptotically bounded by $\sqrt{\rho}$. Further, a clustering positively-correlated with the true communities can be obtained based on the second eigenvector of B in the regime where $\mu_2^2 > \rho$ (i.e., precisely when $(a-b)^2\Phi^{(2)} > 2(a+b)$).

A side-result is that Degree-Corrected Erdős-Rényi graphs asymptotically satisfy the graph Riemann hypothesis, a quasi-Ramanujan property.

In our proof we derive and use a weak law of large numbers for local-functionals on Degree-Corrected Stochastic Block Models, which could be of independent interest.

1.1 Community detection background

In this paper we are interested in community detection: The problem of clustering vertices in a graph into groups of "similar" nodes. In particular, the graphs here are generated according to the DC-SBM and the goal is to retrieve the spin (or group-membership) of the nodes based on a single observation of the DC-SBM.

When the average degree of a vertex grows sufficiently fast with the size of the network (i.e., the average degree is $\Omega(\log(n))$), we speak about dense networks. Community-detection is then well understood and we consider instead sparse graphs where the average degree is bounded by a constant. This setting is more realistic as most real networks are sparse, but is at the same time more challenging. Indeed, traditional methods based on the Adjacency or Laplacian matrix working well in the dense case break down when employed in the sparse case.

In the sparse regime, with high probability, at least a positive fraction of the nodes is isolated. Consequently, one cannot hope to find the community-membership of *all* vertices. We therefore address here the problem of finding a clustering that is positively correlated with the true community-structure.

In [3] it was first conjectured that a detectability phase transition exists in the *ordinary* SBM: When $(a-b)^2 > 2(a+b)$, the belief propagation algorithm would succeed in finding such a positively correlated clustering. Conversely, due to a lack of information, detection would be impossible when $(a-b)^2 \leq 2(a+b)$.

In [18], impossibility of reconstruction when $(a-b)^2 \leq 2(a+b)$ is shown for the SBM. This paper builds further on a tree-reconstruction problem in [4].

The authors of [14] conjectured that detection using the second eigenvector of B would succeed all the way down to the conjectured detectability threshold. Two variants of this so-called spectral redemption conjecture were proven before the work in [2] appeared:

In [16] it is shown that detection based on the second eigenvector of a matrix counting self-avoiding paths in the graph leads to consistent recovery when $\left(\frac{a-b}{2}\right)^2 > \frac{a+b}{2}$.

Independently, in [17], the authors prove the positive side of the conjecture by using a constructing based on counting non-backtracking paths in graphs generated according to the SBM.

More recently, in [2] the spectral redemption conjecture is proved. This work moreover determines the limits of community detection based on the non-backtracking spectrum in the presence of an arbitrary number of communities.

Here we extend the work in [2] to the more general setting of the DC-SBM.

1.2 Quasi Ramanujan property

Following the definition introduced in [15], a k-regular graph is Ramanujan if its second largest absolute eigenvalue is no larger than $2\sqrt{k-1}$. In [9], a graph is said to satisfy the graph Riemann hypothesis if B has no eigenvalues λ such that $|\lambda| \in (\sqrt{\rho_B}, \rho_B)$, where ρ_B is the Perron-Frobenius eigenvalue of B. The graph Riemann hypothesis can be seen as a generalization of the Ramanujan property, because a regular graph satisfies the graph Riemann hypothesis if and only if it has the Ramanujan property [9, 19].

Now, put $a = b = 1$ to obtain a Degree-Corrected Erdős-Rényi graph where vertices u and v are connected by an edge with probability $\frac{\phi_u \phi_v}{n}$. Our results imply that, with high probability, $\rho_B = \Phi^{(2)} + o(1)$, while all other eigenvalues are in absolute value smaller than $\sqrt{\Phi^{(2)}} + o(1)$. Consequently, these Degree-Corrected Erdős-Rényi graphs asymptotically satisfy the graph Riemann hypothesis.

1.3 Outline and main differences with ordinary SBM

We follow the same general approach as in [2]. We focus primarily on the differences and complications here: we often omit or shorten the proof of a statement if it may be proven in a very similar way.

In Section 2 we define the DC-SBM and state the assumptions we make. This is then followed by Theorem 1 on the spectrum of B and its consequences for community detection, Theorem 2.

In Section 3, we give the necessary background on non-backtracking matrices. Further, we give an extension of the Bauer-Fike Theorem, that first appeared in [2].

In Section 4 we give the proof of Theorem 1. It builds on Propositions 4 and 5. Their proofs are deferred to later sections.

In Section 5 we consider two-type branching process where the offspring distribution is governed by a Poisson mixture to capture the weights of the vertices. We associate two martingales to this process and extend limiting results by Kesten and Stigum [12, 13]. Hoeffding's inequality plays an important role here to prove concentrations results for the weights. Further, we define a cross-generational functional on these branching processes that is correlated with the spin of the root.

In Section 6 we state a coupling between local neighbourhoods and the branching process with weights in Section 5. We established this coupling in an earlier work [6], it is technically more involved than the ordinary coupling on graphs with unit weight. It is crucial that the weights in the graph and the branching process are perfectly coupled. We further establish a growth condition on the local neighbourhoods, using a stochastic domination argument that is more involved than its analogue in unweighed graphs.

In Section 7 we define local functionals that map graphs, together with their spins *and* weights to the real numbers. We establish, using Efron-Stein's inequality, a weak law of large numbers for those functionals, which could be of independent interest. Part of the work here is again hidden in the coupling from [6].

In Section 8 we apply those local functionals to establish Proposition 4.

In Section 9 we decompose powers of the matrix B as a sum of products. This technique appeared first in [16] for matrices counting self-avoiding paths and was elaborated in [2]. To bound the norm of the individual matrices occurring in the decomposition, we use the trace method initiated in [5]. In doing so, we need to bound the expectation of products of higher moments of the weights over certain paths. This is a significant complication with respect to the ordinary SBM, see Section 9.2 for a comparison.

In Section 10 we prove that positively correlated clustering is possible based on the second eigenvector of B, i.e., Theorem 2. We use the symmetry present in the two-communities setting here, which gets in general broken in models with more than two communities.

Detailed proofs of the statements in Sections 5, 6, 7, 9 and 10 can be found in Appendices A–E in the detailed version of the underlying article: arXiv:1609.02487.

In each section we give a detailed comparison with the ordinary SBM.

2 Main Results

We define our model more precisely and state the two main theorems.

We consider random graphs on n nodes $V = \{1, \ldots, n\}$ drawn according to the Degree-Corrected Stochastic Block Model [11]. The vertices are partitioned into two clusters of sizes n_+ and n_- by giving each vertex v a spin $\sigma(v)$ from $\{+, -\}$. The vertices have i.i.d. weights

$\{\phi_u\}_{u=1}^n$ governed by some law ν with support in $[\phi_{\min}, \phi_{\max}]$, where $0 < \phi_{\min} \leq \phi_{\max} < \infty$ are constants. We denote the second moment of the weights by $\Phi^{(2)}$. An edge is drawn between nodes u and v with probability $\frac{\phi_u \phi_v}{n} a$ when u and v have the same spin and with probability $\frac{\phi_u \phi_v}{n} b$ otherwise. The model parameters a and b are constant. We assume that for some constant $\gamma \in (0, 1]$,

$$n_{\pm} = \frac{n}{2} + \mathcal{O}(n^{1-\gamma}), \tag{1}$$

i.e., the communities have nearly equal size.

The ordinary SBM on two or more communities was first introduced in [8], which is a generalization of Erdős-Rényi graphs. The Degree-Corrected SBM appeared first in [11]. General inhomogeneous random graphs are considered in [1].

Note that we retrieve the two-communities ordinary SBM by giving all nodes *unit* weight.

Local neighbourhoods in the sparse graphs under consideration are tree-like with high probability. In [6] we showed that these trees are distributed according to a Poisson-mixture two-type branching process, detailed in Section 5 below. We denote the mean progeny matrix of the branching process by

$$M = \frac{\Phi^{(2)}}{2} \begin{pmatrix} a & b \\ b & a \end{pmatrix}. \tag{2}$$

We introduce the orthonormal vectors

$$g_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ and } g_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{3}$$

together with the scalars

$$\rho = \mu_1 = \frac{a+b}{2} \Phi^{(2)}, \text{ and } \mu_2 = \frac{a-b}{2} \Phi^{(2)}. \tag{4}$$

Then, g_k $(k = 1, 2)$ are the left-eigenvectors of M associated to eigenvalues μ_k:

$$g_k^* M = \mu_k g_k^*, \quad k = 1, 2. \tag{5}$$

Note that ρ and μ_2 are also asymptotically eigenvalues of the expected adjacency matrix conditioned on the weights.

Indeed, if A denotes the adjacency matrix, and if ψ_1 and ψ_2 are the vectors defined for $u \in V$ by $\psi_1(u) = \frac{1}{\sqrt{2}} \phi_u$ and $\psi_2(u) = \frac{1}{\sqrt{2}} \sigma_u \phi_u$, then

$$\mathbb{E}\left[A | \phi_1, \ldots, \phi_n\right] = \frac{a+b}{n} \psi_1 \psi_1^* + \frac{a-b}{n} \psi_2 \psi_2^* - a \frac{1}{n} \mathrm{diag}\{\phi_u^2\}.$$

Put $\widehat{\psi}_i = \frac{\psi_i}{\|\psi_i\|_2}$. Then, by the law of large numbers, for $i = 1, 2$,

$$\left\| \mathbb{E}\left[A | \phi_i, \ldots, \phi_n\right] \widehat{\psi}_i - \mu_i \widehat{\psi}_i \right\|_2 \to 0,$$

in probability, as n tends to ∞.

Finally, we define for $k \in \{1, 2\}$,

$$\chi_k(e) = g_k(\sigma(e_2)) \phi_{e_2}, \qquad \text{for } e \in \vec{E}. \tag{6}$$

We show that the *candidate eigenvectors*

$$\zeta_k = \frac{B^\ell B^{*\ell} \check{\chi}_k}{\|B^\ell B^{*\ell} \check{\chi}_k\|} \tag{7}$$

are then, for $\ell \sim \log(n)$, asymptotically aligned with the first two eigenvectors of B. Note the weight in (6), which is *not* present in the ordinary SBM.

▶ **Theorem 1** (Degree-Corrected Extension of Theorem 4 in [2]). *Let G be drawn according to the DC-SBM such that assumption* (1) *holds. Assume that $\ell = C_{min}\log(n)$, with $C_{min} > 0$ a small constant defined in* (9).

If $\mu_2^2 > \rho$, then, with high probability, the eigenvalues λ_i of B satisfy

$$|\lambda_1 - \rho| = o(1), |\lambda_2 - \mu_2| = o(1), \quad \text{and, for } i \geq 3, \quad |\lambda_i| \leq \sqrt{\rho} + o(1).$$

Further, if, for $k \in \{1, 2\}$, ξ_k is a normalized eigenvector associated to λ_k, then ξ_k is asymptotically aligned with ζ_k. The vectors ξ_1 and ξ_2 are asymptotically orthogonal.

If $\rho > 1$, and $\mu_2^2 \leq \rho$, then, with high probability, the eigenvalues λ_i of B satisfy

$$|\lambda_1 - \rho| = o(1), \quad \text{and, for } i \geq 2, \quad |\lambda_i| \leq \sqrt{\rho} + o(1).$$

Further, ξ_1 is asymptotically aligned with ζ_1.

Note that $\mu_2^2 > \rho$ implies $\rho > 1$, so that we consider the DC-SBM precisely in the regime where a giant component emerges, see [1].

In Theorem 2 we show that positively correlated clustering is possible based on the second eigenvector of B when above the feasibility threshold. More precisely, let $\widehat{\sigma} = \{\widehat{\sigma}(v)\}_{v\in V}$ be estimators for the spins of the vertices. Following [3], we say that $\widehat{\sigma}$ has positive overlap with the true spin configuration $\sigma = \{\sigma(v)\}_{v\in V}$ if for some $\delta > 0$, with high probability,

$$\min_p \frac{1}{n} \sum_{v=1}^{n} 1_{\widehat{\sigma}(v) = p\circ\sigma(v)} > \frac{1}{2} + \delta,$$

where p runs over the identity mapping on $\{+, -\}$ and the permutation that swaps $+$ and $-$.

▶ **Theorem 2** (Degree-Corrected Extension of Theorem 5 in [2]). *Let G be drawn according to the DC-SBM such that assumption* (1) *holds and such that $\mu_2^2 > \rho$. Let ξ_2 be the second normalized eigenvector of B.*

Then, there exists a deterministic threshold $\tau \in \mathbb{R}$, such that the following procedure yields asymptotically positive overlap: Put for vertex $v \in V$ its estimator $\widehat{\sigma}(v) = +$ if $\sum_{e:e_2=v} \xi_2(e) > \frac{\tau}{\sqrt{n}}$ and put $\widehat{\sigma}(v) = -$ otherwise.

2.1 Notation

We say that a sequence $(E_n)_n$ of events happens with high probability (w.h.p.) if $\lim_{n\to\infty} \mathbb{P}(E_n) = 1$.

We denote by $\|\cdot\|$ both the euclidean norm for vectors and the operator norm of matrices. I.e., for vectors $x = (x_1, \dots, x_m)$, and a matrix A, $\|x\| = \sqrt{\sum_{u=1}^{m} x_u^2}$, and $\|A\| = \sup_{x,\|x\|=1} \|Ax\|$.

Below we use that the neighbourhoods with a radius no larger than $C_{\text{coupling}} \log_\rho(n)$ can be coupled w.h.p. to certain branching processes, where

$$C_{\text{coupling}} := \frac{\left(\frac{1}{3} - \frac{1}{9}\log(4/e)\right) \wedge \left(\frac{1}{80} \wedge \frac{\gamma}{4}\right)}{\log_\rho(2(a+b)\phi_{\max}^2)}. \tag{8}$$

We put,

$$C_{\min} = \frac{1}{10} C_{\text{coupling}} \tag{9}$$

and consider often neighbourhoods of radius $C_{\min}\log_\rho(n)$.

We denote the k-th moment of the weight distribution ν by $\Phi^{(k)}$. I.e., $\mathbb{E}\left[\phi_1^k\right] = \Phi^{(k)}$.

The non-backtracking property for oriented edges $e, f \in \vec{E}$ is denoted by $e \to f$, i.e., $e_2 = f_1$ and $f_2 \neq e_1$.

In proofs, we often use the symbols $c_1, c_2, \dots$ for suitably chosen constants.

3 Preliminaries

3.1 Background on non-backtracking matrix

We repeat here the most important observations made in [2].

Firstly, for any $k \geq 1$, B^k_{ef} counts the number of non-backtracking paths between oriented edges e and f. A non-backtracking path is defined as an oriented path between two oriented edges such that no edge is the inverse of its preceding edge, i.e., the path makes no backtrack.

Another import observation is that $(B^*)_{ef} = B_{fe} = B_{e^{-1}f^{-1}}$, where for oriented edge $e = (e_1, e_2)$, we set $e^{-1} = (e_2, e_1)$. If we introduce the swap notation, for $x \in \mathbb{R}^{\vec{E}}$,

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

then for any $x, y \in \mathbb{R}^{\vec{E}}$, and integer $k \geq 0$,

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

Denote by P the matrix on $\mathbb{R}^{\vec{E}\times\vec{E}}$, defined on oriented edges e, f as

$$P_{ef} = 1_{f=e^{-1}}.$$

Then, $Px = \check{x}$, $P^* = P$ and $P^{-1} = P$. Further,

$$(B^k P)^* = P(B^*)^k = B^k P,$$

so that we can write the symmetric matrix $B^k P$ in diagonal form: Let $(\sigma_{k,j})_j$ be eigenvalues of $B^k P$ ordered in decreasing order of absolute value, and let $(x_{k,j})_j$ be the corresponding orthonormal eigenvectors. Then,

$$B^k = (B^k P)P = \sum_j \sigma_{k,j} x_{k,j} x^*_{k,j} P = \sum_j \sigma_{k,j} x_{k,j} \check{x}^*_{k,j} = \sum_j s_{k,j} x_{k,j} y^*_{k,j}, \tag{10}$$

where $s_{k,j} = |\sigma_{k,j}|$ and $y_{k,j} = \text{sign}(\sigma_{k,j})\check{x}_{k,j}$. Since P is an orthogonal matrix, $(\check{x}_{k,j})_j$ form an orthonormal base for $\mathbb{R}^{\vec{E}}$ and the term furthest on the right of (10) is thus the spectral value decomposition of B^k. Now, if B is irreducible and if ξ denotes the normalized Perron eigenvector of B with eigenvalue $\lambda_1(B) > 0$, we have $\lambda_1(B) = \lim_{k\to\infty}(\sigma_{k,1})^{1/k}$, and $\lim_{k\to\infty} \|x_{k,1} - \xi\| = 0$.

In [2], the Bauer-Fike Theorem is extended to prove the spectral claims we make here.

3.2 Extension of Bauer-Fike Theorem

Tailored to our needs, we use the following proposition from [2]:

▶ **Proposition 3** (Special case of Proposition 8 in [2])**.** *Let $\ell = C \log_\rho n$, with $C > 0$. Let $A \in M_n(\mathbb{R})$, such that for some vectors $x_1 = x_{\ell,1}, y_1 = y_{\ell,1}, x_2 = x_{\ell,2}, y_2 = y_{\ell,2} \in \mathbb{R}$, some matrix $R_\ell \in M_n(\mathbb{R})$, and some non-zero constants $\rho > \mu_2$ with $\mu_2^2 > \rho$,*

$$A^\ell = \rho^\ell x_1 y_1^* + \mu_2^\ell x_2 y_2^* + R_\ell. \tag{11}$$

Assume there exist $c_0, c_1 > 0$ such that for all $i \in \{1, 2\}$, $\langle y_i, x_i\rangle \geq c_0$, $\|x_i\|\|y_i\| \leq c_1$. Assume further that $\langle x_1, y_2\rangle = \langle x_2, y_1\rangle = \langle x_1, x_2\rangle = \langle y_1, y_2\rangle = 0$ and for some $c > 0$

$$\|R_\ell\| < \rho^{\ell/2} \log^c(n).$$

Let $(\lambda_i)_{1\le i\le n}$, be the eigenvalues of A with $|\lambda_n| \le \ldots \le |\lambda_1|$. Then,

$$|\lambda_1 - \rho| = o(1), |\lambda_2 - \mu_2| = o(1), \quad \text{and, for } i \ge 3, \quad |\lambda_i| \le \sqrt{\rho} + o(1).$$

Further, there exist unit eigenvectors ψ_1, ψ_2 of A with eigenvalues λ_1, respectively λ_2 such that

$$||\psi_i - \frac{x_i}{\|x_i\|}|| = o(1).$$

Proof. This is a special case of Proposition 8 in [2]. In the notation of the latter, we have $\ell' = \ell - 2$, $\theta_1 = \rho$, $\theta_2 = \mu_2$, $\theta = \mu_2$, $\gamma \ge \frac{a+b}{|a-b|} > 1$. Further $\frac{c_0(c_0\gamma^k - c_1)_+}{4c_1} \wedge \frac{c_0^2}{2(\ell\vee\ell')c_1} \sim \frac{1}{\log_\rho n}$, and thus

$$\|R_\ell\| \le \log^c(n)\left(\frac{\sqrt{\rho}}{|\mu_2|}\right)^\ell |\mu_2|^\ell = o(1)\frac{1}{\log_\rho n}|\theta|^\ell. \qquad ◀$$

To prove the case $\mu_2^2 > \rho$ of Theorem 1, we thus need to find candidate vectors x_1, x_2, y_1 and y_2 that meet the conditions in Proposition 3 and further verify that the remainder R_ℓ has small norm. Note that the last condition is true whenever $\|B^\ell x\| \le \rho^{\ell/2}\log^c(n)$ for all normalized x in $\mathrm{span}\{y_1, y_2\}^\perp$.

To address the case $\mu_2^2 \le \rho$ of Theorem 1, we appeal to Proposition 7 in [2], which is very similar in spirit to Proposition 3.

4 Proof of Theorem 1

4.1 The case $\mu_2^2 > \rho$

We start with the case $\mu_2^2 > \rho$. We decompose, for some vectors x_1, y_1, x_2 and y_2 and matrix R_ℓ,

$$B^\ell = \rho^\ell x_1 y_1^* + \mu_2^\ell x_2 y_2^* + R_\ell,$$

and we show that the assumptions of Proposition 3 are met.

Let ℓ be as in Theorem 1 and recall χ_k and ζ_k from (6) and (7). For ease of notation, we introduce for $k \in \{1,2\}$,

$$\varphi_k = \frac{B^\ell \chi_k}{\|B^\ell \chi_k\|}, \quad \text{and } \theta_k = \|B^\ell \check{\varphi}_k\|. \tag{12}$$

Then, $\zeta_k = \frac{B^\ell \check{\varphi}_k}{\theta_k}$.

To prove the main theorem, we need the following two propositions. The proofs are deferred to Section 8 and 9.1. The material in Section 8 builds on ingredients from Sections 6 - 7, where we assume that $\mu_2^2 > \rho$, unless stated otherwise.

▶ **Proposition 4** (Degree-Corrected Extension of Proposition 19 in [2]). *Assume that $\mu_2^2 > \rho$. Let $\ell = C\log_\rho n$ with $0 < C < C_{min}$. For some $b, c > 0$, with high probability,*

(i) $b|\mu_k^\ell| \le \theta_k \le c|\mu_k^\ell|$ *if* $k \in \{1,2\}$,
(ii) $sign(\mu_k^\ell)\langle\zeta_k, \check{\varphi}_k\rangle \ge b$ *if* $k \in \{1,2\}$,
(iii) $|\langle\varphi_1, \varphi_2\rangle| \le (\log n)^3 n^{C-(\frac{\gamma}{2}\wedge\frac{1}{40})}$,
(iv) $|\langle\zeta_j, \check{\varphi}_k\rangle| \le (\log n)^3 n^{\frac{3}{2}C-(\frac{\gamma}{2}\wedge\frac{1}{40})}$ *if* $k \ne j \in \{1,2\}$.
(v) $|\langle\zeta_1, \zeta_2\rangle| \le (\log n)^8 n^{2C-(\frac{\gamma}{2}\wedge\frac{1}{40})}$.

Put $H = \text{span}\{\check{\varphi}_1, \check{\varphi}_2\}$, then

▶ **Proposition 5** (Degree-Corrected Extension of Proposition 20 in [2]). *Let $\ell = C \log_\rho n$ with $0 < C < C_{min}$. For some $c > 0$, with high probability,*

$$\sup_{x \in H^\perp, \|x\|=1} \|B^\ell x\| \leq (\log n)^c \rho^{\ell/2}. \tag{13}$$

Put $\bar{\varphi}_1 = \check{\varphi}_1$, and $\bar{\varphi}_2 = \frac{\check{\varphi}_2 - \langle \check{\varphi}_1, \check{\varphi}_2 \rangle \check{\varphi}_1}{||\check{\varphi}_2 - \langle \check{\varphi}_1, \check{\varphi}_2 \rangle \check{\varphi}_1||}$, then $\bar{\varphi}_1$ and $\bar{\varphi}_2$ are orthonormal and $||\bar{\varphi}_2 - \check{\varphi}_2|| = o(\rho^{-\ell/2})$, due to Proposition 4 (iii).

Let $\bar{\zeta}_1$ be the normalized orthogonal projection of ζ_1 on $\text{span}\{\bar{\varphi}_2\}^\perp$. Similarly, let $\bar{\zeta}_2$ be the normalized orthogonal projection of ζ_2 on $\text{span}\{\bar{\zeta}_1, \bar{\varphi}_1\}^\perp$.

Then $\langle \bar{\zeta}_1, \bar{\zeta}_2 \rangle = 0$ and for $i = 1, 2$, $||\bar{\zeta}_i - \zeta_i|| = o(\rho^{-\ell/2})$, as follows from Proposition 4 (iv) and (v).

We set

$$D = \theta_1 \bar{\zeta}_1 \bar{\varphi}_1^* + \theta_2 \bar{\zeta}_2 \bar{\varphi}_2^* = \rho^\ell \left(\frac{\theta_1}{\rho^\ell} \bar{\zeta}_1 \right) \bar{\varphi}_1^* + \mu_2^\ell \left(\frac{\theta_2}{\mu_2^\ell} \bar{\zeta}_2 \right) \bar{\varphi}_2^*.$$

Note that,

$$\|B^\ell \bar{\varphi}_1\| = \theta_1 = O(\rho^\ell),$$

and

$$\|B^\ell \bar{\varphi}_2\| = \|B^\ell \left((1 + o(1)) \check{\varphi}_2 + o(1) \bar{\varphi}_1 \right) \| = O(\rho^\ell).$$

As a consequence, from Proposition 5,

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

Since $D\bar{\varphi}_i = B^\ell \check{\varphi}_i + \theta_i(\bar{\zeta}_i - \zeta_i)$,

$$\|B^\ell \bar{\varphi}_i - D\bar{\varphi}_i\| \leq \|B^\ell\| \|\bar{\varphi}_i - \check{\varphi}_i\| + \theta_i \|\bar{\zeta}_i - \zeta_i\| = \mathcal{O}\left(\rho^{\ell/2}\right).$$

Let P be the orthogonal projection on $H = \text{span}\{\bar{\varphi}_1, \bar{\varphi}_2\} = \text{span}\{\check{\varphi}_1, \check{\varphi}_2\}$, then $\|B^\ell P - D\| = \mathcal{O}\left(\rho^{\ell/2}\right)$.

Put $R_\ell = B^\ell - D$. Write for $y \in \mathbb{R}^{\vec{E}}$ with unit norm, $y = h + h^\perp$, with $h \in H$ and $h^\perp \in H^\perp$, then

$$\begin{aligned} \|R_\ell y\| &= \|B^\ell h^\perp + (B^\ell - D)h\| \\ &\leq \sup_{x \in H^\perp, \|x\|=1} \|B^\ell x\| + \|B^\ell P - D\| \\ &= \mathcal{O}\left(\log^c(n) \rho^{\ell/2}\right), \end{aligned} \tag{14}$$

as follows from Proposition 5.

We finish by applying Proposition 3 with $x_1 = \frac{\theta_1}{\rho^\ell} \bar{\zeta}_1, y_1 = \bar{\varphi}_1, x_2 = \frac{\theta_2}{\mu_2^\ell} \bar{\zeta}_2$, and, $y_2 = \bar{\varphi}_2$.

4.2 The case $\mu_2^2 \leq \rho$

In case $\mu_2^2 \leq \rho$, Proposition 4 (i) and (ii) continue to hold for $k = 1$. Further, Proposition 4 (iii) as well as Proposition 5 continue to hold. We need however the following bound for $k = 2$:

▶ **Proposition 6.** *Assume that* $\mu_2^2 \le \rho$*. Let* $\ell = C \log_\rho n$ *with* $0 < C < C_{min}$*. For some* $c > 0$*, with high probability,*

$$\theta_2 \le (\log n)^c \rho^{\ell/2}.$$

Using this proposition and $||\bar{\varphi}_2 - \check{\varphi}_2|| = o(\rho^{-\ell/2})$, we get

$$\|B^\ell \bar{\varphi}_2\| \le (\log n)^{c+1} \rho^{\ell/2}.$$

It remains to apply Proposition 7 from [2].

5 Poisson-mixture two-type branching processes

The proofs of the statements in this section can be found in Appendix A in the detailed version of the underlying article (Arxiv:1609.02487).

5.1 A theorem of Kesten and Stigum

We consider the following branching process starting with a single particle, the root o, having spin $\sigma_o \in \{+, -\}$ and weight $\phi_o \in [\phi_{\min}, \phi_{\max}]$ (which we often take random). The root is replaced in generation 1 by Poi $\left(\frac{a}{2}\Phi^{(1)}\phi_o\right)$ particles of spin σ_o and Poi $\left(\frac{b}{2}\Phi^{(1)}\phi_o\right)$ particles of spin $-\sigma_o$. Further, the weights of those particles are i.i.d. distributed following law ν^*, the size-biased version of ν, defined for $x \in [\phi_{\min}, \phi_{\max}]$ by

$$\nu^*([0, x]) = \frac{1}{\Phi^{(1)}} \int_{\phi_{\min}}^{x} y \mathrm{d}\nu(y). \tag{15}$$

For generation $t \ge 1$, a particle with spin σ and weight ϕ^* is replaced in the next generation by Poi $\left(\frac{a}{2}\Phi^{(1)}\phi^*\right)$ particles with the same spin and Poi $\left(\frac{b}{2}\Phi^{(1)}\phi^*\right)$ particles of the opposite sign. Again, the weights of the particles in generation $t+1$ follow in an i.i.d. fashion the law ν^*. The offspring-size of an individual is thus a **Poisson-mixture**.

We use the notation $Z_t = \begin{pmatrix} Z_t(+) \\ Z_t(-) \end{pmatrix}$ for the population at generation $t \ge 1$, where $Z_t(\pm)$ is the number of type $\pm$ particles in generation t. We let $(\mathcal{F}_t)_{t\ge1}$ denote the natural filtration associated to $(Z_t)_{t\ge1}$.

We associate two matrices to the branching process, namely M defined in (2), and, for a root with weight ϕ_o,

$$M_{\phi_o} = \frac{\Phi^{(1)}\phi_o}{\Phi^{(2)}} M. \tag{16}$$

Then, M is the *transition* matrix for generations $t \ge 1$ and later:

$$\mathbb{E}\left[Z_{t+1}|Z_t\right] = M Z_t, \quad \text{for all } t \ge 1, \tag{17}$$

and M_{ϕ_o} describes the transition from the root to the first generation:

$$\mathbb{E}\left[Z_1|Z_0, \phi_o\right] = M_{\phi_o} Z_0, \tag{18}$$

where, by assumption $Z_0 = \begin{pmatrix} 1_{\sigma_o=+} \\ 1_{\sigma_o=-} \end{pmatrix}$. Note that the difference between the root and later generations stems from the fact that the root's weight is deterministic in the conditional expectation, whereas the weight of a particle in any later generation has expectation $\frac{\Phi^{(2)}}{\Phi^{(1)}}$.

Recall from (5) that g_k ($k = 1, 2$) are the left-eigenvectors of M associated to eigenvalues μ_k:

$$g_k^* M = \mu_k g_k^*, \quad k = 1, 2. \tag{19}$$

Note that M_{ϕ_o} has the same left-eigenvectors as M, while the corresponding eigenvalues are given by

$$\mu_{k,\phi_o} = \frac{\Phi^{(1)}\phi_o}{\Phi^{(2)}}\mu_k, \quad k = 1, 2. \tag{20}$$

Theorem 7 shows that a Kesten-Stigum theorem applies to the "classical" branching process obtained after restricting the above process to generations 1 and later. Corollary 8, then, joins this classical branching process to the transition from the root to generation 1.

We further consider the vector $\Psi_t = (\Psi_t(+), \Psi_t(-))$, containing sums of the weights,

$$\Psi_t(\pm) = \sum_{u \in Y_t} 1_{\sigma_u = \pm}\phi_u, \tag{21}$$

where Y_t is the set of particles at distance t from the root, and where ϕ_u and σ_u denote the weight respectively spin of a particle u. Note that $\Psi_t = Z_t$ in case of unit weights.

The martingale Theorem 9 is not present in [2]. We need it to bound the variance of the cross-generational functional defined in Section 5.3.

▶ **Theorem 7** (Degree-Corrected Extension of Theorem 21 in [2])**.** *Assume that $\mu_2^2 > \rho$. Put $\mathcal{F}_t = \{Z_s\}_{s \le t}$. For any $k = 1, 2$,*

$$\left(X_k(t) := \frac{\langle g_k, Z_t \rangle}{\mu_k^{t-1}} - \langle g_k, Z_1 \rangle \right)_{t \ge 1},$$

is an $\mathcal{F}_t$-martingale converging a.s. and in L^2 such that for some $C > 0$ and all $t \ge 1$, $\mathbb{E}\left[X_k(t)\right] = 0$ and $\mathbb{E}\left[X_k^2(t)|Z_1\right] \le C\|Z_1\|_1$.

▶ **Corollary 8.** *Assume that $\mu_2^2 > \rho$. For $k = 1, 2$, with the weight $\phi_o = \psi_o$ of the root fixed, the sequence of random variables $(Y_{k,\psi_o}(t))_{t \ge 1} = \left(\frac{\langle g_k, Z_t \rangle}{\mu_k^{t-1}\mu_{k,\psi_o}} \right)_{t \ge 1}$ converges* almost surely *and in L^2 to a random variable $Y_{k,\psi_o}(\infty)$ with $\mathbb{E}\left[Y_{k,\psi_o}(\infty)|\sigma_o\right] = g_k(\sigma_o)$. Further, the L^2-convergence takes place uniformly over all ψ_o.*

▶ **Theorem 9.** *Assume that $\mu_2^2 > \rho$. Put $\mathcal{G}_t = \{\Psi_s\}_{s \le t}$. For any $k = 1, 2$,*

$$\left(X_k(t) := \frac{\langle g_k, \Psi_t \rangle}{\mu_k^{t-1}} - \langle g_k, \Psi_1 \rangle \right)_{t \ge 1},$$

is an $\mathcal{G}_t$-martingale converging a.s. and in L^2 such that for some $C > 0$ and all $t \ge 1$, $\mathbb{E}\left[X_k(t)\right] = 0$ and $\mathbb{E}\left[X_k^2(t)|Z_1\right] \le C\|Z_1\|_1$.

5.2 Quantitative version of the Kesten-Stigum theorem

We now quantify the growth of the population size. The latter is defined as

$$S_t = \|Z_t\|_1, \quad t \ge 0,$$

i.e., the number of individuals in generation $t \ge 0$. Given S_t, for $t \ge 1$ we have

$$S_{t+1} = \mathrm{Poi}\left(\sum_{l=1}^{S_t} X_t^{(l)} \right), \tag{22}$$

where $\left(X_t^{(l)}\right)_l$ are i.i.d. copies of $\frac{a+b}{2}\Phi^{(1)}\phi^*$, where ϕ^* follows law ν^*.

Note that in the *ordinary* Stochastic Block Model (i.e., when all vertices have unit weight), the argument of the Poisson random variables in (22) is deterministic, contrary to the general case under consideration here. Using (17) recursively in conjunction with (18), it follows that

$$\mathbb{E}\left[S_t|\phi_o\right] = \frac{\Phi^{(1)}\phi_o}{\Phi^{(2)}}\rho^t, \quad \forall t \geq 1.$$

In the following lemma we show that deviations from this average are small. In fact, there exists a constant C such that for each $t \geq 0$, S_t is asymptotically stochastically dominated by an Exponential random variable with mean $C\rho^t$. An important ingredient in the proof below is Hoeffding's inequality, which we use to derive a concentration result for the parameter of the Poisson variable in (22).

▶ **Lemma 10** (Degree-Corrected Extension of Lemma 23 in [2]). *Assume $S_0 = 1$. There exist $c, c' > 0$ such that for all $s \geq 0$,*

$$\mathbb{P}\left(\forall k \geq 1, S_k \leq s\rho^k\right) \geq 1 - c'e^{-cs}.$$

From Theorem 7 and Corollary 8, we know that the different components (expressed in the basis of eigenvectors of M) grow exponentially with rate ρ, respectively μ_2. We now quantify the error. Recall Ψ_t from (21).

5.2.1 The case $\mu_2^2 > \rho$

▶ **Theorem 11** (Degree-Corrected Extension of Theorem 24 in [2]). *Assume that $\mu_2^2 > \rho$. Let $\beta > 0$, $Z_0 = \delta_x$ and $\phi_o = \psi_o$ be fixed. There exists $C = C(x, \beta) > 0$ such that with probability at least $1 - n^{-\beta}$, for all $k \in \{1, 2\}$, all $0 \leq s < t \leq C_{min}\log(n)$, with $0 \leq s < t$,*

$$|\langle g_k, Z_s\rangle - \mu_k^{s-t}\langle g_k, Z_t\rangle| \leq C(s+1)\rho^{s/2}(\log n)^{3/2},$$

and,

$$|\langle g_k, \Psi_s\rangle - \mu_k^{s-t}\langle g_k, \Psi_t\rangle| \leq C\rho^{s/2}(\log n)^{5/2}.$$

5.2.2 The case $\mu_2^2 \leq \rho$

▶ **Theorem 12.** *Assume that $\mu_2^2 \leq \rho$. Let $\beta > 0$, $Z_0 = \delta_x$ and $\phi_o = \psi_o$ be fixed. There exists $C = C(x, \beta) > 0$ such that with probability at least $1 - n^{-\beta}$, for all $t \geq 1$,*

$$|\langle g_2, \Psi_t\rangle| \leq Ct^2\rho^{t/2}(\log n)^2,$$

and,

$$\mathbb{E}\left[|\langle g_2, \Psi_t\rangle|^2\right] \leq Ct^3\rho^t.$$

5.3 $B^\ell B^{*\ell}\check{\chi}_k$ on trees: a cross generation functional

Recall our claim that $B^\ell B^{*\ell}\check{\chi}_k$ are asymptotically aligned with the eigenvectors of B. In the DC-SBM, the local-neighbourhood of a vertex has with high probability a tree-like structure described by the branching process above. In this section we analyse $B^\ell B^{*\ell}\check{\chi}_k$ on trees.

To this end we define a cross-generational functional slightly different from its analogue in [2] due to the presence of weights:

$$Q_{k,\ell} = \sum_{(u_0,\dots,u_{2\ell+1})\in\mathcal{P}_{2\ell+1}} g_k(\sigma(u_{2\ell+1}))\phi_{u_{2\ell+1}}, \tag{23}$$

where $\mathcal{P}_{2\ell+1}$ is the set of paths $(u_0,\dots,u_{2\ell+1})$ (of length $2\ell+1$) in the tree starting from $u_0=o$ with both $(u_0,\dots,u_\ell)$ and $(u_\ell,\dots,u_{2\ell+1})$ non-backtracking and $u_{\ell-1}=u_{\ell+1}$. Note that these paths thus make a back-track exactly once at step $\ell+1$.

Explicitly, we have

$$Q_{1,\ell} = \sum_{(u_0,\dots,u_{2\ell+1})\in\mathcal{P}_{2\ell+1}} \frac{1}{\sqrt{2}}\phi_{u_{2\ell+1}}, \tag{24}$$

and,

$$Q_{2,\ell} = \sum_{(u_0,\dots,u_{2\ell+1})\in\mathcal{P}_{2\ell+1}} \frac{1}{\sqrt{2}}\sigma(u_{2\ell+1})\phi_{u_{2\ell+1}}. \tag{25}$$

Consider a tree $\mathcal{T}'$ and a leaf e_1 on it that has unique neighbour, say, o. Then, if e is the oriented edges from e_1 to o and if $B_{\mathcal{T}'}$ denotes the non-backtracking matrix defined on $\mathcal{T}'$,

$$\left(B^{\ell}_{\mathcal{T}'}B^{*\ell}_{\mathcal{T}'}\check{\chi}_k\right)(e) = Q_{k,\ell} + g_k(\sigma(e_1))\phi_{e_1}\|Z_\ell\|_1, \tag{26}$$

where $Q_{k,\ell}$ and Z_ℓ are defined on the tree $\mathcal{T}$ with root o obtained after removing vertex e_1 from $\mathcal{T}'$.

In the sequel we analyse $Q_{k,\ell}$ on the branching process defined above, starting with a single particle, the root o. Let V indicate the particles of the random tree. Denote the spin of a particle $v\in V$ by $\sigma_v\in\{+,-\}$ and its weight by $\phi_v\in S$.

For $t\geq 0$, let Y_t^v denote the set of particles, *including* their spins and weights, of generation t from v in the subtree of particles with common ancestor $v\in V$. Let $Z_t^v=(Z_t^{v,+},Z_t^{v,-})$ denote the number of $\pm$ vertices in generation t; i.e., $Z_t^{v,\pm}=\sum_{u\in Y_t^v}1_{\sigma(u)=\pm}$. Finally, let $\Psi_t^v=(\Psi_t^{v,+},\Psi_t^{v,-})$, with $\Psi_t^{v,\pm}=\sum_{u\in Y_t^v}1_{\sigma(u)=\pm}\phi_u$.

We rewrite $Q_{k,\ell}$ into a more manageable form: First observe that every path in $\mathcal{P}_{2\ell+1}$, after reaching $u_{\ell+1}$, climbs back to a depth t from which it then again moves down the tree (that is, in the direction away from the root). Let us call the vertex at level t (to which the path climbs back before descending again) u. Then, (if $t\neq 0$) there are two children of u, say v and w such that w lies on the path between u and $u_{\ell+1}$ and v is in between u and $u_{2\ell+1}$. For such fixed v and w in Y_1^u, only the children $u_{2\ell+1}\in Y_t^v$ determine the contribution of a path to (23), regardless of the choice of $u_{\ell+1}\in Y^w_{\ell-t-1}$. Hence, for such fixed u and $v,w\in Y_1^u$ and $u_{2\ell+1}$, there are $|Y^w_{\ell-t-1}|=S^w_{\ell-t-1}$ paths giving the same contribution to (23):

$$Q_{k,\ell} = \sum_{t=0}^{\ell-1}\sum_{u\in Y_t^o} L^u_{k,\ell}, \tag{27}$$

where, for $|u|=t\geq 0$,

$$L^u_{k,\ell} = \sum_{w\in Y_1^u} S^w_{\ell-t-1}\left(\sum_{v\in Y_1^u\setminus\{w\}}\langle g_k,\Psi_t^v\rangle\right). \tag{28}$$

The following theorem is an extension of Theorem 25 in [2]. The important observation is that, again, for $Z_0=\delta_\tau$ fixed, $\left(Q_{2,\ell}/\mu_2^{2\ell}\right)_\ell$ converges to a random variable with mean a constant times τ, that is, the spin of the root. Its proof uses both martingale theorems stated above. We use the second martingale statement, which is not present in the ordinary SBM, to bound the variance of $Q_{k,\ell}$:

▶ **Theorem 13** (Degree-Corrected Extension of Theorem 25 in [2]). *Assume that $\mu_2^2 > \rho$. Let $Z_0 = \delta_x$ and $\phi_o = \psi_o$ be fixed. For $k \in \{1,2\}$, $\left(Q_{k,\ell}/\mu_k^{2\ell}\right)_\ell$ converges in L^2 as ℓ tends to infinity to a random variable with mean $\frac{\Phi^{(3)}}{\Phi^{(2)}}\frac{\rho}{\mu_k^2-\rho}\mu_{k,\psi_o}g_k(x)$. Further, the L^2-convergence takes place uniformly for all ψ_o.*

5.3.1 The case $\mu_2^2 \le \rho$

▶ **Theorem 14.** *Assume that $\mu_2^2 \le \rho$. Let $Z_0 = \delta_x$ and $\phi_o = \psi_o$ be fixed. There exists a constant C such that $\mathbb{E}\left[Q_{2,\ell}^2\right] \le C\rho^{2\ell}\ell^5$.*

5.4 Orthogonality: Decorrelation in branching process

Again, as in [2], $Q_{1,\ell}$ and $Q_{2,\ell}$ are uncorrelated when defined on the branching process above. The proof presented here is simpler than the corresponding one in [2] and uses that for the two communities-case, $Q_{1,\ell}$ and $Q_{2,\ell}$ are *explicitly* known.

The orthogonality of the candidate eigenvectors (i.e., $(iii)-(v)$ in Proposition 4) follows from this fact, see Proposition 24 $(ii), (iii)$ and Proposition 25 (ii) below.

▶ **Theorem 15** (Degree-Corrected Extension of 28 in [2]). *Assume that the spin σ_o of the root is drawn uniformly from $\{+,-\}$. Then for any $\ell \ge 0$,*

$$\mathbb{E}\left[Q_{1,\ell}Q_{2,\ell}|\mathcal{T}\right] = 0.$$

6 Coupling of local neighbourhood

The proofs of the statements in this section can be found in Appendix B in the detailed version of the underlying article (Arxiv:1609.02487).

6.1 Coupling

Here we establish the connection between neighbourhoods in the DC-SBM and the branching process in Section 5. We established this coupling in an earlier paper [6] using an exploration process that we repeat below. Compared to the ordinary SBM, vertices are now weighted, so that two facts need to be verified: At each step of the exploration process, unexplored vertices have a weight drawn from a distribution close in total variation distance to ν. Detected vertices on their turn follow a law close to ν^*.

We distinguish between two different concepts of neighbourhood: the classical neighbourhood that is rooted at a vertex and another neighbourhood that starts with an edge. For the latter, we need the following concept of *oriented* distance $\vec{d}$, which for $e, f \in \vec{E}(V)$ is defined as

$$\vec{d}(e,f) = \min_{\gamma} \ell(\gamma)$$

where the minimum is taken over all self-avoiding paths $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_{\ell+1})$ in G such that $(\gamma_0,\gamma_1) = e$, $(\gamma_\ell, \gamma_{\ell+1}) = f$ and for all $1 \le k \le \ell+1$, $\{\gamma_k, \gamma_{k+1}\} \in E$. and where for such a path γ, $\ell(\gamma) = \ell$. Note that $\vec{d}(e,f) = \vec{d}(f^{-1}, e^{-1})$, i.e., $\vec{d}$ is not symmetric.

We introduce the vector $Y_t(e) = (Y_t(e)(i))_{i\in\{+,-\}}$ where, for $i \in \{+,-\}$,

$$Y_t(e)(i) = \left|\left\{f \in \vec{E} : \vec{d}(e,f) = t, \sigma(f_2) = i\right\}\right|, \tag{29}$$

we denote the number of vertices at oriented distance t from e by

$$S_t(e) = \|Y_t(e)\|_1 = \left|\left\{f \in \vec{E} : \vec{d}(e,f) = t\right\}\right|,$$

and we define vector $\Psi_t(e) = (\Psi_t(e)(i))_{i\in\{+,-\}}$ where, for $i \in \{+,-\}$,

$$\Psi_t(e)(i) = \sum_{f\in\vec{E}:\vec{d}(e,f)=t} 1_{\sigma(f_2)=i}\phi_{f_2}. \tag{30}$$

We denote the classical neighbourhood of radius r rooted at vertex v by $(G,v)_r$ and the neighbourhood around oriented edge $e = (e_1, e_2)$ by $(G,e)_r$. With the definitions above, we then have, $(G,e)_r = (G', e_2)_r$, where G' is the graph G with edge $\{e_1, e_2\}$ removed. In particular,

$$S_t(e) = S'_t(e_2),$$

where S'_t is S_t defined on G'.

The two branching processes that describe the neighbourhoods are almost identical, the only difference lies in the weight of the root: In the classical branching processes, the weight is drawn according to distribution ν. In the branching process starting at an edge oriented towards, say, o, the root o has weight governed by ν^*. See Proposition 16 below.

As a corollary we obtain an analogue of Theorem 11 for local neighbourhoods: the components of $\Psi_t(e)$ grow exponentially, see Corollary 17.

We bound the growth of S_t in Lemma 18. We use a coupling argument to show that the weights of the unexplored vertices and selected vertices are stochastically dominated by variables following law ν, respectively ν^*. This argument is not needed in the ordinary SBM.

Following [17], we need to verify that certain problematic structures, namely *tangles*, are excluded with high probability. We say that a graph H is tangle-free if all its $\ell-$ neighbourhoods contain at most one cycle. If there is at least one $\ell-$ neighbourhood in H that contains more than one cycle, we call H tangled. Note that in the sequel we shall often suppress the dependence on ℓ and simply call a graph tangle-free or tangled; the ℓ dependence is then tacitly assumed.

Following standard arguments we establish in Lemma 19 that the graph is with high probability $\log(n)$-tangle free.

We prepare by recalling the exploration process in [6] starting at a vertex:

At time $m = 0$, choose a vertex ρ in $V(G)$, where G is an instant of the DC-SBM. Initially, it is the only active vertex: $\mathcal{A}(0) = \{\rho\}$. All other vertices are neutral at start: $\mathcal{U}(0) = V(G) \setminus \{\rho\}$. No vertex has been explored yet: $\mathcal{E}(0) = \emptyset$.

At each time $m \geq 0$ we arbitrarily pick an active vertex u in $\mathcal{A}(m)$ that has shortest distance to ρ, and explore all its edges in $\{uv : v \in \mathcal{U}(m)\}$: if $uv \in E(G)$ for $v \in \mathcal{U}(m)$, then we set v active in step $m+1$, otherwise it remains neutral.

At the end of step m, we designate u to be explorated.

Thus,

$$\mathcal{E}(m+1) = \mathcal{E}(m) \cup \{u\},$$

$$\mathcal{A}(m+1) = (\mathcal{A}(m) \setminus \{u\}) \cup (\mathcal{N}(u) \cap \mathcal{U}(m)),$$

and,

$$\mathcal{U}(m+1) = \mathcal{U}(m) \setminus \mathcal{N}(u).$$

▶ **Proposition 16** (Degree-Corrected Extension of Proposition 31 in [2]). *Let $\ell = C\log_\rho(n)$, with $C < C_{coupling}$. Let $\rho \in V$ and $e = (e_1, e_2) \in \vec{E}$. Let (T, o) be the branching process with root o defined in Section 5, where the root has spin $\sigma(v)$ and weight governed by ν. Similarly, Let (T', o) be that same branching process, when the root has spin $\sigma(e_2)$ and weight governed by ν^*. Then, the total variation distance between the law of $(G, v)_\ell$ and $(T, o)_\ell$ goes to zero as $1 - n^{-\left(\frac{\gamma}{2} \wedge \frac{1}{40}\right)}$. The same is true for the difference between the law of $(G, e)_\ell$ and (T', o).*

▶ Remark. Note that with the event $(G, v)_\ell = (T, o)_\ell$, we mean that the graph and tree are equal, **including** their **spins and weights**. See [6] for more details.

▶ **Corollary 17** (Degree-Corrected Extension of Corollary 32 in [2]). *Assume $\mu_2^2 > \rho$. Let $\ell = C\log_\rho n$ with $0 < C < C_{coupling}$. For $e \in \vec{E}(V)$, we define the event $\mathcal{E}(e)$ that for all $0 \le t < \ell$ and $k \in \{1,2\}$: $|\langle g_k, \Psi_t(e)\rangle - \mu_k^{t-\ell}\langle g_k, \Psi_\ell(e)\rangle| \le (\log n)^3 \rho^{t/2}$. Then, with high probability, the number of edges $e \in \vec{E}$ such that $\mathcal{E}(e)$ does not hold is at most $\log(n)\ n^{1-\left(\frac{\gamma}{2}\wedge\frac{1}{40}\right)}$.*

▶ **Lemma 18** (Degree-Corrected Extension of Lemma 29 in [2]). *There exist $c, c' > 0$ such that for all $s \ge 0$ and for any $w \in [n] \cup \vec{E}(V)$,*

$$\mathbb{P}\left(\forall t \ge 0 : S_t(w) \le s\bar{\rho}_n^t\right) \ge 1 - ce^{-c's}.$$

Consequently, for any $p \ge 1$, there exists $c'' > 0$ such that

$$\mathbb{E}\left[\max_{v\in[n], t\ge 0}\left(\frac{S_t(v)}{\bar{\rho}_n^t}\right)^p\right] \le c''(\log n)^p.$$

▶ **Lemma 19** (Degree-Corrected Extension of Lemma 30 in [2]). *Let $\ell = C\log_\rho(n)$, with $0 < C < C_{coupling}$. Then, w.h.p., at most $\rho^{2\ell}\log(n)$ vertices have a cycle in their ℓ - neighbourhood. Further, w.h.p., the graph is ℓ - tangle-free.*

6.2 Geometric growth

Here we show that for $k \in \{1,2\}$, $\langle B^\ell \chi_k, \delta_e\rangle$ grows nearly geometrically in t with rate μ_k. Corollary 21 then establishes a bound for $r \le \ell$ on $\sup_{\langle B^\ell\chi_k, x\rangle = 0, \|x\|=1} \|\langle B^r\chi_k, x\rangle\|$ crucial for the norm bounds in Section 9.

▶ **Proposition 20** (Degree-Corrected Extension of Proposition 33 in [2]). *Assume $\mu_2^2 > \rho$. Let $\ell = C\log_\rho(n)$, with $0 < C < C_{coupling} \wedge \left(\frac{1}{2} - \left(\frac{\gamma}{4}\wedge\frac{1}{80}\right)\right) = C_{coupling}$. For $e \in \vec{E}(V)$, let $\vec{E}_\ell$ be the set of oriented edges such that either $(G, e_2)_\ell$ is not a tree or the event $\mathcal{E}(e)$ (defined in Corollary 17) does not hold. Then, w.h.p. for $k \in \{1,2\}$:*

(i) $|\vec{E}_\ell| \ll (\log n)^2 n^{1-\frac{\gamma}{2}\wedge\frac{1}{40}}$,

(ii) *for all $e \in \vec{E}\backslash\vec{E}_\ell$, $0 \le r \le \ell$,*

$$|\langle B^r\chi_k, \delta_e\rangle - \mu_k^{r-\ell}\langle B^\ell\chi_k, \delta_e\rangle| \le (\log n)^4\rho^{r/2},$$

(iii) *for all $e \in \vec{E}_\ell$, $0 \le r \le \ell$,*

$$|\langle B^r\chi_k, \delta_e\rangle| \le (\log n)^2\rho^r.$$

▶ **Corollary 21** (Degree-Corrected Extension of Corollary 34 in [2]). *Let $\ell = C\log_\rho(n)$, with $0 < C < C_{coupling} \wedge \left(1 - \frac{\gamma}{2}\wedge\frac{1}{40}\right)\wedge\left(\frac{\gamma}{4}\wedge\frac{1}{80}\right) = C_{coupling}$. W.h.p. for any $0 \le r \le \ell - 1$ and $k \in \{1,2\}$:*

$$\sup_{\langle B^\ell\chi_k, x\rangle = 0, \|x\|=1} \|\langle B^r\chi_k, x\rangle\| \le (\log n)^5 n^{1/2}\rho^{r/2}.$$

7 A weak law of large numbers for local functionals on the DC-SBM

The proofs of the statements in this section can be found in Appendix C in the detailed version of the underlying article (Arxiv:1609.02487).

Here we show that a weak law of large numbers applies for local functionals defined on *weighted coloured* random graphs generated according to the DC-SBM.

By a *weighted coloured* graph we mean a graph $G = (V, E)$ together with maps $\sigma : V \to \{+, -\}$ and $\phi : V \to [\phi_{\min}, \phi_{\max}]$. For $v \in V$, we identify $\sigma(v)$ as the spin of v and $\phi(v)$ as its weight. We denote by $\mathcal{G}^*$ the set of *rooted weighted coloured* graphs. We denote an element of $\mathcal{G}^*$ by (G, o): $G = (V, E)$ is then a weighted coloured graph and $o \in V$ is some distinguished vertex. A function $\tau : \mathcal{G}^* \to \mathbb{R}$ is said to be $\ell-$ local if $\tau(G, o)$ depends only on $(G, o)_\ell$.

To derive the claimed weak law when G is drawn according to the DC-SBM, we prepare with a variance bound for $\sum_{v=1}^n \tau(G, v)$, see Proposition 22. The bound follows from the law of total variance,

$$\begin{aligned}\operatorname{Var}\left(\sum_{v=1}^n \tau(G,v)\right) &= \mathbb{E}\left[\operatorname{Var}\left(\sum_{v=1}^n \tau(G,v)\middle|\, \phi_1,\ldots,\phi_n\right)\right] \\ &\quad + \operatorname{Var}\left(\mathbb{E}\left[\sum_{v=1}^n \tau(G,v)\middle|\, \phi_1,\ldots,\phi_n\right]\right),\end{aligned}$$

together with an application of Efron-Stein's inequality to both terms on the right. Note that $\mathbb{E}\left[\sum_{v=1}^n \tau(G,v) |\, \phi_1,\ldots,\phi_n\right]$ is a constant in the *ordinary* SBM, whereas here it needs a careful analysis.

The sample average $\frac{1}{n}\sum_{v=1}^n \tau(G,v)$ concentrates then around $\mathbb{E}\left[\tau(T,o)\right]$, where (T, o) is the branching process from Section 5, with root o having spin drawn uniformly from $\{+, -\}$ and weight governed by ν, see Proposition 23. The coupling, and in particular the matching of the weights, plays an important role in its proof.

In the next section we apply the latter proposition to some specific functionals.

▶ **Proposition 22** (Degree-Corrected Extension of Proposition 35 in [2]). *Let G be drawn according to the DC-SBM. There exists $c > 0$ such that if $\tau, \varphi : \mathcal{G}^* \to \mathbb{R}$ are ℓ-local, $|\tau(G, o)| \le \varphi(G, o)$ and φ is non-decreasing by the addition of edges, then*

$$\operatorname{Var}\left(\sum_{v=1}^n \tau(G,v)\right) \le cn\rho^{2\ell}\left(\mathbb{E}\left[\max_{v\in[n]}\varphi^4(G,v)\right]\right)^{1/2}.$$

▶ **Proposition 23** (Degree-Corrected Extension of Proposition 36 in [2]). *Let G be drawn according to the DC-SBM. Let (T, o) be the branching process from Section 5, with root o having spin drawn uniformly from $\{+, -\}$ and weight governed by ν. Let $\ell = C\log_\rho(n)$, with $C < C_{coupling}$. There exists $c > 0$ such that if $\tau, \varphi : \mathcal{G}^* \to \mathbb{R}$ are ℓ-local, $|\tau(G, o)| \le \varphi(G, o)$ and φ is non-decreasing by the addition of edges, then*

$$\begin{aligned}&\mathbb{E}\left[\left|\frac{1}{n}\sum_{v=1}^n \tau(G,v) - \mathbb{E}\left[\tau(T,o)\right]\right|\right] \\ &\le c_2 n^{-\left(\frac{\gamma}{2}\wedge\frac{1}{40}\right)}\left(\mathbb{E}\left[\max_{v\in[n]}\varphi^4(G,v)\right]^{1/4} \vee \mathbb{E}\left[\varphi^2(T,o)\right]^{1/2}\right) + \mathcal{O}(n^{-\gamma})\end{aligned} \tag{31}$$

7.1 Application with some specific local functionals

Here we consider $\langle B^\ell\chi_1, B^\ell\chi_2\rangle$, $\langle B^{2\ell}\chi_k, B^\ell\chi_j\rangle$, and $\langle B^\ell B^{*\ell}\chi_1, B^\ell B^{*\ell}\chi_2\rangle$, quantities occurring in Proposition 4.

Explicitly, $B^\ell \chi_k(e) = \sum_f B^\ell_{ef} g_k(\sigma(f_2))\phi_{f_2}$, where we recall that B^ℓ_{ef} is the number of non-backtracking walks from e to f. Now, if the oriented $\ell-$ neighbourhood of e is a *tree*, then $B^\ell \chi_k(e) = \langle g_k, \Psi_\ell(e)\rangle$. With this intuition in mind, we analyse likewise expressions in Proposition 24 below.

Inspired by (26), which expresses $B^\ell B^{*\ell}\chi_k$ on *trees* in terms of the operator $Q_{k,\ell}$, we extend the latter to an operator defined on general graphs. First, for $e \in \vec{E}(V)$ and $t \geq 0$, set $\mathcal{Y}_t(e) = \{f \in \vec{E} : \vec{d}(e,f) = t\}$. Then, for $k \in \{1,2\}$, we set

$$P_{k,\ell}(e) = \sum_{t=0}^{\ell-1} \sum_{f \in \mathcal{Y}_t(e)} L_k(f), \tag{32}$$

with

$$L_k(f) = \sum_{(g,h)\in\mathcal{Y}_1(f)\setminus\mathcal{Y}_t(e); g\neq h} \langle g_k, \tilde{\Psi}_t(g)\rangle \tilde{S}_{\ell-t-1}(h),$$

where $\tilde{\Psi}_t(g)$, $\tilde{S}_{\ell-t-1}(h) = \|\tilde{Y}_{\ell-t-1}(h)\|_1$ are the variables $\Psi_t(g)$, respectively $S_{\ell-t-1}(h)$, defined on the graph G where all edges in $(G, e_2)_t$ have been removed. Note that, if $(G,e)_{2\ell}$ is a tree, then $\tilde{\Psi}_s(g) = \Psi_s(g)$ for $s \leq 2\ell - t$. Compare $P_{k,\ell}$ to $Q_{k,\ell}$ in (23) and $L_k(f)$ to $L^u_{k,\ell}$ in (28).

Finally, define

$$S_{k,\ell}(e) = S_\ell(e) g_k(\sigma(e_1))\phi_{e_1}. \tag{33}$$

We then have an extension of (26), when $(G, e_2)_{2\ell}$ is a tree:

$$B^\ell B^{*\ell}\check{\chi}_k(e) = P_{k,\ell}(e) + S_{k,\ell}(e). \tag{34}$$

We analyse (34) in Proposition 25 below.

7.1.1 The case $\mu_2^2 > \rho$

▶ **Proposition 24** (Degree-Corrected Extension of Proposition 37 in [2]). *Assume that $\mu_2^2 > \rho$. Let $\ell = C\log_\rho n$ with $0 < C < C_{coupling}$.*

(i) *For any $k \in \{1,2\}$, there exists $c'_k > 0$ such that, in probability,*

$$\frac{1}{n}\sum_{e\in\vec{E}} \frac{\langle g_k, \Psi_\ell(e)\rangle^2}{\mu_k^{2\ell}} \to c'_k.$$

(ii) *For any $k \in \{1,2\}$, there exists $c''_k > 0$ such that, in probability,*

$$\frac{1}{n}\sum_{e\in\vec{E}} \frac{\langle g_k, Y_\ell(e)\rangle^2}{\mu_k^{2\ell}} \to c''_k.$$

(iii)

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{e\in\vec{E}}\langle g_1, \Psi_\ell(e)\rangle\langle g_2, \Psi_\ell(e)\rangle\right|\right] \leq (\log n)^3 n^{2C-\left(\frac{\gamma}{2}\wedge\frac{1}{40}\right)} + n^{-\gamma}.$$

(iv) *For any $k \neq j \in \{1,2\}$,*

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{e\in\vec{E}}\langle g_k, \Psi_{2\ell}(e)\rangle\langle g_j, \Psi_\ell(e)\rangle\right|\right] \leq (\log n)^3 n^{3C-\left(\frac{\gamma}{2}\wedge\frac{1}{40}\right)} + n^{-\gamma}.$$

(v) *For any $k \in \{1,2\}$, in probability*

$$\frac{1}{n}\sum_{e\in\vec{E}}\frac{\langle g_k, \Psi_{2\ell}(e)\rangle\langle g_k, \Psi_\ell(e)\rangle}{\mu_k^{3\ell}} \to c_k'''.$$

▶ **Proposition 25** (Degree-Corrected Extension of Proposition 38 in [2]). *Assume that $\mu_2^2 > \rho$. Let $\ell = C\log_\rho n$ with $C < C_{coupling}$.*

(i) *For any $k \in \{1,2\}$, there exists $c_k'''' > 0$ such that in probability*

$$\frac{1}{n}\sum_{e\in\vec{E}}\frac{P_{k,\ell}^2(e)}{\mu_k^{4\ell}} \to c_k''''.$$

(ii)

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{e\in\vec{E}}(P_{1,\ell}(e)+S_{1,\ell}(e))(P_{2,\ell}(e)+S_{2,\ell}(e))\right|\right] \le (\log n)^8 n^{4C-\left(\frac{\gamma}{2}\wedge\frac{1}{40}\right)}$$

7.1.2 The case $\mu_2^2 \le \rho$

Most of the above claims continue to hold if $\mu_2^2 \le \rho$. We treat the exceptions here.

▶ **Proposition 26.** *Assume that $\mu_2^2 \le \rho$. Let $\ell = C\log_\rho n$ with $0 < C < C_{coupling}$. There exists some $c > 0$, such that w.h.p.,*

$$\frac{1}{n}\sum_{e\in\vec{E}}\frac{\langle g_2, \Psi_\ell(e)\rangle^2}{\rho^\ell} \ge c.$$

▶ **Proposition 27.** *Assume that $\mu_2^2 \le \rho$. Let $\ell = C\log_\rho n$ with $C < C_{coupling}$. There exists $c > 0$ such that w.h.p.,*

$$\frac{1}{n}\sum_{e\in\vec{E}}\frac{P_{2,\ell}^2(e)}{\rho^{2\ell}\log^5(n)} \le c.$$

8 Proof op Propositions 4 and 6

We introduce for $k \in \{1,2\}$ the vector $N_{k,\ell}$, defined on $e \in \vec{E}$ as

$$N_{k,\ell}(e) = \langle g_k, \Psi_\ell(e)\rangle.$$

If $(G, e_2)_\ell$ is a tree, then

$$N_{k,\ell}(e) = \langle B^\ell \chi_k, \delta_e\rangle,$$

and we have a similar expression for $B^\ell B^{*\ell}\check{\chi}_k$ in (34). Now, at most $\rho^{2\ell}\log(n)$ vertices have a cycle in their ℓ-neighbourhood (see Lemma 19). Therefore:

▶ **Lemma 28** (Degree-Corrected Extension of Lemma 39 in [2]). *Let $\ell = C\log_\rho n$ with $0 < C < C_{min}$. Then, w.h.p. $\|B^\ell\chi_k - N_{k,\ell}\| = O\left((\log n)^{5/2}\rho^{2\ell}\right) = o\left(\rho^{\ell/2}\sqrt{n}\right)$, $\|B^\ell B^{*\ell}\check{\chi}_k - P_{k,\ell} - S_{k,\ell}\| = O((\log n)^4\rho^{4\ell})$ and $\|B^\ell B^{*\ell}\check{\chi}_k - P_{k,\ell}\| = O(\rho^\ell\sqrt{n})$.*

Proof. The proof of Lemma 39 in [2] can be easily adapted to the current setting. The key idea is pointed out above. It thus remains to bound $|(B^\ell \chi_k - N_{k,\ell})(e)|$ and $|(B^\ell B^{*\ell} \check{\chi}_k - P_{k,\ell})(e)|$ on edges e for which $(G, e_2)_\ell$ is not a tree. For this, use that with high probability the graph is 2ℓ-tangle free so that there are at most two non-backtracking paths between e and any edge at distance ℓ. ◀

We can thus in our calculations replace $B^\ell \chi_k$ by $N_{k,\ell}$ and $B^\ell B^{*\ell} \check{\chi}_k$ by $P_{k,\ell}$. From Propositions 24 and 25, Proposition 4 then follows:

Proof of Proposition 4. This proof follows the corresponding proof in [2]. We give the key observations: (i) From Proposition 24 (i), $\|N_{k,\ell}\| \sim \sqrt{n}\mu_k^\ell$ and from Proposition 25 (i), $\|P_{k,\ell}\| \sim \sqrt{n}\mu_k^{2\ell}$.

(ii) From Proposition 24 (v), $|\langle N_{k,\ell}, N_{k,2\ell}\rangle| \sim n\mu_k^{3\ell}$.

(iii) From Proposition 24 (iii), $|\langle N_{1,\ell}, N_{2,\ell}\rangle| \sim (\log n)^3 n^{3C-\left(\frac{\gamma}{2}\wedge\frac{1}{40}\right)}$.

(iv) From Proposition 24 (iv), $|\langle N_{k,2\ell}, N_{j,\ell}\rangle| \sim (\log n)^3 n^{4C-\left(\frac{\gamma}{2}\wedge\frac{1}{40}\right)}$.

(v) From Proposition 25 (ii), $|\langle P_{1,\ell} + S_{1,\ell}, P_{2,\ell} + S_{2,\ell}\rangle| \sim (\log n)^8 n^{5C-\left(\frac{\gamma}{2}\wedge\frac{1}{40}\right)}$. ◀

Proposition 6 follows similarly from the case $\mu_2^2 \le \rho$ treated in Section 7.1:

Proof of Proposition 6. This follows from Propositions 26 and 27 in conjunction with Lemma 28. ◀

9 Norm of non-backtracking matrices

The proofs of the statements in this section can be found in Appendix D in the detailed version of the underlying article (Arxiv:1609.02487).

In this section the product over an empty set is defined to be one.

It is convenient to extend matrix B and vector χ_k to the set of directed edges on the *complete* graph, $\vec{E}_K(V) = \{(u,v) : u \neq v \in V\}$: For $e, f \in \vec{E}_K(V)$, B_{ef} is then extended to

$$B_{ef} = A_e A_f 1_{e_2 = f_1} 1_{e_1 \neq f_2}, \tag{35}$$

where A is the adjacency matrix. For each $e \in \vec{E}_K(V)$ we set $\chi_k(e) = g_k(\sigma(e_2))\phi_{e_2}$.

For integer $k \ge 1$, $e, f \in \vec{E}_K(V)$, we let Γ_{ef}^k be the set of non-backtracking walks $\gamma = (\gamma_0, \dots, \gamma_k)$ of length k from $(\gamma_0, \gamma_1) = e$ to $(\gamma_{k-1}, \gamma_k) = f$ on the *complete* graph with vertex set V.

By induction it follows that

$$(B^k)_{ef} = \sum_{\gamma \in \Gamma_{ef}^{k+1}} \prod_{s=0}^{k} A_{\gamma_s \gamma_{s+1}}. \tag{36}$$

Indeed, note that $\prod_{s=0}^k A_{\gamma_s\gamma_{s+1}}$ is one when γ is a path in G and zero otherwise.

To each walk $\gamma = (\gamma_0, \dots, \gamma_k)$, we associate the graph $G(\gamma) = (V(\gamma), E(\gamma))$, with the set of vertices $V(\gamma) = \{\gamma_i, 0 \le i \le k\}$ and the set of edges $E(\gamma) = \{\{\gamma_i, \gamma_{i+1}\}, 0 \le i \le k-1\}$.

From Lemma 19, the graphs following the DC-SBM are tangle-free with high probability. Hence, it makes sense to consider the subset $F_{ef}^{k+1} \subset \Gamma_{ef}^{k+1}$ of tangle-free non-backtracking walks on the *complete* graph. Indeed, if G is tangle-free, we need only consider the tangle-free paths in the summation (36):

$$(B^{(k)})_{ef} = \sum_{\gamma \in F_{ef}^{k+1}} \prod_{s=0}^{k} A_{\gamma_s \gamma_{s+1}}, \tag{37}$$

and $B^k = B^{(k)}$ for $1 \le k \le \ell$.

Define for $u \ne v$ the *centred* random variable

$$\underline{A}_{uv} = A_{uv} - \frac{\phi_u \phi_v}{n} W_{\sigma_u \sigma_v}, \tag{38}$$

where

$$W = \begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

Compare this to the SBM *without* degree-corrections in Section 10.1 of [2]: $\phi_u = 1$ for all u in the latter model.

Using $\underline{A}$ we shall attempt to center B^k when the underlying graph G is tangle-free through considering

$$\Delta_{ef}^{(k)} = \sum_{\gamma \in F_{ef}^{k+1}} \prod_{s=0}^{k} \underline{A}_{\gamma_s \gamma_{s+1}}. \tag{39}$$

Further, we set

$$\Delta_{ef}^{(0)} = 1_{e=f} \underline{A}_e \quad \text{and} \quad B_{ef}^{(0)} = 1_{e=f} A_e. \tag{40}$$

To decompose (37), following a decomposition that appeared first in [16], we use

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

with $x_s = A_{\gamma_s \gamma_{s+1}}$ and $y_s = \underline{A}_{\gamma_s \gamma_{s+1}}$ on a path $\gamma \in F_{ef}^{k+1}$:

$$\prod_{s=0}^{\ell} A_{\gamma_s \gamma_{s+1}} = \prod_{s=0}^{\ell} \underline{A}_{\gamma_s \gamma_{s+1}} + \sum_{t=0}^{\ell} \prod_{s=0}^{t-1} \underline{A}_{\gamma_s \gamma_{s+1}} \left(\frac{\phi_{\gamma_t} \phi_{\gamma_{t+1}}}{n} W_{\sigma_{\gamma_t} \sigma_{\gamma_{t+1}}} \right) \prod_{s=t+1}^{\ell} A_{\gamma_s \gamma_{s+1}}.$$

Summing over all $\gamma \in F_{ef}^{\ell+1}$ then gives

$$\begin{aligned} B_{ef}^{(\ell)} &= \sum_{\gamma \in F_{ef}^{\ell+1}} \prod_{s=0}^{\ell} \underline{A}_{\gamma_s \gamma_{s+1}} \\ &\quad + \sum_{t=0}^{\ell} \sum_{\gamma \in F_{ef}^{\ell+1}} \prod_{s=0}^{t-1} \underline{A}_{\gamma_s \gamma_{s+1}} \left(\frac{\phi_{\gamma_t} \phi_{\gamma_{t+1}}}{n} W_{\sigma_{\gamma_t} \sigma_{\gamma_{t+1}}} \right) \prod_{s=t+1}^{\ell} A_{\gamma_s \gamma_{s+1}} \\ &= \Delta_{ef}^{(\ell)} + \sum_{t=0}^{\ell} \sum_{\gamma \in F_{ef}^{\ell+1}} \prod_{s=0}^{t-1} \underline{A}_{\gamma_s \gamma_{s+1}} \left(\frac{\phi_{\gamma_t} \phi_{\gamma_{t+1}}}{n} W_{\sigma_{\gamma_t} \sigma_{\gamma_{t+1}}} \right) \prod_{s=t+1}^{\ell} A_{\gamma_s \gamma_{s+1}}. \end{aligned} \tag{41}$$

Consider the two products in the summation over $F_{ef}^{\ell+1}$ on the right of (41): We can, for $1 \le t \le \ell - 1$, replace the summation over $F_{ef}^{\ell+1}$ by summing over all pairs $\gamma' = (\gamma_0, \dots, \gamma_t) \in F_{eg}^{t}$ and $\gamma'' = (\gamma_{t+1}, \dots, \gamma_{\ell+1}) \in F_{g'f}^{\ell-t}$ for some $g, g' \in \vec{E}(V)$ such that there exists a non-backtracking path with one intermediate edge, on the *complete* graph, between oriented edges g and g' (we denote this property by $g \xrightarrow{2} g'$). However caution is needed, as this summation also includes *tangled* paths, namely those in the sets $\{F_{t,ef}^{\ell+1}\}_{t=0}^{\ell}$. Where, for $1 \le t \le \ell - 1$,

$F_{t,ef}^{\ell+1}$ is defined as the collection of all *tangled* paths $\gamma = (\gamma_0, \ldots, \gamma_{\ell+1}) = (\gamma', \gamma'') \in \Gamma_{ef}^{\ell+1}$ with γ' and γ'' as above. For $t = 0$, $F_{0,ef}^{\ell+1}$ consists of all non-backtracking *tangled* paths (γ', γ'') with $\gamma' = (e_1)$ and $\gamma'' \in F_{g'f}^{\ell}$ for any g' such that $g'_1 = e_2$. For $t = \ell$, $F_{\ell,ef}^{\ell+1}$ is the set of non-backtracking *tangled* paths (γ', γ'') such that $\gamma'' = (f_2)$ and $\gamma' \in F_{eg}^{\ell}$ for some $g \in \vec{E}(V)$ with $g_2 = f_1$. We rewrite (41) as

$$B^{(\ell)} = \Delta^{(\ell)} + \frac{1}{n} K B^{(\ell-1)} + \frac{1}{n} \sum_{t=1}^{\ell-1} \Delta^{(t-1)} K^{(2)} B^{(\ell-t-1)} + \frac{1}{n} \Delta^{(\ell-1)} \widehat{K} - \frac{1}{n} \sum_{t=0}^{\ell} R_t^{(\ell)}, \tag{42}$$

where for $e, f \in E_K$,

$$K_{ef} = 1_{e \to f} \phi_{e_1} \phi_{e_2} W_{\sigma(e_1)\sigma(e_2)}, \tag{43}$$

the *weighted* non-backtracking matrix on the *complete* graph (recall that $e \to f$ represents the non-backtracking property),

$$\widehat{K}_{ef} = 1_{e \to f} \phi_{f_1} \phi_{f_2} W_{\sigma(f_1)\sigma(f_2)}, \tag{44}$$

$$K_{ef}^{(2)} = 1_{e \overset{2}{\to} f} \phi_{e_2} \phi_{f_1} W_{\sigma(e_2)\sigma(f_1)}, \tag{45}$$

and where

$$(R_t^{(\ell)})_{ef} = \sum_{\gamma \in F_{t,ef}^{\ell+1}} \prod_{s=0}^{t-1} \underline{A}_{\gamma_s \gamma_{s+1}} \phi_{\gamma_t} \phi_{\gamma_{t+1}} W_{\sigma(\gamma_t)\sigma(\gamma_{t+1})} \prod_{s=t+1}^{\ell} A_{\gamma_s \gamma_{s+1}}. \tag{46}$$

Indeed,

$$\begin{aligned}
\left(\sum_{t=1}^{\ell-1} \Delta^{(t-1)} K^{(2)} B^{(\ell-t-1)} \right)_{ef} &= \sum_{t=1}^{\ell-1} \sum_{g,g'} \Delta_{eg}^{(t-1)} K_{gg'}^{(2)} B_{g'f}^{(\ell-t-1)} \\
&= \sum_{t=1}^{\ell-1} \sum_{g,g'} \sum_{\gamma' \in F_{eg}^{t}} \sum_{\gamma'' \in F_{g'f}^{\ell-t}} \prod_{s=0}^{t-1} \underline{A}_{\gamma'_s \gamma'_{s+1}} 1_{g \overset{2}{\to} g'} \phi_{\gamma'_t} \phi_{\gamma''_0} \\
&\quad \cdot W_{\sigma(\gamma'_t)\sigma(\gamma''_0)} \prod_{s=0}^{\ell-t-1} A_{\gamma''_s \gamma''_{s+1}},
\end{aligned} \tag{47}$$

$$\left(K B^{(\ell-1)} \right)_{ef} = \sum_{g} \sum_{\gamma'' \in F_{gf}^{\ell}} 1_{e \to g} \phi_{e_1} \phi_{e_2} W_{\sigma(e_1)\sigma(e_2)} A_{e_2, g_2} \prod_{s=1}^{\ell-2} A_{\gamma''_s \gamma''_{s+1}} A_{f_1 f_2}, \tag{48}$$

and,

$$\left(\Delta^{(\ell-1)} \widehat{K} \right)_{ef} = \sum_{g} \sum_{\gamma' \in F_{eg}^{\ell}} \underline{A}_{e_1 e_2} \prod_{s=1}^{\ell-2} \underline{A}_{\gamma'_s \gamma'_{s+1}} \underline{A}_{g_1 f_1} 1_{g \to f} \phi_{f_1} \phi_{f_2} W_{\sigma(f_1)\sigma(f_2)} \tag{49}$$

that is exactly the splitting described just below (41), where we also pointed out the need to compensate for *tangled* paths occuring in (47), which is precisely the role of $R_t^{(\ell)}$ in (42).

To bound (42), we introduce

$$\overline{W} = \frac{2}{\Phi^{(2)}} \left(\rho \chi_1 \check{\chi}_1^* + \mu_2 \chi_2 \check{\chi}_2^*\right) = \left(\phi_{e_2} \phi_{f_1} W_{\sigma(e_2)\sigma(f_1)}\right)_{ef}, \tag{50}$$

and,

$$L = K^{(2)} - \overline{W}. \tag{51}$$

Note the presence of weights in (50), hence our choice for the candidate eigenvectors.

Further, we set for $1 \leq t \leq \ell - 1$,

$$S_t^{(\ell)} = \Delta^{(t-1)} L B^{(\ell-t-1)}. \tag{52}$$

We then have:

▶ **Proposition 29** (Degree-Corrected Extension of Proposition 13 in [2]). *If G is tangle-free and $x \in \mathbb{C}^{\vec{E}(V)}$ with norm smaller than one, we have*

$$\begin{aligned}\|B^\ell x\| \leq \|\Delta^{(\ell)}\| + \frac{1}{n}\|KB^{(\ell-1)}\| + \frac{1}{n}\sum_{j=1,2}\frac{2\mu_j}{\Phi^{(2)}}\sum_{t=1}^{\ell-1}\|\Delta^{(t-1)}\chi_j\|\|\langle\check{\chi}_j, B^{\ell-t-1}x\rangle\| \\ + \frac{1}{n}\sum_{t=1}^{\ell-1}\|S_t^{(\ell)}\| + \phi_{max}^2(a \vee b)\|\Delta^{(\ell-1)}\| + \frac{1}{n}\sum_{t=0}^{\ell}\|R_t^{(\ell)}\|.\end{aligned}$$

Proof. Due to the tangle-freeness, $B^\ell = B^{(\ell)}$. Further $K^{(2)} = L + \overline{W}$ and $||K|| \leq \phi_{\max}^2(a \vee b)n$. ◀

In Appendix D in the detailed version of the underlying article (Arxiv:1609.02487) we prove the following bounds on the matrices in Proposition 29:

▶ **Proposition 30** (Degree-Corrected Extension of Proposition 14 in [2]). *Let $\ell = C \log_\rho n$ with $C < 1$. With high probability, the following norm bounds hold for all k, $0 \leq k \leq \ell$, and $i = 1, 2$:*

$$\|\Delta^{(k)}\| \leq (\log n)^{10}\rho^{k/2}, \tag{53}$$

$$\|\Delta^{(k)}\chi_i\| \leq (\log n)^5\rho^{k/2}\sqrt{n}, \tag{54}$$

$$\|R_k^{(\ell)}\| \leq (\log n)^{25}\rho^{\ell-k/2}, \tag{55}$$

$$\|KB^{(k)}\| \leq \sqrt{n}(\log n)^{10}\rho^k, \tag{56}$$

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

$$\|S_k^{(\ell)}\| \leq \sqrt{n}(\log n)^{20}\rho^{\ell-k/2}. \tag{57}$$

9.1 Proof of Proposition 5

From Propositions 29 and 30, the geometric growth in Corollary 21 together with the tangle-freeness due to Lemma 19, the proof of Proposition 5 follows:

Let $j \in \{1, 2\}$. If, for some vector x, $\langle\check{\varphi}_j, x\rangle = 0$, then $\langle B^\ell\chi_j, \check{x}\rangle = 0$. Therefore, using Corollary 21,

$$\begin{aligned}\sup_{\|x\|=1,\langle\check{\varphi}_j,x\rangle=0}\langle\check{\chi}_j, B^{\ell-t-1}x\rangle &= \sup_{\|x\|=1,\langle B^\ell\chi_j,\check{x}\rangle=0}\langle B^{\ell-t-1}\chi_j, \check{x}\rangle \\ &= \sup_{\|\check{x}\|=1,\langle B^\ell\chi_j,\check{x}\rangle=0}\langle B^{\ell-t-1}\chi_j, \check{x}\rangle \\ &\leq \log^2(n)n^{1/2}\rho^{\frac{\ell-t-1}{2}}.\end{aligned} \tag{58}$$

With high probability, the graph is $\ell-$ tangle free (Lemma 19). Thus, invoking Propositions 29 and 30, with high probability,

$$\begin{aligned}\sup_{x\in H^{\perp},\|x\|=1} \|B^{\ell}x\| &\leq \log^{10}(n)\rho^{\frac{\ell}{2}} + n^{-1/2}\log^{10}(n)\rho^{\ell-1} \\ &\quad + c_1\log^{8}(n)\rho^{\frac{\ell}{2}} + n^{-1/2}\log^{21}(n)\rho^{\ell} \\ &\quad + c_2\log^{10}(n)\rho^{\frac{\ell}{2}} + n^{-1}\log^{26}(n)\rho^{\ell} \\ &\leq \log^{c}(n)\rho^{\frac{\ell}{2}},\end{aligned} \tag{59}$$

since $C < 1$.

9.2 Comparison with the Stochastic Block Model in [2]

Putting $\phi_u = 1$ for all u, we retrieve exactly **the same bounds as in the Stochastic Block Model**, that is equations (30) – (34) in [2].

Below we use the trace method and therefore path counting combinatorial arguments to establish Proposition 30. In particular, we bound the expectation of expressions of the form

$$\mathbb{E}\left[\prod_{i=1}^{2m}\prod_{s=1}^{k}\underline{A}_{\gamma_{i,s-1}\gamma_{i,s}}\right], \tag{60}$$

for certain paths $\gamma = (\gamma_1, \ldots, \gamma_{2m})$ with $\gamma_i = (\gamma_{i,0}, \cdots, \gamma_{i,k}) \in V^{k+1}$, where $\underline{A}$ is defined in (38).

In bounding (60) the following term occurs:

$$\prod_{u\in V(\gamma)} \Phi^{(d_u)},$$

where $(d_u)_u$ are the degrees of the vertices in a specific tree (or forest) spanning the path γ. See, for instance, (D.4) and (D.17) in the detailed version of the underlying article (Arxiv:1609.02487). Here lies a major complication with respect to the Stochastic Block Model: those terms are not present in the latter model. In (D.8) and (D.19) in the detailed version of the underlying article, we find

$$\prod_{u=1}^{|V(\gamma)|} \Phi^{(d_u)} \leq C_2^{\sum_{u:d_u>2}(d_u-2)}\left(\Phi^{(2)}\right)^{|V(\gamma)|-n_{\mathcal{C}}},$$

where $C_2 > 1$ is some constant and where $n_{\mathcal{C}} \geq 1$ is the number of components on the path γ. To compare this term with powers of $\Phi^{(2)}$ (which are present in powers of $\rho = \frac{a+b}{2}\Phi^{(2)}$), we bound $\sum_{u:d_u>2}(d_u - 2)$, see in particular Lemma (D.2) and (D.5) in the detailed version of the underlying article.

10 Detection: Proof of Theorem 2

The proofs of the statements in this section are deferred to Appendix E in the detailed version of the underlying article (Arxiv:1609.02487).

We need the following special case of a lemma in [2]:

▶ **Lemma 31** (Special case of Lemma 40 in [2]). *Assume that there exists a function $F : V \to \{0,1\}$ such that in probability, for any $i \in \{+,-\}$,*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=i}F(v) = \frac{f(i)}{2},$$

where $f : \{+,-\} \to [0,1]$ is such that $f(+) > f(-)$. Then, assigning to each vertex a label $\widehat{\sigma}(v) = +$ if $F(v) = 1$ and $\widehat{\sigma}(v) = -$ if $F(v) = 0$, yields asymptotically positive overlap with the true spins.

Recall the eigenvector ξ_2 from Theorem 1. Below we use the function $F : v \mapsto 1_{\sum_{e:e_2=v}\xi_2(e) > \frac{\tau}{\sqrt{n}}}$ or $F : v \mapsto 1_{\sum_{e:e_2=v}\xi_2(e) \leq \frac{\tau}{\sqrt{n}}}$ for some fixed parameter τ. We verify also that ξ_2 is aligned with $P_{2,\ell}$. It is therefore useful to introduce the vector I_ℓ, defined element-wise by

$$I_\ell(v) = \sum_{e\in\vec{E}:e_2=v} P_{2,\ell}(e), \tag{61}$$

for $v \in V$.

Further, put

$$\widehat{c} = \frac{a+b}{2}\frac{(\Phi^{(1)})^2\Phi^{(3)}}{\Phi^{(2)}}\frac{\rho}{\mu_2^2-\rho}\mu_2$$

The following lemma shows that I_ℓ is correlated with the spins:

▶ **Lemma 32** (Degree-Corrected Extension of Lemma 41 in [2]). *Let $\ell = C\log_\rho n$ with $C < C_{coupling}$ and $i \in \{+,-\}$. There exists a random variable Y_i such that $\mathbb{E}[Y_i] = 0$, $\mathbb{E}[|Y_i|] < \infty$ and for any continuity point t of the distribution of Y_i, in L^2,*

$$\frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=i}1_{I_\ell(v)\mu_2^{-2\ell}-\widehat{c}g_2(i)\geq t} \to \frac{1}{2}\mathbb{P}\left(Y_i \geq t\right).$$

Recall from Theorem 1 that the eigenvector ξ_2 is asymptotically *aligned* with

$$\frac{B^\ell B^{*\ell}\check{\chi}_2}{\|B^\ell B^{*\ell}\check{\chi}_2\|}, \tag{62}$$

where $\ell \sim \log_\rho(n)$. Hence, for some unknown sign ω, the vector $\xi_2' = \omega\xi_2$ is asymptotically *close* to (62). From Lemma 28 we know that $B^\ell B^{*\ell}\check{\chi}_2$ and $P_{2,\ell}$ are asymptotically close. Consequently, properly renormalizing ξ_2' will make it asymptotically close to $P_{2,\ell}$, so that we can replace $P_{2,\ell}$ in (61) by ξ_2'. That is, we set for $v \in V$,

$$I(v) = \sum_{e:e_2=v} s\sqrt{n}\xi_2'(e),$$

with $s = \sqrt{c_2''''}$ the limit in Proposition 25. Then, I and $I_\ell/\mu_2^{2\ell}$ are close, which leads to the following lemma:

▶ **Lemma 33** (Degree-Corrected Extension of Lemma 42 in [2]). *Let $i \in \{+,-\}$ and $\widehat{Y}_i$ be as in Lemma 32. For any continuity point t of the distribution of $\widehat{Y}_i$, in L^2,*

$$\frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=i}1_{I(v)-\widehat{c}g_2(i)\geq t} \to \frac{1}{2}\mathbb{P}\left(\widehat{Y}_i \geq t\right).$$

Put for $i \in \{+,-\}$, $X_i = \widehat{Y}_i + \widehat{c}g_2(i) = \widehat{Y}_i + \frac{1}{\sqrt{2}}\widehat{c}i$. Then, for all $t \in \mathbb{R}$ that are continuity points of the distribution of X_i, the following convergence holds in probability

$$\frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=i} 1_{I(v)>t} \to \frac{1}{2}\mathbb{P}\left(X_i > t\right).$$

Since $\mathbb{E}\left[X_+\right] > 0$, the argument below (90) in [2] establishes the existence of a continuity point $t_0 \in \mathbb{R}$ such that $\mathbb{P}\left(X_+ > t_0\right) > \mathbb{P}\left(X_- > t_0\right)$.

Further, we note that X_+ is in distribution equal to $-X_-$, a fact that we use below.

We are now in a position to apply Lemma 31 and thereby finishing the proof of Theorem 2:

If $\omega = 1$, then we define F, for $v \in V$, by

$$F(v) = 1_{\sum_{e:e_2=v} \xi_2(e) > \frac{t_0}{s\sqrt{n}}} = 1_{I(v)>t_0}.$$

Then,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=+}F(v) = \frac{1}{2}\mathbb{P}\left(X_+ > t_0\right) =: \frac{f(+)}{2},$$

and,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=-}F(v) = \frac{1}{2}\mathbb{P}\left(X_- > t_0\right) =: \frac{f(-)}{2},$$

so that $f(+) > f(-)$ and Lemma 31 applies.

If, however, $\omega = -1$, then we define F, for $v \in V$, by

$$F(v) = 1_{\sum_{e:e_2=v} \xi_2(e) \le \frac{t_0}{s\sqrt{n}}} = 1_{-I(v)\le t_0}.$$

Then, this time,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=+}F(v) = \lim_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=+}1_{I(v)>-t_0} = \frac{1}{2}\mathbb{P}\left(X_+ > -t_0\right) =: \frac{f(+)}{2},$$

since $-t_0$ is a continuity point of X_+, which follows from the fact that X_+ is in distribution equal to $-X_-$ and t_0 is a continuity point of X_-.

Similarly,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1_{\sigma(v)=-}F(v) = \frac{1}{2}\mathbb{P}\left(X_- > -t_0\right) =: \frac{f(-)}{2}.$$

Now,

$$f(+) = \mathbb{P}\left(X_+ > -t_0\right) = 1 - \mathbb{P}\left(X_- > t_0\right) > 1 - \mathbb{P}\left(X_+ > t_0\right) = \mathbb{P}\left(X_- > -t_0\right) = f(-),$$

exactly the setting of Lemma 31.

References

1 B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Algorithms*, 31(1):3–122, August 2007.

2 C. Bordenave, M. Lelarge, and L. Massoulié. Non-backtracking spectrum of random graphs: community detection and non-regular ramanujan graphs. *arXiv preprint 1501.06087*, 2015.

3 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, Dec 2011.

4 W. Evans, C. Kenyon, Y. Peres, and L. Schulman. Broadcasting on trees and the ising model. *Ann. Appl. Probab.*, 10(2):410–433, 05 2000.

5 Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

6 L. Gulikers, M. Lelarge, and L. Massoulié. An impossibility result for reconstruction in a degree-corrected planted-partition model. *arXiv preprint 1511.00546*, 2015.

7 L. Gulikers, M. Lelarge, and L. Massoulié. A spectral method for community detection in moderately-sparse degree-corrected stochastic block models. *arXiv preprint 1506.08621*, 2015.

8 P. Holland, K. Laskey, and S. Leinhardt. Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137, June 1983.

9 M. Horton, H. Stark, and T. Terras. What are zeta functions of graphs and what are they good for? *Contemporary Mathematics, Quantum Graphs and Their Applications;*, 415:173–190, 2006.

10 K. i. Hashimoto. Zeta functions of finite graphs and representations of p-adic groups. In *Automorphic Forms and Geometry of Arithmetic Varieties*, volume 15 of *Advanced Studies in Pure Mathematics*, pages 211 – 280. Academic Press, 1989.

11 B. Karrer and M. Newman. Stochastic blockmodels and community structure in networks. *Phys. Rev. E*, 83:016107, Jan 2011.

12 H. Kesten and B. P. Stigum. Additional limit theorems for indecomposable multidimensional galton-watson processes. *Ann. Math. Statist.*, 37(6):1463–1481, 12 1966.

13 H. Kesten and B. P. Stigum. A limit theorem for multidimensional galton-watson processes. *Ann. Math. Statist.*, 37(5):1211–1223, 10 1966.

14 F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang. Spectral redemption in clustering sparse networks. *Proceedings of the National Academy of Sciences*, 110(52):20935–20940, 2013.

15 A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

16 L. Massoulié. Community detection thresholds and the weak ramanujan property. *ACM Symposium on the Theory of Computing (STOC)*, 2014.

17 E. Mossel, J. Neeman, and A. Sly. A proof of the block model threshold conjecture. *arXiv preprint 1311.4115*, 2015.

18 E. Mossel, J. Neeman, and A. Sly. Reconstruction and estimation in the planted partition model. *Probability Theory and Related Fields*, 162(3):431–461, 2015.

19 R. Murty. Ramanujan graphs. *J. Ramanujan Math. Soc.*, 18(1):1–20, 2003.