Non-Backtracking Spectrum of Degree-Corrected Stochastic Block Models

Lennart Gulikers, Marc Lelarge, Laurent Massoulié

September 9, 2016

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 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 $\frac{a \phi_u \phi_v}{n}$ and the inter-cluster connection probability is $\frac{b \phi_u \phi_v}{n}$.

We show that with high probability, the following holds: The leading eigenvalue of the non-backtracking matrix $B$ is asymptotic to $\rho = a + b \frac{\Phi^{(2)}}{2}$. The second eigenvalue is asymptotic to $\mu_2 = a - b \frac{\Phi^{(2)}}{2}$ 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.

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.

Motivated by community detection problems, we are interested in the spectrum of $B$ when $G$ is a random graph. We characterise its leading eigenvalues and corresponding eigenvectors when the number of vertices in $G$ tends to infinity. Above a certain threshold, the second eigenvalue of $B$ is correlated with the underlying communities.

The random graph model that we consider is the Degree-Corrected Stochastic Block Model (DC-SBM) [11], 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. 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 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 \( \frac{\phi_u \phi_v}{2n} \) and the inter-cluster connection probability is \( \frac{\mu_2}{2n} \), for two constants \( a, b > 0 \). Note that we obtain the ordinary SBM by setting \( \phi_1 = \ldots = \phi_n = 1 \).

The ordinary SBM is already known to contain a phase-transition in its sparse regime: In [18], impossibility of reconstruction when \( (\frac{a+b}{n})^2 \leq \frac{a+b}{2n} \) is shown. Above the threshold (i.e., \( (\frac{a+b}{n})^2 > \frac{a+b}{2n} \)), positively-correlated reconstruction can be obtained by thresholding the second-eigenvector of \( B \) [2].

We show here that the DC-SBM exhibits a similar behaviour. Specifically, we show that the same algorithm as in [2] succeeds in finding a positively correlated clustering when \( (\frac{a+b}{n}\Phi^{(2)})^2 > \frac{a+b}{2n}\Phi^{(2)} \). This is best possible as we demonstrated in an earlier work [6]: detection is impossible when \( (\frac{a+b}{n}\Phi^{(2)})^2 \leq \frac{a+b}{2n}\Phi^{(2)} \).

This method is robust in the sense that it succeeds all the way to the detectability-threshold. We emphasize that no modification of the non-backtracking matrix is needed to perform community detection in the DC-SBM (compare this to the adjacency matrix, which needs to be adapted to the degree-corrected setting [7]). In particular, the same algorithm succeeds: it does not need to know any information on the weights.

The same goes for the strategy in [2]: although the proofs presented here are technically more involved, the general idea remains the same.

Informally, we have the following results: With high probability, the leading eigenvalue of the non-backtracking matrix \( B \) is asymptotic to \( \rho = \frac{a+b}{2n}\Phi^{(2)} \). The second eigenvalue is asymptotic to \( \mu_2 = \frac{a+b}{2n}\Phi^{(2)} \) 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 \( \mu_2^2 > \rho \).

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.

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 \( (\frac{a+b}{n})^2 > \frac{a+b}{2n} \), 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 \( (\frac{a+b}{n})^2 \leq \frac{a+b}{2n} \).

In [18], impossibility of reconstruction when \( (\frac{a+b}{n})^2 \leq \frac{a+b}{2n} \) 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 \( \frac{a+b}{2} \)^2 > \( \frac{a+b}{4} \).

Later, in [17] (a work established independently from [16]), 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.

Most 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 \( \lambda \) has no eigenvalues \( \lambda \) such that \( \left| \lambda \right| \in (\sqrt{\rho_B},\rho_B) \), where \( \rho_B \) 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 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 2.1 on the spectrum of \( B \) and its consequences for community detection, Theorem 2.2.

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 2.1. It builds on Propositions 4.1 and 4.2. 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].

3
In Section 8 we apply those local functionals to establish Proposition 4.1.

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 2.2. We use the symmetry present in the two-communities setting here, which gets in general broken in models with more than two communities.

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 $n_+ = \frac{n}{2} + O(n^{1/2})$, and $n_- = \frac{n}{2} - O(n^{1/2})$. 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

$$M = \frac{\Phi^{(2)}}{2} \begin{pmatrix} a & b \\ b & a \end{pmatrix}. \quad (2.2)$$

We introduce the orthonormal vectors

$$g_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \text{and} \quad g_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (2.3)$$

together with the scalars

$$\rho = \mu_1 = \frac{a + b}{2} \Phi^{(2)}, \quad \text{and} \quad \mu_2 = \frac{a - b}{2} \Phi^{(2)}. \quad (2.4)$$

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

$$g_k^* M = \mu_k g_k^*, \quad k = 1, 2. \quad (2.5)$$
Note that $\rho$ and $\mu_2$ are also asymptotically eigenvalues of the expected adjacency matrix conditional on the weights. Indeed, if $A$ denotes the adjacency matrix and $\psi_k$ is the vector defined for $u \in V$ by $\psi_k(u) = g_k(\sigma(u))\phi_u$, then
\[
\mathbb{E}[A|\phi_1, \ldots, \phi_n] = \frac{a + b}{n} \psi_1 \psi_1^* + \frac{a - b}{n} \psi_2 \psi_2^* - \frac{1}{n} \text{diag}(\phi_n^2).
\]
So that, by abusively notating a vector of norm $O(1)$ by $O(1)$,
\[
\mathbb{E}[A|\phi_1, \ldots, \phi_n] \psi_1 = \left(\frac{a + b}{n} \sum_{u=1}^n \phi_1^2\right) \psi_1 + \left(\frac{a - b}{n} \sum_{u=1}^n \sigma_u \phi_1^2\right) \psi_2 + O(1) \to \rho \psi_1,
\]
and
\[
\mathbb{E}[A|\phi_1, \ldots, \phi_n] \psi_2 = \left(\frac{a - b}{n} \sum_{u=1}^n \phi_1^2\right) \psi_2 + \left(\frac{a + b}{n} \sum_{u=1}^n \sigma_u \phi_1^2\right) \psi_1 + O(1) \to \mu_2 \psi_2,
\]
by the law of large numbers and (2.1).

Finally, we define for $k \in \{1, 2\}$,
\[
\chi_k(e) = g_k(\sigma(e))\phi_{e_2}, \quad \text{for } e \in \overline{E}.
\] (2.6)

We show that the candidate eigenvectors
\[
\xi_k = \frac{B^T B^{(i)} \chi_k}{\|B^T B^{(i)} \chi_k\|}
\] (2.7)
are then, for $\ell \sim \log(n)$, asymptotically aligned with the first two eigenvectors of $B$. Note the weight in (2.6), which is not present in the ordinary SBM.

**Theorem 2.1** (Degree-Corrected Extension of Theorem 4 in [2]). Let $G$ be drawn according to the DC-SBM such that assumption (2.1) holds. Assume that $\ell = C_{\text{min}} \log(n)$, with $C_{\text{min}} > 0$ a small constant defined in (2.9).

If $\mu_2 > \rho$, then, with high probability, the eigenvalues $\lambda_i$ 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 $k \in \{1, 2\}$, $\xi_k$ is a normalized eigenvector associated to $\lambda_k$, then $\xi_k$ is asymptotically aligned with $\xi_k$. The vectors $\xi_1$ and $\xi_2$ are asymptotically orthogonal.

If $\mu_2 \leq \rho$, then, with high probability, the eigenvalues $\lambda_i$ of $B$ satisfy
\[
|\lambda_1 - \rho| = o(1), \quad \text{and, for } i \geq 2, \quad |\lambda_i| \leq \sqrt{\rho} + o(1).
\]
Further, $\xi_1$ is asymptotically aligned with $\xi_1$.

In Theorem 2.2 we show that positively correlated clustering is possible based on the second eigenvector of $B$ when above the feasibility threshold. More precisely, let $\hat{\sigma} = \{\hat{\sigma}(v)\}_{v \in V}$ be estimators for the spins of the vertices. Following [3], we say that $\hat{\sigma}$ has positive overlap with the true spin configuration $\sigma = \{\sigma(v)\}_{v \in V}$ if for some $\delta > 0$, with high probability,
\[
\min_p \frac{1}{n} \sum_{v=1}^n 1_{\hat{\sigma}(v) = p \sigma(v)} > \frac{1}{2} + \delta,
\]
where $p$ runs over the identity mapping on $\{+, -\}$ and the permutation that swaps $+$ and $-$. 

**Theorem 2.2** (Degree-Corrected Extension of Theorem 5 in [2]). Let $G$ be drawn according to the DC-SBM such that assumption (2.1) holds and such that $\mu_2 > \rho$. Let $\xi_2$ 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 $\hat{\sigma}(v) = +$ if $\sum_{e: e_2=v} \xi_2(e) > \sqrt{\rho}$ and put $\hat{\sigma}(v) = -$ otherwise.
2.1 Notation

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

We denote by \(\| \cdot \|\) both the euclidean norm for vectors and the operator norm of matrices. I.e., for vectors \(x = (x_1, \ldots, x_m)\), 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
\[
C_{\text{coupling}} := \left(\frac{1}{3} - \frac{1}{9} \log(4/e)\right) \land \left(\frac{1}{80} \land \frac{\gamma_4}{\log(\rho(2(a + b)\sigma_{\text{max}}))}\right)
\]
(2.8)

We put,
\[
C_{\min} = \frac{1}{10} C_{\text{coupling}}
\]
(2.9)
and consider often neighbourhoods of radius \(C_{\min} \log \rho(n)\).

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

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 \(c_1, c_2, \ldots\) for constants, without explicitly mentioning.

3 Preliminaries

3.1 Background on non-backtracking matrix

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

Firstly, for any \(k \geq 1\), \(B_{e,f}^k\) 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 \((B^*)_{e,f} = B_{f,e} = B_{e^{-1},f^{-1}}\), where for oriented edge \(e = (e_1, e_2)\), we set \(e^{-1} = (e_2, e_1)\). If we introduce the swap notation, for \(x \in \mathbb{R}^\vec{E}\),
\[
\hat{x}_e = x_{e^{-1}}, \quad e \in \vec{E},
\]
then for any \(x, y \in \mathbb{R}^\vec{E}\), and integer \(k \geq 0\),
\[
\langle y, B^k x \rangle = \langle B^k \hat{y}, \hat{x} \rangle.
\]

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, \(P x = \hat{x}, P^* = P\) 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,

$$B^k = (B^k P) = \sum_j \sigma_{k,j} x_{k,j}^* x_{k,j} P = \sum_j \sigma_{k,j} x_{k,j}^* \tilde{x}_{k,j} = \sum_j s_{k,j} x_{k,j} y_{k,j},$$  \hspace{1cm} (3.1)

where $s_{k,j} = |\sigma_{k,j}|$ and $y_{k,j} = \text{sign}(\sigma_{k,j}) \tilde{x}_{k,j}$. Since $P$ is an orthogonal matrix, $(\tilde{x}_{k,j})_j$ form an orthonormal base for $\mathbb{R}^F$ and the term furthest on the right of (3.1) is thus the spectral value decomposition of $B^k$. Now, if $B$ is irreducible and if $\xi$ 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 3.1** (Special case of Proposition 8 in [2]). Let $\ell = C \log_p n$, with $C > 0$. Let $A \in M_n(\mathbb{R})$, such that for some vectors $x_1 = x_{1,1}, y_1 = y_{1,1}, x_2 = x_{1,2}, y_2 = y_{1,2} \in \mathbb{R}$ and some matrix $R_\ell \in M_n(\mathbb{R})$,

$$A^\ell = \rho^\ell x_1 y_1^* + \mu^\ell x_2 y_2^* + R_\ell. \hspace{1cm} (3.2)$$

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

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

Further, there exist unit eigenvectors $\psi_1, \psi_2$ of $A$ with eigenvalues $\lambda_1$, respectively $\lambda_2$ such that

$$\|\psi_1 - x_1\|_1 = o(1).$$

**Proof.** This is a special case of Proposition 8 in [2]. In the notation of the latter, we have $\ell' = \ell - 2$, $\theta_1 = \rho$, $\theta_2 = \mu_2$, $\gamma = \mu_2 \geq \frac{\alpha + \beta}{|\alpha - \beta|} > 1$. Further $\frac{c_0(c_0 x_0^k - c_1)}{4c_1} \leq \frac{c_0^2}{2(\sqrt{\rho})c_1} < 1$, and thus

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

We thus need to find candidate vectors $x_1, x_2, y_1$ and $y_2$ that meet the conditions in Proposition 3.1 and further verify that the remainder $R_\ell$ has small norm. Note that the last condition is true whenever $\|B^k x\| \leq \rho^{\ell/2} \log^c(n)$ for all normalized $x$ in span$\{y_1, y_2\}^\perp$. 


4 Proof of Theorem 2.1

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

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

and show that the assumptions of Proposition 3.1 are met.

Let $\ell$ be as in Theorem 2.1 and recall $\chi_k$ and $\zeta_k$ from (2.6) and (2.7). For ease of notation, we introduce for $k \in \{1, 2\}$,

$$\varphi_k = \frac{B^\ell \chi_k}{\|B^\ell \chi_k\|} \quad \text{and} \quad \theta_k = \|B^\ell \varphi_k\|. \quad (4.1)$$

Then, $\zeta_k = \frac{B^\ell \varphi_k}{\theta_k}$.

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 4.1** (Degree-Corrected Extension of Proposition 19 in [2]). Let $\ell = C \log_p n$ with $0 < C < C_{\min}$. For some $b, c > 0$, with high probability,

(i) $b \mu_k^2 \leq \theta_k \leq c \mu_k^2$ if $k \in \{1, 2\}$,

(ii) $\text{sign}(\mu_k^\ell)(\zeta_k, \varphi_k) \geq b$ if $k \in \{1, 2\}$,

(iii) $|\langle \zeta_1, \varphi_2 \rangle| \leq (\log n)^3 n C (\frac{2}{\rho^4} + \frac{1}{n})$,

(iv) $|\langle \zeta_1, \varphi_2 \rangle| \leq (\log n)^3 n C (\frac{2}{\rho^4} + \frac{1}{n})$ if $n \neq 1, 2$.

(v) $|\langle \zeta_1, \zeta_2 \rangle| \leq (\log n)^3 n C (\frac{2}{\rho^4} + \frac{1}{n})$.

Put $H = \text{span}(\varphi_1, \varphi_2)$, then

**Proposition 4.2** (Degree-Corrected Extension of Proposition 20 in [2]). Let $\ell = C \log_p n$ with $0 < C < C_{\min}$. For some $c > 0$, with high probability,

$$\sup_{x \in H^\perp, \|x\| = 1} \|B^\ell x\| \leq (\log n)^c \rho^{\ell / 2}. \quad (4.2)$$

Put $\varphi_1 = \varphi_2$, and $\varphi_2 = \frac{\varphi_2 - \langle \varphi_1, \varphi_2 \rangle \varphi_1}{\|\varphi_2 - \langle \varphi_1, \varphi_2 \rangle \varphi_1\|}$, then $\varphi_1$ and $\varphi_2$ are orthonormal and $|\|\varphi_2 - \varphi_2\| = o(\rho^{\ell / 2})$, due to Proposition 4.1 (iii).

Let $\bar{\zeta}_1$ be the normalized orthogonal projection of $\zeta_1$ on $\text{span}\{\varphi_2\}^\perp$. Similarly, let $\bar{\zeta}_2$ be the normalized orthogonal projection of $\zeta_2$ on $\text{span}\{\zeta_1, \varphi_2\}^\perp$.

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 4.1 (iv) and (v).

We set

$$D = \theta_1 \bar{\zeta}_1 \varphi_1^* + \theta_2 \bar{\zeta}_2 \varphi_2^* = \rho \left( \frac{\theta_1}{\rho^\ell \bar{\zeta}_1} \right) \bar{\zeta}_1 \varphi_1^* + \rho \left( \frac{\theta_2}{\rho^\ell \bar{\zeta}_2} \right) \bar{\zeta}_2 \varphi_2^*.$$

Note that,

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

and

$$\|B^\ell \bar{\zeta}_2\| = \|B^\ell ((1 + o(1))\varphi_2 + o(1)\varphi_1)\| = O(\rho^\ell).$$

As a consequence, from Proposition 4.2,

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

Since $D \varphi_1 = B^\ell \varphi_1 + \theta_i (\zeta_i - \bar{\zeta}_i)$,

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

8
Let $P$ be the orthogonal projection on $H = \text{span}\{\vec{\varphi}_1, \vec{\varphi}_2\} = \text{span}\{\vec{\varphi}_1, \vec{\varphi}_2\}$, then 
\[ \|B^t P - D\| = O\left(\rho^{t/2}\right). \]

Put $R_t = B^t - D$. Write for $y \in \mathbb{R}^E$ with unit norm, $y = h + h^\perp$, with $h \in H$ and $h^\perp \in H^\perp$, then
\[ \|R_t y\| = \|B^t h^\perp + (B^t - D)h\| \leq \sup_{x \in H^\perp, \|x\| = 1} \|B^t x\| + \|B^t P - D\| \quad (4.3) \]
\[ = O\left(\log^c(n)\rho^{t/2}\right), \]
as follows from Proposition 4.2.

We finish by applying Proposition 3.1 with $x_1 = \frac{\mu_1}{\rho} \xi_1$, $y_1 = \vec{\varphi}_1$, $x_2 = \frac{\mu_2}{\rho} \xi_2$, and, $y_2 = \vec{\varphi}_2$.

In case $\mu_2^2 \leq \rho$, Proposition 4.1 (i) and (ii) continue to hold for $k = 1$. Further Proposition 4.2 holds with $H$ redefined as $H = \{\vec{\varphi}_1\}$. Redefine $R_t$ as $R_t = B^t - \rho^t x_1 y_1$, with again $x_1 = \frac{\mu_1}{\rho} \xi_1$ and $y_1 = \vec{\varphi}_1$. Apply Proposition 7 from [2].

## 5 Poisson-mixture two-type branching processes

### 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$\left(\frac{2}{5} \Phi^{(1)}(\phi_o)\right)$ particles of spin $\sigma_o$ and Poi$\left(\frac{2}{5} \Phi^{(1)}(\phi_o)\right)$ particles of spin $-\sigma_o$. Further, the weights of those particles are i.i.d. distributed following law $\nu^*$, the size-biased version of $\nu$, defined for $x \in [\phi_{\min}, \phi_{\max}]$ by
\[ \nu^*(x) = \frac{1}{\Phi^{(1)}} \int_{\phi_{\min}}^x y \nu(y). \quad (5.1) \]

For generation $t \geq 1$, a particle with spin $\sigma$ and weight $\phi^*$ is replaced in the next generation by Poi$\left(\frac{2}{5} \Phi^{(1)}(\phi^*)\right)$ particles of the same spin and Poi$\left(\frac{2}{5} \Phi^{(1)}(\phi^*)\right)$ particles of the opposite sign. Again, the weights of the particles in generation $t + 1$ follow in an i.i.d. fashion the law $\nu^*$. The offspring-size of an individual is thus a Poisson-mixture.

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

We associate two matrices to the branching process, namely $M$ defined in (2.2), and, for a root with weight $\phi_o$,
\[ M_{\phi_o} = \frac{\Phi^{(1)}(\phi_o)}{\Phi^{(2)}} M. \quad (5.2) \]

Then, $M$ is the transition matrix for generations $t \geq 1$ and later:
\[ \mathbb{E}[Z_{t+1}|Z_t] = MZ_t, \quad \text{for all} \ t \geq 1, \quad (5.3) \]
and $M_{\phi_o}$ describes the transition from the root to the first generation:
\[ \mathbb{E}[Z_1|Z_0, \phi_o] = M_{\phi_o}Z_0, \quad (5.4) \]
where, by assumption $Z_0 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. 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 \( \frac{q^2}{\mu_{j}^2} \).

Recall from (2.5) that \( g_k \) \((k = 1, 2)\) are the left-eigenvectors of \( M \) associated to eigenvalues \( \mu_k \):

\[
g_k^* M = \mu_k g_k^*, \quad k = 1, 2. \tag{5.5}
\]

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

\[
\mu_k, \phi_o = \frac{\Phi^{(1)}(\phi_o)}{\Phi^{(2)}(\phi_o)} \mu_k, \quad k = 1, 2. \tag{5.6}
\]

Theorem 5.1 shows that a Kesten-Stigum theorem applies to the "classical" branching process obtained after restricting the above process to generations 1 and later. Corollary 5.2, then, joins this classical branching process to the transition from the root to generation 1.

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

\[
\Psi_t(\pm) = \sum_{u \in Y_t} 1_{\sigma_u = \pm} \phi_u, \tag{5.7}
\]

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

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

**Theorem 5.1 (Degree-Corrected Extension of Theorem 21 in [2]).** Put \( \mathcal{F}_t = \{ Z_s \}_{s \leq t} \).

For any \( k = 1, 2 \),

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

is an \( \mathcal{F}_t \)-martingale a.s. and in \( L^2 \) such that for some \( C > 0 \) and all \( t \geq 1 \),

\[
\mathbb{E} [X_k(t)] = 0 \quad \text{and} \quad \mathbb{E} [X_k^2(t) | Z_1] \leq C\|Z_1\|_1.
\]

**Proof.** For \( 1 \leq q < t \), we have

\[
Z_t - M^{t-s} Z_s = \sum_{u=s}^{t-1} M^{t-u-1} (Z_{u+1} - M Z_u),
\]

consequently, \( g_k^* M = \mu_k g_k^* \),

\[
\frac{\langle g_k, Z_t \rangle}{\mu_k^{t-1}} = \frac{\langle g_k, Z_q \rangle}{\mu_k^{q-1}} + \sum_{u=q}^{t-1} \frac{\langle g_k, Z_{u+1} - M Z_u \rangle}{\mu_k^u}, \tag{5.8}
\]

compare to (55) in [2]. Hence, \( (X_k(t))_{t \geq 1} \) is an \( \mathcal{F}_t \)-martingale with mean 0. We shall invoke Doob's martingale convergence theorem to prove the assertion. That is, we shall show that for some \( C > 0 \) and all \( t \geq 1 \),

\[
\mathbb{E} [X_k^2(t) | Z_1] \leq C\|Z_1\|_1.
\]

Let, for \( i, j \in \{+, -\}, Z_{s+1}(i, j) \) denote the number of type \( i \) individuals in generation \( s + 1 \) which descend from from a type \( j \) particle in the \( s \)-th generation. Then,

\[
\mathbb{E} [\|Z_{s+1} - M Z_s\|_2^2 | Z_s] = \sum_{i, j \in \{+, -\}} \mathbb{E} [(Z_{s+1}(i, j) - M_{ij} Z_s(j))^2 | Z_s(j)]. \tag{5.9}
\]
We calculate first, for some integer $z \geq 0$,
\[
E \left[ (Z_{s+1}(i, j) - M_{i,j}Z_s(j))^2 \mid Z_s(j) = z \right] = E \left[ \left( \sum_{t=1}^{z} (Y_t(i, j) - M_{i,j}) \right)^2 \mid Z_s(j) = z \right]
\]
\[
= \sum_{t=1}^{z} E \left[ (Y_t(i, j) - M_{i,j})^2 \right],
\]  
(5.10)

where $(Y_t(i, j))_{t=1}^{z}$ are i.i.d. copies of Pois\(\left(1+\frac{k}{2}x\phi^{(1)}(\phi^{*})\right)\), where $\phi^{*}$ follows the biased law $\nu^{*}$.

Put $c_1 = \max_{i,j \in \{+, -\}} E \left[ (Y_t(i, j) - M_{i,j})^2 \right] < \infty$. Then, plugging (5.10) into (5.9), we obtain
\[
E \left[ \|Z_{s+1} - MZ_s\|_2^2 \mid Z_s \right] \leq 2c_1 \|Z_s\|_1.
\]
Consequently,
\[
E \left[ \|Z_{s+1} - MZ_s\|_2^2 \mid Z_1 \right] = E \left[ E \left[ \|Z_{s+1} - MZ_s\|_2^2 \mid Z_s \right] \mid Z_1 \right]
\]
\[
\leq 2c_1 E \left[ \|Z_s\|_1 \mid Z_1 \right]
\]
\[
= 2c_1 \rho^{s-1} \|Z_1\|_1.
\]  
(5.11)

Combining the above with (5.8) for $q = 1$, we obtain
\[
E \left[ X_s^2(t) \mid Z_1 \right] = \sum_{s=1}^{t-1} E \left[ \left( g_k, (Z_{s+1} - MZ_s)^2 \right) \mid Z_1 \right]
\]
\[
\leq \|g_k\|_2^2 \sum_{s=1}^{t-1} E \left[ \|Z_{s+1} - MZ_s\|_2^2 \mid Z_1 \right]
\]
\[
\leq 2c_1 \|g_k\|_2^2 \sum_{s=1}^{\infty} \rho^{s-1} \|Z_1\|_1.
\]  
(5.12)

The assertion now follows upon noting that
\[
C := 2c_1 \max_{k \in \{+, -\}} \|g_k\|_2^2 \sum_{s=1}^{\infty} \rho^{s-1} \frac{\mu_k^2}{\mu_k^2} < \infty,
\]
since $\rho < \mu_k^2$.

**Corollary 5.2.** For $k = 1, 2$, with the weight $\phi_0 = \psi_0$ of the root fixed, the sequence of random variables $(Y_{k, \psi_0}(t))_{t \geq 1}$ converges almost surely and in $L^2$ to a random variable $Y_{k, \psi_0}(\infty)$ with $E [Y_{k, \psi_0}(\infty) | \sigma_0] = g_k(\sigma_0)$. Further, this convergence takes place uniformly over all $\psi_0$.

**Proof.** From Theorem 5.1 we know that there exists a random variable $X_k(\infty)$ such that
\[
X_k(t) := \frac{\langle g_k, Z_t \rangle}{\mu_k^2} - \frac{\langle g_k, Z_1 \rangle}{\mu_k^2},
\]
as $t \to \infty$. Now,
\[
\langle g_k, Z_1 \rangle = \mu_k \phi_0 \langle g_k, Z_0 \rangle + \langle g_k, Z_1 - M\psi_0, Z_0 \rangle.
\]

We combine this with the definition of $X_k(t)$ to obtain
\[
\frac{\langle g_k, Z_t \rangle}{\mu_k^2} = \langle g_k, Z_0 \rangle + \frac{\langle g_k, Z_1 - M\psi_0, Z_0 \rangle}{\mu_k \phi_0} + \frac{X_k(t)}{\mu_k \phi_0}.
\]
where the right hand side is seen to converge in both senses to the random variable

\[ Y_{k,\psi_0}(\infty) = \langle g_k, Z_0 \rangle + \frac{\langle g_k, Z_1 - M_{\psi_0} Z_0 \rangle}{\mu_{k,\psi_0}} + X_k(\infty). \]

Indeed,

\[ \left| \frac{\langle g_k, Z_0 \rangle}{\mu_{k,\psi_0}} - Y_{k,\psi_0}(\infty) \right| \leq \frac{1}{\mu_{k,\phi_{\min}}} |X_k(t) - X_k(\infty)|, \]

for all \( \psi_0 \).

**Theorem 5.3.** Put \( \mathcal{G}_t = \{\Psi_s\}_{s \leq t} \). For any \( k = 1, 2 \),

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

is an \( \mathcal{G}_t \)-martingale converging a.s. and in \( L^2 \) such that for some \( C > 0 \) and all \( t \geq 1 \),

\[ \mathbb{E}[X_k(t)] = 0 \quad \text{and} \quad \mathbb{E}[X_k^2(t) | Z_1] \leq C\|Z_1\|. \]

**Proof.** For \( 1 \leq q < t \), we have again

\[ \frac{\langle g_k, \Psi_1 \rangle}{\mu_k^{-1}} = \frac{\langle g_k, \Psi_q \rangle}{\mu_k^{-1}} + \sum_{s=q}^{t-1} \frac{\langle g_k, \Psi_{s+1} - M\Psi_u \rangle}{\mu_k} \tag{5.13} \]

Since \( \mathbb{E}[\Psi_{s+1} | \Psi_u] = M\Psi_u \), \( (X_k(t))_{t \geq 1} \) is an \( \mathcal{G}_t \)-martingale with mean 0. We show again that for some \( C > 0 \) and all \( t \geq 1 \),

\[ \mathbb{E}[X_k^2(t) | Z_1] \leq C\|Z_1\|. \]

Let, for \( i, j \in \{+,-\} \), \( \Psi_{s+1}(i, j) \) denote the sum over the weights of type \( i \) individuals in generation \( s + 1 \) which descend from a type \( j \) particle in the \( s \)-th generation. Then,

\[ \mathbb{E} [\|\Psi_{s+1} - M\Psi_s\|_2^2 | Z_s] = \sum_{i,j \in \{+,-\}} \mathbb{E} [(\Psi_{s+1}(i, j) - M_{ij}\Psi_s(j))^2 | Z_s(j)]. \tag{5.14} \]

We calculate first, for some integer \( z \geq 0 \),

\[ \mathbb{E} [\left( \sum_{l'=1}^{Y_{i}(j)} \phi_{ll'} - M_{ij}\phi_i^j \right)^2 | Z_s(j) = z] = \mathbb{E} \left[ \left( \sum_{l=1}^{z} \left( \sum_{l'=1}^{Y_{i}(j)} \phi_{ll'} - M_{ij}\phi_i^j \right) \right)^2 | Z_s(j) = z \right] \tag{5.15} \]

where \( \phi_{ll'} \) and \( \phi_i^j \) are all independent and governed by the biased law \( \nu^* \), and where \( (Y_{i}(j))_{i=1}^{Z_s(j)} \) are i.i.d. copies of \( \text{Poi}\left( \frac{1-\nu^*+1/2}{2}\Phi^{(1)}(1) \phi_i^j \right) \), with \( \phi^* \) governed by \( \nu^* \). Thus the summands indexed by \( l \) are independent. We have

\[ \mathbb{E} \left[ \sum_{l'=1}^{Y_{i}(j)} \phi_{ll'} - M_{ij}\phi_i^j \middle| Z_s(j) \right] = \frac{1-\nu^*+1/2}{2}\Phi^{(1)}(1) \phi_i^j - M_{ij}\phi_i^j = 0 \]

Therefore,

\[ \mathbb{E} [\left( \sum_{l'=1}^{Y_{i}(j)} \phi_{ll'} - M_{ij}\phi_i^j \right)^2 | Z_s(j) = z] = \sum_{l=1}^{z} \mathbb{E} \left[ \left( \sum_{l'=1}^{Y_{i}(j)} \phi_{ll'} - M_{ij}\phi_i^j \right)^2 \right]. \tag{5.16} \]

Put \( c_1 = \max_{i,j \in \{+,-\}} \mathbb{E} \left[ \left( \sum_{l'=1}^{Y_{i}(j)} \phi_{ll'} - M_{ij}\phi_i^j \right)^2 \right] < \infty \). Then, plugging (5.16) into (5.14), we obtain

\[ \mathbb{E} [\|\Psi_{s+1} - M\Psi_s\|_2^2 | Z_s] \leq 2c_1\|Z_s\|. \]
5.2 Quantitative version of the Kesten-Stigum theorem

We now quantify the growth of the population size. The latter is defined as
\[ S_t = \|Z_t\|_1, \quad t \geq 0, \]
i.e., the number of individuals in generation \( t \geq 0 \). Given \( S_t \), for \( t \geq 1 \) we have
\[ S_{t+1} = \text{Poi} \left( \sum_{l=1}^{S_t} X_l^{(l)} \right), \tag{5.17} \]
where \( \left( X_l^{(l)} \right) \) are i.i.d. copies of \( \frac{a+b}{2}\Phi^{(1)}\phi^* \), where \( \phi^* \) follows law \( \nu^* \).

Note that in the ordinary Stochastic Block Model (i.e., when all vertices have unit weight), the argument of the Poisson random variables in (5.17) is deterministic, contrary to the general case under consideration here. Using (5.3) recursively in conjunction with (5.4), it follows that
\[ \mathbb{E}[S_t|\phi_0] = \Phi^{(1)}\phi_0 \Phi^{(2)}\rho_t, \quad \forall t \geq 1. \]

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 5.17.

**Lemma 5.4** (Degree-Corrected Extension of Lemma 23 in [2]). Assume \( S_0 = 1 \). There exist \( c, c' > 0 \) such that for all \( s \geq 0 \),
\[ \mathbb{P} \left( \forall k \geq 1, S_k \leq s\rho^k \right) \geq 1 - c'e^{-cs}. \]

**Proof.** For \( k \geq 1 \), put
\[ \epsilon_k = \rho^{-k/2}\sqrt{k} \quad \text{and} \quad f_k = \prod_{\ell=1}^{k}(1+\epsilon_{\ell}). \]

Due to convergence of \( (f_k) \), there exist constants \( c_0, c_1 > 0 \) such that for all \( k \geq 1 \),
\[ c_0 \leq f_k \leq c_1 \quad \text{and} \quad \epsilon_k \leq c_1, \quad (5.18) \]

exactly as (57) in [2].

Recall the law of \( S_{k+1} \) from (5.20). We shall firstly derive a concentration result for \( \sum_{l=1}^{S_k} X_k^{(l)} \), by using Hoeffding’s inequality. Note that by definition \( X_k^{(l)} \in \frac{a+b}{2}\Phi^{(1)}[\phi_{\min}, \phi_{\max}] \). Put \( \gamma = \left( \frac{a+b}{2}\Phi^{(1)} \right)^2(\phi_{\max} - \phi_{\min})^2 \), then Hoeffding’s equality reads
\[ \mathbb{P} \left( \left\| \sum_{l=1}^{n} X_k^{(l)} - n\rho \right\| \geq t \right) \leq 2 \exp \left( -\frac{2t^2}{n\gamma} \right). \]

Hence, in particular,
\[ \mathbb{P} \left( \sum_{l=1}^{sf_k^k} X_k^{(l)} - sf_k^k \rho \geq sf_k^k \rho \frac{\epsilon_{k+1}}{2} \right) \leq 2 \exp \left( -\frac{sf_k^k(k+1)}{2\gamma} \right) \leq 2 \exp \left( -\frac{sf_k^k(k+1)}{2\gamma} \right), \quad (5.19) \]
for some $c_2 > 0$, due to (5.18). We use the last result to obtain

$$
\mathbb{P}\left(S_{k+1} > s_f k^{k+1} | S_k \leq s_f k^k\right) \leq \mathbb{P}\left(\text{Poi}\left(\frac{s_f k^k}{\sum_{i=1}^k X_i^{(1)}}\right) > s_f k^{k+1}\right)
$$

$$
\leq \mathbb{P}\left(\text{Poi}\left(s_f k^{k+1} \left(1 + \frac{\epsilon_{k+1}}{2}\right)\right) > s_f k^{k+1} \left(1 - 2e^{-c_2s}\right) + 2e^{-c_2s}\right).
$$

We bound

$$
s_f k^{k+1} = s_f k^{k+1} \left(1 + \frac{\epsilon_{k+1}}{2}\right) + \frac{\epsilon_{k+1}}{2}
$$

where $c_3 = \frac{1}{2 \max(1, c_2/2)} > 0$. Combining the last estimate with (5.20) and the inequality

$$
\mathbb{P}(\text{Poi}(\lambda) \geq \lambda s) \leq e^{-\lambda I(s)},
$$

where

$$
I : x \mapsto \left\{ \begin{array}{ll}
\log x - x + 1 & \text{if } x > 0; \\
\infty & \text{if } x \leq 0,
\end{array} \right.
$$

entails that

$$
\mathbb{P}\left(S_{k+1} > s_f k^{k+1} | S_k \leq s_f k^k\right) \leq \exp\left(-s_f k^{k+1} \left(1 + \frac{\epsilon_{k+1}}{2}\right) I(1 + c_3\epsilon_{k+1})\right) + 2e^{-c_2s}.
$$

It remains to bound $I(1 + c_3\epsilon_k)$ from below. But, due to the form of $I$, there exists a $\theta > 0$ such that for $x \in [0, c_3\max_k \epsilon_k]$, $I(1 + x) \geq \theta_x^2$. Consequently

$$
\mathbb{P}\left(S_{k+1} > s_f k^{k+1} | S_k \leq s_f k^k\right) \leq 3e^{-c_4s},
$$

for some constant $c_4 > 0$. Hence,

$$
\mathbb{P}\left(\exists k : S_k > sc_1 s^k\right) \leq \sum_{k=1}^{\infty} 3e^{-c_4s} = \frac{3}{1 - e^{-c_4s}} e^{-c_4s},
$$

from which the statement follows.

From Theorem 5.1 and Corollary 5.2, we know that the different components (expressed in the basis of eigenvectors of $M$) grow exponentially with rate $\rho$, respectively $\mu_2$. We now quantify the error. Recall $\Psi_t$ from (5.7).

**Theorem 5.5** (Degree-Corrected Extension of Theorem 24 in [2]). Let $\beta > 0$, $Z_0 = \delta_x$ and $\phi_0 = \psi_0$ 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 $0 \leq s < t$,

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

and,

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

**Proof.** We claim that there exist constants $c, c’ > 0$ such that for any $s \geq 0$

$$
\mathbb{P}\left(\|Z_{t+1} - MZ_t\|_2 > s\|Z_t\|_1^{1/2} \bigg| \mathcal{F}_s\right) \leq c’ e^{-c(s^2 + s^2)}.
$$

(5.22)

To prove (5.22), we shall employ Hoeffding’s inequality to establish a concentration result for

$$
\lambda^+ = \frac{\phi^{(1)}}{2} \left( a \sum_{i=1}^t \Phi_i^+ + b \sum_{i=1}^t \Phi_i^- \right),
$$

(5.23)
and,

\[ \lambda^+ = \frac{\Phi^{(1)}_1}{2} \left( b \sum_{i=1}^{Z^+_t} \Phi^+_i + a \sum_{i=1}^{Z^-_t} \Phi^-_i \right) \]  \hfill (5.24)

around their respective means \( y^+ = \mathbb{E}_+ [\lambda^+] \) and \( y^- = \mathbb{E}_- [\lambda^-] \), where \((\Phi^+_i)_i\) are i.i.d. random variables with law \( \nu^+ \), and where \( \mathbb{E}_+ [\cdot] = \mathbb{E} [\cdot | Z_t] \). This in conjunction with the classical tail bound for \( Y \overset{d}{=} \text{Pois}(\lambda) \):

\[ \mathbb{P} (|Y - \lambda| > \lambda s) \leq 2 e^{-\lambda^s(s)} \]  \hfill (5.25)

where \( \delta : x \mapsto I(1-x) \land I(1+x) \), with \( I \) defined in (5.21), shall allow us to prove concentration of \( \left( \frac{Z^+_t}{Z^-_t} \right) = \left( \text{Pois} (\lambda^+) \text{Pois} (\lambda^-) \right) \) around \( \mathbb{E}_+ \left( \left( \frac{Z^+_t}{Z^-_t} \right) \right) = \left( \frac{y^+}{y^-} \right) = M Z_t \).

Let \( t^+, t^- > 0 \). Then, Hoeffding’s inequality gives

\[ \mathbb{P}_+ \left( \left| \sum_{i=1}^{Z^+_t} \Phi^+_i - Z^+_t \frac{\Phi^{(2)}_1}{\Phi^{(1)}_1} \right| \geq t^+ \right) \leq 2 \exp \left( -\frac{2(t^+)^2}{Z^+_t} \right), \]  \hfill (5.26)

where \( \gamma = (\phi_{\text{min}} - \phi_{\text{max}})^2 \), and where \( \mathbb{P}_+ (\cdot) = \mathbb{P} (\cdot | Z_t) \).

Hence,

\[ \mathbb{P}_+ \left( |\lambda^+ - y^+| \leq \frac{\Phi^{(1)}_1}{2} (at^+ + bt^-) \right) \geq \mathbb{P}_+ \left( \left| \sum_{i=1}^{Z^+_t} \Phi^+_i - Z^+_t \frac{\Phi^{(2)}_1}{\Phi^{(1)}_1} \right| \leq t^+ \right) \]

\[ \geq \left( 1 - 2 \exp \left( -\frac{2(t^+)^2}{Z^+_t} \right) \right) \left( 1 - 2 \exp \left( -\frac{2(t^-)^2}{Z^-_t} \right) \right). \]

Plugging \( t^+ = \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \) and \( t^- = \frac{\sqrt{y^-}}{\sqrt{\Phi^{(1)}_1}} \) into the last equation leads to

\[ \mathbb{P}_+ \left( |\lambda^+ - y^+| \leq \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \frac{1}{2} \right) \geq \left( 1 - 2 \exp \left( -\frac{4}{\Phi^{(1)}_1} \frac{y^+}{\sqrt{\Phi^{(1)}_1}} \frac{1}{2} s^2 \right) \right) \left( 1 - 2 \exp \left( -\frac{4}{\Phi^{(1)}_1} \frac{y^-}{\sqrt{\Phi^{(1)}_1}} \frac{1}{2} s^2 \right) \right) \]

\[ \geq \left( 1 - 2 e^{-c_0 s^2} \right)^2 \]

\[ \geq 1 - 4 e^{-c_0 s^2}, \]  \hfill (5.28)

for some constant \( c_0 > 0 \), since \( \frac{y^+}{2} \) is bounded away from zero by some constant.

We use the last inequality to obtain

\[ \mathbb{P}_+ \left( Z^+_{t+1} - y^+ > s \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right) \leq \mathbb{P}_+ \left( \text{Pois} \left( \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right) - \left( \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right) > \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right) + 4 e^{-c_0 s^2}. \]

We continue by invoking (5.25),

\[ \mathbb{P}_+ \left( \text{Pois} \left( \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right) - \left( \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right) > \delta \left( \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right) \right) \leq 2 \exp \left( -(y^+ + \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}}) \delta \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right). \]

We note the existence of a \( \theta > 0 \) such that for all \( x \in [0,1], \delta(x) \geq \theta x^2 \), so that

\[ (y^+ + \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}}) \delta \left( \frac{s}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \right) \geq \theta \frac{s^2}{2} \frac{\sqrt{y^+}}{\sqrt{\Phi^{(1)}_1}} \geq c_2 (s^2 \land s), \]
for some constant $c_2 > 0$, because $y^+ + \frac{s}{2}\|y\|_1^{1/2} \leq \max\{2y^+, s\|y\|_1^{1/2}\}$.

Similarly, to bound $\mathbb{P}_*\left(Z^+_{t+1} - y^+ \geq -s\|y\|_1^{1/2}\right)$ from above, we need to estimate

$$\mathbb{P}_*\left(\text{Poi}\left(y^+ - \frac{s}{2}\|y\|_1^{1/2}\right) - \left(y^+ - \frac{s}{2}\|y\|_1^{1/2}\right) \leq -\frac{s}{2}\|y\|_1^{1/2}\right) \leq 2\exp\left(-\left(y^+ - \frac{s}{2}\|y\|_1^{1/2}\right)\delta\left(\frac{s}{2}\|y\|_1^{1/2}\right)\right),$$

(5.30)

when $y^+ > \frac{s}{2}\|y\|_1^{1/2}$ (if $y^+ \leq \frac{s}{2}\|y\|_1^{1/2}$, then $Z^+_{t+1} - y^+ \geq \frac{s}{2}\|y\|_1^{1/2}$, so that $\mathbb{P}_*\left(Z^+_{t+1} - y^+ \leq -s\|y\|_1^{1/2}\right) = 0$).

We distinguish between two cases: Firstly, when $y^+ - \frac{s}{2}\|y\|_1^{1/2} > \frac{s}{2}\|y\|_1^{1/2}$, we have

$$\left(y^+ - \frac{s}{2}\|y\|_1^{1/2}\right) \delta\left(\frac{s}{2}\|y\|_1^{1/2}\right) \geq \frac{\theta}{2}\|y\|_1^{1/2} \geq \frac{\theta}{y^+ - \frac{s}{2}\|y\|_1^{1/2}} \geq \frac{\theta}{s^2/4} \geq c_4s^2,$$

(5.31)

for some constant $c_4$.

Combining (5.29) - (5.32), leads to

$$\mathbb{P}\left(|Z_{t+1}^+ - y^+| > s\|y\|_1^{1/2}\right) \leq 2\left(e^{-c_2(s^2\wedge s)} + e^{-c_4s} + e^{-c_4s^2} + 8e^{-c_0s^2}\right) \leq c_5e^{-c_0(s^2\wedge s)}.$$

(5.33)

An identical bound holds (after possibly redefining the values of $c_5$ and $c_6$ for $|Z_{t+1}^+ - y^-|$.

Finally, noting that $\|y\|_1 = \rho\|Z_t\|_2$, we have

$$\mathbb{P}\left(\|Z_{t+1} - MZ_t\|_2 > s\|Z_t\|_1^{1/2} \mid \mathcal{F}_t\right) \leq \mathbb{P}\left(|Z_{t+1}^+ - y^+| \geq \frac{s}{\sqrt{2}}\|Z_t\|_1^{1/2} \mid \mathcal{F}_t\right) + \mathbb{P}\left(|Z_{t+1}^+ - y^-| \geq \frac{s}{\sqrt{2}}\|Z_t\|_1^{1/2} \mid \mathcal{F}_t\right) \leq c'e^{-c(s^2\wedge s)},$$

(5.34)

that is exactly claim (5.22).

We are now in a position to derive a similar bound as (59) in [2]:

$$\mathbb{P}\left(\forall t \geq 1 : \|Z_{t+1} - MZ_t\|_2 \leq u(t + 1) \log n\|Z_t\|_1^{1/2}\right) \geq 1 - c''\sum_{t \geq 1} e^{-ct\log n} \geq 1 - C'n^{-C'u}.$$

(5.35)

Recalling (5.8), we have, for $s \geq 1$,

$$\left|\langle y_k, Z_s \rangle - \mu^{s-t}_k \langle y_k, Z_t \rangle\right| \leq \mu^{s-t}_k \|y_k\|_2 \sum_{u=s}^{t-1} \frac{\|Z_{u+1} - MZ_u\|_2}{\mu_k^u},$$

From Equation (5.35) we know that, for all $u \geq 1$,

$$\|Z_{u+1} - MZ_u\|_2 \leq c_0(\log n)(u + 1)\|Z_u\|_1^{1/2},$$

(5.36)

where $c_0$ is so large that 5.36 holds with probability $1 - n^{-\beta}$. Further, $\|Z_h\|_1$ itself is bounded by Lemma 5.4:

$$\|Z_h\|_1 \leq c_{10}(\log n)^{\rho^h},$$

(5.37)
also with probability at least $1 - n^{-\beta}$.

With the same probability, for $k \in \{1, 2\}$,

$$\langle g_k, Z_s \rangle - \mu_k^{s-t}(g_k, Z_s) \leq c_{11}(\log n)^{3/2} \mu_k^{s-1} \sum_{u=\hat{u}}^{t-1}(u+1)\frac{\gamma_u}{\mu_k} \leq c_{12}(\log n)^{3/2}(s+1)\rho^{s/2}.$$  \hfill (5.38)

The proof the last claim, write

$$\langle g_k, \Psi_s \rangle - \mu_k^{s-t}(g_k, \Psi_s) = \frac{\Phi(2)}{\Phi(1)}(\langle g_k, Z_s \rangle - \mu_k^{s-t}(g_k, Z_s)) + \epsilon_s - \mu_k^{s-t}\epsilon_s,$$  \hfill (5.39)

where, for $s \geq 1,$

$$\epsilon_s = g_k(+)\left(\Psi_s(+) - Z_s^+\frac{\Phi(2)}{\Phi(1)}\right) + g_k(-)\left(\Psi_s(-) - Z_s^-\frac{\Phi(2)}{\Phi(1)}\right).$$

We bound $\epsilon_s$ using (5.26),

$$\mathbb{P}\left(\forall t \geq 1 : \epsilon_t \leq t \log n\|Z_t\|_{1/2}^{1/2}\right) \geq 1 - c_{13} \sum_{t \geq 1} e^{-c_{14}t^2 \log^2 n} \geq 1 - C' n^{-C_u}.$$  \hfill (5.40)

So that, with probability $1 - n^{-\beta},$

$$|\epsilon_s - \mu_k^{s-t}\epsilon_s| \leq c_{15} \log^{5/2}(n)\left(\rho^{s/2} + |\mu_k|^{s-t}\rho^{1/2}\right) \leq c_{16} \log^{5/2}(n)\rho^{s/2},$$

since $|\mu_k| > \rho^{1/2}.$ \hfill \square

### 5.3 $B^\ell B^\tau \bar{\chi}_k$ on trees: a cross generation functional

Recall our claim that $B^\ell B^\tau \bar{\chi}_k$ 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 $B^\ell B^\tau \bar{\chi}_k$ on trees.

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

$$Q_{k,t} = \sum_{(u_0, \ldots, u_{2\ell+1}) \in \mathcal{P}_{2\ell+1}} g_k(\sigma(u_{2\ell+1}))\phi_{u_{2\ell+1}},$$  \hfill (5.41)

where $\mathcal{P}_{2\ell+1}$ is the set of paths $(u_0, \ldots, u_{2\ell+1})$ (of length $2\ell+1$) in the tree starting from $u_0 = o$ with both $(u_0, \ldots, u_{\ell})$ and $(u_{\ell}, \ldots, u_{2\ell+1})$ 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

$$Q_{1,t} = \sum_{(u_0, \ldots, u_{2\ell+1}) \in \mathcal{P}_{2\ell+1}} \frac{1}{\sqrt{2}}\phi_{u_{2\ell+1}},$$  \hfill (5.42)

the cardinality of $\mathcal{P}_{2\ell+1}$ and

$$Q_{2,t} = \sum_{(u_0, \ldots, u_{2\ell+1}) \in \mathcal{P}_{2\ell+1}} \frac{1}{\sqrt{2}}\sigma(u_{2\ell+1})\phi_{u_{2\ell+1}}.$$  \hfill (5.43)

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 $B_{\tau e}$ denotes the non-backtracking matrix defined on $\mathcal{T}'$, \hfill (5.44)

$$\left(B^\ell_{\tau e} B^\tau_{\tau e} \bar{\chi}_k\right)(e) = Q_{k,t} + g_k(\sigma(e_1))\phi_{e_1}\|Z_t\|_{1/2},$$  \hfill (5.44)
where $Q_{k,\ell}$ and $Z_{\ell}$ 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 $\phi_v \in \mathbb{S}$.

For $t \geq 0$, let $Y^o_t$ 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., $Z^t_v = \sum_{u \in V^t} 1_{\sigma(u) = \pm}$. Furthermore, let $\Psi^t_v = (\Psi^t_v, \Psi^t_v)$, with $\Psi^t_v = \sum_{u \in V^t} 1_{\sigma(u) = \pm} \phi_u$.

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 $w$ is in between $u$ and $u_{2\ell+1}$. For such fixed $v$ and $w$ in $Y^o_t$, only the children $u_{2\ell+1} \in Y^o_t$ determine the contribution of a path to (5.41), regardless of the choice of $u_{\ell+1} \in Y^o_{\ell-1}$. Hence, for such fixed $u$ and $v$, $w \in Y^o_t$ and $u_{2\ell+1}$, there are $|Y^o_{\ell-1}| = S^o_{\ell-1}$ paths giving the same contribution to (5.41):

$$Q_{k,\ell} = \sum_{t=0}^{\ell-1} \sum_{u \in Y^o_t} L^u_{k,\ell},$$

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

$$L^u_{k,\ell} = \sum_{w \in Y^o_{\ell-1}} S^o_{\ell-1} \left( \sum_{v \in V^o_{\ell-1} \setminus \{w\}} (g_k, \Psi^o_v) \right) \tag{5.45}$$

The following theorem is an extension of Theorem 25 in [2]. The important observation is that, again, for $Z_0 = \delta_o$ and $\phi_o = \psi_o$ be fixed. For $k \in \{1, 2\}$, $(Q_{k,\ell}/\mu_k^2)_{\ell}$ converges to a random variable with mean a constant times $\tau$, 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 5.6 (Degree-Corrected Extension of Theorem 25 in [2]).** Let $Z_0 = \delta_o$ and $\phi_o = \psi_o$ be fixed. For $k \in \{1, 2\}$, $(Q_{k,\ell}/\mu_k^2)$ converges in $L^2$ as $\ell$ tends to infinity to a random variable with mean $\frac{\theta(3)}{\theta(2)} \frac{\rho}{\mu_k^2} - \rho \mu_{k,\psi_o} Y_{k,\psi_o}(\infty)$. Further, the $L^2$-convergence takes place uniformly for all $\psi_o$.

**Proof.** We start by calculating the expectation and variance of $\sum_{u \in Y^o_t} L^u_{k,\ell}$ conditional on $\mathcal{F}_t$. We use this to show that, as $\ell \to \infty$, uniformly for all $\psi_o$,

$$Q_{k,\ell} = \sum_{t=0}^{\ell-1} \sum_{u \in Y^o_t} L^u_{k,\ell}. \tag{5.46}$$

The latter is reminiscent of

$$Q_{k,\ell} = \sum_{t=0}^{\ell-1} \sum_{u \in Y^o_t} L^u_{k,\ell}.$$
and we show that $\bar{Q}_{k,\ell}$ and $Q_{k,\ell}$ are in fact close in $L^2$-distance:
\[ \|\bar{Q}_{k,\ell} - Q_{k,\ell}\| = o(\mu_k^{2\ell}). \]

Consider for $t \geq 0$ and $\ell \geq t + 2$,
\begin{align*}
E_{F_t, Y^t, Y^t_u} L^t_{k,\ell} &= \sum_{(v, w) \in Y^t, \nu \neq w} E_{F_t, Y^t, Y^t_u, \nu}^{t-1} \sum_{(x, y) \in Y^t, \nu} (g_k, \Psi^t_u) \\
&= \sum_{(v, w) \in Y^t, \nu \neq w} \rho_w \rho^{t-2} \phi_w (g_k, Z_0^u) \tag{5.48}
\end{align*}
where $\rho_w = \frac{a+b}{\Phi(1)} \phi_w$, with $\phi_w$ a random variable that follows law $\nu^*$. The second equality in (5.48) follows after calculating
\[ E[\Psi_t^u | Y_t^u] = \Phi(2)^{\ell} \Phi(1) \Phi(2)^{t-2} \phi_w (g_k, Z_0^u) \]
where the factor $\frac{\Phi(1)}{\Phi(2)^{2t}}$ accounts for the fact that the "parental" vertex $v$ has deterministic type $\phi_w$ (and transitions are thus given by $M_{\phi_w} = \Phi(1) \Phi(2)^{2t} M$), whereas vertices in the later generations have i.i.d. weights (for which $M$ is the transition matrix). Now,
\begin{align*}
E_{F_t, Y^t, Y^t_u} L^t_{k,\ell} &= E_{F_t, Y^t, Y^t_u}^{t-1} E_{F_t, Y^t, Y^t_u} (g_k, Z_0^u) \\
&= \rho^{t-2} \mu_k \rho^2 \phi^*(g_k, Z_0^u) \tag{5.49}
\end{align*}
where $\phi^*$ has law $\nu^*$, $\rho^*$ is an i.i.d. copy of $\frac{a+b}{\Phi(1)} \phi^*$ and $\sigma^* = \sigma_u$ with probability $\frac{a}{a+b}$, and $\sigma^* = -\sigma_u$ with probability $\frac{b}{a+b}$ (further, $\rho^*$, $\phi^*$ and $\sigma^*$ are independent).

We thus have
\begin{align*}
E_{F_t, Y^t, Y^t_u} L^t_{k,\ell} &= \rho^{t-2} \mu_k \rho^2 \phi^*(g_k, Z_0^u) \tag{5.50}
\end{align*}
where $\rho_u = \frac{a+b}{\Phi(1)} \phi_u$ (with $\phi_u$ the weight of $u$) and for $(x, y) \in \{+, -\} \times \{+, -\}$, $c(x, y) = \frac{a}{a+b}$ if $x = y$ and $c(x, y) = \frac{b}{a+b}$ otherwise.

Now, as $g_k$ is an eigenvector of $M$ with eigenvalue $\mu_k$, we have
\[ (g_k(1) c(\sigma_u, +) + g_k(2) c(\sigma_u, -)) = \frac{2}{a+b} \frac{\mu_k}{\Phi(2)^t} (g_k, Z_0^u) = \frac{\mu_k}{\rho} (g_k, Z_0^u). \]
Together with (5.49) this gives
\[ E_{F_t, Y^t, Y^t_u} L^t_{k,\ell} = \rho^{t-2} \mu_k \rho^2 \phi^*(g_k, Z_0^u). \tag{5.51} \]
Summing over $u \in Y^t$, using the last equation yields
\begin{align*}
E_{F_t} \sum_{u \in Y^t} L^t_{k,\ell} &= E_{F_t} \sum_{u \in Y^t} E_{F_t, Y^t_u}^{t-1} E_{F_t, Y^t_u} L^t \tag{5.52}
\end{align*}
where $E_{F_t, Y^t_u}^{t-1} E_{F_t, Y^t_u} L^t = (g_k(1) c(\sigma_u, +) + g_k(2) c(\sigma_u, -)) = \frac{2}{a+b} \frac{\mu_k}{\Phi(2)^t} (g_k, Z_0^u) = \frac{\mu_k}{\rho} (g_k, Z_0^u)$.
We leave it to the reader to verify that the same inequality holds for \( l = t + 1 \).

We continue by bounding the variance of \( L^u_{k,\ell} \):

\[
\text{Var}_{\mathcal{F}_t} L^u_{k,\ell} \leq \mathbb{E}_{\mathcal{F}_t} (L^u_{k,\ell})^2 \\
= \mathbb{E}_{\mathcal{F}_t} \sum_{(v,w) \in Y^u_t, v \neq w} \sum_{(v',w') \in Y^u_t, v' \neq w'} S^v_{t-1} S^{w'}_{t-1} \langle g_k, \Psi^v \rangle \langle g_k, \Psi^{w'} \rangle \\
\leq \mathbb{E}_{\mathcal{F}_t} |Y^u_t|^2 \mathbb{E}_{\mathcal{F}_t} S^2_{t-1} E_{\infty}(g_k, \Psi^1_t)^2,
\]

where \( E_{\infty} ([\cdot]) = \max_{t' \in \{t, \ldots, \infty\}} \mathbb{E} [\phi_{\sigma_{t'}} = \phi_{\max_{t'}, \sigma_{t'}} = \tau'] \). Now, \( \mathbb{E}_{\mathcal{F}_t} |Y^u_t|^2 \leq c_0 \). From Lemma 5.4, we know that \( S_k \leq \text{Exp} (c_1 \rho^{k}) \), hence \( E_{\infty} S^2_{t-1} \leq 2c_3^2 (\rho^{t-1})^2 \). To bound \( E_{\infty}(g_k, \Psi^1_t)^2 \), recall from Theorem 5.3 that

\[
E \left[ \left( \frac{\langle \phi_k, \Psi_1 \rangle}{\mu_k^{\ell-1}} - \langle g_k, \Psi_1 \rangle \right)^2 \right] \leq C_2 \| Z_1 \|_1.
\]

Consequently, as \( E [\| Z_1 \|_1] \) is bounded,

\[
E_{\infty}(g_k, \Psi^1_t)^2 \leq c_3 \mu_k^{2t}.
\]

Returning to (5.53), we have

\[
\text{Var}_{\mathcal{F}_t} \sum_{u \in Y^u_t} L^u_{k,\ell} \leq c_4 \mu_k^{2t} \rho^{2(\ell-t)} S_t.
\]  

(5.54)

We have

\[
\bar{Q}_{k,\ell} = \sum_{t=0}^{\ell-1} \mathbb{E}_{\mathcal{F}_t} \sum_{u \in Y^u_t} L^u_{k,\ell} \\
= \rho^t \mu_k \langle g_k, Z_0 \rangle \psi_0 + \sum_{t=1}^{\ell-1} \rho^{t-t'} \mu_k^{t+1} \langle g_k, Z_t \rangle \frac{\Phi^{(3)}}{\Phi^{(2)}} \\
= \rho^t \mu_k \langle g_k, Z_0 \rangle \psi_0 + \Phi^{(3)} \frac{\Phi^{(3)}}{\Phi^{(2)}} \sum_{t=1}^{\ell-1} \rho^{t-t'} \mu_k \mu_{k,\psi_0} Y_{k,\psi_0}(t),
\]

where \( Y_{k,\psi_0}(t) \) is defined in Corollary 5.2.

We consider

\[
\frac{\bar{Q}_{k,\ell}}{\mu_k^{2\ell}} = o(1) + \frac{\Phi^{(3)}}{\Phi^{(2)}} \sum_{t=1}^{\ell-1} \left( \frac{\mu_k}{\rho} \right)^{t-\ell} \mu_{k,\psi_0} Y_{k,\psi_0}(t),
\]

(5.56)

and verify our claim (5.47). To do so, split for arbitrary fixed \( \epsilon > 0 \),

\[
\sum_{t=1}^{\ell-1} r^{t-t'} Y_k(t) = \sum_{t=1}^{T_{\epsilon}} r^{t-t'} Y_k(t) + \sum_{t=T_{\epsilon}}^{\ell-1} r^{t-t'} Y_k(t),
\]

where \( r = \frac{\mu_k}{\rho} \), \( Y_k \) is shorthand notation for \( Y_{k,\psi_0} \), and where

\[
T_{\epsilon} = \min \{ t : \forall s \geq t, |Y_k(\infty) - Y_{k}(s)| < \epsilon \}.
\]

Then,

\[
\sum_{t=1}^{T_{\epsilon}-1} r^{t-t'} Y_k(t) \leq |\sup_{t} Y_k(t)| r^{-T_{\epsilon}} T_{\epsilon}^{a_k} = 0.
\]
as $\ell \to \infty$, since $(Y_k(t))_t$ is convergent (uniformly in $\psi_o$) and hence bounded. Further,

$$\sum_{t=0}^{\ell-1} r^{t-\ell} Y_k(t) = \sum_{t=0}^{\ell-1} (Y_k(\infty) + O(\epsilon))$$

(5.57)

where the limit is taken for $\ell \to \infty$. Since $\epsilon > 0$ was arbitrary, (5.47) follows.

$L^2$-convergence follows from [2](this convergence takes place uniformly for all $\psi_o$ due to Theorem 5.1).

Further, that $\|Q_{k,\ell} - Q_{k,0}\| = o(\mu_k^2)$ can be established by following the proof in [2]. Indeed, from the latter proof we know that, for some constant $c_0$ independent of $\psi_o$,

$$\|Q_{k,\ell} - Q_{k,0}\|_2 \leq c_0 \sum_{t=0}^{\ell} \left( \text{Var}_{\mathcal{F}_t} \left( \sum_{u \in Y_k^t} L^t_k,\ell \right) \right)^{1/2} \leq c_0 \sum_{t=0}^{\ell} \mu_k^t \rho^{t-\ell} \|\sqrt{S_t}\|_2$$

(5.58)

due to the variance bound in (5.54) and Lemma 5.4.

Finally, combining the uniform bounds (5.47) and (5.58), entails that

$$\left\| \frac{Q_{k,\ell}}{\mu_k^2} - \frac{\Phi^{(3)}}{\Phi^{(2)}} \right\|_2 \to 0,$$

uniformly for all $\psi_o$. \hfill \Box

### 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 4.1) follows from this fact, see Proposition 7.3 (ii), (iii) and Proposition 7.4 (ii) below.

**Theorem 5.7** (Degree-Corrected Extension of 28 in [2]). Assume that the spin $\sigma_o$ of the root is drawn uniformly from $\{+,-\}$. Then for any $\ell \geq 0$,

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

**Proof.** Recall the explicit expressions for $Q_{1,\ell}$ and $Q_{2,\ell}$ from (5.42), respectively (5.43). Now, conditional on $\mathcal{T}$ and the weights (denoted by $T_o$), $\mathcal{P}_{2\ell+1}$ is deterministic, hence

$$\mathbb{E}[Q_{1,\ell} Q_{2,\ell} | \mathcal{T}, T_o] = Q_{1,\ell} \sum_{(u_0, \ldots, u_{2\ell+1}) \in \mathcal{P}_{2\ell+1}} \phi_{u_{2\ell+1}} \mathbb{E}[\sigma(u_{2\ell+1}) | \mathcal{T}] = 0,$$

because, $\mathbb{E}[\sigma(u) | \mathcal{T}, \sigma_o] = \left( \frac{a-b}{a+b} \right)^{|u|} \sigma_o$, for a vertex $u$ at distance $|u|$ from the root, by construction of the branching process. \hfill \Box
6 Coupling of local neighbourhood

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 \( \nu \). Detected vertices on their turn follow a law close to \( \nu^* \).

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 \( \hat{d} \), which for \( e, f \in \hat{E}(V) \) is defined as
\[
\hat{d}(e, f) = \min_\gamma \ell(\gamma)
\]
where the minimum is taken over all self-avoiding paths \( \gamma = (\gamma_0, \gamma_1, \cdots, \gamma_{\ell+1}) \) in \( G \) such that \( (\gamma_0, \gamma_1) = e, (\gamma_1, \gamma_{\ell+1}) = f \) and for all \( 1 \leq k \leq \ell + 1 \), \( \{\gamma_k, \gamma_{k+1}\} \in E \), and where for such a path \( \gamma \), \( \ell(\gamma) = \ell \). Note that \( \hat{d}(e, f) = \hat{d}(f^{-1}, e^{-1}) \), i.e., \( \hat{d} \) is not symmetric.

We introduce the vector \( Y_t(e) = (Y_t(e)(i))_{i \in \{+, -\}} \) where, for \( i \in \{+, -\} \),
\[
Y_t(e)(i) = \left\{ f \in \hat{E} : \hat{d}(e, f) = t, \sigma(f_2) = i \right\},
\]
we denote the number of vertices at oriented distance \( t \) from \( e \) by
\[
S_t(e) = \|Y_t(e)\|_1 = \left\{ f \in \hat{E} : \hat{d}(e, f) = t \right\},
\]
and we define vector \( \Psi_t(e) = (\Psi_t(e)(i))_{i \in \{+, -\}} \) where, for \( i \in \{+, -\} \),
\[
\Psi_t(e)(i) = \sum_{f_2 \in \hat{E}, \hat{d}(e, f) = t} 1_{\sigma(f_2) = i} \phi(f_2).
\]
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 \( \nu \). In the branching process starting at an edge oriented towards, say, \( o \), the root \( o \) has weight governed by \( \nu^* \). See Proposition 6.1 below.

As a corollary we obtain an analogue of Theorem 5.5 for local neighbourhoods: the components of \( \Psi_t(e) \) grow exponentially, see Corollary 6.3.

We bound the growth of \( S_t \) in Lemma 6.4. We use a coupling argument to show that the weights of the unexplored vertices and selected vertices are stochastically dominated by variables following law \( \nu \), respectively \( \nu^* \). This argument is not needed in the ordinary SBM.

Following [17], we need to verify that certain problematic structures, namely tangents, 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 $\ell$ and simply call a graph tangle-free or tangled; the $\ell$ dependence is then tacitly assumed.

Following standard arguments we establish in Lemma 6.5 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 $v \in V(G)$, where $G$ is an instance of the DC-SBM. Initially, it is the only active vertex: $A(0) = \{v\}$. All other vertices are neutral at start: $U(0) = V(G) \setminus \{v\}$. No vertex has been explored yet: $E(0) = \emptyset$.

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

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

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

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

and

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

**Proposition 6.1** (Degree-Corrected Extension of Proposition 31 in [2]). Let $\ell = C \log_{\rho}(n)$, with $C < C_{\text{coupling}}$. Let $\rho \in V$ and $e = (e_1, e_2) \in E$. 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 $v$. Similarly, let $(T', o)$ be that same branching process, when the root has spin $\sigma(e_2)$ and weight governed by $\nu'$. Then, the total variation distance between the law of $(G, v)_{\ell}$ and $(T, o)_{\ell}$ goes to zero as $n^{-\frac{1}{2} \log(4/e)}$. The same is true for the difference between the law of $(G, e)_{\ell}$ and $(T', o)_{\ell}$.

**Remark 6.2.** 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.

**Proof.** The second statement follows from the first after recalling that $(G, e)_{\ell} = (G', e_2)_{\ell}$, where $G'$ is the graph $G$ with edge $\{e_1, e_2\}$ removed. Since $e \in E, e_2$ then has a biased weight governed by $\nu'$.

In [6], we established a coupling between the branching process and the DC-SBM where the spins are drawn uniformly from $\{+, -, \}$, with error probability $n^{-\frac{1}{2} \log(4/e)}$.

Thus, we are done if we couple the neighbourhoods in the latter graph to the DC-SBM with deterministic spins under consideration here.

Now, with probability at least $1 - e^{-\Omega(n^{1/2})}$ we can couple the graphs such that at most $c_1 n^{\frac{1}{2} \log(1 - \epsilon)}$ have unequal spins (call the corresponding set of vertices $S$) and all weights are equal. Further, we may assume that the subgraphs obtained after removing $S$ are identical.

The $\ell$-neighbourhoods in both graphs are exactly the same if they are both disjoint with $S$. Conditional on $|S|$ and $|G|$, this happens with probability at least $1 - c_2 |G| |S| / n$.

From [6], we know that with probability $1 - n^{-\log(4/e)}$, $|G| < n^{\frac{1}{2} \log(1 - \epsilon)}$.

Thus, conditional on the bounds for $|S|$ and $|G|$, the neighbourhoods are the same with probability at least $1 - C_3 n^{\frac{1}{2} \log(1 - \epsilon)}$.

All together, $P((G, v)_{\ell} = (T, o)_{\ell}) \geq 1 - c_4 n^{\frac{1}{2} \log(1 - \epsilon)}$. \hfill \Box

**Corollary 6.3** (Degree-Corrected Extension of Corollary 32 in [2]). Let $\ell = C \log_{\rho}(n)$ with $0 < C < C_{\text{coupling}}$. For $e \in E(V)$, we define the event $E(e)$ that for all $0 \leq t < \ell$ and $k \in \{1, 2\}$: $|\langle g_k, \Psi(t) \rangle - \mu_k^{t-\ell} \langle g_k, \Psi(e) \rangle| \leq (\log n)^3 \rho^{t/2}$. Then, with high probability, the number of edges $e \in E$ such that $E(e)$ does not hold is at most $\log(n) n^{\frac{1}{2} \log(1 - \epsilon)}$. 

23
Proof. This proof follows the proof of Corollary 32 in [2]. Indeed (although with a slightly different probability) the graph neighbourhood \((Y_t(e))_{0 \leq t \leq T}\) and branching process \((Z_t)_{0 \leq t \leq T}\) coincide again, and moreover, the weights are equal in both processes.

\[\mathbb{P}\left(\forall t \geq 0 : S_t(w) \leq s^*_t \right) \geq 1 - ce^{-c's_t}.\]

Consequently, for any \(p \geq 1\), there exists \(c'' > 0\) such that

\[\mathbb{E}\left[\max_{v \in [n], t \geq 0} \left(\frac{S_t(v)}{\bar{p}_t}\right)^p \right] \leq c''(\log n)^p.\]

Proof. As observed in [2], the second statement follows from the first.

Adapting our paper [6], at step \(m\) in the exploration process, the weights of the vertices in \(U(m)\) are independent, and those with spin \(\tau\) have weight governed by \(\nu_l(m)\), where

\[d\nu_\nu^\langle m \rangle(\psi) = \frac{g_\nu(\psi)}{\int_{\phi_{\min}}^{\phi_{\max}} g_\nu(\psi')d\nu(\psi')}d\nu(\psi),\]

where \(g_\nu(\cdot) = \prod_{i=1}^{n} \left(1 - \frac{d(\nu, \tau)}{\nu}\right), \) with \(\nu_u = \nu_u^0\phi_u\) the types of the already explored vertices.

We claim that variables following \(\nu_l(m)\) are stochastically dominated by variables governed by \(\nu\). Indeed, use that for any non-decreasing \(f, h : \mathbb{R} \to \mathbb{R}\) and any random variable \(X\) we have \(\mathbb{E}[f(X)h(X)] \geq \mathbb{E}[f(X)]\mathbb{E}[h(X)]\). Then, for \(x \geq 0\),

\[\nu_l(m)(\psi) = \frac{-\mathbb{E}[g_{\nu}(\phi) \cdot 1_{\phi \leq \psi}]}{\mathbb{E}[g_{\nu}(\phi)]} \geq \frac{\mathbb{E}[g_{\nu}(\phi)]\mathbb{E}[1_{\phi \leq \psi}]}{\mathbb{E}[g_{\nu}(\phi)]} = \nu(\psi),\]

with \(\phi \sim \nu\).

Secondly, we claim that the weight of a vertex when it is just discovered is stochastically dominated by variables governed by \(\nu^*\). To prove this, let \(m \geq 0\) and assume the claim to hold for all \(l \leq m\) and vertex \(v\). Consider vertex \(v\) explored in step \((m+1)\) (itself discovered in step, say, \(l \leq m\)) with weight \(\phi_v^{(l)}\). Its children are selected from the set \(U(m)\) in which they have independent weights \((\phi^*_u)_{u \in U(m)}\) all stochastically dominated by \(\nu\).

We claim that the weight of a particle with weight \(\phi^* \sim \nu^*\) has its children selected following the same rules from a reservoir of \(|U(m)|\) particles with spins as in \(U(m)\) and i.i.d. weights \((\phi^*_u)_{u \in U(m)} \sim \nu\). Due to the assumed stochastic domination, there exists a coupling of the exploration process and the setting \(S\), such that pointwise \(\phi^{(l)}_v \leq \phi^*\) and \(\phi^*_v \leq \phi^*_u\) for all \(u\). To decide whether \(u \in U(m)\) is selected as a child, we can draw uniformly from \([0,1]\) a number \(U_u\) and include \(u\) in the exploration process exactly when \(\frac{(1_{\phi^*_u = \phi^*} + 1_{\phi^*_u - \phi^*})}{\nu^*U_u} \geq \frac{(\phi^{(l)}_v - \phi^*_v)}{\phi^{(l)}_v}\).

Denote the vertices in \(S_t\) by \(1, \ldots, S_t\) and their weights by \((\hat{\phi}^*_u)_{v \in S_t}\).

We shall use the same strategy as in Lemma 5.4 to bound

\[S_{t+1} = \sum_{v=1}^{S_t} \hat{D}_v^*,\]

where \(\hat{D}_v^*\) is the offspring-size of \(v\). In particular, to use large deviation theory as in (5.19), we shall calculate for \(\theta \geq 0\)

\[\mathbb{E}\left[ e^{\theta \sum_{v=1}^{S_t} \hat{D}_v^*} | S_t, (\hat{\phi}^*_u)_{u \in S_t}\right].\]
Caution is needed here as the variables \((\hat{D}_v^*)_{v \in S_t}\) are not independent. To circumvent this issue, we use the stochastic domination established above: If vertex \(v\) is explored in step \(m + 1\), then
\[
\hat{D}_v^* = \sum_{u \in H(m)} \text{Ber}\left(1_{\sigma_u = \sigma_v} a + 1_{\sigma_u = -\sigma_v} b \frac{\phi_u^{(m)}}{n}\right),
\]
where we recall that \(\phi_u^{(m)}\) is stochastically dominated by \(\nu\). Hence, using that \(1 + y \leq e^y\) for all \(y \in \mathbb{R}\),
\[
\mathbb{E}\left[e^{\theta \hat{D}_v^*}\right] \leq \mathbb{E}\left[\prod_u \left(1 + \frac{\phi_u^{(m)}}{n}(1_{\sigma_u = \sigma_v} a + 1_{\sigma_u = -\sigma_v} b)(e^{\theta} - 1)\right)\right]
\leq \left(1 + \frac{\phi_v^{(1)}}{n}(e^{\theta} - 1)\right)^n \left(1 + \frac{\phi_v^{(1)}}{n}(e^{\theta} - 1)\right)^{-n - \tau_v}.
\]
where \(r_n = \max\{\frac{a + a + n + b}{n} \frac{a - a + n + b}{n}\}\). Iterating (6.3), we obtain
\[
\mathbb{E}\left[e^{\theta \sum_{t=1}^{S_t} \hat{D}_v^*}\right| S_t, (\hat{\phi}_v^*)_{v \in S_t}\} \leq e^{\sum_{t=1}^{S_t} \hat{\phi}_v^{(1)}(e^{\theta} - 1)}.
\]

We use that \(r_n \phi_v^{(1)} \sum_{t=1}^{S_t} \hat{\phi}_v^d \leq r_n \phi_v^{(1)} \sum_{t=1}^{S_t} \hat{\phi}_v^\ast\), where \((\hat{\phi}_v^\ast)_{v \in S_t}\) are i.i.d. with law \(\nu^\ast\), to establish as in (5.19) an upper bound for \(S_t \rho_n - r_n \phi_v^{(1)} \sum_{t=1}^{S_t} \hat{\phi}_v^\ast\). Conditional on the latter being sufficiently small, we see that the rate function of \(\sum_{v=1}^{S_t} \hat{D}_v^*\), i.e.,
\[
x \mapsto \sup_{\theta \geq 0} \{\theta x - \log \theta \sum_{v=1}^{S_t} \hat{D}_v^*\}
\]
pointwise dominates the rate function of a Poisson random variable (which for Poi(\(\lambda\)) is for \(x \geq 0\) equal to \(\lambda I(x/\lambda)\) with \(I\) defined in (5.21)). Hence upon replacing \(\rho\) by \(\rho_n\) in the proof of Lemma 5.4, the arguments used there after line (5.21) establish the claim here.

**Lemma 6.5** (Degree-Corrected Extension of 30 in [2]). Let \(\ell = C \log_2(n)\), with \(0 < C < C_{\text{coupling}}\). Then, w.h.p., at most \(\rho^2 \log(n)\) vertices have a cycle in their \(\ell\) - neighbourhood. Further, w.h.p., the graph is \(\ell\) - tangle-free.

**Proof.** Fix a vertex \(v\). Let \(m \geq 0\) be the smallest integer such that all vertices within distance \(R\) of \(v\) have been revealed at step \(m\) of the exploration process. Now, the exploration process constructs a spanning tree \(T_m\) for \(G_R(v)\). However, edges between vertices in \(\partial G_{\ell}\) (\(r \leq \ell\)) are not inspected, and neither is it verified whether two vertices in \(\partial G_{\ell}\) share a common neighbour in \(\partial G_{\ell+1}\) (\(r \leq R - 1\)). The number of those uninspected edges is bounded by \(|G_{\ell}|^2\). Hence, among them at most \(\text{Bin}(|G_{\ell}|^2, c_1)\) are actually present in \(G_{\ast}\). Thus, using twice Markov’s inequality in conjunction with Lemma 6.4, for some \(c_2 > 0\),
\[
\mathbb{P}\left(G_{\ell}(v) \text{ is not a tree}\right) \leq \mathbb{E}\left[|G_{\ell}|^2\right] \frac{c_1}{n} \leq \frac{c_2 \rho^2 \ell}{n},
\]
and,
\[
\mathbb{P}\left(\sum_v 1_{G_{\ell}(v) \text{ is not a tree}} \geq \rho^2 \log(n)\right) \leq \frac{c_4}{\log(n)}.
\]

25
For the other claim, if the graph is tangled, then there is a vertex such that among its uninspected edges in the exploration process at step $m$, at least two are in fact present. Now,

$$\Pr \left( \text{Bin} \left( \left| G_r \right|^2, \frac{c_1}{n} \right) \geq 2 \right) \leq \left( \frac{c_1}{n} \right)^4 \mathbb{E} \left[ \left| G_r \right|^4 \right] \leq \frac{c_5 \rho^t}{n^t}. $$

A union bound over all vertices then gives

$$\Pr \left( G \text{ tangled} \right) \leq \frac{c_6 \rho^{4t}}{n^t} = o(1).$$

\[ \square \]

### 6.2 Geometric growth

Here we show that for $k \in \{1, 2\}, \langle B^r \chi_k, \delta_e \rangle$ grows nearly geometrically in $t$ with rate $\mu_k$. Corollary 6.7 then establishes a bound for $r \leq \ell$ on $\sup_{\langle B^r \chi_k, x \rangle = 0, \|x\| = 1} \|\langle B^r \chi_k, x \rangle\|$ crucial for the norm bounds in Section 9.

**Proposition 6.6** (Degree-Corrected Extension of Proposition 33 in [2]). Let $\ell = C \log_p(n)$, with $0 < C < C_{\text{coupling}} \wedge \left( \frac{1}{2} - \left( \frac{1}{2} \wedge \frac{1}{36} \right) \right) = C_{\text{coupling}}$. For $e \in \tilde{E}(V)$, let $\tilde{E}_e$ be the set of oriented edges such that either $(G, e_2) \notin$ is a tree or the event $E(e)$ (defined in Corollary 6.3) does not hold. Then, w.h.p. for $k \in \{1, 2\}$:

1. $|\tilde{E}_e| \ll (\log n)^2 n^{-1/2} \wedge \frac{1}{8}$,
2. for all $e \in \tilde{E} \setminus \tilde{E}_e$, $0 \leq r \leq \ell$,

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

3. for all $e \in \tilde{E}_e$, $0 \leq r \leq \ell$,

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

**Proof.** (i) follows from Lemma 6.4 and Corollary 6.3.

To prove (ii), recall that $B_{\tilde{e}g}$ is the number of non-backtracking paths of length $r$ (i.e., containing $r + 1$ edges) between $\tilde{e}$ and $\tilde{g}$. Further, if $G_r(e_2)$ is a tree, then there is exactly one path between $e$ and any edge $g$ on the tree. Hence

$$\langle B^r \chi_k, \delta_e \rangle = \langle \phi_k, \Psi_r(e) \rangle.$$

An appeal to Corollary 6.3 then establishes (ii).

Further, (iii) follows from the fact that $G$ is $\ell$-tangle-free with high probability, so that there are at most two non-backtracking walks of length $r$ between any edges $\tilde{e}$ and $\tilde{f}$. Thus,

$$|\langle B^r \chi_k, \delta_e \rangle| \leq 2 \|g_k\| \max S_l(e) \leq \log^2(n)\rho^r,$$

with probability at least $1 - e^{-\Omega(n)}$, due to Lemma 6.4. $\square$

**Corollary 6.7** (Degree-Corrected Extension of Corollary 34 in [2]). Let $\ell = C \log_p(n)$, with $0 < C < C_{\text{coupling}} \wedge \left( \frac{1}{2} - \left( \frac{1}{2} \wedge \frac{1}{36} \right) \right)$ and $k \in \{1, 2\}$:

$$\sup_{\langle B^r \chi_k, x \rangle = 0, \|x\| = 1} \|\langle B^r \chi_k, x \rangle\| \leq (\log n)^5 n^{1/2} \rho^{r/2}.$$
Proof. Using that $\langle B^r \chi_k, x \rangle = 0$ and Proposition 6.6 (iii), we write,

$$\langle B^r \chi_k, x \rangle = \left| \sum_{e \in E^r_l} x_e \langle B^r \chi_k, \delta_e \rangle + \sum_{e \not\in E^r_l} x_e \langle B^r \chi_k, \delta_e \rangle - \mu_k^{r-\ell} \left( \sum_{e \in E^r_l} x_e \langle B^r \chi_k, \delta_e \rangle + \sum_{e \not\in E^r_l} x_e \langle B^r \chi_k, \delta_e \rangle \right) \right|$$

$$\leq (\log n)^2 \rho^r \sqrt{|E^r_l|} + \sum_{e \not\in E^r_l} |x_e| \| \langle B^r \chi_k, \delta_e \rangle - \mu_k^{r-\ell} \langle B^r \chi_k, \delta_e \rangle \| + \mu_k^{r-\ell} \log(n)^2 \rho^r \sqrt{|E^r_l|}.$$ 

(6.4)

Now, $|\mu_k| > 1$ and for $e \not\in E^r_l$, bound (ii) in Proposition 6.6 applies, so that w.h.p.

$$\langle B^r \chi_k, x \rangle \leq 2 \rho^r (\log n)^2 \sqrt{|E^r_l|} + \rho^r (\log n)^3 \sqrt{|E^r_l|}$$

$$\leq \rho^r (\log n)^3 n^{r/2} + \rho^r (\log n)^{5/2}$$

$$\leq \rho^r (\log n)^{5/2} n^{1/2},$$

since $\rho^r = n^{C} \ll n^{\frac{r}{2}}$. \qed

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

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 $G^*$ the set of rooted weighted coloured graphs. We denote an element of $G^*$ by $(G, o): G = (V, E)$ is then a weighted coloured graph and $o \in V$ is some distinguished vertex. A function $\tau : G^* \to \mathbb{R}$ 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 7.1. The bound follows from the law of total variance,

$$\text{Var} \left( \sum_{v=1}^{n} \tau(G, v) \right) = \mathbb{E} \left[ \text{Var} \left( \sum_{v=1}^{n} \tau(G, v) \mid \phi_1, \ldots, \phi_n \right) \right] + \text{Var} \left( \mathbb{E} \left[ \sum_{v=1}^{n} \tau(G, v) \mid \phi_1, \ldots, \phi_n \right] \right),$$

together with an application of Efron-Stein’s inequality to both terms on the right. Note that $\mathbb{E} \left[ \sum_{v=1}^{n} \tau(G, v) \mid \phi_1, \ldots, \phi_n \right]$ 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 $\mathbb{E} [\tau(T, o)]$, where $(T, o)$ is the branching process from Section 5, with root $o$ having spin drawn uniformly from $\{+, -\}$ and weight governed by $\nu$, see Proposition 7.2. 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 7.1 (Degree-Corrected Extension of Proposition 35 in [2]).** Let $G$ be drawn according to the DC-SBM. There exists $c > 0$ such that if $\tau, \varphi : G^* \to \mathbb{R}$ are $\ell$-local, $|\tau(G, o)| \leq \varphi(G, o)$ and $\varphi$ is non-decreasing by the addition of edges, then

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

27
Proof. We start by using the law of total variance for $Y = \sum_{v=1}^n \tau(G, v)$:
\[
\text{Var}(Y) = \mathbb{E}[\text{Var}(Y|\phi_1, \ldots, \phi_n)] + \text{Var}(\mathbb{E}[Y|\phi_1, \ldots, \phi_n]),
\]
and shall apply Efron-Stein’s inequality on both terms.

Define the function $h$ for $(\psi_1, \ldots, \psi_n) \in [\phi_{\min}, \phi_{\max}]^n$ as $h(\psi_1, \ldots, \psi_n) = \mathbb{E}[\phi_1 = 1, \ldots, \phi_n = 1]$. We need to bound $|h(\psi_1, \ldots, \psi_k, \psi_{k+1}, \ldots, \psi_n) - h(\psi_1, \ldots, \psi_k, \psi_{k+1}, \ldots, \psi_n)|^2$ for arbitrary $\psi_k \in [\phi_{\min}, \phi_{\max}]$. Denote by $G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}$ the random graph $G$, conditional on $\phi_1 = 1, \ldots, \phi_n = 1$. Assume without loss of generality that $\psi_k \geq \psi_k'$, then, there exists a coupling of $G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}$ and $G_{\psi_1, \ldots, \psi_k', \ldots, \psi_n}$ such that $G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}$ is a subgraph of $G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}$ obtained after removing some edges between $k$ and its neighbours in the latter graph. For this coupling, $|\tau(G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}) - \tau(G_{\psi_1, \ldots, \psi_k', \ldots, \psi_n})|$ is nonzero only if $u \in V(G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}, k)$, and it is bounded by $\max_v \varphi(G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}, v) + \max_v \varphi(G_{\psi_1, \ldots, \psi_k', \ldots, \psi_n}, v)$. Consequently,
\[
|h(\psi_1, \ldots, \psi_{k-1}, \psi_k, \psi_{k+1}, \ldots, \psi_n) - h(\psi_1, \ldots, \psi_{k-1}, \psi_k', \psi_{k+1}, \ldots, \psi_n)|^2 \\
\leq \mathbb{E}[|V(G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}, k)| \left( \max_v \varphi(G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}, v) + \max_v \varphi(G_{\psi_1, \ldots, \psi_k', \ldots, \psi_n}, v) \right)^2] \\
\leq \mathbb{E}[|V(G_{k, \infty}, k)|^2] \cdot \left( \max_v \varphi(G_{\psi_1, \ldots, \psi_k, \ldots, \psi_n}, v) + \max_v \varphi(G_{\psi_1, \ldots, \psi_k', \ldots, \psi_n}, v) \right) \\
\cdot \left( \mathbb{E} \left[ \max_v \varphi^2(G, v) \bigg| \phi_1 = 1, \ldots, \phi_k = \psi_k, \ldots, \phi_n = \psi_n \right] + \mathbb{E} \left[ \max_v \varphi^2(G, v) \bigg| \phi_1 = 1, \ldots, \phi_k = \psi_k', \ldots, \phi_n = \psi_n \right] \right)
\]
where $G_{k, \infty}$ is the random graph $G$ conditioned on $\phi_k = \phi_{\max}$, and where we used Hölder’s inequality and the fact that $(x + y)^2 \leq 3(x^2 + y^2)$ for any $x, y \in \mathbb{R}$. Hence, using again Hölder’s inequality, Efron-Stein’s inequality becomes
\[
\text{Var}(\mathbb{E}[Y|\phi_1, \ldots, \phi_n]) \leq \frac{1}{2} \sum_{k=1}^n \mathbb{E}[|h(\phi_1, \ldots, \phi_k, \ldots, \phi_n) - h(\phi_1, \ldots, \phi_k', \ldots, \phi_n)|^2] \\
\leq 3 \sum_{k=1}^n \sqrt{\mathbb{E}[|V(G_{k, \infty}, k)|^2]} \sqrt{\mathbb{E} \left[ \max_v \varphi^4(G, v) \right]},
\]
where $(\phi_k')_k$ is an i.i.d. copy of $(\phi_k)_k$. Now, due to 6.4, $\mathbb{E}[|V(G_{k, \infty}, k)|^2] \leq \frac{\rho^4}{\rho^{4\ell}}$. Thus,
\[
\text{Var}(\mathbb{E}[Y|\phi_1, \ldots, \phi_n]) \leq c_{2n} \rho^{2\ell} \sqrt{\mathbb{E} \left[ \max_v \varphi^4(G, v) \right]}.
\]
To bound $\mathbb{E}[\text{Var}(Y|\phi_1, \ldots, \phi_n)]$ we use again Efron-Stein’s inequality. Define for $1 \leq k \leq n$, $X_k = \{1 \leq v \leq k : (v, k) \in E\}$, where $E$ is the edge set of $G$. Then, conditional on the weights, $(X_k)_k$ are independent. Let $(X'_k)_k$ be an independent copy of $(X_k)_k$ and define $G_k$ as the graph on vertex set $V$ with edge set $\cup_{v \in k} X_v \cup X'_v$. Thus, conditional on the weights, $G_k$ equals $G$ except for the edges in $\{1 \leq v \leq k\}$ which are redrawn independently. Now, for some function $F$,
\[
\mathbb{E} \left[ \sum_{k=1}^n \tau(G, v)|\phi_1 = 1, \ldots, \phi_n = 1 \right] = F(X_1, \ldots, X_n),
\]
and hence,
\[
\mathbb{E} \left[ \sum_{k=1}^n \tau(G_k, v)|\phi_1 = 1, \ldots, \phi_n = 1 \right] = F(X_1, \ldots, X'_n, \ldots, X_n).
\]
Proceeding as above, we obtain

\[
\text{Var}(Y|\phi_1, \ldots, \phi_n) \leq \frac{1}{2} \sum_{k=1}^{n} \mathbb{E} \left[ |F(X_1, \ldots, X_k, \ldots, X_n) - F(X_1, \ldots, X'_k, \ldots, X_n)|^2 \right]
\]

\[
\leq \frac{1}{2} \sum_{k=1}^{n} \sqrt{\mathbb{E} \left[ |V(G, k)|^4 \right] \mathbb{E} \left[ \max_{v} \varphi(G, v) + \max_{v} \varphi(G_k, v) \right]^4}
\]

\[
\leq c_3 n^2 \rho^2 \sqrt{\mathbb{E} \left[ \max_{v} \varphi^4(G, v) \right]}. 
\]

(7.3)

**Proposition 7.2** (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 \( v \). Let \( \ell = C \log_{n}(n) \), with \( C < C_{\text{coupling}} \). There exists \( c > 0 \) such that if \( \tau, \varphi : \mathbb{G}^* \rightarrow \mathbb{R} \) are \( \ell \)-local, \( |\tau(G, o)| \leq \varphi(G, o) \) and \( \varphi \) is non-decreasing by the addition of edges, then

\[
\mathbb{E} \left[ \frac{1}{n} \sum_{\tau=1}^{n} \tau(G,v) - \mathbb{E} [\tau(T,o)] \right] \leq c_2 n^{-\left(\frac{7}{2}+\frac{1}{2n}\right)} \left( \mathbb{E} \left[ \max_{v \in [n]} \varphi^4(G,v) \right]^{1/4} \vee \mathbb{E} \left[ \varphi^2(T,o) \right]^{1/2} \right) + \mathcal{O}(n^{-\gamma})
\]

(7.4)

**Proof.** We recall that the coupling between neighbourhoods and branching processes is such that, in case of success, the weights are equal in both processes. Therefore, as in the proof of Proposition 36 in [2], we obtain

\[
\mathbb{E} \left[ \frac{1}{n} \sum_{\tau=1}^{n} \tau(G,v) \right] = \mathbb{E} [\tau(T,o)] + \epsilon(n),
\]

where

\[
\epsilon(n) = \mathcal{O}(n^{-\gamma}) + c_1 n^{-\left(\frac{7}{2}+\frac{1}{2n}\right)} \mathbb{E} \left[ \max_{v \in [n]} \varphi^2(G,v) \right] \vee \mathbb{E} \left[ \varphi^2(T,o) \right].
\]

This error stems from the probability for the coupling to fail.

Hence,

\[
\mathbb{E} \left[ \frac{1}{n} \sum_{\tau=1}^{n} \tau(G,v) - \mathbb{E} [\tau(T,o)] \right] \leq \sqrt{ \text{Var} \left( \frac{1}{n} \sum_{\tau=1}^{n} \tau(G,v) \right) + \epsilon(n) }.
\]

An appeal to Proposition 7.1 then finishes the proof.

### 7.1 Application with some specific local functionals

Here we consider \( \langle B^\ell \chi_1, B^\ell \chi_2 \rangle, \langle B^{2\ell} \chi_k, B^\ell \chi_2 \rangle, \) and \( \langle B^\ell B^\ell \chi_1, B^\ell B^\ell \chi_2 \rangle \), quantities occurring in Proposition 4.1.

Explicitly, \( B^\ell \chi_k(e) = \sum_f B^\ell_{ef} g_k(\sigma(f)) \phi_{f,c} \), where we recall that \( B^\ell_{ef} \) is the number of non-backtracking walks from \( e \) to \( f \). Now, if the oriented \( t- \) neighbourhood of \( e \) is a tree, then \( B^\ell \chi_k(e) = \langle g_k, \Psi(f) \rangle \). With this intuition in mind, we analyse likewise expressions in Proposition 7.3 below.

Inspired by (5.44), which expresses \( B^\ell B^\ell \chi_k \) on trees in terms of the operator \( Q_{k,t} \), we extend the latter to an operator defined on general graphs. First, for \( e \in \vec{E}(V) \) and \( t \geq 0 \), set \( \mathcal{Y}_t(e) = \{ f \in \vec{E} : d(e,f) = t \} \). Then, for \( k \in \{1,2 \} \), we set

\[
P_{k,t}(e) = \sum_{t=0}^{t-1} \sum_{f \in \mathcal{Y}_t(e)} L_k(f). 
\]

(7.5)
with

\[ L_k(f) = \sum_{(g,h) \in Y_k(f) \setminus \{e\} : g \neq h} \langle g_k, \Psi_t(g) \rangle \tilde{S}_{t-1}(h), \]

where \( \Psi_t(g), \tilde{S}_{t-1}(h) = ||\tilde{Y}_{t-1}(h)||\) are the variables \( \Psi_t(g) \), respectively \( \tilde{S}_{t-1}(h) \), defined on the graph \( G \) where all edges in \((G, e_2)\) have been removed. Note that, if \((G, e)_{2t}\) is a tree, then \( \Psi_s(g) = \Psi_s(g) \) for \( s \leq 2t - t \). Compare \( P_{k,t} \) to \( Q_{k,t} \) in (5.41) and \( L_k(f) \) to \( L_{k} \) in (5.46).

Finally, define

\[ S_{k,t}(e) = S_t(e)g_k(\sigma(e))\phi_{e_i}. \quad (7.6) \]

We then have an extension of (5.44), when \((G, e_2)_{2t}\) is a tree:

\[ B^tB^{t'} \chi_k(e) = P_{k,t}(e) + S_{k,t}(e). \quad (7.7) \]

We analyse (7.7) in Proposition 7.4 below.

**Proposition 7.3** (Degree-Corrected Extension of Proposition 37 in \([2]\)). Let \( \ell = C\log n \) with \( 0 < C < C_{\text{coupling}} \).

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

\[ \frac{1}{n} \sum_{e \in \bar{E}} \frac{\langle g_k, \Psi_t(e) \rangle^2}{\mu_k^2} \to c'_k. \]

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

\[ \frac{1}{n} \sum_{e \in \bar{E}} \frac{\langle g_k, \Psi_t(e) \rangle^2}{\mu_k^2} \to c''_k. \]

(iii)

\[ \mathbb{E} \left[ \frac{1}{n} \sum_{e \in \bar{E}} \langle g_1, \Psi_t(e) \rangle \langle g_2, \Psi_t(e) \rangle \right] \leq (\log n)^3 n^{-2C - \left( \frac{q}{4} \wedge \frac{q}{4} \right)} + n^{-\gamma}. \]

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

\[ \mathbb{E} \left[ \frac{1}{n} \sum_{e \in \bar{E}} \langle g_k, \Psi_{2t}(e) \rangle \langle g_j, \Psi_t(e) \rangle \right] \leq (\log n)^3 n^{-2C - \left( \frac{q}{4} \wedge \frac{q}{4} \right)} + n^{-\gamma}. \]

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

\[ \frac{1}{n} \sum_{e \in \bar{E}} \frac{\langle g_k, \Psi_{2t}(e) \rangle \langle g_k, \Psi_t(e) \rangle}{\mu_k^3} \to c''_k. \]

**Proof.** We give the key steps used to prove Proposition 37 in \([2]\) together with the main differences in the current setting. For (i), consider the branching process defined in Section 5, which we denote again by \( Z_t(\pm) \). We denote the associated random rooted tree by \((T, o)\).

Put \( \tau(G, v) = \sum_{e \in E, e_1 = v} \frac{\langle g_k, \Psi_t(e) \rangle^2}{\mu_k^2} \). Then, \( \frac{1}{n} \sum_v \tau(G, v) = \frac{1}{n} \sum_{e \in E} \frac{\langle g_k, \Psi_t(e) \rangle^2}{\mu_k^2} \) and \( \tau(G, v) \leq \varphi(G, v) := \phi_{\max} \delta_{\alpha}^2 \). It follows from Lemma 6.4 that \( \mathbb{E} \left[ \max_{v \in [n]} \varphi(G, v) \right] = O \left( (\log n)^{\frac{q}{2}} n^{-\frac{q}{2}} \right) \).

We have \( \tau(T, o) = \sum_{v \in Z^T} \frac{\langle g_k, \Psi_t(e) \rangle^2}{\mu_k^2} \). Theorem 5.3 says that \( \left( \frac{\langle g_k, \Psi_t(e) \rangle}{\mu_k} \right) \) converges in \( L^2 \) and so does it conditional on \( ||Z^T|| = 1 \). Hence, \( \mathbb{E} [\tau(T, o)] \) converges.
An appeal to Proposition 7.2 in conjunction with the triangle inequality then establishes that \( \frac{1}{n} \sum \tau(G, v) \) converges to a constant, say \( c_k' \).

Statement (ii) follows similarly.

The statements (iii) – (v) follow after properly choosing local functionals. We further use that \( \mathbb{E}[\phi_u \phi_v g_1(\sigma_u)g_2(\sigma_v)|T] = \mathbb{E}[\phi_u \phi_v \bar{\sigma}_v|T] = 0 \), for any two nodes \( u, v \). Further, on the branching process, \( \mathbb{E}[\phi_g \phi_j (g_j, \Psi_j)\Psi_j] = (g_j, \Psi_j) (g_k, M' \Psi_k) = \mu^0_k \langle g_k, \Psi \rangle (g_j, \Psi) \).

\[ \text{Proposition 7.4 (Degree-Corrected Extension of Proposition 38 in [2]).} \]

Let \( \ell = C \log n \) with \( C < C_{\text{coupling}} \).

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

\[
\frac{1}{n} \sum_{e \in E} P_0^{2k,e}(e) \mu_{k}^{2k} \rightarrow c_k'' .
\]

(ii) \( \mathbb{E} \left[ \frac{1}{n} \sum_{e \in E} (1) \sum_{i} (S_1) \sum_{j} (S_2) \right] \leq (\log n)^3 n^{4C - (\frac{2}{\gamma} + \frac{3}{\gamma})} \]

\[ \text{Proof.} \]

Starting with (i), we define the local function \( \tau \) as \( \tau(G, v) = \sum_{e \in E, e_1 = v} P_0^{2k,e}(v) \mu_{k}^{2k}, \) for a rooted graph \( (G, v) \). Let

\[ M(v) = \max_{u \leq t \leq t \in (G, v)} \max_{v \in E, e_1 = v} (S_i(u)/\rho^s). \]

By monotonicity, the statement of Lemma 6.4 holds also for \( \tilde{S}_{t-1}(h) \) and \( \tilde{S}_{t}(g) \). We use this fact to bound powers of \( M(v) \) in the following calculation:

\[
\begin{align*}
\tau(G, v) & \leq \rho^{-2\ell} \sum_{e \in E, e_1 = v} \left( \sum_{t = 0}^{\ell-1} \sum_{f \in \mathcal{Y}_t(e)} \|g_k\|_{\text{max}} \tilde{S}_t(f) \tilde{S}_{t-1}(f) \right)^2 \\
& \leq c_1 \rho^{-2\ell} \sum_{e \in E, e_1 = v} \left( \sum_{t = 0}^{\ell-1} \sum_{f \in \mathcal{Y}_t(e)} M^2(v) \rho^{t+1} \rho^{-t} \right)^2 \\
& = c_1 \left( M^2(v) \rho^2 \right)^2 \sum_{e \in E, e_1 = v} \left( \sum_{t = 0}^{\ell-1} S_t(e) \right)^2 \\
& = c_2 \left( M^2(v) \rho^2 \right)^2 \sum_{e \in E, e_1 = v} \left( M(v) \rho^s \right)^2 \\
& \leq c_2 M^2(v) \rho^{2\ell}.
\end{align*}
\]

We put \( \varphi(G, v) = c_2 M^7(v) \rho^{2\ell} \). Then, \( \mathbb{E} \left[ \max_{v} \varphi(G, v)^4 \right] = O((\log n)^2 \rho^{4\ell}) \), and the same bound holds for \( \varphi(T, o) \). From Proposition 7.2, we then know that

\[
\mathbb{E} \left[ \frac{1}{n} \sum_{e \in E} P_0^{2k,e}(e) \mu_{k}^{2k} - \mathbb{E} [\tau(T, o)] \right] \leq c_3 n^{-\left(\frac{2}{\gamma} + \frac{3}{\gamma}\right)} (\log n)^7 \rho^{2\ell}, \tag{7.8}
\]

where

\[
\begin{align*}
\tau(T, o) & = \frac{1}{\mu^{2\ell}_k} \sum_{e \in \mathcal{Y}_o^{\ell}} P_0^{2k,e}(o \rightarrow v) \\
& = \frac{1}{\mu^{2\ell}_k} \sum_{e \in \mathcal{Y}_o^{\ell}} (Q_{k,e})^2, \tag{7.9}
\end{align*}
\]

31
where $Q_{k,\ell}^e$ is equal to $Q_{k,\ell}$ defined on the subtree of all vertices with common ancestor $v$.

We need to show that the expectation of $\tau(T, o)$ converges for $\ell \to \infty$. Conditional on $\sigma_o$, and $\{\gamma^\ell_{V,E}\}$ for $\ell = 1, 2, 3$ are independent copies of $Q_{k,\ell}$ defined on the branching process in Section 5 where the root has spin $\sigma_o$ and random weight governed by the biased law $\nu^\ell$. The uniform $L^2$ convergence in Theorem 5.6 establishes the claim.

We now prove (ii). Put $\tau(G, v) = \sum_{u \in E} (P_{1,\ell}(v) + S_{1,\ell}(v))(P_{2,\ell}(v) + S_{2,\ell}(v))$. We claim that $\mathbb{E}[\tau(T, o)] = 0$. Consider $\tau(T, o) = \sum_{v \in Z_T} (P_{1,\ell}(o \rightarrow v) + S_{1,\ell}(o \rightarrow v))(P_{2,\ell}(o \rightarrow v) + S_{2,\ell}(o \rightarrow v))$. Firstly, for $k \in \{1, 2\}$, $P_{k,\ell}(o \rightarrow v) = Q_{k,\ell}^e$. Now, it follows from Theorem 5.7, that $\mathbb{E}[Q_{1,\ell}^eQ_{2,\ell}^e] = 0$, since $\sigma_v$ is drawn uniformly from $\{+, -\}$.

Secondly, $S_{1,\ell}(o \rightarrow v)S_{2,\ell}(o \rightarrow v) = \frac{1}{2} \phi^2 \sigma_o S^2_t(o \rightarrow v)$ has also zero expectation.

Thirdly,

$$Q_{1,\ell}^eS_{2,\ell}(o \rightarrow v) = \frac{1}{2} \sum_{(u_0, \ldots, u_{2\ell+1}) \in P_{2\ell+1}} \phi_{u_{2\ell+1}} \phi_o \sigma_o S_t(o \rightarrow v),$$

(7.10)

where $P_{2\ell+1}^e$ is $P_{2\ell+1}$ from (5.41) defined on the subtree of all vertices with common ancestor $v$. The expectation of $Q_{1,\ell}^eS_{2,\ell}(o \rightarrow v)$ is thus zero since $\sigma_o$ is independent of all other terms in (7.10).

Lastly, $Q_{2,\ell}^e = \sum_{(u_1, \ldots, u_{2\ell+1}) \in P_{2\ell+1}} \sigma_{u_{2\ell+1}}$, is seen to have zero expectation.

Those four statements combined establish $\mathbb{E}[\tau(T, o)] = 0$. As above, we calculate $\mathbb{E}[\max_v \varphi(G, v)^5] = O((\log n)^{20} \rho^{15\delta})$.

8 Proof of Proposition 4.1

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

$$N_{k,\ell}(e) = (g_k, \Psi_k(e)).$$

If $(G, e_2)$ is a tree, then

$$N_{k,\ell}(e) = (B^\ell \chi_k, \delta_k),$$

and we have a similar expression for $B^\ell B^\ell \chi_k$ in (7.7). Now, at most $\rho^{2\ell}(\log n)$ vertices have a cycle in their $\ell$-neighbourhood (see Lemma 6.5). Therefore:

**Lemma 8.1** (Degree-Corrected Extension of Lemma 39 in [2]). Let $\ell = C \log n$ with $0 < C < C_{\text{min}}$. Then, w.h.p. $\|B^\ell \chi_k - N_{k,\ell}\| = O((\log n)^{5/2} \rho^{2\ell}) = o\left(\rho^{\ell/2} \sqrt{n}\right)$, $\|B^\ell B^\ell \chi_k - P_{k,\ell} - S_{k,\ell}\| = O((\log n)^3 \rho^{3\ell})$, and $\|B^\ell B^\ell \chi_k - P_{k,\ell}\| = O(\rho^{\ell/2} \sqrt{n})$.

**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 $\|B^\ell \chi_k - N_{k,\ell}(e)\|$ on edges for which $(G, e_2)$ is not a tree. For this, use that with high probability the graph is $2\ell$-tangle free so that there are at most two non-backtracking paths between $e$ and any edge at distance $\ell$.

We can thus in our calculations replace $B^\ell \chi_k$ by $N_{k,\ell}$ and $B^\ell B^\ell \chi_k$ by $P_{k,\ell}$. From Propositions 7.3 and 7.4, Proposition 4.1 then follows:

**Proof of Proposition 4.1.** This proof follows the corresponding proof in [2]. We give the key observations:

(i) From Proposition 7.3 (i), $\|N_{k,\ell}\| \sim \sqrt{\mu_{k}^e}$ and from Proposition 7.4 (i), $\|P_{k,\ell}\| \sim \sqrt{\mu_{k}^e}$.

(ii) From Proposition 7.3 (v), $\|N_{k,\ell} - N_{k,2\ell}\| \sim \eta_{k,\ell}^e$.

(iii) From Proposition 7.3 (iii), $\|N_{k,\ell} - N_{2,\ell}\| \sim (\log n)^3 \eta_{k,\ell}^e\chi_k - \frac{2\ell}{3\delta} \frac{3\ell}{4}.$

(iv) From Proposition 7.3 (iv), $\|N_{k,2\ell} - N_{k,\ell}\| \sim (\log n)^3 \eta_{k,\ell}^{2\ell} \chi_k - \frac{2\ell}{3\delta} \frac{3\ell}{4}.$

(v) From Proposition 7.4 (ii), $\|P_{k,\ell} + S_{1,\ell} + P_{2,\ell} + S_{2,\ell}\| \sim (\log n)^3 \eta_{k,\ell}^{2\ell} \chi_k - \frac{2\ell}{3\delta} \frac{3\ell}{4}.$

32
9 Norm of non-backtracking matrices

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

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

$$B_{ef} = A_e A_f 1_{e_2 = f_1} 1_{e_1 \neq f_2}, \quad (9.1)$$

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

For integer $k \geq 1$, $e, f \in \tilde{E}_K(V)$, we let $\Gamma_{ef}^k$ be the set of non-backtracking walks $\gamma = (\gamma_0, \ldots, \gamma_k)$ 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

$$(B^{(k)})_{ef} = \sum_{\gamma \in \Gamma_{ef}^k} \prod_{s=0}^{k} A_{\gamma_s \gamma_{s+1}}. \quad (9.2)$$

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

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

From Lemma 6.5, 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 (9.2):

$$(B^{(k)})_{ef} = \sum_{\gamma \in F_{ef}^{k+1}} \prod_{s=0}^{k} A_{\gamma_s \gamma_{s+1}}, \quad (9.3)$$

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

Define for $u \neq v$ the centred random variable

$$A_{uv} = A_{uv} - \frac{\phi_u \phi_v}{n} W_{\phi_u \phi_v}, \quad (9.4)$$

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

$$\Delta^{(k)}_{ef} = \sum_{\gamma \in F_{ef}^{k+1}} \prod_{s=0}^{k} A_{\gamma_s \gamma_{s+1}}. \quad (9.5)$$

Further, we set

$$\Delta^{(0)}_{ef} = 1_{e = f} A_e \quad \text{and} \quad B^{(0)}_{ef} = 1_{e = f} A_e. \quad (9.6)$$

To decompose (9.3), following a decomposition that appeared first in [16], we use

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

33
with \( x_s = A_{\gamma_s\gamma_{s+1}} \) and \( y_s = A_{\gamma_s\gamma_{s+1}} \) on a path \( \gamma \in F^{\ell+1}_e \):

\[
\prod_{s=0}^{\ell} A_{\gamma_s\gamma_{s+1}} = \prod_{s=0}^{\ell} \Delta_{\gamma_s\gamma_{s+1}} + \sum_{t=0}^{\ell-1} \prod_{s=0}^{\ell-t} \Delta_{\gamma_s\gamma_{s+1}} \left( \frac{\phi_{\gamma_s\gamma_{s+1}}}{n} W_{\sigma_{\gamma_s\gamma_{s+1}}} \right) \prod_{s=t+1}^{\ell} A_{\gamma_s\gamma_{s+1}}.
\]

Summing over all \( \gamma \in F^{\ell+1}_e \) then gives

\[
P^{(\ell)}_{e,f} = \sum_{\gamma \in F^{\ell+1}_e} \prod_{s=0}^{\ell} \Delta_{\gamma_s\gamma_{s+1}} + \sum_{t=0}^{\ell} \prod_{s=0}^{\ell-t} \Delta_{\gamma_s\gamma_{s+1}} \left( \frac{\phi_{\gamma_s\gamma_{s+1}}}{n} W_{\sigma_{\gamma_s\gamma_{s+1}}} \right) \prod_{s=t+1}^{\ell} A_{\gamma_s\gamma_{s+1}} (9.7)
\]

Consider the two products in the summation over \( F^{\ell+1}_e \) on the right of (9.7): We can, for \( 1 \leq t \leq \ell - 1 \), replace the summation over \( F^{\ell+1}_e \) by summing over all pairs \( \gamma' = (\gamma_0, \ldots, \gamma_t) \in F^{\ell+1}_e \) and \( \gamma'' = (\gamma_{t+1}, \ldots, \gamma_{\ell+1}) \in F^{\ell+1}_e \) for some \( g, g' \in \hat{E}(V) \) 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 \( g \rightarrow g' \)). However caution is needed, as this summation also includes tangled paths, namely those in the sets \( \{ F^{\ell+1}_{e,f} \}_{t=0}^{\ell} \). Where, for \( 1 \leq t \leq \ell - 1 \), \( F^{\ell+1}_{e,f} \) is defined as the collection of all tangled paths \( \gamma = (\gamma_0, \ldots, \gamma_{t+1}) = (\gamma', \gamma'') \in \Gamma^{\ell+1}_{e,f} \) with \( \gamma' \) and \( \gamma'' \) as above. For \( t = 0 \), \( F^{\ell+1}_{e,f} \) consists of all non-backtracking tangled paths \( (\gamma', \gamma'') \) with \( \gamma' = (e_1) \) and \( \gamma'' \in F^{\ell+1}_{e,f} \) for any \( g' \) such that \( g'_1 = e_2 \). For \( t = \ell \), \( F^{\ell+1}_{e,f} \) is the set of non-backtracking tangled paths \( (\gamma', \gamma'') \) such that \( \gamma'' = (f_2) \) and \( \gamma' \in F^{\ell+1}_{e,f} \) for some \( g \in \hat{E}(V) \) with \( g_2 = f_1 \). We rewrite (9.7) as

\[
P^{(\ell)} = \Delta^{(\ell)} + \frac{1}{n} KB^{(\ell-1)} + \frac{1}{n} \sum_{t=1}^{\ell} \Delta^{(t-1)} K^2 B^{(\ell-t-1)} + \frac{1}{n} \Delta^{(\ell-1)} \hat{K} - \frac{1}{n} \sum_{t=0}^{\ell} R^{(t)}_{e,f}, (9.8)
\]

where for \( e, f \in E_K \),

\[
K_{e,f} = 1_{e \rightarrow f} \phi_{e_1} \phi_{f_2} W_{\sigma(e_1)\sigma(e_2)}, \tag{9.9}
\]

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

\[
\hat{K}_{e,f} = 1_{e \rightarrow f} \phi_{f_1} \phi_{f_2} W_{\sigma(f_1)\sigma(f_2)}, \tag{9.10}
\]

and where

\[
(R^{(t)}_{e,f})_{e,f} = \sum_{\gamma \in \Gamma^{\ell+1}_{e,f}} \prod_{s=0}^{\ell-t} A_{\gamma_s\gamma_{s+1}} \phi_{\gamma_s\gamma_{s+1}} W_{\sigma(\gamma_s)\sigma(\gamma_{s+1})} \prod_{s=t+1}^{\ell} A_{\gamma_s\gamma_{s+1}}. \tag{9.12}
\]
Indeed,

\[
\left( \sum_{t=1}^{\ell-1} \Delta^{(t-1)} K(2) B^{(\ell-t-1)} \right)_{ef} = \sum_{t=1}^{\ell-1} \sum_{g,g'} \Delta^{(t-1)} K(2) B^{(\ell-t-1)}_{g'f} \\
= \sum_{t=1}^{\ell-1} \sum_{g,g'} \sum_{\gamma' \in F'_g} \sum_{\gamma'' \in F'_{g'}} \prod_{s=0}^{t-1} \Delta_{\gamma'\gamma''+1} \phi_{\gamma'\gamma''} W_{\sigma(\gamma'') \sigma(\gamma'')} \\
\cdot \prod_{s=0}^{\ell-t-1} A_{\gamma'\gamma''+1},
\]

(9.13)

and

\[
\left( KB^{(\ell-1)} \right)_{ef} = \sum_{g} \sum_{\gamma'' \in F'_{g'}} 1_{g \rightarrow g} \phi_{e1} \phi_{e2} W_{\sigma(e1)\sigma(e2)} A_{e1,e2} \prod_{s=1}^{\ell-2} A_{\gamma''+1},
\]

(9.14)

that is exactly the splitting described just below (9.7), where we also pointed out the need to compensate for tangled paths occurring in (9.13), which is precisely the role of \( R(t) \) in (9.8).

To bound (9.8), we introduce

\[
W = \frac{2}{\Phi(2)} \langle \rho \chi_1^* \chi_1^* + \mu_2 \chi_2 \chi_2^* \rangle = \langle \phi_{e2} \phi_{f1} W_{\sigma(e2)\sigma(f1)} \rangle_{ef},
\]

(9.16)

and,

\[
L = K^{(2)} - W.
\]

(9.17)

Note the presence of weights in (9.16), hence our choice for the candidate eigenvectors.

Further, we set for \( 1 \leq t \leq \ell - 1, \)

\[
S^{(t)} = \Delta^{(t-1)} L B^{(\ell-t-1)}.
\]

(9.18)

We then have:

**Proposition 9.1** (Degree-Corrected Extension of Proposition 13 in [2]). If \( G \) is tangle-free and \( x \in C^{L(V)} \) with norm smaller than one, we have

\[
\|B^x\| \leq \|\Delta^{(t)}\| + \frac{1}{n} \|KB^{(\ell-1)}\| + \frac{1}{n} \sum_{j=1,2} \frac{2\mu_j}{\Phi(2)} \sum_{t=1}^{\ell-1} \|\Delta^{(t-1)} \chi_j\| \|\langle \hat{x}_j, B^{(\ell-t-1)} x\rangle\| \\
+ \frac{1}{n} \sum_{t=1}^{\ell-1} \|S^{(t)}\| + \phi_{\max}^2(a \vee b) \|\Delta^{(\ell-1)}\| + \frac{1}{n} \sum_{t=0}^{\ell} \|R^{(t)}\|.
\]

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

Below we prove the following bounds on the matrices in Proposition 9.1:

\]

35
Proposition 9.2 (Degree-Corrected Extension of Proposition 14 in [2]). Let \( \ell = C \log_p n \) with \( C < 1 \). With high probability, the following norm bounds hold for all \( k \), \( 0 \leq k \leq \ell \), and \( i = 1, 2 \):

\[
\| \Delta^{(k)} \| \leq (\log n)^{10} \rho^{k/2}, \\
\| \Delta^{(k)} \chi_i \| \leq (\log n)^{5} \rho^{k/2} \sqrt{n}, \\
\| B_k^{(i)} \| \leq (\log n)^{25} \rho^{\ell-k/2}, \\
\| KB_k^{(i)} \| \leq \sqrt{n}(\log n)^{10} \rho, \\
\text{and the following bound holds for all } k, 1 \leq k \leq \ell - 1:\n\| S_k^{(i)} \| \leq \sqrt{n}(\log n)^{20} \rho^{\ell-k/2}.
\]

9.1 Proof of Proposition 4.2

From Propositions 9.1 and 9.2, the geometric growth in Corollary 6.7 together with the tangle-freeness due to Lemma 6.5, the proof of Proposition 4.2 follows:

Let \( j \in \{1, 2\} \). If, for some vector \( x \), \( (\hat{\varphi}_j, x) = 0 \), then \( (B_j^t \chi_j, \hat{x}) = 0 \). Therefore, using Corollary 6.7,

\[
\sup_{\| x \|=1, (\varphi_j, x)=0} \langle \hat{\chi}_j, B^t \rangle \leq \sup_{\| x \|=1, (B^t \chi_j, \hat{x})=0} \langle B^t \chi_j, \hat{x} \rangle = \sup_{\| x \|=1, (B^t \chi_j, \hat{x})=0} \langle B^t \chi_j, \hat{x} \rangle \leq \log^2(n) n^{1/2} \rho^{-\ell-1}.
\]

With high probability, the graph is \( \ell \)-tangle free (Lemma 6.5). Thus, invoking Propositions 9.1 and 9.2, with high probability,

\[
\sup_{x \in \mathcal{H}^+, \| x \|=1} \| B^t x \| \leq \log^{10}(n) \rho^{\frac{\ell}{2}} + n^{-1/2} \log^{10}(n) \rho^{\ell-1} + c_1 \log^8(n) \rho^{\frac{\ell}{2}} + n^{-1/2} \log^{21}(n) \rho^{\ell} + c_2 \log^{10}(n) \rho^{\frac{\ell}{2}} + n^{-1} \log^{26}(n) \rho^{\ell} \leq \log^\epsilon(n) \rho^{\frac{\ell}{2}},
\]

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 9.2. In particular, we bound the expectation of expressions of the form

\[
\mathbb{E} \left[ \prod_{i=1}^{2m} \prod_{s=1}^{k} \delta_{\gamma_i,s-1\gamma_i,s} \right],
\]

for certain paths \( \gamma = (\gamma_1, \ldots, \gamma_{2m}) \) with \( \gamma_i = (\gamma_{i,0}, \ldots, \gamma_{i,k}) \in V^{k+1} \), where \( \delta \) is defined in (9.4).

In bounding (9.26) the following term occurs:

\[
\prod_{u \in V(\gamma)} \Phi(x_u)
\]

36
where \((d_u)\) are the degrees of the vertices in a specific tree (or forest) spanning the path \(\gamma\). See, for instance, (9.30) and (9.43) below. Here lies the main difference with the Stochastic Block Model: those terms are not present in the latter model. In (9.34) and (9.45) we find

\[
\prod_{u=1}^{\vert V(\gamma)\vert} \Phi^{(d_u)} \leq C_2 \sum_{u:d_u>2} (d_u-2) \Phi^{2(d_u-2)},
\]

where \(C_2 > 1\) is some constant and where \(n_c \geq 1\) is the number of components on the path \(\gamma\). To compare this term with powers of \(\Phi^{(2)}\) (which are present in powers of \(\rho = \frac{\Phi^{(2)}}{2}\)), we bound \(\sum_{u:d_u>2} (d_u-2)\), see in particular Lemma 9.4 and Lemma 9.7.

### 9.3 Bound on \(\|\Delta^{(k)}\|\)

We set

\[
m = \left\lfloor \frac{\log n}{13 \log(\log n)} \right\rfloor.
\]

We bound the norm of \(\|\Delta^{(k)}\|\) by using the trace method. Following (36) in [2] (which remains true for the DC-SBM), we obtain

\[
\|\Delta^{(k-1)}\|^{2m} \leq \sum_{\gamma \in W_{k,m}} \prod_{i=1}^{2m} \prod_{s=1}^{k} A_{\gamma_{i,s-1}\gamma_{i,s}},
\]

(9.27)

where \(W_{k,m}\) is the collection containing all sequences of paths \(\gamma = (\gamma_1, \ldots, \gamma_{2m})\) such that for all \(i:\)

- \(\gamma_i = (\gamma_i, 0, \ldots, \gamma_i, k) \in V^{k+1}\) is a non-backtracking tangle-free path of length \(k\), and
- \((\gamma_{i,k-1}, \gamma_i, k) = (\gamma_{i+1,1}, \gamma_{i+1,0})\),

where we put \(\gamma_0 = (\gamma_{2m})\).

Recall the notation \(G(\gamma) = (V(\gamma), E(\gamma))\). Further introduce the notation \(E_\phi(\cdot) = E[\cdot|\phi_1, \ldots, \phi_n]\). We bound, for a given \(\gamma \in W_{k,m}\),

\[
E_\phi \left( \prod_{i=1}^{2m} \prod_{s=1}^{k} A_{\gamma_{i,s-1}\gamma_{i,s}} \right) = \prod_{e \in E(\gamma)} E_\phi \left( A_{e_{1\gamma}e_{2\gamma}}^{(\gamma)} \right),
\]

(9.28)

where for \(e \in E(\gamma)\), \(p_{e_{1\gamma}e_{2\gamma}}^{(\gamma)}\) denotes the number of times the edge \(e\) is traversed on the walk \(\gamma\). In (9.28) we used that \(A\) is symmetric and that, conditional on the weights, edges are independently present. Note that for any edge \(uw\), and integer \(p\),

\[
E_\phi \left( A_{uw}^{p} \right) \leq \phi_u \phi_w \frac{W_{\sigma(u)\sigma(w)}}{n}.
\]

Below in Lemma 9.4, we construct a spanning tree \(T(\gamma) = (V(\gamma), E_T(\gamma))\) of \(\gamma\). In particular, for the \(e - (v - 1)\) edges not present in \(T\), we have \(\phi_u \phi_w \frac{W_{\sigma(u)\sigma(w)}}{n} \leq \frac{a}{n}\), with \(c_1 = \phi_{\max}^n (a \lor b)\). Putting this into (9.28), we get

\[
\prod_{e \in E(\gamma)} E_\phi \left( A_{e_{1\gamma}e_{2\gamma}}^{(\gamma)} \right) \leq (c_1/n)^{v+1} \prod_{e \in E_T(\gamma)} \phi_{e_{1\gamma}} \phi_{e_{2\gamma}} \frac{W_{\sigma(e_1)\sigma(e_2)}}{n},
\]

(9.29)

\[
= (c_1/n)^{v+1} \prod_{u \in V(\gamma)} \phi_u^{d_u} \prod_{e \in E_T(\gamma)} \frac{W_{\sigma(e_1)\sigma(e_2)}}{n},
\]

37
where \( d_u \) is the degree of \( u \) in the spanning tree. Consequently,

\[
\mathbb{E} \left( \prod_{e \in E(\gamma)} A_{e}^{(\tau)} \right) \leq (c/n)^{\nu - 1} \prod_{v \in V(\gamma)} \Phi(d_v) \prod_{\tau \in E_T(\gamma)} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n}.
\]  

(9.30)

Let \( \tau : [v(\gamma)] \mapsto V(\gamma) \) be the bijection describing the order the vertices are visited for the first time. I.e., for \( 1 \leq u \leq v(\gamma) - 1 \), \( \tau(u) \) is seen for the first time, before \( \tau(u + 1) \).

We shall say that a path \( \gamma_c \) is canonical if \( V(\gamma_c) = [v(\gamma_c)] \) and the vertices are first visited in the order \( 1, \ldots, v(\gamma_c) \). With every path \( \gamma \) there corresponds (through the bijection \( \tau \)) a canonical path \( \gamma_c \). Consequently, if \( W_{k,m}(v, e) \) denotes the set of canonical paths in \( W_{k,m} \) with \( v \) vertices and \( e \) edges, and \( I_{\gamma_c} \) the set of all injections from \([v(\gamma_c)]\) to \([n]\),

\[
\mathbb{E} \left[ \sum_{\gamma \in W_{k,m}} \prod_{i=1}^{2m} \prod_{t=1}^{k} A_{\sigma(\tau_i)\sigma(\tau_2)} \right] \leq \sum_{v=3}^{k+1} \sum_{e=v-1}^{km} \sum_{\gamma \in W_{k,m}(v, e)} \sum_{\tau \in I_{\gamma_c}} \mathbb{E} \left[ \prod_{e \in E(\gamma)} A_{e}^{(\tau)} \right].
\]  

(9.31)

because any non-backtracking path has at least 3 vertices, and \( v - 1 \leq e \leq km \), since (9.28) is non-zero only if each edge is traversed at least twice.

We now bound the term \( \sum_{\tau \in I_{\gamma_c}} \mathbb{E} \left[ \prod_{e \in E(\gamma)} A_{e}^{(\tau)} \right] \) in (9.31). Using (9.30), we have,

\[
\sum_{\tau \in I_{\gamma_c}} \mathbb{E} \left[ \prod_{e \in E(\gamma)} A_{e}^{(\tau)} \right] \leq \left( c_1/n \right)^{\nu - 1} \sum_{u=1}^{v(\gamma_c)} \Phi(d_u) \prod_{\tau \in I_{\gamma_c}} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n}
\]  

(9.32)

Our objective is to compare \( \sum_{u=1}^{v(\gamma_c)} \Phi(d_u) \prod_{\tau \in I_{\gamma_c}} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n} \) with \( n^o(\nu - 1) \).

We start by analysing the term containing the spins:

**Lemma 9.3.** For any canonical path \( \gamma_c \in W_{k,m} \),

\[
\sum_{\tau \in I_{\gamma_c}} \prod_{e \in E_T(\gamma_c)} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n} \leq (1 + o(1))n \left( \frac{a + b}{2} \right)^{\nu - 1}.
\]  

(9.33)

**Proof.** Let \( l \) be any leaf on the tree with unique neighbour \( g \). Then, writing \( \tau_u = \tau(u) \) for \( u \in \{1, \ldots, v\} \),

\[
\sum_{\tau \in I_{\gamma_c}} \prod_{e \in E_T(\gamma_c)} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n} \leq \prod_{\tau_1 = 1}^{n} \prod_{\tau_2 = 1}^{n} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n}.
\]

Keeping \( \tau_g \) fixed,

\[
\prod_{\tau_1 = 1}^{n} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n} = \prod_{e \in E_T(\gamma_c) \setminus \{g,l\}} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n} \sum_{\tau_1 = 1}^{n} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n}
\]

\[
= \prod_{e \in E_T(\gamma_c) \setminus \{g,l\}} \frac{W_{a(\tau_1)\sigma(\tau_2)}}{n} \left( \frac{a + b}{2} + O(n^{-\gamma}) \right),
\]

due to assumption (2.1).

Repeating inductively this procedure (by removing leaves from the tree) proves the assertion.  \( \square \)
It remains to bound $\prod_{u=1}^{v(\gamma)} \Phi^{(d_u)}$. To do so, we note that, since the weights are assumed to be bounded,

$$\Phi^{(d_u)} \leq C_2^{-2d_u-2} \Phi^{(2)} \left( \Phi^{(1)} \right)^{d_u-2}$$

if $d_u \geq 2$, with $C_2 = \frac{\Delta_{\max}}{\Phi^{(1)}} > 1$. Consequently,

$$\prod_{u=1}^{v(\gamma)} \Phi^{(d_u)} \leq C_2^{\sum_{u,d_u \geq 2} (d_u-2)} \prod_{u,d_u > 2} \Phi^{(2)} \left( \Phi^{(1)} \right)^{d_u-2} \prod_{u,d_u \leq 2} \Phi^{(d_u)}$$

$$\leq C_2^{\sum_{u,d_u > 2} (d_u-2)} 
\left( \Phi^{(2)} \right)^{\frac{1}{2}} \prod_{u=1}^{v} d_u$$

$$= C_2^{\sum_{u,d_u > 2} (d_u-2)} \left( \Phi^{(2)} \right)^{v-1},$$

(9.34)

where we used that by Jensen’s inequality $\left( \Phi^{(1)} \right)^2 \leq \Phi^{(2)}$.

Now, the sum $\sum_{u,d_u > 2} (d_u - 2)$ is small for a tree spanning a path in $W_{k,m}$:

**Lemma 9.4.** For any $\gamma \in W_{k,m}$, with $v$ vertices and $e$ edges, there exists a tree spanning $\gamma$ with degrees $(d_u)_{u=1}^{v}$ such that:

$$\sum_{u,d_u > 2} (d_u - 2) \leq e - (v - 1) + 2m.$$  

(9.35)

**Proof.** We construct a spanning tree, while traversing $\gamma$. We denote by $p(t)$ the graph constructed at step $t \geq 0$. Put $p(0) = \{\gamma_1, 0, 2\}$ and $r = s = 0$ (the meaning of these two counters becomes clear in the algorithm below). Consider edge $f$ traversed in step $t + 1$ of the walk: If $f$ or $f$ has already been traversed, then continue with step $t + 2$. Otherwise, if both $f$ and $f$ have not yet been traversed, distinguish between the following cases:

1. $f_1$ is a leaf of $p(t)$ and
   a) $p(t)$ contains a cycle, then if $f_2 \notin p(t)$, put $p(t + 1) = p(t) \cup f$, otherwise, if $f_2 \in p(t)$, put $p(t + 1) = p(t)$;
   b) $p(t)$ does not contain a cycle, then put $p(t + 1) = p(t) \cup f$. If $f_2 \notin p(t)$, then put $e_c = f$;

2. $f_1$ is not a leaf of $p(t)$ and
   a) $p(t)$ contains a cycle, then put $p(t + 1) = (p(t) \setminus e_c) \cup f$. If $f_2 \notin p(t)$, put $e_c = f$. Increase the value of $r$ with one.
   b) $p(t)$ does not contain a cycle, then put $p(t + 1) = p(t) \cup f$. If $f_2 \notin p(t)$, put $e_c = f$. Otherwise, if $f_2 \notin p(t)$, increase the value of $s$ with one.

Once the path is completely traversed, remove $e_c$ to obtain a spanning tree.

Note that at each stage of the construction, the graph contains at most one cycle and in this case, removing $e_c$ will make the graph into a tree.

Further, cases 1.a and 1.b do not contribute to $\sum_{u,d_u > 2} (d_u - 2)$, since the leave in $p(t + 1)$ becomes a vertex with degree 2 in $p(t)$. A cycle formed in step $t + 1$ will temporarily increase the degree of the vertex that is merged by the leaf, however this edge $e_c$ will later be removed.

In case 2.a, the degree of vertex $f_1$ increases with one, however, at the same time an edge is removed. The number of times 2.a happens, $r$, is thus bounded by the number of times an edge is removed: $r \leq e - (v - 1)$.

In case 2.b, we need only to consider the case where no cycle is formed. But, before arriving at such a vertex considered in 2.b, the path must have made a backtrack. Hence
s \leq 2m$. (In fact, between two subsequent occurrences of event 2, the walk should at least either make a backtrack or 'get back to the tree' by forming a cycle: giving the same bound for $s + r$).

All together,
\[ \sum_{u, d_u > 2} (d_u - 2) \leq r + s \leq e - (v - 1) + 2m. \]

Finally, we recall the bound on the cardinality of $W_k,m$ from [2]:

**Lemma 9.5** (Lemma 17 in [2]). Let $W_{k,m}(v,e)$ be the set of canonical paths with $v(\gamma) = v$ and $e(\gamma) = e$. We have
\[ |W_{k,m}(v,e)| \leq k^2m(2km)^{6m(e-v+1)}. \] (9.36)

Hence, combining (9.27), (9.31) - (9.36),
\begin{align*}
\mathbb{E} \left[ \| \Delta^{(k-1)} \|^{2m} \right] &\leq \sum_{v=3}^{km+1} \sum_{e,v-1}^{km} |W_{k,m}(v,e)| \left( \frac{e}{n} \right)^{e-(v-1)} nC^{e-(v-1)+2m} \rho^{e-v-1} \\
&\leq n c_{v} \rho^{km} \sum_{v=3}^{km+1} \sum_{e,v-1}^{km} \ell^{2m} \left( \frac{c_{v}(2e\ell m)^{6m}}{n} \right)^{e-(v-1)} \\
&\leq n c_{v} \rho^{km} \ell^{2m} \sum_{s=0}^{\infty} \left( \frac{c_{v}(2e\ell m)^{6m}}{n} \right)^{s} \\
&\leq n(c_{v} \log n)^{m} \log^{2} n \rho^{km} \\
&\leq (c_{v} \log n)^{16m} \rho^{km},
\end{align*}
(9.37)

where we used the bound on $m$, in particular to derive convergence of the series, and the fact that $n^{1/m} = o(\log n)^{14}$.

We finish by using Markov's inequality.

### 9.4 Bound on $\| \Delta^{(k)} \chi_i \|$ 

We point out the differences with bound (31) in [2]: Here, we have
\begin{align*}
\mathbb{E} \left[ \| \Delta^{(k-1)} \chi_i \|^2 \right] &= \mathbb{E} \left[ \sum_{e,f,g} \Delta^{(k-1)}_{e,f} \Delta^{(k-1)}_{e,g} \xi_i(f) \xi_i(g) \right] \\
&\leq \phi_{\text{max}}^2 \mathbb{E} \left[ \sum_{e,f,g} \Delta^{(k-1)}_{e,f} \Delta^{(k-1)}_{e,g} \right] \\
&\leq \phi_{\text{max}}^2 \sum_{\gamma \in \mathcal{W}^\prime_{k,1}} \mathbb{E} \left[ \prod_{i=1}^{2} \prod_{s=1}^{k} \Delta^{(r_{i,s-1},r_{i,s})} \right],
\end{align*}
(9.38)

where $\mathcal{W}^\prime_{k,1}$ is defined in [2]. In the latter paper it is also shown that the same bound, Lemma 9.5 holds for the cardinality of $W_{k,1}''$. Hence, using the penultimate line of (9.37) with $m = 1$, gives
\[ \mathbb{E} \left[ \| \Delta^{(k-1)} \chi_i \|^2 \right] \leq c_{v} n \log^{3}(n) \rho^{k}. \]
9.5 Bound on $\|R_k^{(\ell)}\|

Put

$$m = \left\lfloor \log n \over 25 \log(\log n) \right\rfloor.$$  

We apply the same strategy as above: for $0 \leq k \leq \ell - 1$, we have the bound

$$\|R_k^{(\ell-1)}\|^{2m} \leq \text{tr} \left\{ \left( R_k^{(\ell-1)} R_k^{(\ell-1)*} \right)^m \right\}$$

$$= \sum_{\gamma \in T_{x,m,k}} \prod_{l=1}^{m} \prod_{s=1}^{k} A_{\gamma_i,s-1,\gamma_i,s} \phi_{\gamma_i,k} \phi_{\gamma_i,k+1} W_{\sigma(\gamma_i,k)\sigma(\gamma_i,k+1)} \prod_{s=k+2}^{\ell} A_{\gamma_i,s-1,\gamma_i,s}$$

$$\leq c_1^m \sum_{\gamma \in T_{x,m,k}} \prod_{l=1}^{m} \prod_{s=1}^{k} A_{\gamma_i,s-1,\gamma_i,s} \prod_{s=k+2}^{\ell} A_{\gamma_i,s-1,\gamma_i,s},$$

(9.39)

where $c_1 = \phi_{\text{max}}(a \lor b)^2$, and where $T_{x,m,k}$ is the collection containing all sequences of paths $\gamma = (\gamma_1, \ldots, \gamma_{2m})$ such that

- for all $i$: $\gamma_i = (\gamma_1^i, \gamma_2^i)$, where $\gamma_1^i = (\gamma_i, 0, \cdots, \gamma_i,k)$ and $\gamma_2^i = (\gamma_i,k+1, \cdots, \gamma_i,\ell)$ are non-backtracking tangle-free;

- for all odd $i$: $(\gamma_i,0, \gamma_i,1) = (\gamma_i,0, \gamma_i,1)$ and $(\gamma_i,\ell-1, \gamma_i,\ell) = (\gamma_i,\ell-1, \gamma_i,\ell)$, with the convention that $\gamma_0 = \gamma_{2m}$.

To calculate the expectation of $\|R_k^{(\ell-1)}\|^{2m}$, we note that

$$E \left[ \prod_{l=1}^{m} \prod_{s=1}^{k} A_{\gamma_i,s-1,\gamma_i,s} \prod_{s=k+2}^{\ell} A_{\gamma_i,s-1,\gamma_i,s} \right]$$

is non-zero only if (for $i$ fixed) each edge $(\gamma_i,s-1, \gamma_i,s)$ for $1 \leq s \leq k$ appears more than once in the $2(\ell-1)m$ pairs $\{(\gamma_j,s-1, \gamma_j,s)\}_{j=1,s \neq k+1}^{j=2m}$. Hence,

$$E \left[ \|R_k^{(\ell-1)}\|^{2m} \right] \leq c_1^m \sum_{\gamma \in T_{x,m,k}} E \left[ \prod_{l=1}^{m} \prod_{s=1}^{k} A_{\gamma_i,s-1,\gamma_i,s} \prod_{s=k+2}^{\ell} A_{\gamma_i,s-1,\gamma_i,s} \right],$$

(9.40)

where

$$T_{x,m,k} = \{ \gamma \in T_{x,m,k} \mid v(\gamma) \leq c(\gamma) \leq km + 2m(\ell - 1 - k) \}.$$  

(9.41)

Similarly as in establishing the bound on $\|\Delta_k^{(b)}\|$, we say that a path $\gamma_c$ is canonical if $V(\gamma_c) = (v(\gamma_c))$ and the vertices are first visited in order. We denote by $T_{x,m,k}(v,e)$ the set of canonical paths in $T_{x,m,k}(v,e)$ with $v$ vertices and $e$ edges. Then:

$$E \left[ \|R_k^{(\ell-1)}\|^{2m} \right] \leq c_1^m \sum_{v=1}^{m(2\ell-2-k)} \sum_{e=1}^{m(2\ell-2-k)} \sum_{\gamma_c \in I_{v,e}} E \left[ \prod_{e \in E(\gamma_c)} A_{\gamma_c}^{(\gamma_c)} \prod_{(v(e), e)} A_{\gamma_c}^{(\gamma_c)} \right],$$

(9.42)

where $I_{v,e}$ is defined as above, $I_{v,e}^{(\gamma)}(\gamma)$ is the number of times edge $\{e_1, e_2\}$ occurs in $\{\{\gamma_j,s=1, \gamma_j,s\}\}_{s=1,j=1}^{s=2m}$, and $P_{v,e}\gamma_c$ denotes the number of times edge $\{e_1, e_2\}$ occurs in the remainder of the collection of edges, $\{\{\gamma_j,s=1, \gamma_j,s\}\}_{s=k+2,j=1}^{s=2m}$.

Now, again,

$$E \left[ \prod_{e \in E(\gamma)} A_{\gamma_c}^{(\gamma_c)} \right] \leq \phi_{\gamma_c}^{(\gamma)} \frac{W_{\sigma(\gamma_c)\sigma(\gamma_c)} n}{\phi_{\gamma_c}^{(\gamma)}}.$$

Below we construct a spanning forest $F = (V(\gamma), E_F(\gamma))$ of $\gamma$ (i.e., $F$ is the disjoint union of trees, each spanning another component of $G(\gamma)$).
Let \( n_C \leq m \) denote the number of components of \( G(\gamma) \). Then,

\[
\mathbb{E} \left[ \prod_{\tau \in \mathcal{E}(\gamma)} \mathbb{P}_{\mathcal{E}(\gamma)}^{(\gamma)}(\tau | \gamma_{c}) \right] \leq \left( \frac{e}{n} \right)^{e-(v-n_C)} \prod_{u \in V(\gamma)} \Phi^{(d_u)} \prod_{e \in E(\gamma)} \frac{W_{\sigma(\epsilon_1)\sigma(\epsilon_2)}}{n},
\]

(9.43)

with \( d_u \) the degree of vertex \( u \) in the forest \( F \), compare to (9.30).

Now, this time,

**Lemma 9.6.** For any canonical path \( \gamma_c \in \mathcal{T}_{k,n,m}(v,e) \),

\[
\sum_{\tau \in \mathcal{T}_{k,n,m}(\gamma_c)} \mathbb{E} \left[ \prod_{\tau \in \mathcal{E}(\gamma_c)} \mathbb{P}_{\mathcal{E}(\gamma_c)}^{(\gamma_c)}(\tau | \gamma_{c}) \right] \leq (1 + o(1))n^n \left( \frac{a+b}{2} \right)^{v-n_C}.
\]

(9.44)

**Proof.** Apply Lemma 9.3 subsequently to the different components of \( F \).

Further, applying (9.34) to different components in \( F \) gives

\[
\prod_{u=1}^{v(\gamma_c)} \Phi^{(d_u)} \leq C_2 \sum_{u=d_u \geq 2} (d_u-2) \left( \Phi(2) \right)^{v-n_C}.
\]

(9.45)

Together,

\[
\sum_{\tau \in \mathcal{T}_{k,n,m}(\gamma_c)} \mathbb{E} \left[ \prod_{\tau \in \mathcal{E}(\gamma_c)} \mathbb{P}_{\mathcal{E}(\gamma_c)}^{(\gamma_c)}(\tau | \gamma_{c}) \right] \leq \left( \frac{e}{n} \right)^{e-(v-n_C)} C_2 \sum_{u=d_u \geq 2} (d_u-2) \left( \Phi(2) \right)^{v-n_C}.
\]

(9.46)

Again, we bound \( \sum_{u=d_u \geq 2} (d_u-2) \):

**Lemma 9.7.** For any \( \gamma \in \mathcal{T}_{k,n,m} \), with \( v \) vertices and \( e \) edges, there exists a forest spanning \( \gamma \) with degrees \( (d_u)_{u=1}^{v(\gamma)} \) such that:

\[
\sum_{u=d_u \geq 2} (d_u-2) \leq 18m + e - (v - n_C).
\]

(9.47)

**Proof.** As in Lemma 9.3, we construct the spanning forest, while traversing \( \gamma \). Again \( p(t) \) denotes the graph constructed at step \( t \geq 0 \), with \( p(0) = \{ \gamma_1, \emptyset \} \). Further, we introduce three counters: \( r = s = q = 0 \), together with \( e_c = \emptyset \) (below, \( e_c \) is either equal to \( \emptyset \) or it is an edge such that \( p(t) \) contains one cycle, but \( p(t) \setminus e C \) is a forest). At any step \( t \), we let \( C_1, \ldots, C_{\#\text{components}} \) be the components of \( p(t) \).

Consider step \( t + 1 \) of the walk: if the step consists in jumping to a vertex \( w \), then put \( p(t+1) = (p(t) \setminus e C) \cup \{ w \} \).

Else, if the step consists in traversing an edge \( f = f_1f_2 \), then: If \( f \) or \( \tilde{f} \) has already been traversed, continue with step \( t+2 \). Otherwise, if both \( f \) and \( f \) have not yet been traversed, distinguish between the following cases:

1. \( f_1 \) is a leave or an isolated vertex of component \( C_i \) of \( p(t) \) and
   a. \( C_i \) does not contain a cycle, then put \( p(t+1) = p(t) \cup f \). Further, distinguish between the following cases:
      i) \( f_2 \notin p(t) \);
      ii) \( f_2 \in C_i \), then put \( e_c = f \);
      iii) \( f_2 \notin C_i \), then increase the value of \( s \) with one.
   b. \( C_i \) contains a cycle, then distinguish between the following cases:
      i) \( f_2 \notin p(t) \), then put \( p(t+1) = p(t) \cup f \);
      ii) \( f_2 \in C_i \), then put \( p(t+1) = p(t) \);
iii) \( f_2 \in C_{j \neq i} \), then put \( p(t + 1) = p(t) \cup f \) and increase the value of \( s \) with one.

2. \( f_1 \) in component \( C_i \) has degree at least 2 in \( p(t) \), then distinguish between the following cases:
   a. \( C_i \) does not contain a cycle, then put \( p(t + 1) = p(t) \cup f \). Further, distinguish between the following cases:
      i) \( f_2 \not\in p(t) \), then increase the value of \( q \) with one;
      ii) \( f_2 \in C_i \), then put \( e_c = f \);
      iii) \( f_2 \in C_{j \neq i} \), then increase the value of \( s \) with two.
   b. \( C_i \) contains a cycle, then put \( p(t + 1) = (p(t) \setminus e_c) \cup f \). Further, distinguish between the following cases:
      i) \( f_2 \not\in p(t) \), then increase the value of \( r \) with one;
      ii) \( f_2 \in C_i \), then put \( e_c = f \);
      iii) \( f_2 \in C_{j \neq i} \), then increase the value of \( s \) with two.

Once the path is completely traversed, remove \( e_c \) to obtain a spanning tree.

The only cases that contribute to \( \sum_{u : d_u \geq 2} (d_u - 2) \) are 1.a.iii, 1.b.iii, 2.a.i, 2.a.iii, 2.b.i and 2.b.iii.

Now, \( s \) counts the contribution of 1.a.iii, 1.b.iii, 2.a.iii and 2.b.iii. But, in all those 4 cases, two components are merged, hence \( s \leq 6 \#\text{merges} \leq 12m \).

By definition of the event 2.b.i, \( r \) is an upper bound for the number of edges that are removed: \( r \leq e - (v - n_c) \).

To bound \( q \) (which counts the occurrence of 2.a.i), note that between two subsequent occurrences of the event 2.a.i, the walk makes at least one of the following: a backtrack, a jump or a merge. Hence \( q \leq 2m + 2m + 2m = 6m \).

Adding the bounds for \( r, q \) and \( s \) establishes (9.47).

Returning to (9.46), we get, since \( n_c \leq 2m \):

\[
\sum_{r \in T_{\gamma}} \mathbb{E} \left[ \prod_{e \in E(\gamma_c)} \mathcal{A}_{\gamma_c}(e_1)^{\frac{m(\gamma_c)}{2}} \mathcal{A}_{\gamma_c}(e_2)^{\frac{m(\gamma_c)}{2}} \right] \leq (c_1/n)^{e-v} C_2^{18m + e-v + n_c} \rho^{v-n_c} \\
\leq \left( \frac{c_3}{n} \right)^{e-v} c_4^m \rho^{v-n_c}.
\] (9.48)

Putting this into (9.42), we obtain

\[
\mathbb{E} \left[ \|R_k^{(t-1)}\|^{2m} \right] \leq c_5^m \sum_{v=1}^{m} \sum_{e=v}^{m} \sum_{\gamma \in T_{\ell,m,k}(v,e)} \left( \frac{c_3}{n} \right)^{e-v} \rho^{m(2t-k)}.
\] (9.49)

Now the cardinality of \( T_{\ell,m,k}(v,e) \) is bounded in the following lemma:

**Lemma 9.8** (Lemma 18 in [2]). Let \( T_{\ell,m,k}(v,e) \) be the set of canonical paths in \( T_{\ell,m,k} \) with \( v(\gamma) = v \) and \( e(\gamma) = e \). We have

\[
|T_{\ell,m,k}(v,e)| \leq (4fm)^{12m(e-v+1)+8m}.
\]
Hence,
\[
\mathbb{E} \left[ \| R_k^{(l-1)} \|^2 \right] \leq c_2^m \rho^{m(2l-k)} \sum_{v=1}^{m(2l-2-k)} \sum_{c=v}^{m(2l-2-k)} (4\ell m)^{12m(c-v+1)+8m} \left( \frac{e_3}{n} \right)^{c-v}
\]
\[
\leq \rho^{m(2l-k)} c_3 (4\ell m)^{20m} \sum_{v=1}^{m(2l-2-k)} \sum_{s=0}^{\infty} \left( \frac{e_3 (4\ell m)^{12m}}{n} \right)^s
\]
\[
\leq \rho^{m(2l-k)} c_3 (4\ell m)^{20m} 2\ell m \cdot \mathcal{O}(1)
\]
\[
\leq \rho^{m(2l-k)} (c_5 \log(n))^{42m}.
\]

We used that, due to our choice of \( m, (4\ell m)^{12m} \leq n^{24/25} \).

We use (9.50) together with Markov's inequality:
\[
P \left( \| R_k^{(l)} \| > (\log(n))^{25} \rho^{l-k/2} \right) \leq \mathbb{E} \left[ \| R_k^{(l)} \|^2 \right] / (\log(n))^{50m} \rho^{m(2l-k)}
\]
\[
\leq (c_6 \log(n))^{-8m} \to 0.
\]

### 9.6 Bound \( \| KB^{(k)} \| \)

Put
\[
m = \left\lfloor \frac{\log n}{13 \log(\log n)} \right\rfloor.
\]

We have, with the convention that \( e_{2m+1} = e_1 \),
\[
\| KB^{(k-2)} \|^2 \leq \text{tr} \{ (KB^{(k-2)} KB^{(k-2)})^m \}
\]
\[
= \sum_{e_1, \ldots, e_{2m}} \prod_{i=1}^{m} (KB^{(k-2)})_{e_{2i-1}, e_{2i}} (KB^{(k-2)})_{e_{2i+1}, e_{2i}}.
\]

Now,
\[
(KB^{(k-2)})_{ef} = \sum_g K_{eg} E_g^{(k-2)}
\]
\[
= \sum_g 1_{e \to g} \phi_{\gamma_1} \phi_{\gamma_2} W_{g(e_1)} W_{g(e_2)} \prod_{s=0}^{k-2} A_{\gamma_s, \gamma_{s+1}}
\]
\[
\leq c_1 \sum_g 1_{e \to g} \prod_{s=0}^{k-2} A_{\gamma_s, \gamma_{s+1}}.
\]

Hence,
\[
\| KB^{(k-2)} \|^2 \leq c_2^m \sum_{e_1, \ldots, e_{2m}} \prod_{i=1}^{m} \left( \sum_g 1_{e_{2i-1} \to g} \prod_{s=0}^{k-2} A_{\gamma_s, \gamma_{s+1}} \right) \left( \sum_g 1_{e_{2i+1} \to g} \prod_{s=0}^{k-2} A_{\gamma_s, \gamma_{s+1}} \right)
\]
\[
= c_2^m \sum_{\gamma \in \mathcal{W}_{k,m}} \prod_{s=1}^{k-1} \prod_{s=1}^{k} A_{\gamma_s, \gamma_{s+1}}.
\]

where \( \mathcal{W}_{k,m} \) is the collection containing all sequences of paths \( \gamma = (\gamma_1, \ldots, \gamma_{2m}) \) with \( \gamma_s = (\gamma_{i,0}, \ldots, \gamma_{i,k}) \in V^{k+1} \) is non-backtracking such that
• for all i: \((\gamma_{i+1}, \gamma_{i+1}) = (\gamma_{i+1}, \gamma_{i+1})\),
• for all odd i: \((\gamma_{i,1}, \cdots, \gamma_{i,k})\) is tangle-free,
• for all even i: \((\gamma_{i,0}, \cdots, \gamma_{i,k})\) is tangle-free,

with the convention that \(2m+1 = 1\).

Recall the definition of \(W_k,m\) and note that \(W_k,m \subset \overline{W}_k,m\). Fix \(\gamma \in \overline{W}_k,m \setminus W_k,m\) and consider

\[ S_\gamma := \{ \gamma \in \overline{W}_k,m \setminus W_k,m \} \]

odd \(i: (\gamma_{i,1}, \cdots, \gamma_{i,k}) = (\gamma_{i,1}, \cdots, \gamma_{i,k})\), even \(i: (\gamma_{i,0}, \cdots, \gamma_{i,k-1}) = (\gamma_{i,0}, \cdots, \gamma_{i,k-1})\).

Then \(|S_\gamma| \leq k^m\). Indeed, if for odd \(i\), \(\gamma_i\) is not tangle-free then necessarily \(\gamma_i,0 \in \{\gamma_{i,1}, \cdots, \gamma_{i,k}\}\), i.e., \(\gamma_i,0\) can be chosen in at most \(k\) different ways. A similar argument works in case \(i\) is even.

Now, there always exists \(\gamma \in W_k,m\) such that for all odd \(i: (\gamma_{i,1}, \cdots, \gamma_{i,k}) = (\gamma_{i,1}, \cdots, \gamma_{i,k})\) and for all even \(i: (\gamma_{i,0}, \cdots, \gamma_{i,k-1}) = (\gamma_{i,0}, \cdots, \gamma_{i,k-1})\).

As a consequence of these two observations, we have

\[ ||KB^{(k-2)}||_{2m} \leq c_2^m (1 + k^m) \sum_{i=1}^{m} \prod_{i=1}^{k} A_{\gamma_{2i-1},s-1,\gamma_{2i-1},s} \prod_{s=1}^{k-1} A_{\gamma_{2s},s-1,\gamma_{2s},s}. \]  \hspace{1cm} (9.56)

To proceed following the method used to bound \(\Delta^{(k)}\), note that the product in (9.56) is taken over a path, consisting of 2m non-backtracking tangle-free subpaths of length \(k-1\), that makes at most 2m backtracks. Hence Lemma’s 9.3 and 9.4 may be adapted to the current setting (for instance the right hand side of (9.35) becomes \(e - (v - m - 1) + 2m\), entailing

\[ \mathbb{E}[||KB^{(k-2)}||_{2m}] \leq c_2^m (1 + k^m) \sum_{v=3}^{2km+1} c_4^m \sum_{v=3}^{2km+1} \left| W_{k,m} \right| \left( \frac{c_5}{n} \right)^{e-(v-1)-m} c_4^m \left( \frac{c_6}{n} \right)^{e-(v-1)+2m} \rho^{v-1} \]

\[ \leq c_2^m (1 + k^m) n^{m+1} \sum_{v=3}^{2km+1} \left| W_{k,m} \right| \left( \frac{c_5}{n} \right)^{e-(v-1)} \rho^{v-1} \]

\[ \leq c_2^m (1 + k^m) n^{m+1} \rho^{2km} \left( \frac{c_7}{e} \right)^{2km} \left( \frac{c_8}{n} \right)^{m+1} \rho^{2km} \]

\[ \leq c_2^m (1 + k^m) n^{m+1} \rho^{2km} \leq (c_9 \log n)^{19m} n^m \rho^{2km}, \]  \hspace{1cm} (9.57)

where we used our choice for \(m\) several times. An appeal to Markov’s inequality finishes the proof.

9.7 Bound on \(\|S_k^{(l)}\|\)

This proof follows almost line-to-line the proof used in [2] to establish bound (34) there. We restrict ourselves here to the differences:

Observe that \(L_{ef} = 0\) unless \(e \rightarrow f\) does not hold, that is \(e = f, e \rightarrow f, f^{-1} \rightarrow e\) or \(e \rightarrow f^{-1}\), in which cases \(L_{ef} = -\phi_{e} \phi_{f} W_{\sigma(e),\sigma(f)}\). Hence, we have the decomposition

\(L = -I^* - K^*\),

where \((I^*)_ef = \delta_{e,f} \phi_{e} \phi_{f} W_{\sigma(e),\sigma(f)}\) and where \((K^*)_ef = \phi_{e} \phi_{f} W_{\sigma(e),\sigma(f)}\) if \(e \rightarrow f\), \(f^{-1} \rightarrow e\) or \(e \rightarrow f^{-1}\) and \((K^*)_ef = 0\) otherwise.

Thus

\[ \|S_k^{(l)}\| \leq \phi_{max}^2 (a \lor b) \left( \|\Delta^{(k-1)}\| \|B^{(l-k-1)}\| + \|\Delta^{(l-1)} K'\| \|B^{(l-k-1)}\| \right), \]

where \(K'\) is defined in [2]. The rest of the proof follows after applying the arguments used in [2] and following the procedure set out above to obtain the bound on \(KB^{(k)}\).
10 Detection: Proof of Theorem 2.2

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

**Lemma 10.1** (Special case of Lemma 40 in [2]). Assume that there exists a function $F : V \rightarrow \{0, 1\}$ 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 $f : \{+, -\} \rightarrow [0, 1]$ is such that $f(+) > f(-)$. Then, assigning to each vertex a label $\hat{\sigma}(v) = +$ if $F(v) = 1$ and $\hat{\sigma}(v) = -$ if $F(v) = 0$, yields asymptotically positive overlap with the true spins.

**Proof.** Since $\hat{\sigma}(v) = +$ if and only if $F(v) = 1$, it follows that

$$\frac{1}{n} \sum_{v=1}^{n} 1_{\sigma(v) = +} 1_{\hat{\sigma}(v) = \sigma(v)} = \frac{1}{n} \sum_{v=1}^{n} 1_{\sigma(v) = +} F(v) \rightarrow \frac{f(+)}{2},$$

and

$$\frac{1}{n} \sum_{v=1}^{n} 1_{\sigma(v) = -} 1_{\hat{\sigma}(v) = \sigma(v)} = \frac{1}{n} \sum_{v=1}^{n} 1_{\sigma(v) = -} (1 - F(v)) \rightarrow \frac{1 - f(-)}{2}.$$ 

Consequently,

$$\frac{1}{n} \sum_{v=1}^{n} 1_{\hat{\sigma}(v) = \sigma(v)} \rightarrow \frac{1 + f(+) - f(-)}{2} > \frac{1}{2},$$

because $f(+) > f(-)$ by assumption.

Recall the eigenvector $\xi_2$ from Theorem 2.1. Below we use the function $F : v \mapsto 1_{\sum_{v, e \in v} \xi_2(e) > \frac{\tau}{2 \rho}}$ or $F : v \mapsto 1_{\sum_{v, e \in v} \xi_2(e) \leq \frac{\tau}{2 \rho}}$ for some fixed parameter $\tau$. We verify also that $\xi_2$ is aligned with $P_{2, \ell}$. It is therefore useful to introduce the vector $I_\ell$, defined element-wise by

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

for $v \in V$.

Further, put

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

The following lemma shows that $I_\ell$ is correlated with the spins:

**Lemma 10.2** (Degree-Corrected Extension of Lemma 41 in [2]). Let $\ell = C \log_{\rho} n$ with $C < C_{\text{coupling}}$ 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 $L^2$,

$$\frac{1}{n} \sum_{v=1}^{n} 1_{\sigma(v) = i} I_\ell(v) \mu_2^{-2\ell - \bar{\xi}_{g2}(i) \geq t} \rightarrow \frac{1}{2} \mathbb{P}(Y_i \geq t).$$

**Proof.** We use Proposition 7.2 with

$$\tau(G, v) = 1_{\sigma(v) = i} I_\ell(v) \mu_2^{-2\ell - \bar{\xi}_{g2}(i) \geq t}.$$ 

Denote by $(T, o)$ the branching process defined in Section 5 where the root has spin $\sigma_o$ uniformly drawn from $\{+, -\}$. Denote the number of offspring of the root by $D$ and let $Q_\ell(v)$ be equal to $Q_{2, \ell}$ defined on the tree $T^v$ obtained after removing the subtree attached to $v$ from $T$. Then,

$$\tau(T, o) = 1_{\sigma_o = i} I_\ell \mu_2^{-2\ell - \bar{\xi}_{g2}(i) \geq t}. $$
where

\[ J_\ell = \sum_{v=1}^{D} Q_\ell(v) = (D - 1)Q_{2,\ell} - L_{2,\ell}. \]

with \( L_{2,\ell} \) defined in (5.46).

We need to calculate \( \lim_{r \to \infty} E[\tau(T,o)] \). To this end, we first show that, conditional on \( \sigma_o = i, \frac{J_\ell}{\mu^2} - \hat{c}g_2(i) \) converges in probability to some centered random variable \( \hat{Y}_i \).

We first calculate \( E_i[J_\ell|\phi_o] \), where \( E_i[\cdot] = E[\cdot|\sigma_o = i] \). Put \( r_o = \frac{4\delta}{\Phi(i)} \phi_o \), then

\[
E_i[J_\ell|\phi_o] = \sum_{n=0}^{\infty} \frac{\mu_2^2}{\mu_2} E_i[Q_{2,\ell}|D = n, \phi_o] P(D = n|\phi_o)
= \sum_{n=0}^{\infty} n E_i[Q_{2,\ell}|D = n, \phi_o] \frac{1}{n!} n^r_e^2 \rho_n e^{-r_o} \rho_n
= r_o \sum_{n=1}^{\infty} E_i[Q_{2,\ell}|D = n - 1, \phi_o] \frac{1}{(n - 1)!} n^r_e^2 \rho_n e^{-r_o} \rho_n
= r_o E_i[Q_{2,\ell}|\phi_o].
\]

Recall from Theorem 5.6 that, uniformly for all \( \psi_o \),

\[
E_i \left[ \frac{Q_{2,\ell}}{\mu_2^2} \bigg| \phi_o = \psi_o \right] \to \frac{\Phi(3)}{\Phi(2)} \frac{\rho}{\mu_2^2} - \rho^{2,\psi_o} g_2(i)
\]
as \( n \to \infty \). Hence, \( \sup_{n,\psi_o} E_i \left[ \frac{Q_{2,\ell}}{\mu_2^2} \bigg| \phi_o = \psi_o \right] < \infty \), so that we can apply Lebesgue’s dominated convergence theorem:

\[
\frac{E_i[J_{2,\ell}]}{\mu_2^2} = \mathbb{E} \left[ r_o E_i \left[ \frac{Q_{2,\ell}}{\mu_2^2} \bigg| \phi_o \right] \right] \to \hat{c}g_2(i),
\]
as \( n \to \infty \).

We now combine the right hand side of (10.2), (10.4), and Theorem 5.6 (and in particular (5.53) which implies that \( L_{2,\ell}/\mu^2 \to 0 \) as \( n \to \infty \)) to establish the claim that, conditional on \( \sigma_o = i, \frac{J_\ell}{\mu^2} - \hat{c}g_2(i) \) converges in probability to some centered random variable \( \hat{Y}_i \).

In particular, conditional on \( \sigma_o = i, \frac{J_\ell}{\mu^2} - \hat{c}g_2(i) \) converges in distribution to \( \hat{Y}_i \). So that, for \( t \) as in the statement,

\[
E[\tau(T,o)] = \frac{1}{2} P \left( \frac{J_\ell}{\mu^2} - \hat{c}g_2(i) \geq t \big| \sigma_o = i \right) \to \frac{1}{2} P \left( \hat{Y}_i \geq t \right),
\]
as \( n \to \infty \).

Finally, noting that the error term in Proposition 7.2 is \( O \left( n^{-\frac{2}{3} + \frac{4}{15}} \right) = o(1) \) finishes the proof.

Recall from 2.1 that the eigenvector \( \xi_2 \) is asymptotically aligned with

\[
\frac{B^fB^f\hat{\chi}_2}{\|B^fB^f\hat{\chi}_2\|},
\]

where \( \ell \sim \log_o(n) \). Hence, for some unknown sign \( \omega \), the vector \( \xi'_2 = \omega \xi_2 \) is asymptotically close to (10.5). From 8.1 we know that \( B^fB^f\hat{\chi}_2 \) and \( P_{2,\ell} \) are asymptotically close.

Consequently, properly renormalizing \( \xi'_2 \) will make it asymptotically close to \( P_{2,\ell} \), so that we can replace \( P_{2,\ell} \) in (10.1) by \( \xi'_2 \). That is, we set for \( v \in V \),

\[
I(v) = \sum_{e : e = v} s\sqrt{n}\xi'_2(e),
\]

47
with \( s = \sqrt{\frac{c}{2}} \) the limit in Proposition 7.4. Then, \( I \) and \( I_i/\mu_2^d \) are close, which leads to the following Lemma:

**Lemma 10.3** (Degree-Corrected Extension of Lemma 42 in [2]). Let \( i \in \{+,-\} \) and \( \hat{Y}_i \) be as in Lemma 10.2. For any continuity point \( t \) of the distribution of \( \hat{Y}_i \), in \( L^2 \),

\[
\frac{1}{n} \sum_{v=1}^{n} 1_{\sigma(v)=i} 1_{I(v)-\hat{g}_2(i) \geq t} \to \frac{1}{2} \mathbb{P} \left( \hat{Y}_i \geq t \right).
\]

**Proof.** This follows after repeating the proof in [2] in conjunction with Lemma 10.2 established here.

Put for \( i \in \{+,-\} \), \( X_i = \hat{Y}_i + \hat{g}_2(i) = \hat{Y}_i + \frac{1}{\sqrt{2}} \hat{c}_i \). Then, for all \( t \in \mathbb{R} \) that are continuity points of the distribution of \( X_i \), the following convergence holds in probability

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

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

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 10.1 and thereby finishing the proof of Theorem 2.2:

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

\[
F(v) = 1_{\sum_{u \in x} \xi_2(u) \geq \frac{1}{\sqrt{n}}} = 1_{I(v) > t_0}.
\]

Then,

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

and,

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

so that \( f(+) > f(-) \) and Lemma 10.1 applies.

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

\[
F(v) = 1_{\sum_{u \in x} \xi_2(u) \leq \frac{1}{\sqrt{n}}} = 1_{I(v) \leq t_0}.
\]

Then, this time,

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

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,

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

Now,

\[
f(+) = \mathbb{P} (X_+ > -t_0) = 1 - \mathbb{P} (X_- > t_0) > 1 - \mathbb{P} (X_+ > t_0) = \mathbb{P} (X_- > -t_0) = f(-),
\]

exactly the setting of Lemma 10.1.
References

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